

---

**Molecular Genetic Characterization of Photoperiodic genes in  
Cassava (*Manihot esculenta* Crantz) and attempts to manipulate their  
expression to promote floral induction**

Inaugural-Dissertation

zur

Erlangung des Doktorgrades  
der Mathematisch-Naturwissenschaftlichen Fakultät  
der Universität zu Köln

vorgelegt von

Sarah Adeyemo  
aus Ibadan, Nigeria

Köln 2009

---

---

Die vorliegende Arbeit wurde am Max-Planck-Institut für Züchtungsforschung, Köln, in der Arbeitsgruppe von Dr Seth J. Davis, Abteilung für Entwicklungsbiologie der Pflanzen (Direktor Prof. Dr. George Coupland) angefertigt und am Internationale Mitte für Tropische Landwirtschaft von Dr Martin Fregene, Abteilung für Genetik der Maniok, Kolumbien.

Berichtersteller: Prof. Dr George Coupland

Prof. Dr. Marcel Bucher

Prüfungsvorsitzender: Prof. Dr Ulf-Ingo Flügge

Tag der mündlichen Prüfung: 25 Juni 2009

---

---

*unto the King Eternal, Immortal, Invisible,  
the ONLY WISE GOD,  
be Glory and Honor and Power Forever and Ever,  
AMEN*

---

---

## ABSTRACT

Flowering and its control is a critical developmental transition in angiosperms. The developmental timing of flowering has been widely studied in the temperate species *Arabidopsis*, which was well aided by the sequencing of its genome. However, the timing of flowering is not well understood in tropical species. Efforts are being made to extend the understanding of the molecular pathways controlling flowering from those of *Arabidopsis* to that of Cassava by taking advantage of genomics information from this temperate species. This should be an invaluable tool for understanding and exploiting floral timing in a tropical plant. Cassava is a crop that is critically important for food security in the tropics and sub-tropical regions of the world. Because, it is an orphan crop of limited research interest, almost nothing is known regarding its molecular basis of flowering. Therefore, it remains unclear what molecular pathways are implicated in flowering in Cassava and how this can be used to improve breeding.

In this thesis, I provide insight to the discovery of photoperiod genes in Cassava and exploit this pathway to manipulate floral timing. Several genes of interest were sought for in this tropical plant, ranging from clock genes, to photoperiod genes, to the floral integrator, and finally, floral meristem-identity genes. Several genes were found and characterized. From there, transgenic approaches revealed that reproductive timing can be modified in Cassava.

In Chapter 3, I identified and characterized three classes of photoperiod genes. The first gene isolated and characterized was *MeGI*, which revealed sequence similarity and conservation regions when compared with *Arabidopsis*. The comparison with *GIGANTEA*-related sequences from other species revealed by phylogeny that Cassava *GI* clearly clades with dicots and is more closely related in sequence to Castor bean *GI*. The *MeGI* expression was measured in Cassava and it was shown to track dusk. I found that *MeGI* peaks in expression at dusk under both long day or short day growth conditions. This is consistent with similar studies from other species. Another set of genes I found in Cassava was the CO-like genes, which I named *MeCOL1*, *MeCO*, and *MeCOL2*, respectively. The predicted protein encoded by *MeCOL2* possesses two adjacent zinc-finger motifs, which is specific for genes in this family. *COL2* was shown by phylogeny as the closest to the three *MeCO*-related genes. The expression of *MeCO*-like genes in Cassava showed they anticipated dawn. They were found to be acutely induced by light at dawn. Under short day growth conditions, they peaked two hours before dawn and this peak sharply increased at dawn under long days. This expression was decreased over the light period of the day. Thus, *MeCOL* genes have pre-dawn expression. Finally, I studied in genetic detail a Cassava orthologue of *Arabidopsis* *ELF4*. *MeELF4* was found to complement *elf4* by restoring circadian-rhythm defects of this *Arabidopsis* mutant.

---

Therefore, I successfully characterized in Cassava the *GI*, members of the *CO*-like gene family, and the *ELF4* orthologue, all of which are candidate genes for photoperiodic regulators.

I bridged the gap between basic and applied research by exploiting the photoperiod- integrator gene *FLOWERING LOCUS T (FT)* as a tool in the manipulation of floral timing in Cassava. To generate this tool, I placed the Arabidopsis *FT* gene under the control of an ethanol-inducible promoter. I showed that the construct was functional, as it promoted flowering in a late-flowering genotype of Arabidopsis after ethanol application. This technology was then transferred to Cassava. I transformed a shy flowering Cassava genotype with Arabidopsis *FT* under the control of an ethanol-inducible promoter. *FT* expression was found to increase after spraying the Cassava transgenics with ethanol. These plants then flowered. This system is hoped to be applied in conventional breeding programs in order to be able to induce flowering at will, and thus make desired crosses that will lead to the improvement of the genetic basis of the crop.

---

## ZUSAMMENFASSUNG

Die Kontrolle des Blühzeitpunktes ist ein kritischer Entwicklungsschritt in Angiospermen. Das entwicklungsbiologische Timing des Blühzeitpunktes wurde, auch begünstigt durch die Sequenzierung ihres Genoms, weitgehend in *Arabidopsis thaliana* untersucht, einer Pflanze aus gemäßigten Zonen der Erde. Bei tropischen Arten hingegen ist die Kontrolle des Blühzeitpunktes noch wenig erforscht. Es wird daher versucht, das Verständnis der molekularen Signalwege, die das Blühen kontrollieren, ausgehend von *Arabidopsis* zu denen von Maniok (engl. Cassava; *Manihot esculenta*) zu erweitern, indem die Genominformationen aus *Arabidopsis* genutzt werden. Dies sollte ein wertvolles Werkzeug sein, um das Blühtiming einer tropischen Pflanze zu verstehen und entsprechend zu nutzen.

Maniok ist eine Nutzpflanze, die für die Ernährungssicherheit in den Tropen und den subtropischen Regionen der Welt von entscheidender Bedeutung ist. Da es sich um eine vernachlässigte Ertragspflanze von begrenztem Forschungsinteresse handelt, ist fast nichts über die molekularen Grundlagen ihres Blühens bekannt. Daher sind die molekularen Signalwege, die eine Rolle beim Blühen von Maniok spielen, ebenso unklar wie die Frage, wie diese Signalwege zur züchterischen Verbesserung von Maniok genutzt werden könnten.

Im Rahmen dieser Doktorarbeit gewähre ich Einblicke in die Entdeckung von Genen des photoperiode-abhängigen Blüh-Signalweges in Maniok und nutze Gene dieses Signalweges, um den Blühzeitpunkt zu manipulieren. Verschiedene Gene wurden gefunden und näher charakterisiert, darunter Gene der circadianen Uhr, des photoperiode-abhängigem Blüh-Signalweges, Blühzeitpunktgene und letztendlich Gene, die für die Identität des Blütenmeristems verantwortlich sind. Davon ausgehend konnte ich über transgene Pflanzen zeigen, dass der Blühzeitpunkt in Maniok modifiziert werden kann.

In Kapitel 3 habe ich drei Klassen photoperiodischer Gene identifiziert und analysiert. Das erste isolierte und charakterisierte Gen war ein *GIGANTEA* Ortholog (*MeGI*), das eine hohe Sequenzähnlichkeit im Vergleich zu *Arabidopsis* zeigte. Der Vergleich mit *GIGANTEA*-Orthologen aus anderen Arten offenbarte durch Phylogenie, dass *MeGI* eindeutig dem Stamm der Dicotyledonen angehört und am engsten mit *GI* von Rizinus verwandt ist. Die *MeGI* Expression wurde in Maniok gemessen und es zeigte sich, dass die *MeGI* Expression sowohl unter Langtag- als auch unter Kurztag-Wachstumsbedingungen jeweils am Abend am höchsten ist. Dies stimmt mit ähnlichen Studien in anderen Organismen überein.

---

Eine andere Gruppe von Genen, die ich in Maniok analysierte, waren die CONSTANS-ähnlichen Gene, die ich *MeCOL1*, *MeCO* und *MeCOL2* nannte. Das vorhergesagte Protein von *MeCOL2* besitzt zwei benachbarte Zink-Finger Motive, die spezifisch für Gene dieser Familie sind. Phylogenetische

Untersuchungen zeigten, dass *COL2* den drei *MeCO*-verwandten Genen am nächsten steht. Die Expression von *CO*-ähnlichen Genen in Maniok zeigte, dass sie die Morgendämmerung antizipieren. (Es wurde festgestellt, dass sie durch Licht bei Tagesanbruch stark induziert werden). Unter Kurztagbedingungen erreichte die Expression den Höhepunkt zwei Stunden vor Tagesanfang, während das Expressionsmaximum unter Langtagbedingungen genau beim Tagesanfang lag. Diese Expression nahm während der Lichtperiode des Tages ab. Dies zeigt, dass *MeCOL*-Gene eine Expression vor Tagesanbruch haben. Schlussendlich studierte ich das Maniok Ortholog von *Arabidopsis* *ELF4* im Detail. Es stellte sich heraus, dass *MeELF4* die *elf4*-Mutation in *Arabidopsis* komplementiert und die circadianen Rhythmus-Defekte dieser Mutante beheben kann.

Daher habe ich in Maniok erfolgreich *GI*, Mitglieder der *CO*-ähnlichen Genfamilie und das *ELF4* Ortholog charakterisiert, welche allesamt Kandidatengene für photoperiodische Regulatoren sind.

Ich schloss die Lücke zwischen Basis- und angewandter Forschung, indem ich das photoperiodische Integratorgen *FLOWERING LOCUS T (FT)* als Werkzeug für die Manipulation des Blühzeitpunkts in Maniok nutzte. Um dieses Werkzeug zu erzeugen, setzte ich das *Arabidopsis* *FT*-Gen unter die Kontrolle eines Ethanol-induzierbaren Promotors. Ich zeigte die Funktionalität dieses Konstrukts, da es das Blühen in einem spät-blühenden Genotyp von *Arabidopsis* nach Ethanolbehandlung förderte. Diese Technologie wurde dann auf Maniok übertragen. Ich transformierte einen spät blühenden Genotyp von Maniok mit *Arabidopsis* *FT* unter der Kontrolle eines Ethanol-induzierbaren Promotors. Es wurde gezeigt, dass sich die *FT*-Expression nach dem Besprühen der transgenen Maniokpflanzen mit Ethanol erhöhte. Daraufhin begannen diese Pflanzen zu blühen.

Es besteht die Hoffnung, dass dieses System in konventionellen Züchtungsprogrammen Anwendung finden wird, um den Beginn des Blühens zu einem gewünschten Zeitpunkt zu induzieren. Dies wird erwünschte Kreuzungen ermöglichen und somit dafür sorgen, dass die genetische Basis der Nutzpflanze Maniok weiter verbessert wird.

---

# TABLE OF CONTENTS

ABSTRACT	i
ZUSAMMENFASSUNG	iii
TABLE OF CONTENTS	v
ABBREVIATIONS	viii
LIST OF FIGURE ELEMENTS	x
<b>CHAPTER 1. INTRODUCTION</b>	<b>1</b>
1.0 Introduction	2
1.1 Introduction to Cassava	2
1.2 Economic Importance of Cassava	2
1.3 Origin and distribution of Cassava	3
1.4 Cassava Morphology and Physiology	4
1.5 Genome structure of Cassava	8
1.6 Cassava Production	8
1.7 Constraints of Cassava	9
1.8 Cassava <i>in vitro</i> Regeneration (Methods of Cassava transformation)	11
1.8.1 Friable Embryogenic Callus	11
1.8.2 Somatic Embryogenesis	11
1.9 History of Genetic Transformation in Cassava	12
1.9.1 Promoters Choice in Plant Biotechnology	12
1.9.1.1 The AlcA promoter system transferred into plants	13
1.10 Flowering Control Pathways in Arabidopsis	13
1.10.1 Day length/Photoperiod in Arabidopsis	14
1.11 Photoperiod Perception In Cassava	15
1.12 Aim of thesis	15
<b>CHAPTER 2. MATERIALS AND METHODS</b>	<b>16</b>
2.1 Materials	17
2.1.1 Chemicals	17
2.1.2 Enzymes	17
2.1.3 Bacterial Strains	17
2.1.3.1 <i>Escherichia coli</i>	17
2.1.3.2 <i>Agrobacterium tumefaciens</i>	17
2.1.4 Vectors	18



---

2.1.5 Oligonucleotides	19
2.1.6 Media	22
2.1.6.1 Media for bacteria	22
2.1.6.2 Media for plants	22
2.1.7 Buffers and other reagents	26
2.2 Methods	28
2.2.1 Plant Material	28
2.2.2 Molecular Biology Techniques	29
2.2.3 Plant Procedures	34
2.3.0 Software, databases, and other internet resources	39

**CHAPTER 3. DISCOVERY, PHYLOGENETIC RELATIONSHIPS, AND CHARACTERIZATION OF SEVERAL PHOTOPERIODIC GENES OF CASSAVA** 40

3.1.0 Introduction/ In Quest for Photoperiod Cassava Genes	41
3.2.0 Results	42
3.2.1 Computational Searches for Photoperiodic genes	42
3.3.0 Identification of the Cassava <i>GIGANTEA</i> gene	42
3.3.1 Sequence alignment of encoded GI proteins	44
3.3.2 Phylogenetic relationships of encoded GI protein among plant species and their relationships.	46
3.3.3 <i>MeGI</i> expression in wild-type Cassava under LD and SD conditions	48
3.4.0 Identification of <i>CONSTANS</i> -Like Genes in Cassava	49
3.4.1 Phylogenetic relationship of <i>CONSTANS</i> -Like genes from both Arabidopsis and Cassava	50
3.4.2 Sequence alignment of <i>CONSTANS</i> -Like Proteins	52
3.4.3 Phylogenetic relationships of <i>CONSTANS</i> -Like proteins among plant species and their relationships	54
3.4.4 <i>MeCOL</i> expression in wild-type Cassava under LD and SD conditions	56
3.5.0 Identification and Characterization of the <i>ELF4</i> Gene in Cassava	57
3.5.1 Sequence alignment of ELF4 proteins from other species	58
3.5.2 Phylogenetic relationships of ELF4 proteins among plant species and their relationships	60
3.6.0 Complementation of <i>elf4-1</i> mutant with <i>MeELF4</i>	62
3.6.1 Results from hypocotyl elongation measurement	62
3.6.2 Results from leaf movement measurement	63
3.6.3 Molecular confirmation by imaging	64

**CHAPTER 4. EXPLOITING ARABIDOPSIS *FT* UNDER THE CONTROL OF AN ETHANOL- INDUCIBLE SYSTEM TO PROMOTE FLOWERING IN ARABIDOPSIS AND CASSAVA** 67

4.1 Introduction	68
4.2 Results	70
4.2.1 Arabidopsis <i>FLC/FRI</i> experiments	70

---

---

4.2.1.1 Conventional PCR confirmation of <i>FLC/FRI</i> transgenic events	70
4.2.1.2 Ethanol induction	71
4.2.1.3 Flowering time measurement	73
4.2.2.0 Floral promotion from Arabidopsis <i>FLOWERING LOCUS T</i> in shy flowering genotype of Cassava	75
4.2.2.1 Genomic PCR analysis of <i>AtFT</i> in Cassava transgenics	76
4.2.2.2 Gene expression of <i>FT</i> transgene in Cassava by RT-PCR analysis	77
4.2.2.3 Observations in the greenhouses	78
4.3 Conclusion	81
<b>CHAPTER 5. GENERAL CONCLUSIONS AND PERSPECTIVES</b>	82
5.0 Introduction	83
5.1 <i>MeGI</i> Summary	83
5.1.1 Future perspectives on <i>MeGI</i>	83
5.2 Summary of the <i>COLs</i>	84
5.2.1 Future perspectives on <i>COL</i> genes	84
5.2.2 <i>MeELF4</i> Summary	85
5.2.3 Future perspectives on <i>MeELF4</i> gene	85
5.3 Perspective on genes of interest and Cassava flowering model	86
5.3.1 Floral repressors	86
5.4 Summary of Arabidopsis Floral induction	87
5.4.1 Future perspectives on the Arabidopsis <i>FLC/FRI</i> genotypes	88
5.5 Summary of Cassava Floral induction	88
5.5.1 Future perspectives on the Cassava transgenic lines	88
<b>CHAPTER 6. LITERATURE</b>	92

---

## ABBREVIATIONS

2,4- D	2,4-dichlorophenoxyacetic acid
AG	AGAMOUS
AGL24	AGAMOUS-LIKE 24
API	APETALA 1
B <sub>5</sub>	Vitamin B <sub>5</sub>
BAP	6-benzyl amino purine
CAB	CHLOROPHYLL A/B BINDING
CCA1	CIRCADIAN CLOCK ASSOCIATED 1
CCR2	COLD CIRCADIAN REGULATED 2
CO	CONSTANS
COL	CONSTANS LIKE
CGIAR	Consultative Group on International Agricultural Research
CIAT	Centro Internacional de Agricultura Tropical
CRYs	Cryptochromes
DNA	Deoxyribonucleic acid
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária Assessoria de Comunicação Social Ministério da Agricultura, Pecuária e Abastecimento
ELF4	EARLY FLOWERING 4
ELF3	EARLY FLOWERING 3
FeNaEDTA	Iron sodium ethylenediaminetetraacetate
FLC	FLOWERING LOCUS C
FRI	FRIGIDA
FT	FLOWERING LOCUS T
GA <sub>3</sub>	Gibberellic acid 3
GD	Gresshof and Doy
GI	GIGANTEA
IBA	Indole-3-butyric acid
LD	Long day
LFY	LEAFY
LHY	LATE ELONGATED HYPOCOTYL
LL	Continuous red and blue light
LUC	LUCIFERASE
MS	Murashige and Skoog medium

---

MCol 2215	Cassava cultivar Manihot columbia 2215
μL	microliter
μM	micro molar
NAA	Naphthalene acetic acid
ORF	Open Reading Frame
PCR	Polymerase chain reaction
PHYs	Phytochromes
PRR9	PSEUDO-RESPONSE REGULATOR 9
RNA	Ribonucleic acid
RT-PCR	Reverse transcription- polymerase chain reaction
SD	Short day
SOC1	SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1
T-DNA	Transfer DNA
TMS 60444	Cassava cultivar Tropical Manihot Selection 60444
TOC1	TIMING OF CAB EXPRESSION1
ZT	<i>Zeitgeber</i> time

---

## LIST OF FIGURE ELEMENTS

### Figures

	Page
Figure 1.1 Cassava growing in the field and the harvested tuber	2
Figure 1.2 The Cassava Distribution Map	4
Figure 1.3 Cassava's Floral Parts	7
Figure 2.1 The Vector Map of pNewFT used for the transformation of both Arabidopsis and Cassava	18
Figure 2.2 Cassava Transformation from Agrobacterium-mediated transfer to somatic embryo formation	37
Figure 2.3 Cassava somatic embryo germination to green house establishment	38
Figure 3.1 Partial Protein-Sequence alignment of Arabidopsis and Cassava GI	43
Figure 3.2 The multiple alignment of GI protein sequences from other species showing conserved amino-acid regions	45
Figure 3.3 Phylogenetic tree showing GI relationship within species	47
Figure 3.4 Expression of <i>MeGI</i> under SD (8hL, 16hD) conditions	48
Figure 3.5 Expression of <i>MeGI</i> under LD (16hL, 8hD) conditions	49
Figure 3.6 Phylogenetic tree showing <i>CO</i> -like genes from both Arabidopsis and Cassava	51
Figure 3.7 Sequence alignment of <i>CO</i> -like proteins from Arabidopsis and Cassava	53
Figure 3.8 Phylogenetic tree showing <i>CO</i> -like proteins from different species	55
Figure 3.9 Expression of <i>MeCO</i> -like under SD (8hL, 16hD) conditions	56
Figure 3.10 Expression of <i>MeCO</i> -like under LD (16hL, 8hD) conditions	57
Figure 3.11 The multiple alignment of <i>ELF4</i> protein sequences and their conservation	59
Figure 3.12 Phylogenetic tree of <i>ELF4</i> proteins from different species	61
Figure 3.13 Hypocotyl-length elongation confirmed complementation of <i>MeELF4</i>	63
Figure 3.14 Representative leaf movement data for lines expressing a clock phenotype	64
Figure 3.15 <i>CCA1::LUC</i> in various <i>ELF4</i> genotypes	65
Figure 3.16 <i>CCR2::LUC</i> in various <i>ELF4</i> genotypes	65
Figure 4.1 The ethanol inducible system driving the expression of the <i>FT</i> gene	69
Figure 4.2 Genotyping by PCR of the 5 transgenic lines of Arabidopsis	71
Figure 4.3 <i>FT</i> was found to be expressed in <i>FLC/FRI</i> transgenic <i>FT</i> genotype after spraying with 1% ETOH	72
Figure 4.4 <i>FT</i> is NOT-expressed in <i>FLC/FRI</i> transgenic genotype after spraying with distilled water	73
Figure 4.5 Ethanol application promotes flowering in some pNewFT <i>FLC/FRI</i> transgenic lines	74

---

Figure 4.6 FT-5A transgenic line flowering along side with the empty control and <i>FLC/FRI</i> non-transgenic genotype after 3 weeks after the ethanol induced <i>FT</i> was activated by 1% ethanol	75
Figure 4.7 Genomic PCR of Arabidopsis <i>FT</i> in Cassava transgenics	76
Figure 4.8 <i>FT</i> expression in Cassava transgenics 3,7, and 15 days after induction with 2.5% ethanol	77
Figure 4.9 Cassava plants in the glass house in Colombia	79
Figure 4.10 Cassava plants in the green house in Germany	80
Figure 4.11 Cassava transgenics flowered after induction of the ethanol controlled <i>FT</i> transgene system by ethanol	81
Figure 5.1 A Comparison of the flowering model of Arabidopsis and a comparison to Cassava	87
Figure 5.2 Southern blot hybridization on Cassava transgenic lines with <i>FT</i> probe	90

## Tables

Table 1.1 Primer sequences designed for cloning, genotyping, and gene discovery	19
Table 3.1 The source of each <i>GI</i> amino-acid sequence, and their corresponding name	44

---

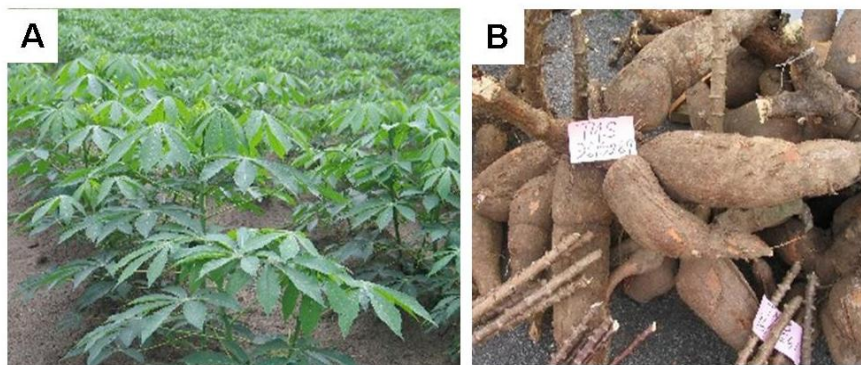
## **CHAPTER 1. INTRODUCTION**

---

## 1.0 Introduction

### 1.1 Introduction to Cassava

Cassava (*Manihot esculenta* Crantz) is the fifth most important source of food energy. It is ranked below wheat, rice, maize, and sorghum in terms of global calorie consumption. Furthermore, it is the most important staple food crop in the sub-Saharan Africa (FAO 2001). Cassava is a member of the dicotyledonous family Euphorbiaceae (Alves, 2002), whose genus *Manihot* has been reported to have approximately 98 species. Cassava is the only domesticated and agronomically important species of this genus (Allem, 2002). It is a famine reserve crop because of its ability and capacity to yield well in drought-prone, marginal wastelands, where other crops would fail. It is a shrub reaching 1-4 m in height (Figure 1.1). It provides a significant role as an edible source of carbohydrate in many tropical countries, as it accumulates and stores starch within large swollen secondary root structures. It is commonly known as Cassava in English, yuca in Spanish, tapioca and manioc in French, and mandioca in Portuguese.



**Figure 1.1 Cassava growing in the field and the harvested tuber (A)** A typical Cassava plant of about 2 months old growing in the field (B) Cassava tubers and stakes displayed at a farmer's festival. Pictures adapted from a web source.

### 1.2 Economic Importance of Cassava

Cassava is an important staple food in Africa and is the most important source of energy for over 200 million people in sub-Saharan Africa. There, it is extensively cultivated for human consumption (Dahniya, 1994). Processing and cooking serves as the main method to extend shelf-life of the root, which rapidly deteriorates after harvest. This also improves palatability and lowers water content to reduce transportation costs. As food source, the root is cooked, or converted to roasted or steamed granules, made into flour used in the bakeries, dry cut into chunks, made into fermented pastes, processed into drinks, and processed into many other variations. The leaves can also be eaten. They are high in protein, vitamin C, iron, and calcium. The leaves are generally



---

cooked for an extended period of time. Cassava leaves are an important vegetable in Congo and in Tanzania, but are little used in Uganda and West Africa (Nweke *et al.*, 2002). In Latin America, Brazil is the principal country that makes use of Cassava leaves. These leaves serve as silage for animal feed, whereas in industry, the tuberous roots are processed for the manufacture of starch and starch-derived products, alcohol, and high fructose-glucose syrups. The main benefits for Cassava as a crop plant are its starch content, nutritional value, and post-harvest storage characteristics. Collectively, Cassava provides five main roles: (1) famine reserve; (2) rural food staple; (3) urban food staple; (4) livestock feed and industrial raw material; and (5) earner of foreign exchange. Cassava therefore has a number of economic and security uses. Some regions use Cassava for multiple purposes, whereas others do not (Nweke *et al.*, 2002).

### **1.3 Origin and distribution of Cassava**

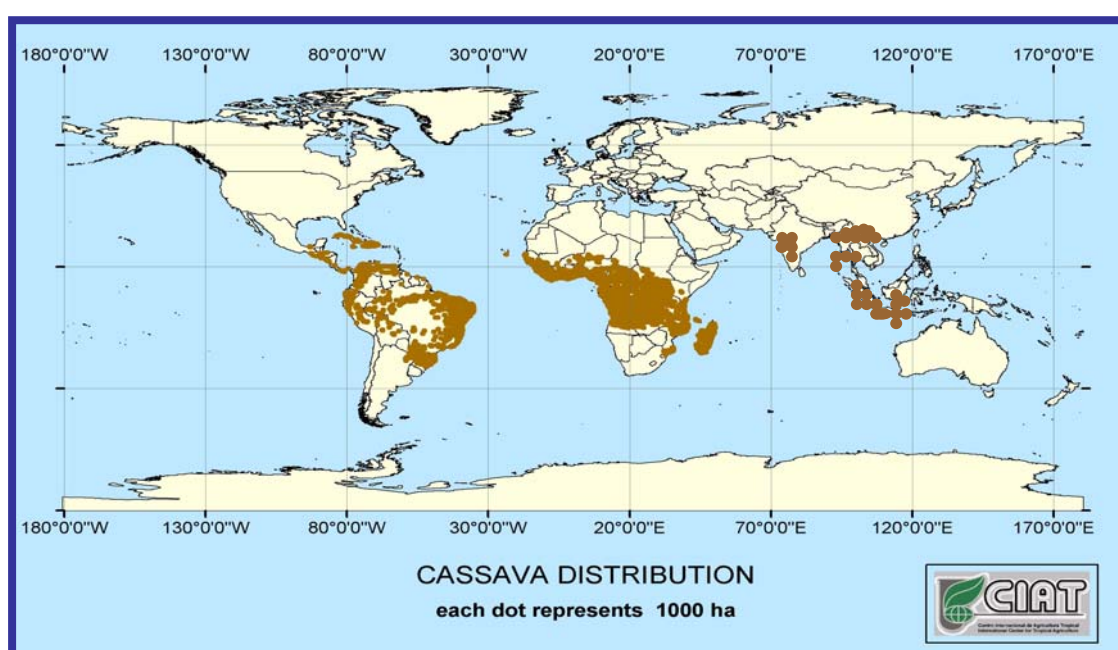
Cassava is widely cultivated throughout the tropics, and is generally grown in environments with minimal agronomic modifications. Fertilizer, irrigation, or other inputs are generally not used and the crop is subject to a wide variation of environmental factors. Among the most important of these environmental factors are temperature, photoperiod, light intensity, water, relative humidity, and soil characteristics (Alves, 2002). Variations are greatest across geographical areas, but can also be substantial across time within a given site.

The origin of domestication of Cassava had been disputed for many years. With botanical, genetic, and geographical evidence, its origin is now understood to be from South America. This pointed to the Amazon region as the center of Cassava domestication (Allem, 1994; Hillocks, 2002). The domestication of Cassava was further investigated using a phylogeographic study. In this work, the locus of the single-copy nuclear-gene glyceraldehyde 3-phosphate dehydrogenase (*G3dph*) provided high levels of non-coding sequence variation in Cassava and its wild relatives, with 28 haplotypes identified among 212 individuals, which demonstrated that Cassava was domesticated from wild Cassava populations along the southern border of the Amazon basin (Olsen and Schaal, 1999).

According to Rogers and Appan (1973), the Portuguese were the first to import Cassava from the eastern coast of Brazil to Africa. After the 16<sup>th</sup> century, the species gradually spread through the various regions of sub-Saharan Africa. The inhabitants of this region then imported Cassava to Madagascar and Zanzibar (Jennings, 1976). Cassava was introduced from the western coast of Central America by Spanish explorers to Southeast Asia and the Philippines. It was also introduced to Southern India from the East African coast. Thus, Cassava is now present in all tropical regions.

---

The inland spread of Cassava cultivation within Africa was by African traders who were drawn to Cassava for its ability to provide security against famine and its drought tolerant properties. The importance of Cassava as a staple in the African community has prompted the Consultative Group on International Agricultural Research (CGIAR) to inaugurate the International Institute of Tropical Agriculture (IITA) with its headquarters in Ibadan, Nigeria in 1972 to oversee the development of Cassava in Africa. Together with the Centro Internacional de Agricultura Tropical (CIAT, its Spanish acronym) in Colombia, IITA has the global mandate for Cassava improvement and is responsible for developing the crop in Africa. (Hillocks, 2002). Today, Cassava is spread in large areas of Africa, Asia, and South America and grown in over 60 countries most of which are developing countries (Figure 1.2).



**Figure 1.2 The Cassava Distribution Map.** The graph shows the widespread of Cassava. The brown dots represent Cassava cultivation points that are over 1000 ha. Picture was adopted from the graphics department of CIAT.

#### 1.4 Cassava Morphology and Physiology

Cassava has simple palmate leaves with 3 to 9 lobes. They are responsible for both photosynthesis and tuberization. The leaves are subtended by a long petiole and arranged spirally around the stem, which is important in setting the leaf for light interception. The lobes are varied in color, shape, size and number. The leaf structure includes epidermal tissues with cuticles; mesophyll, with palisade and spongy parenchyma; stomata; and pigments. The coloring of the terminal bud is an important feature of cultivars, which can be purple, green, or copper (EMPRABA, 2005). When grown at elevated temperatures ( $> 24^{\circ}\text{C}$ ), the time from appearance to full expansion of a leaf takes about two weeks. The size of fully expanded leaves increases with the age of the plant. At low temperatures, the maximum size is smaller and the product of the largest leaves is delayed. There are large varietal and environmental effects on leaf area (Irikura *et al.*, 1979). Drought stress

---

---

(Connor and Cock, 1981) and a limited supply of nutrients (CIAT, 1979) can both greatly reduce leaf size.

Cassava employs the C<sub>3</sub> carbon-fixation metabolic pathway for carbon fixation in photosynthesis (Edwards *et al.*, 1990; Angelov *et al.*, 1993; Ueno and Agarie, 1997). The maximum photosynthetic rates varies from 13 to 24  $\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-2}$  under greenhouse or growth chamber conditions (Mahon *et al.*, 1977, Edwards *et al.*, 1990) and from 20 to 35  $\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-2}$  in the field (El-Sharkawy and Cock, 1990). Cassava has a high CO<sub>2</sub> compensation point, typical of C<sub>3</sub> plants. The optimal temperature for photosynthesis for field grown Cassava is 35°C, but the range for optimal photosynthesis is 25 to 45°C (El-Sharkawy and Cock 1990). Thus, Cassava is adapted to the tropical environment.

Under long days of 16 hour photoperiod, total plant weight tends to remain the same or decrease, and the proportion of root weight to total plant weight (harvest index) generally decreases (Bolhuis, 1966; CIAT, 1982). Photoperiod also influences reproductive development, and while not directly associated with yield formation, has implications for canopy (*e.g.* branching habit) as well as seed production in a breeding program. A long photoperiod induces flower initiation, and consequently branching, in many genotypes (Cunha and Conceição, 1975; Bruijn, 1977; and Keating *et al.*, 1982a). Intensity of light received by individual leaves can vary as a result of cloud cover, competition for light by an intercropped species, or intraplant/ interspecific shading. Intercropping effects can be especially pronounced when Cassava is completely shaded. Cassava appears to be highly sensitive to reduced light intensity (El-Sharkawy and Cock, 1990).

Cassava grows in the tropics from sea level to about 2,200 m elevation, in areas receiving more than 400 mm average annual rainfall. In the subtropics, maximum elevation is somewhat lower. The species is cultivated between 30° North latitude and 30° South latitude. The largest plantation of this crop are concentrated between the parallels of 20° N and 20° S (EMBRAPA, 2005). This is due to the need for a long growing season, and also to the difficulty of storing the Cassava planting material for extended periods of time. Here, it is noted that Cassava is difficult to store during an extended cold winter. This weather condition inhibits its storability for planting.

Cassava is sensitive to frost. It has a growing season of up to a year and more, which essentially limits its range to the tropics and subtropics. When grown in frost-prone areas, such as winter season in the subtropics, it is generally harvested or cut back before winter (Ng and Ng, 2002). Minimum mean temperature for growth is about 17°C, where the absolute minimum should not fall below 10°C. Below these temperatures, stake sprouting is extremely delayed and may fail completely. From this temperature reduction, growth and yield drop off markedly (Cock and Rosas, 1975; Alves, 2002). The upper limits for temperature adaptation appears to be within the

---

---

range of most tropical environments, but the ideal average annual temperature for its development and production is between 25 and 28 °C (Alves, 2002).

Cassava is highly tolerant to drought. There are regions where farmers grow it in areas that have less than 500 mm of rainfall per year. In higher rainfall areas, the crop can withstand sustained dry periods (Pardales and Esquibel, 1996). The species combines several mechanisms that allow it to withstand both short and prolonged water stress periods. The principal response to moderate water shortage is a reduction of leaf area. The plant maintains normal root growth or may even increase it. (Connor *et al.*, 1981; Connor and Cock, 1981). The combination of smaller leaf size and slower leaf formation rate, rather than leaf fall-off is important to Cassava's ability to maintain high root yield under stress. The ability of Cassava to regulate its stomata to maintain high midday leaf water potentials and prevent water loss is a key mechanism for tolerance to prolonged drought (El-Sharkawy *et al.*, 1984a; 1984b). Although partial closure of the stomata restricts O<sub>2</sub> supply to the leaf, it leads to a stable leaf water potential during stress. A high percentage of Cassava is produced on soils with low pH (often resulting from high aluminium saturation in highly leached soils, or in the organic tropical peat soils), and generally low levels of major nutrients (Howeler, 2002).

Little is known about the physiology and the biochemistry of flowering in Cassava. Some clones have never been seen to flower (Alves, 2002). It appears that moderate temperatures (approximately 24°C) are suitable for flowering in Cassava (Alves, 2002). Branching correlates to the onset of flowering, which is promoted in long days by some cultivars (Alves, 2002). Usually, the apical meristem becomes reproductive when branching occurs, but abortion of flowers is common. Keating *et al.* (1982a), evaluated Cassava at 12 different planting dates at a high latitude (27°C 37'S), where photoperiods ranged from 14.8 to 11.2 hours. Flowering and forking was found to occur under photoperiods > 13.5 hours. This result is consistent with similar findings from Bruijn (1977) and Cunha and Conceição (1975), who suggested flowering in Cassava may be promoted by increasing day length.

Cassava's inflorescence occurs along the juncture between branches, branching is a precondition for flowering. The flowers are monoecious, producing both male (pistillate) and female (staminate) flowers on the same plant, in the same inflorescence. The flowers have a simple unisexual form, and display an indefinite structure called a perianth, which is comprised of five tepals that can be yellow, red, or purple. The flowers are arranged in chymes; the male flowers are smaller and more numerous than the female ones. The flowers lack corolla or calyx, but have a perianth with five tepals which can be yellow, red or purple, depending on the cultivar. The staminate flower is located on the lower part of the branches and the pistillate flower on the upper part of the branch (Alves, 2002).

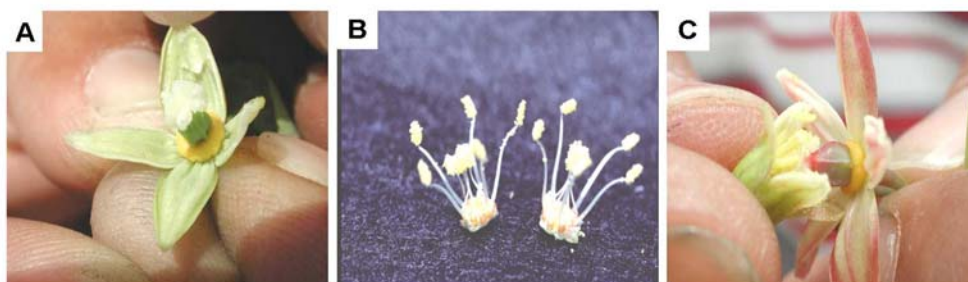
---

---

The male flower has a delicate pedicel (supporting stalk) and five overlapping tepals, which are sometimes colored on their ventral side (Figure 1.3B). On the inside of the male flower, there is a disk divided into 10 lobules, in the center of which there are the beginnings of an ovary. Between these lobules ten stamens or filaments grow, arranged in two series: five internal and five external, with the latter being more developed. On the stamens there are anthers that face the center of the flower. The pollen is yellow or orange and the pollen grains range in size from 122-148  $\mu\text{m}$  (Ghosh *et al.*, 1988).

The female flower has a less developed pedicel and a perianth divided down to the base, with colorful tepals that, numbering five, are always colored on their ventral side (Figure 1.3A). The superior ovary has three carpels, each of which houses an individual ovule, with a micropyle (opening) facing upward. The central disc, located at the base of the ovary, has five lobes, each of which has two lobules. The stigma is wide, crinkly, fleshy, and of varying color. The female flower has a ten-lobed basal disk while the ovary is tricarpeillary with six ridges. The three locules each contain one ovule which matures into carunculate seeds. The female flowers normally open between 10-14 days before the males on the same branch. Together, this makes the plant prone to cross-pollination, which might explain the high level of heterozygosity in Cassava (Alves, 2002).

Cross-pollination of Cassava plants is often made possible by insects (Alves, 2002). Self-pollination is somewhat uncommon because of protogyny, by which female flowers open before male flowers in the same raceme limiting synchronized flowering (Alves, 2002). However, in rare cases of simultaneous maturing of male and female flowers in different racemes of the same plant, self-pollination can occur. Manual pollination for breeding purposes involves the female and the male floral parts (Figure 1.3C).



**Figure 1.3 Cassava's Floral Parts** (A) The Cassava female flower (B) The Cassava male flower (C) A type of controlled pollination

Cassava roots are rich in starch. They have various shapes depending on the variety and environmental conditions in which the plant develops. They can be cylindrical, conical, fusiform,

---

cylindrical-conical, or globular; crooked roots, constricted roots, and lateral offshoots are also common. The plant has on average, between 5 and 12 roots, whose sizes depend on growing conditions. The root can grow over one meter long. In addition to its tuberous roots, Cassava plants have fibrous root network, which is responsible for water and nutrient absorption as well as providing stability to the plant. The tissue that constitute the tuberous roots are: the peel, the pulp and the central fibres. The peel is made up of the periderm, which is sometimes smooth or wrinkled, with varying shades of white, cream, or brown; and the phelloderm, a layer of varying colour with cells rich in starch. The phloem is located in the cortex and is characterized by vessels that contain latex rich in cyanogenic glycosides. The pulp consists of secondary xylem tissue and storage parenchyma cells constituting a fleshy part that is rich in starch. In the center of the center cylinder, there are hard xylem and sclerenchyma vessels, which form the central fibre bundles- thick, white cords of cellulose nature that run longitudinally along the roots (EMPRABA, 2005).

### **1.5 Genome structure of Cassava**

Cassava has a diploid genome ( $2n=36$ ). The nucleic acid content of diploid Cassava is 1.67 picogram per nucleus that is 772 mega base pairs in the haploid genome (Awoleye *et al.*, 1994). Some authors have described it as an allotetraploid with basic chromosome number  $x=9$ . Jos and Nair, (1979) however, conducted studies on the meiotic behavior of several Cassava genotypes and observed regular 18 bivalent formation of the chromosomes typical of its diploid ( $2n=2x=36$  chromosomes). In 1997, a molecular genetic linkage map of the Cassava genome was constructed (Fregene *et al.*, 1997). However, not much information is known about individual genes that make up the genome. The genomes of more than 180 genomes of organisms have been sequenced, ranging from bacteria to plants to animals (Goodner *et al.*, 2001; Dietrich *et al.*, 2004; AGI, 2000, Venter *et al.*, 2001; Yu *et al.*, 2002). There are still a lot of on-going genome sequencing projects. Cassava falls within the category of organisms whose genome has not been sequenced. Sequencing the Cassava genome is important to be able to know and understand how genes in this crop plant work together to direct the growth, development and maintenance of the tropical plant.

### **1.6 Cassava Production**

Production of Cassava in Africa makes up to 54% of the world total, with the other two major producing continents being South America and Asia (FAO 2003). This 54%, however, is cultivated on an estimated 65% of the total area available to Cassava (FAO 2003). The reason for this apparent mismatch is that the yields in Africa, averaging 8.8t/ha, is only 70% of that in South America and 61% of that in Asia (Legg and Thresh, 2003).

Cassava production expanded broadly throughout the lowlands tropics in the twentieth century, mainly on the less-fertile, poor-quality agricultural lands (Hillocks, 2002). In traditional, low-input cropping systems, Cassava is often an end-of-cropping-phase where it is grown and yield well on

---

low fertility soils. Its ability to withstand locust attacks and drought, and its low cost of production motivated the farmers to replace other traditional root crops, such as yams (Hillocks, 2002). In areas where population growth has caused a reduction of the rotation pattern shifting culture and a commensurate decline in soil fertility, Cassava is one of the few crops that can thrive without supplied inputs, provided some form of rotation remains. Similarly, in the Southern India and Java (both with very high population densities), Cassava is increasingly relegated to low-quality land. In one of the most notable agricultural success stories of recent years, the area planted to Cassava increased fivefold in Thailand. This development in the 1970s was to meet an export opportunity in Europe. However, the production and commercialization of Cassava is limited due to a host of abiotic and biotic constraints.

### **1.7 Constraints of Cassava**

Cassava is plagued by a host of insect pests and diseases, a major production constraint which can cause up to 70% to 100% yield loss (Abate *et al.*, 2000). These pests include the Cassava hornworm, Cassava mites, thrips, and gall midges. There are also prevalent diseases of the Cassava Mosaic and the African Mosaic Virus, both of which are transmitted by the insect vector-white fly (Bellotti *et al.*, 1999). Other types of diseases of Cassava include brown streak, stem rot, anthracnose, bacterial blight, frog skin and the brown leaf spot (Hahn and Keyser, 1985; Thomas *et al.*, 1986; Guthrie, 1987; Silvestre, 1989; Alaux and Fauquet, 1990; Hong *et al.*, 1993; Swanson and Harrison, 1994; Legg and Thresh, 2003). The combat against pests and diseases of Cassava continues to be a major breeding challenge.

