CONVERTING TOXIC CYANIDE INTO VALUABLE AMINOACIDS: ISOLATION OF β-CYANOALANINE SYNTHASE IN CASSAVA (*Manihot esculenta* Crantz) AND EVALUATION OF ITS PHYSIOLOGICAL ROLE

By

Josean David Marrero Degró

A Thesis Submitted in Partial
Fulfillment of the Requirements for the
Degree of

MASTER OF SCIENCE

in

BIOLOGY

UNIVERSITY OF PUERTO RICO
MAYAGÜEZ CAMPUS
2009

Approved by:

_____	_____
Fernando Souto, Ph.D. Member, Graduate Committee	Date
_____	_____
Carlos Ríos Velázquez, Ph.D. Member, Graduate Committee	Date
_____	_____
Dimuth Siritunga, Ph.D. President, Graduate Committee	Date
_____	_____
Linda W. Beaver, Ph.D. Representative of Graduate Studies	Date
_____	_____
Nanette Diffoot Carlo, Ph.D. Chairman, Biology Department	Date

ABSTRACT

Cassava is a perennial shrub cultivated in the tropical and sub-tropical regions of the world and is characterized by its ability to develop secondary roots of high starch storing capacity. This root crop shows certain advantages such as the capability of growing in marginal conditions of drought periods and poor and acidic soils. It also shows resistance to certain herbivore pests and can persist in the soil for 8-24 months without decaying. Certain African regions contribute to the main inputs in global cassava production, where it is mainly used as a food source.

Cassava is known to accumulate linamarin in all of its tissues. The enzymatic action of linamarase and HNL cause cyanide release from linamarin mainly after tissue injury, serving as chemical defense molecule against herbivores. In addition, it has been considered the implication of cyanoglycosides as storing and transportable forms of reduced nitrogen. β-cyanoalanine synthase (β-CAS) provides the first step in the cyanide detoxification pathway, allowing its incorporation into the amino acid biosyntheses pathways. This enzyme belongs to the β-substituted alanine synthases (*Bsas*) family of enzymes, being close-related to cysteine synthase (CS). Despite the importance of the β-CAS detoxification pathway in cyanide metabolism in cassava, the enzymes involved in it have not been identified at the molecular level. This project aims to isolate the cassava *Bsas* genes and determine their role in cyanide metabolism.

Three cDNAs were isolated using RACE-PCR, which were first identified as *Bsas*-specific genes harboring sequences of different molecular weight. Sequencing information led to the conclusion that two of these contained identical sequences, each one of which encodes a short and a long version of the same cDNA. Sequence analysis tools allowed the prediction of the amino acid sequences encoded by these genes, being MANes;BsasA the protein encoded by the two identical sequences and MANes;BsasB the remaining gene. Sequence cluster analyses grouped both proteins among the *Bsas* enzymes targeted to the mitochondria, the cell compartment where β-CAS is found. However, one of the sub-cellular location estimating tools identified MANes;BsasA as a possible plastidic enzyme, which together with the cytoplasm, comprise the two main organelles for cysteine biosynthesis.

Cloning and over-expression of these genes in a bacterium of *Bsas* mutant background provided concrete evidence of the *in vitro* kinetic properties of these genes. MANes;BsasB showed higher β-CAS activity levels than MANes;BsasA and ARAth;Bsas3;1, which was used as a β-CAS positive control. In the other hand, MANes;BsasA showed remarkably higher CS activity levels than ARAth;Bsas3;1, but more importantly than MANes;BsasB. This information clarifies that MANes;BsasA truly encodes a CS isoform, whose shorter version lacking the signal peptide might encode the cytoplasmic isoform. The longer cDNA could encode an enzyme of plastidic or mitochondrial sub-cellular location, since both compartments have been shown to have both enzymes. In the other hand MANes;BsasB encodes a mitochondrial β-CAS enzyme, since no other evidence has been shown that a true β-CAS protein is targeted outside this compartment.

An attempt of transformation was done by cloning these genes in a plant binary vector for transformation of cassava somatic embryos. A total of 2,806 and 2,513 explants were transformed harboring the ARAth;Bsas3;1 and MANes;BsasA genes, respectively, producing a total of 80 and 212 lines that endured antibiotic selection. However, no line showed a repetitive pattern of PCR amplification, when screening these lines using different primer combinations. Moreover, no line showed increased β-CAS activity. Apparently the selection protocol employed did not guaranteed a proper selection of transgenic lines with the T-DNA integrated into its genome.

RESUMEN

La yuca es un arbusto perenne cultivado en regiones tropicales y tiene la capacidad para desarrollar raíces secundarias con alta capacidad de almacenar almidón. Este cultivo muestra ciertas ventajas como la capacidad de crecer en condiciones marginales (sequía y suelos pobres y ácidos). También muestra resistencia a ciertas plagas de herbívoros, mientras que es capaz de persistir en el suelo por un período de 8-24 meses sin que decaiga su calidad. Ciertas regiones africanas contribuyen a los principales insumos en la producción mundial de yuca, donde se utiliza principalmente como fuente de alimento.

La yuca acumula linamarina en todos sus tejidos. La acción enzimática de linamarasa y HNL causan la liberación de cianuro a partir de linamarina, sobre todo después de la lesión tisular, actuando así como molécula de defensa química contra los herbívoros. Además, también se ha considerado la implicación de los cianoglicósidos como formas de almacenamiento y transporte de nitrógeno reducido. La enzima sintetasa de β-cyanoalanina (β-CAS) provee la catálisis del primer paso en la ruta de desintoxicación del cianuro, lo que permite su incorporación a las vías de síntesis de ciertos amino ácidos. Esta enzima pertenece a la familia de sintetasas de alaninas β-sustituidas (*Bsas*) en la cual también se encuentra cisteína sintasa (CS), una enzima a la cual β-CAS está próximamente relacionada. A pesar de la importancia de estas enzimas en el metabolismo de cianuro en yuca, las mismas no se han identificado a nivel molecular. El presente proyecto tiene por objeto aislar los genes que codifican enzimas de la familia *Bsas* en yuca y determinar el rol que juega cada una en el metabolismo de cianuro.

Se lograron aislar tres cDNAs utilizando RACE-PCR, dos de estos conteniendo secuencias idénticas (MANes;BsasA) y otro gen con algunas diferencias en la secuencia de nucleótidos. Análisis de cluster utilizando estas secuencias agruparon a ambas proteínas entre las enzimas *Bsas* específicas del mitocondrio. Sin embargo, uno de los programas utilizados para la estimación de la localización sub-celular demostró que MANes;BsasA podría ser posiblemente una enzima localizada en los cloroplastos.

La sobre-expresión de estos genes en una bacteria mutante, negativa en la expresión de genes de la familia *Bsas*, proveyó evidencia cinética de estas proteínas. MANes;BsasB mostró mayores niveles de actividad β-CAS al comparar con los niveles mostrados por MANes;BsasA y ARAth;Bsas3;1, el cual fue utilizado como un control positivo para β-CAS. Por otra parte, MANes;BsasA mostró un nivel de catalización de la reacción de CS notablemente superior a ARAth,Bsas3,1, pero una diferencia más drástica fue observada al comparar este valor con el que obtuvo MANes;BsasB. Esta información aclara que MANes;BsasA realmente codifica una isoforma de CS, cuya versión más corta carece del péptido señal, codificando la isoforma citoplasmática. Por otra parte, MANes;BsasB codifia la enzima β-CAS

Se hizo un intento transformación de embriones somáticos de yuca con éstos genes. Un total de 2,806 y 2,513 explantes se transformaron con constructos genéticos albergando los genes ARAth;Bsas3;1 y MANes;BsasA, respectivamente. Estos explantes produjeron un total de 80 y 212 líneas respectivamente, que soportaron selección con antibiótico. Sin embargo, ninguna línea mostró un patrón repetitivo de amplificación de PCR al amplificar utilizando diferentes combinaciones de cebadores específicos para cada constructo. Además, ninguna línea mostró aumento de la actividad enzimática de β-CAS en sus tejidos. Al parecer, el protocolo de selección empleados no garantizó una adecuada selección de las líneas transgénicas conteniendo el T-DNA integrado en su genoma.

DEDICATION

To Karen, Kamyla and Sebastián.

“La mentira es siempre arma de cobardes”.

“To lie is always a weapon of cowards”.

San Josemaría Escrivá de Balaguer (Forja n. 905).

ACKNOWLEDGEMENTS

First of all I want to say that my perception of life as a Christian has driven me to be what I am, so I want to especially give thanks to God for all the good things He has given me during all my life and also thanks for all the things I have judged as not so good, because if they happen to occur is due to something.

Also I want to give thanks to my family, since I am their reflection. To my parents Jose Antonio, my mother Carmen Rosa and to Pochy and Lester (my brothers) and Marie Carmen, Emily and Rocío del Carmen (my sisters) for their example as how to be responsible in life without taking advantage of others especially if they are in a disadvantageous position. All of you have a piece of my heart. Especially I dedicate this work to my sister Karen Gisselle who died in a car accident just when I was in the middle of the writing of my thesis. Thank you for your example as a hard working human being and struggling against those things sometimes you don't understand. I also dedicate this work to her daughter and son Kamyla and Sebastián.

I give special thanks to my housemates at the Residencia Universitaria Almenares. To Ricky Martínez, Enrique Saavedra, Chú Muñiz and Padre Andrés Eiroa, since they have become into my family, and to all the people that has been in contact with us, helping in our activities aiming the professional, human and spiritual education of university students. Thanks for believing in us.

Regarding this work I want to give thanks to my graduate committee. To Dimuth Siritunga, for giving me this opportunity to develop high quality research; to Carlos Ríos Velázquez and Fernando Souto, since each one of you have contributed to my formation as a scientist. To Linda Beaver who was the representative of Graduate School during the oral exam. Thanks also to other faculty members and employees of the Department of Biology, who helped in providing equipment or any other supplies. Special thanks to María Mendez and Mary Jiménez for always remining me everything concerning paperwork.