Cassava contains high cyanide because it accumulates the cyanogenic glucoside compounds linamarin and lotaustralin. Linamarin is synthesized in the leaves and transported to the roots (Wheatley and Chuzel, 1993; Siritunga and Sayre, 2003). The hydrolysis of linamarin and breakdown of acetone cyanohydrin by hydroxynitrile lyase (HNL) or elevated pH (> 5.0) leads to the release of hydrogen cyanide (HCN), which is poisonous (Cooke and Coursey, 1981). The total amount of cyanogenic glycoside in Cassava plant is dependent on cultivar, cultural practice, environmental conditions, and plant age (McMahon *et al.*, 1995). Processing of Cassava roots before consumption will hydrolyze the cyanogenic glycosides by hydroxynitrile lyase to acetone and cyanide (White *et al.*, 1998). Improvement of Cassava varieties with high cyanogenic glycoside compounds can be achieved through breeding with elite cultivar with that of a variety with low levels of the compound. As well, genetic transformation events has been used to successfully reduce the amount of the cyanogenic glycosides in Cassava by antisense downregulation of cytochrome P450 genes (CYP79D1/D2) involved in cyanogenic glycoside synthesis (Siritunga and Sayre, 2003; White *et al.*, 1998, Siritunga *et al.*, 2004).

---

The Cassava root is high in starch content. Around 80% of Cassava dry root matter is starch, and thus it suffers from low protein content (El-Sharkawy, 2004). Less than 2% of the dry weight of the root is protein, as compared to 9.1% in potato (Bushway *et al.*, 1980). There is therefore a need to increase the protein content in roots of Cassava. Reports of crosses between Cassava and *Manihot tristis* revealed root protein content of more than 8% in F1 hybrids (Bolhuis 1953; Asiedu *et al.*, 1992). Unfortunately the high protein content was lost during back crossing to recover the desired characteristics and high root yield of Cassava (Asiedu *et al.*, 1992). Hence, there is need of alternative measures of improving the genetic basis of nutrition of this crop.

Cassava roots undergo rapid post harvest physiological deterioration which causes blue/black discoloration of the vascular tissues from the synthesis of simple phenolic compounds as a result of wound response (Wheatley and Chuzel, 1993). Cassava has the shortest post-harvest life compared to any of the other major root crops (Ghosh *et al.*, 1988). The roots become perishable and become inedible within 24 to 72 hours after harvest. This can cause up to 20% post harvest loss (Egesi *et al.*, 2007). Cassava post-harvest deterioration prevents exportation of Cassava roots and reduces generation of income to farmers. Conventional breeding programs have been successful in making significant contributions to reducing post-harvest deterioration of Cassava (Cortés *et al.*, 2002).

Plant breeding is an indispensable tool in producing superior yielding Cassava (Hershey and Jennings, 1992). For any breeding scheme to be effective, knowledge on the floral biology of Cassava should first be considered. Cassava seeds are genetically diverse due to segregation and recombination from sexual reproduction (Halsey *et al.*, 2008). Traditional breeding programs in Cassava have been successful in introducing improved cultivars (Hershey and Jennings, 1992). However, the high degree of heterozygosity in this allotetraploid plant, long growing season, irregular flowering, low seed set, and variable germination rates have impeded faster progress via classical breeding (González *et al.*, 1998).

Unlike many of the world's major crop plants, genetic improvement of Cassava through sexual crosses has been difficult. Many varieties flower rarely and seed production is often low. Further, early flowering is associated with heavy branching, which tends to lead to low harvest index and yield (Cock *et al.*, 1979; Cock and El Sharkaway, 1988). Conversely high yielding genotypes either do not branch or branch late, and the first branches formed often do not produce fertile flowers (Ceballos *et al.*, 2002; Jennings and Iglesias, 2002). This leads to a dilemma for Cassava breeders. They must produce shy flowering types, but in order to reduce generation times, and produce many progeny through controlled crosses, they require profuse early- flowering types. Furthermore, synchronized flowering is essential to perform the cross. In the field, Cassava is typically propagated clonally by stem cuttings (Iglesias *et al.*, 1994). This vegetative propagation

---



---

strategy is ideal for molecular approaches to crop improvement, since an individual clone does not experience gene segregation through outcrossing. Cassava plants obtained from seeds produce genetically diverse offsprings, and stem cutting propagation produces clones of the same genotypic background (Iglesias *et al.*, 1994).

In light of the constraints described above, plant breeding and genetic transformation can be essential tools to produce improved Cassava genotypes. This should overcome the array of agronomic constraints that currently exists. For effective breeding programs, an understanding and manipulation of the timing of flowering in Cassava becomes an indispensable tool. This is notably true given that many Cassava varieties flower late, while some have been known never to flower (Alves, 2002). Further relevance occurs given the need for synchronization of flowering, because different Cassava varieties flower at different times in their breeding cycles. This makes conventional breeding difficult. The problem of the long growth cycle makes the evaluation of successful breeding programs prolonged, and this delays release of elite crops generated. It literally takes a decade to generate a new cultivar by the conventional breeding method (Fregene and Ponti-Kaerlas, 2002). Thus, an alternative floral manipulation strategy becomes critical to reduce breeding time.

### **1.8 Cassava *in vitro* Regeneration (Methods of Cassava transformation)**

Transformation provides an alternative for Cassava improvement. However, for successful manipulation of the Cassava genome by molecular-genetic transformation, an efficient and reproducible regeneration system of Cassava is required. The now established regeneration system of Cassava includes the use of the Friable Embryogenic Callus system (FECs) and Somatic Embryogenesis (SE) (Roca, 1984).

#### **1.8.1 Friable Embryogenic Callus**

Friable embryogenic callus are organized somatic embryos which are generated from leaf-lobe explants. They can be used to regenerate whole plants that are derived from single cells. The friable embryogenic callus (FEC) differentiates into thousands of homogeneous sub-millimeter sized, pro-embryogenic units per gram of tissue. (Taylor *et al.*, 1996; Taylor *et al.*, 2001). Many plants have been successfully regenerated from this embryogenic callus (Raemakers *et al.*, 1996; Schöpke *et al.*, 1996; Taylor *et al.*, 2001, González *et al.*, 1998; Zhang and Pounti-Kaerlas, 2000).

#### **1.8.2 Somatic Embryogenesis**

Somatic embryogenesis involves the production of embryo-like structures from young leaf lobes and cotyledons (Stamp and Henshaw, 1982, 1986, 1987a, b). These embryos develop further and germinate into plantlets through developmental steps that correspond to zygotic embryos. The

---

direct mode of somatic embryogenesis involves the formation of an asexual embryo from a single cell or a group of cells on a part of the explant tissue without an intervening callus phase. Plants have been successfully regenerated from these tissues (Stamp and Henshaw, 1982, 1986, 1987a, b; Stamp, 1987; Szabados *et al.*, 1987; Taylor and Henshaw, 1993; Mathews *et al.*, 1993; Raemakers *et al.*, 1993; Konan *et al.*, 1994b).

### **1.9 History of Genetic Transformation in Cassava**

In 1996, the first stable genetic transformations of Cassava was reported using *Agrobacterium tumefaciens* (Li *et al.*, 1996) and microparticle-mediated delivery of DNA to plants (Schöpke *et al.*, 1996). Additional reports of the genetic transformation of Cassava have followed (Sarría *et al.*, 2000; Zhang *et al.*, 2000a, b). However, only recently have transgenic plants been generated with enhanced agronomic traits (Zhang *et al.*, 2003, Ihemere *et al.*, 2006).

In the last 6 years, there have been several reports of genetically modified Cassava with enhanced agronomic traits. In 2003, Siritunga and Sayre reported the introduction of an antisense CYP79D1 and CYP79D2 construct into Cassava to suppress the expression of the cytochrome P450s that catalyse the first dedicated step in cyanogenic glycoside synthesis. Transgenic plants having less than 1% of the normal root cyanogens levels were generated. However, these plants were unable to grow without supplemental reduced nitrogen (Siritunga and Sayre, 2003). In the same year, Zhang *et al.* reported the misexpression of an artificial storage protein gene (ASP1) in Cassava leaves and roots to increase protein content (Zhang *et al.*, 2003). Its altered transgene expression had little effect on the overall amino-acid composition of leaf proteins. More recently, Siritunga *et al.* (2004) have reported the overexpression of hydroxynitrile lyase gene in Cassava tuberous roots, and this led to accelerated cyanogen removal and food detoxification. In another report, Ihemere *et al.* (2006), produced Cassava with 2.6-fold increase in total tuberous root biomass by altering the rate-limiting step in starch biosynthesis, overexpressing the *E.coli* gene *glg C*, encoding AGPase (Ihemere *et al.*, 2006). From these success stories, there is every reason to be hopeful that transgenic Cassava will be part of the solution to break breeding constraints.

#### **1.9.1 Promoters Choice in Plant Biotechnology**

Genetic manipulations in plant biotechnology involve the introduction of one or more transgenes which can turn on or off the desired trait in plants. A number of promoters have been used to transcribe the genes of interest. Among such promoters are constitutive promoter CaMV35S and chemical-inducible systems for gene-regulated expression (Zuo and Chua, 2000). The chemical-inducible system in plants is based on de-repression, inactivation, and activation of transcription of the target gene (Zuo and Chua, 2000). One class of promoters of particular interest is the activation of transcription of the target gene which can be directed to “turn on” genes of interest. One such inducible promoter system will be described.

---

The ethanol-utilization pathway of the filamentous *Aspergillus nidulans* is a well characterized positive-operon system. It controls the cellular response to ethanol and related alcohol and ketones (Felenbok *et al.*, 1988, Felenbok, 1991). The first enzyme in the path of ethanol utilization is alcohol dehydrogenase I (Adh-1), which is encoded by the alc gene (Felenbok, 1991). The transcription activator protein AlcR binds target sequences within the alcA gene promoter in the presence of ethanol, or other alcohols or ketones. These compounds act as inducer of gene expression.

#### **1.9.1.1 The AlcA promoter system transferred into plants**

The ethanol system was transferred into plants as a two-component inducible system. This requires the alcR gene, which encodes the transcriptional activator protein AlcR, and the alcA promoter (Lockington *et al.*, 1987). The alcR gene can be placed under the control of a strong constitutive promoter, such as CaMV 35S. Thus, the transcriptional factor for induction is continuously present. In the absence of the ethanol inducer, the transcriptional protein AlcR cannot bind the specific sequences of the modified alcA promoter, which is linked to a gene of interest. In the presence of ethanol, the transcriptional protein AlcR is activated and binds to the alcA promoter which drives the expression of the gene of interest to which it is fused (Kulmburg *et al.*, 1992).

The AlcR/AlcA system has an advantage over a wide variety of other inducible- promoter systems of plants is that:

1. Ethanol is a simple alcohol, which is inexpensive and readily available, and its toxicity is mild.
2. Induction of the system can be achieved by low doses of ethanol.
3. Under normal growth conditions, the levels of the natural inducers (ethanol) and related alcohols produced are low, such that they do not induce alcR expression.
4. Expression can be induced at different stages of development.

The use of this gene switch has previously been described in tobacco (*Nicotiana tabacum*), *Arabidopsis thaliana*, potato (*Solanum tuberosum*), tomato (*Lycopersicon esculentum*), oil-seed rape (*Brassica napus*), and Poplar (Caddick *et al.*, 1998; Salter *et al.*, 1998; Roslan *et al.*, 2001; Sweetman *et al.*, 2002; Garoosi *et al.*, 2005; Filichkin *et al.*, 2006). From the success of the use of the ethanol inducible systems in these biological systems, it is hopeful that the AlcR/AlcA will work well in Cassava.

### **1.10 Flowering Control Pathways in Arabidopsis**

Flowering and its control is one of the most important developmental aspects of crop plants. Control of flowering in *Arabidopsis* has been extensively studied and several pathways have been implicated in the control of flowering. These are photoperiodic, vernalization, gibberellin and

---

autonomous pathways (Blazquez *et al.*, 1998, 2002; Levy *et al.*, 2002; Yanovsky and Kay, 2003; Hayama and Coupland, 2004; Searle and Coupland, 2004; Amasino, 2005). The inductive ability of the photoperiod pathway to promote flowering in “shy” genotypes is particularly noted (Jaeger *et al.*, 2006).

### 1.10.1 Day length/Photoperiod in Arabidopsis

Plants in nature perceive light-dark stimulus as a result of the rotation of the earth around its axis over a period of 24 hours of the day, which causes day and night as a result from the position of the earth's surface facing towards or away from the sun. This light-dark transition is a potent environmental cue that affects angiosperm flowering time (Garner and Allard, 1920 and 1923). Arabidopsis is a facultative long day plant responsive to inductive long day photoperiod, which causes it to flower. The timing of floral transition is attenuated under short days (Searle and Coupland, 2004).

Various Scientists have proposed that the measurement of photoperiod was dependent on an endogenous diurnal rhythm. The first published accounts on this was that from the French Astronomer, Jean Jacques d'Ortous Marian (1729), who was intrigued by the daily folding of the leaves and leaflets of the ‘sensitive’ mimosa plants during the night and their re-opening during the day to study whether this biological behavior was as a response to sunlight. Placing the plants in constant darkness, he observed the opening and the closing of the leaves persisted as though the plants were perceiving/seeing day and night. He concluded the involvement of an internal rhythm (de Marian, 1729). This internal rhythm is termed as the circadian clock from the Latin *circa*, and *dies*, which means "approximately one day." Modern photoperiod research extends from this, and it is now understood that flowering time of many plant species is regulated by day length (Bünning, 1960 and Searle and Coupland, 2004).

Transition to flowering is one of the most important phases of development. In Arabidopsis, flowering is regulated by an integrated network of several genetic pathways which monitor both the developmental state of the plants and environmental cues, such as day length, light and temperature. The photoperiodic flowering is known to be regulated by *GIGANTEA (GI)*, a circadian clock-controlled gene (Fowler *et al.*, 1999). *GI* is placed between the clock and *CONSTANS (CO)* (Mizoguchi *et al.*, 2005). *CO* plays a central role in the photoperiodic promotion of flowering under long day conditions. It was shown that *CO* acts between the circadian clock and genes controlling meristem identity (Samach *et al.*, 2000; Suárez- López *et al.*, 2001). The flowering time genes *FLOWERING LOCUS T (FT)* and *SUPPRESSOR OF OVEREXPRESSION OF CO1 (SOC1)*, together with the floral meristem identity gene *LEAFY (LFY)*, are three essential regulators integrating floral signals from multiple pathways in Arabidopsis (Yu *et al.*, 2002). The interaction among these genes is mediated by *AGAMOUS-*

---

*LIKE 24 (AGL24)*, a putative transcription factor which is activated in the shoot apical meristems during the floral transition (Yu *et al.*, 2002). Together, it is clear that a number of floral inducers exists in Arabidopsis.

Interestingly, Kardailsky *et al.*, 1999, reported that *FT* and *SOCI* are among the most potent activators of flowering, by causing extreme flowering when overexpressed. *FT* and *SOCI* are at the point of convergence of several flowering-time pathways and are often therefore described as floral integrators (Mouradov *et al.*, 2002; Simpson and Dean, 2002). Corbesier *et al.*, 2007, demonstrated a strong correlation between the expression of *CO* in the light, increased expression of the downstream gene *FT*, and early flowering. This has been coupled also with the evidence provided by Corbesier *et al.*, 2007, that *FT* protein is a major component of florigen and that the gain-of-function of *FT* leads to early flowering and furthermore, its loss of function results in late flowering. More recently *FT* has been shown to be mobile in its floral inductive abilities (Jaeger and Wigge, 2007; Giakountis and Coupland, 2008). Taken together, a linear flowering pathway has been proposed, and *FT* is a key factor of the photoperiod pathway that ultimately functions as a positive factor in the expression of identity genes.

### **1.11 Photoperiod Perception In Cassava**

Flowering and photoperiodism in Cassava has not been a major focus of study. At present there is no strong research entity which focuses its attention on flowering of tropical species, and the applied research agencies that are required to respond to requests for systems to control flowering of tropical species have only a limited capacity to respond. Given the similarities in the floral-inductive pathways of Arabidopsis and rice (Andersen *et al.*, 2004), gene sequences from Arabidopsis show promise to manipulate floral timing in Cassava.

### **1.12 Aim of thesis**

My objectives for conducting this research were to isolate and characterize Cassava photoperiodic genes and as a consequence, modify flowering time in Cassava with the Arabidopsis mobile and graft transmissible floral promoter, *FT*. It is hoped that my success will consequently improve food security in the sub Saharan community by creating a genetic system that would be useful for inducing floral initiation at will for the purpose of conventional breeding. This might open up “locked,” but desirable genes in Cassava, and as a consequence, lead to increased food supply and income in the rural areas.

---

## **CHAPTER 2. MATERIALS AND METHODS**

---

## 2.1 Materials

### 2.1.1 Chemicals

Laboratory grade chemicals were purchased from Amersham Biosciences (Freiburg), Merck (Darmstadt), Fluka (Neu-Ulm), Sigma-Aldrich (Taufkirchen), Serva (Heidelberg), Duchefa (Haarlem, Netherlands), and Roth (Karlsruhe), Invitrogen (Karlsruhe).

### 2.1.2 Enzymes

Restriction enzymes were purchased from New England Biolabs (Frankfurt a.M) and Fermentas (St.Leon-Rot).

Other enzymes:

Taq DNA polymerase Peqlab (Erlangen)

Pfu Turbo HotStart DNA polymerase Stratagene (Heidelberg)

T4 DNA ligase Fermentas (St. Leon-Rot)

Klenow enzyme Roche (Mannheim)

DNaseI Roche (Mannheim)

RNaseH Fermentas (St. Leon-Rot)

RNase inhibitor (Roche)

Superscript II Reverse Transcriptase Invitrogen (Karlsruhe)

BP-clonase (GATEWAY®-Technology) Invitrogen (Karlsruhe)

LR-clonase (GATEWAY®-Technology) Invitrogen (Karlsruhe)

### 2.1.3. Bacterial Strains

#### 2.1.3.1 *Escherichia coli*

DH5 $\alpha$ , (Invitrogen, Karlsruhe) genotype supE44 lacU169( $\phi$ 80, lacZ M15) hsdR17 recA1 endA1 gyrA96thi-1 relA1

XL 10-Gold (Stratagene)

DH5 (Invitrogen, Karlsruhe)

DB3.1 (Invitrogen, Karlsruhe) RR1 gyrA462 endA (recA-)

XL10-Gold (Stratagene, Heidelberg)

#### 2.1.3.2 *Agrobacterium tumefaciens*

ABI (Davis *et al.*, 2009)

GV3101 strain with pMP90RK (Koncz and Schell, 1986)

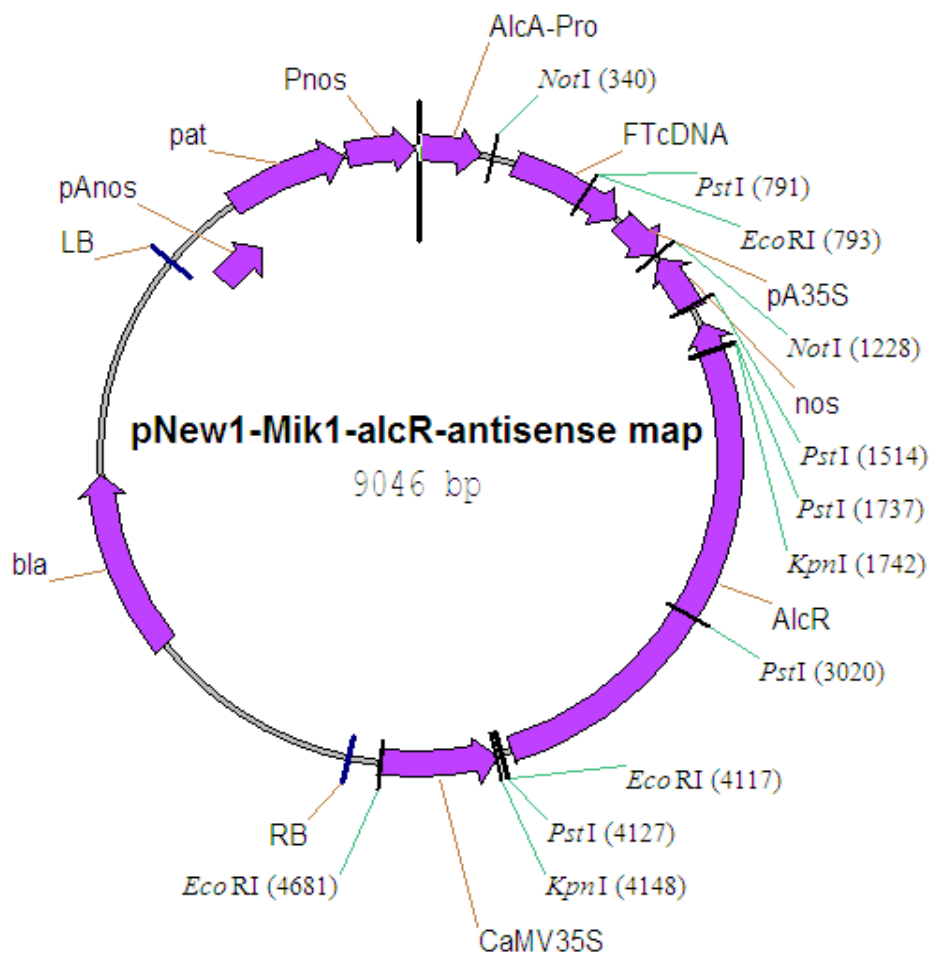
---

#### 2.1.4. Vectors

pDONR207 Gateway Cloning Donor vector (Invitrogen, Karlsruhe)

pNew-Mik1-antisense GATEWAY-compatible plant expression to express cDNA under the control of CaMV35S promoter and an ethanol inducible system (Bekir Ülker, MPIZ). The pNew-FT vector was obtained from LR reaction

pJalee4-Destination vector harboring the *ELF4* promoter.



**Figure 2.1 The Vector Map of pNewFT used for the transformation of both Arabidopsis and Cassava.** The AlcR transcription factor under the constitutive promoter is activated in the presence of ethanol, binds to the AlcA-Pro, the promoter which drives the expression of the *FT* cDNA fused to it.



### 2.1.5. Oligonucleotides

5'-3' Oligonucleotides were synthesized by Invitrogen (Karlsruhe) and Sigma (Steinheim).

**Table 1.1:** Primer sequences designed for cloning, genotyping, and gene discovery.

<b>Gene</b>	<b>Primer Name</b>	<b>Primer sequence 5'- 3'</b>
<i>MeELF4</i>	GFWF-elf4fwprimer	GGGACAAGTTTGTACAAAAAAGCAGGCTCCATGA ACAACAATCCAATCACAAATC
	GWRV-elf4rvprimer	GGGACCACCTTTGTACAAGAAAGCTGGGTCTTACGA CCTTCCACTACTGTT ACTGTTGCCG
<i>AtFT</i>	FTGFWF	GGGACAAGTTTGTACAAAAAAGCAGGCTCCATGT CTATAAATATAAGAGACCCTC
	FTGWRV	GGGACCACCTTTGTACAAGAAAGCTGGGTCTAAAGT CTTCTTCCTCCGCAGCCA
	FTf1	CCAAGTCCTAGCAACCCTCA
	FTr1	TACTGTTTGCCTGCCAAG
	L33S-FTSBF1	AAGCAGAGTTGTTGGAGACG
	R508-FTSBR1	AGCCACTCTCCCTCTGACAA
	L104-FTSBF1	GAGAGGTGACTAATGGCTTGG
	L17-FTF1	CTATGAAAGCCCACGACCAT
	R205-FTR1	CCATCGGTGGGAATTCATAG
<i>AtTUBULIN</i>	TUBF	ACTCGTTGGGAGGAGGAACT
	TUBR	ACACCAGACATAGAGCAGAAATCAAG
<i>MeCO</i>	CO51	GCGGTTGCAGCAAGAAAGGAGAAAGG
	CO52	GGAGAAAGGC AAAATCTGAGG
	CO53	GAACTATAAAAAAGCTACACAACCATAG
	CO54	CAACCATAGACTAGAAATTCT
	CO31	ATCAAACACAATACTTAGATGTGTATAG
	CO32	GATGTGTATAGCTCATATAA
	CO33	CTCCATTTCAATCAATGGATGGGGGAGTGG
	CO34	GGGGGAGTGGTACCTGATTC
	CO5A	GACTGCTCTCTGCTTTCCTCTCAGCTCCC
	CO5B	CTCTCAGCTCCCACTAAAAAAACCATTGCC
	CO5C	CTCCACCACCAAACAAGAAGCCATTATTGTTCTGC
	CO5D	GCCATTATTGTTCTGCCCCGTTGCTATTCTTCGC
	CO3A	CCTTTGAAACTATATTTTGATGCAGGACTTAG

<i>MeCO</i>	CO3B	GCAGGACTTAGAATGATGGGACAACGCCA
<i>continued</i>	CO3C	TGCAGAAGCAGCAAGAAAAGATCAACTGCATC
	CO3D	GCATCATCAATACCATAATTTCCACTTGGGGTTGG
<i>MeGI</i>	GI5 1	GGCTTGTATGCATGCTCTCTCTGTTTTG
	GI5 2	CTGTTTTGATGCGGTGCAAG
	GI5 3	TGCCATAGAAGTGAGACAACAGAGAAAC
	GI5 4	GAAACATAACTCCTGTTTCTC
	GI3 1	AATTTCTTAACAATGGACAGGCACATTG
	GI3 2	GCACATTGGATTNNNCTTCAG
	GI3 3	CTCTATATCAGCTTGCCAATCGATGACATC
	GI3 4	GATGACATCTACCTTGAATAAC
<i>MeAPI</i>	AP1 1	GGGCACAGAGAGAAGCTAGAAGCAG
	AP1 2	AGCAGCACGCGCTGCAACTG
	AP1 3	AACCTTCTTTGCAAGTTGGTTGTTTTG
	AP1 4	GGTTGTTTTGCTCCTGCAATG
<i>MeELF4</i>	EL41	GCCATTAATTACTGCCTGACCTCCCTC
	EL42	GACCTCCCTCTGCTCTACGC
	EL43	TTTATTCATTTGTTTATTCCAAATTGCG
	EL44	CCAAATTGCGAACAATAACCCCAT
<i>MeLHY</i>	LH51	GAGAACTCCTTTTCGATGCTCCTTTTTC
	LH52	GATGCTCCTTTTTCCTTCTTTC
	LH53	AAAAACTTCTGTGCATGACTCCTGATC
	LH54	CCTGATCTGCACAGCAGTCTTTG
	LH31	ACTGGCTTTTCGAGCTCTCTTCTCCAG
	LH32	CTCTTCTCCAGGGAGATATTG
	LH33	TTTTTTTAAACAAACACTGCCTTAG
	LH34	CACTGCCTTAGAAAGAAAAGGATTAT
<i>MeSOC1</i>	SOC1	ATCAGATTTTCAGTTACTTCTTGC
	SOC2	CTTCTTGCTTTATCTTTCTTGTTTTGG
	SOC3	TGCAGGCACCCAACTTCTTCCCAGTAG
	SOC4	CCAGTAGTTTCTTTTTGCAATTTT
<i>RcFT</i>	FT51	TTAGGGATCCTCTTGTGTTGGGGG
	FT52	TCTTGTGTTGGGGGTGTGATAGGA
	FT53	AGGAGAAGGAGAGTGGAGATGATCA

<i>RcFT</i>	FT54	GGAGATGATCATTACCAAAGTATA
<i>continued</i>	FT31	TTTAACTAAATCACTTATCATCTAT
	FT32	ATCATCTATTTGTTGTTAATATATC
	FT33	CTAAATTAATAAAGACTCAAATTAA
	FT34	TCAAATTAATAATATTGAATCAAATTATC
<i>RcCO</i>	Co51	AGAGAAGCAGCATCTGCCTTGACAA
	Co52	TCTGCCTTGACAAAGAAGGCAGCCGG
	Co53	ACGAAACTCGTAGTAATCATGTTG
	Co54	ATCATGTTGAAAGAAGAGAACACTG
	Co31	GAAGGCTTAGAATGATGGAACAATGC
	Co32	AATGATGGAACAATGCCATATCCAG
	Co33	CACAGCTCAGTCCAAGGGACAGGGAGGCA
	Co34	ACAGGGAGGCAAGGGTTCTAAGATAC
<i>RcGI</i>	G1F1	ATATTTACCGATGAAATAAACGTGC
	G1R1	TAAACGTGCATATGGCTCCAGAGC
	G2F1	TTCATCTCATGGATCGGGGAAG
	G2F2	GATCGGGGAAGCATCCTCAACTCATGC
	G3F1	TAACTATTGATGCAACAGCTTTGCTGTG
	G3F2	CAGCTTTGCTGTGAATATCAATGAG
	G3R1	AAAGATAAAAAACAAGGTCCTGTTGCAG
	G3R2	CTGTTGCAGCATTGATTCTTATGTTCTTG
	G4F1	TGTCAAATGGATATAGTACAACCTAATTC
	G4R1	GATATAAATGGCATTTCGTGGTCCATCATATC

---

## 2.1.6. Media

### 2.1.6.1. Media for bacteria

Media for bacteria were prepared as described (Sambrook and Russell, 2001).

<b>Luria Bertani (LB)</b>	<b>YEBS</b>	<b>Circle Grow-purchased from Qbiogene</b>
10 g/L bactotryptone 5 g/L yeast extract 5 g/L NaCl 1% agar pH 7.5	5 g/L beef extract 5 g/L peptone 5 g/L sucrose 1 g/L yeast extract 0.5 g/L MgSO <sub>4</sub> 1% agar pH 7.0	<b>Antibiotics</b> Ampicillin 100 mg/ml Carbenicillin 100 mg/ml H <sub>2</sub> O Gentamicin 25 mg/ml H <sub>2</sub> O Kanamycin 100 mg/ml H <sub>2</sub> O Phosphinothricin (PPT) 12mg/ml H <sub>2</sub> O Rifampicin 100 mg/ml methanol Spectinomycin + Streptomycin 50 mg/l each (100 mg/l in total)

### 2.1.6.2. Media for plants

#### Arabidopsis

<b>MSO - Murashige and Skoog (MS) growth media</b>	<b>MS3</b>
2.2 g/L MS salt, 0.5g/L MES, 2-(N-Morpholino) Ethane Sulfonic acid 1.2% phytoagar (Duchefa) pH 5.7 adjusted with KOH	4.4 g/L MS 0.5 g/L MES 30 g/L sucrose 1.5% phytoagar pH 5.7 adjusted with KOH

---

## Cassava

All MS salts (Murashige and Skoog, 1962) have same stock concentration, except when otherwise stated.

### Murashige & Skoog's medium salts Stocks (Murashige and Skoog, 1962)

<b>MS Stock Solution I (X 50)</b>	<b>MS Stock Solution 2(X 1000)</b>
8.5 g/L KH <sub>2</sub> PO <sub>4</sub> 95 g/L KNO <sub>3</sub> 18.5 g/L MgSO <sub>4</sub> ·7H <sub>2</sub> O 82.5 g/L NH <sub>4</sub> NO <sub>3</sub>	25 mg/L CoCl <sub>2</sub> ·5H <sub>2</sub> O 25 mg/L CuSO <sub>4</sub> ·5H <sub>2</sub> O 6.2 g/L H <sub>3</sub> BO <sub>3</sub> 16.9 g/L MnSO <sub>4</sub> ·H <sub>2</sub> O 250 mg/L NaMoO <sub>4</sub> ·2H <sub>2</sub> O 8.6 g/L ZnSO <sub>4</sub> ·7H <sub>2</sub> O
<b>MS Stock Solution 3(X 1000)</b>	<b>MS Stock Solution 4 (X 100)</b>
830 mg/L KI	43.9 g/L CaCl <sub>2</sub> ·2H <sub>2</sub> O
<b>MS Stock Solution 5</b>	<b>MS Vitamins ( X 100)</b>
7.45 g/L Na <sub>2</sub> -EDTA 5.57 g/L FeSO <sub>4</sub> ·7H <sub>2</sub> O	200 mg/L Glycine 10 g/L myo-inositol 50 mg/L nicotinic acid 50 mg/L pyridoxine HCl 10 mg/L thiamine HCl

<b>Embryo Induction</b>	<b>MS2-IBA-BAP (Organogenesis)</b>	<b>MS4-Somatic Embryo Cotyledon Transformation</b>
20ml/L Stock 1 1ml/L Stock 2 1ml/L Stock 3 10ml/L Stock4 5ml/L Stock 5 10ml/L Vitamin B <sub>5</sub> 250µl/L Picloram 2% Sucrose 1ml/L 2mM CuSO <sub>4</sub> 50mg/L Casein hydrolysate pH 5.6-5.7 0.2% Gel rite	20ml/L Stock 1 1ml/L Stock 2 1ml/L Stock 3 10ml/L Stock4 5ml/L Stock 5 MS Vitamins 100mg/L M-inositol 1ml/L Vitamin Stock 1000X 2% Sucrose 1ml CuSO <sub>4</sub> 500µl BAP Stock 2000ppm 500µl IBA Stock 1000ppm pH 5.8 0.6% Agar	20ml/L Stock 1 1ml/L Stock 2 1ml/L Stock 3 10ml/L Stock4 5ml/L Stock 5 1ml/L MS Vitamins 100mg/L Inositol 250µl/L 4mg 2,4-D 2% Sucrose 1ml/L 2µM CuSO <sub>4</sub> 50mg/L Casein hydrolysate pH 5.6-5.7

<b>MS2-BAP (Maturation Medium)</b>	<b>MS2-1µM NAA</b>	<b>MS3-Activated charcoal</b>
<b>MS Salts</b> 20ml/L Stock 1 1ml/L Stock 2 1ml/L Stock 3 10ml/L Stock 4 5ml/L Stock 5 MS Vitamins 100mg/L M-inositol 1ml/L Vitamins without inositol Sucrose 2% 1ml 2µM CuSO <sub>4</sub> 50µl 0.1mg BAP pH 5.8 0.45% Agar	<b>MS Salts</b> 20ml/L Stock 1 1ml/L Stock 2 1ml/L Stock 3 10ml/L Stock 4 5ml/L Stock 5 MS Vitamins 100mg/L M-inositol 1ml/L Vitamins without inositol 2% Sucrose 0.465ml α-NAA (naphthalene acetic acid) pH 5.6-5.7 Agar: Gel rite 3:1	<b>MS Salts</b> 20ml/L Stock 1 1ml/L Stock 2 1ml/L Stock 3 10ml/L Stock 4 5ml /L Stock 5 1ml Vitamin B <sub>5</sub> 100mg/L myo-inositol 3% Sucrose 1.4g/L Glutamine 5.0g/L Activated carbon 200mg/L Casein hydrolysate pH 5.6-5.7 3.0g/L phytigel

---

**GRESSHOFF AND DOY (GD) MEDIUM (1972) Basal Salt Mixture**

<b>Macro 100 X Stock Solution in g/L</b>	<b>Micro 1000X in mg/l</b>	<b>Vitamins GD 100 X mg/L</b>
NH <sub>4</sub> NO <sub>3</sub> 100g/l KNO <sub>3</sub> 100g/l MgSO <sub>4</sub> ·7H <sub>2</sub> O 3.49g/l KH <sub>2</sub> PO <sub>4</sub> 30g/l KCl 6.5g/l	300mg/L H <sub>3</sub> BO <sub>3</sub> 1000mg/L MnSO <sub>4</sub> ·H <sub>2</sub> O 300mg/L ZnSO <sub>4</sub> ·H <sub>2</sub> O 25mg/L Na <sub>2</sub> MoO <sub>4</sub> ·2H <sub>2</sub> O 25mg/L CuSO <sub>4</sub> ·5H <sub>2</sub> O 25mg/L CoCl <sub>2</sub> ·6H <sub>2</sub> O 800mg/L KI	400mg/L Glycine 100mg/L pyridoxine HCl 10000mg/L myo-inositol 1000mg/L thiamine HCl 100mg/L nicotinic acid
<b>Ca(NO<sub>3</sub>)<sub>2</sub>·4H<sub>2</sub>O 100 X</b> 24.130g/l		
<b>GD2-50Pi</b>		
10ml/L Macro 10ml/L Ca(NO <sub>3</sub> ) <sub>2</sub> 1ml/L Micro 36.7mg/L FeNaEDTA 10ml/L Vit. GD 100X 2% Sucrose 250 µl/L 50µM Picloram Stock 200mM pH 5.7-5.8 4.5g/L Agar		

<b>4E- Propagation Medium MS Salts</b>	<b>17N-Rooting Medium</b>
20ml/L Stock 1	1/3 MS Salts
1ml/L Stock 2	6.66ml/L Stock 1
1ml/L Stock 3	0.33ml/L Stock 2
10ml/L Stock 4	0.33ml/L Stock 3
5ml/L Stock 5	3.33ml/L Stock 4
10ml/L Thiamine 1ppm	1.66ml/L Stock 5
Stock 1000ppm	2% Sucrose
100mg/L Myo-inositol 100ppm	10ml/L Thiamine (100ppm)
2% Sucrose	100mg/L Inositol
40µl BAP 0.04ppm Stock 1000ppm	25.0µl/L NAA (0.01mg) 400ppm
55.5µl GA <sub>3</sub> 0.05ppm Stock 900ppm	10.0µl/L GA <sub>3</sub> (0.01mg) 1000ppm
20µl NAA 0.02ppm Stock 1000ppm	25mg/L Plant Product pp 10:52:10
1ml/L CuSO <sub>4</sub> 2µM Stock 2mM	pH 5.7-5.8
pH 5.7-5.8	4.5g/L Agar Duchefa
0.45% Agar Duchefa	

### **2.1.7 BUFFERS AND OTHER REAGENTS**

#### **DNA EXTRACTION BUFFER**

##### **CTAB**

140mM Sorbitol

220mM Tris-HCl pH 8.0

22Mm EDTA pH 8.0

800mM NaCl

1%<sup>w/v</sup> Sarkosyl

0.8% CTAB pH 8.0, autoclaved 121°C for 20mins

##### **TE**

10mM Tris-HCl

1mM EDTA pH 8.0

#### **MEDIA FOR GEL ELECTROPHORESIS**

##### **10X TAE**

48.4g/L Tris Base



---

11.42ml of glacial acetic acid  
20ml of 0.5M EDTA,  
Add up to 1000ml with dH<sub>2</sub>O  
adjust pH to 8.5 with KOH

***25X TBE buffer***

67.23 g/L Tris  
34.31 g/L boric acid  
37.22 g/L EDTA  
pH 8.0

***DNA visualization dye***

Ethidium bromide (Stock 10 mg/ml in H<sub>2</sub>O)

***SOLUTIONS FOR SOUTHERN BLOT***

***Depurination solution***

11ml HCl  
989ml of distilled water

***Denaturation Buffer***

87.66g/L NaCl  
20g NaOH/L

***Neutralization Buffer***

87.66g/L NaCl  
60.5g/L Trizma base  
Adjust pH to 7.5 with concentrated HCl

***Nucleic acid transfer buffer (20X SSC)***

88.23g/L Tris-Sodium Citrate  
175.32g/L NaCl  
pH 7-8

***100X Denhardt's Solution***

2.0g/100ml Bovine Serum Albumin  
2.0g/100ml Ficoll 400

---

---

2.0g/100ml Polyvinylpyrrolidone

Stored at -15 °C to -30 °C

### **10% SDS**

100 g SDS in 900 ml distilled H<sub>2</sub>O

### ***Hybridization Buffer***

5 x SSC

5 x Denhardt's solution

0.5% <sup>w/v</sup> SDS

### ***Bleach solution***

33% (v/v) KLORIX® (commercial sodium hypochlorite solution) in 0.02% (v/v) Triton X-100

### ***Firefly D-luciferin 50 mM stock***

1 g firefly D-luciferin(D-[4,5-dihydro-2-(6-hydroxy-2-benzothiazolyl)-4-thiazole-carboxylic acid] (LABTECH INTERNATIONAL) was dissolved in 71.3 ml 1 M triphosphate buffer (Na<sub>2</sub>HPO<sub>4</sub>/NaH<sub>2</sub>PO<sub>4</sub>) pH 8.0 to give a 50 mM luciferin solution. 1.5 ml aliquots were stored at -80°C. The luciferin stock was diluted to a 5 mM luciferin working solution with 0.01% (w/v) Triton-X100.

## **2.2 Methods**

### **2.2.1 Plant Material**

#### **Arabidopsis**

Arabidopsis lines *elf4-1* was in Ws-2 ecotype (Doyle *et al.*, 2002)

*FLC/FRI* was in the Columbia (Col-0) background (Michaelis and Amasino, 1999)

*CCA1::LUC* and *CCR2::LUC* was in Ws-2 ecotype (Doyle *et al.*, 2002)

*ft-10* was in the Columbia background (Yoo *et al.*, 2005)

#### **Cassava**

All Cassava lines were regenerated *in vitro* and the wild-type controls were from the Genetic Resource Unit, CIAT, Colombia.

---

## 2.2.2 Molecular Biology Techniques

### ***Cloning***

*E. coli* cells were transformed by heat shock at 42°C for 30 secs according to the manufacturers' instructions (INVITROGEN). Transformation of *Agrobacterium* was performed by electroporation, according to Wen-Jun and Forde (1989). Briefly, 1 µl of plasmid DNA (~100ng/µl) was used to transform 50 µl of cells. Transformation was performed at field strength of 12.5 kV/cm, at capacitance of 25 µF and resistance of 400 to 600 ohms for 12msec in a 1mm cuvette. Transformed bacteria were selected on YEBS agar plates containing appropriate antibiotics. Plates were incubated at 28°C for 24 to 72 hours, at which time colonies were visible.

### ***GATEWAY® constructs***

The 35S promoter fragment of the binary vector pJawohl (Kan<sup>R</sup>, gift from Bekir Ulker, MPIZ) was replaced with the *ELF4* promoter using *ClaI* and *AscI* sites to create pJawohl/ELF4p. Subsequent restriction with *AscI* and *SpeI* enabled exchange of the promoter-GATEWAY® cassettes of pJawohl/ELF4p and pLeela (Basta<sup>R</sup>, gift from Marc Jakoby, MPIZ), to give pJalee4 (Kolmos thesis, 2007).

### ***GATEWAY BP reaction:***

attB-PCR product 25 fmol  
GATEWAY, BP clonase 1 µl  
BP reaction buffer (5X) 1 µl  
Destination vector (150 ng/µl) 0.5 µl  
dH<sub>2</sub>O to 5 µl

### ***LR reaction:***

Entry clone (50 ng/µl) 1.25 µl  
GATEWAY, LR clonase 0.5 µl  
LR reaction buffer (5X) 0.5 µl  
Destination vector (50 ng/µl) 0.25 µl  
Reactions were carried out in room temperature between 1 hour to overnight incubation.

### ***E.coli Plasmid Isolation***

Qiagen QIAprep Spin Mini preparation Kit was used according to the instruction manual.

---

### ***Plant DNA Manipulation***

Qiagen DNeasy Plant Mini Kit was used according to the instruction manual.

### ***Polymerase Chain Reaction (PCR)***

The basic PCR performed was with Taq polymerase (PeqLab, Erlangen)

2µl 2.5 mM dNTP mix

2µl Buffer S (PeqLab)

4µl Enhancer solution (PeqLab)

1.6µl 25 mM MgCl<sub>2</sub>

0.5µl 10µM forward primer

0.5µl 10µM reverse primer,

0.1µl (5U/µl) Taq polymerase

7.3µl dH<sub>2</sub>O for a total reaction volume of 20µl

### ***PCR Thermal profile***

Initial denaturation step of 94°C for 2mins

Denaturation at 94°C for 15 sec

Annealing 50-60°C for 30 sec

Extension 72°C for 30s-2 min

Final extension time at 72°C for 5-10 min

PCR products were separated on TBE-agarose gels containing ethidium bromide and visualized using the BIO-RAD GEL DOC 2000 system (QUANTITY ONE, version 4.6.2 software). The denaturation to extension steps were repeated up to 29-39 times.

### ***PCR fragment extraction from Agarose Gels***

Qiagen QIAquick PCR purification kit was used, according to the instruction manual.

### ***DNA sequencing***

Sequencing was performed by the MPIZ DNA core facility (ADIS) on Applied Biosystems (Weiterstadt, Germany) ABI Prism 377, 3100 and 3730 sequencers using BigDye-terminator chemistry.

### ***RNA Extraction Procedure***

Qiagen RNeasy Kit was used, according to the instruction manual.

---

### ***cDNA synthesis***

2 µg of total RNA was used for cDNA synthesis. Prior to the synthesis, RNA was treated with 10U/µl DNase I (Roche) with DNase 1 Buffer (MgCl<sub>2</sub> included) and incubated at 37°C for 30 minutes to remove any residual genomic DNA. The reaction was continued as follows:

1µl 25mM EDTA, 10mins incubation at 65°C

1µl 0.5µg oligo(dT)<sub>12-18</sub>, 10mins incubation at 65°C

Incubate at least 1 min on ice

4 µl of 5 x RT buffer (Invitrogen)

1 µl of 10 mM dNTP mix

2 µl of 0.1 M DTT

0.5 µl of RNase inhibitor 40U/ul (Roche)

Incubation at 42°C, 2 min

1 µl of Superscript II RT (200U/µl) (Life Technologies, Rockville, MD)

Incubation at 42°C, 60 min

Incubation at 70°C, 15 min

90µl dH<sub>2</sub>O was added to the reaction

### ***Reverse transcriptase -PCR (RT-PCR) analysis***

10µl iQ-SYBR Green Supermix (Biorad), 2µl 10µM forward primer, 2µl 10µM reverse primer, 4µl dH<sub>2</sub>O and 2µl cDNA.

All PCR analyses were performed with the iCycler iQ5<sup>TM</sup> Multicolor Real-time PCR Detection System (Bio-Rad), according to manufacturers instructions.

### ***DNA Manipulation for Southern Blot***

Genomic DNA was isolated from young leaves of greenhouse-grown plants, according to the CTAB protocol, with modifications based on Doyle and Doyle 1987.

Three to five grams of leaf tissue was frozen and ground to fine powder, and then 20 ml of CTAB buffer was added to the powder at room temperature. The ground tissue was incubated at 65°C for 20 minutes, with occasional mixing by inverting. This was followed by addition of 10 ml of Chloroform. This mixture was shaken for 20 minutes at room temperature. Then the sample was centrifuged for 10 minutes at 2,500 g. To the aqueous phase, in a fresh 50 ml tube, 17 ml of isopropanol was added. This was mixed and placed on ice 10 minutes and centrifuged 5 minutes. The liquid was drained off without drying the pellet. To this, 4 ml of TE was added and the pellet were dissolved by gentle shaking. The solution of DNA was precipitated again with 4 ml of lithium acetate, this mixture was incubated on ice for 20 minutes. It was centrifuged for 10 minutes and the supernatant was transferred to a fresh tube. To this solution, 16 ml of 100%

---

ethanol was added, and this mixture was incubated on ice for another 20 minutes. It was then centrifuged for 5 mins and the precipitate was collected. The liquid was drained off and 4 ml of TE was added to resuspend the pellet in solution. RNase (100 mg/ml) was added to digest any remaining contaminant RNA. To remove the protein, first 400  $\mu$ l of sodium acetate was added and this mixture was transferred to a 15 ml tube containing 4.5 ml phenol. The mixture was inverted and then centrifuged for 10 minutes. 4.5 ml phenol:chloroform (1:1), was added to the aqueous layer, and this was mixed by inverting. The sample was then centrifuged for 10 minutes. To the aqueous layer from the preceding centrifugation step, 4.5 ml of chloroform was added. This was then mixed and then the sample was centrifuged at 2500 g for 10 minutes. The supernatant was transferred to a fresh 50 ml tube and 2 volumes of 100% ethanol was added. This was incubated on ice for 20 minutes. The incubated mixture was centrifuged at 2500 g for 15 minutes to collect the DNA as the precipitant. The liquid was drained off, the tube was inverted and the pellet was collected onto a piece of clean parafilm. The pellet was then carefully placed into a 1.5 ml eppendorf tube. This was again centrifuged for 5 minutes at 15,000 g. The residual liquid was carefully removed and the pellet dried with vacuum desiccator for about 10 minutes. The pellet was resuspended in 200  $\mu$ l TE. The concentration of the DNA was determined using the Pqlab Nanodrop ND-1000UV visible spektralphotometer, according to the manufacturer's instructions.

Southern-blot analysis of transgenic cassava and its wild-type was carried out according to the methods of Sambrook *et al.* (1989). Genomic DNA (20  $\mu$ g) was digested with *HindIII* (20U/ $\mu$ l) and 5U/ $\mu$ g was used for the digestion. The samples were separated by 0.8% (w/v) agarose gel, run first at 100 volts for 1 hour, and then at 30 volts overnight. This electrophoretic run was in 1 X TAE buffer. The gel was depurinated for 10-20 minutes with gentle agitation. This was followed by 30 minutes of denaturation, then 30 minutes of neutralization with gentle agitation. The gel was blotted to the Hybond-NX membrane (Amersham, GE Healthcare, Buckinghamshire,UK) by means of the capillary transfer, according to Southern (1975), with 20 x SSC used as the transfer buffer. The transfer was carried out over night at room temperature. The DNA was then immobilized to the membrane by crosslinking with Stratalinker UV crosslinker at 180,000 microjoules/cm<sup>2</sup>. The membrane was prehybridized in hybridization buffer (Amersham, GE Healthcare, Buckinghamshire,UK) for 1-2 hours at 42°C.

Following prehybridization, a probe was made by PCR amplification of the *AtFT* gene from the plasmid pNewFT (Figure 2.1), with the primers L104 and R508, giving a product size of 405 bp. This PCR product was purified with Qiagen PCR kit (Qiagen In, Valencia, CA). The PCR product was made up to 45  $\mu$ l with 1X TE to a concentration of 25 ng. This was boiled at 100°C for 5 minutes, quickly cooled on ice for 5 min after denaturation. The tube was briefly centrifuged, and this denatured DNA was added to the reaction tube (GE Healthcare Amersham Rediprime II Random Prime Labelling System). To this 5  $\mu$ l of [ $\alpha$ -<sup>32</sup>P] dCTP was added and mixed. This was

---

---

incubated at 37°C for 10 minutes and then the labeled DNA was denatured by heating at 100°C for 5 minutes. This was immediately cooled on ice for 5 minutes. The sample was centrifuged to collect and carefully added into the hybridization buffer. This hybridization was left for 1-2 days at 42°C.

After hybridization was completed, the membrane was washed twice with 2 X SSC, 0.1% (w/v) SDS for 5 minutes at 37°C, a single wash with 1 X SSC, 0.1% (w/v) SDS for 15 minutes at 37°C and finally with 0.1X SSC, 0.1% SDS for 2 X 10 minutes at 37°C. The membrane was dried at room temperature, wrapped with Saran wrap, placed a film within the cassette in the dark and exposed for 1-3 days at -80°C after which the signals were visualized after developing.

### ***ESTs searches***

BLAST searches (Altschul *et al.*, 1990) were made on both Cassava ESTs database in NCBI and that of Castor bean for using the Arabidopsis clock genes, *LONG HYPOCOTYL (LHY)*, *EARLY FLOWERING 3 (ELF3)*, *EARLY FLOWERING 4 (ELF4)*, the photoperiodic genes, *GIGANTEA (GI)*, and *CONSTANS (CO)*, their direct targets, the integrator genes, *FLOWERING LOCUS T (FT)*, *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1(SOC1)*, and *FD*, the direct target of *FT*, the floral meristem identity genes *APETALA 1 (API)*, *AGAMOUS- LIKE 24 (AGL24)*, *AGAMOUS* and *LEAFY (LFY)*, and the floral repressors, *FLC* and *FRI* (Onouchi *et al.*, 2000; Borner *et al.*, 2000; Samach *et al.*, 2000; Ratcliffe *et al.*, 2001; Kardailsky *et al.*, 1999; Kobayashi *et al.*, 1999; Blazquez *et al.*, 1997; Hepworth *et al.*, 2002; Lee *et al.*, 2000; Abe *et al.*, 2005). The sequence information obtained from the ESTs databases were used to design primers with the Primer 3 and the Invitrogen Vector NTI Advance<sup>TM</sup> 10 program (<http://frodo.wi.mit.edu/>). These primers were used for PCR against Cassava cDNA and genomic DNA. The sequences obtained from sequencing the fragments from the gel purified products of PCR amplification were first analyzed by comparing these sequences again with the Arabidopsis TAIR blast using the search “TAIR8 Proteins.” When these sequences were retrieved from BLAST searches, they are allocated an e-score which indicates the degree of similarity between the subject and the query. If the e-value is closer to 0, it shows a higher degree of similarity between the two sequences. Nucleotide sequences were used for this search against a protein database using a BLASTX search. This program translates the nucleotide sequence in all six reading frames and compares the output against all the translated sequences in the database. This program maximizes the potential for retrieving sequences similar to the gene of interest (Altschul *et al.*, 1997), which is the main reason for which amino-acid database were used in this study. These steps were carried out for all these genes of interest, except for those genes that were not amplified by PCR or those whose sequences with the highest e-value showed no similarity to the corresponding gene of interest within the Arabidopsis database. The sequence output used for primer design were from only those sequences which were identified and recognized as corresponding or likely orthologous gene with those in the Arabidopsis database (TAIR-BLAST) (<http://www.arabidopsis.org/blast>).

---

---

### **Contig assembly and Phylogenetic analysis**

Candidate sequences were identified using the BASIC LOCAL ALIGNMENT SEARCH TOOL (BLAST) (Altschul *et al.*, 1990) in the Cassava ESTs database, Castor bean genome and other genome databases (DOE JOINT GENOME INSTITUTE, PHYSCOBASE, TIGR). Contigs were assembled from sequenced gene fragments using the program Vector NTI Advance 10<sup>TM</sup>. A consensus sequence was determined.

The derived protein sequences and partial sequences from public databases, were aligned using the multiple alignment tool CLUSTALW2 EMBL-EBI,

<https://www.ebi.ac.uk/Tools/clustalw2/index.html> (Larkin *et al.*, 2007), according to the Neighbour joining method (Saitou and Nei, 1987). Those, alignments were visualized after processing with Java<sup>TM</sup> 6 Standard edition. Trees were generated from the alignment outputs by calculating average distance using % identity.

### **2.2.3 Plant Procedures**

#### ***Seed sterilization and germination in Arabidopsis***

Small aliquots of seeds (up to 400 µl) were surface-sterilized. First the seeds were rinsed in 500 µl ethanol. After removal of ethanol, the seeds were incubated in 500 µl of bleach solution for 1-3 minutes. The bleach was removed and the seeds were rinsed with sterile water. The seeds were suspended in 0.1% agar water before they were plated on appropriate MS agar medium. The seeds were stratified at 4°C in the dark for 2-3 days before transferred to the appropriate growth cabinet.

#### ***Flowering time***

Seeds were stratified in agar water in the dark at 4°C for 3-4 days before they were transferred to soil. The plants were grown in a long-day (16 h light, 8 h dark) greenhouse. 100-150 µE white light intensity and a constant temperature of 22°C. Flowering time was scored as a number of rosette leaves at flowering. Five to six plants per genotype were analyzed in each experiment.

#### ***Hypocotyl elongation***

Seeds were plated on MS3 and stratified for 3-4 days at 4°C in the dark. Plates were illuminated for 3-4 hours before transferred to the appropriate cabinet with a short-day regime (8L:16D, 60 µE white light). The hypocotyls were measured after 9 days by imaging the seedlings with a flatbed scanner (Microtex artixScan 1100), and then region measurement feature of METAMORPH (MOLECULAR DEVICES CORPORATION, Downingtown, PA) was used to calculate the length of the hypocotyls.



---

### ***Cotyledon movement***

Rhythmic growth was measured and calculated as described (Hanano *et al.*, 2008). For this, seeds were surface-sterilized, placed on MS3, and stratified at 4°C in the dark for 3 days. Plates were transferred to long day under 12L:12D (100 µE white light) for 5 days before transferred to 100 mm square plates with 25 compartments (BIBBY STERILIN, UK). The seedlings were maintained in the germination medium to avoid damage of the hypocotyl. 1-cm-square agar blocks containing single seedlings were transferred to the plate, which was kept in vertical position. A few drops of sterile water was added inside the plate and the plates were sealed to avoid moisture loss. The plates were transferred to continuous light (low intensity white light, average 15 µE, lighting from the sides) and constant temperature of 22°C at dusk the next day. First leaf movements were monitored for 7 days using video cameras. The images were recorded every 30 minutes using METAMORPH. The rhythms of the cotyledon movements were analyzed in METAMORPH. Regions were defined for each leaf and the (x,y) pixel coordinates corresponding to the central position of the leaves were measured (Edwards *et al.*, 2005). The data was analyzed as described (Edwards *et al.*, 2005 and Hanano *et al.*, 2008).

### ***Generation of transgenic Arabidopsis plants***

Plants were transformed with the Agrobacterium harboring the construct of interest by the floral-dip method, as described (Davis *et al.*, 2009).

### ***Imaging with a CCD camera***

Seeds were surface-sterilized, sown on MS3-medium, stratified for 2-3 days at 4°C in darkness, and transferred to long day conditions (16L:8D). After 7 days, 8 seedlings per genotype were transferred to fresh MS3-medium in a 96-well micro-titre plate. All the plants were sprayed with 5 mM luciferin to inactivate accumulated luciferase and plates sealed before taken to the single-photon-counting liquid-nitrogen cooled CCD camera. The 5 mM Luciferin used was in excess to ensure that the detected activity of the luciferin is proportional to that of the promoter. The plates were transferred to the CCD camera for the monitor and measurement of the emitted luminescence (Visitron Systems). Each image was for a 15 minutes picture. These images were taken 1 per hour for a period of 6 days. The imaging of the plants during exposure and data were analyzed (Gould *et al.*, 2006), using the MetaMorph Imaging Series 6.1 software (Universal Imaging Corporation, Dawnington, PA, USA).

### ***Cassava Transformation Procedure***

The Friable Embryogenic Callus (FECs) used for Cassava transformation were obtained from nodal cuttings from *in-vitro* plantlets cultured in 4E media at a density of 25 cuttings per Erlenmeyer glass flask. The node cuttings were grown for 2-3 months, then the buds were

---

---

extracted from the explants and cultured for 25-30 days in 5 ml MS4 liquid media -Murashige and Skoog (1962) salts and vitamins supplemented with 4 mg/l 2,4-Dinitrophenol for somatic embryo induction. The somatic embryos formed were excised from the rest of the tissue and sub-cultured on Gresshoff and Doy basal solid medium in the presence of 4-amino-3,5,6-trichloro-picolinic acid (Picloram) for 3 months to induce the formation of FECs from which highly totipotent embryogenic suspension cultures were established. The suspension were plated on GD2-50Pi +tyrosin (400 $\mu$ M) media and after 30-45 days, developed FECs in the clusters were subcultured into fresh GD2-50Pi +tyrosin media for proliferation (maximum of 9 clusters per dish). The FECs are ready for transformation.

The ethanol inducible pNewFT plasmid was introduced into *Agrobacterium* ABI by electroporation and transferred to Cassava FECs by *Agrobacterium*-mediated-transfer protocol, as described by Schöpke *et al.* 1996, with modifications made at CIAT known to promote transformation in several independent transgenic events (CIAT 2002).

Pure FECs used for transformation were obtained from the Cassava variety TMS 60444. Bacterial cultures from ABI-pNewFT was grown overnight at 250 rpm in 50 ml bacterial induction medium LB containing 19.6  $\mu$ l of acetosyringone (100 mg/ml) with the selective antibiotics Kanamycin 50 mg/ml, Chloramphenicol 30 mg/ml, and Carbenicillin 100 mg/ml at 28°C until the Optical Density at O.D 560 was between 0.5-1.0 (Figure 2.2A). Acetosyringone was added as a phenolic inducer which helps trigger the activation of the genes in the virulence region of the Ti plasmid of the *Agrobacterium* which are necessary for the initiation of the transfer of the T-region DNA from the Ti plasmid to the plant chromosomes (Sheikhholeslam & Weeks, 1987). The bacteria were pelleted, the supernatant discarded, and resuspended in 10 ml LB media. To this, 9.6 $\mu$ l of acetosyringone was again added and 10  $\mu$ l of the bacterial culture was used to inoculate the FECs by suspending the bacterial culture in them (Figure 2.2B). Prior to this stage, the FECs had been re-distributed in clusters of 5 mm width of about 20 clusters per dish of solid media GD2-50Pi + Acetosyringone, each cluster weighed about 0.082 g. The plates were placed in a vacuum apparatus to remove air trapped within the FECs to ensure direct and total contact of the *Agrobacterium* culture with all the FEC units. The co-cultivation of construct containing *Agrobacterium* and FECs was carried out in darkness for 48 hours at 21°C.

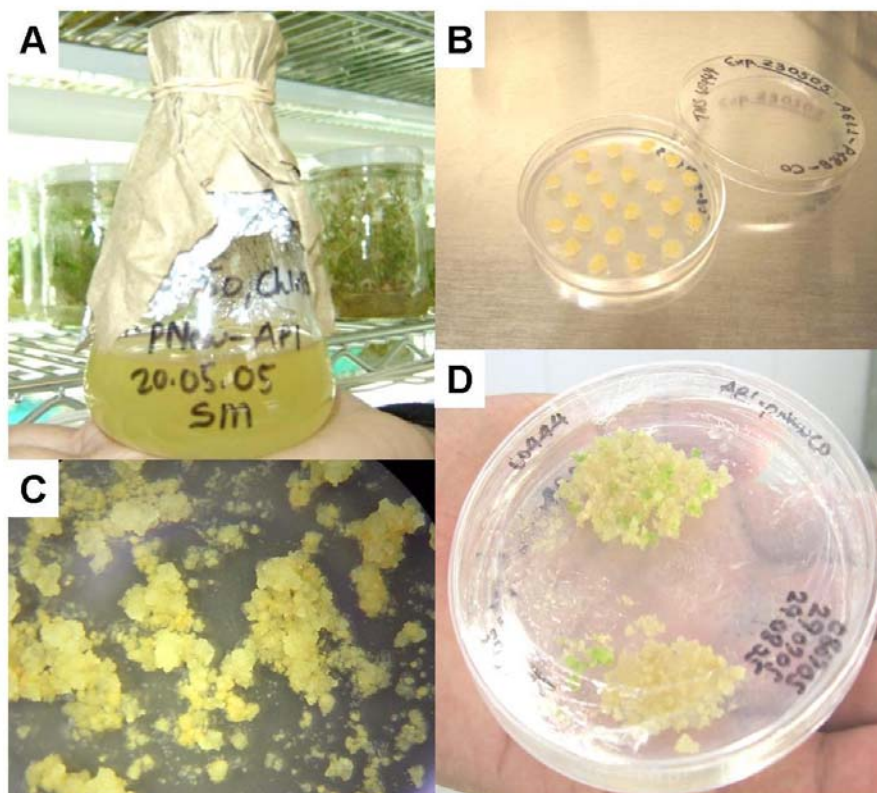
Fifteen FEC-containing plates were inoculated. After co-cultivation, those FECs already showing increased infection by bacteria were collected with a sterile spatula, and these were washed 4 times with GD2-50Pi liquid media supplemented with Cefotaxin (0.5mg/ml), to kill residual *Agrobacterium* and prevent future contamination of the FECs. The washing procedures were repeated each day for one week, and then washed for an additional week with 10 mg/L of Finale (active ingredient phosphinothricin). After the washing steps were complete, individual cell lines

---

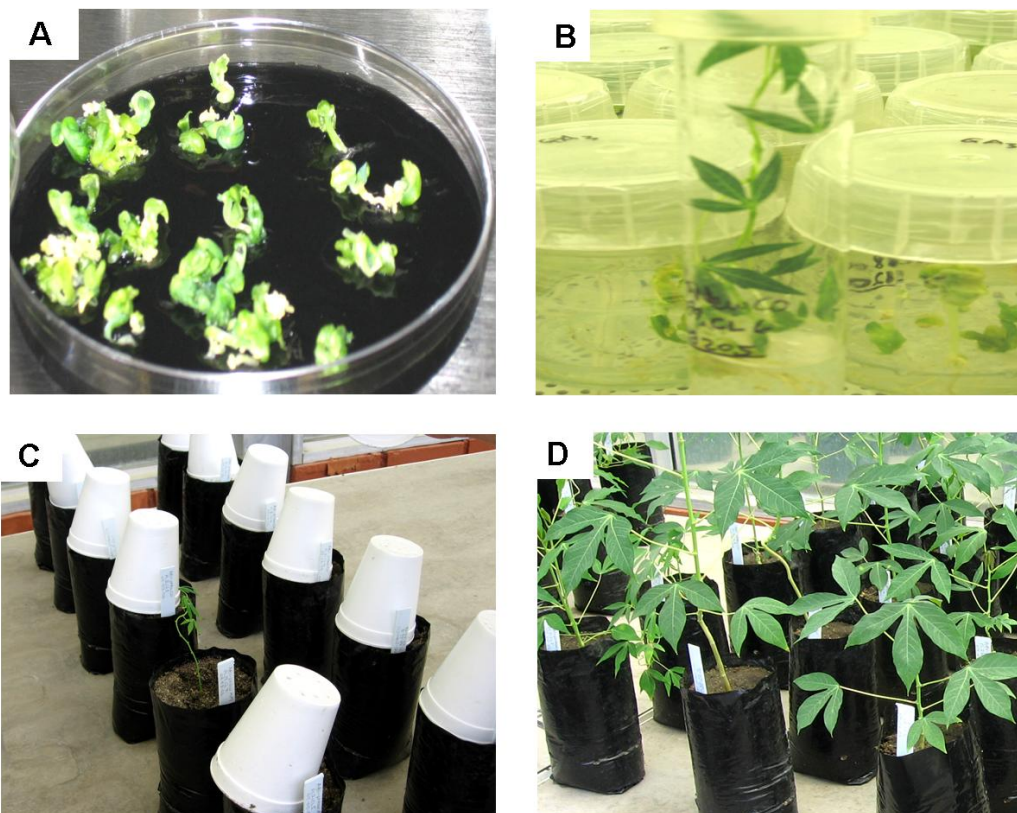
---

of infected FECs were plated on solid GD2-50Pi selection media for proliferation by differentiation of the FECs for 4 weeks after which the medium was changed twice after every 2 weeks each under the appropriate selection pressure (Figure 2.2C).

The FECs were continued in selection on MS2-1 $\mu$ M NAA ( $\alpha$ -Naphthalene Acetic Acid) for more than 5 weeks to allow for the induction of the growth and development of somatic embryos. Somatic embryo development is however asynchronous, since the time of embryogenesis vary from cell to cell (Figure 2.2D). The green somatic embryo from these FEC lines were transferred from the maturation medium to MS2-0.5% activated charcoal (Figure 2.3A). FECs undergoing somatic embryogenesis were transferred to this media of activated charcoal, to further improve cell growth and development by enhancing morphogenesis and decrease toxic metabolites (exudates phenolic compounds accumulation and promote adsorption of vitamins and growth regulators). This growth step took about one month. Young matured green transgenic lines were transferred to elongation media MS2-0.2 $\mu$ M GA<sub>3</sub> (Gibberellin) for another period of elongation, which took about 3 and more weeks. After elongation, the shoots were transferred to 17N (Figure 2.3B) for the development of roots for another 4 weeks and maintained *in vitro*.



**Figure 2.2 Cassava Transformation from Agrobacterium-mediated transfer to somatic embryo formation** (A) Agrobacterium culture ready for transforming FECs (B) Cassava FECs co-cultivated with Agrobacterium harboring the pNewFT construct, after 48 hours in the dark (C) Close up, microscopic view of transformed FECs in GD2-50Pi proliferation medium (D) Cassava transformed FECs in somatic embryo maturation medium.



**Figure 2.3 Cassava somatic embryo germination to green house establishment** (A) FECs has just completed embryogenesis and germinating in activated charcoal (B) Putative transgenic cassava lines differentiating into organ root (C) Cassava putative transgenic lines freshly transferred to soil (D) Cassava putative transgenic events well established in the glass house ready for molecular analysis.

### ***Greenhouse establishment***

The *in vitro* maintained putative transgenic Cassava lines were grown up to 4 weeks.

After this 4 weeks period, when the plantlets had well established roots, they were transferred to the green house. Plantlets were carefully removed from test tubes, washed with running tap water to eliminate excesses of phytoagar, which could serve as bait for fungal growth, and planted in sterile soil. The planted plantlets were covered with transparent water proof bags or inverted with stryophyle cups and punctured to maintain humidity (Figure 2.3C). The bags/cups were removed after the plantlets were established on soil. Plantlets were carefully watered every other day, or as the need arose (Figure 2.3D). These plants were maintained in the glass house with temperature ranging from 25-30°C, under long days (16h Light and 8h Dark) green house with light intensity ranging from 120-153  $\mu\text{mol s}^{-1} \text{m}^{-2}$  and under short days (8h Light, and 16h Dark) growth chamber with light intensity from 140-190  $\mu\text{mol s}^{-1} \text{m}^{-2}$  and at 22°C. These plants were used for subsequent molecular and physiological analysis.

---

### ***Somatic embryo cotyledon transformation***

Transformation of somatic embryo cotyledons using the Agrobacterium ABI carrying the flowering constructs were introduced to cassava cultivars CM3306-4 and Col 2215. About 300 explants were transformed. None of which survived the aggressive attack of the Agrobacterium infection on these explants. Additional efforts were made to introduce the flowering constructs into a less aggressive Agrobacterium LBA4404 were futile due to incompatibility between the plasmid and LBA4404. This was because LBA4404 is sensitive to Carbenicillin, which was used to select for positive transformants after the plasmid is introduced to the Agrobacterium.

### **2.3 Software, databases, and other internet resources**

#### Databases for genomic sequences of Arabidopsis:

<http://www.arabidopsis.org>

<http://www.arabidopsis.org/Blast/>

<http://blast.ncbi.nlm.nih.gov/Blast.cgi>

#### Database for genomic sequences of Ricinus communis:

(<http://blast.jcvi.org/er-blast/index.cgi?project=rca1>)

#### Database for genomic sequences of Cassava:

<http://www.ncbi.nlm.nih.gov/mapview/static/MVPlantBlast.shtml?3983>

<http://www.jgi.doe.gov/sequencing/statusreporter/psr.php?projectid=16112>

#### Primer design:

Primer 3 (<http://frodo.wi.mit.edu/>) (Rozen and Skaletsky, 2000).

Vector NTI Advance 10<sup>TM</sup>

Invitrogen corporation, 1600 Faraday avenue

Carlsbad, California.

#### Cluster Analysis and Tree:

<https://www.ebi.ac.uk/Tools/clustalw2/index.html>

---

**CHAPTER 3. DISCOVERY, PHYLOGENETIC RELATIONSHIPS, AND  
CHARACTERIZATION OF SEVERAL PHOTOPERIODIC GENES OF  
CASSAVA**

---

### 3.1 Introduction/ In Quest for Photoperiod Cassava Genes

Cassava is a tropical crop plant whose photoperiodic pattern has not been well studied. Daylength changes over the seasons in the tropical region are small, varying from 10 to 14 ( $\pm 2$  hours) throughout the year. These small differences are probable as to why studies on photoperiodic effects on tropical plants are uncommon (Alves, 2002). In limited studies, Cassava was reported to initiate floral formation at photoperiods  $> 13.5$  hours (Keating *et al.*, 1982a). Furthermore, Bruijin (1977) documented that these plants are sensitive to small differences in daylength, but flower initiation occurred under non-inductive conditions after 200 days of planting. As well, Keating *et al.* (1985), showed that growth under a long photoperiod increases leaf area, shoot growth, and storage root growth. It is thus critical to study and understand the molecular genetic response of Cassava to photoperiod, the genes that are implicated in the photoperiodic pathway of Cassava and lastly to define how these photoperiodic genes directly affect flowering time in this tropical plant.

To assist gene identification in Cassava, some limited description of the Cassava genome has been performed. A linkage map was developed from molecular markers (Gomez *et al.*, 1996 and Fregene *et al.*, 1997). This genetic map has been exploited in the positional cloning of genes (Mba *et al.*, 2001, Fregene *et al.*, 2001). The discovery of more genes within the Cassava genome will aid in the description of the genetic repertoire of this critically important tropical plant.

It is noted that Castor bean (*Ricinus communis*) is related to Cassava in that they both belong to the family of *Euphorbiaceae* (Mathews *et al.*, 1993). It is thus expected that Cassava orthologues of photoperiod genes will be most similar in sequence to that of this plant. Since the genome of Castor bean has been sequenced, much of the sequence information available in its database will be used with a similarity approach (Lifschitz *et al.*, 2006; Böhlenius *et al.*, 2006; Turner *et al.*, 2005; Hayama and Coupland, 2004). Therefore, I exploited the previously existing information available in the Cassava ESTs database, alongside with that of the genome of Castor bean to isolate several photoperiod genes in Cassava, and compared these Cassava sequences with corresponding genes in other species.

Nothing has been reported in Cassava on the physiology, biochemistry, and genetics of the clock genes, integrator genes, floral meristem identity genes, and the floral repressor genes. The identification, isolation, and function of these genes will present a giant stride towards the capacity to study molecular events of photoperiodism in the Cassava community. Such studies will contribute greatly to the genetic repertoire of this crop. The sequence information gathered from these Cassava genes were studied in detail, as reported in the subsequent sections.

---

---

Expressed Sequence Tags (ESTs) searches were made from the Cassava database and Castor bean genome site for sequences of interest. The sequence information obtained from these ESTs databases were used to design genomic-amplification primers in different orientation and they were used in PCR to isolate genes from Cassava genomic DNA and cDNA generated from pools of samples harvested from different time points of the day. Following the sequencing of the amplification fragments, sequence information was retrieved for *MeGI*, *COL*, *AGL24*, *ELF4*, and *LHY/CCA1*, all of which consequently showed extended identity with the Arabidopsis equivalent genes. It is noted that the sequences obtained lacked the complete ORF. The singular exception was *MeELF4* gene, which has the complete ORF. The characterization of these genes is described below.

## 3.2 Results

### 3.2.1 Computational Searches for Photoperiodic genes

Searches were made at both the Cassava ESTs database and at the Castor bean genome site for related sequences to the already characterized genes in Arabidopsis that are of interest to photoperiodism. These genes were the clock genes *LONG HYPOCOTYL (LHY)*, *EARLY FLOWERING 3 (ELF3)*, *EARLY FLOWERING 4 (ELF4)*, the photoperiodic output genes *GIGANTEA (GI)*, and *CONSTANS (CO)*, their direct targets, the integrator genes *FLOWERING LOCUS T (FT)*, *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1)*, and *FD*, the direct target of *FT*, and finally, the floral-meristem identity genes *APETALA 1 (API)*, *AGAMOUS-LIKE 24 (AGL24)*, *AGAMOUS*, and *LEAFY (LFY)*. Additionally, I searched for the sequences related to the floral repressors *FLC* and *FRI* (Onouchi *et al.*, 2000; Borner *et al.*, 2000; Samach *et al.*, 2000; Ratcliffe *et al.*, 2001; Kardailsky *et al.*, 1999; Kobayashi *et al.*, 1999; Blazquez *et al.*, 1997; Hepworth *et al.*, 2002; Lee *et al.*, 2000; Abe *et al.*, 2005). In the Cassava ESTs database, sequences similar to Arabidopsis *ELF4*, *LHY/CCA1*, *GI*, *COL*, *SOC1*, *AGL24*, and *API*, were identified and additionally from the Castor bean database, *FT*, *GI* and *COL* genes were found. To date *ELF3*, *FD*, *AG*, *FLC*, *FRI* and *LFY* were not identified in either databases. Therefore, molecular studies were carried out on the discovered genes.

### 3.3 Identification of the Cassava *GIGANTEA* gene

Circadian rhythms are believed to exist among organisms from bacteria to plants to animals, and green plants are responsive to photoperiod in part because of the clock (Young and Kay, 2001). The degree to which clock genes are conserved in plants is not yet clear, and the way photosynthesizing organisms respond to photoperiod varies across species (Garner and Allard, 1920). The *GIGANTEA (GI)* gene functions both in the regulation of the clock and alongside with *CO* and *FT* to induce flowering under long day conditions (Mizoguchi *et al.*, 2005). I suspect that the sequence and expression pattern of *GI* may be conserved in Cassava.



The *GI* sequence was identified in the Cassava ESTs database. The sequence was confirmed to be related to Arabidopsis *GI* by TBLASTX (Altschul *et al.*, 1990) searches in TAIR. Primers were designed from the EST sequences to PCR amplify *GI* from Cassava genomic DNA. The DNA amplification fragments were purified and directly sequenced. The sequencing results from the different *GI* fragments in different orientations, which were as a result of different primer combinations led to a contig that was assembled. The resulting 1624 bp consensus was analyzed. This contig was searched against the Arabidopsis TAIR BLAST 2.2.8 (see methods) with BLASTX: NT, this queries with a nucleic-acid search against a protein data set. The program search identified the *MeGI* sequence as sharing high similarity with Arabidopsis *GI* protein. The result showed *MeGI* protein sequence in relation to Arabidopsis *GI* protein in the database as 69% identical, 78% similar with no gaps. In Figure 3.1 below, the conserved amino-acid sequence region between the *MeGI* and *GI* proteins were compared. Thus, I have been able to show that there is conservation of the encoded GIGANTEA protein from the encoded polypeptide from the Cassava protein.

Query: 1505-1296

KKQGPVAAFDSYVLAAVCALACELQLFPFVSSGSNHSSNDLGTAKSVKLNCGNGSSREFQSSFDSAVH  
 K QGPVAAFDSYVLAAVCALACE+QL+P +S G N S+S GT+ K VK+NG SS+E+ + DSA+  
 KGQGPVAAFDSYVLAAVCALACEVQLYPMISGGGNFNSAVAGTITKPKVING---SSKEYGAGIDSAIS  
 Subject: 640-706

Query: 1315-1086

HTRRILAILALEALFSLKPSSVGTSSWSYSSNEIVAAAMVAHVSELFRRSKACMHALSVMRCKWDNEITYTR  
 HT RILAILALEALFSLKPSSVGT WSYSS+EIVAAAMVAAH+SELFRRSKA HALS LMRCKWD EI+ R  
 HTRRILAILALEALFSLKPSSVGT PWSYSSSEIVAAAMVAHI SELFRRSKALTHALSGLMRCKWDKEIHKR  
 Subject: 707-776

Query: 1085-876

ASSLFNLIDIHASKAVASIVTKAEPLAHLQ-VPVWKDSLVCDFGKRQNKKLSIRCFDSGQSASQCVEST  
 ASSL+NLLID+HSK VASIV KAEPLA+L+ PV KDS+ C + K++N S CFD+ +SAS+ +  
 ASSLYNLIDVHSKVVASIVDKAEPLAYLKNTPVQKDSVTCLNWKQENTCASTTCFDTAVTSASRTEMNP  
 Subject: 77-846

Query: 875-666

HSDTTIITERSSCSKEGSGSTLGKDIAAFPLDASDLANFLTMDRHIGFNLCGAVVLLRRTLVEKQELCFVS  
 + S EGS K I F LDASDLANFLT DR GF G LLR++L EK EL FSV  
 RGNHXYARH---SDEGSRPSEKGIKDFLLDASDLANFLTADRLAGFYCGTQKLLRSVLAEKPELSFSV  
 Subject: 845-912

Query: 665-456

VSLLWHKMI AAPETQPSAVSTSAQQGWRQVVDALCNVVSASETKAATAIVLQAERELQPWIAKDDDQGQK  
 VSLLWHK+IAAPE QP+A STSAQQGWRQVVDALCNVVS+A P KAA A+VLQAERELQPWIAKDD++GQK  
 VSLLWHKLI AAPETQPTAESTSAQQGWRQVVDALCNVVSATPAKAAA+VLQAERELQPWIAKDDDEEGQK  
 Subject: 913-982

Query: 455-246

LWRINQRIVRLMVELMRNHDTPESLVILASASDLLLRATDGMLVDGEACTLPQLELLEATARA+QPVLEW  
 +W+INQRIV+++VELMRNHD PESLVILASASDLLLRATDGMLVDGEACTLPQLELLEATARA+QPVL W  
 MWKINQRIVKVLVELMRNHD RPESLVILASASDLLLRATDGMLVDGEACTLPQLELLEATARA+QPVLAW  
 Subject: 983-1052

Query: 245-37

GESGFVADGLSNLLKCRPATIRCLSHPSAHVRALSTSVLRSILHTCSI---KPTANQADINDIRGPS  
 G SG AV DGLSNLLKCRPATIRCLSHPSAHVRALSTSVLR I++ SI P + N + PS  
 GPSGLAVVDGLSNLLKCRPATIRCLSHPSAHVRALSTSVLRD IMNQSSIPIKVTTPKLPTEKNGMNSPS  
 Subject: 1053-1122

---

Query: 36-1  
 YQLFKVDVIDWQ  
 Y+ F IDW+  
 YRFENAASIDWK  
 Subject: 1123-1134

**Figure 3.1 Partial Protein-Sequence alignment of Arabidopsis and Cassava GI.** The shaded purple region shows the regions of conservation between both proteins, the + sign signifies similar amino acid and the empty space between protein sequences represents non-similar amino-acid sequence. The query signifies the *MeGI* nucleotide sequence translated into protein and the corresponding nucleotide number while the subject represents the GI amino acid sequence in the database and the corresponding amino-acid position.

### 3.3.1 Sequence alignment of encoded GI proteins

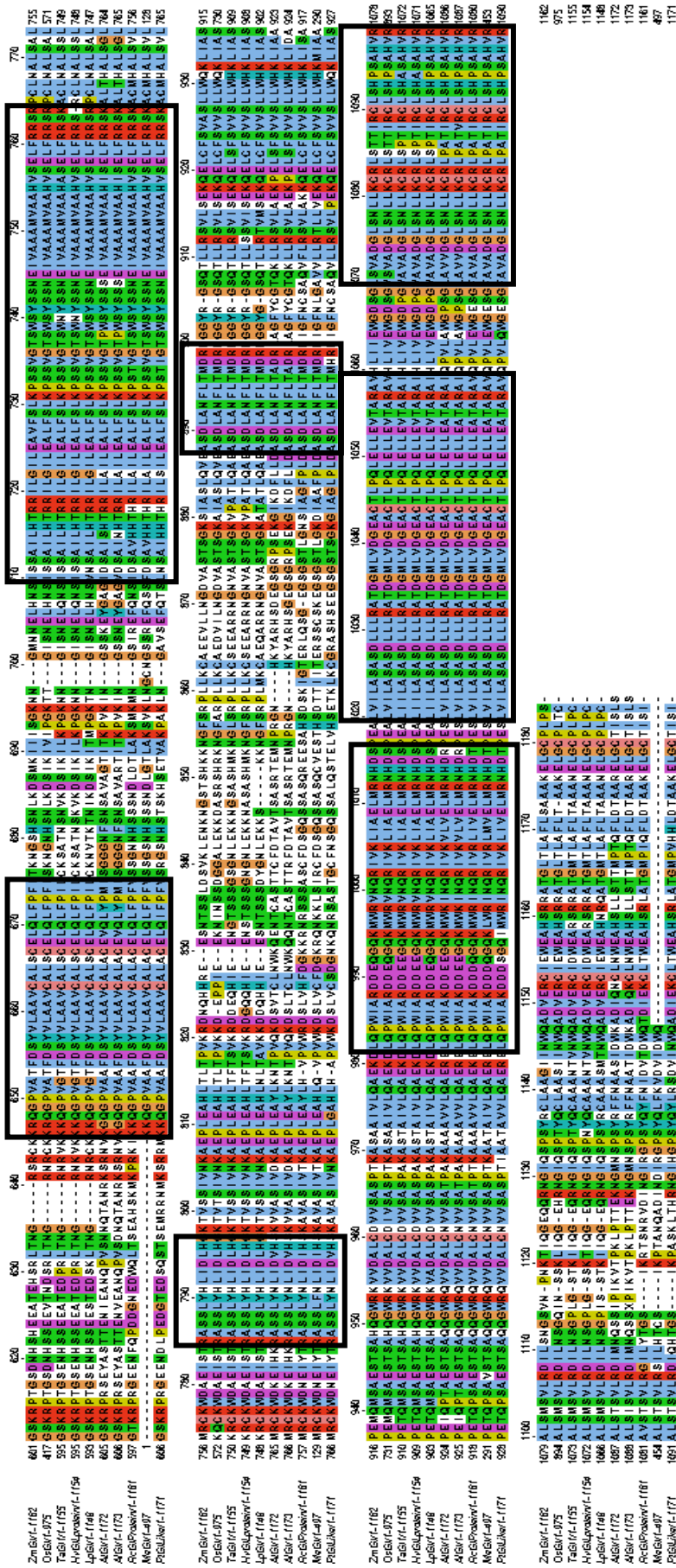
The conservation of amino-acid sequence of the Arabidopsis GI protein to the partially defined Cassava GI protein signifies the degree to which the conservation of the protein in species of related ancestors are preserved. The conservation and relatedness of the protein could infer that the protein may perform similar functions. However, before a general statement can be made about the putative function of the *MeGI* incomplete protein sequence, it is important to investigate the function of the GI protein among other species for analytical comparison.

I placed a query of the *MeGI* protein on the search gene from the NCBI genome database which revealed several candidates for GI orthologues. In addition, the genome database for *Ricinus communis* was queried to find its putative GI homologue. Table 3.1 below shows the plant species and assessments from where the GI proteins were obtained. A multiple alignment of the derived amino-acid predicted from the encoded genes was generated (Figure 3.2). The alignment showed highly conserved regions within these GI sequences. By visual inspection of multiple alignment, the regions of conservation are boxed, which includes slight differences in the amino-acid structure.

**Table 3.1** The source of each GI amino-acid sequence, and their corresponding name

Plant species	Amino-acid length	Assession number	Coined Name
<i>Arabidopsis thaliana</i>	1172	2199685	<i>AtGI</i>
<i>Arabidopsis lyrata</i>	1173	ABP96481	<i>AlGI</i>
<i>Zea mays</i>	1162	ABZ81992	<i>ZmGI</i>
<i>Triticum aestivum</i>	1155	AAQ11738	<i>TaGI1</i>
<i>Lolium perenne</i>	1148	ABF83898	<i>LpGI</i>
<i>Oryza sativa</i>	975	CAB56058	<i>OsGI</i>
<i>Hordeum vulgare</i>	1155	AAW66945	<i>HvGIL</i> protein
<i>Populus trichocarpa</i>	1171	EEE94512	<i>PtGILike</i>
<i>Ricinus communis</i>	1161	29863.m001097	<i>RcGI</i> protein
<i>Manihot esculenta</i>	497-partial sequence	This thesis	<i>MeGI</i>

---



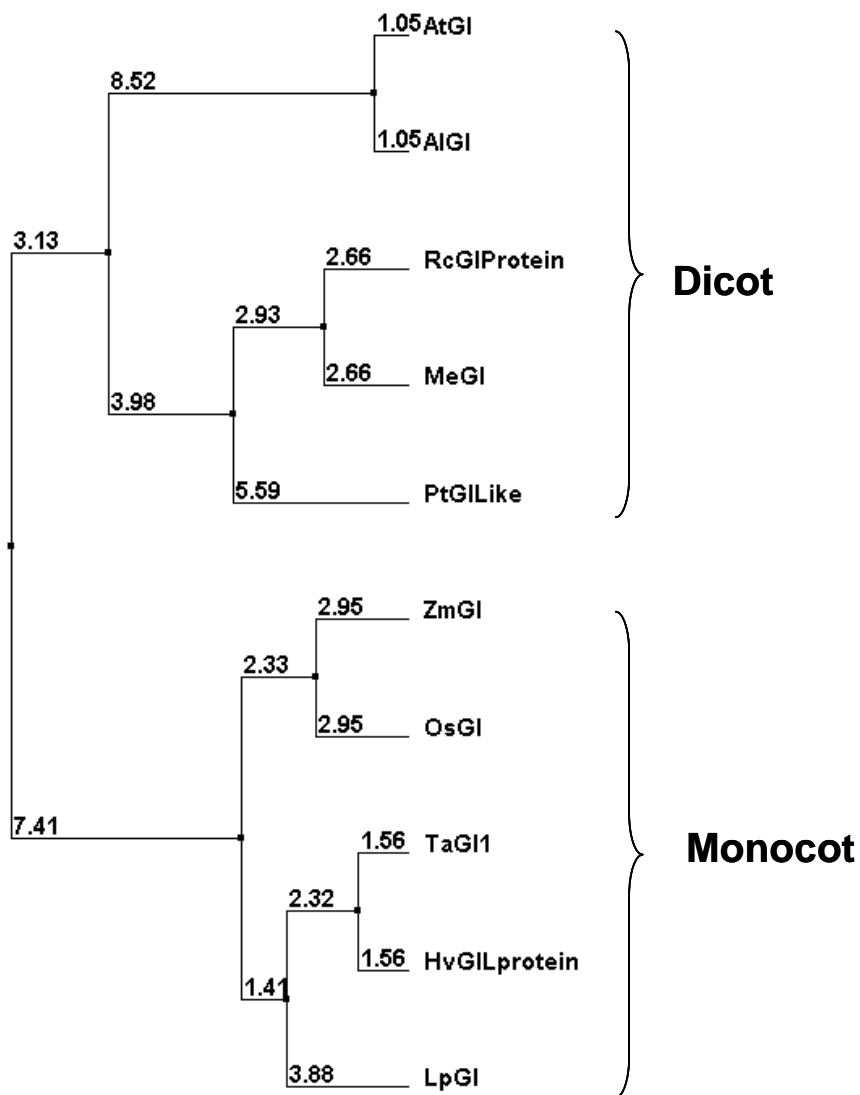
See Legend on next page.