Thanks to all the fellow members of our laboratory who were always in close contact during the development of this research. To Morgan Echeverry and Ana María Correa for their support as scientists and friends. To the undergraduate students that helped in this project but especially to Joan Marcano, who worked with me on the characterization of the enzymes and for providing nice music to the lab. To Milly Montero, Andrés Salcedo, Lorraine Rodríguez, Barbara Sánchez, Amaury Ortiz and other members of the lab during all these years.

Last but not least to all fellow Grad Students for their friendship and support, or for just been there when I just needed to talk or have a good time at the lobby.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION:
LITERATURE REVIEW AND OBJECTIVES

1 Introduction

1.1 Cassava (*Manihot esculenta* Crantz)

1.1.1 The biology of cassava, taxonomy and origin

Cassava (*Manihot esculenta* ssp. *esculenta* Crantz) is a perennial shrub cultivated in the tropics and sub-tropics, where it can grow to about 1-5 meters tall and is characterized for having lobed leaves and starchy roots (figure 1.1). Cassava is propagated using vegetative stem cuttings and although the plant can develop seeds, this type of propagation is preferred for breeding purposes. This sub-species does not exist as a wild plant and it has been proposed to be the product of human domestication, being close related to the wild sub-species *M. esculenta* ssp. *flabellifolia* and *M. esculenta* ssp. *peruviana* within the *Euphorbiaceae* family (reviewed by Allem, 2002). Cassava cultivation is thought to have started 5,000-7,000 years BC in South America, from where Portuguese traders took it to Africa and Asia during the XVI century. Differences between cassava and its wild relatives rely mainly on: i) the ability of the crop to develop secondary roots from stem cuttings at a faster rate; ii) the formation of tuberous roots; iii) being less toxic for food and animal feeding; and iv) having shorter cultivation periods (reviewed by Allem, 2002).

Figure 1.1: The cassava plant. A, one month old cassava growing in a pot (approximately 0.5 m tall); B, cassava starchy roots of a plant grown in the field; C, field grown cassava (approximately four month old).

1.1.2 Advantages of cassava as a crop

Cassava is characterized by its ability to develop secondary roots of high starch storing capacity. These tuberous roots develop during the first months of the plant as fibrous roots, which initially grow longitudinally, providing root length and anchorage. Later on these roots start radial growth and thickening due to starch accumulation, developing into tuberous roots (Ceballos and De la Cruz, 2002). Root length varies between cultivars, reaching up to 2.5 m in length, depending on consumer and farmer preferences. In its ability to accumulate carbohydrates lies the economical value of cassava roots. Cassava roots comprise 90% of their dry weight as starch, stored mainly as granules in the amyloplasts of root parenchyma (Buitrago, 1990).

Cassava also shows the capability of growing in marginal conditions, such as drought and poor and acidic soils. It has been shown that its high drought tolerance is due to stomata closing, resulting in a low stomatal conductance that impedes water loss under drought conditions (Mejía de Tafur, 2002). Several studies have reported that yield is only slightly affected at drought conditions (reviewed by Mejía de Tafur, 2002). However, López (2002) states that even though cassava is known for being a low input crop, previous studies show that irrigation practices increase yield up to 60%.

Another advantage of cassava is that it has developed natural barriers against herbivores. Physical mechanisms, such as the presence of foliar pilosity, insect induced calli (phloem cicatrisation) and the cell wall, interfere with the establishment of aphids in developing leaf meristems (Calatayud and Múnera, 2002; Calatayud et al., 1996). Chemical mechanisms against arthropods include secondary metabolites, such as cyanogenic glycosides that provide resistance to root eating insects (Belloti et al., 1999), and glycosilated flavonoids such as rutin, which

together with the aforementioned foliar pilli, constitute the main constraining agents against cassava mealy bug (*Phenacoccus manihoti*) establishment (reviewed by Calatayud and Múnera, 2002).

In addition, cassava can persist in the soil for 8-24 months without decaying, providing product security during drought periods, serving as a good source of staple food to subsistence farmers (Nweke et al., 2002). This feature also helps to select the harvest time, avoiding the rapid physiological post-harvest deterioration of cassava roots, which together with its low root protein content (1-2% dry weigh, compared to 20-25% dry weigh in leaves) constitute two of the main disadvantages of the crop.

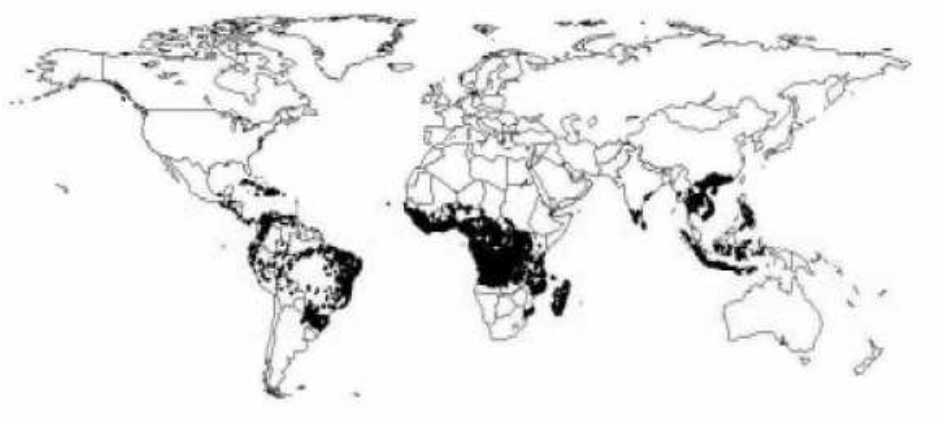

Figure 1.2: Major producing regions of cassava in the world. Source: "Centro Internacional de Agricultura Tropical" (CIAT).

1.1.3 Global economical impact

Cassava is cultivated in tropical America, Asia and more extensively in certain African regions, extending from Madagascar in the Southeast to Senegal in the Northwest (figure 1.2), forming what has been sometimes called the cassava belt (Bokanga, 1994). Ghana, Nigeria, the Democratic Republic of Congo and Tanzania are countries that contribute to main outputs in cassava harvesting (FAO/IFAD, 2000). Annually approximately 165 million tons of cassava

roots are harvested worldwide, being forth among the main sources of calories in the tropics, after rice, wheat and maize (Ceballos, 2002). According to statistics published by the Food and Agriculture Organization (FAO) of the United Nations, cassava production increased by a rate of about 2.50% per year during the period from 1965 to 1995 (FAO/IFAD, 2000; Ceballos, 2002), apparently mainly due to an increase in its extensive cultivation in Africa. These values compare to trends in other crops such as wheat (4.32%), potato (4.00%), maize (3.94%), yams (3.90%), rice (2.85%) and sweet potato (1.07%). Moreover, 60% of cassava cultivated lands worldwide belong to this continent, providing 50% of total cassava production (Ceballos, 2002; figure 1.3).

Cassava has many uses in different regions of the world, but in sub-Saharan Africa it serves mainly as a food source, preventing famine during drought periods. The clear increase in cassava production worldwide can be associated to an increase in its use for food (FAO/IFAD, 2000; figure 1.3), serving as a food reserve when other basic crops fail. Even though countries such as Ghana have develop strategies for the exploitation of cassava roots at industrial level, in Africa it is traditionally cultivated by small farmers, who mainly culture it in small farms and backyards, last in rotation with other crops before soil fallow. This practice provides the crop to benefit from residual fertilizer, but could probably account for yields far below optimum (FAO/IFAD, 2000).