---

**Figure 3.2 The multiple alignment of GI protein sequences from other species showing conserved amino-acid regions.** ClustalW2 multiple alignment of GI sequences. Names of plants as described in Table 3.1. Amino-acid positions are noted. The boxes are selected regions of conservation. The numbers above the aligned proteins are the positions relative to GI from *Zea mays*. Amino acids in red are small and hydrophobic (including aromatic-Y), in blue are acidic, magenta is basic and green are the hydroxyl, and basic-glutamine (Q) amino-acids.

### **3.3.2 Phylogenetic relationships of encoded GI protein among plant species and their relationships.**

A phylogenetic tree of encoded GI proteins was constructed from the sequence information obtained from NCBI (<http://www.ncbi.nlm.nih.gov/sites/entrez>). The NCBI search revealed several candidates for GIGANTEA orthologues in various species. GI orthologues from seven species were selected (Appendix 1; Table 3.1). In addition, the genome databases for *Populus trichocarpa* (NCBI) and *Ricinus communis* were queried to find putative GI orthologues. With the aim to define the conserved structural information from the GI primary sequences, a multiple alignment of the GI protein sequences was generated. Two subgroups could be identified within the alignment, two main clusters of monocotyledonous (monocot) and dicotyledonous (dicot). In this phylogenetic tree, two sub-clades were evident in both the dicot and the monocot groups. The MeGI-encoded protein showed high conservation within the dicot group and shares a higher degree of conservation with the Castor bean GI protein.

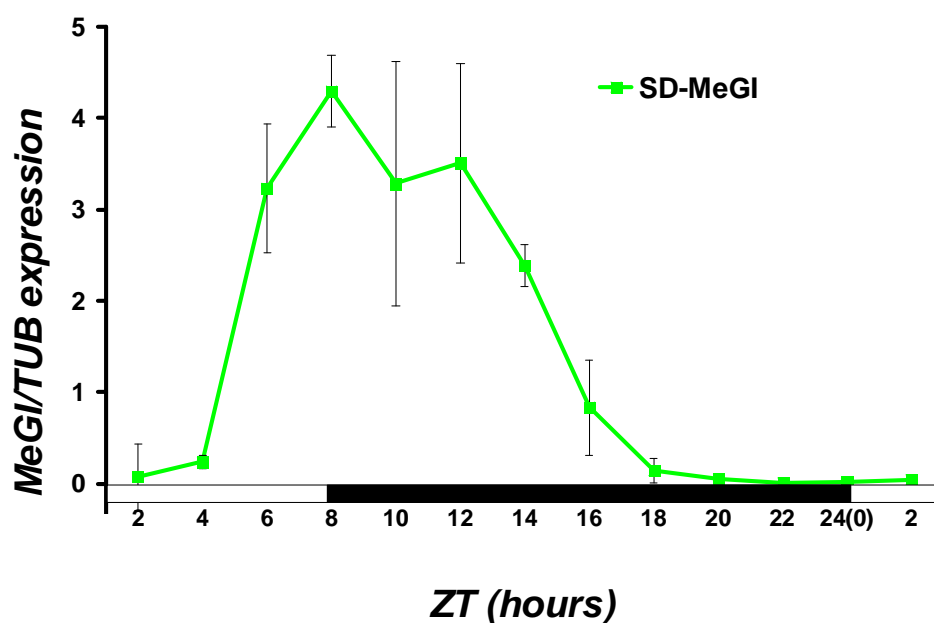


**Figure 3.3 Phylogenetic tree showing GI relationship within species.** This tree was generated using the Neighbor joining (NJ) method (Saitou and Nei, 1987). Identity of sequences as in Table 3.1. The tree classified the species into distinct groups of Dicot comprising of the *Arabidopsis thaliana*, *Ricinus communis*, *Populus trichocarpa* and the Monocot comprising of *Zea mays*, *Oryza sativa*, *Triticum aestivum*, *Hordeum vulgare*, *Lolium perenne*. The branch lengths are proportional to the amount of inferred evolutionary change.

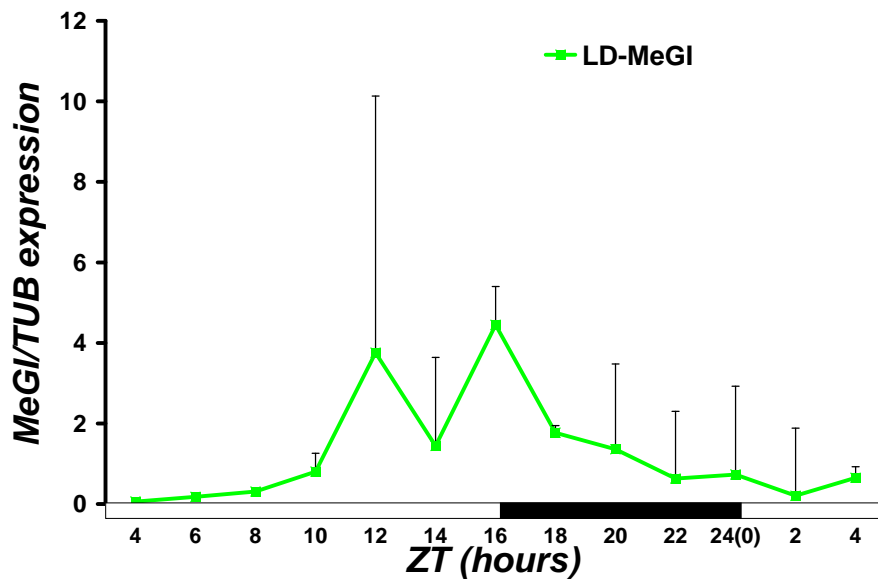
---

### 3.3.3 *MeGI* expression in Wild type Cassava under LD and SD conditions

The identification of the *GI* homologue of Cassava provided sequence information to examine its transcriptional pattern of expression. I determined the transcript expression in Cassava of *MeGI*, under long day conditions and short day conditions, respectively. Under short days, *MeGI* transcript peaked around dusk (4pm). It reduced in expression in the subsequent dark interval, and had an expression minima at around dawn (Figure 3.4). In contrast, under long day conditions, *GI* peaked at 10 pm (Figure 3.5). It is noted that whereas the time *GI* peaked changed, the peak coincided with dusk under both conditions. *MeGI* expression thus tracked dusk. This is similar to that observed in the Arabidopsis *GI* expression (Fowler *et al.*, 1999).



**Figure 3.4 Expression of *MeGI* under SD (8hL, 16hD) conditions.** Tissue harvesting started at ZT-2 (10 am) and continued every 2 hours for 24 hours. RNA (2 $\mu$ g) was extracted from apices. *MeGI/TUB* expression was measured by quantitative real-time PCR. The y-axis represents normalized gene expression between *MeGI* and *TUBULIN*. The white box represents the light phase and the dark box represents the night phase. Error bars are  $\pm$ SEM, where n=3



**Figure 3.5 Expression of *MeGI* under LD (16hL, 8hD) conditions.** Tissue harvesting started at ZT-4 (10 am) and continued every 2 hours for 24 hours. RNA (2 $\mu$ g) was extracted from apices. *MeGI/TUB* expression was measured by quantitative real-time PCR. The y-axis represents normalized gene expression between *MeGI* and *TUBULIN*. The white box represents the light phase and the dark box represents the night phase. Error bars are  $\pm$ SEM, where n=3

### 3.4 Identification of *CONSTANS*-Like Genes in Cassava

The phase transition from vegetative to floral growth is an important process under circadian-clock control. *CONSTANS* (*CO*) plays a central role in the output of photoperiodic promotion of flowering under long day conditions (Suarez-Lopez *et al.*, 2001). I found genomic regions of Cassava that contain sequence identity of *CO*. The *CO* sequence identified in the Cassava ESTs database was confirmed to be related to Arabidopsis *COL* genes by TBLASTX searches in TAIR. Primers were designed based on the sequence identity of the *CO* gene within the Cassava EST database. Amplification fragments by PCR of the Cassava genomic DNA was followed by sequencing. The sequencing results from *MeCO*-like genes were queried against the Arabidopsis sequence database. The program revealed series of sequence identities to *COL2*, *CO*, and *COL1*, respectively. Primers were designed from these identities and the primers were used to amplify cDNA. The cDNA was from plant tissues harvested from wild-type Cassava plant at every 2 hours of the day, under both long and short day conditions. These cDNA were pooled and used as DNA template in PCR. The results of the sequenced amplification fragments produced three different contigs. These contigs were searched against the Arabidopsis TAIR BLAST 2.2.8 (see methods) with BLASTX: NT. The results from the highest hits showed sequence identity to *COL1*, *COL2*, and *CO* with 49%, 54% and 43% identities, respectively. I annotated these genes as *MeCOL1*, *MeCOL2* and *MeCO*, respectively. The discovery of these *MeCO*-like genes imply that there is an

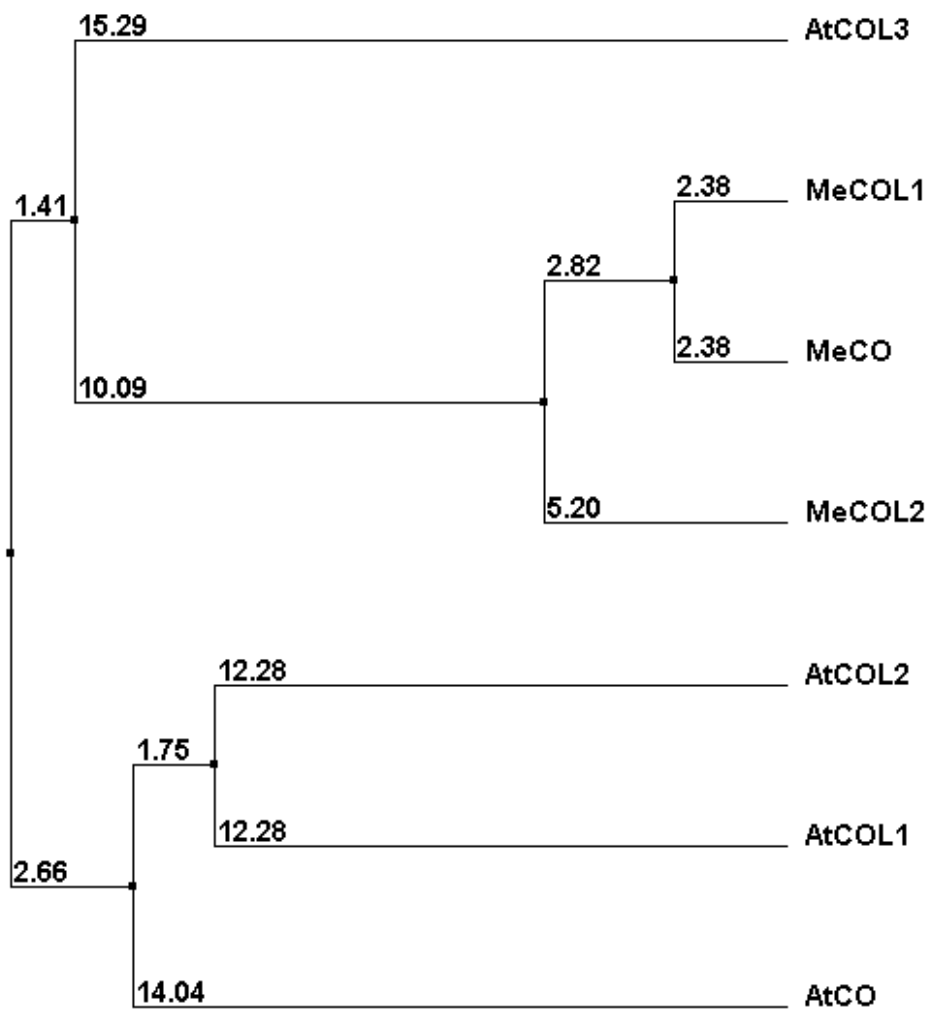
---

existence of a form of *CO*-like expression within the genome of Cassava, the exact and complete structure and function is unclear and yet to be resolved.

### **3.4.1 Phylogenetic relationship of *CONSTANS*-Like genes from both Arabidopsis and Cassava**

A phylogenetic tree was constructed from *CO*-like sequences from Arabidopsis and the identified *CO*-related sequences from Cassava to determine the relationship of these genes between the two species. It showed that the Arabidopsis *COL2*, *COL1*, and *CO* shared greater sequence similarity and had the same point of divergence. All Arabidopsis *CO*, *COL1* and *COL2* were categorized within a single clade (Figure 3.6). The *MeCOL1*, *MeCO*, and *MeCOL2* diverged at the same point as *COL3*, and this was classified as the outgroup (Figure 3.6). However, whereas *MeCOL1*, *MeCOL2*, and *MeCO* had same divergence point, both *MeCOL1* and *MeCO* belonged to the same clade and had the same distance to the sister clade with *MeCOL2*. This defines that *MeCOL1* and *MeCO* represent nearly identical genes (Figure 3.6). *MeCOL2* could be seen as closely related to both *COL2* or *COL3*.





**Figure 3.6 Phylogenetic tree showing *CO*-like genes from both *Arabidopsis* and *Cassava*.** This tree was generated using the Neighbor joining (NJ) method (Saitou and Nei, 1987). The identity of sequence as in Appendix II. The *COL2*, *COL1* and *CO* belong to the same clade, and the *MeCOL1* and *MeCO* were classified into another clade. The *MeCOL2* and *COL3* are the outgroups. The branch lengths are proportional to the amount of inferred evolutionary change.

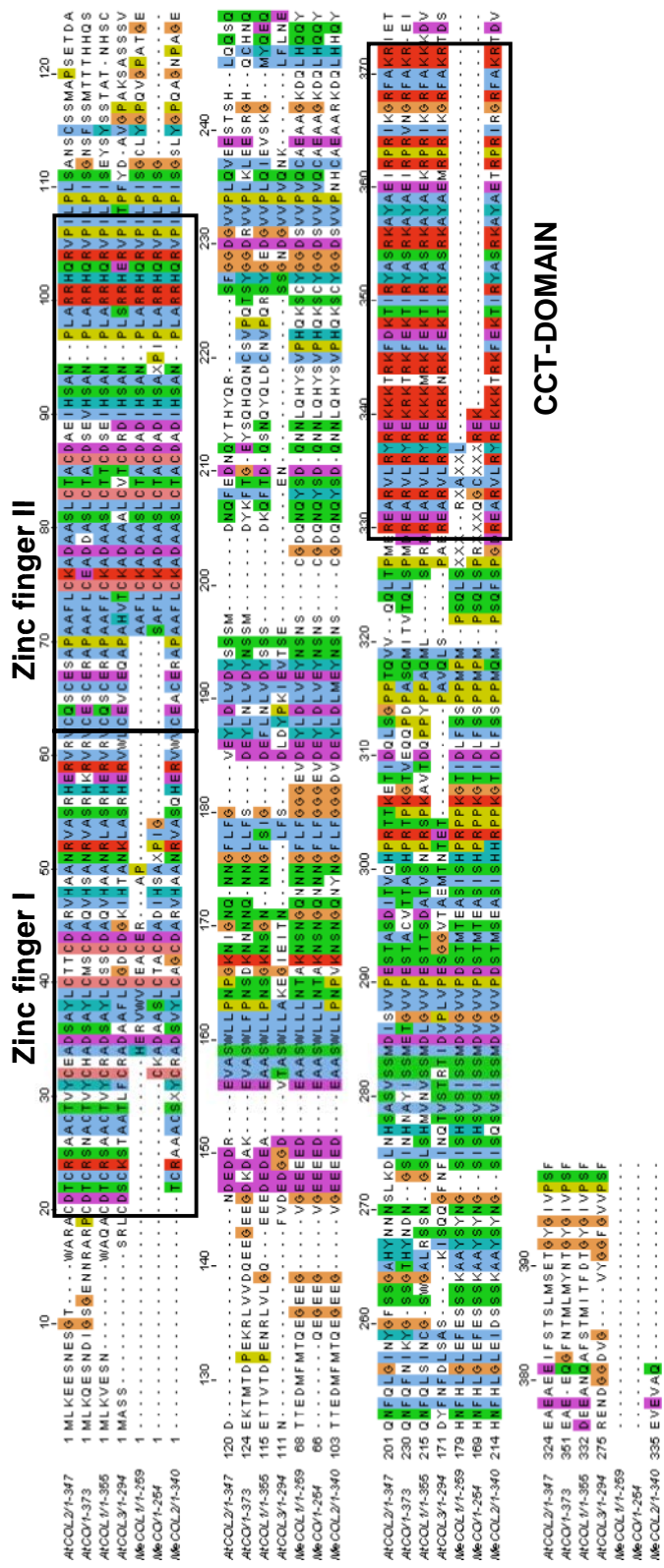
---

### 3.4.2 Sequence alignment of CONSTANS-Like Proteins

The *MeCOL* proteins (*MeCOL1*, *MeCOL2* and *MeCO*) were examined for domain structure. The three *MeCOL* proteins and the Arabidopsis CO, COL1, COL2, and COL3 proteins were aligned by cluster analysis. It is known that two regions of the derived CO proteins are well conserved, which is an N-terminal region with two putative zinc-fingers and a C-terminal region with the presence of CCT (CO, CO-like, TOC) domain. The CCT domain may contain a nuclear localization signal and it is found to be common to all CO-family (Griffiths *et al.*, 2003; Suarez-Lopez *et al.*, 2001). Genes were defined by Lagercrantz and Axelsson (2000) as members of the COL family if they contained two adjacent zinc-finger motifs fitting the consensus sequences  $CX_2CX_{16}CX_2C_1$ .

The resulting alignment showed only *MeCOL2* as the closest to share or to possess the two adjacent zinc-finger motifs fitting the consensus sequence  $XCX_{16}CX_2C$ , the first two amino acids were missing from my contig, this may be due to sequence deletion or incomplete sequence information. The details of these regions are highlighted in Figure 3.7. In the amino-acid sequences of the *MeCOL1* and *MeCO*, however, the zinc-finger motifs were found to be absent. The *MeCOL1* and *MeCO*'s motif appeared as  $X_6CX_2C$ , and  $X_8CX_2C$ , respectively (Figure 3.7). The incomplete ORF information of these conserved zinc-finger motifs in both *MeCOL1* and *MeCO* made it difficult to conclude whether these proteins lacked these domains.

The three encoded *MeCOL* proteins were also examined for the presence of CCT domain. CCT domains were detected in all three *MeCOLs* (Figure 3.7). *MeCOL2* had a complete CCT domain, but *MeCOL1* and *MeCO* were truncated in my contigs. However, upon the complete sequence information of these genes, the zinc finger motifs and the CCT-domain are expected to be present and complete, because these regions are conserved across species. Thus, all three *MeCOL* proteins have the same domain architecture as Arabidopsis COL1, COL2, COL3 and CO proteins.



See Legend on next page.

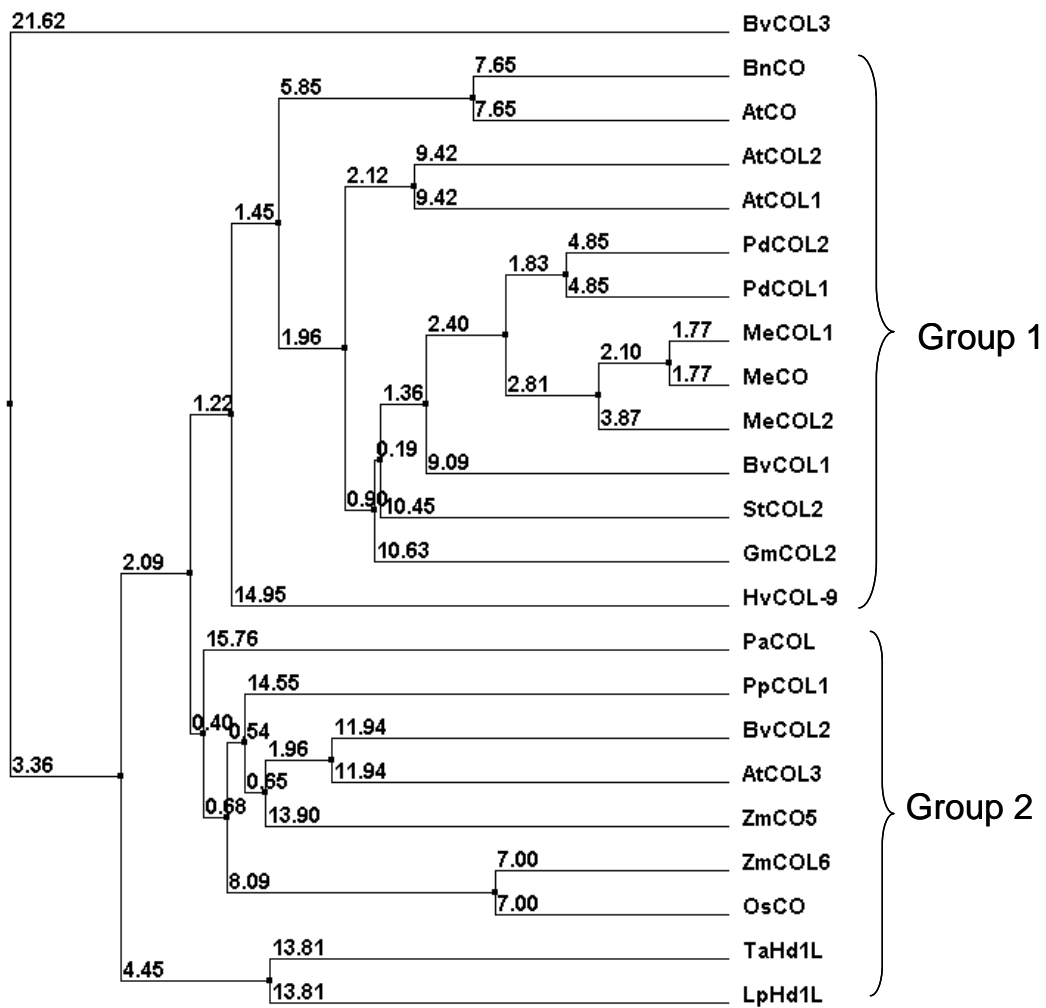
---

**Figure 3.7 Sequence alignment of CO-like proteins from Arabidopsis and Cassava.**

ClustalW2 multiple alignment of CO-like sequences. Amino-acid positions are indicated. The numbers above the aligned proteins are the positions relative to CO from COL3. The colors of the amino acids in red signify the small and hydrophobic (including aromatic-Y) amino acids, the amino acids in blue are acidic, those in magenta are basic and green ones are the hydroxyl, amine, and basic-glutamine (Q) amino acids.

**3.4.3 Phylogenetic relationships of CONSTANS-Like proteins among plant species and their relationships**

An expanded phylogenetic tree of various CO-like proteins was generated to assess the relations of *MeCO*-like proteins to those of other plants. The tree was constructed based on the alignment of four Arabidopsis CO and CO-like proteins and CO-like proteins from other plant species, including the characterized Cassava CO-like proteins. The tree showed a high degree of conservation between CO-like encoded proteins across all species. It revealed that COL3 protein from dicot plant *Beta vulgaris* is an outgroup, while all other sequences diverged into a major clade (Figure 3.8). Two sister clades resulted from this. I classified them as groups 1 and 2 respectively. The group two comprised of CO-like sequences from monocots, but included the Arabidopsis COL3 protein. In group 2, all the dicots clustered together as the second clade but the *Hordeum vulgare* COL-9 protein diverged as an outgroup from this clade of dicots (Figure 3.8). In the sub-clade within which *MeCO*-like proteins were found, sister clades with *Populus deltoids* COL1 and COL2 proteins. This may be due to high degree of conservation among the two families of malpighiales. The three *MeCO* encoded proteins clustered together but *MeCOL2* separately diverged from the *MeCO* and *MeCOL1*, this may indicate a functional role for *MeCOL2* as was seen from its structure. Conclusively, it can be inferred that these CO-like proteins from various species are conserved and they shared common ancestors and evolved over the years.

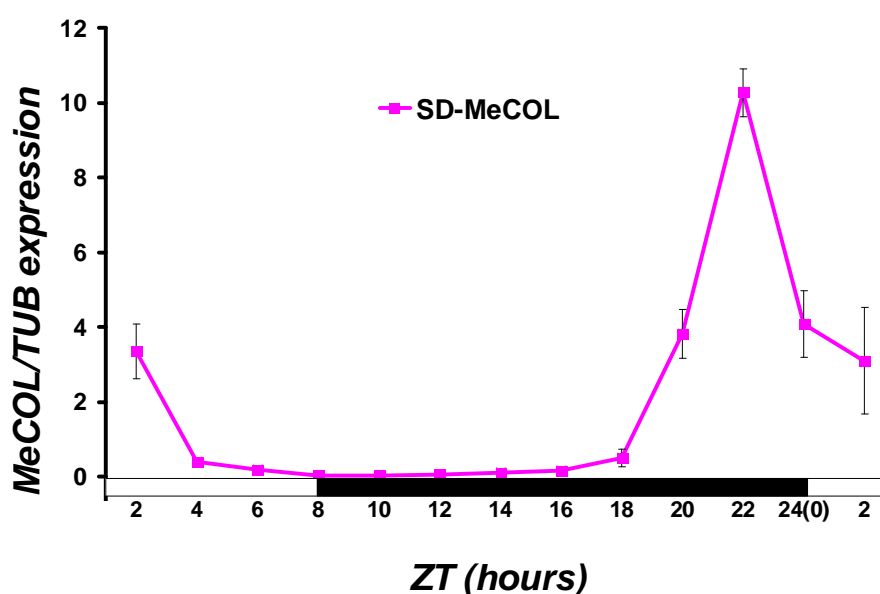


**Figure 3.8 Phylogenetic tree showing CO-like proteins from different species.** This tree was generated using the Neighbor joining (NJ) method (Saitou and Nei, 1987). The sequence used in the generation of the tree are *Populus deltoids*, *Lolium perenne*, *Physcomitrella patens*, *Glycine max* *Zea mays*, *Triticum aestivum*, *Picea abies*, *Oryza sativa*, *Brassica napus*, *Beta vulgaris*, *Solanum tuberosum*, *Arabidopsis thaliana*, *Manihot esculenta*, and *Hordeum vulgare*. The Branch lengths are proportional to the amount of inferred evolutionary change. Group 1 and 2 represents the two major clades with high sequence similarity.

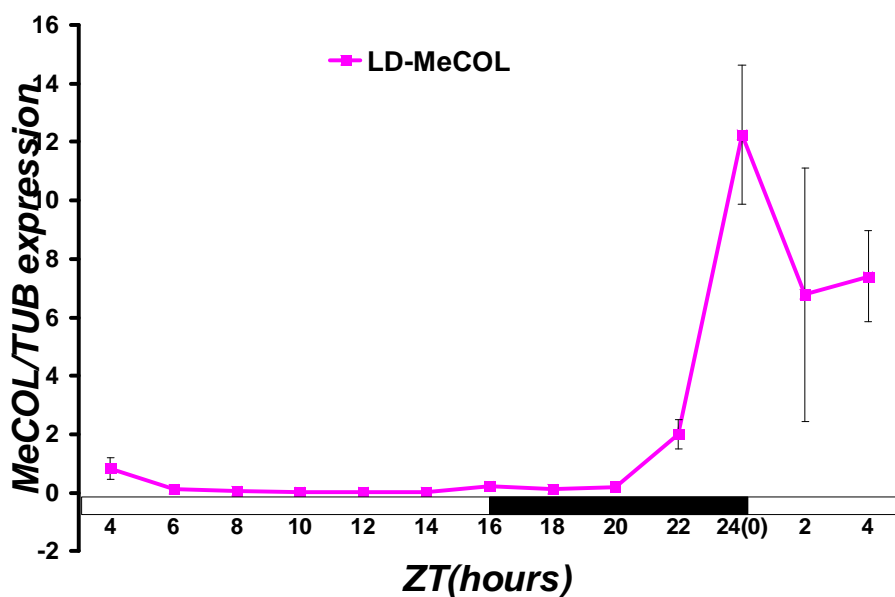
---

### 3.4.4 *MeCOL* expression in wild-type Cassava under LD and SD conditions

The identification of the *CO*-like genes in Cassava provided sequence information to examine their joint pattern of expression. The transcript expression of the *MeCOLs* was measured under long day conditions and short day conditions, respectively. The primers used for this expression study was from the conserved regions between all the three types of *MeCOL* isolated from the previous section. The combined *MeCOL* transcript peaked at 6 am, under short day conditions (Figure 3.9). Here, evidence is provided that the *MeCOLs* expression preceded the light phase, as the transcript level peaked in the dark 2 hours before dawn. This expression reached a trough at dawn. In contrast, under long day conditions, the *MeCOLs* peaked at 6 am. Their expression was found to be higher than that expressed under short day conditions. This higher expression under long days was found around dawn (Figure 3.10). The *MeCOLs* were thus predominantly expressed in the dark under both subjective conditions of short and long days, respectively.



**Figure 3.9 Expression of *MeCO*-like under SD (8hL, 16hD) conditions.** Tissue harvesting started at ZT-2 (10 am) and continued every 2 hours for 24 hours. RNA (2 $\mu$ g) was extracted from apices. *MeCO*-like/*TUB* expression was measured by quantitative real-time PCR. The y-axis represents normalized gene expression between *MeCO*-like and *TUBULIN*. The white box represents the light phase and the dark box represents the night phase. Error bars are  $\pm$ SEM, where n=3



**Figure 3.10 Expression of *MeCO*-like under LD (16hL, 8hD) conditions.** Tissue harvesting started at ZT-4 (10 am) and continued every 2 hours for 24 hours. RNA (2 $\mu$ g) was extracted from apices. *MeCO*-like/*TUB* expression was measured by quantitative real-time PCR. The y-axis represents normalized gene expression between *MeCO*-like and *TUBULIN*. The white box represents the light phase and the dark box represents the night phase. Error bars are  $\pm$ SEM, where n=3

### 3.5 Identification and Characterization of the *ELF4* Gene in Cassava

The Arabidopsis *EARLY FLOWERING 4 (ELF4)* gene has been described to be involved in photoperiod perception and circadian regulation. It promotes clock precision and is required for sustained rhythms in the absence of daily light/dark cycles (Doyle *et al.*, 2002; McWatters *et al.*, 2007). The *ELF4* gene is predicted to encode 111-amino acid protein having no significant homology to the proteins of known function, it belongs to a small but highly conserved Arabidopsis gene family (Khanna *et al.*, 2003, Doyle *et al.*, 2002).

These numerous clock role of the *ELF4* prompted me to examine the Cassava ESTs for the *MeELF4* gene homologue. The Cassava *ELF4* gene was identified, sequenced, and the open reading frame was isolated and cloned into a gateway pJalee4 vector fused to the *AtELF4* promoter. In this section, I describe the transcriptional regulation of the Cassava ortholog of Arabidopsis *ELF4* by reverse genetics using the loss-of-function mutants *elf4-1* harboring the *CCA1: luciferase (LUC)* and *CCR2: luciferase* markers (Doyle *et al.*, 2002).

---

### 3.5.1 Sequence alignment of ELF4 protein from other species

Investigations conducted in the past have reported ELF4 homologues from ice plant, sorghum and rice (Khanna *et al.*, 2003; Doyle *et al.*, 2002; Boxall *et al.*, 2005). There also exists EST notations from soybean and tomato (Boxall *et al.*, 2005).

In this sub-section, I report the characterization of Cassava *ELF4* sequence. Arabidopsis *ELF4* sequence identified homologous Cassava sequences within the Cassava ESTs by TBLASTX searches in TAIR (Altschul *et al.*, 1990). Primers were designed from this EST sequence and Cassava genomic DNA was amplified by PCR. The amplification product was sequenced and a 491 amino-acids sequence contig was formed, 111 of which amino-acid sequence was the ORF consensus. To examine the Cassava ELF4 (*MeELF4*) protein sequence with known ELF4-like EST clones from different plant species for conserved structural information from these primary sequences, a multiple alignment was generated (Figure 3.11). All amino-acid sequences are most similar in the central part of the proteins, suggesting this domain may be important for function.



Line	Text	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180	190	200	210	220	230
1	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
10	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
21	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
22	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
23	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
24	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
25	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
26	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
27	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
28	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
29	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
30	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
31	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
32	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
33	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
34	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
35	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
36	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
37	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
38	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
39	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
40	MEDS	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

See Legend  
on next page

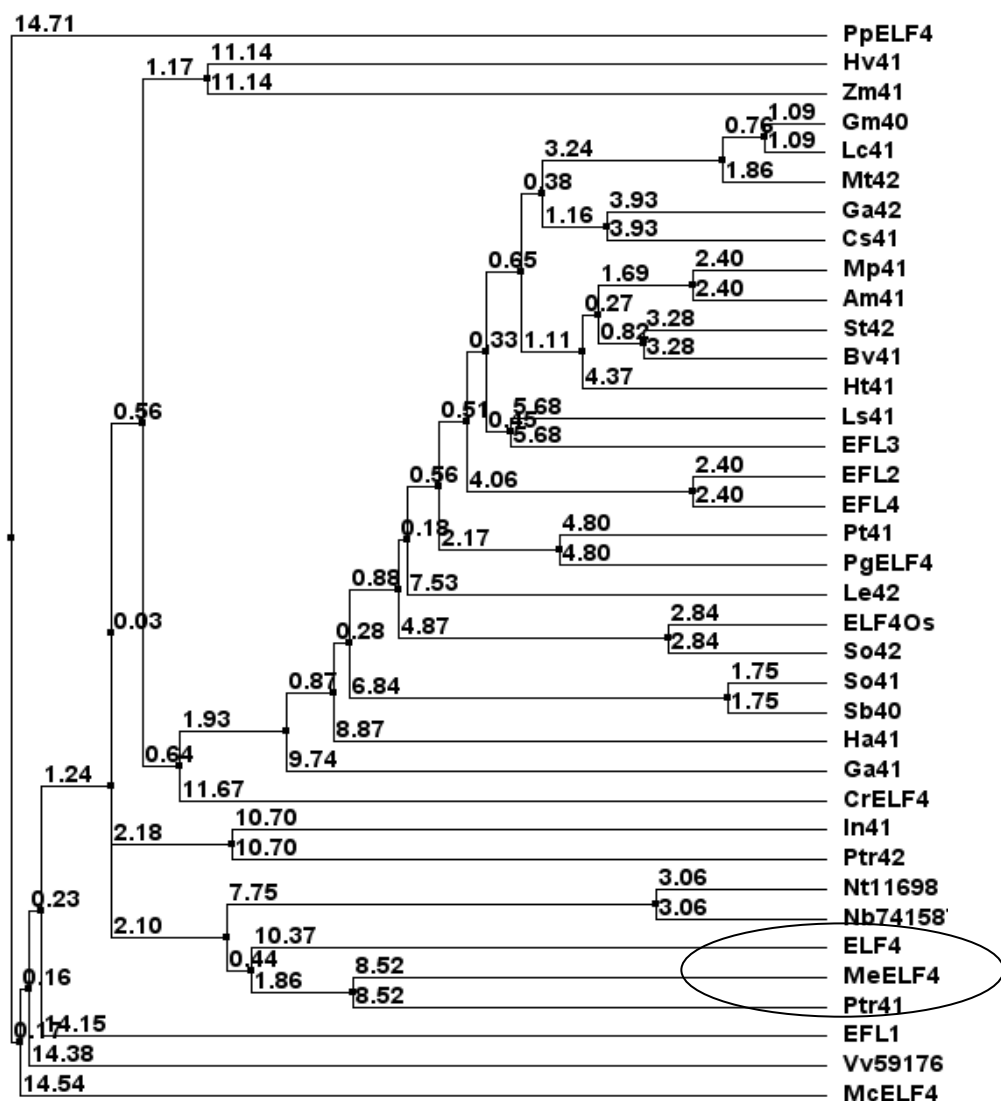
---

**Figure 3.11 The multiple alignment of ELF4 protein sequences and their conservation.**

ClustalW2 multiple alignment of ELF4 sequences. ELF4 and related sequences were retrieved from the thesis of Kolmos, 2007. Amino acid positions are noted. The numbers above the aligned proteins are the positions relative to ELF4 *Hordeum vulgare*. Amino acids in red are small and hydrophobic (including aromatic-Y), in blue are acidic, magenta is basic and green are the hydroxyl, amine and basic-glutamine (Q) amino acids.

**3.5.2 Phylogenetic relationships of ELF4 proteins among plant species and their relationships.**

A phylogenetic tree based on the alignment of the ELF4 full-length and partial ORF sequences from different species was calculated (Figure 3.12). This revealed that the *Physcomitrella* sequence *Pp*ELF4 could be defined as an outgroup. In this refined rooted tree, a sister clade of interest is notified in a circle among the numerous sub-clades (Figure 3.12). ELF4, *Me*ELF4, and *Ptr*41 constituted the clade of interest. This alignment showed that there is a high degree of conservation between Cassava ELF4 and Arabidopsis ELF4, due to the close clustering. This degree of conservation leads to the conclusion that the *Me*ELF4 and Arabidopsis sequences can be orthologues. To investigate the hypothesis of conserved function of *Me*ELF4 protein, complementation tests were conducted with Arabidopsis *elf4* mutants.



**Figure 3.12 Phylogenetic tree of ELF4 proteins from different species.** This tree was generated using the Neighbor joining (NJ) method (Saitou and Nei, 1987). Branch lengths are proportional to the amount of inferred evolutionary change. The circle identifies MeELF4. Note that it is in the same clade as Arabidopsis ELF4.

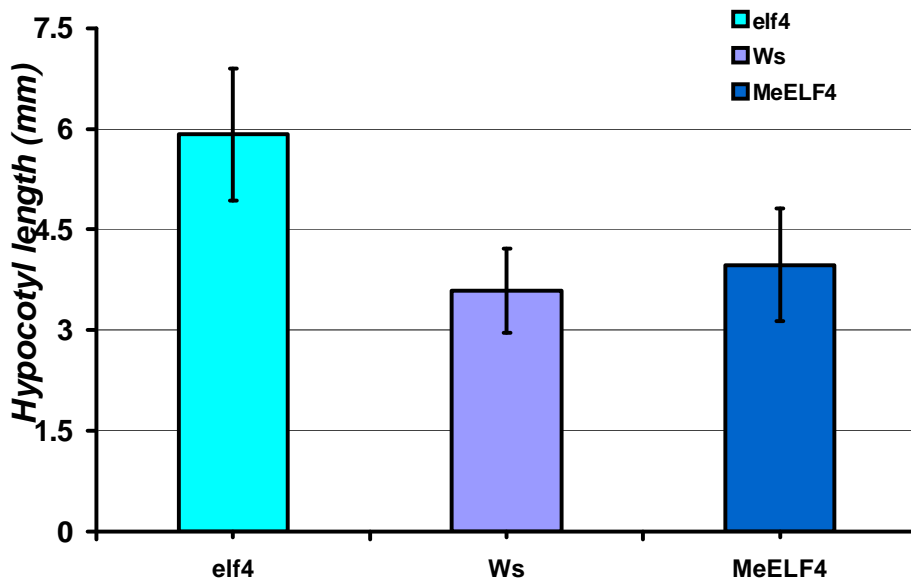
---

### 3.6.0 Complementation of *elf4-1* mutant with *MeELF4*

*ELF4* plays a role in the circadian clock of the Arabidopsis plant, as the loss of function results in clock arrhythmicity and accelerated flowering and long hypocotyl (Doyle *et al.*, 2002). The successful isolation and sequencing of the complete Open Reading Frame of the *MeELF4* presents an opportunity to investigate its role in the sustenance of normal clock function, and thus be able to conclude whether the *MeELF4* protein function is conserved. For this, a complementation experiment of *MeELF4* gene to *elf4-1* mutant was performed. The *MeELF4* complete ORF was cloned into a gateway vector pJalee4, generously provided by Elsebeth Kolmos (Kolmos thesis, 2007). The *MeELF4* gene was fused to Arabidopsis *ELF4* promoter and transformed into *elf4-1* mutants harboring the *CCA1::LUC* and *CCR2::LUC*, respectively. These markers are morning and evening expressed respectively (Doyle *et al.*, 2002). I obtained 48 independent T<sub>1</sub> transgenic complementation lines of *MeELF4* gene harboring the *CCA1* and *CCR2::Luciferase* markers.

### 3.6.1 Results from hypocotyl elongation measurement

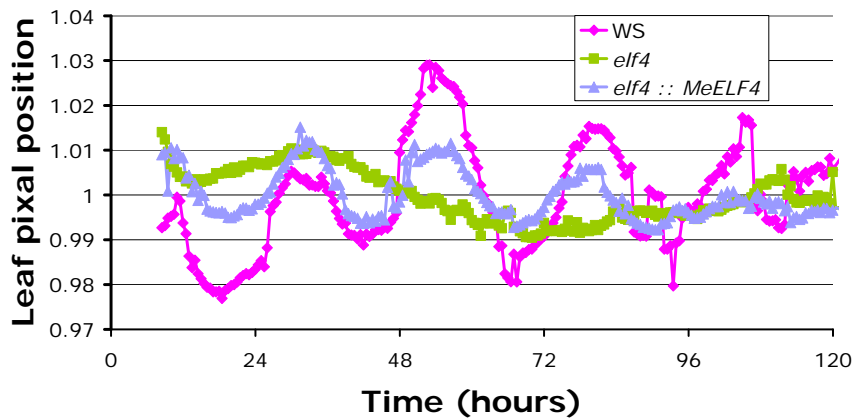
The growth rate of the hypocotyl is one of many aspects of the plant physiology exhibiting circadian behavior. It has been shown that circadian dysfunction causes aberrant hypocotyl elongation patterns in Arabidopsis (Dowson-Day *et al.*, 1999). Therefore, hypocotyl length of the T<sub>1</sub> transgenic complementation lines of *MeELF4* in *elf4-1* were measured for complementation of the hypocotyl-elongation phenotype. This was the first test to determine whether *MeELF4* can rescue the *elf4-1* mutant. The *elf4-1* mutant, *MeELF4*, and the wild-type Arabidopsis seedlings were grown for 10 days under a 8L:16D photoperiod, and the hypocotyls were measured and compared. The result showed the *elf4-1* mutant had an elongated hypocotyl, whereas the *MeELF4* rescue lines showed hypocotyl length comparable to the wild-type (Figure 3.13). Therefore, the phenotype expressed by the *MeELF4* complementation lines restored rhythmic hypocotyl growth.



**Figure 3.13 Hypocotyl-length elongation confirmed complementation of *MeELF4*.** Hypocotyl length of 10 days old seedlings of *Arabidopsis* transformants grown under short day 8L:16D conditions. *elf4* is the mutant line, *Ws* is the wild-type non-transgenic line in the *elf4-1* background and *MeELF4* represents the Cassava *ELF4* complementing lines. Data are averages of hypocotyl length in mm. Error bars represent  $\pm$ SEM.

### 3.6.2 Results from leaf movement measurement.

The circadian system drives pervasive biological rhythms in plants. Circadian clocks integrate endogenous timing information with environmental signals (Salathia *et al.*, 2006). Plant leaves exhibit circadian behavior in response to changes in light signals during the day. In order to confirm the clock role of the *MeELF4* in the *elf4-1* mutant background, genetic complementation tests by the leaf movement measurement of the leaves from T<sub>1</sub> transgenic lines of the *MeELF4*, the *elf4-1* mutant, and the wild-type was observed. The leaf movement measured under constant conditions of light after the seedlings were entrained to a 12L:12D photoperiod. The lines were then shifted to constant conditions and the leaf positions were measured for 5 days (Figure 3.14). The lines with sustained circadian rhythm is expected to be in horizontal position during the day and a more vertical position during the night. The mutant line was expected to be in vertical position irrespective of the day time. Leaf movement was then measurement from these *MeELF4* complementing lines, the wild-type and the *elf4-1* mutant. Figure 3.14 shows the different lines expressing different clock phenotypes. The WS wild-type and the *elf4::MeELF4* complementation lines expressed a circadian rhythm, whereas the *elf4* line did not (Figure 3.14). This shows that the Cassava *ELF4* could complement the leaf movement phenotype of *elf4*.