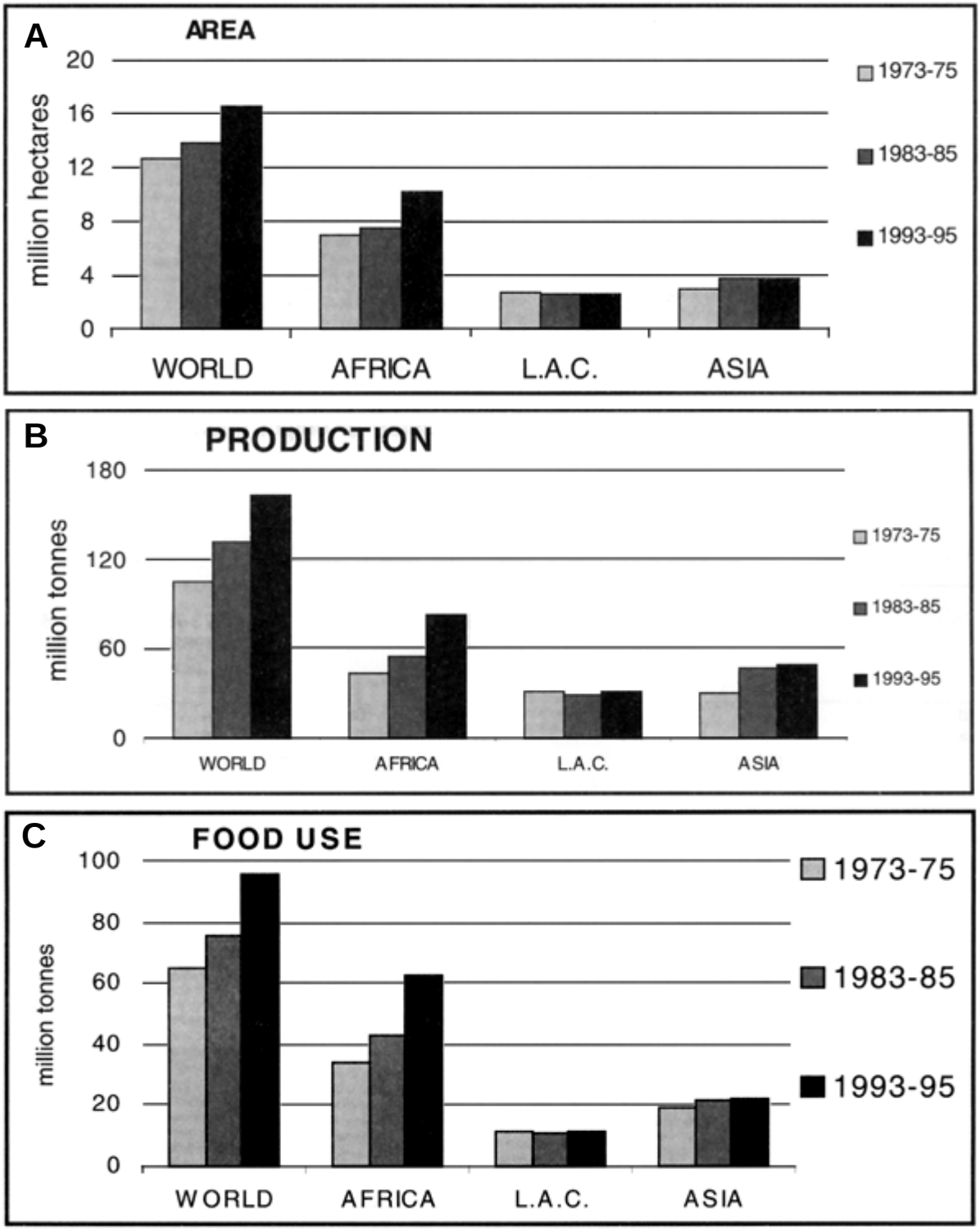

Figure 1.3: Trends in cassava production according to the FAO (FAO/IFAD, 2000). A, area dedicated to cassava cultivation; B, cassava production (million tons of roots); C, cassava used for human consumption (million tons of roots); L.A.C. stand for Latin America and the Caribbean.

The processed cassava roots are eaten in various forms: boiled, baked, fried and as flour. In Western Africa, cassava roots are commonly eaten as gari, a cassava flour paste made by peeling and grating the roots until a mash is made, followed by fermentation for 1-2 days in bags that allow its water loss and drying by pressing the bags. Further heating allows complete drying and roasting, ending in the formation of the granular gari, which can be further ground to fine flour. In Eastern Africa countries, cassava is mainly fried. Some African cultures also consume cassava leaves, which provide a good source of protein and vitamins (Bokanga, 1994).

1.2 Cyanide in cassava

1.2.1 Linamarin synthesis

As mentioned above, cassava is known to accumulate cyanogenic glycosides in all of its tissues except seeds. These are a group of nitrile-containing secondary compounds in plants that yield free cyanide upon their enzymatic breakdown, in a series of reactions collectively called cyanogenesis. This process can also occur in plants due to cyanolipid hydrolysis and as a by-product of ethylene synthesis (Poulton, 1990), but its release due to cyanogenic glycoside breakdown generally provide toxic quantities for human consumption in this tropical crop.

Linamarin is the most abundant cyanogenic glycoside in cassava, accounting for 95% of total cyanogens in the plant (Cock, 1985). Extensive research has been done in unraveling what are the biosynthesis pathways involved in linamarin synthesis, which is known to be synthesized from valine, in reactions that occur in the vacuoles of plant cells (figure 1.4; McMahon et al., 1995; Koch et al., 1992). Initially, two sequential hydroxylation reactions of valine occur, catalyzed by the enzymes CYP79D1 and CYP79D2, which are cytochromes P-450 (Andersen et al., 2000) that allow the formation of the two subsequent intermediates *N*-hydroxy valine and 2-methyl-propanal oxime (figure 1.4). Further addition of oxygen by a putative CYP71E takes

place, producing the non-stable intermediate acetone cyanohydrin. CYP71E has been proposed to be present in cassava by Andersen et al. (2000), in homology to dhurrin synthesis in sorghum. Acetone cyanohydrin is finally glycosylated by a vacuolar uridine 5'-diphosphoglucose (UDPG)-glucosyl transferase, forming linamarin (McMahon et al., 1995; figure 1.4).

The vacuolar localization of these reactions has been suggested by evidence showing CYP79D1/D2 to be targeted to the tonoplast membrane (Koch et al., 1992). In addition, the quantities of linamarin in isolated vacuolar fractions account for that present in intact cells (White et al., 1994). Furthermore, the localization of cyanogenic glycosides in vacuoles has been also suggested in other plants such as sorghum and rubber tree (reviewed by McMahon et al., 1995).

It is generally accepted that linamarin is synthesized in leaves and transported to roots by mechanisms that to date are not clearly understood (McMahon et al., 1995). The two linamarin metabolizing enzymes linamarase and hydroxynitrile lyase, which trigger cyanide release (figure 1.4), are located in the cell wall, requiring a symplastic transport of linamarin, or in default, a partial modification of linamarin that prevents its undesired metabolization, if an apoplastic pathway for its translocation is preferred. Various works have suggested linustatin, the glycosylated form of linamarin, as the translocable form, since its presence in other cyanogenic plants has been proved (Koch et al., 1992; Selmar et al., 1988; Selmar, 1994). However, the levels of linustatin in cassava tissues are apparently too low to account for the only source of linamarin in roots; moreover, certain studies have documented that roots are capable of certain levels of linamarin synthesis (White et al., 1994; McMahon et al., 1995).

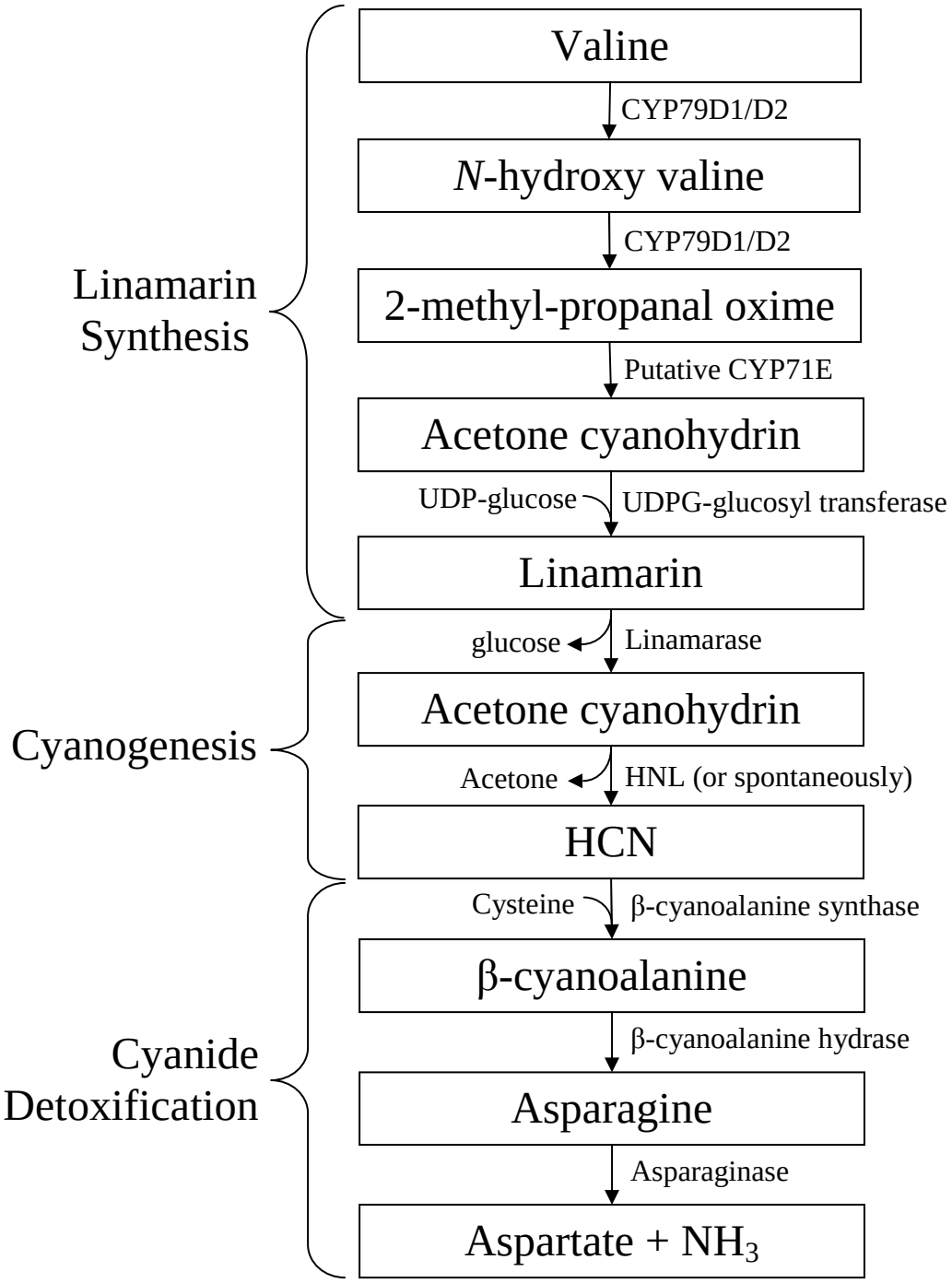

Figure 1.4: Metabolic pathways of linamarin synthesis, cyanogenesis and cyanide detoxification in cassava.

Transgenic approaches demonstrated the leaf specific silencing of the genes encoding CYP79D1 and CYP79D2 involved in the first steps of linamarin synthesis, resulting in a significant reduction in leaf and root linamarin content, while a root specific silencing meant no change in linamarin accumulation in any tissue (Siritunga and Sayre, 2003, 2004). This information clarifies that linamarin is at least mainly synthesized in the leaves, which serve as a main source of linamarin to other tissues.

1.2.2 Cyanogenesis

Cyanogenesis involves a series of reactions that yield free cyanide. The first reaction is the deglycosilation of linamarin to produce acetone cyanohydrin and glucose by linamarase (figure 1.4; Mkpong et al., 1990), a very stable β-glucosidase enzyme. This step is broadly accepted to be limiting for cyanogenesis to occur, since the produced aglycone acetone cyanohydrin can be subjected to spontaneous decomposition to cyanide and acetone at pH > 5.0 and/or temperatures greater than 35ºC (figure 1.4; White et al., 1998). Linamarase has a narrow range of substrate specificity compared to other β-glucosidases, as it is capable of hydrolyzing linamarin and other non-physiological monoglycosidic substrates, but not the diglycosides amygdalin and linustatin (McMahon et al., 1995), being one of the proofs of linustatin serving as the linamarin translocated form.

Cyanide can also be released enzymatically from acetone cyanohydrin by hydroxynitrile lyase (HNL; figure 1.4). This is the critical step that releases cyanide, being acetone an additional product. HNL, like linamarase, is very stable, showing no reduction in its activity after its storing for one month at 4ºC (McMahon et al., 1995), a feature particularly common to extracellular enzymes. Moreover, the highest HNL activity levels have been found in protein extracts prepared from the apoplast, suggesting a localization of the enzyme at the outer inter-cellular

space (White et al., 1994), similar to linamarase (Mkpong et al., 1990; Selmar et al., 1988). This information has driven the commonly accepted statement that these two enzymes are localized in the cell wall (McMahon et al., 1995).

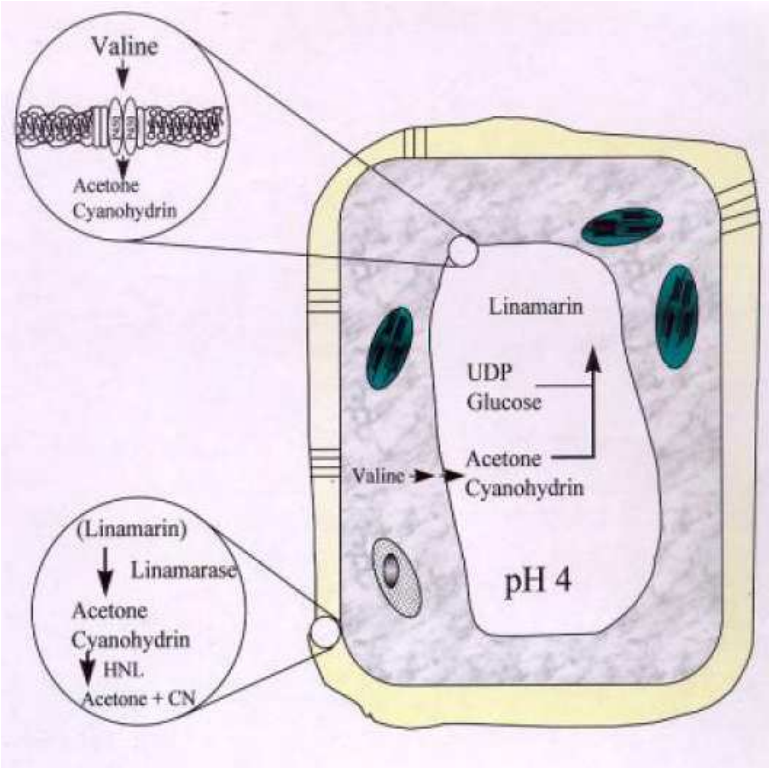

Figure 1.5: Sub-cellular distribution of linamarin and the enzymes involved in its synthesis and cyanogenesis.

1.2.3 Biological role of linamarin in cassava

Cyanogenic glycosides are thought to have two main functions: as chemical defense molecules against herbivores, and as transportable and storage forms of reduced nitrogen. The idea of implicating cyanogens in plant defense has been widely supported (Bellotti et al., 1999; Calatayud and Múnera, 2002; Kakes, 1994; Morant et al., 2007; Poulton, 1990; Vetter, 2000; Zagrobelny et al., 2004), since cyanogenesis is known to be triggered by tissue injury. The different sub-cellular localizations of both the cyanogenic glycosides and their catabolizing enzymes evidences this statement. In cassava, linamarin is synthesized and stored in vacuoles,

while linamarase and HNL are targeted to the cell wall (Figure 1.5). In this case, cyanogenesis can only be stimulated when the vacuolar contents are allowed outside the cell, which is what occurs when the plant tissue is disrupted by mechanical means such as herbivore action (Poulton, 1990; Vetter, 2000). Moreover, the entire pathway for dhurrin synthesis, the main cyanoglycoside present in sorghum, has been engineered in the non-cyanogenic plant *Arabidopsis thaliana*, resulting in a 4% (dry weigh) dhurrin accumulation which triggered 80% reduction in the consumption of the plant by the flea beetle *Phyllotreta nemorum*, which naturally feeds from this plant (Tattersall et al., 2001; Morant et al., 2007).

Various authors have also considered the implication of cyanoglycosides as storing and transportable forms of reduced nitrogen. Nartey (1969) published the first evidence when reported that two week old cassava seedlings, when fed with ^{14}C-radiolabeled cyanide, accumulated 49% of the label in asparagine. Since a direct reaction leading to the formation of asparagine from cyanide and serine or cysteine seemed unlikely, this author proposed the formation of β-cyanoalanine as an intermediate molecule that is rapidly converted to asparagine, in accordance with other works that documented the formation of this non-protein amino acid specifically from cyanide (figure 1.4; Nartey, 1969; Dunhill and Fowden, 1965; Floss et al., 1965). Later on, the formation of asparagine from β-cyanoalanine was confirmed to occur in a reaction catalyzed by β-cyanoalanine hydrolase (Castric et al., 1973). Asparagine, in addition to serve as a protein precursor, is a key compound in nitrogen storage and transport, due to its stability and high nitrogen/carbon ratio (Taiz and Zeiger, 2002a). Asparagine is further metabolized by asparaginase, which catalyzes its de-amination, converting it to aspartate, releasing ammonia. The latter is the utilizable form of nitrogen for amino acid synthesis, being incorporated into the glutamine and glutamate biosynthesis pathways. Nartey (1969) observed

that a significant part of the label (6%) was also found in aspartate, confirming the possible role of cyanogens as alternate sources of reduced nitrogen, other than nitrate and nitrite reduction to ammonia. In the other hand, asparaginase activity levels were found to be lower in cassava roots than in leaves, in contrast to the other enzymes involved in asparagine synthesis β-cyanoalanine synthase and β-cyanoalanine hydrolase, which showed higher activity in roots, suggesting that asparagine metabolization occurs mainly in leaves (Elias et al., 1997b)

Furthermore, Nartey (1969) also observed that the levels of asparagine increased after seed germination. This was also observed by Selmar and co-workers (1988) whom evidenced the metabolization of linamarin in germinating rubber tree seedlings. Apparently the accumulation of asparagine was accompanied by a decrease in linamarin content in the endosperm and an increase in linustatin content. At the same time, activities of the linustatin diglucosidase and β-cyanoalanine synthase increased in seedlings, indicating that linamarin was somehow converted to linustatin and translocated from the maternal endosperm to the germinating embryo, following its metabolism to asparagine, through the formation of β-cyanoalanine. Later on, Poulton (1990) stated in his review about plant cyanogenic glycosides that the utilization of linamarin in cassava as an alternate nitrogen source seemed plausible.

Recent experiments aimed at decreasing cyanogen content in cassava tissues using transgenic approaches provided further evidence of their role in nitrogen transport and supply (Siritunga and Sayre, 2004; Siritunga and Sayre, 2007). Transgenic cassava plants developed to specifically silence the CYP79D1/D2 gene expression (see section 1.2.1) showed no accumulation of linamarin and these plants apparently failed to produce roots in medium devoid of ammonium.

1.2.4 Cyanide detoxification: β-cyanoalanine synthase (β-CAS)

There are three possible cyanide detoxification pathways in living organisms that target cyanide to be enzymatically metabolized. These are rhodanase, formamide hydrolyase (FHL) and β-cyanoalanine synthase (β-CAS). Rhodanase is more likely to be present in insects and at a very low level in higher plants but is not present in cassava (Poulton, 1990; Miller and Conn, 1980). FHL is not present in any cyanide metabolism pathway in plants (Miller and Conn, 1980). The most common enzyme involved in cyanide detoxification in plants is β-CAS, which catalyzes the incorporation of cyanide into L-cysteine to form β-cyanoalanine, releasing sulfide (figure 1.6).

β-CAS was first purified in 1969 from blue lupine seedlings by Hendrickson and Conn and previous evidence showed that protein extracts prepared from this plant, as well as from sorghum, common vetch and *Escherichia coli* were capable of synthesizing β-cyanoalanine from cyanide (Blumenthal et al., 1963; Blumenthal et al., 1968; Dunhill and Fowden, 1965; Floss et al., 1965). The reaction was determined to occur at a faster rate when cysteine and cyanide were used as substrates and not serine as originally suggested (Hendrickson and Conn, 1969).

β-CAS has been included among the β-substituted alanine synthases (*Bsas*) family of enzymes, which share the common feature of catalyzing reactions of β-replacement leading to the biosynthesis of non-protein alanine-derived amino acids (Ikegami and Murakoshi, 1994; figure 1.5). In general, all members of the *Bsas* family of enzymes are capable of using *O*-acetyl-serine as the donor of the alanyl moiety, but are also capable of employing closely related molecules such as L-serine or L-cysteine. Figure 1.6 shows the general reaction catalyzed by all *Bsas* enzymes, in which the formation of an alanine-related amino acid is achieved, having an R-

group bound to the β-carbon of alanine. These proteins are pyridoxal phosphate (PLP) dependent enzymes requiring this molecule as a co-factor for amino acid binding.