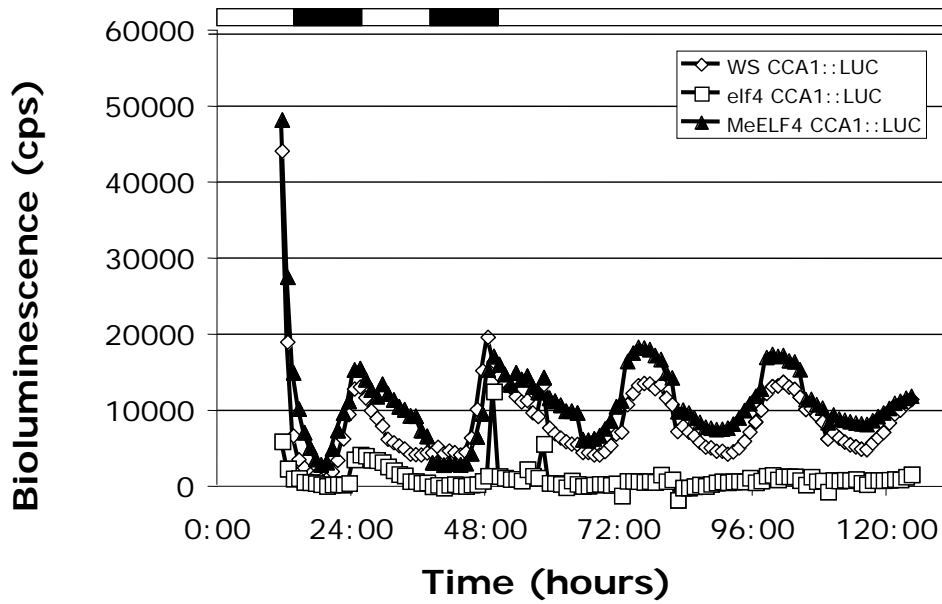
**Figure 3.14. Representative leaf movement data for lines expressing a clock phenotype.** Leaf movement rhythms were assayed under constant light for approx. 5 days (n = 14–28). The y-axis is the arbitrary leaf pixal position. The x-axis represents time in hours. Representative traces of rhythmic leaf movement of Ws is wild-type, *elf4* is the mutant line, and *elf4::MeELF4* is the complementing *MeELF4* line. Note that *elf4::MeELF4* was as rhythmic as the wild-type.

### 3.6.3 Molecular confirmation by imaging

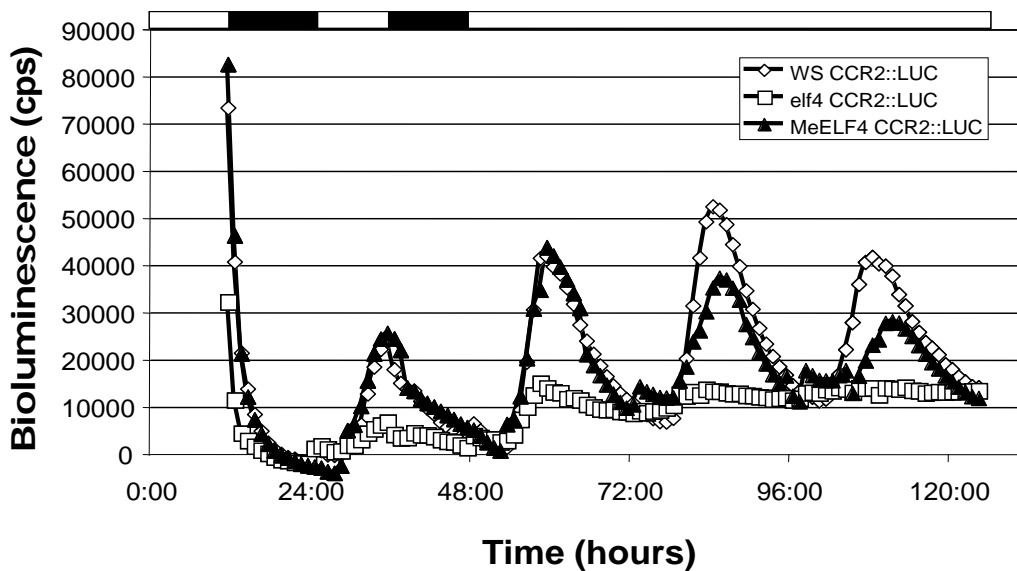
In a final complementation test, the ability of *MeELF4* to restore rhythms was tested by molecular imaging of clock-controlled gene expression. The *MeELF4* complementing lines were assayed for bioluminescence under diurnal cycles, and under free running conditions, using the CCD imaging camera. The experiment included the positive control, *ELF4p::ELF4*; the negative control, *ELF4p::no\_insert*; the *elf4-1*, parental line, and the Ws, wild-type. The 48 T<sub>1</sub> transgenic lines were directly compared. The transgenic lines were in respective *CCA1* and *CCR2::LUC* backgrounds. Both markers were tested for bioluminescence.

Under light-dark cycle of 12L:12D regime, the *MeELF4* gene fully complemented the *elf4-1* loss of function phenotype. The *MeELF4* gene restored *CCA1* expression and the circadian rhythm pattern of *elf4-1*. This complementation pattern was similar to what could be observed for the wild-type Ws *CCA1* non-transgenic control (Figure 3.15). The *elf4-CCA1::LUC* was found to be arrhythmic (Figure 3.15). The *MeELF4* gene complements the *elf4-CCR2::LUC* mutant by restoring amplitude and rhythmicity of the *CCR2* expression. The *elf4-CCR2::LUC* control remained arrhythmic under these conditions, as expected (Figure 3.16).

Under constant light, the *MeELF4* lines sustained circadian rhythm with amplitude similar to that of the wild-type, whereas the *elf4-1* mutants with either the *CCR2* or the *CCA1* markers, showed a loss of rhythmic expression (Figures 3.15 and 3.16). *elf4-ELF4* complementing *CCR2::LUC*, positive control, sustains rhythmic patterns while *elf4:- empty CCA1::LUC*, negative control, was arrhythmic. Thus, it can be concluded that the *ELF4* gene cloned from Cassava is the true orthologue of Arabidopsis *ELF4*.



**Figure 3.15 *CCA1::LUC* in various *ELF4* genotypes.** Plants were entrained under 12L:12D conditions for 2 days and then shifted to constant light (LL) for 3 days. *Ws* is the wildtype, *elf4* is the mutant, and *MeELF4* is the Cassava complementation in *elf4-1* Arabidopsis background, all expressed by the *CCA1::LUC* marker. Time is in hours. White bars indicate light intervals and dark bars indicate darkness.



**Figure 3.16 *CCR2::LUC* in various *ELF4* genotypes.** Plants were entrained under 12L:12D conditions for 2 days and then shifted to constant light (LL) for 3 days. *Ws* is the wildtype, *elf4* is the mutant, and *MeELF4* is the Cassava complementation in *elf4-1* Arabidopsis background, all expressed by the *CCR2::LUC* marker. Time is in hours. White bars indicate light intervals and dark bars indicate darkness.

---

In conclusion, in this chapter, I reported the discovery and successful characterization of several photoperiod genes in Cassava. Firstly, I showed that *MeGI* gene is conserved when compared across species, and that its expression pattern was similar to that observed in other species. A phylogenetic study showed *MeGI* to be closely related to GI. Next, I found several *COL* genes and showed that these *COLs* were comparable to *CO* and *CO*-like genes. A phylogenetic study showed the *MeCOLs* to be related more to Arabidopsis COL2 and COL3, than its *CO* counterpart. The encoded polypeptides of *MeCOLs* showed characteristic domain of *CO* family members with the presence of two amino-terminal zinc finger motifs and a CCT domain at the carboxyl terminal. This is similar to observations in barley and rice. The transcript study of *MeCOLs* indicates dawn expression pattern. Lastly, I was able to identify in Cassava a clock gene, *MeELF4*, which performs a clock function. I found it to be a true orthologue of *ELF4* in Cassava. *MeELF4* restored circadian rhythm in *elf4-1* mutants and sustained rhythmicity under constant conditions. Therefore, I have characterized several photoperiod and clock genes in Cassava.



---

**CHAPTER 4. EXPLOITING ARABIDOPSIS *FT* UNDER THE CONTROL OF  
AN ETHANOL-INDUCIBLE SYSTEM TO PROMOTE FLOWERING IN  
ARABIDOPSIS AND CASSAVA**

---

#### 4.1 Introduction

A major developmental transition for many green plants is that to flower. Plants detect seasonal conditions in a coordinated manner to maximize reproductive success. Extensive studies on this photoperiodic effect have been reported in *Arabidopsis*, and this provides insight into the understanding of the molecular pathways involved floral promotion (Chuck and Hake, 2005; Poethig, 2003). Flowering in Cassava is an important agronomic trait whose control is not optimal due to Cassava's recalcitrant flowering nature (Gonzalez *et al.*, 1998). Induction of flowering in Cassava through hormone application, or photoperiod manipulation is difficult, cumbersome, and sometimes expensive (Gonzalez *et al.*, 1998). Development of a low-cost alternative means of controlled and inductive flowering in Cassava would greatly overcome the synchronization problems being faced by breeders to thus enhance the genetic improvement of the crop. The success of controlled and inductive flowering in Cassava will make the elite genotypes, with "locked-up" desirable traits, readily accessible for conventional breeding.

The hypothesis for which this investigation is based is built on the success achieved from the overexpression of floral-integrator pathway genes and floral meristem-identity genes, to by pass floral repressors, and thus induce flowering. For example, in trees, it was demonstrated that mis-expression of floral meristem-identity genes induced flowering. Peña *et al.*, 2001 showed the overexpression of *LEAFY (LFY)* or *APETALA 1* induced flowering in transgenic citrus and this reduced the generation time from 8-10 years to 3 months. Similarly, the Aspen tree was genetically modified with *LFY* mis-expression and this promoted flowering time (Weigel and Nilsson, 1995). This suggests that an array of plants can be manipulated to flower faster. Such a technology could be applied to control flowering in tropical species.

More is now known as to the signal cues from the photoperiodic and the vernalization-dependent pathways, and how they are integrated transcriptionally, to coordinate the timing of flowering in *Arabidopsis*. One key target gene is of this pathway's *FLOWERING LOCUS T (FT)*. It acts as a mobile and graft-transmissible floral promoter that acts in the photoperiod-dependent pathway (Kardailsky *et al.*, 1999). Turck *et al.* (2008) provided evidence that FT protein is a major component of florigen and that gain-of-function of FT leads to early flowering, and that its loss-of-function causes late flowering (Corbesier *et al.*, 2007; Giakountis and Coupland, 2008). *FT* could act generally to promote flowering.

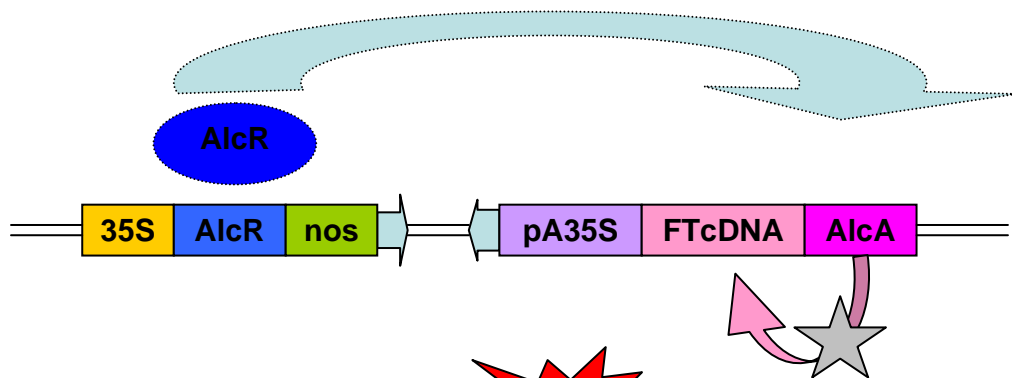
*FT* is known to promote flowering in a range of plants. It was recently investigated in cucurbits (squash plants) that the presence of a FT-Like protein is highly correlated with the onset of

---

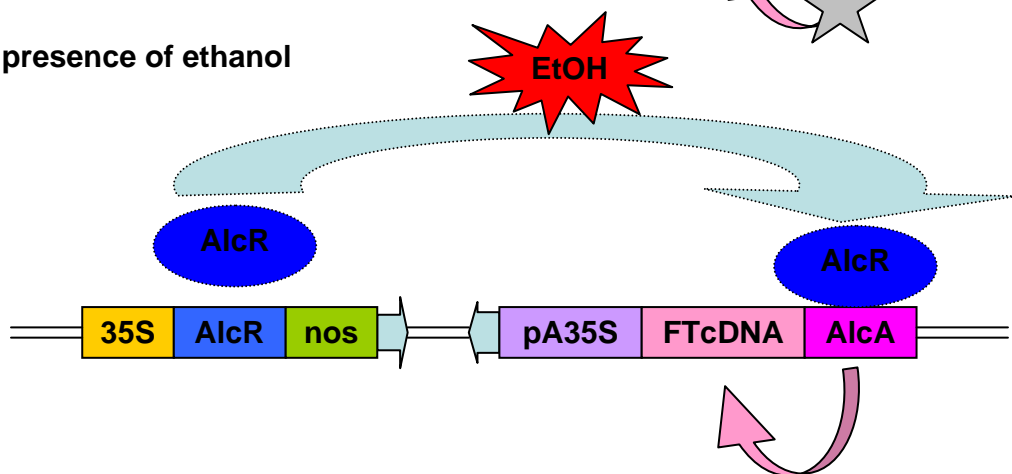
flowering (Lin *et al.*, 2007). *FT* induced flowering in wheat and barley (Turner *et al.*, 2005), in tomato, in tobacco (Lewis and Kernodle, 2009), in rice (Tamaki *et al.*, 2007; Izawa *et al.*, 2002; Kojima *et al.*, 2002), and in Arabidopsis (Corbesier *et al.*, 2007). *FT* appears to be a good candidate for manipulation to promote flowering in virtually any plant.

Promoter systems can be used based on the varying needs of misexpression studies. One useful inducible system is the AlcR/AlcA system (Felenbok, 1991). The AlcA ethanol-inducible promoter has been successfully used in several plants (Caddick *et al.*, 1998; Roslan *et al.*, 2001; Filichkin *et al.*, 2006; Garoosi *et al.*, 2005). Advantages of ethanol as a chemical inducer are that it is cheap, common, and readily available, and that this chemical is particularly assessible to farmers. It is a storable compound that could be used in the field, and only small quantities are required to induce gene expression (Roslan *et al.*, 2001). The AlcR/AlcA system function has been shown to work in controlled breeding environments (Sweetman *et al.*, 2002). The system comprises of the AlcR transcriptional factor and an AlcA promoter.

#### In the absence of ethanol



#### In the presence of ethanol



**Figure 4.1** The ethanol inducible system driving the expression of the *FT* gene. In the absence of ethanol, the system is in-active. In the presence of ethanol however, the AlcR transcriptional factor is activated which binds to the AlcA promoter. The AlcA promoter now drives the expression of the *FT* gene fused to it. 35S is the constitutive promoter, AlcR is the transcriptional factor, nos is the terminator, AlcA is the ethanol-inducible promoter, *FT* is the *FLOWERING LOCUS T* gene, pA35S is the terminator. Pnos is the promoter of pat (phosphoacetyltransferase) and pAnos is the terminator. The arrow under the AlcA denotes no expression in the absence of ethanol, and expression in its

---

It is my hypothesis that the AlcR/AlcA system will be suitable for Cassava because the *FLOWERING LOCUS T (FT)* gene is activated upon application of ethanol to the system.

In this chapter, I generated a pNewFT vector that manipulates the expression of floral-integrator gene *FT*, as a tool to control the timing of flowering. First, I confirmed the function and efficiency of this pNewFT construct (which harbors the AlcAFT) in Arabidopsis *FLC/FRI* late flowering genotype. This was a model system for a "shy" flowering genotype under the control of an ethanol-inducible system. I then transferred this system to a Cassava shy-flowering genotype (Figure 4.1). Here, I generated in Cassava, 12 independent putative transgenic events under the control of an ethanol-inducible pNewFT construct. The transgenic events in both Arabidopsis and Cassava were confirmed by PCR. The induction of the *FT* transcript after the application of ethanol was monitored by reverse transcriptase PCR, and the corresponding phenotypic change of floral induction was observed. I was thus able to show that the *FT* gene under the control of an ethanol-inducible system can cause *FT* transcript accumulation and that this could promote both plants to flower.

## 4.2 Results

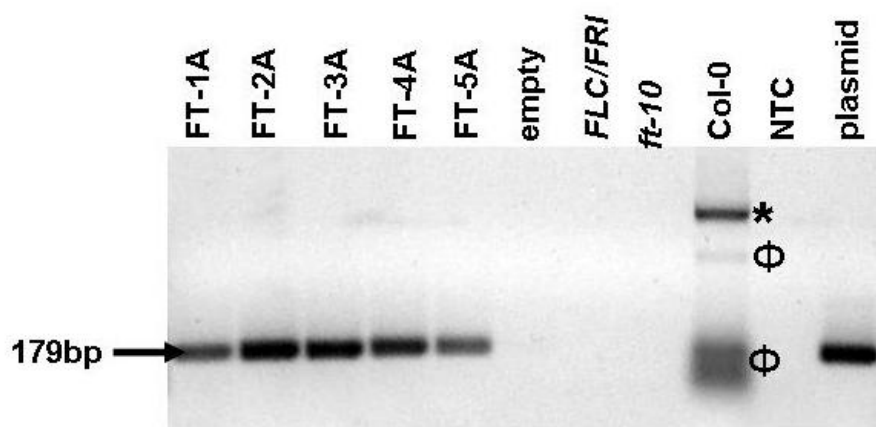
### 4.2.1 Arabidopsis *FLC/FRI* experiments

#### 4.2.1.1 Conventional PCR confirmation of *FLC/FRI* transgenic events

In the context of my experiment, a flowering-time gene was fused to the AlcA promoter, which drives the expression of the flowering-time gene in transgenic plants. Upon induction of the system by exogenous ethanol, the transcriptional factor AlcR binds the AlcA promoter, which in turn drives the expression of the floral timing gene (Figure 4.1). I investigated the promotive role of *FT* in the *FLC/FRI* late-flowering genotype of Arabidopsis. The pNewFT construct was introduced into this Arabidopsis "shy" flowering genotype, which was genetically active for *FLC* and *FRI*. The presence of *FLC* and *FRI* strongly represses inductive flowering (Michaels and Amasino, 1999). In my experiments, the pNewFT *FLC/FRI* genotype was grown under long day conditions on soil, and they were herbicide selected. This eliminated false-positive transformants. The resistant lines were transferred to fresh soil. Leaf samples were collected for genomic assays, in order to confirm the presence of the *FT* transgene. For this, PCR primers were designed from the coding sequence of the *FT* gene. In Figure 4.2, one can see the detection of the presence of the transgene. In particular, I could find the expected 179 bp product size from the *FT* coding region (Figure 4.2). This fragment size was absent in the empty control, as expected. This empty control has the pNew construct backbone, but lacks the *FT* gene. The 179 bp fragment was also absent in the *FLC/FRI* non-transformed genotype, as expected. In the wild-type genotype, a larger fragment was amplified, and this represents the genomic *FT* locus with introns. The 179 bp fragment was

---

also noted in the plasmid positive control, as expected. Therefore, it was evident that the *FT* gene was incorporated into the Arabidopsis genome in the desired transgenics.

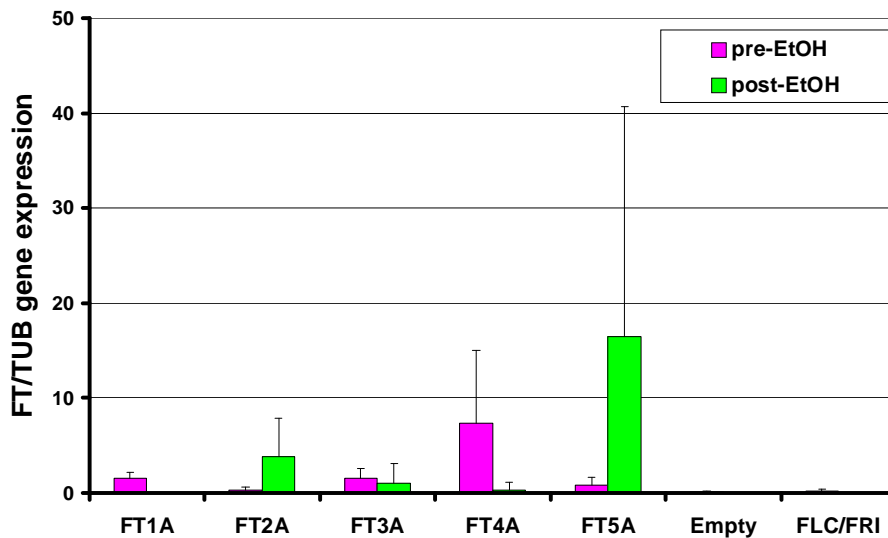


**Figure 4.2 Genotyping by PCR of the 5 transgenic lines of Arabidopsis.** The transgenic lines are numbered *FT-1A* to *FT-5A*. The empty signifies the absence of the *FT* gene, the *FLC/FRI* is the wild-type line, the *ft-10* was a T-DNA insertion line which lacks *FT* gene. Col-O is the Arabidopsis Columbia ecotype, the NTC was the water control, and the plasmid from which the transgenic lines were transformed. The amplification product size is 179 bp, the \* is over *FT* intron amplification in Col-O and  $\Phi$  are the non-specific amplification products.

#### 4.2.1.2 Ethanol induction

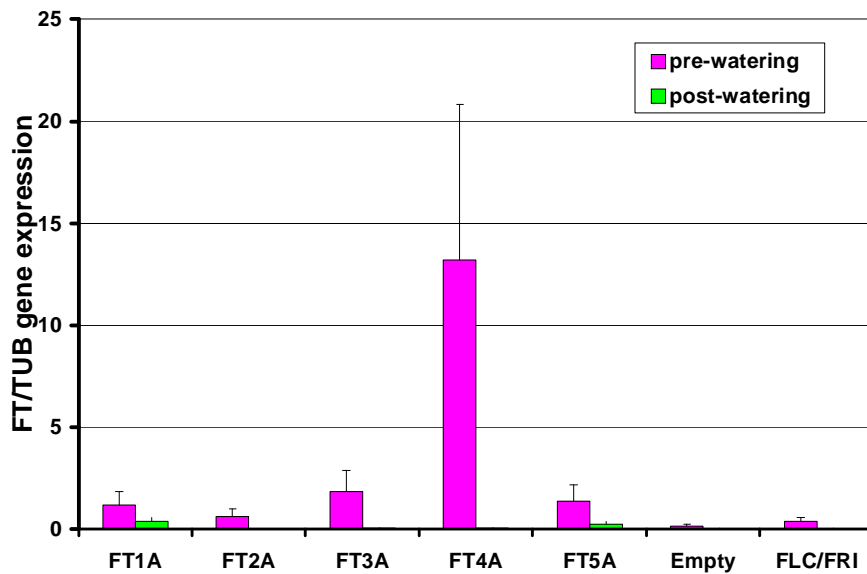
To the above herbicide resistant and PCR positive transgenics, ethanol was applied. The ethanol spray was found to drive the expression of *FT* gene (Figure 4.3). Without exogenous ethanol, the system was generally not induced. Therefore, I sprayed 1% ethanol solution from the 10<sup>th</sup> day after transplanting to soil. This spraying was continued for three and half days, at 4 different time points of the day, which were 9 am, 12 noon, 3 pm, and 6 pm. On the last day, apical tissues were harvested at 10 am, for *FT* expression analysis. The *FT* transcript level accumulation was monitored before and after induction (Figure 4.3). In the transgenic FT1A line, *FT* transcript level was low after 3 days of induction. FT2A, showed a 3 fold *FT* transcript accumulation after 3 days of induction. FT3A showed reduced *FT* expression after 3 days. *FT* transcript level in line FT4A was reduced in expression after the 3 days of ethanol application. FT5A showed the highest transcript accumulation level, and this was at 15 fold levels. The *FT* expression level in the empty control, and *FLC/FRI* lines were essentially undetectable as expected (Figure 4.3). Thus, *FT* regulation by ethanol induction increased *FT* transcript in lines FT2A and FT5A, whereas for lines FT1A, FT3A, and FT4A, increased *FT* transcript levels were not seen after induction (Figure 4.3). This shows that the system of induction by ethanol can be effective in some transgenic lines.

---



**Figure 4.3** *FT* was found to be expressed in *FLC/FRI* transgenic *FT* genotypes after spraying with 1%EtOH. FT1A, FT2A, FT3A, FT4A and FT5A were the ethanol induced transgenic lines used for the experiment. Empty is the vector without the *FT* transgene and *FLC/FRI* is the late flowering non-transformed genotype. The y-axis represents the normalized *FT* and *TUBULIN* expression. x-axis represents the different transgenic and control lines. Error bars represent  $\pm$  SEM.

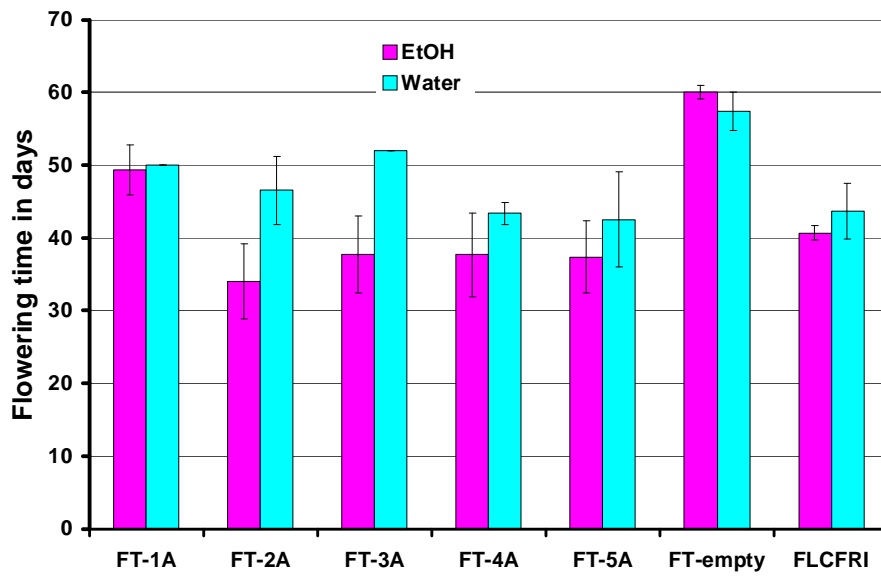
A parallel watering-control experiment was set up with the replicate plants that were the same genotypes. The transgenic lines were sprayed with distilled water to examine the requirement of ethanol induction on the system. The results from this watering-control experiment (Figure 4.4) indicated that lines FT1A, FT2A, FT3A, and FT5A had no induction by water addition. Line 4A had high *FT* levels before water spraying, and this mirrored what was seen in Figure 4.3. The controls empty, and *FLC/FRI* non-transgenic lines showed no detectable *FT* transcript accumulation. Therefore, the results from the watering-control experiments (Figure 4.4) showed that water was not capable of inducing the AlcR/AlcA system. This suggested to me that the ethanol-inducible system could be tightly controlled in some lines.



**Figure 4.4: *FT* is NOT-expressed in *FLC/FRI* transgenic genotype after spraying with distilled water.** FT1A, FT2A, FT3A, FT4A and FT5A were the transgenic lines used for the experiment. Empty is the vector without the *FT* transgene and *FLC/FRI* is the late flowering non-transformed genotype. The y-axis represents the normalized *FT* and *TUBULIN* expression. x-axis represents the different transgenic and control lines. Error bars represents  $\pm$  SEM.

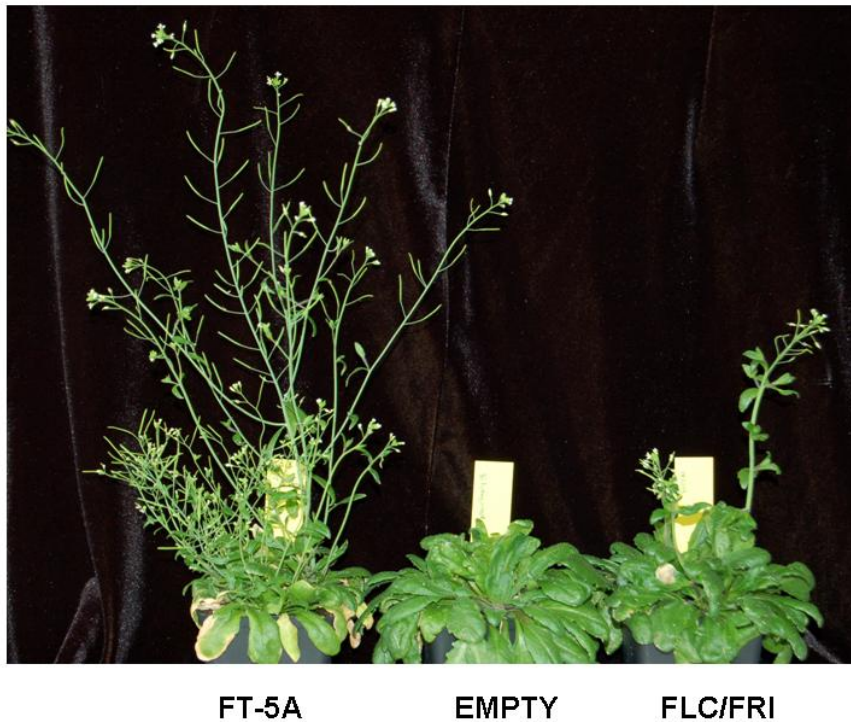
#### 4.2.1.3 Flowering time measurement

The 1% ethanol spray was completed after three and half days, and the plants were closely monitored to measure the timing of flowering. The plants started to flower as from a few days after spraying. Flowering time was measured by documenting the bolting dates for each transgenic lines, the control empty *FT* gene transgenic lines, and non-transgenic *FLC/FRI* genotypes, relative to germination dates (Figure 4.5). This flowering-time data was also compared to the data generated from the parallel watering-control experiments. Transgenic lines FT-2A, FT-3A, FT-4A, and FT-5A flowered earlier than the *FT*-empty and *FLC/FRI* controls. Figure 4.6 shows phenotypic expression of flowering in lines FT-5A, *FT*-empty and *FLC/FRI* 3 weeks after ethanol spraying. FT-1A showed a late-flowering phenotype, for reasons that are as of yet unclear to me (Figure 4.5). The watering-control plants all flowered later, when compared to the ethanol sprayed lines, as expected (Figure 4.5). Hence, the controls *FT*-empty and *FLC/FRI* flowered later than transgenic lines FT-2A to FT-5A in ethanol induced lines whereas FT-1A flowered later than the controls. Overall, ethanol was an effective inducer of some pNewFT lines and the *FT* gene was apparently able to by-pass the floral repressor *FLC/FRI*.



**Figure 4.5 Ethanol application promotes flowering in some pNewFT *FLC/FRI* transgenic lines.** FT1A to FT5A are the Arabidopsis *FT* transgenics. FT-empty is *FT* absent transgenic control and *FLC/FRI* is the non-transgenic control, all plants were under the *FLC/FRI* background. The y-axis is the flowering time in days measured as days from planting to first flower. The x-axis represents the samples. The pink bars are the ethanol sprayed lines while the blue are the water sprayed plants. Error bars represents  $\pm$  SEM.





**Figure 4.6: FT-5A transgenic line flowering along side with the empty control and *FLC/FRI* non-transgenic genotype 3 weeks after the ethanol induced *FT* was activated by 1% ethanol. *FT-5A* is the transgenic line representative, seen flowering profusely. The empty control in the middle has not started flowering and *FLC/FRI* non-transgenic has just started to flower.**

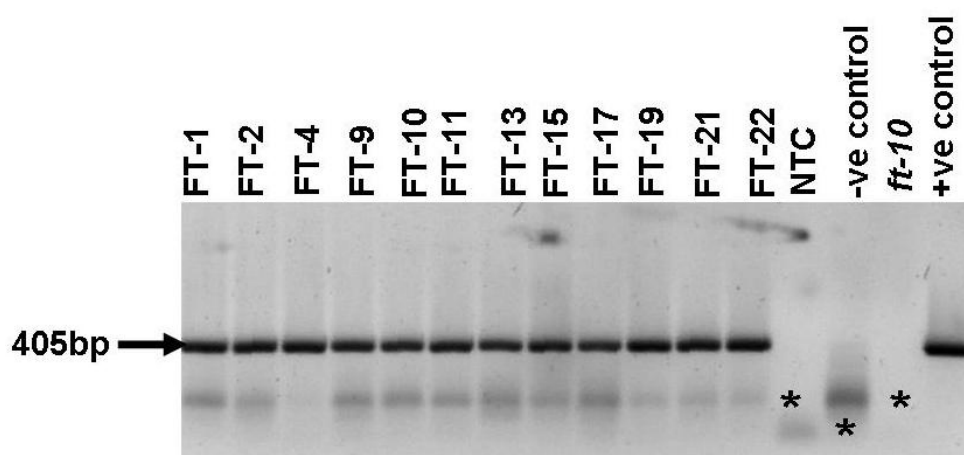
#### **4.2.2.0 Floral promotion from Arabidopsis *FLOWERING LOCUS T* in shy flowering genotype of Cassava**

The success of the ethanol-inducible system in the late flowering Arabidopsis genotype gave me the confidence to use the same ethanol-inducible *FT* construct to induce flowering in shy flowering genotype of Cassava. For this, Friable Embryogenic Callus created from the Cassava cultivar TMS 60444 was transformed by Agrobacterium-mediated transfer of the T-DNA carrying the same pNewFT construct described above. See transformation details from the Materials and Methods of this thesis (Section 2.2.1). Fifteen independent transgenic events were regenerated. These transgenic plants were generated by a closely examined, time consuming, and laborious selection. The transformed FECs were under stringent selection from the herbicide selectable marker. The selection period was for 9 months from the friable embryogenic callus stage, through to the different plant regeneration phases. Finally, plantlets were generated. Only twelve independent transgenic lines survived to the green house for physiological and molecular analysis.

---

#### 4.2.2.1 Genomic PCR analysis of *AtFT* in Cassava transgenics

The establishment of the twelve independent events was attained in the green house. Leaf samples were collected for molecular analysis of the putative-transformed lines by conventional PCR. This was necessary to confirm if these putative transgenic lines carried the transgene from the T-DNA insertion, and that there was stable integration of the *FT* gene within the genome of Cassava. DNA was extracted from all twelve independent events, the DNA preparation for a non-transgenic Cassava wild-type controls, and Arabidopsis *ft-10* mutant, were produced. From this, PCR was performed with primers against Arabidopsis *FT* with these Cassava DNA samples, along-side the control of the plasmid DNA of pNewFT, which was the construct from which all the Cassava transgenic lines were generated. After PCR, all twelve independent transgenic events showed the expected 405 bp amplification fragment. Thus, the *FT* cDNA was incorporated within their genomes. As expected, the wild-type Cassava plant did not generate an amplification fragment (Figure 4.7). The *ft-10* Arabidopsis mutant, as expected, failed to generate the 405 bp fragment, as expected. The plasmid DNA sample generated the expected 405 bp fragment. These results demonstrated the presence of the *FT* transgene in these 12 transgenic Cassava lines. Therefore, the regeneration of the Cassava lines from FECs and the incorporation of the *FT* within the T-DNA insertion was a success.

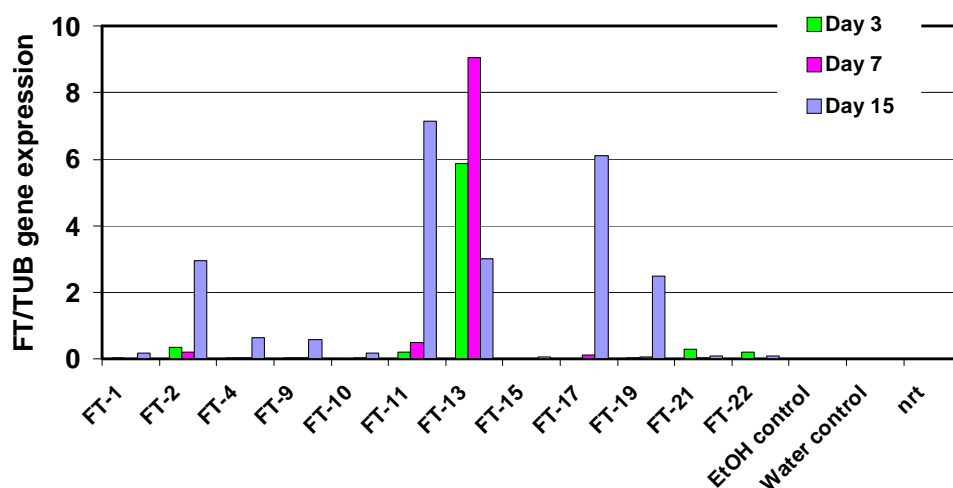


**Figure 4.7 Genomic PCR of Arabidopsis *FT* in Cassava transgenics** FT-1 to FT-22 represents the 12 Cassava transgenic lines, the NTC is water control used which lacked any DNA template, -ve is the negative control from the non-transgenic Cassava line, *ft-10* is an *FT* mutant and +ve control is PCR against the plasmid from which the transgenic plants were generated. \* shows the non-specific amplifications in these samples.

---

#### 4.2.2.2 Gene expression of *FT* transgene in Cassava by RT-PCR analysis

The Cassava transgenic lines were grown for 3 months and then the induction system was activated. The method of induction of the *FT* transgene was by ethanol watering and spraying, to drive the *AtFT* expression. For this, 50 ml of 2.5% EtOH was soil drenched once in a day at 12 noon and 2.5% ethanol was sprayed copiously at 12 noon, 3 pm, and at 9 pm for 15 days. The transcript level of *FT* accumulation was measured in a progressive manner by reverse transcriptase PCR (RT-PCR). This was done by analyzing leaf samples from the apical part of all the transgenic lines. The samples were harvested on the 3<sup>rd</sup>, the 7<sup>th</sup>, and the 15<sup>th</sup>, day respectively, following the commencement of the last daily application of ethanol. The result from the (RT-PCR) data shows that there was little to no *FT* expression detected after the spraying of the 2.5% ethanol on the 3<sup>rd</sup> and 7<sup>th</sup> day of the start of the induction. However, on the 15<sup>th</sup> day, the expression level of *FT* was readily detected for most of the Cassava transgenics (Figure 4.8). It is worthy of note that transgenic line FT-15 showed low expression, compared to the control lines (Figure 4.8). Strangely, *FT*-13 expression at day 3 was higher than the expression at day 15, and it records the highest level of expression at Day 7 after commencement of spraying. The reason for this anomaly is unknown (Figure 4.8).



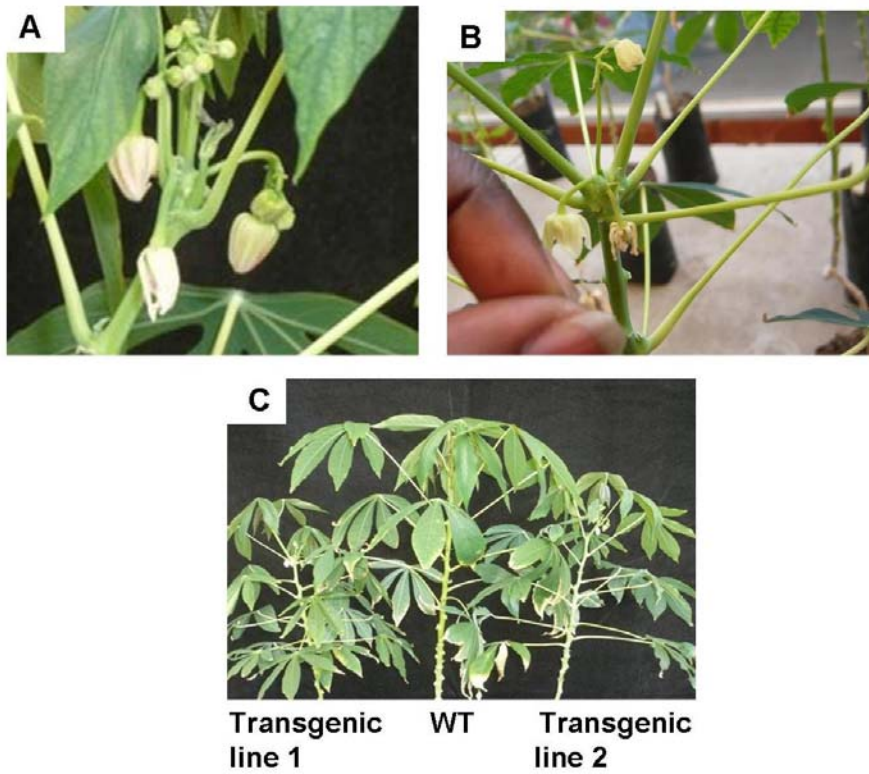
**Figure 4.8 *FT* expression in Cassava transgenics 3,7, and 15 days after induction with 2.5% ethanol.** Gene expression of *FT* transgenic lines 1 to 12 after 3,7 and 15 days of ethanol spray, EtOH control was the non-transgenic line sprayed with ethanol, Water control was the non-transgenic Cassava plant inducing with water and nrt was cDNA not transcribed by reverse transcriptase.

---

#### 4.2.2.3 Observations in the greenhouses

Flowering is a phenotypic expression of integrated network of endogenous and environmental cues. For Cassava to flower in the green house is very uncommon. The conducive flowering environment known to Cassava is under field conditions (Byrne , 1984). The independent *FT* transgenic lines under the control of an ethanol-inducible promoter were grown in the glass house at two different locations, in two different continents, and under different physiological growth conditions. I made phenotypic observation of the Cassava transgenic lines in glass house in Cali, Colombia and in a green house in Cologne, Germany. The glass house in Cali was subjected to natural day length of essentially a seasonal constant of 12 hours. Cali is essentially at the equator, at 4°N latitude and Cologne is at 51°N latitude. I observed that the transgenic lines in the glass house in Colombia started flowering mostly at 8 months after planting, some even without ethanol application (Figure ). In Cologne, the plants in the green house, which were genetic replicates as the transgenic lines in Cali, flowered comparatively earlier. The temperature of the German greenhouse and the photoperiod was regulated, with the temperature being between 25-30°C and the photoperiod being at 16 hours of light and 8 hours of darkness. It was observed that most of the transgenic lines started flowering by the third month under these German green- house conditions (Figure 4.12). This is consistent with previous observations that long day photoperiods promote flowering in Cassava (Keating *et al.*, 1982a, 1982b).

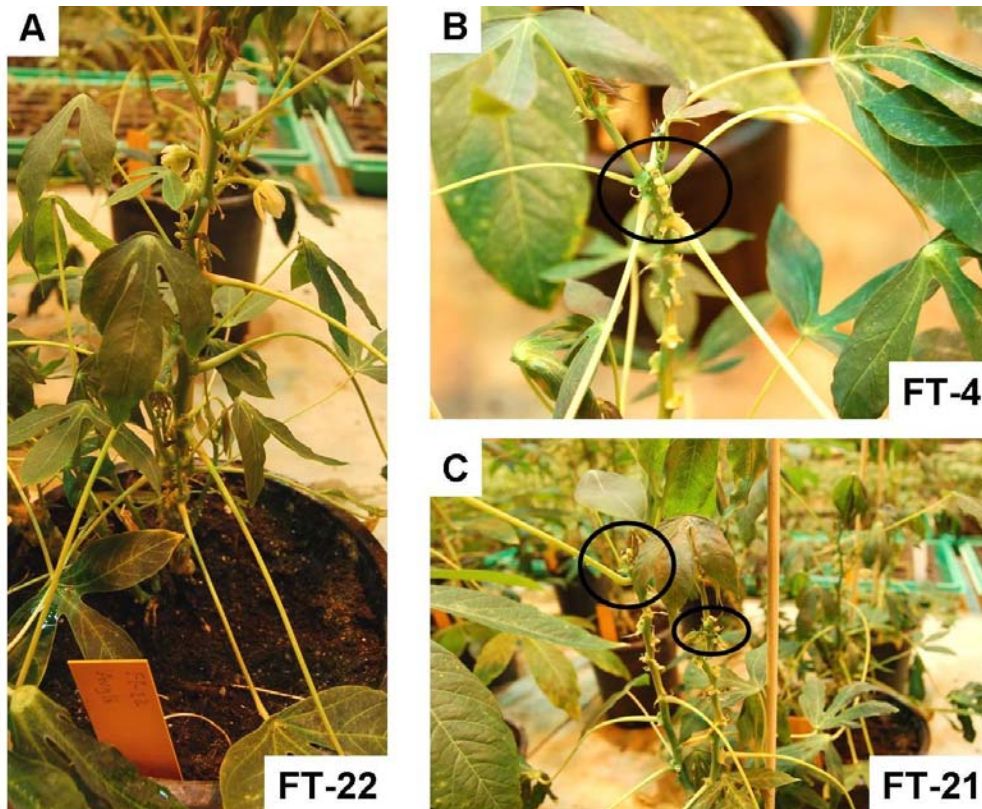
Ethanol spraying promoted flowering in some Cassava transgenics. Notably, in lines FT-4, FT-21 and FT-22, floral initiation was observed 2 weeks after the completion of the ethanol spraying (Figure 4.11). The floral behavior of these 3 lines is typical of the developmental response expected as a result of ethanol-induction of the system. This flowering plants after 2 weeks of ethanol spray presents evidence of the efficiency of the system.