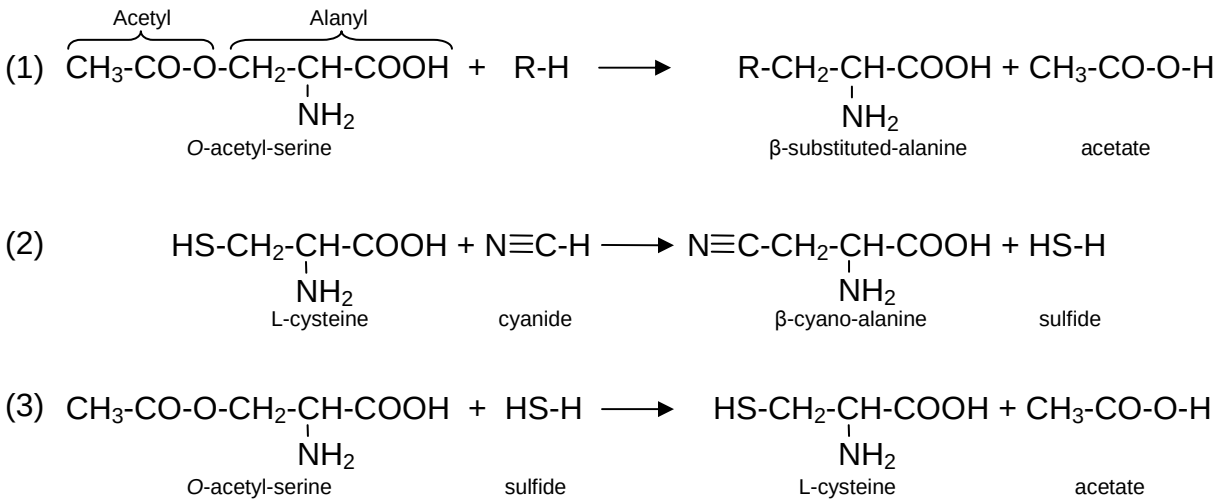

Figure 1.6: Biosynthesis of β-substituted alanines in plants (Ikegami and Murakoshi, 1994). (1) The generalized reaction of the synthesis of β-substituted alanines involves the replacement of organic groups (R-) with *O*-acetate, serving *O*-acetyl-serine as the donor of the alanyl moiety; (2) Reaction catalyzed by β-Cyanoalanine Synthase, using L-cysteine as the alanyl donor molecule and cyanide, resulting in the formation of β-Cyanoalanine and sulfide; (3) Reaction catalyzed by Cysteine Synthase, using *O*-acetyl-serine and sulfide releases L-cysteine and acetate.

Ever since the first report of the isolation of the enzyme, protein fractions showing high β-CAS activity have also been shown capable of synthesizing cysteine, using *O*-acetyl-serine and sulfide as substrates in a reaction more similar to the generalized scheme of the synthesis of β-substituted alanines (figure 1.6; Hendrickson and Conn, 1969). Cysteine Synthase (CS; also called *O*-acetyl-serine(thiol) lyase), another key member of the *Bsas* family of enzymes, mainly catalyzes this reaction which provides the key step in sulfur assimilation into organic compounds (Taiz and Zeiger, 2002b). Due to the high degree of sequence homology and hence structural and functional similarities between β-CAS and CS, it was not until the late 1990's that more concrete explanations arose about the identity of each enzyme.

It has been reported that the synthesis of β-substituted alanines is catalyzed by reaction-specific enzymes, but that CS is capable of their biosynthesis (Ikegami and Murakoshi, 1994).

As expected, CS has been found to be capable of synthesizing β-cyanoalanine (Ikegami et al., 1988, 1989a, 1989b). In addition to the first report of the isolation of β-CAS (Hendrickson and Conn, 1969), various works have documented that β-CAS is capable of catalyzing the cysteine synthesis reaction (Hatzfeld et al., 2000; Maruyama et al., 2000, 2001; Warrillow and Hawkesford, 1998, 2000; Yamaguchi et al., 2000). However, each enzyme is capable of catalyzing its corresponding activity to a higher extent, hence receiving its name by their kinetic properties (Warrillow and Hawkesford, 1998, 2000; Hatzfeld et al., 2000).

The role of β-CAS in cyanide metabolism received special attention when cyanide was identified as a by product of ethylene biosynthesis and that a ubiquitous distribution of β-CAS was apparent (Wurtele et al., 1984; Wurtele et al., 1985). Ethylene is a plant growth regulator primarily involved in senescence and fruit ripening, meaning that its biosynthesis is expected in the vast majority of higher plants. Compared to other *Bsas* enzymes, β-CAS was found in cyanogenic as well as non-cyanogenic plants, a behavior that is also observed also in CS (Ikegami and Murakoshi, 1994).

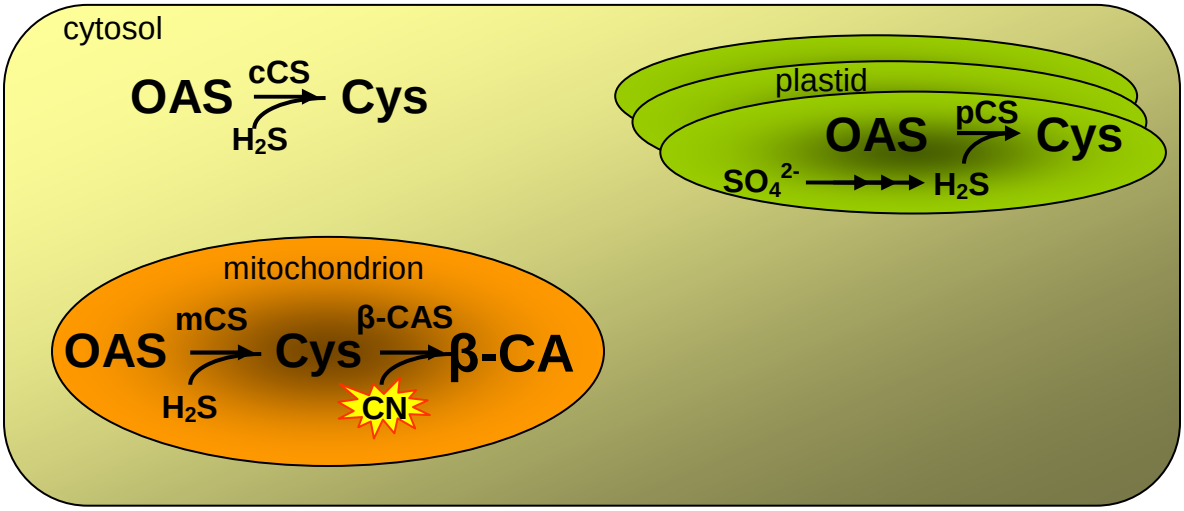

Figure 1.7: Cellular location and physiological roles of the different *Bsas* members in *Arabidopsis thaliana* (Watanabe et al., 2008a). Legend: OAS, *O*-acetyl-serine; Cys, cysteine; β-CA, β-cyanoalanine; cCS, cytosolic CS; pCS, plastidic CS; mCS, mitochondrial CS.

Regarding the dual function of both enzymes in cysteine biosynthesis and cyanide metabolism, more recently it was documented that the actual level at which β-CAS is able to catalyze the CS reaction is much lower than the contrary: i.e. CS is more efficient in catalyzing the β-CAS reaction, than β-CAS catalyzing the CS reaction (Watanabe et al., 2008a). This work focused on elucidating the physiological role of all the genes encoding *Bsas* in *Arabidopsis thaliana* by assaying the ability of knockdown gene mutant plants in both cysteine synthesis and cyanide metabolism.

Another feature that has been used extensively for isoform characterization other than kinetic properties is sub-cellular localization. Mitochondrial fractions have shown to have higher β-CAS activity (Hendrickson and Conn, 1969; Wurtele et al., 1985; Warrillow and Hawkesford, 1998). β-CAS is located in the mitochondria because of its role in cyanide detoxification. Cyanide interferes with the electron transport chain of the respiratory pathway by inhibiting cytochrome oxidase. On the other hand, CS is predominantly found in plastids and the cytosol, but the mitochondria are known to be capable of a certain level of cysteine biosynthesis (figure 1.7; Lunn et al., 1990; Takahashi and Saito, 1996). This evidence took to the assumption that β-CAS has CS activity, though for example in *Arabidopsis thaliana* two mitochondrial *Bsas* isoforms have been found, each one being β-CAS and CS according to their kinetic properties (figure 1.7; Watanabe et al., 2008a). The presence of CS in the mitochondria is justified by serving as a cysteine supply for the β-CAS reaction. In addition, serine acetyl transferase (SAT), the enzyme involved in *O*-acetyl-serine synthesis (needed for cysteine synthesis) was found also in these three compartments (Watanabe et al., 2008b), mainly because CS and SAT apparently work together in sulfur assimilation by structurally interacting in a macromolecular complex that

assures transferring biosynthetic intermediates, resulting in what has been called metabolic channeling (figure 1.7; Berkowitz et al., 2002; Winkel, 2004).

Apparently higher plants have evolved β-CAS as a CS mitochondrion targeted isoform, that has a higher affinity for cyanide, rather than for the original substrate H_2S. According to sequence alignments, the major differences in sequence between all the *Bsas* isoforms present in a single plant lies in a signal peptide present in β-CAS or in plastidic CS, which have been shown to be a mitochondrion or a chloroplastic targeting sequence (Maruyama et al., 2001). The cytosolic CS does not have any signal peptide in its amino terminal.

1.3 Cyanide-related health disorders

Linamarin is the most abundant cyanogenic glycoside in cassava, accounting for more than 95% of total cyanogens in the plant (Poulton, 1990). It has been reported that linamarin levels is a genotype dependent characteristic, varying between cultivars. There is no acyanogenic cassava clone and only one member of the genus has been found not to contain cyanogens (McMahon et al., 1995). Even though the biosynthesis of linamarin occurs in leaves, where the highest amount has been reported, linamarin levels in cassava roots (10-500 mg CN equivalents/kg dry weight) are higher than the maximum recommended cyanide levels (10 mg CN equivalents/kg dry weight) in foods established by the FAO.