**Figure 4.9 Cassava plants in the glass house in Colombia.** (A and B) Displays two of the transgenic lines flowering without ethanol induction after 8 months. (C) The non-flowering control in the middle of two Cassava flowering transgenics.



**Figure 4.10** Cassava plants in the green house in Germany. (A) The non-flowering control after 3 months in the green house. (B) One transgenic line flowering without ethanol induction (C) Close up view of the non-flowering apices of a non-transgenic control.



**Figure 4.11 Cassava transgenics flowered after induction of the ethanol controlled *FT* transgene system by ethanol.** Plants were grown in the green house in Cologne. FT-22, FT-4, and FT-21 are the transgenic lines that flowered 2 weeks after 2.5% ethanol was sprayed on them to induce the system. The circle represents the floral zones.

### 4.3 Conclusion

One dream for breeders would be that flowering could be induced at will in any plant system, given the appropriate molecular tools. In a tropical species like Cassava, which is known to be recalcitrant to flowering, I showed that flowering time can be manipulated by the ethanol-inducible system of the pNewFT construction. I showed that the *FT* transcript accumulated after 15 days of induction by ethanol. This led to floral initiation in Cassava. Although the physiological processes responsible for flowering in some of the Cassava transgenic lines before ethanol induction is not yet clear. This pre-ethanol induction of flowering may be as a result of photoperiodic response of *FT* transgene, or that the ethanol system had leaky expression. Still, in some transgenic events, flowering time was initiated by chemical induction. The success of being able to induce flowering at will in Cassava by this transgenic approach becomes an invaluable tool for the Cassava breeders to be able to synchronize flowering for Cassava improvement.

---

**CHAPTER 5. GENERAL CONCLUSIONS AND PERSPECTIVES**



---

## 5.0 Introduction

In this thesis, my primary aim was to characterize photoperiod aspects of flowering in Cassava, a tropical plant in which very little is known about these response pathways. I identified several photoperiod genes. The knowledge of the characterized photoperiod genes was further used in the aim of modifying flowering time in Cassava. This is important owing to its recalcitrant flowering nature. Inducible flowering was achieved by genetically modifying Cassava with the floral integrator pathway gene *FLOWERING LOCUS T (FT)*, which is known to be implicated in the Arabidopsis flowering pathway. A detailed study of the characterization of photoperiod genes and the manipulation of flowering time in Cassava is summarized.

### 5.1 *MeGI* Summary

In Chapter 3, I provided molecular information on the *GIGANTEA* homologue in Cassava. I isolated a partial *MeGI* sequence and showed the encoded protein has extensive sequence conservation with GI from other plants. A multiple alignment of the GI encoded-protein sequences from other species, including that of *MeGI*, showed a clear separation of the dicots from the monocots (Figure 3.3). *MeGI* claded closely with the GI of Castor bean. I also measured the expression level of this photoperiod gene in Cassava and found the transcript level of *MeGI* to peak at 12 and 16 hours after dawn under subjective long day conditions. Under short day conditions, it peaked 8 hours and 12 hours after dawn (Figures 3.4 and 3.5). This finding is similar to observations in *Hordeum vulgare*, a long-day monocot plant, where *GI* was found to peak at 6 h and 9 h after dawn under short day conditions and 15 h after dawn under a 16 h long day period (Griffiths *et al.*, 2003). Therefore, *MeGI* sequences are conserved across species and share similar expression patterns.

#### 5.1.1 Future perspectives on *MeGI*

The first future need in the characterization of the *MeGI* gene is to complete the amplification and sequencing of the 5' end. This could be done either by optimization of a 5' RACE (Random Amplification of cDNA ends) protocol (Roche), which was earlier used and/or to continue to amplify the Cassava genome, or to assemble small fragments generated through inverse PCRs. Upon completion of the sequencing of the whole *GI* gene in Cassava, a functional analysis by complementation with *gi* mutants in Arabidopsis should be conducted. Here, I propose that *MeGI* gene is transformed into the *gi* mutants of Arabidopsis. Experiments would include measurements of flowering time, circadian function, and light signaling of all known phenotypes of this mutant (Oliverio *et al.*, 2007). This will serve to predict its function and make conclusions on the functional properties of the *MeGI*, if it is indeed a true orthologue of Arabidopsis *GI*. Studies

---

---

should be conducted and to examine if *MeGI* is also a nuclear protein in Cassava, which could define the generality of its localization. Collectively, this will start to assign a photoperiodic role for the *MeGI* in Cassava.

## 5.2 Summary of the *COLs*

In Chapter 3, I isolated partial sequences of three of the *CO*-like genes in Cassava. I compared their sequences with the corresponding *CO* genes of Arabidopsis. The results showed gene conservation across species (Figure 3.7), conserved domains have been shown in other *CO*-like genes across species (Lagercrantz and Axelsson, 2000; Holefors *et al.*, 2009; Robert *et al.*, 1998; Song *et al.*, 1998; Yano *et al.*, 2000; Serrano *et al.*, 2009; Jeong *et al.*, 1999; Zobell *et al.*, 2005; Chia *et al.*, 2008 and Liu *et al.*, 2001). Next, I measured in Cassava the transcript level of these *CO*-like genes, and I was able to show that *COL* transcript levels peaked at dawn, in anticipation of daylight under long day conditions, and under short day conditions, *COL* transcription levels were found to be elevated 2 hours before dawn (Figures 3.9 and 3.10). This pattern of expression in Cassava is similar to the related *COL* transcript levels in Arabidopsis (Suárez-López *et al.*, 2001), again showing sequence conservation between both Cassava and Arabidopsis. Perhaps there is a conservation of *CO*-like expression in the timing of flowering control. To test for this, it would be worthy to examine the *GI-CO-FT* relationship in Cassava. Here, physiological experiments should be conducted to see if *GI* expression correlates with the expression of any of the *COL* genes, and whether such *COL* expression under a light phase can induce *FT*. Cassava *FT* awaits full characterization for this to be completed. Since in Arabidopsis, *GI* expression activates the transcription of *CO* mRNA, and when *CO* transcript is under the light phase, this directs the transcription of *FT* mRNA which triggers flowering (Suárez-López *et al.*, 2001; Searle and Coupland, 2004; Jaeger and Wigge, 2007), similar processes might be involved in the transition to flowering in Cassava. As photoperiod advances can promote flowering in Cassava (Keating *et al.*, 1985), this is plausible.

### 5.2.1 Future perspectives on *COL* genes

We understand from the literature that the photoperiodic control of flowering is an important adaptive characteristic in plants (Laurie *et al.*, 2004). Studies from the model dicot Arabidopsis showed that the *AtCO* (*CONSTANS*) is important in the photoperiod pathway (Turck *et al.*, 2008). It is not surprising that Cassava *COL* transcripts were not detectable above background levels during the day phase, and this may be consistent with the findings of Suárez-López (2001) that *CO* in Arabidopsis is low during the day. Here, I reported three partial *MeCOLs* sequences, and I recommend that the existing ORFs should be completed and characterized. Notably, it is unclear how the effect of the *COL* expression under short day and long day influences flowering in Cassava. It was noticed that there is no detectable transcript of the *MeCOL* during the light phase

---

of the day (Figure 3.9 and 3.10). It is known in Arabidopsis that transcript accumulation of *CO* during the light influences flowering by upregulating *FT* expression (Searle and Coupland, 2004). Therefore, a detailed study of the function of the Cassava *COL* gene expression patterns and their influence on flowering time in Cassava is recommended.

### 5.2.2 *MeELF4* Summary

In Chapter 3, I discovered and characterized the complete *MeELF4* gene. First, I identified this gene, and concluded by phylogeny that *MeELF4* is closely related to the Arabidopsis *ELF4*. I used it in complementation tests in Arabidopsis to confirm orthology, and this was conducted with the *elf4-1* mutant. Here, restoration of the expression patterns of *CCA1:LUC* and *CCR2:LUC* markers of clock function was found (Figures 3.14 and 3.15). My studies led to the conclusion that *MeELF4* rescued *elf4-1*, as rhythmicity in the *MeELF4:elf4-1* line was sustained under free run (Figures 3.15 and 3.16). Data from assays of hypocotyl-length elongation also showed that *MeELF4* restored growth to a level comparable to the wild-type, as the *MeELF4:elf4-1* lines were shorter than the *elf4-1* mutant (Figure 3.16). Leaf movement studies also confirmed this hypothesis, as leaves of *MeELF4:elf4-1* plants showed rhythmic leaf movements under circadian cycles, while that of the *elf4-1* mutants showed arrhythmicity (Figure 3.14). Thus, I found the Cassava orthologue of Arabidopsis *ELF4*.

Seeds from the T<sub>1</sub> generation of *MeELF4:elf4-1* were planted under long day conditions to observe the flowering-time physiology. These transgenic lines flowered under long-day conditions, but there was no difference in flowering-time amongst the *MeELF4:elf4-1* lines, the positive Ws-2 wild-type controls and the *elf4-1* mutant. Further tests under short day conditions would reveal further complementation of flowering-time, as the *elf4-1* mutant is most phenotypically perturbed under such an environment (Doyle *et al.*, 2002). This investigation should reveal the ability of *MeELF4* to complement the early flowering defect of *elf4-1*.

### 5.2.3 Future Perspectives on *MeELF4* gene

*ELF4* is implicated in the circadian clock (McWatters *et al.*, 2007; Khanna *et al.*, 2003; Doyle *et al.*, 2002; Doyle, 2003). It is recommended that a detailed flowering-time analysis of the *MeELF4* complementation of *elf4-1* lines under short day conditions be investigated.

There is a significant difference in flowering-time in Arabidopsis when grown under long days and short-days given that the floral transition occurs earlier when plants are grown in long days than when they are grown in short days (Martinez-Zapater, 1994). *ELF4* plays a role in the sensing of long day and short days (Doyle *et al.*, 2002; McWatters *et al.*, 2007). It is recommended that a replicate biological experiments of the *elf4-1* mutants with the *CCA1:LUC* and *CCR2:LUC* markers be repeated to confirm the complementation results described in Chapter 3. It is also recommended that the T<sub>2</sub> segregating population of the *MeELF4* transgenic lines be investigated

---

---

further. This is necessary for further confirmation of the accuracy and consistency of the molecular and genetics findings from the T<sub>1</sub> generation using the “luciferase platform.” The leaf movements and hypocotyl elongation experiments may be repeated for confirmation of the reported results in this thesis. Therefore, more investigation is recommended in the confirmation of *MeELF4* as a true orthologue of Arabidopsis *ELF4*.

### 5.3 Perspective on genes of interest and Cassava flowering model

The expressed sequence tags and genome sequence databases searches used to identify flowering related gene sequences from Arabidopsis and Castor bean generated molecular information for *MeAPI*, *FT*, and *SOC1*. These molecular sequencing information was not enough to generate contigs. Therefore, the expression levels of *MeAPI*, *FT*, and *SOC1* was not measured in Cassava. It is recommended that further efforts be made to discover *FT*, *API* and *SOC1* genes in Cassava.

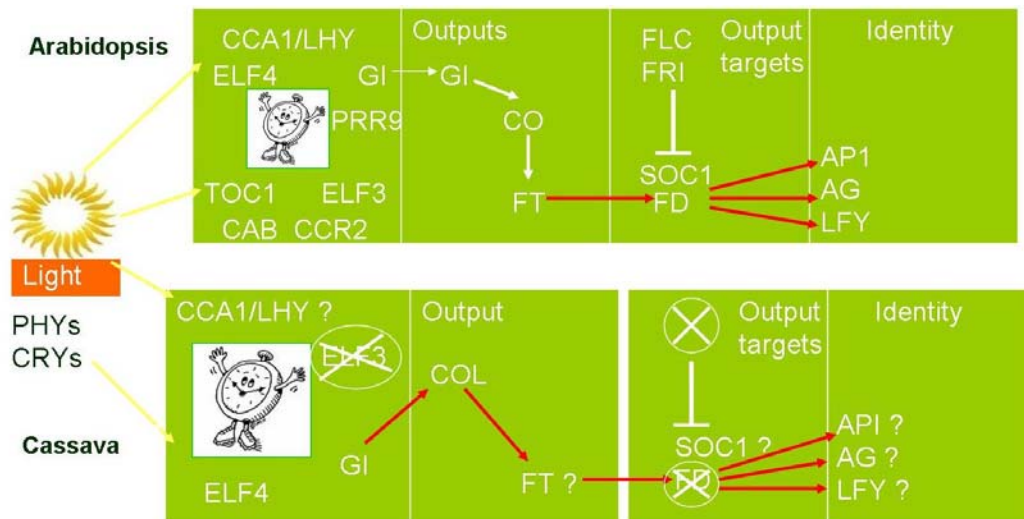
I was unable to identify the bZIP transcriptional factor *FD* (Figure 5.1), which mediates floral signal from the integrator genes *FT* to induce flowering (Wigge *et al.*, 2005). The *ELF3* clock gene (Zagotta *et al.*, 1996) was also unidentified (Figure 5.1). Therefore, for my interest in direct comparison of the physiological genetics of flowering of the model system of Arabidopsis with that of Cassava which lacks molecular information on flowering, a comparative flowering model for Cassava and Arabidopsis was of yet unclear (Figure 5.1).

Several photoperiod-related genes were detected in Cassava and their characterization could be furthered. A gene sequence with identity to the Arabidopsis *LHY/CCA1* was detected in Cassava. The sequence information was incomplete (data not shown). Similarly, the floral integrator pathway gene *AGL24* (Yu *et al.*, 2002) was detected. I obtained partial sequence information and this led to an incomplete ORF (data not shown). Both *MeLHY/CCA1* and *MeAGL24* thus are potential homologues to the related genes of Arabidopsis (sequence details not shown). A more accurate molecular technique for identifying and sequencing the 5' and the 3' ends of these genes is recommended. Upon the complete sequencing of these genes, a detailed characterization should be thoroughly investigated in order to be able to conclude their gene for gene and/or gene by function, for example by complementation tests against the *lhy* mutants of Arabidopsis. The mutant lines rescued by the *MeLHY/CCA1* and *MeAGL24* genes might display phenotype similar to that of the wild-type, which is evidence of successful complementation.

#### 5.3.1 Floral repressors

I was unable to detect any sequences related to the *FLC* and *FRI* genes in Cassava. *FRI* encodes a protein of an unknown function (Johanson *et al.*, 2000), and *FLC* is a MADS box transcriptional regulator that quantitatively represses flowering (Michaels and Amasino, 2001). *FRI* represses

flowering only in the presence of *FLC* activity (Koorneef, 1994). Studies have shown that the *FLC* and are only present in the Brassicaceae family (Roux *et al.*, 2006), which might explain why I could not detect them in Cassava. Perhaps the strong floral repression in this tropical plant is genetically distinct from *FLC* repression of temperate Brassicaceae.



**Figure 5.1 A Comparison of the flowering model of Arabidopsis and a comparison to Cassava.** The model is in comparative parts. Each part is composed of the clock, its output, output targets, and identity genes. The yellow and red arrows signify activators of expression, the white arrows, repressors and the question mark shows sequence or function uncertainty. The circles signify genes not found in Cassava. The genes are annotated in the abbreviations. In the Cassava model, *ELF4*, *GI*, *COL*, *CCA1/LHY*, *SOC1* complete and partial sequences have been found, while *ELF3*, *FT*, *FD*, *API*, *AG*, *LHY*, *FLC*, and *FRI* have not been found.

#### 5.4 Summary of Arabidopsis Floral induction

In Chapter 4, I was successful in inducing flowering in the *FLC/FRI* genotype (Michaels and Amasino, 2000) with *FT* gene under the control of an ethanol-inducible promoter (Caddick *et al.*, 1998). This Arabidopsis genotype under long day conditions flowered after the inducible system was activated by the exogenous application of 1% ethanol solution (Figure 4.7). This ethanol-inducible system was well regulated in some Arabidopsis. A concurrent watering-control experiment with the ethanol-inducible test sample plants showed late-flowering phenotype when compared to its ethanol-induced counterparts (Figures 4.3 and 4.4). This showed that water could not induce this ethanol system of induction. In conclusion, the ethanol inducible *AlcA/AlcR FT* construct I designed was functional and could bypass the “shy” flowering phenotype of *FLC/FRI*. These pNewFT *FLC/FRI* Arabidopsis lines could serve as a useful resource to understand the competition of *FLC* and *FT* on their target gene *SOC1*.

---

#### **5.4.1 Future perspectives on the *Arabidopsis FLC/FRI* genotypes**

I recommend that the experiment with the *FLC/FRI* genotypes be repeated under short day conditions. The flowering-time measurement was conducted under a subjective long day conditions known to induce flowering. It is important to observe the timing of floral induction under non-inductive photoperiods, because the inductive photoperiodic effects on the flowering time will be physiologically suppressed.

To further confirm the biological activity of the ethanol-inducible FT within the pNew construct, complementation tests against the *ft-10* mutants are proposed. This could be done by transforming pNewFT into *ft-10* and measure flowering time before and after ethanol application. *ft-10* is a mutant line (Yoo *et al.*, 2005), and is a null allele in the Col. background. This will confirm that my pNewFT construct can rescue this different late flowering *ft-10* mutant and would further verify the efficiency of the ethanol promoter inducible system. (Caddick *et al.*, 1998; Roslan *et al.*, 2001; Filichkin *et al.*, 2006; Garoosi *et al.*, 2005).

#### **5.5 Summary of Cassava Floral induction**

In Chapter 4, I generated 12 independent transgenic lines of Cassava transformed with the *FT* gene under the control of the ethanol-inducible AlcA promoter. I showed that *FT* gene was incorporated into the Cassava genome by conventional PCR against genomic DNA and by detection of *FT* transcript expression (Figures 4.8 and 4.9). The gene expression trend of *FT* transcript accumulation after application of 2.5% ethanol solution over a period of 15 days showed dramatic *FT* transcript accumulation (Figure 4.9), except for line FT-13 which showed no stable pattern of transcript accumulation after induction (Figure 4.9). Cassava transgenic lines FT-4, FT-21, and FT-22 had not started flowering at the time the 2.5% ethanol spraying began, but were seen to flower 2 weeks after the 15 days of ethanol spraying had ended (Figure 4.9). The timing of flowering observed in lines FT-4, FT-21, and FT-22 is typical of what was expected from this system of induction, such that floral initiation be “turned on” only after the application of the inducer. Lines FT-4, FT-21, and FT-22 are the best of all 12 independent events. Therefore, I was able to make a shy flowering Cassava genotype flower after induction by ethanol (Figure 4.12).

#### **5.5.1 Future perspectives on the Cassava transgenic lines**

I was been able to show that Cassava can be induced to make flowers through the use of *FT* transgenic approach. It is worthy of note that in the 20 years history of Cassava study at the International Center for Tropical Agriculture (CIAT), flowering in the green-house has never been observed (Fregene, personal communication). There, Cassava flowers in tropical environments only under field conditions, and then, only when the plants are over 6 months in age. I noticed that 75% of the 12 transgenic lines tested, flowered by the third month in the green house, and before

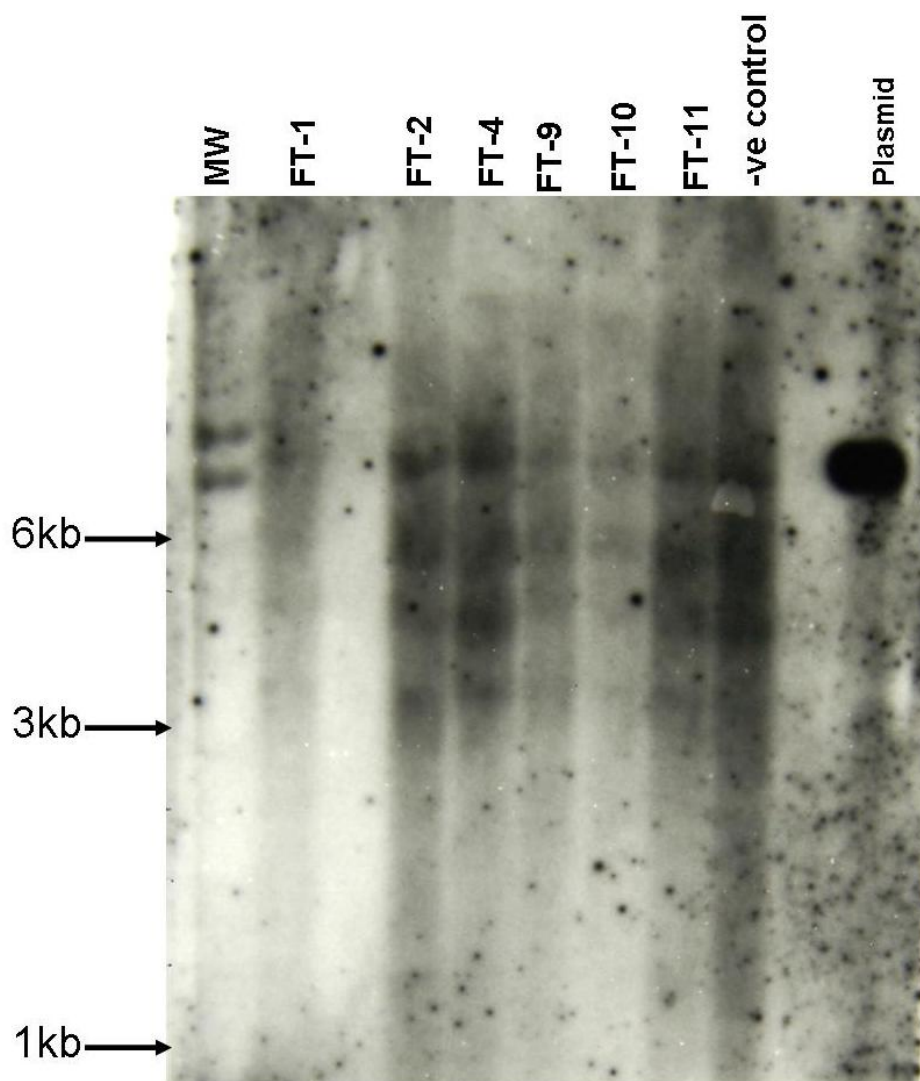
---

---

the ethanol induction was started with 2.5% ethanol solution. Therefore, transgenic *FT* Cassava can flower in the green house in response to photoperiod.

It was a general trend amongst the transgenic Cassava lines that the *FT* was induced after ethanol addition on the 15<sup>th</sup> day of spraying (Figure 4.9). One exception was line FT-13, whose expression level did not increase progressively. It recorded the highest *FT* expression level on the 7<sup>th</sup> day, followed by the expression on the 3<sup>rd</sup> day and the lowest of its expression was on the 15<sup>th</sup> day in which the *FT* expression was expected to be the highest. This could be due to the leakiness of the promoter in this particular transgenic line. It was reported in tobacco that the ethanol-inducible promoter system appears to have negligible levels of leaky expression, even though they recorded the system was not induced by endogenous alcohol production in response to water stress (Roslan *et al.*, 2001). This anomaly in this single line could also be due to the number of copies incorporated with the genome.

In order to verify the incorporation of the *FT* gene from *Arabidopsis* within the genome of the Cassava transgenic lines, Southern-blot hybridization was conducted. Cassava genomic DNA was digested with *HindIII*. The *HindIII* is not predicted to have any restriction site within the T-DNA region. The PCR product described from the genomic PCR amplification described in Figure 4.8 was labeled and used as the probe. The result showed many bands per line. However, the banding pattern of the hybridization signal were indistinguishable from the wild type (Figure 4.10). This may suggest the detection of the presence of homologous *FT* genes in Cassava (Fig. 5.2). To establish the integrity of the transgenic lines, a detailed examination of the T-DNA insertion is proposed as a method to distinguish the wild-type from the putative transgenic lines. The T-DNA is 5670 bp and the *HindIII* cuts once between the selectable marker gene (*pat*) and *AlcA* promoter border, which generates fragments at the expected size of 2.2 and 3.4 kb, respectively. This was also seen to be present in the wild-type, although this may suggest that the *FT* of Cassava would have homologous sequences of the *Arabidopsis FT*. Therefore, I recommend that the Southern-blot hybridization experiment be repeated using a probe from the T-DNA derived sequences of Left Border or the Right Border of the T-DNA insertion regions of the pNewFT plasmid. This will provide more specificity of hybridization due to the absence of these borders in the Cassava wild-type and confirm if these Cassava lines are transgenic or not.



**Figure 5.2 Southern blot hybridization on Cassava transgenic lines with *FT* probe.** 50 µg of DNA was digested by 125 units of *HindIII*. DNA was separated by electrophoresis, and a blot was made. It was hybridized to *FT* probe by radiolabelled  $^{32}\text{P}$  dCTP. MW is the 1kb ladder, FT-1 to FT-11 are the transgenic lines representative, -ve control is the wild-type Cassava non-transgenic, and the +ve control is the digested plasmid with *HindIII*. The arrows signify the corresponding molecular weight at each level.

I recommend that a detailed study of the interactions between the endogenous *MeFT* and the *FT* from Arabidopsis be conducted. One could investigate the photoperiodic influence on both the transgenic and wild-type Cassava plants. This investigation may include, first the isolation and then the sequencing of *MeFT* genes in Cassava. The sequencing results can be used to design primers for the *MeFT* transcript profiling in the wild-type Cassava and in the transgenic lines that harbor Arabidopsis *FT*. In the latter, it will be interesting to see if the induction of Arabidopsis *FT* in the pNewFT transgenic Cassava cause changes in the expression of endogenous Cassava *FT*



---

genes. Furthermore, the *MeFT* transcript level could be correlated to *MeGI* and *MeCOL* transcript levels, respectively, in the photoperiodic pathway to investigate the induction of flowering in Cassava.

I recommend the use of lines FT-4, FT-21, and FT-22 for future field trials as they flowered in response to the ethanol induction of the transgene. This implies that these lines will not flower before they are induced by ethanol, and thus, the breeder can be sure the plants will flower after 2 weeks of ethanol induction in order to make desired crosses. It would also be useful to attempt crosses with these transgenics to test whether the flowers are fertile and can form fruits and viable seeds. Collectively, I suggest that ethanol spraying using a mechanical approach under field conditions is appropriate for even and homogenous distribution of the ethanol.

In conclusion, there is a need to identify more photoperiod genes in Cassava. Furthermore, the incomplete molecular information on the genes I discovered from this study should be completely characterized. In this era, more than 180 genomes of organisms have been sequenced, ranging from bacteria to plants to animals (Goodner *et al.*, 2001; Dietrich *et al.*, 2004; AGI, 2000, Venter *et al.*, 2001; Yu *et al.*, 2002), and still on-going genome sequencing projects, of which Cassava is one. Since the sequencing of the Cassava genome is aimed at the identification and study of gene function, this molecular information generated from this study will contribute to the understanding of these genes and complement efforts in understanding the growth, development, and maintenance of this tropical plant, as a consequence of understanding the molecular and genetic detail of the photoperiodic effect of flowering in Cassava. This could lead to a breakthrough in understanding Cassava's transition from vegetative propagation to reproductive development. Finally, the physiological and morphological characterization of the Cassava transgenic lines generated from this investigation under field conditions will demonstrate the potential of the improvement and its invaluable resource for the breeders.

---

## **CHAPTER 6. LITERATURE**

- 
- Abate, T., Huis, A., van Ampofo, J.K.O.** (2000) Pest management strategies in traditional agriculture, An African perspective. *Annual Review of Entomology*, **45**, 631-659.
- Abe, M., Kobayashi, Y., Yamamoto, S., Daimon, Y., Yamaguchi, A., Ikeda, Y., Ichinoki, H., Notaguchi, M., Goto, K., Araki, T.** (2005) *FD*, a bZIP protein mediating signals from the floral pathway integrator *FT* at the shoot apex. *Science*, **309**, 1052-1056.
- Alaux, J.P., Fauquet, C.** (1990) African Cassava Mosaic Disease: from knowledge to control. Yamoussoukro, *Ivory Coast 4-8 May 1987, Seminar summary report.*, 50.
- Allem, A.C.** (1994) The origin of *Manihot esculenta* Crantz (Euphorbiaceae). *Genetic Resour. Crop Evol.*, **41**, 133-150.
- Allem, A.C.** (2002) The origin and taxonomy of cassava. In: Hillocks, R. J., M.J. Thresh and A.C. Bellotti (eds.) *Cassava: Biology, production and utilisation*. CABI International, Oxford, 1-16.
- Alves, A.A.A.** (2002) Cassava botany and physiology, In: Hillocks, R. J., M.J. Thresh and A.C. Bellotti (eds.) *Cassava: Biology, production and utilisation*. CABI International, Oxford, 67-89.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J.** (1990) Basic Local Alignment Search Tool. *Journal of Molecular Biology*, **215**, 403-410.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J.H., Zhang, Z., Miller, W., Lipman, D.J.** (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, **25**, 3389-3402.
- Amasino R.M.** (2005) Vernalization and flowering, *Curr. Opin. Biotech.*, **16**, 154-158.
- Andersen, C. H., Jensen, C.S., Petersen, K.** (2004) Similar genetic switch systems might integrate the floral inductive pathways in dicots and monocots. *Trends in Plant Science*, **9**, 105-107.
- Angelov, M.N., Sun, J., Byrd, G.T., Brown, H., Black, C.** (1993) Novel characteristics of cassava (*Manihot esculenta* Crantz), a C3-C4 intermediate photosynthesis species. *Photosyn Res*, **38**, 61-72.
- Asiedu, R., Bai, K.V., Terauchi, R., Dixon, A.G.O., Hahn, S.K.** (1992) Status of wide crosses in Cassava and Yam. In: Thottapilly, G. (ed.) *Biotechnology; enhancing research on tropical crops in Africa: Proceedings of an international conference held at the International Institute of Tropical Agriculture*, 26-30 November 1990, IITA, Ibadan, Nigeria.
- Awolaye, F., Van Duren, M., Dolezel, J., Novak, F.J.** (1994) Nuclear DNA content and in vitro induced somatic polyploidization (*Manihot esculenta* Crantz) Cassava breeding. *Euphytica*, **76**, 195-202.
- Bellotti, A.C., Smith, L., Lapointe, S.L.** (1999) Recent advances in cassava pest management. *Annu. Rev. Entomol.* **44**, 343-370.
- Blazquez, M.A., Soowal, L.N., Lee, I., Weigel, D.** (1997) *LEAFY* expression and flower initiation in *Arabidopsis*. *Development*, **124**, 3835-3844.
- Blazquez, M.A., Green, R., Nilsson, O., Sussman, M.R. and Weigel, D.** (1998) Gibberellins promote flowering of *Arabidopsis* by activating the *LEAFY* promoter. *Plant Cell*, **10**, 791-800.
- Blazquez, M.A., Trenor, M., Weigel, D.** (2002) Independent control of gibberellin biosynthesis and flowering time by the circadian clock in *Arabidopsis*. *Plant Physiol*, **130**, 1770-1775.
- Böhlenius, H., Huang, T., Charbonnel-Campaa, L., Brunner, A.M., Jansson, S., Strauss, S.H., Nilsson, O.** (2006) CO/FT Regulatory Module Controls Timing of Flowering and Seasonal Growth Cessation in Trees. *Science*, **412**, 1040-1043.
- Bolhuis, G.G.** (1966) Influence of length of the illumination period on root formation in cassava (*Manihot utilisima* Pohl). *Netherlands Journal of Agricultural Science*, **14**, 251-254.
- Bolhuis, G.G.**, (1953) A survey of some attempts to breed Cassava with a high content of proteins in the roots. *Euphytica*, **2**, 107-112.
-

- 
- Borner, R., Kampmann, G., Chandler, J., Gleissner, R., Wisman, E., Apel, K., Melzer, S.** (2000) A MADS domain gene involved in the transition to flowering in *Arabidopsis*. *Plant J*, **24**, 591-599.
- Boxall, S.F., Foster, J.M., Bohnert, H.J., Cushman, J.C., Nimmo, H.G., Hartwell, J.** (2005) Conservation and divergence of circadian clock operation in a stress-inducible crassulacean acid metabolism species reveals clock compensation against stress. *Plant Physiol.*, **137**, 969-982.
- Bruijn, G.H.** (1977) Influence of daylength on the flowering of cassava. *Tropical Root Tuber Crops Newsletter*, **10**, 1-3.
- Byrne, D.** (1984) Breeding cassava. In: Janick, J. (ed.) *Plant Breeding Reviews*, **2**, 73-134.
- Bünning, E.** (1960) Circadian rhythms and time measurement in photoperiodism. *Cold Spring Harbor Symp. Quant. Biol.* **25**:249-256.
- Bushway, A.A., Bushway, A.W., Belyea, P.R., Bushway, R.J.** (1980) The proximate composition and glycoalkaloid content of three potato meals. *American Journal of Potato Research*, **57**, 167-171.
- Caddick, M.X., Greenland, A.J., Jepson, I., Krause, K.P., Qu, N., Riddell, K.V., Salter, M.G., Schuch, W., Sonnewald, U., Tomsett, A.B.** (1998) An ethanol inducible gene switch for plants used to manipulate carbon metabolism. *Nat Biotechnol*, **16**, 177-180.
- Ceballos, H., Fregene, M., Roca, W., Tohme, J.** (2002) Cassava: An Ideal crop for biotechnology tools. In: Taylor, N.J., F. Ogbe and C.M. Fauquet (eds.) *Cassava, An Ancient Crop for Modern Times Food, Health, Culture. Proceedings of Fifth International Scientific Meeting of the Cassava Biotechnology Network* (Abstracts) Danforth Plant Science Center St Louis, Missouri, USA.
- Chia, T.Y., Muller, A., Jung, C., Mutasa-Gottgens, E.S.** (2008) Sugar beet contains a large CONSTANS-LIKE gene family including a CO homologue that is independent of the early-bolting (B) gene locus. *J Exp Bot*, **59**, 2735-2748.
- Chuck, G., Hake, S.** (2005) Regulation of developmental transitions, *Current Opinion in Plant Biology*, **8**, 67-70.
- CIAT** (1979) Annual Report of the Centro Internacional de Agricultura Tropical, Cali, *Colombia*.
- CIAT** (1982) Annual Report of the Centro Internacional de Agricultura Tropical, Cali, *Colombia*.
- CIAT** (2002) Annual Report Project SB2, Assessing and Utilizing Agrobiodiversity through Biotechnology, CIAT, Cali, Colombia, pp 203-207.
- CIAT** (2002) Annual Report Project SB2, Assessing and Utilizing Agrobiodiversity through Biotechnology, CIAT, Cali, Colombia, pp 212-214.
- Cock, J.H., Franklin, D., Sandoval, D., Juri, P.** (1979) The ideal cassava plant for maximum yield. *Crop Science*, **19**: 271-279.
- Cock, J.H., El-Sharkawy, M.A.** (1988) Physiological characteristics for cassava selection. *Experimental Agriculture*, **24**, 443-448.
- Cock, J.H., Rosas, S.** (1975) Ecophysiology of Cassava, In: (eds) *Symposium on Ecophysiology of Tropical Crops. Communications Division of CEPLAC*, Ilhéus, Bahia, Brazil., 1-14.
- Connor, D.J., Cock, J.H.** (1981) Response of Cassava to water shortage. II, Canopy dynamics. *Field Crops Res.*, **4**, 285-296.
- Connor, D.J., Cock, J.H., Parra, G.E.** (1981) Response of Cassava to water shortage. I. Growth and yield. *Field Crops Research*, **4**, 181-200.
- Cooke R.D., Coursey, D.G.** (1981) Cassava: a major cyanide-containing food crop In: Vennesland, B., Conn, E.E. and Knowles, C.J. (eds) *Cyanide in Biology*. Academic Press, New York.
-

- 
- Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., Giakountis, A., Farrona, S., Gissot, L., Turnbull, C., Coupland, G.** (2007) FT protein movement contributes to long-distance signaling in floral induction of *Arabidopsis*. *Science*, **316**, 1030-1033.
- Cortés, D.F., Reilly, K., Okogbenin, E., Beeching, J.R., Iglesias, C., Tohme, J.** (2002) Mapping wound-response genes involved in post-harvest physiological deterioration (PPD) of cassava (*Manihot esculenta* Crantz). *Euphytica*, **128**, 47–53.
- Cunha, H.M.P., Conceição, A. J. da** (1975) Indução ao florescimento da mandioca (*Manihot esculenta* Crantz) –Nota prévia. In: *Projecto Mandioca. Universidade Federal de Bahia. Escola de Agronomia*. Cruz das Almas. Convenio UFBA.
- Dahniya, M.T.** (1994) An overview of Cassava in Africa. *African Crop Science Journal*, **2**, 337-343.
- Davis, A.M., Hall, A., Millar, A.J., Darrah, C., Davis, S.J.** (2009) Protocol: Streamlined sub-protocols for floral-dip transformation and selection of transformants in *Arabidopsis thaliana*. *Plant Methods*, **5**, 3.
- De Marian, J.J.** (1729) Observation botanique-Histoire de l'Academie Royale de Sciences (Paris), p35.
- Dietrich, F.S., Voegeli, S., Brachat, S., Lerch, A., Gates, K., Steiner, S., Mohr, C., Pohlmann, R., Luedi, P., Choi, S.D., Wing, R.A., Flavier, A., Gaffney, T.D., Phillippsen, P.** (2004) The *Ashbya gossypii* genome as a tool for mapping the ancient *Saccharomyces cerevisiae* genome. *Science*, **304**, 304-307.
- Dowson-Day, M.J., Millar, A.J.** (1999) Circadian dysfunction causes aberrant hypocotyl elongation patterns in *Arabidopsis*. *Plant J.*, **17**, 63-71.
- Doyle, J.J., J.L. Doyle.** (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.*, **19**, 11-15.
- Doyle, M. R., Davis, S. J., Bastow, R. M., McWatters, H. G., Kozma-Bognár, L., Nagy, F., Millar, A. J., Amasino, R. M.** (2002) The *ELF4* gene controls circadian rhythms and flowering time in *Arabidopsis thaliana*. *Nature*, **419**: 74-77.
- Doyle, M.R** (2003) The cloning and characterization of *EARLY FLOWERING 4*: A gene involved in circadian regulation and control of flowering time in *Arabidopsis thaliana*, PhD thesis, *University of Wisconsin*, Madison, U.S.A.
- Edwards G.E., Sheta E., Moore B.D., Dai Z., Fransceschi V.R., Cheng S-H, Lin C-H, Ku M.S.B.** (1990) Photosynthetic characteristics of cassava (*Manihot esculenta*) a C3 species with chloronchymatous bundle sheath cell. *Plant Cell Physiol.*, **31**, 1199-1206.
- Edwards, K.D., Lynn, J.R., Gyula, P., Nagy, F., Millar, A.J.** (2005) Natural allelic variation in the temperature-compensation mechanisms of the *Arabidopsis thaliana* circadian clock. *Genetics*, **170**, 387-400.
- Egesi, C., Cuambe, C., Rosero, A., Sanchez, T., Morante, N., Ceballos, H., Fregene, M.** (2007) Genetic Mapping Of Delayed Post-Harvest Physiological Deterioration In Cassava In Backcross Derivatives Of *Manihot walkerae* . Abstract in PAGXVI San Diego, CA P673.
- El-Sharkawy, M.A., Cock, J.H., Cadena, G.D.** (1984a) Stomatal characteristics among cassava cultivars and their relation to gas exchange. *Experimental Agriculture*, **20**, 67-76.
- El-Sharkawy, M.A., Cock, J.H., Held, K.A.** (1984b). Water use sufficiency of cassava II. Differing sensitivity of stomata to air humidity in cassava and other warm-climate species. *Crop Sci*, **24**, 503-507.
- El-Sharkawy, M.A.** (2004) Cassava biology and physiology, *Plant Mol. Biol.*, **56**, 481-501.
- El-Sharkawy, M.A., Cock, J.H.** (1990) Photosynthesis of cassava (*Manihot esculenta*). *Experimental Agriculture*, **26**, 325-340.
- EMBRAPA** (2005) Cassava: the bread of Brazil = Yuca, el pan de Brasil. Brasilia, DF, 280.
-

- 
- FAO (2001) FAO production yearbook for 2000, Rome, Italy, FAO.
- FAO (2003) Cassava Production Statistics 2002. <http://www.fao.org>.
- Felenbok, B. (1991) The ethanol utilization regulon of *Aspergillus nidulans*, the Alca-Alcr system as a tool for the expression of recombinant proteins. *J Biotechnol*, **17**, 11–17.
- Felenbok, B., Sequeval, D., Mathieu, M., Sibley, S., Gwynne, D.I., Davies, R.W. (1988) The ethanol regulon in *Aspergillus nidulans*: characterization and sequence of the positive regulatory gene alcR. *Gene*, **73**, 385-396.
- Filichkin, S. A., Meilan, R., Busov, V. B., Ma, C., Brunner, A. M., Strauss, S. H. (2006) Alcohol-inducible gene expression in transgenic *Populus*. *Plant Cell Rep*, **25**, 660–667.
- Fowler, S., Lee, K., Onouchi, H., Samach, A., Richardson, K., Morris, B., Coupland, G., Putterill, J. (1999) *GIGANTEA*: a circadian clock-controlled gene that regulates photoperiodic flowering in *Arabidopsis* and encodes a protein with several possible membrane-spanning domains. *Embo J*, **18**, 4679-4688.
- Fregene, M., Angel, F., Gomez, R., Rodriguez, F., Chavarriaga, P., Roca, W., Tohme, J., Bonierbale, M. (1997) A molecular genetic map of cassava (*Manihot esculenta* Crantz). *Theor Appl Genet*, **95**, 431-441.
- Fregene, M., Okogbenin, E., Mba, C., Angel, F., Suarez, M.C., Janneth, G., Chavarriaga, P., Roca, W., Bonierbale, M., Tohme, J. (2001) Genome mapping in cassava improvement: Challenges, achievements and opportunities. *Euphytica*, **120**, 159-165.
- Fregene, M., Ponti-Kaerlas, J. (2002) Cassava Biotechnology. In CAB International Cassava: Biology Production and Utilization (eds. R.J. Hillocks, J.M. Thresh and A.C. Belloti).
- Garner, W.W., Allard, H.A. (1920) Effect of the relative length of day and night and other factors of the environment on growth and reproduction in plants. *J. Agric. Res.*, **18**, 553-606.
- Garner, W.W., Allard, H.A. (1923) Further studies in photoperiodism, the response of the plant to the relative length of day and night. *J. Agric. Res.*, **23**, 871-920.
- Garooosi, G.A., Salter, M.G., Caddick, M.X., Tomsett, A.B. (2005) Characterization of the ethanol- inducible alc gene expression system in tomato. *J Exp Bot*, **56**, 1635-1642.
- Ghosh, S.P., Ramanujam, T., Jos, S., Moorthy, S.N., Nair, R.G. (1988) Tuber Crops. Oxford & IBH Publishing Co., *New Dehli*, 3-146.
- Giakountis, A., Coupland, G. (2008) Phloem transport of flowering signals. *Curr Opin Plant Biol*, **11**, 687-694.
- Gomez, R., Angel, F., Bonierbale, M.W., Rodriguez, F., Tohme, J., Roca, W.M. (1996) Inheritance of random amplified polymorphic DNA markers in cassava (*Manihot esculenta* Crantz). *Genome*, **39**, 1039-1043.
- González, A.E., Schöpke, C., Taylor, N.J., Beachy, R.N., Fauquet, C.M. (1998) Regeneration of transgenic Cassava plants (*Manihot esculenta* Crantz) through *Agrobacterium*-mediated transformation of embryogenic suspension cultures. *Plant Cell Rep.*, **17**, 827-831.
- Goodner, B., Hinkle, G., Gattung, S., Miller, N., Blanchard, M., Quorollo, B., Goldman, B.S., Cao, Y.W., Askenazi, M., Halling, C., Mullin, L., Houmiel, K., Gordon, J., Vaudin, M., Iartchouk, O., Epp, A., Liu, F., Wollam, C., Allinger, M., Doughty, D., Scott, C., Lappas, C., Markelz, B., Flanagan, C., Crowell, C., Gurson, J., Lomo, C., Sear, C., Strub, G., Cielo, C. and Slater, S. (2001) Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. *Science*, **294**, 2323-2328.
- Gould, P.D., Locke, J.C., Larue, C., Southern, M.M., Davis, S.J., Hanano, S., Moyle, R., Milich, R., Putterill, J., Millar, A.J., Hall, A. (2006) The molecular basis of temperature compensation in the *Arabidopsis* circadian clock. *Plant Cell*, **18**, 1177-1187.
- Gresshoff, P.M., Doy, H.M (1972) Haploid *Arabidopsis thaliana* callus and plants from anther culture. *Aust. J. Biol. Sci.*, **25**, 259-264.
-

- 
- Griffiths, S., Dunford, R.P., Coupland, G., Laurie, D. A.** (2003) The Evolution of *CONSTANS*-Like Gene Families in Barley, Rice, and Arabidopsis. *Plant Physiology*, **131**, 1855–1867.
- Guthrie, E.J.** (1987) African cassava mosaic virus disease and its control. *Proceedings of the international Seminar of African Cassava Mosaic Virus Diseases and its control*. Yamoussoukro, Cote D'Ivoire., 1-9.
- Hahn, S.K., Keyser, J.** (1985) Cassava: a basic food of Africa. *Outlook on Agriculture*, **14**, 95-99.
- Halsey, M.E., Olsen, K.M., Taylor, N.J., Chavarriaga-Aguirre, P.** (2008) Reproductive Biology of Cassava (*Manihot esculenta* Crantz) and Isolation of Experimental Field Trials. *Crop Sci*, **48**, 49-58.
- Hanano, S., Stracke, R., Jakoby, M., Merkle, T., Domagalska, M. A., Weisshaar, B., Davis, S. J.** (2008) A systematic survey in Arabidopsis thaliana of transcription factors that modulate circadian parameters. *BMC Genomics*, **9**:182.
- Hayama R., Coupland, G.** (2004) The molecular basis of diversity in the photoperiodic flowering responses of Arabidopsis and rice. *Plant Physiology*, **135**, 677–684.
- Hepworth, S.R., Valverde, F., Ravenscroft, D., Mouradov, A., Coupland, G.** (2002) Antagonistic regulation of flowering-time gene *SOC1* by *CONSTANS* and *FLC* via separate promoter motifs. *Embo J*, **21**:4327-4337.
- Hershey, C.H., Jennings, D.L.** (1992) Progress in breeding cassava for adaptation to stress. *Plant Breeding Abstracts*, **62**, 823-831.
- Hillocks, R.J.** (2002) Cassava in Africa. In: Hillocks, R.J., J.M. Thresh and A.C. Bellotti (eds.) *Cassava: Biology, Production and Utilization*. CABI Publishing Oxon, UK and New York, USA., 40-54.
- Holefors, A., Opseth, L., Ree Rosnes, A.K., Ripel, L., Snipen, L., Fossdal, C.G., Olsen, J.E.** (2009) Identification of PaCOL1 and PaCOL2, two *CONSTANS*-like genes showing decreased transcript levels preceding short day induced growth cessation in Norway spruce. *Plant Physiol Biochem*, **47**, 105-115.
- Hong, Y. G., Robinson, D. J., Harrison, B. D.** (1993) Nucleotide sequence evidence for the occurrence of three distinct whitefly transmitted geminiviruses in cassava. *Journal of General Virology*, **74**, 2437- 2443.
- Howeler, R.H.** (2002) Cassava Mineral Nutrition and Fertilization. In: Hillocks, R.J., J.M. Thresh and A.C. Bellotti (eds.) *Cassava: Biology, Production and Utilization*. CABI Publishing Oxon, UK and New York, USA., 115-130.
- Iglesias, C., Hershey, C., Calle, F., Bolanos, A.** (1994) Propagating Cassava (*Manihot esculenta*) by sexual seed. *Expl.Agric*, **30**, 283-290.
- Ihemere, U., Arias-Garzon, D., Lawrence, S., Sayre, R.** (2006) Genetic modification of cassava for enhanced starch production. *Plant Biotechnol J*, **4**, 453-465.
- Irikura, Y., Cock, J.H., Murria, B.G.** (1979) The physiological basis of genotype-temperature interactions in cassava. *Field Crops Research*, **2**, 227-239.
- Izawa, T., Oikawa, T., Sugiyama, N., Tanisaka, T., Yano, M., Shimamoto, K.** (2002) Phytochrome mediates the external light signal to repress FT orthologs in photoperiodic flowering of rice. *Genes & Development*, **16**, 2006-2020.
- Jaeger, K.E., Graf, A., Wigge, P.A.** (2006) The control of flowering in time and space. *J Exp Bot*, **57**, 3415-3418.
- Jaeger, K.E., Wigge, P.A.** (2007) FT protein acts as a long-range signal in Arabidopsis. *Curr Biol*, **17**, 1050-1054.
- Jennings, D.L.** (1976) Cassava, (*Manihot esculenta* Euphorbiaceae) In: *Simmonds N (ed) Evolution of Crop Plants*. Longman, London, 81-84.
-

- 
- Jennings, D.L., Iglesias, C.** (2002) Breeding for crop improvement. In: Hillocks, R. J., M.J. Thresh and A.C. Bellotti (eds.) *Cassava: Biology, production and utilisation*. CABI International, Oxford, 149-166.
- Jeong, D. H., Sung, S. K., An, G.** (1999) Molecular cloning and characterization of CONSTANS-like cDNA clones of the Fuji apple. *J. Plant Biol.* **42**, 23-31.
- Johanson, U., West, J., Lister, C., Michaels, S., Amasino, R. and Dean, C.** (2000) Molecular analysis of FRIGIDA, a major determinant of natural variation in Arabidopsis flowering time. *Science*, **290**, 344-347.
- Jos, J.S., Nair, S. G.** (1979) Pachytene pairing in relation to pollen fertility in five cultivars of Cassava. *Cytologia*, **44**, 813-820.
- Kardailsky, I., Shukla, V.K., Ahn, J.H., Dagenais, N., Christensen, S.K., Nguyen, J.T., Chory, J., Harrison, M.J., Weigel, D.** (1999) Activation tagging of the floral inducer FT. *Science*, **286**, 1962-1965.
- Kaul, S., Koo, H.L., Jenkins, J., Rizzo, M., Rooney, T., Tallon, L.J., Feldblyum, T., Nierman, W., Benito, M.I., Lin, X.Y., Town, C.D., Venter, J.C., Fraser, C.M., Tabata, S., Nakamura, Y., Kaneko, T., Sato, S., Asamizu, E., Kato, T., Kotani, H., Sasamoto, S., Ecker, J.R., Theologis, A., Federspiel, N.A., Palm, C.J., Osborne, B.I., Shinn, P., Conway, A.B., Vysotskaia, V.S., Dewar, K., Conn, L., Lenz, C.A., Kim, C.J., Hansen, N.F., Liu, S.X., Buehler, E., Altafi, H., Sakano, H., Dunn, P., Lam, B., Pham, P.K., Chao, Q., Nguyen, M., Yu, G.X., Chen, H.M., Southwick, A., Lee, J.M., Miranda, M., Toriumi, M.J., Davis, R.W., Wambutt, R., Murphy, G., Dusterhoft, A., Stiekema, W., Pohl, T., Entian, K.D., Terryn, N., Volckaert, G., Salanoubat, M., Choisine, N., Rieger, M., Ansorge, W., Unseld, M., Fartmann, B., Valle, G., Artiguenave, F., Weissenbach, J., Quetier, F., Wilson, R.K., de la Bastide, M., Sekhon, M., Huang, E., Spiegel, L., Gnoj, L., Pepin, K., Murray, J., Johnson, D., Habermann, K., Dedhia, N., Parnell, L., Preston, R., Hillier, L., Chen, E., Marra, M., Martienssen, R., McCombie, W.R., Mayer, K., White, O., Bevan, M., Lemcke, K., Creasy, T.H., Bielke, C., Haas, B., Haase, D., Maiti, R., Rudd, S., Peterson, J., Schoof, H., Frishman, D., Morgenstern, B., Zaccaria, P., Ermolaeva, M., Pertea, M., Quackenbush, J., Volfovsky, N., Wu, D.Y., Lowe, T.M., Salzberg, S.L., Mewes, H.W., Rounsley, S., Bush, D., Subramaniam, S., Levin, I., Norris, S., Schmidt, R., Acarkan, A., Bancroft, I., Quetier, F., Brennicke, A., Eisen, J.A., Bureau, T., Legault, B.A., Le, Q.H., Agrawal, N., Yu, Z., Martienssen, R., Copenhagen, G.P., Luo, S., Pikaard, C.S., Preuss, D., Paulsen, I.T., Sussman, M., Britt, A.B., Selinger, D.A., Pandey, R., Mount, D.W., Chandler, V.L., Jorgensen, R.A., Pikaard, C., Juergens, G., Meyerowitz, E.M., Theologis, A., Dangl, J., Jones, J.D.G., Chen, M., Chory, J., Somerville, M.C.** (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, **408**, 796-815.
- Keating, B.A., Evenson, J.P., Fukai, S.** (1982a) Environmental effects on growth and development of Cassava (*Manihot esculenta* Crantz) I. Crop development. *Field Crops Res.*, **5**, 271-281.
- Keating, B.A., Evenson, J.P., Fukai, S.** (1982b) Environmental effects on growth and development of cassava (*Manihot esculenta* Crantz) II. Crop growth rate and biomass yield. *Field Crop Research*, **5**, 283-292.
- Keating, B.A., Wilson, G. L., Evenson, J.P.** (1985) Effect of Photoperiod on Growth and Development of Cassava (*Manihot esculenta* Crantz). *Aust. J. Plant Physiol.*, **12**, 621-630.
- Khanna, R., Kikis, E.A., Quail, P.H.** (2003) *EARLY FLOWERING 4* functions in phytochrome B-regulated seedling de-etiolation. *Plant Physiol.*, **133**, 1530-1538.
- Kobayashi, Y., Kaya, H., Goto, K., Iwabuchi, M., Araki, T.** (1999) A pair of related genes with antagonistic roles in mediating flowering signals. *Science*, **286**, 1960-1962.
- Kojima, S., Takahashi, Y., Kobayashi, Y., Monna, L., Sasaki, T., Araki, T., Yano, M.** (2002) Hd3a, a rice ortholog of the *Arabidopsis* FT gene, promotes transition to flowering downstream of Hd1 under short-day conditions. *Plant Cell Physiol.*, **43**, 1096-1105.
-



- 
- Kolmos, E.** (2007) Genetic analysis of two evening genes in *Arabidopsis thaliana* circadian clock, *Ph.D thesis, MPIZ, Cologne, Germany*.
- Konan, N.K., Sangwan, R.S., Sangwan-Norreel, B.S.** (1994b) Somatic embryogenesis from cultured mature cotyledons of cassava (*Manihot esculenta* Crantz). *Plant Cell Tiss. Org. Cult.* **37**, 91-102.
- Koncz, C., Schell, J.** (1986) The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimeric genes carried by a novel type of *Agrobacterium* binary vector. *Mol. Gen. Genet.*, **204**, 383-396.
- Koornneef, M., Blankestijn-de Vries, H., Hanhart, C., Soppe, W., Peeters, T.** (1994) The phenotype of some late-flowering mutants is enhanced by a locus on chromosome 5 that is not effective in the Landsberg *erecta* wild-type. *Plant J*, **6**, 911-919.
- Kulmburg P., Judewicz, N., Mathieu, M., Lenouvel, F., Sequeval, D., Felenbok, B.** (1992) Specific binding sites for the activator protein, ALCR, in the alcA promoter of the ethanol regulon of *Aspergillus nidulans*. *J Biol Chem.*, **267**, 21146-53.
- Lagercrantz, U., Axelsson, T.** (2000) Rapid Evolution of the Family of CONSTANS LIKE Genes in Plants *Mol. Biol. Evol.*, **17**, 1499-1507.
- Laurie, DA, Griffiths, S., Dunford, R.P., Christodoulou, V., Taylor, S.A., Cockram, J., Beales, J., Turner, A.** (2004). Comparative genetic approaches to the study of flowering time in temperate cereals. *Field Crops Research*, **90**: 87-99.
- Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm A., Lopez R., Thompson, J.D., Gibson, T.J., Higgins, D.G.** (2007) ClustalW and ClustalX version 2. *Bioinformatics*, **23**, 2947-2948.
- Lee, H., Suh, S.S., Park, E., Cho, E., Ahn, J.H., Kim, S.G., Lee, J.S., Kwon, Y.M., Lee, I.** (2000) The AGAMOUS-LIKE 20 MADS domain protein integrates floral inductive pathways in *Arabidopsis*. *Genes Dev*, **14**, 2366-2376.
- Legg, J.P., Thresh, J.M.** (2003) Cassava virus diseases in Africa. In: *Plant Virology in sub-Saharan Africa*, 517-550. <http://www.iita.oorg/info/virology.htm>.
- Levy, Y.Y., Mesnage, S., Mylne, J.S., Gendall, A.R., Dean, C.** (2002) Multiple roles of *Arabidopsis* VRN1 in vernalization and flowering time control. *Science*, **297**, 243-246.
- Lewis, R.S., Kernodle, S.P.** (2009) A method for accelerated trait conversion in plant breeding. *Theor Appl Genet.*, **118**, 1499-1508.
- Li, H.Q., Sautter, C., Potrykus I., Puonti-Kaerlas, J.** (1996) Genetic transformation of Cassava (*Manihot esculenta* Crantz). *Nature Biotechnol.*, **14**, 736-740.
- Lin, M. K., Belanger, H., Lee, Y.J., Varkonyi-Gasic, E., Taoka, K.I., Miura, E., Xoconostle-Cázares, B., Gendler, K., Jorgensen, R.A., Phinney, B., Lough, T.J., Lucas, W.J.** (2007) FLOWERING LOCUS T protein may act as the long-distance florigenic signal in the cucurbits. *Plant Cell*, **19**, 1488-1506.
- Lifschitz, E., Eviatar, T., Rozman, A., Shalit, A., Goldshmidt, A., Amsellem, Z., Alvarez, J.P., Eshed, Y.** (2006) The tomato FT ortholog triggers systemic signals that regulate growth and flowering and substitute for diverse environmental stimuli. *Proc Natl Acad Sci U.S.A.*, **103**, 6398-6403.
- Liu, J., Yu, J., McIntosh, L., Kende, H., Zeevaart, J.A.** (2001) Isolation of a *CONSTANS* ortholog from *Pharbitis nil* and its role in flowering. *Plant Physiology*, **125**, 1821-1830.
- Lockington, R., Scazzocchio, C., Sequeval, D., Mathieu, M., Felenbok, B.** (1987) Regulation of alcR, the positive regulatory gene of the ethanol utilization regulon of *Aspergillus nidulans*. *Mol Microbiol.*, **1**, 275-281.
-

- 
- Mahon, J.D., Lowe, S.B., Hunt, L.A., Thiagarajah, M.** (1977) Environmental effects on photosynthesis and transpiration attached leaves of cassava (*Manihot esculenta* Crantz). *Photosynthetica*, **11**, 121–130.
- Martinez-Zapater, J.M., Coupland, C., Dean, C., Koornneef, M.** (1994) The transition to flowering in *Arabidopsis*. In *Arabidopsis*. Edited by Somerville, C.R. and Meyerowitz, E.M. pp. 403-434. Cold Spring Harbor Laboratory Press, New York.
- Mathews, H., Schopke, C., Carcamo, R., Chavarriaga, P., Fauquet, C., Beachy, R.N.** (1993) Improvement of Somatic Embryogenesis and Plant-Recovery in Cassava. *Plant Cell Reports*, **12**, 328-333.
- Mba, R.E.C., Stephenson, P., Edwards, K., Melzer, S., Nkumbira, J., Gullberg, U., Apel, K., Gale, M., Tohme, J., Fregene, M.** (2001) Simple sequence repeat (SSR) markers survey of the cassava (*Manihot esculenta* Crantz) genome: towards an SSR-based molecular genetic map of cassava. *Theoretical and Applied Genetics*, **102**, 21-31.
- McMahon, J.M., White, W.L.B., Sayre, R.T.** (1995) Cyanogenesis in cassava (*Manihot esculenta* Crantz). *Journal of Experimental Botany*, **46**, 731-714.
- McWatters, H. G., Kolmos, E., Hall, A., Doyle, M. R., Amasino, R. A., Gyula, P., Nagy, F., Millar, A. J., and Davis, S. J.** (2007) ELF4 is required for oscillatory properties of the circadian clock. *Plant Phys.*, **144**, 391-401.
- Michaels, S.D., Amasino, R.M.** (1999) FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell*, **11**, 949-956.
- Michaels S. D., Amasino R. M.** (2000) Memories of winter: vernalization and the competence to flower. *Plant, cell and environment*, **23**, 1145-1153.
- Michaels S. D., Amasino R. M.** (2001) Loss of FLOWERING LOCUS C activity eliminates the late-flowering phenotype of FRIGIDA and autonomous pathway mutations but not responsiveness to vernalization. *Plant Cell*, **13**, 935-941.
- Mizoguchi, T., Wright, L., Fujiwara, S., Cremer, F., Lee, K., Onouchi, H., Mouradov, A., Fowler, S., Kamada, H., Putterill, J., Coupland, G.** (2005) Distinct roles of GIGANTEA in promoting flowering and regulating circadian rhythms in *Arabidopsis*. *Plant Cell*, **17**, 2255-2270.
- Mouradov, A., Cremer, F., Coupland, G.** (2002) Control of flowering time: interacting pathways as a basis for diversity. *Plant Cell*, **14 Suppl**, S111-130.
- Murashige, T., Skoog, F.** (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol. Plant.* **15**: 473-497.
- Ng, N.Q., Ng, S.Y.C.** (2002) Genetic Resources Conservation, In: Hillocks, R. J., M.J. Thresh and A.C. Bellotti (eds.) Cassava: Biology, production and utilisation. CABI International, Oxford, 167.
- Nweke, F.I., Spenser, D.S.C., Lynam, J.K.,** (2002) The cassava transformation: Africa's best kept secret. Michigan State University Press, Lansing, Michigan, USA, 273.
- Oliverio, K.A., Crepy, M., Martin-Tryon, E.L., Milich, R., Harmer, S.L., Putterill, J., Yanovsky, M.J., Casal, J.J.** (2007) GIGANTEA Regulates Phytochrome A-Mediated Photomorphogenesis Independently of Its Role in the Circadian Clock. *Plant Physiology*, **144**, 495-502.
- Olsen, K.M., Schaal, B.A.** (1999) Evidence on the origin of cassava: phylogeography of *Manihot esculenta*. *Proc Natl Acad Sci U.S.A.*, **96**, 5586-5591.
- Onouchi, H., Igeno, M.I., Perilleux, C., Graves, K., Coupland, G.** (2000) Mutagenesis of plants overexpressing CONSTANS demonstrates novel interactions among *Arabidopsis* flowering-time genes. *Plant Cell*, **12**, 885-900.
- Pardales, J.R., Esquibel, C.B.** (1996) Effect of Drought during the Establishment Period on the Root System Development of Cassava. *Jpn. J. Crop Sci.*, **65**, 93-97.
-

- 
- Peña, L., Martín-Trillo, M., Juárez, J., Pina, J.A., Navarro, L., Martínez-Zapater J.M.** (2001) Constitutive expression of *Arabidopsis* *LEAFY* or *APETALA1* genes in citrus reduces their generation time. *Nature Biotechnology*, **19**, 263-267.
- Poethig, R.S.** (2003) Phase change and the regulation of developmental timing in plants. *Science*, **301**, 334-336.
- Raemakers, C.J.J.M., Amati, M., Staritsky, G., Jacobsen, E., Visser, R.G.F.** (1993) Cyclic somatic embryogenesis in cassava. *Ann. Bot.*, **71**, 289-294.
- Raemakers, C.J.J.M., Bessembinder, J., Staritsky, G., Jacobsen, E., Visser, R.G.F.** (1993) Induction, germination and shoot development of somatic embryos in cassava. *Plant Cell Tiss. Org. Cult.*, **33**, 151-156.
- Raemakers, C.J.J.M., Sofiari, E., Taylor, N., Henshaw, G., Jacobsen, E., Visser, R.G.F.** (1996) Production of transgenic cassava (*Manihot esculenta* Crantz) plants by particle bombardment using luciferase activity as selection marker. *Mol. Breeding*, **2**, 339-349.
- Ratcliffe, O.J., Nadzan, G.C., Reuber, T.L., Riechmann, J.L.** (2001) Regulation of flowering in *Arabidopsis* by an FLC homologue. *Plant Physiol.*, **126**, 122-132.
- Robert, L.S., Robson, F., Sharpe, A., Lydiate, D., Coupland, G.** (1998) Conserved structure and function of the *Arabidopsis* flowering time gene *CONSTANS* in *Brassica napus*. *Plant Mol. Biol.*, **37**:763-772.
- Roca, W.M.** (1984) Cassava. In: Sharp W, Evans D, Ammirato P and Yamada Y (eds.), *Handbook of plant cell culture*. MacMillan, New York, 269-301.
- Rogers, D.J. and Appan, S.G.** (1973). *Manihot, Manihotoides. Flora Neotrópica*. Monograph No. 13. Hafner Press, New York, NY, USA, 272.
- Roslan, H.A., Salter, M.G., Wood, C.D., White, M.R., Croft, K.P., Robson, F., Coupland, G., Doonan, J., Laufs, P., Tomsett, A.B., Caddick, M.X.** (2001) Characterization of the ethanol-inducible alc gene-expression system in *Arabidopsis thaliana*. *Plant J*, **28**, 225-235.
- Roux, F., Touzet, P., Cuguen, J., Le Corre, V.** (2006) How to be early flowering: an evolutionary perspective. *Trends Plant Sci*, **11**, 375-381.
- Rozen, S., Skaletsky, H.** (2000) Primer3 on the www for general users and for biologist programmers. *Methods Mol. Biol.*, **132**, 365-386.
- Saitou, N., Nei, M.** (1987) The Neighbor-joining method-a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, **4**, 406-425.
- Salathia, N., Davis, S.J. Lynn, J.R., Michaels, S.D., Amasino, R.M., Millar, A.J.** (2006) *FLOWERING LOCUS C*-dependent and -independent regulation of the circadian clock by the autonomous and vernalization pathways. *BMC Plant Biol.*, **6**, 10.
- Sambrook, J., Russell, D.W.** (2001) *Molecular Cloning: A Laboratory Manual* (3rd ed.). Cold Spring Harbor Laboratory Press.
- Salter, M.G., Schuch, W., Sonnewald, U., Tomsett, A.B.** (1998) An ethanol inducible gene switch for plants used to manipulate carbon metabolism. *Nat Biotechnol*, **16**, 177-180.
- Samach, A., Onouchi, H., Gold, S.E., Ditta, G.S., Schwarz-Sommer, Z., Yanofsky, M.F., Coupland, G.** (2000) Distinct roles of *CONSTANS* target genes in reproductive development of *Arabidopsis*. *Science*, **288**, 1613-1616.
- Sambrook, J., Fritsch, E.F., Maniatis, T.** (1989) *Molecular cloning. A laboratory manual*. Coldspring Harbor Laboratory Press, 1.53-1.105.
- Sarría, R., Torres, E., Angel, F., Chavarriaga P, Roca W.M.** (2000) Transgenic plants of cassava (*Manihot esculenta*) with resistance to Basta obtained by *Agrobacterium*-mediated transformation. *Plant Cell Rep.*, **19**, 339-344.
-

- 
- Schöpke, C., Taylor, N., Carcamo, R., Konan, N.K., Marmey, P., Henshaw, G.G, Beachy, R.N, Fauquet, C.** (1996) Regeneration of transgenic cassava plants (*Manihot esculenta* Crantz) from microbombarded embryogenic suspension cultures. *Nature Biotechnol.*, **14**, 731-735.
- Searle, I., Coupland, G.** (2004) Induction of flowering by seasonal changes in photoperiod. *Embo J*, **23**, 1217-1222.
- Serrano, G., Herrera-Palau, R., Romero, J.M., Serrano, A., Coupland, G., Valverde, F.** (2009) Chlamydomonas CONSTANS and the evolution of plant photoperiodic signaling. *Curr Biol*, **19**, 359-368.
- Sheikholeslam, S.N., Weeks, D.P.** (1987) Acetosyringone promotes high-efficiency transformation of *Arabidopsis thaliana* explants by *Agrobacterium tumefaciens*. *Plant Mol Biol*, **8**, 291-298.
- Silvestre, P.** (1989) Cassava. The Tropical Agriculturist, CTA, **82**.
- Simpson, G.G., Dean, C.** (2002) Arabidopsis, the Rosetta stone of flowering time? *Science*, **296**, 285-289.
- Siritunga, D., Arias-Garzon, D., White, W., Sayre, R.T.** (2004) Over-expression of hydroxynitrile lyase in transgenic cassava roots accelerates cyanogenesis and food detoxification. *Plant Biotechnology Journal.*, **2**, 37-43.
- Siritunga, D., Sayre, R.T.** (2003) Generation of cyanogens-free transgenic cassava. *Planta*, **217**, 367-373.
- Song, J., Yamamoto, K., Shomura, A., Itadani, H., Zhong, H.S., Yano, M., Sasaki, T.** (1998) Isolation and mapping of a family of putative zinc-finger protein cDNAs from rice. *DNA Res*, **5**, 95-101.
- Southern, E.M.** (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol.*, **98**, 503-517.
- Stamp, J.A.** (1987) Somatic embryogenesis in cassava: the anatomy and morphology of the regeneration process. *Annn. Bot.*, **59**, 451-459.
- Stamp, J.A., Henshaw, G.G.** (1982) Somatic embryogenesis in cassava. *Zeitschrift fur Pflanzenphysiologie*, **105**, 183-187.
- Stamp, J.A., Henshaw, G.G.** (1986) Adventitious regeneration in cassava In: Withers, L.A. and Alders, P.G. (eds) Plant tissue culture and its agricultural applications. Butterworths, London, 149-157.
- Stamp, J.A., Henshaw, G.G.** (1987a) Secondary somatic embryogenesis and plant regeneration in cassava. *Plant Cell Tissue and Organ Culture*, **10**, 227-233.
- Stamp, J.A., Henshaw, G.G.** (1987b) Somatic embryogenesis from clonal leaf tissues of cassava. *Annn. Bot.*, **59**, 445-450.
- Suarez-Lopez, P., Wheatley, K., Robson, F., Onouchi, H., Valverde, F., Coupland, G.** (2001) CONSTANS mediates between the circadian clock and the control of flowering in *Arabidopsis*. *Nature*, **410**, 1116-1120.
- Swanson, M.M., Harrison, B. D.** (1994) Properties, relationships and distribution of cassava mosaic geminiviruses. *Trop Sci*, **34**, 15-25.
- Sweetman J.P., Chu, C., Qu, N., Greenland, A.J., Sonnewald, U., Jepson, I.** (2002) Ethanol vapor is an efficient inducer of the alc gene expression system in model and crop plant species. *Plant Physiol*, **129**, 943-948.
- Szabados, L., Hoyos, R., Roca, W.** (1987) *In vitro* somatic embryogenesis and plant regeneration of cassava. *Plant Cell Rep*, **6**, 248-251.
- Tamaki, S., Matsuo, S., Wong, H.L., Yokoi, S., Shimamoto, K.** (2007) Hd3a protein is a mobile flowering signal in rice. *Science*, **316**, 1033-1036.
-

- 
- Taylor, N.J., Henshaw, G.G.** (1993). The induction of somatic embryogenesis in fifteen African and one South American cassava cultivars. *In: WM Roca and AM Thro (eds). Proceedings First International Scientific Meeting of the Cassava Biotechnology Network Cartagena Colombia* pp 134-140.
- Taylor, N.J., Edwards, M., Kiernan, R.J., Davey, C.D., Blakesley, D., Henshaw, G.G.** (1996) Development of friable embryogenic callus and embryogenic suspension culture systems in cassava (*Manihot esculenta* Crantz). *Nat Biotechnol*, **14**, 726-730.
- Taylor, N. T., Masona, M. V., Carcamo, R., Ho, T., Schöpke, C., Fauquet, C. M.** (2001). Production of embryogenic tissue and regeneration of transgenic plants in cassava (*Manihot esculenta* Crantz). *Euphytica*, **120**, 25-34.
- Thomas, J. E., Massalski, P. R., Harrison, B. D.** (1986) Production of monoclonal antibodies to African cassava mosaic virus and differences in their reactivities with other whitefly-transmitted geminiviruses. *Journal of General virology*, **67**, 2739–2748.
- Turck, F., Fornara, F., Coupland, G.** (2008) Regulation and identity of florigen: FLOWERING LOCUS T moves center stage. *Annual Review of Plant Biology*, **59**, 573-594.
- Turner, A., Beales, J., Faure, S., Dunford, R.P., Laurie, D.A.** (2005) The pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in barley. *Science*, **310**, 1031-1034.
- Ueno, O., Agarie, S.** (1997) The intercellular distribution of glycine decarboxylase in leaves of cassava in relation to the photosynthetic mode and leaf anatomy. *Japanese Journal of crop Science*, **66**, 268-278.
- Venter, J.C.** (2001) The sequence of the human genome, *Science*, **292**, 1838-1838.
- Wen-Jun, S., Forde, B.G.** (1989) Efficient transformation of *Agrobacterium* spp. by high voltage electroporation. *Nucleic Acids Res*, **17**, 8385.
- Weigel, D., Nilsson, O.** (1995) A Developmental Switch Sufficient for Flower Initiation in Diverse Plants. *Nature*, **377**, 495-500.
- Wheatley, C.C., Chuzel, G.** (1993) Cassava: The nature of the tuber and use as a raw material. *In: Macrae, R., R.K. Robinson and M.J. Sadler (eds.) Encyclopaedia of Food Science, Food Technology, and Nutrition.* Academic Press, San Diego, California, 734-743.
- White, W., Arias-Garzon, D., McMahon, J.M., Sayre, R.** (1998) Cyanogenesis in cassava: the role of hydroxynitrile lyase in root cyanide production. *Plant Physiology*, **116**, 1219-1225.
- Wigge, P.A., Kim, M.C., Jaeger, K.E., Busch, W., Schmid, M., Lohmann, J.U., Weigel, D.** (2005) Integration of spatial and temporal information during floral induction in *Arabidopsis*. *Science*, **309**, 1056-1059.
- Yano, M., Katayose, Y., Ashikari, M., Yamanouchi, U., Monna, L., Fuse, T., Baba, T., Yamamoto, K., Umehara, Y., Nagamura, Y., Sasaki, T.** (2000) Hd1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. *Plant Cell*, **12**, 2473-2484.
- Yanovsky, M.J., Kay, S.A.** (2003) Living by the calendar: how plants know when to flower. *Nat Rev Mol Cell Biol.*, **4**, 265-275.
- Yoo, S.K., Chung, K.S., Kim, J., Lee, J.H., Hong, S.M., Yoo, S.J., Yoo, S.Y., Lee, J.S., Ahn, J.H.** (2005) *CONSTANS* activates *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* through *FLOWERING LOCUS T* to promote flowering in *Arabidopsis*. *Plant Physiology*, **139**, 770–778.
- Young, M.W., Kay, S.A.** (2001) Time zones: a comparative genetics of circadian clocks. *Nat Rev Genet*, **2**, 702-715.
- Yu, J., Hu, S., Wang, J., Wong, G.K., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., Zhang, X., Cao, M., Liu, J., Sun, J., Tang, J., Chen, Y., Huang, X., Lin, W., Ye, C., Tong, W., Cong, L.,**
-

---

Geng, J., Han, Y., Li, L., Li, W., Hu, G., Huang, X., Li, W., Li, J., Liu, Z., Li, L., Liu, J., Qi, Q., Liu, J., Li, L., Li, T., Wang, X., Lu, H., Wu, T., Zhu, M., Ni, P., Han, H., Dong, W., Ren, X., Feng, X., Cui, P., Li, X., Wang, H., Xu, X., Zhai, W., Xu, Z., Zhang, J., He, S., Zhang, J., Xu, J., Zhang, K., Zheng, X., Dong, J., Zeng, W., Tao, L., Ye, J., Tan, J., Ren, X., Chen, X., He, J., Liu, D., Tian, W., Tian, C., Xia, H., Bao, Q., Li, G., Gao, H., Cao, T., Wang, J., Zhao, W., Li, P., Chen, W., Wang, X., Zhang, Y., Hu, J., Wang, J., Liu, S., Yang, J., Zhang, G., Xiong, Y., Li, Z., Mao, L., Zhou, C., Zhu, Z., Chen, R., Hao, B., Zheng, W., Chen, S., Guo, W., Li, G., Liu, S., Tao, M., Wang, J., Zhu, L., Yuan, L. and Yang, H. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. indica). *Science*, **296**, 79-92.

Yu, H., Xu, Y., Tan, E.L., Prakash, P. (2002) *AGAMOUS-LIKE 24*, a dosage-dependent mediator of the flowering signals, *Proc Natl Acad Sci U.S.A.*, **99**, 16336–16341.

Zagotta, M.T., Shannon, S., Jacobs, C., Meekswagner, D.R. (1992) Early-flowering mutants of *Arabidopsis thaliana*. *Austr. J. Plant Physiol.*, **19**, 411-418.

Zhang, P., Puonti-Kaerlas, J. (2000) PIG-mediated cassava transformation using positive and negative selection. *Plant Cell Rep.*, **19**, 1041-1048.

Zhang, P., Potrykus, I., Puonti-Kaerlas, J. (2000a) Efficient production of transgenic cassava using negative and positive selection. *Transgenic Res.*, **9**, 405-415.

Zhang P, Legris G, Coulin, P., Puonti-Kaerlas, J. (2000b) Production of stably transformed cassava plants via particle bombardment. *Plant Cell Rep.*, **19**, 939-945.

Zhang, P., Bohl-Zenger, S., Puonti-Kaerlas, J, Potrykus, I., Gruijssem, W. (2003) Two cassava promoters related to vascular expression and storage root formation. *Planta*, **218**, 192–203.

Zobell, O., Coupland, G., Reiss, B. (2005) The family of CONSTANS-like genes in *Physcomitrella patens*. *Plant Biol.*, **7**, 266-275.

Zuo, J., Chua, N.H. (2000) Chemical-inducible systems for regulated gene expression of plant genes. *Current opinion in biotechnology*, **11**, 146-151.

---

## APPENDICES

<b>APPENDIX I</b> GI SEQUENCES.....	B
<b>APPENDIX II</b> CO SEQUENCES.....	F
<b>APPENDIX III</b> ELF4 SEQUENCE.....	J
<b>ACKNOWLEDGEMENTS</b> .....	K
<b>ERKLÄRUNG</b> .....	M
<b>TAGUNGEN</b> .....	N
<b>LEBENS LAUF</b> .....	O

---

## Appendix I GI sequences

Consensus sequences from generated from *MeGI* contig and from other species as described in Chapters 2 and 3

### *MeGI* sequence

```
NTCTATNTTNNNNCTTGCCAATCGATGACATCTACCTTGAATAACTGATAAGAGGGTCCGC
GAATATCATTTATGTCTGCTTGGTTAGCTGTAGGTTTTATCGAACAGGTATGCAAAATACT
CGAAGAAGTATGTGCTAAGAGCACGGACATGGGCACTTGGGTGCGAAAGGCATCGGA
TTGTAGCTGGTAGACGACACTTCAAAGATTTCGAGAGGCCATCTGCAACTGCAAAATCCAG
ATTCTCCCCACTCTAGCACAGGCTGAACTGCTCTAGCTGTTGCTTCCAGTAACTCCAGCTG
TGGTAAAGTGCAAGCTTCTCCATCCACAAGCATCCCATCTGTGGCCCGTAAAAGTAGGTCA
GACGCACTTGCCAAAATTACCAATGATTCTGGTGTATCATGATTTCTCATCAGTTCTACCA
TTAATCTAACTATTCTCTGGTTGATTCTCCACAGCTTCTGACCCTGGTCGTCATCTTTGGCA
ATCCAAGGTTGCAATTCCTCTCCGCCTGAAGAACAATTGCTGTAGCTGCTTTTGTGGGCG
ATGCTGATAACAACATTGCAAAGTGCATCCACAACCTGTCTCCATCCTTGTGTGCAGAAGT
GCTTACTGCACTAGGTTGAGTTTCTGGTGTGCAATCATTTTGTGCCATAGAAGTGAGACA
ACAGAGAAACATAACTCCTGTTTCTCCACAAGCAAGGTTCTTAAAAGAACTACTGCACCC
AAATTGAATCCAATGTGCCTGTCCATTGTTAAGAAATTGGCCAAATCTGAAGCATCTAATG
GGAAAGCTGCAATATCCTTGCCCAAGGTAATCCTGAACCTTCTTTCGAGCATGATGATCT
CTCAGTTATAATAGTAGTGTCTGAATGAGTGGATTCTACACACTGTGAAGCAGATGACTGA
CCTGAATCAAAGCAGCGAATGCTTAATTTTTTATTTTGCCTTTTGGCCATCAAACACACAA
GGGAATCTTTCCAAACAGGTAATGCAAGTGTGCCTCTAAAGGCTCAGCTTTTGTAACTAT
GGATGCAACAGCTTTGCTGTGAATATCAATGAGGTTAAATAATGATGAAGCCCTGGTGTG
AATTCATTATCCCACCTTGCAACCGCATCAAACAGACAGAGCATGCATACAAGCCTTTGAC
CGTCGGAATAGTTCAGAAACATGAGCAGCAACCATAGCAGCAGCCACTATCTCATTGAA
CTGTAACTCCAAGAGGTACCAACAGAAGACGGTTTCAAAGAAAAAAGTGCCTCCAAGATT
GCTAAGATTCTATGAGTATGATGAACAGCAGAATCAAACACTACTTTGGAACCTCCGACTA
GATCCATTGCATCCATTTAATTTACAGACTTGGCCAAAGTACCTAAATCATTAGAAGTTG
AATGATTACTGCCACTTGAACAAAAGGGAACAATTGAAGCTCACAGGCAAGAGCACATA
CAGCAGCAAGAACATAAGAATCAAATGCTGCAACAGGACCTTGTTTTTTTATCTNNNNN
NNNNNCNCGGNNCTTCCNCCGCCACCANGGCCCTTCTTCACCGCNCNTTGTTCGCANCA
CGGTGCTCTTCCANTTNNNNNTTNNNNNNNAAAAANNNNNNNANN
```

### >*AtGI*

```
MASSSSSERWIDGLQFSSLLWPPPRDPQHKDQVVAYVEYFGQFTSEQFPDDIAELVRHQYPS
TEKRLDDVLAAMFVLHHPHGHAVILPIISCLIDGSLVYSKEAHPFASFISLVCPSSENDYSEQW
ALACGEILRILTHYNRPIYKTEQQNGDTERNCLRKATTSGSPTSEPKAGSPTQHERKPLRPLSPW
ISDILLAAPLGIRSDYFRWCSGVMGKYAAGELKPPTIASRGSKGHPQLMPSTPRWAVANGAGV
ILSVCDDDEVARYETATLTAVAVPALLPPPTSLDEHLVAGLPALPEPYARLFHRYAYAIATPSAT
QRLLLGLLEAPPSWAPDALDAAVQLVELLRAEDYASGVRLPRNWMHLHFLRAIGIAMSMR
AGVAADAAAALLFRILSQPALLFPPLSQVEGVEIQHAPIGGYSSNYRKQIEVPAEATIEATAQ
GIASMLCAHGPEVEWRICTIWEAAAYGLIPLNSSAVDLPEIIVATPLQPPILSWNLYIPLLKVLEYL
PRGSPSEACLMKIFVATVETILSRTFPPESSRELTRKARSSFTTRSATKNLAMSELRAMVHALFL
ESCAGVELASRLLFVVLTVCVSHEAQSSGSKRPRSEYASTTENIEANQPVSNNQTANRKS RN
V
KGQGPVAAAFDSYVLAAVCALACEVQLYPMISGGGNFSNSAVAGTITKPKVINGSSKEYGAGID
SAISHTRRILAILFSLKPSSVGPWSYSSSEIVAAAMVAAHISELFRRSKALTHALSGLMRC
KWDKEIHKRASSLYNLIDVHSKVVASIVDKAEPLA YLKNTPVQKDSVTCLNWKQENTCAST
TCFDTAVTSASRTEMNPRGNHKEYARHSDEGSGRPSEKGIKDFLLDASDLANFLTADRLAGFYC
GTQKLLRSVLAKEPELSFSVVSLLWHKLIAPAIEQPTAESTSAQQGWRQVVDALCNVVSATPA
KAAA AVVLQAERELQPWIAKDDEEGQKMWKINQRIVKVLVELMRNHRDPESL VILASASDLL
LRATDGM LVDGEACTLPQLELLEATARA IQPVLAWGPSGLAVVDGLSNLLKCRLPATIRCLSH
PSAHV RALSTSVLRDIMNQSSIPIKVTPKLPTEKNGMNSPSYRFFNAASIDWKADIQNCLNWE
AHSLLSTTMPTQFLD TAARELGCTISLS
```



---

>A/GI

MASSSSSSERWIDGLQFSSLLWPPPRDPQQHKDQVVA YVDYFGQFTSEQFPDDIAELVRHQYP  
STEKRLDDVFLAMFVLHHPHGHAVILPIISCLIDGSLVYSKEAHPFASFISLVCPSSENDYSEQ  
WALACGEILRILTHYNRPIYKTEQQNGETERNCFSKATTSGSPTSEPKAVSPTQHERKPLRPLSP  
WISDILLAAPLGI RSDYFRWC SGVMGKYAAGELKPPTIASRGS GKHPQLMPSTPRWAVANGA  
GVILSVCDEEVARYETATLTA VAVPALLLPPTTSLDEHLVAGLPALPEPYARLFHRY YAIATPS  
ATQRLLGLLEAPPSWAPDALDAAVQLVELLRAAEDYASGVRLPRNWMHLHFLRAIGIAMS  
RAGVAADAAAALLFRILSQPALLFPPLSQVEGVEIQHAPIGGYSSNYRKQIEVPAAEATIEATAQ  
GIASMLCAHGPEVEWRICTIWEAA YGLIPLNSSAVDLPEIIVATPLQPPILSWNLYIPLLKVLE YF  
PRGSPSEACLMKIFVATVETILSRTPPESSREHTRKARSSFTTRSATKNLVMAELRAMVHALFL  
ESCAGVELASRLLFVVLTVCVSHEAQSSGSKRPRSEYASTTENVEANQPVS DNQTANRKS RNV  
KGQGPVAAFDSYVLAAVCALACEVQLYPMISGGNFNSAVARTITKPKVINGSSNEYGAGV  
DSAINHTRRILAILEALFSLKPSVGTWPSYSSSEIVAAAMVA AHISELFRRSKALHTALSGLMR  
CKWDKEIHARSSLYNLIDVHSKVVASIVDKAEPL EAYLKNTPVQKDSLTCLNWKQDNTCAS  
TTRFDTA VTSASRTE MNPRGNHKYARHSGEGSGRPSSEKGIKDFLLDASDLANFLTADRLAGFY  
CGTQKLLRSVLAEKPELSFSVVSLLWHKLIDAPEIQPTAESTSAQQGWRQVVDALCNVVSATP  
AKAAA AVVLQAERELQPWIAKDDEEGQKMWQINQRIVKVLVELMRNHDRPESLVILASASDL  
LLRATDGMLVDGEACTLPQLELLEATARA IQPVLA WGPSGLAVVDGLSNLLKCR LPATVRCL  
SHPSAHVRALSISVLRDIMNQSSXPIK VTPKLPTEKNGMNSPSYRFFNAATIDWKADIQKCLN  
WEAHSLLSTTMPTQFLDTAARELGCTISLS