Cyanide is toxic to most living organisms mainly by the inhibition of cytochrome oxidase (complex IV), a mitochondrial enzyme involved in oxygen reduction in the respiratory electron transport chain. In addition it can cause the inhibition of plastocyanin reduction in photosynthesis and the inhibition of catalase. Linamarin enters the human body, where its metabolism to actetone cyanohydrin and cyanide mainly by bacteria in the intestine causes toxicity. Additionally, linamarin can be allowed to reach the brain by a glucose transporter. In an *in vitro*

study using neural cell lines, linamarin caused cell death to considerably high levels (Sreeja et al., 2003).

In Africa, certain health disorders and even death have been associated with poorly processed cassava (Aregheore and Agunbiade, 1991). Apparently, bitter cassava varieties are preferred mainly due to product security, since it could serve in protection against unplanned harvest by family members, theft and animal spoilage. Women in Malawi have mentioned these as the main reasons for growing bitter cassava (Chiwona-Karktun et al., 2000). The preparation of cassava as the granular gari and similar practices involving its fermentation and further sun-drying and/or baking allows effective levels of cyanide volatilization (Padmaja, 1995). These are practices broadly adopted in sub-Saharan Africa, mainly because farmers know about the high toxicity levels of the cassava cultivars employed in their food (Chiwona-Karktun et al., 2000). The level of harm caused by these disorders depends on the level and frequency of cyanogens exposure and the state of nutrition of the consumer. Symptoms are presented at a higher level in nutritionally compromised individuals who have a low protein intake in their diets, especially of sulfur containing amino acids, which are key molecules involved in cyanide assimilation into thiocyanate, which allows a certain level of cyanide detoxification (Tylleskar et al., 1992).

Western African cultures show a high incidence of tropical ataxic neuropathy (TAN), caused by long-term consumption of cassava-based diets with relatively low cyanogen content (Nhassico et al., 2008). Nigerian communities consume cassava daily, but the processing of cassava to gari allows the release of most of the cyanide, thus not resulting in chronic cyanide exposure (Nhassico et al., 2008; Oluwole et al., 2000; Oluwole et al., 2002). The prevalence of TAN in Western African people is higher amongst seniors of 60 years or older, and also more frequent in women, possibly because of their cultural role in cassava culturing and processing,

thus being more prone to cyanide poisoning (Nhassico et al., 2008; Oluwole et al., 2000). TAN is a progressive disease, causing unsteady walking, loss of vision and sensation in hands, deafness and weakness (Nhassico et al., 2008).

Exposure of relatively high toxic cyanide diets is revealed in the paralytic disorder Konzo (Sreeja et al., 2003; Tylleskar et al., 1992). Konzo is an upper neuron disease of sudden onset that causes irreversible leg paralysis, especially amongst children and young women, (Howlett, 1994). South, central and eastern African countries have a higher incidence of Konzo, where larger epidemics and outbreaks have been reported during drought periods, since this abiotic condition causes an increase in the total cyanogens present in cassava roots (Nhassico et al., 2008). Since cyanide intoxication depends on the weight of the individual, it affects mainly children, who also present a deficient and unbalanced diet, consuming daily up to 900 g of cassava derived foods, compared to Western African regions where cassava based foods are limited to 150-200 g daily within children (Nhassico et al., 2008).

1.4 Cassava genetic improvement

1.4.1 Main goals and methods in cassava improvement

During the XXth century, plant breeding has contributed to major increases in crop yield and productivity. Cassava genetic improvement started later than in other crops of agronomical importance such as wheat and maize, whose economical exploitation depends mainly on first world countries where the private sector of the seed industry finds an attractive market. It was not until the second half of the 1970's that the "Centro Internacional de Agricultura Tropical" (CIAT) in Cali, Colombia and the International Institute of Tropical Agriculture (IITA) in Ibadan, Nigeria developed the first scientific programs for gathering and improving cassava

germplasm (Ceballos et al., 2004). These initiatives were due to the need to reduce hunger in the third world, rather than based on the economical impact of the crop.

Productivity and stability of production are among the main priorities in cassava breeding. An increase in starch and dry matter quantity is important for the industrial exploitation of cassava (Ceballos et al., 2004). On the other hand, the improvement of cassava lines for human consumption requires the modification of starch composition and other physiological and morphological traits involved in food quality, such as cyanogen, carotene (food color) and protein content, as well as post-harvest physiological deterioration. Apparently the last two traits show limited genetic variability within the known cassava germplasm, thus requiring the employment of close-related species and subspecies within the *Manihot* genus in order to attempt the introgression of desired alleles (Ceballos et al., 2004; Jennings and Iglesias, 2002).

Meanwhile, stability of production involves the improvement of cassava to resist certain biotic and abiotic stresses. Among the major biotic agents constraining cassava production are viral (cassava mosaic disease mainly in Africa and Asia), bacterial (cassava bacterial blight in LAC), fungal and insect pests (reviewed by Ceballos et al., 2004). Certain abiotic factors affecting productivity such as drought and low fertility soils have driven breeding programs to suggest certain characteristics of the plant involved in its adaptation to these environments. These characteristics include leaf longevity, optimum leaf area index and ideal plant architecture (Ceballos et al., 2004).

However, cassava breeding is dificult and time consuming, mainly due to difficulties in synchronization of flowering between parental lines, high frequency of self pollination, scarcity of flowers in the inflorescence due to farmer selection (since vegetative propagation is preferred)

and time required for the seed to mature: it takes one year to obtain seeds from a planned cross (Ceballos et al., 2004). Moreover, cassava is characterized for having high levels of heterozygosity within populations that make difficult to obtain pure homozygous individuals for the desired characters. Due to all these constraints, genetic engineering approaches have been suggested as additional alternatives for the introgression of genes into parental lines of agronomical importance (Taylor et al., 2004a; Taylor et al., 2004b).

Genetic transformation offers the advantage of the incorporation of a single gene of particular interest into cultivars of agronomical importance, without the problems of gene linkage and whole genome sharing that occur when traditional breeding is employed in crop genetic improvement. Nevertheless, the use of these technologies depends mainly on two things: a method for the delivery and integration of exogenous DNA molecules into the plant genome and a method for the *de novo* regeneration of plants in *in vitro* conditions. DNA transfer technologies require the establishment of successful protocols, but apparently this is not the limiting factor in plant transgenesis. On the other hand, not only the development of a regeneration method, but also its compatibility with a selection system has been a matter of intense research among scientists attempting the implementation of these technologies in plants, sometimes remaining far from its acomplishment.

1.4.2 Plant regeneration methods in cassava

To date, embryogenic suspensions developed from friable embryogenic callus (FEC) and germinated somatic embryos (GSE) through cyclic embryogenesis are the two most successful methods for the regeneration of plants in cassava (Zhang and Puonti-Kaerlas, 2004; Fregene and Puonti-Kaerlas, 2002). These two methods are genotype, tissue and developmental stage dependent, thus their employment may vary depending on the ability to control these conditions

in suitable variety candidates. Both methods depend on a callus phase developed from meristematic tissue of somatic origin such as axillary buds, young apical leaves and leaf lobes in a medium containing auxin.

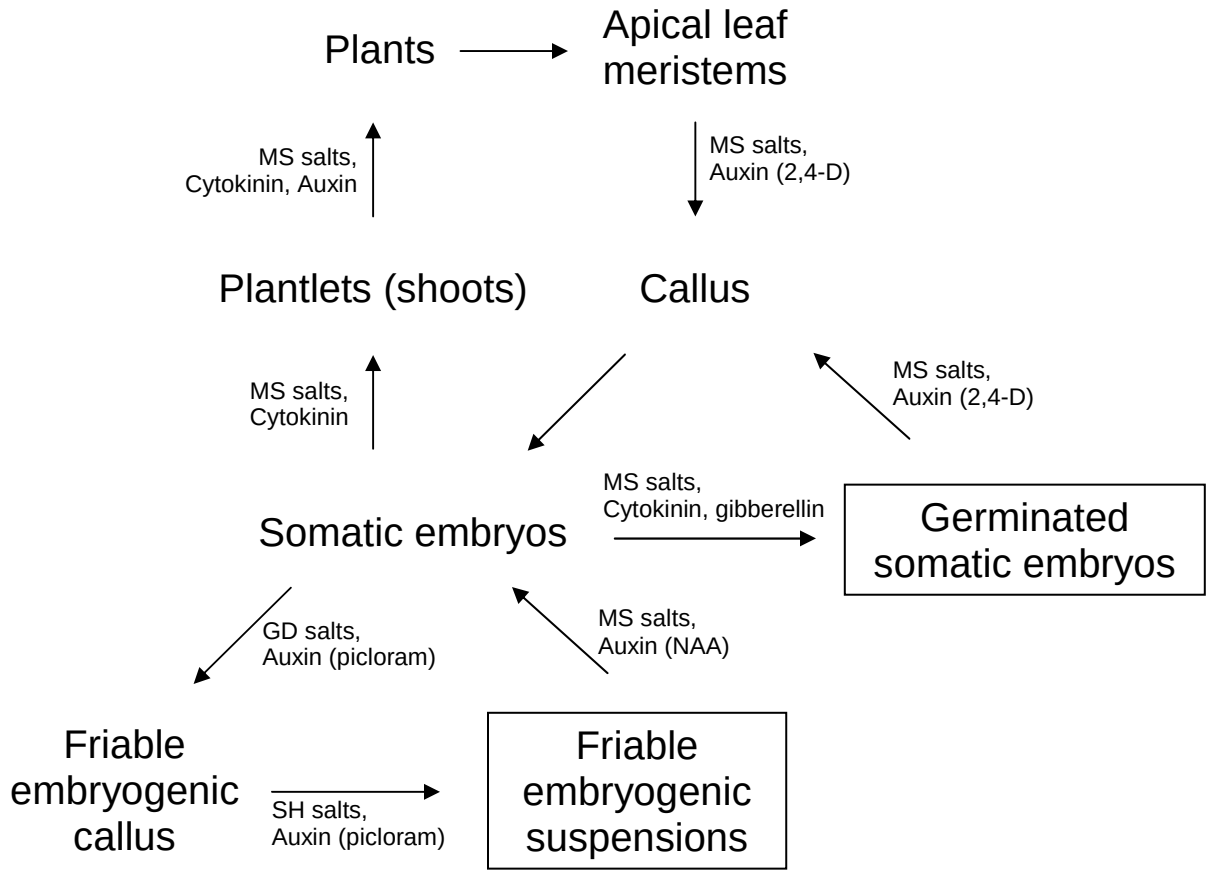

Figure 1.8: Scheme of plant regeneration methods in cassava. The two tissues suitable for transformation are enclosed in rectangles. MS, Murashige and Skoog (1962) basal salt medium; GD, Gresshoff and Doy (1974) basal salt medium; SH, Stamp and Henshaw (1987) basal salt medium; 2,4-D, 2**,**4-Dichlorophenoxyacetic acid; NAA, Naphthalene acetic acid.

Evidence has shown that the development of either FEC or GSEs in cassava depends greatly on the mineral composition of the medium and type of auxin employed (Taylor et al., 1996). Development of somatic embryos is routinely achieved using Murashige and Skoog (MS; 1962) basal salts and vitamins supplemented with the synthetic auxin 2**,**4-Dichlorophenoxyacetic acid (2,4-D), while the employment of picloram and the basal salt mixture of Gresshoff and Doy

(GD; 1974) has been proven to be more effective in the development of FEC. These FEC can be used for further establishment of embryogenic suspensions in picloram-containing Stamp and Henshaw (SH; 1987) liquid medium (figure 1.8; Taylor et al., 1996; Zhang and Puonti-Kaerlas, 2004).

Initial culture of meristematic tissues in MS salts supplemented with 2,4-D (figure 1.8) triggers the development of two types of callus within a period of two to four weeks: white hard callus and brown friable callus. Young meristems form white hard callus more easily than developed leaf lobes do. Eventually, after more than six weeks, white hard callus tends to convert into brown friable especially near the borders, but this tissue does not show embryogenic properties. Primary somatic embryos are of multicellular origin (Fregene and Puonti-Kaerlas, 2002; Zhang and Puonti-Kaerlas, 2004) and typically appear in clumps within white hard callus, though occasionally individual embryos can be observed. Each organized embryogenic unit is characterized by resembling a bi-polar structure with an independent vascular system from the maternal tissue and can be properly germinated into a plantlet through developmental stages resembling that of an embryo of sexual origin when grown in cytokinin-containing medium (figure 1.8; Raemakers et al., 1993; Stamp and Henshaw, 1987).

Somatic embryos can also be stimulated to develop enlarged cotyledonary leaves using gibberellin together with cytokinin. These cotyledons of the GSEs have been found highly embryogenic when employed for further acquisition of secondary somatic embryos using auxins, thus providing a cyclic method for *de novo* regeneration of cassava plantlets in a process commonly known as cyclic somatic embryogenesis (figure 1.8; Raemakers et al., 1993). While primary somatic embryogenesis is known to be genotype dependent, certain reports indicate that

secondary somatic embryogenesis is almost independent and rather simple once the cyclic process is established (Raemakers et al., 1993, 2001).

Some cultivars, such as the Nigerian 60444 are useful for the development of embryogenic suspensions, since it can produce pure friable embryogenic tissue. Taylor et al. (1996) reported the first evidence of the establishment of these suspensions in cassava when somatic embryos (developed as previously explained) were subsequently cultured in GD salts supplemented with picloram. The friable consistency of this callus allows the release of small globular embryo-like structures able to grow as stable embryogenic clustered suspensions that proliferate rapidly in SH liquid medium with picloram (Taylor et al., 1996). Plants have been regenerated from these suspensions when cultured back in MS salts containing the auxin Naphthalene acetic acid (NAA), which stimulates the development of somatic embryos that are capable of shoot organogenesis by further culture using cytokinins (figure 1.8; Taylor et al., 1996; Taylor et al., 2001; Zhang and Puonti-Kaerlas, 2004). However, the ability of friable suspension cultures to develop into somatic embryos does not seem to be as successful as meristem-dependent (primary) and GSEs-dependent (secondary) somatic embryogenesis.

1.4.3 Cassava genetic transformation

Both embryogenic suspensions from FECs and explants from GSEs constitute the main tissues employed in the genetic transformation of cassava (figure 1.8; Fregene and Puonti-Kaerlas, 2002; Taylor et al., 2004a; Zhang and Puonti-Kaerlas, 2004; Ibrahim et al., 2008). Among the DNA transfer technologies used are particle bombardment, *Agrobacterium*-mediated transformation and electroporation, though the former two have been used more extensively (Fregene and Puonti-Kaerlas, 2002; Taylor et al., 2004a). Moreover, Taylor et al. (2004a) commented that there has been a general consensus among laboratories dedicated to cassava

genetic transformation in preferring *Agrobacterium*-mediated transformation mainly due to its capability of producing single insertions vs multiple insertions that occur using biolistics, which is more prone to gene silencing.

In the case of embryogenic suspensions, the method consists of developing fully transgenic suspensions from co-cultivated or bombarded friable suspensions selected using paromomycin or hygromicin (depending on the selective gene within the construct) in liquid medium that promotes the continuous proliferation of only transgenic tissue (Schöpke et al., 1996; Taylor et al., 2001; Raemakers et al., 2001; Zhang and Puonti-Kaerlas, 2004). In addition, visual selection using reporter genes such as the firefly *luc* encoding luciferase have been employed (reviewed in Raemakers et al., 2001). Once the establishment of entirely transgenic suspensions has been achieved, plants can be regenerated by the method described above (see last paragraph of section 1.4.2 and figure 1.8).

The first successful method for the regeneration of fully transgenic plants of cassava with FEC was published in 1996, using negative selection with paromomycin (Schöpke et al., 1996) or luciferase-positive visual selection (Raemakers et al., 1996). In addition, Zhang and Puonti-Kaerlas (2000) showed that positive selection using mannose is possible in cassava FEC when transforming cells with a gene encoding phosphomannose isomerase and growing in a medium containing mannose, since non-transgenic cells are not capable of metabolizing this compound (Wenck and Hansen, 2004).

Transgenic lines of agronomical importance developed using embryogenic suspensions include: increased protein content by expression of an endogenous storage protein (Zhang et al., 2003), virus resistance (Chellapan et al., 2004), decreased starch content and increased monosaccharides by silencing the gene encoding ADPG-pyrophosphorylase (developed by the

Richard Visser's group and reviewed by Taylor et al., 2004a). However, literature is still emerging, encouraging the scientific community about overcoming the difficulties and time consuming efforts employed in the development of embryogenic suspensions (Hankoua et al., 2006; Ibrahim et al., 2008). In addition, this technique has been found to be prone to somaclonal variation resulting in morphological differences between the maternal tissue and the regenerated material, mainly due to the culture regime to which the plant material is subjected (Raemakers et al., 2001).

Explants from GSEs were the first tissue proposed to provide regeneration of transgenic cassava plants and Li et al. (1996) reported for the first time the regeneration of fully transgenic plants with β-glucoronidase activity, using shoot organogenesis. Later on Zhang et al. (2000) followed by transforming also with the *uidA* gene encoding β-glucoronidase. These methods employed negative selection with the *hpt* gene for hygromicin resistance.

CIAT published the development of the first transgenic plants of agronomical importance developed from GSEs, having the *bar* gene integrated into its genome (Sarria et al., 2000). This gene confers resistance to the commercial herbicide Basta. However, no report has been published about the use of these lines in commercial practices. Another successful work using GSEs as the regeneration method reported transgenic cassava plants with modified starch quality (Ihemere et al., 2006). Several works have used GSEs for transforming cassava in order to decrease cyanogens (reviewed in Morant et al., 2007 and Siritunga and Sayre, 2007). However, due to the multicellular origin of somatic embryos, some authors have identified friable embryogenic suspensions, which are known to be of single cell origin, as a better option for cassava transgenesis, since it provides a method devoid of the generation of chimerical individuals (Schreuder et al., 2001; Fregene and Puonti-Kaerlas, 2002; Zhang and Puonti-

Kaerlas, 2004; Ibrahim et al., 2008). Taylor et al. (2004a) proposed that each regeneration method has weak and strong points that have been and will be subject to further optimization.

1.4.4 Improving cassava for decreased cyanogenic content

Iglesias and co-workers reported in 2002 a breeding experiment attempting to gather information about a possible statistical correlation between cyanogen content in cassava roots and linamarase activity by crossing cassava clones of high and low linamarin quantity (Iglesias et al., 2002). No positive correlation was found. The lowest linamarin content found in any clone was 120 mg CN equivalents/ kg dry weight, surpassing ten fold the highest value for human consumption established by FAO (see section 1.3). On the other hand, the development of a cyanogen-free cassava plant has been one of the most important issues in transgenic approaches in this tropical crop (Taylor et al., 2004a). Koch et al. proposed in 1994 that biotechnological approaches can help us to better understand the physiological implications of cyanogens in plants, hence providing possible ways of improving food security. Furthermore, various successful attempts have developed transgenic cassava plants with decreased amounts of cyanide content (Jørgensen et al., 2005; Siritunga and Sayre, 2007).

Siritunga and Sayre (2003, 2004) were able to transform cassava using GSEs and developed Mcol 2215 (a Colombian cultivar) derived lines with *cyp79d1/d2* decreased transcript levels. This first attempt used anti-sense RNA technology to knockout the expression of these genes, which encode cytochrome P450, the enzyme involved in the first-dedicated step in linamarin biosynthesis (see section 1.2.1). These plants had decreased linamarin levels as well, showing up to 99% reduction in linamarin content in roots compared to wild type plants. However, they apparently failed to grow in medium devoid of ammonium or potting soil (Siritunga, 2002), thus showing a physiological dependence on cyanogens. These authors

proposed that linamarin supports in reduced nitrogen supply by the β-CAS detoxification pathway (Siritunga and Sayre, 2004, 2007; see section 1.2.4). Jørgensen et al. (2005) also reported an attempt to reduce cyanogens in cassava by silencing the *cyp79d1/d2* genes in transgenic plants using RNAi silencing technology. This group was also able to significantly reduce the levels of linamarin, and showed that the transgenic plants were able to grow in soil. Nevertheless, they reported no molecular evidence (i.e. southern blot) that these plants were transgenic.