>TaGI

MSVSNKWKIDGLQFSSLFWPPPHDVQKQQAQILAYVEYFGQFTSDSEQFPEDVAQLIQSCYPS  
KEKRLVDEVLATFVLHHPHGHAVVHPILSRIIDGTL SYDSHGSPFNSFISLFTQSSEKEYSEQW  
ALACGEILRVLTHYNRPIFKVADCNHQIRPGH SKLFCTEKAITLPGNEPEGKPLRPLSPWITDIVL  
TAPLGI RSDYFRWCGGVMGKYAAGGELKPPTTAYSRGAGKHPQLMPSTPRWAVANGAGVIL  
SVCDEEVARYETANLTA AAVPALLLPPTTPLDEHLVAGLPPEPYARLFHRY YAIATPSATQR  
LLFGLLEAPPSWAPDALDAAVQLVELLRAAEDYATGMRLPKNWLHLHFLRAIGTAMSMRAGI  
AADTAAALLFRILSQPMLLFPPLRHAEGVEVQHEPLGGYVSSYKRQLEVPASETTIDATAQGIA  
SLLCAHGPDVEWRICTIWEAA YGLLPLNSSAVDLPEIIVAAAPLQPPTLSWSLYLPLLKVFEYLP  
RGSPSEACLMRIFVATVEAILRRTFPSETSESSKRPRSQSKNLAV AELRTMIHSLFVESCAS MNL  
ASRLLFVVLTVCVSHQALPGGSKRPTGSENHSSEEATEDPRLTNGRNRVKKKQGPVGT FDSYV  
LAAV CALSCELQLPILCKSATNSNVKDSIKILKPGKNGISNELQNSISSAILHTRRILGILEALF  
SLKPSVGTSWNYSSNEIVAAAMVA AHASELFRRSKACLNALSSLKRCKWD AEISTRASSLYH  
LIDLHGKTVSSIVNKAEPLEAHLTFTSVKR DDEQHIENGTSSSGSGNLEKNGSASHMKNGLS  
RPLLKCSSEARRNGNVASTSGKVPATLQAEASDLANFLTMDRNGGYRGSQTLLRSVISEKQEL  
SFSVVSLLWHKLIASPETQMSAESTSAHQGWRKVVDALCDVVSASPAKASTAIVLQAEKDLQP  
WIARDDEEGQKMWRVNRIVKLI AELMRNHDSPEAL IILASASDLLRATDGMLVDGEACTLP  
QLELLEVTARAIHLIVEWGDPGVAVADGLSNLLKCR LSPTIRCLSHASAHVRALSMSVLRDILN  
SGPLGSTKIIQGEQRNGIQSPTYQCAAANTVNWQADVERCIDWEARSRRATGMTLAFLTAAA  
NELGCPLPC

>ZmGI

MSESNVKWIDGLQFTSLYWPPLDAEQKQAQILAYVEYFGQFTADTDQFPEDIAQLIQSSYPSK  
ENRLVGEVLATFVLHHPHGHAV AHPILSRIIDGTLCYDRHGPPFSSFISLFSHNSEQEYSEQWA  
LACGEILRVLTHYNRPIFKVERQHTEAECSS TSDQATSSDSTDKRSNNSPGNESDRKPLRPLTP  
WITDILLAAPLGI RSDYFRWCGGVMGKYAAGGELKPPTTACSRGS GKHPQMPSTPRWAVAN  
GAGVILSVCDEEVARYETANLTA AAVPALLLPPTTPLDEHLVAGLPPEPYARLFHRY YAIAT  
PSATQRLLFGLLEAPPSWAPDTLDAAVQLVELLRAAEDYASGMRLPKNWMHLHFLRAIGTAM  
SMRAGIAADTAAALLFRILSQPTLLFPPLRHAEGVEVHHEPLGGYVSSYKKQLEVPASEATIDA  
TAQGIASLLCAHGPDVEWRICTIWEAA YGLLPLSSSAVDLPEIIVAAAPLQPPTLSWSLYLPLLK  
VFEYLPRXSPSEACLMRIFVATVEAILRRTFPSETSEQPRKPRSQSKNLAV AELHTMIHSLFVES  
CASMDLASRLLFVVLTVCVSHQALPGGSKRPTGSDNHSHEEATEHSRLTNGRSRCKKRQGPV  
ATFDSYVLAAVCALSCELQLPFITKNGSHSNLKD SMKIIISGKNNGMNNE LHNSISSAILHTRRI  
LGILEAVFSLKPSVGTWSYSSNEIVAAAMVA AHVSELFRRSR PCLNALSALMRCKWD AEIST  
RASSLYHLIDLHGKTVSSIVNKAEPLEAHLTLTPVKRDNQHHREESNTSSLD SVKLENKNGSTS  
HKNGFSRPLLKCAEEVLLNGDVASTSGKSIASLQVEASDLANFLTMDRNGGYRGSQTLLRSV  
LSEKQELCFV ASLLWQKLIASPEMQMSAESTSAHQGWRKVVDALCDVVSASPTKASAAIVL  
QAEKDLQPWIARDDEEQGQKMWRVNRIVKLI AELMRNHDSPEALVILASASDLLRATDGML

---

---

VDGEACTLPQLELLEVTARAVHLIIEWGDSGLSVADGLSNLLKCRSLSTTIRCLSHPSAHVRALS  
MSVLRDILNSGSVNPNTIQGEQQRNGIQSPSYRCLAAGIINWQADVERCIEWEAHSRRATGLT  
LAFLSAAAKELGCPLPS

>LpGI

MSVSNKGWIDGLQFSSLFWPPPHDAQQKQAQTLAYVEYFGQFTSDSEQFPEDVAQLIQSYYP  
KEKRLVDEVLATFVLHHPHGHAVVHPILSRIIDGSLSYDRHGSPFNSFISLFTQTAEKEYSEQW  
ALACGEILRVLTHYNRPFKVAECNDTSDQATTSYSLHDKANSSPENEPERKPLRPLSPWITDIL  
LNAPLGIIRSDYFRWCGGVMGKYAAGGELKPPTTAYSRGAGKHPQLMPSTPRWAVANGAGVI  
LXVCDEEVARYETANLTAAPALLPPPTTPLDEHLVAGLPPLEPYARLFHRYAIAATPSATQ  
RLLFGLLEAPPSWAPDALDAAVQLVELLRAAEDYATGMRLPKNWLHLHFLRAIGTAMSMRA  
GMAADTAAALLFRILSQPTLLFPPLRHAEGVVQHEPLGGYVSSYKRQLEIPASETTIDATAQGI  
ASLLCAHGPDVEWRICTIWEAAAYGLLPLNSSAVDLPEIVVAAPLQPPTLSWSLYLPLLKVFEYL  
PRGSPSEACLMRIFVATVEAILRRTFPSETEPSKKPRSPSKSLVAELRTMIHSLFVESCASMN  
ASRLXFVVLTVSXSXQALPGGSKRPTGSENHSSEESTEDSKLTNGRNRCKKKQGPVGTFD  
SYVLAAVCALSCQLFPILCKNVTKNIKDSIKITMPGKTNGISNELHNSVNSAILHTRRILGILEAL  
FSLKPSSVGTSSWSYSSNEIVAAAMVAAHVSELFRSRPCLNALSALKRCKWDAEISTRASSLY  
HLIDLHGKTVSSIVNKAEPLEAHLNLTAVKKDDQHHEESNTSSSDYGNLEKSKKNGFSRPL  
MKCAEQARRNGNVASTSGKATATLQAEASDLANFLTMDRNGGYGGSQTLRLVMSEKQELC  
FSVVSLWHLKLIASPETQMSAESTSAHQGWRKVADALCDVVSASPAKASTAIVLQAEKDLQ  
WIARDDEQGQKMWRVNRIVKLIAEMLRNHDSPEALILASASDLLLRATDGMLVDGEACTL  
PQLELLEVTARAIHLIVEWGDVAVADGLSNLLKCRSLSTTIRCLSHPSAHVRALSMSVLRDIL  
NSGPISSTKIIQGEQQRNGIQSPSYRCAAASMTNWQADVERCIEWEAHNRQATGMTLAFLLAAA  
NELGCPLPC

>OsGI

RKPLRPLSPWITDILLAAPLGIIRSDYFRWCGGVMGKYAAGGELKPPTTAYSRGSGKHPQLMPS  
TPRWAVANGAGVILSVCDEEVARYETANLTAAPALLPPPTTPLDEHLVAGLPPLEPYARL  
FHYAIAATPSATQRLFLGLLEAPPSWAPDALDAAVQLVELLRAAEDYDSGMRLPKNWMLH  
HFLRAIGTAMSMRAGIAADTSAALLFRILSQPTLLFPPLRHAEGVELHHEPLGGYVSSYKRQLE  
VPASEATIDATAQGIASMLCAHGPDVEWRICTIWEAAAYGLLPLSSAVDLPEIVVAAPLQPPTL  
SWSLYLPLLKVFEYLPRGSPSEACLMRIFVATVEAILRRTFPSETSESRKPRSQSKNLVAELR  
TMIHSLFVESCASMDLASRLLFVVLTVCVSHQALPGGSKRPTGSDNHSSEEVNTDSRLTNGRN  
RCKKRQGPVATFDYSVLAAVCALSCQLFPFISKNGNHSNLKDSIKIVIPGKTTGISNELHNSIS  
SAILHTRRILGILEALFSLKPSSVGTSSWSYSSNEIVAAAMVAAHVSELFRSRPCLNALSALKQC  
KWDAEISTRASSLYHLIDLHGKTVTSIVNKAEPLEAHLTLTPVKKDEPPIEKNINSSDGGALEK  
KDAASRSHRKNGFARPLLKCAEDVILNGDVASTSGKAIASLQVEASDLANFLTMDRNGGYRGS  
QTLRLSVLSEKQELCFSVVSLWQKLIASPEMQMSAESTSAHQGWRKVVDALCDIVSASPTKA  
SAAIVLQAEKDLQWIARDDEQGQKMWRVNRIVKLIAEMLRNHDSPEALVILASASDLLLR  
ATDGMLVDGEACTLPQLELLEVTARAVHLIVEWGDGVSADGLSNLLKCRSLSTTIRCLSHPS  
AHVRALSMSVLRDILNSGQINSSKLIQGEHRNGIQSPTYQCLAASIINWQADVERCIEWEAHSR  
RATGLTLAFLTAAKELGCPLTC

>HvGILprotein

MSASNGKWIDGLQFSSLFWPPPHDAQQKQAQILAYVEYFGQFTSDSEQFPEDVAQLIQTCYPS  
KEKRLVDEVLATFVLHHPHGHAVVHPILSRIIDGTLSDSHGSPFNSFISLFTQSSEKEYSEQW  
ALACGEILRVLTHYNRPFKVADCNNTSDQATTSCSAQEKANYSFGNEPERKPLRPLSPWITDIL  
LTAPLGIIRSDYFRWCGGVMGKYAAGGELKPPTTAYSRGAGKHPQLMPSTPRWAVANGAGVI  
LSVCDEEVARYETANLTAAPALLPPPTTPLDEHLVAGLPPLEPYARLFHRYAIAATPSATQ  
RLLFGLLEAPPSWAPDALDAAVQLVELLRAAEDYATGMRLPKNWLHLHFLRAIGTAMSMRA  
GIAADTAAALLFRILSQPTLLFPPLRHAEGVEVQHEPLGGYVSSYKRQLEVPASETTIDATAQGI  
ASLLCAHGPDVEWRICTIWEAAAYGLLPLNSSAVDLPEIVVAAPLQPPTLSWSLYLPLLKVFEYL  
PRGSPSEACLMRIFVATVEAILRRTFPSETSESKRPRSQSKNLVAELRTMIHSLFVESCASMN  
LASRLLFVVLTVCVSHQALPGGSKRPTGSENHSSEATEDPRLTNGRNVKVKKQGPVGTFD  
SYVLAAVCALSCQLFPILCKSATNSKVKDSIKILKPGKNNGISNELQNSISSAILHTRRILGILEAL  
FSLKPSSVGTSSWNYSSNEIVAAAMVAAHVSELFRSRCLNALSALKRCKWDAEISTRASSLYH  
LIDLHGKTVSSIVNKAEPLEAHLTFTSVKRDGQQHIEENSTSSSGNGNLEKKNASASHMKN  
GFSRPLLKCSSEARRNGNVASTSGKVPATLQAEASDLANFLTMDRNGGYRGSQTLSSVISEKQEL  
CFSVVSLWHLKLIASPETQMSAESTSAHQGWRKVVDALCDVVSASPAKASTAIVLQAEKDLQ  
PWIARDDEEGQKMWRVNRIVKLIAEMLRNHDSPEALILASASDLLLRATDGMLVDGEACTL

---

---

PQLELLEVTARAIHLVWGDGVA VADGLSNLLKCRLSPTIRCLSHASAHVRALSMSVLRDIL  
NSGPLGSSKTIQGEQRNGIQSPNYQCAAANTVNWQADVERCIDWEARSRRATGMTLAFLTAA  
ANELGCPLPC

>PtGILike

MASSSERWIDGLQFSSLFWPPPQDAQQRKAQITAYVDYFGQCTSEHFPPDDISELIRNRYPSKD  
KRLFDDVLATFVLHHPEHGHA VVLPPIISCIIDGTLVYDRSSPPFASFISLVCPSSENEYSEQWAL  
ACGEILRILTHYNRPIYKREQQNNETDRSSSDSHATSSESAEGKSTSMPLVQKERKPFRLSPWI  
TDILLAAPLGRSDYFRWCSGVMGKYAAGELKPPTTSSRGSGKHPQLIPSTPRWAVANGAGV  
ILSVCDEEVARYETATLTAAAVPALLPPPTTALDEHLVAGLPALPEPYARLFHRYAIAIATPSAT  
QRLLLGLLEAPPSWAPDALDAAVQLVELLRAAEDYASGIRLPRNWMHLHFLRAIGTAMSMRA  
GIAADAAAALLFRILSQALLFPPLRQVEGVEVQHEPLGGYISCYRKQIEVPAAEATIEATAQGI  
ASMLCAHGPEVEWRICTIWEAA YGLIPLSSAVDLPEIIVATPLQPPLSWNL YIPLLKVLEYLP  
RGSPSEACLMIKIFVATVEAILQRTFPPEASREQTRRTRYFSSLGPASKNLAV AELRTMVHSLFLE  
SCASVELASRLLFVVLTVCVSHEAHSRGSKRPRGEENDLPEDGTEDSQSTSEMRRNMKSRRM  
KKQGPVAAFDSYVLAAVCALACELQIFPFVSRGNSHSTSKHSETVAKPAKLNAGAVSEFQTSLN  
SAIHHTHRLSILEALFSLKPSTIGTWSYSSNEIVAAAMVAAHVSELFRRSKACMHALSVMR  
CKWDNEIYTRASSLYNLIDVHSAKAVASIVNKAEP LGAHLHAPVWKDSLVCSDGNKQNRSAST  
GCFNSGQSSALQSTELVHSETKLCGRASHSEEGSGSTSGKGIAGLPLDASDLANFLTMHRHIG  
FNCSAQVLLRSVLPEKQELCFSVVSLLWQKLIASPETQPSAESTSAQQGWRQVVDALCNVVS  
SPTIAATAVVLQAERELQPWIAKDDDSGQIMWRINQRIVKLIVELMRNHDTPELVILASASDL  
LLRATDGMMLVDGEACTLPQLELLEATARA VQPVLEWGESGFAVADGLSNLLKCRLPATIRCL  
SHPSAHVRALSTSVLRDIQHTGSIKPASKLTHRNHIGPSYQYLRSDVINWQADIEKCLTWEAH  
SRLATGMPVHHLDTAAKELGCTISI

>RcGIProtein

MASSERWIDGLQFSSLFWPPPQDAQQRKAQITAYVEYFGQFTSEQFPDDIAEVTASHFFSSNPA  
TFVLHHPEHGHA VVLPPIISCLIDGTLVYDRSTPPFASFISLVCPSSENEYSEQWALACGEILRVLT  
HYNRPIYKKEQQKSETEKSGGGEDAVNGGLADGESSHTPPAQKERKPLRPLSPWITDILLTAPL  
GIRSDYFRWCSGVMGKYAGGELKPPTTASSHGSGKHPQLMPSTPRWAVANGAGVILSVCDDDE  
VARYETATLTAAAVPALLPPPTTALDEHLVAGLPALPEPYARLFHRYAFAFATPSATQRLLLGL  
LEAPPSWAPDALDAAVQLVELLRAAEDYASGIRLPRNWMHLHFLRAIGIAMSMRAGIAADAA  
AALLFRILSQALLFPPLRQVEGMEVHHEPLGAYSSSYRKQIEVPAAEATIEATAQGIASMLCA  
HGPEVEWRICTIWEAA YGLLPLGSSAVDLPEIIVAAPLQPPILSWNL YIPLLKVLEYLPRGSPSEA  
CLIKIFVATVEAILQRTFPPESSREQTRKAKYLFGLGSASKNLAV AELRTMVHSLFLKSCASVEL  
ASRLLFVVLTVCVSHEAQSNGTKRPRGEENFQPDGDNEDWLTSEAHSKMKPRKIKKQGPVA  
AFDSYVLAAVCALACELQIFPFVSSGNHSSSNDLDTLAKSMKMNGSIREFQNSIDSAVHHHTH  
RILAILEALFSLKPSTVGTWSYSSNEIVAAAMVAAHVSELFRRSKACMHALSVMRCKWDNE  
IYTRASSLYNLIDIHSAKAVASIVTKAEPLA YLHVPVWRDSL VHF DGKKRNRSSASCDFSGQS  
SASQREESAHSDSKIGTERLQSGEGSGSTLGNSIAGFPLDASDLANFLTMDRHIGFNCSAQVFLR  
SVLAKKQELCFSVVSLLWHKLISAPETQPSAESTSAQQGWRQVVDALCNVVSATPTKAAA  
AVVLQA EKELQPWIAKDDDDQGQKMWRINQRIVRLIVELMRNHDTPELVILASASDLLLRATDG  
MLVDGEACTLPQLELLEATARA VQPVLEWGESGFAVADGLSNLLKCRLPATIRCLSHPSAHVR  
AVSTSVLRGILYTGSIKRTSNRVDINGIRGPSYQYFNIDVTDWQTDIEKCLTWEAHSRLATGMPI  
QFLDTAAKELGCTISI

>MeGI

KKQGPVAAFDSYVLAAVCALACELQIFPFVSSGSNHSSSNDLGTAKSVKLNCGNGSSREFQS  
SFDSA VHHHTHRLAILEALFSLKPSSVGTWSYSSNEIVAAAMVAAHVSELFRRSKACMHALS  
VMRCKWDNEIYTRASSLYNLIDIHSAKAVASIVTKAEPLAHLQVPVWKDSLVCDFDGKRQNK  
SIRCFDSGQSSASQCVSTHSDTTITERSSCSKEGSGSTLGKDIAAFPLDASDLANFLTMDRHIG  
FNLGAVVLLR TLLVEKQELCFSVVSLLWHKMIAAPETQPSAVSTSAQQGWRQVVDALCNVVS  
ASPTKAATAIVLQAERELQPWIAKDDDDQGQKLWRINQRIVRLMVELMRNHDTPELVILASAS  
DLLLRATDGMMLVDGEACTLPQLELLEATARA VQPVLEWGESGFAVADGLSNLLKCRLPATIR  
CLSHPSAHVRALSTSVLRSILHTCSIKPTANQADINDIRGPSYQLFKVDVIDWQ

---

## Appendix II CONSTANS and CO-like sequences

Encoded CO protein sequences described in Chapters 2 and 3

### >*Bn*CO

MFKQESNNIGSEENNTGPRACDTGSGTICTVYCHADSAYLCNSCDAQVHSANRVASRHKRVR  
VCESCERAPAAFMCEADDVSLCTACDLEVHSANPLARRHQRVPVVPITGNSCSSLATANHTTV  
TEPEKRVVLVQEDAKETASWLFPKNSDNHNNNNQNNELLSDDYLDLADYNSSMDYKFTGQ  
YNQPTQHKQDCTVPEKNYGGDRVVPVLEETRGNLHHKQHNITYGSSGSHYNNNGSINHNA  
NPSMETDFVPEQTAPDKTVSHPKTHKGKIEKLPEPLIQILSPMDREARVRLRYREKKRKRKFEKT  
IRYASRKAYAERRPRINGRFAKISETEVEDQEYNTMLMYDYDTGYGIVPSFYGQK

### >*Bv*COL2

MGGGLMAAKLCDSCKSATATIFCRADTAYLCISCDAKIHAANKLASRHARVWVCEVCEHAP  
ATVTCKADAAHLCACTCDRIHSANPLARRHERVPLTPFYDPLSPPNTTNNNNDDSDSSATAAA  
AAKSAAINKFLGDEYYSDADEAEAASWLLPNPNKTDEPKSIDYLFSSSGNDGDDIDPYLDLDF  
GAEAKPDPDLSSDGVVPDPDQKGVHHHHLTTLQHPAASMFLSSYHHHHHHHHVSNNGHF  
DGFENSSAACKPFALSSYHTQPSLSHSVSSSSLDGFGVVPDASNITDVASTGFDKQQQMKIIGMD  
REARVRLRYREKRKNRKFECTIRYASRKAYAETRPRI

### >*Bv*COL1

MMKEEVSGSDTNSWARVCDTCRAAPCTVYCRADSAFLCTSCDARIHAANQVASRHERVWVC  
EACERAPAAFLCKADAASLCATCDAEIHSANPLARRHQRVPIMPVAGCVYGPQGGRMSEDRF  
LTLPEGDDHTTDHEGDEDEAASWLLNPVKNNSNNQNTNGFLTGGGEVDEYLDLLEYNSGAD  
NQLCEQYNQQQEFKVEKNCGGDSVVPVQCREAKDHQIQYQNFLFGMECETKSGYTYNTSIS  
QSVSVSSMDVGVVPESAMSDISMSHPRPPKGTIDLFSSPPMQVPTQLSPLDREARVMRYREKK  
KNRKFECTIRYASRKAYAETRPRIKGRFAKRTDVEAEMDQMFTNSLMADSGYGIVPSY

### >*Bv*COL3

MKLECDYCKNAAVLYCEADSANLCLLCDRIHSANSLSLKHIRIPRFGISNPNSEPKSAIDGCP  
SASELAPFWGINDLVVPCLGDDEKDAVIQQLVKLSKWDLEGREYSSEIGPGTPSLDAGEGDGS  
ELLFQNTNFTSLLMKEGETVSVEDRDFMWFADADYQPPQEWGQCGRPFDLDLEESSCASPR  
VSGDVCEMQCSNILEGSESRRDDGGNQTVDNDIFVTLENCVPSIGLSSDAEMIKPEAIDGEQNF  
QVMEWPYWRKPLGNTDMEQLAENRGKAMRLRYKEKKKTRRYDKHIRYESRKARADIRQRVK  
GRFVKASDTPDEGSGL

### >*Pa*COL

MVKEEDCKVPKEAGIVKEFQAWTMPKPCNVCRIASASLYCRADSAYLCSGCDVKVHGANK  
LASRHERVWLCEVCEQAPAAVTCKADAASLCVSCDADIHSANPLARRHDRVPIVPFYECA  
SVAKTFLPPPPPTSSLQSDVVGTLDYEDDDEDDEIYAAEAASWLLPNPKSSAEGAKN  
CDDGGSCFGVDAGPPVNAAGGYFSVVDLFPDVPYLDLDYASPLEATGGTDSVVPVQSN  
VSSQDGAVSTPSDFDTEKATYSYTTTTSLSHSVSSSSLDVGVVPDATLSDMSRPLNRGV  
FELANPGVVNVGIQYVQLDREARVRLRYKEKRKNRKFECTIRYASRKAYAETRPRIKGRFA  
KRVDADVAQMYTSAELSYGLVPSF

### >*Ta*Hd1L

MFMNCNFNSNLEKEAGRTSFPWARPCDGCCHAAPSAVYCCADAAYLCASCDTQVHSANRVA  
SRHERVRCETCESAPAVLACHADAAALCTACDAQVHSANPIAQRHQRVPVLPALPAVALPAA  
SGFVEAEASVTAHGDKKEEGEVDSWLLRRNSDDNCKANKIDRYFNLVGYNMYDNTCDPRP  
EEQYRMQEQQHVQNRVIEKEGCECVVPPQVVMASEQQESDYGTIGAGQAASVTAMTSTYTA  
SISNDISFSSMEVGIVPDNTRPDISNSNILTSSEAMELSGHSQMPVHFNSMDREARVRLRYKEKK  
QTRKFQKTIRYATRKYAEARPRIKGRFAKRSIDIEHEEDHMLSPPALQDTSSYNTAPWF

### >*Zm*CO5

MDTAAELELGLELEQKPAAGYWSVVGARPCDACAAPARLHCRADGAFLCPGCDARAHGA  
GSRHARVWLCEVCEHAPAVVTCRADAAALCAACDADIHSANPLARRHERLPIAPLFGALADA  
PQFPSPALAAAAGAEAPAPTPAQGEVAEDYGSSEAEAAASWLLPEPDNSHEDSAADTFFAES

---

---

DAYLGADLDFARCMDGVKAIGVPVAPPELDIGAGSFCYPEHSMNHILSSSSEVAVVPDAQAAG  
LPVVVVVSRGEEREARLMRYREKRKNRRFDKTIRYASRKAYAETRPRIKGRFAKRRAEGEDE  
ALEHEEGACFSPAGSAPAASDGVVPSLC

>*Zm*COL6

MELHKYWGVGGRRRCGSCEGAPAAVHCRTC VGG SFLCTTCDARPAHARLGHERVWMCEVCE  
LAPAAVTCKADAAVLCAACDS DIHDANPLARRHARVPVAPIGSEAAAAAVEAMLFGTGEAA  
ASEADEQHAAAEHAHAHALNLNVEAKDMKLDYLFSELDPYLSVEIPRFQHADS VVPNGAGA  
AVELDFTCGIGVKHSSYSSYTATSLAHSGSSSEVGVVPEAFGGSGSGGGSFELDFTRPKPQAYM  
PYTGTPQSHSVPSADVEVVPERGDLAAVRPVPLMGESREARLMRYREKRKNRRFEKTIRYASR  
KAYAETRPRIKGRFAKRADHDGDADADDAEAEAEAEAAVPMSYVLD FGYGVVPSF

>*Gm*COL2

MLEGQATTPTWPRMCDTCRSVPSTVFCRSHTAFLCATCDTRLHVSLTWHERVVWCEACER  
APAAFLCKADAASLCASCDADIHAANPLASRHRVPILPIAAANNNNNDDDDVADVDED  
ETASWLLL NPIKSATVPNTNNNNNNNGFLYNGEVDEYLDLVDNCNSCGDNNHFASAAATT  
DHYAQH QHFAGVSQKSYAGDSVVPVQQHQHFQLGLDFDNSKPAFSYNGSVS QSVSVSSMD  
IGVVPESMRDVSIAHTRPPKGTIDLFSGPPIQVPSHFSPMDREARV LRYREKKMRKFE  
KTIRYASRKAYAETRPRIKGRFAKRTDVEAEVDQMFSSTLITEVGYGIVPSF

>*Pp*COL1

MPKPCDACHVSSAAVFCRADAAYLCVCGDGK VHGANKLASRHERVWMCEVCEVAPAVVTC  
KADAASLCVACDTDIHSANPLAQRHERVPTPLFESASPLRGPDFCVL VSENGCHDLLKGCED  
ASVVEAVSWLLPHPKISTNSIIRGSAADMGSSPFHDRPFSPKPKKQKVELPADIFSDVDPFLD  
LDDATVTGIQPSLVPVHMPECESED TSLAHSMDPSFTKFPLSAKSGYSYGTSTLTQSISCSLD  
AAVVPDSSSLDISTPYLDSSQDMSARLPHQTGGPIDTVDREARV LRYKEKRQKRKFEKTIRY  
ASRKAYAESRPRIKGRFAKRTDSDMEQFGSVDS SFGVVPSF

>*Lp*Hd1L

MKSNSSTIYEEAVGQEGSW SRLCDGCCMVPSV VYCHADSAYLCASCDVRIHSANRVASRHE  
RVCLSEAHEHAPALLQCRTDAVASCAA YEAQAHYANLLAGMHQCVPV VSHPATAIPTASLLA  
EAAVTTTILSCKEEEEASWLLSKNSANHNCSGDNRRSSSTYFGEVDEYFDL VGYN SYD SRMN  
NNRAQYVMQEQHLQPMQKEYAEKEGSECVVPSQFATASKPQQSGYALVGAEQAASMTAG  
VSVYTDSVNNISFSSMEGGIVPDNTVVDLPYSIIPTAGASSLHSGPPLQ MPLHFSSMDREAKV  
LRYKEKKKTRTFEKTTRYATR KAYAEARPRIKGRFAKISEAEMEVDQMFSAAALSDSSYSTVP  
WFQ

>*Pd*COL2

MLKQESSGSGGDNRRARLCTCRAAACTVYCRADSAYLCAGCDARVHAANRVASRHERVW  
VCES CERAPAALLCKADAASLCTACDADIHSANPLARRHQ RVPILPISGCLHGSQVGPAAGETE  
DRFTTQEGEETISEEEEEDEEASWLLLNPKNSKNQNNNGFLFEGEVDEYLDLVEYNSCTE  
NQCSDQYNQQHYCVPPKSYGGDRVVP IQYGEKDHQQQRQYHNFQLGLE YEPSKAAYSYNG  
LISQSVSMSSMDVGVVPESTMSEISISQHRTPKRTIELFSSTAIQMPSQLSPMDREARV LRYREK  
KKTRKFEKTIRYASRKAYAETRP RVKGRFAKRKDVEVEDDRTFSSTLMAGTGCGIVPSF

>*Pd*COL1

MPRFTSLILSSPLVFFLQVSLILATNAHSLSYIAKDSIIEIGIGREREEV MLKEESGGSGGVVNNW  
ARVCDTCRAAACTVYCRADSAYLCAGCDARVHAANRVASRHERVRVCEACERAPAALLCK  
ADAASLCTACDADIHSANPLARRHQ RVPILPISGYLYGTQVGPAAGETEDQFMTQEGEETIGEE  
DEDEAASWLLLNPAKNSNNQNNNGFLFGGEVDEYLDIVEYNSCAENQYSDQYNQQHY SVPP  
KSCGGDSVVP IQYGEKDHQQQQQQYHNFQLGLE YEPAKAAYSYDGSVSQGVSMSSMDVG  
VVPESAMSEISISHQSASRG TIDLFSPPIQMPSQLSPMEREARV LRYREKKKARKFEKTIRYASR  
KAYAETRPRIKGRFAKRTDVDVEVDQMFSSTLMAETAYGIVPSF

>*At*COL2

MLKEESNESGTWARACDTCR SAACTVYCEADSAYLCTTCDARVHAANRVASRERVRVCQSC  
ESAPAAFLCKADAASLCTACDAEIH SANPLARRHQ RVPILPLSANS CSMAPSETDADNDEDDR  
EVASWLLPNPGKNIGNQNNNGFLFGVEYLDLVDYSSSMDNQFEDNQYTHYQRSFGGDGVVPL  
QVEESTSHLQQSQQNFQLGINYGFSSGAHYNNNSLKD LNHSASVSSMDISV VPESTASDITVQH

---

PRTTKETIDQLSGPPTQVVQLTPMEREARVLR YREKKKTRKFDKTIRYASRKAYAEIRPRIKGR  
FAKRIETEAEAEEIFSTSLMSETGYGIVPSF

>A<sub>t</sub>COL3

MASSRLCDSCKSTAATLFCRADAAFLCGDCDGKIHTANKLASRHERVWLCEVCEQAPAHVT  
CKADAAALCVTCDRDIHSANPLSRRHERVPITPFYDAVGPASASSSVNFVDEDGGDVTASWL  
LAKEGIEITNLFSDLDYPKIEVTSEENSSGNDGVVVPVQNKLFNEDYFNFDLSASKISQQGFNFI  
NQTVSTRITIDVPLVPESGGVTAEMTNTETPAVQLSPAEREARVLR YREKRKRNKFEKTIRYAS  
RKAYAEMRPRIKGRFAKRTDSRENDGGDVGVYGGFGVVPSF

>A<sub>t</sub>CO

MLKQESNDIGSGENNRARPCDTCRSNACTVYCHADSAYLCMSCDAQVHSANRVASRHKRVR  
VCESERAPAAFLCEADDASLCTACDSEVHSANPLARRHQRVPILPISGNSFSSMTTTHHQSEK  
TMTDPEKRLVVDQEEGEEGDKAKEVASWLFPSNDKNNNNQNNGLLFSDEYLNLDYNSSM  
DYKFTGEYSQHQQNCSVPQTSYGGDRVPLKLEESRGHQCHNQNFQFNKYGSSGTHYNDN  
GSINHNAYISSMETGVVPESTACVTTASHPRTPKGTVEQQPDPASQMITVTQLSPMDREARVLR  
YREKRKTRKFEKTIRYASRKAYAEIRPRVNGRFAKREIEAEEQGFNTMLMYNTGYGIVPSF

>A<sub>t</sub>COL1

MLKVESNWAQACDTCRSAACTVYCRADSAYLCSSCDAQVHAANRLASRHERVRVCQSCERA  
PAAFFCKADAASLCTTCDSEIHSANPLARRHQRVPILPISEYSYSSTATNHSCETTVDPENRLV  
LGQEEDEDEAEAAASWLLPNSGKNSGNNNGFSIGDEFNLVDYSSSDKQFTDQSNQYQLDCN  
VPQRSYGEDGVVPLQIEVSKGMYQEQQNFQLSINCGSWGALRSSNGLSHMVNVSMDLGVV  
PESTTSDATVSNRSPKAVTDQPPYPPAQMLSPRDREARVLR YREKKKMRKFEKTIRYASRKA  
YAEKRPRIKGRFAKKKDVDEEANQAFSTMITFDTGYGIVPSF

>S<sub>t</sub>COL2

MLKKEKSGGFDGSSNNWARVCDSCHSATCTVYCRADSAYLCADCDARIHAASLMASRHERV  
WVCEACERAPAAFLCKADAASLCAASCDADIHSANPLARRHHRVPIMPPIGTLYGPPAVHTVSG  
GSMIGGTTGEGTEDDGFLSLTQDADDTIDEEDENEAAASWLLNPPVKNNNKNNNNNNNN  
QNNNYGMLFGGEVVDEYLDLAEYGGDSQFNDQYSVNQQQQHYSVPQKSYVEDSVVPVQNG  
QRKSLILYHQPPQQQQQQSHHLNFQLGMEYDNSNTGYGYPASLSHSVSISSMDVSVVPES  
ALSETSNSHPRPPKGTIDLFSGPPIQIPQLTPMDREARVLR YREKKKRNKFEKTIRYASRKAYA  
ETRPRIKGRFAKRTDVKAQVDMFSTQLMTDSSY GIVPSF

>H<sub>v</sub>COL9

XFNDGSIYENFCVDDAGLTFENYEELFGTPHIQTEQLFDDAGIDSYFGMKEMPAADCNEQLKP  
MQPECSNAVSADSSLCVPAQQAISISLFSFGFTGEGNAGDHQDCGVSPWLHPGPEGSSASGSR  
GSALSRYMEKKRRKFDKIRYASRKARADVRRKRVKGRFVKAGEAYDYDPLKKTRKFEKT  
IRYASRKAYAETRPRVKGRFAKRDVEVEDDRFSSSTLMAGTGCGIVPSF

>O<sub>s</sub>CO

MELRKYWVGGRRCGACEASPAAVHCRGCGGVYLCTACDARPGHARAHERVWVCEVCE  
VAPAAVTCKADA AVLCAACDADIHDANPLARRHARVPVAPIGSAAAAVAEAMLFVAAA  
GAEAEAVEDKAAAEHHHHQQRQHGALNLNVEAKDMKLDYLFSDLDPYLNVEFARFPHAD  
SVVPNGAGAGAAIELDFTCGLGVGVGAKQSYSSYTATDLAHS GSSEVGVVPEAMCGGGG  
AIDLDFTRPKPQPYMPYTATPPPSHSVVAQMSVVVDVGVVPERAAAMGEGREARLMRYRE  
KRKNRRFEKTIRYASRKAYAETRPRIKGRFAKRADHDADDADADADDPAAVPSSYMLDFGY  
GVVPSF

>MeCOL1

HERVWVCEACERAPAAFLCKADAASLCTACDADIHSANPLARRHQRVPILPISGCLYGPQVGP  
ATGETTEDMFMTQEGEEGVGEEEEEDEAAASWLLNNTAKNSNGQNNNGFLFGGGEVDEYLDLV  
EYNSNSCGDQNYSDQNNLQHYSPHQKSCYGGDSVVPVQCAEAAGKDQLHQYHNFHLG  
LEFESSKAAYSYNGSISHSVSISSMDVGVVDPSTMTTEASISHPRPPKGTIDLFSPPMPMPSQLS  
XXXRXAXXL

>MeCOL2

TCRAAACSYCRADSVYLCAGCDARVHAANRVASQHERVWVCEACERAPAAFLCKADAAS  
LCTACDADIHSANPLARRHQRVPILPISGSLYGPQAGNPAGETTEDMFMTQEGEEGVGEEEEE

---

---

EAASWLLPNPVKNSNGQNYNGFLFGGGDVDEYLDLMEYNSNSCGDQYSDQNNLQHYSVP  
HQBKSCYGGDSVVPNHCAEAARKDQLHHQYHNFHLGLEIDSSSKAAYSNGSISQSVSISMDV  
GVVPDSTMSEASISHHRPPKGTIDLFSSPPMQMPSQFSPGDREARVLR YREKKKTRKFEKTIRY  
ASRKAYAETRPRIRGRFAKRTDVEVEVAQ

>MeCO

CKADAASLCTACDADIHSAXPIGSAFLCKADAASLCTACDADIHSAXPIPLARRHQRVPILPISG  
QEGEEGVGEEEEDEAASWLLLNTAKNSNGQNNNGFLFGGGEVDEYLDLVEYNSNSCGDQYSDQNNLQHYSVPHQKSCYGGDSVVPVQCAEAAGKDQLHQQYHNFHLGLEFESSKAAYS  
NGSISHSVSISSMDVGVVPDSTMTEASISHPRPPKGTIDLFSSPPMPMPSQLSPRXXXQGCXXXR  
EK

---

**Appendix III ELF4 ORF sequence of Cassava**

MNNNSNHKSSRRRHKHTADDDDDDEGDPEVWSTFNSTFRQVQSVLDRNRNLIQQVNENHQSRI  
PDNMVKNVALIQELNGNISKVVSLYDLNSNFTTNYQQRNGSGSGNSNSSGRS



---

## ACKNOWLEDGEMENTS

*Better is the end of a thing than the beginning thereof. The completion of this Ph.D thesis has been a result of four and half years of research work across continents. Now that I have completed this race in one piece, I will like to say thank you to all the people who contributed to this success.*

*First of, I want to acknowledge and thank George for being my Betreuer, and also Prof Bucher for agreeing to be my second examiner. Many thanks to Prof Flüggé for accepting to chairman this thesis defense. Dr Agim Ballvora, thanks a lot for being my Beitzeister. I really appreciate all of you.*

*I am saying a big thank you to the Rockefeller foundation who provided funding for this research work. To Dr Martin Fregene, for this great opportunity, you believed in me that I could carry the vision of this project work for my Ph.D. I am sure I have not let you down. Thanks for helping me discover myself in research and for the good times we shared in Cali, alongside with your family. I will be forever grateful to you.*

*I am also thankful to Seth for your supervision, guidance, invaluable suggestions, scientific grooming, ideas, inspiration and the endless scientific conversations. Thanks for helping me and pushing me to attain this height, although I am still learning “not to trust Seth,” rule number 1! I will not forget your avoidance of the depressing administrative issues, those memories linger on but I came out matured and experienced. Thanks.*

*To all the Davis group lab members, both old and new, Mark, Eva, Jieun, Alfredo, Usman, Nora, Gosia, Elsebeth, Eleni, Zhaojun, Chiarina for making the environment conducive for research in spite of our differences and backgrounds. I gained a lot in science from Gosia, my first “floral dip” experience was from her, to Elsebeth, who introduced me to rt-pcr, Eleni helped me a lot with the university issues. Thanks Ela for being so sweet and always willing to help. I appreciate you all.*

*To Mandi, for her administrative prowess of always getting things done, the readiness to help with research supplies, for the good times and of course the bad moments, it’s part of life.*

*I will also like to thank my old and new scientific coordinators, Ralf Petri and Olof Persson for their constant support with all immigration issues. I am very grateful.*

*I also want to appreciate the research team in CIAT, Cali, all the members of the Biotech unit, Dr Tohme and Paul Chavarriaga for the Cassava transformation experience and challenges. To the Colombian crew who has made my stay in Colombia forever to remember. Thanks to ‘Chucho’, ‘Mona’, Danilo, Magdalena, Norma, Janeth, Paola, Cesar. *Que pena con la gente que el espacio no**

---

---

*me dejar a mencionar, te agradezco mucho. Many thanks also to Dr Caldas, Marcela, Andres, Eleonora and the capacitation people.*

*To my friends, Seun Oyekola, Tobi Oluwafemi, Yemi Adenle, Damaris, Camila, Takayuki, Roxana and Derry, for your constant encouragement, help, love and concern. For the times we shared. Thanks a whole bunch. Special thanks to my friend Tayo, thank you for being there for me.*

*Thanking my senior colleagues Drs Okogbenin, Egesi and Akinbo for your brotherly advise everytime I need encouragement.*

*Words cannot describe how thankful I am to all my family members, my mom, sisters, in-laws, nieces and nephew for your understanding, patience, support, prayers, and encouragement, you have been my source of strength through-out the years. I am deeply grateful to you all and i love you all. Momi ese ganni ti e pemi saye, adura yin ti gba lori mi, iku, esu ati ese ko ni ya wa o, amin.*

---

## **ERKLÄRUNG**

*„Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit –einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie – abgesehen von unten angegebenen Teil-publicationen – noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. George Coupland betreut worden.“*

*Max-Planck-Institut für Züchtungsforschung,*

*Köln: 12<sup>th</sup> of May 2009*

*Signature,*

*Sarah Adeyemo*

---

## **TAGUNGEN**

### Oral Präsentationen:

Adeyemo, O.S., Florez, C.F., Lentini, Z., Chavarriaga, P., Tohme, J., Fregene, M., Davis S.J. Genetic Transformation of Cassava and Mango for Improved Flowering.

*The Third General Meeting on Biotechnology, Breeding, and Seed Systems for African Crops.* Maputo, Mozambique, March 2007

Adeyemo O.S., Davis S.J., Fregene M. Modification of Flowering In Cassava.

*International Max Planck Research School (IMPRS) Retreat,* Monschau, Germany, March 2009

### Poster Präsentationen:

Adeyemo, O.S., Fregene, M., Chavarriaga, P., Tohme, J., Davis S.J. Modification of Flowering In Cassava Using A Transgenic Approach.

*16th International Conference on Plant and Animal Genome,* San Diego, California, January 2008

Adeyemo, O.S., Fregene, M., Chavarriaga, P., Tohme, J., Davis S.J. Strategies For Enhancing Flowering In Cassava Using Molecular Tools: Towards A More Efficient Breeding Programme. *Global Cassava Partnership-1 Ghent, Belgium, July 2008*

---

## LEBENS LAUF

### Persönlicher Informationen

Name: Adeyemo  
Vorname: Oluwabusayo Sarah

Geburtstag: 22.März 1978  
Geburtsort: Ibadan, Nigeria  
Nationalität: Nigerianischen

### Ausbildung

- 1990-1995 High School St Louis Grundschule, Mokola, Ibadan, Nigeria
- 1995-1997 Studium die Biochemie, Ogunstadt Universität, AgoIwoye, Nigeria.
- 1997-2000 Studium die Biochemie (B.Sc), Universität von Ibadan, Nigeria.  
*Diplomarbeit: Estimation of Genetic Diversity and relationships within a collection of Yam (Dioscorea sp) germplasm using Random Amplified Polymorphic Deoxyribonucleic acids (RAPDs) markers*  
Betreuer: Dr J.K. Ikea
- 2001-2003 Studium der Umwelt Biologie (M.Sc.). Universität von Ibadan, Fakultät für Landwirtschaft, Pflanzenschutz und Umweltbiologie. *Diplomarbeit: Morphological and Biochemical Characterization of foliage colour mutants of Cowpea (Vigna unguiculata (L) Walp) and their putative parents.*  
Betreuer: Prof. I. Fawole
- January -05 Beginn der Doktorarbeit am Max Planck Institut für Züchtungsforschung, Köln, in der Arbeitsgruppe von Dr. Seth J. Davis