Later on the Sayre group also developed plants with increased HNL levels (Siritunga et al., 2004). This strategy allows the presence of linamarin in the plant's tissues, so that the proposed role as a source of reduced nitrogen is supplied, while accelerating the rate of liberation of cyanide from acetone cyanohydrin, which is then volatilized and eliminated from the plant. HNL specific activity was increased approximately 13-fold in roots and a higher rate of cyanide elimination was reported after tissue homogenization compared to wild type plants. The actual physiological status and agronomical importance of this technology is now being tested in field trials in Isabela, PR (Dimuth Siritunga, personal communication).

1.5 Objectives

Despite the importance of the β-CAS detoxification pathway in cyanide metabolism in cassava, the enzymes involved have not been identified at the molecular level. Elias et al. (1996a) reported the presence of the enzyme in cassava by partially purifying leaf and root protein extracts, but no molecular information about the presence of this and other *Bsas* members has been reported. This information evidences the presence of the enzyme in cassava, but due to the high homology to other *Bsas* members, this published kinetic information does not address information about the specific role of each potential member in cyanide metabolism and any other close related metabolic event such as the synthesis of cysteine and/or other β-substituted alanines. As a general objective, this project aims to determine the role of the cassava β-CAS enzyme in cyanide metabolism. In order to achieve this goal, the genes encoding any of the possible *Bsas* members were isolated using PCR. Further characterization using sequence analysis provided insights about their possible identity. The genes were tested for their kinetic properties in *in vitro* conditions and an *in vivo* analysis of the β-CAS encoding gene was attempted by cassava transformation.

MATERIALS AND METHODS

2 Materials and Methods

2.1 β-CAS cDNA isolation from cassava and copy number determination

2.1.1 Primers design

All primers used for PCR were designed using the Primer3 tool (Rozen and Skaletsky, 1998) and the OligoCalc Oligonucleotide Properties Calculator tool (Kibbe, 2007). A partial cassava β-CAS sequence (accession CK644310; Lopez et al., 2004) and the complete β-CAS mRNA of rubber tree (*Hevea brasiliensis*; accession AY207389) were used for primer design. The rubber tree sequence of β-CAS was used due to its high degree of homology to the partial cassava β-CAS sequence. All primers were synthesized by Integrated DNA Technologies (IDT) and are listed in Table A.1 (see appendix A). Their approximate priming sites within the β-CAS mRNA are shown in Figure 3.1.

2.1.2 RNA isolation

Approximately 50 mg of leaves were harvested from two different *in vitro* cassava varieties: Mcol 2215, which is a sweet (low cyanide level) variety, and TMS 60444, which is a bitter-sweet (mid cyanide level) variety. Leaves were covered with aluminum foil and frozen in liquid nitrogen. Liquid nitrogen was poured in a mortar, allowing it to chill and the tissue was ground with a pestle until the liquid nitrogen evaporated completely. Total RNA was extracted from the tissue using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA) following the instructions of the manufacturer. RNA stability was checked by electrophoresis (0.8 % (w/v) agarose in 1X TAE (40 mM Tris acetate and 1 mM ethylenediaminetetraacetic acid (EDTA); pH 8.5) prepared with 0.1 % (v/v) diethyl pyro-carbonate (DEPC) water). RNA concentration was

measured in a Genesis 10 UV Scanning spectrophotometer (Thermo Electron Corporation) at OD_{260}. The RNA samples were stored at –80ºC until use.

2.1.3 Rapid Amplification of cDNA Ends (RACE)-related reactions and cDNA synthesis

RACE PCR was done using the GeneRacer kit (Invitrogen, San Diego, CA, USA). For 5'RACE, three reactions were done before cDNA synthesis: dephosphorylation, 7-methyl guanosine cap structure removal and RNA oligo ligation to 5' end.

The dephosphorylation of RNA was done in a 10 µl reaction containing 5 µg of total leaf RNA, 1X CIP buffer (50 mM Tris-HCl pH 8.5 and 0.1 mM EDTA), 10 U RnaseOut (Rnase inhibitor) and 10 U Calf Intestinal Phosphatase (CIP), in DEPC water. The reaction was incubated for 1 hour at 50ºC followed by the purification of RNA using the GeneRacer kit, following the instructions of the manufacturer (using a phenol:chloroform extraction protocol).

The reaction for the removal of the 7-methyl guanosine cap structure was done in a 10 µl reaction containing all the RNA purified from the dephosphorylation reaction (<5 µg), 1X TAP buffer (50 mM sodium acetate pH 6.0, 1 mM EDTA, 0.1% β-mercaptoethanol and 0.01% Triton X-100), 10 U RnaseOut and 0.5 U Tobacco Acid Pyrophosphatase (TAP), in DEPC water. The reaction was incubated for 1 hour at 37ºC and after this period, the RNA was purified using the GeneRacer kit, following the instructions of the manufacturer.

The final reaction was the RNA oligo-ligation to the 5' end. In this, the purified RNA from the previous reaction was ligated to 0.25 µg of lyophilized GeneRacer RNA oligo in a 10 µl reaction containing 1X ligase buffer (33 mM Tris-Acetate pH 7.8, 66 mM potassium acetate, 10 mM magnesium acetate and 0.5 mM DTT), 10 mM ATP, 10 U RnaseOut and 5 U of T4 RNA ligase, in DEPC water. The reaction was incubated for 1 hour at 37ºC and after this period, the RNA was purified using the GeneRacer kit, following the instructions of the manufacturer.

Retro-transcription was done using the Omniscript RT kit (Qiagen) in a 20 µl reaction mixture containing 2 µg of total RNA (or ligated RNA), 1X RT buffer (Qiagen), 0.5 mM dNTP mix, 1 µM oligo dT primer (IDT) or 2.5 µM GeneRacer oligo dT primer (Invitrogen; only for 3'RACE), 10 units Recombinant RNasin Ribonuclease Inhibitor (Promega, San Luis Obispo, CA, USA) and 4 units Omniscript Reverse Transcriptase, in DEPC water. The samples were incubated for 2 hours at 37ºC and stored at –20ºC until use.

2.1.4 Polymerase Chain Reaction (PCR)

PCR was carried out in 20-50 µl reaction mixtures containing: 50-500 ng template cDNA, 0.2 mM dNTP mix, 0.4 µM each primer, 1X *Taq* buffer (10 mM Tris-HCl pH 9.0, 50 mM KCl and 1.5 mM $MgCl_2$) and 0.5 U *Taq* DNA polymerase (Fisher).

An initial amplification was done using GeneRacer 5' or 3' primer (Invitrogen) and β-CAS specific primers (see Table A.1 and Figure 3.1) with the following program in a Mastercycler thermal-cycler (Eppendorf, Hamburg, Germany):

Temperature	Time	Cycles
94ºC	2 min	1
94ºC	30 sec	
72ºC	1 min, 30 sec	5
94ºC	30 sec	
70ºC	1 min, 30 sec	5
94ºC	30 sec	
60ºC	30 sec	20
72ºC	1 min, 30 sec	
72ºC	10 min	

A second amplification was performed to reduce non-specific amplification, by coupling touchdown and nested PCR, using as template DNA the PCR products from the first reaction and primers that prime internal regions of these amplified products (GeneRacer 5' or 3' nested primers and β-CAS specific primers). The program for touchdown PCR was as follows:

Temperature	Time	Cycles
94ºC	2 min	1
94ºC	30 sec	20, each cycle with a 0.5ºC decrease in annealing temp.
60ºC	30 sec	
72ºC	1 min	
94ºC	30 sec	15 cycles
50ºC	30 sec	
72ºC	1 min	
72ºC	10 min	

Amplified PCR products were analyzed in 1.0-1.5 % agarose gel prepared in 1X TAE buffer. PCR products and bands extracted from the gels were purified using the QIAquick PCR Purification kit and the QIAquick Gel Extraction kit (Qiagen), respectively. DNA was quantified with agarose gel electrophoresis using the 1 Kb Plus DNA Ladder (Invitrogen).

2.1.5 Cloning of MANes;BsasA and MANes;BsasB

Cleaned PCR products were cloned in pCR2.1 vector using the TA Cloning kit (Invitrogen). Ligation conditions were as follows in 10 μl reaction mixtures: 1X Ligation buffer (6 mM Tris-HCl pH 7.5, 6 mM $MgCl_2$, 5 mM NaCl, 0.1 mg/ml BSA, 7 mM β-mercaptoethanol, 0.1 mM ATP, 2 mM DTT and 1 mM spermidine), 50 ng pCR2.1 vector, 19 ng insert (MANes;BsasA or MANes;BsasB at a 1:1 vector/insert ratio) and 4 Weiss units of T4 DNA ligase (Invitrogen). The reaction was incubated overnight at 4ºC.

Chemically competent cells were prepared by the $CaCl_2$ method (Sambrook and Russell, 2001) with certain modifications. An overnight culture of *Escherichia coli* DH5α cells in LB medium (10 g/l tryptone, 5 g/l yeast extract and 10 g/l NaCl; pH 7.0) was diluted five times in 5 ml cultures and grown to an OD_{600} of 0.4-0.6. Cells were harvested by centrifugation at 800 x g for 5 min in a microfuge 5415D (Eppendorf) at room temperature. The supernatant was discarded and the bacterial pellet resuspended in 250 µl of ice chilled 50 mM $CaCl_2$. After incubation for 30 min in ice, the cells were centrifuged at 800 g for 5 min at room temperature. The supernatant was discarded and the bacterial pellet was resuspended in 200 µl of ice chilled 100 mM $CaCl_2$. The cells were immediately used for transformation or frozen using liquid nitrogen (after addition of glycerol to a concentration of 10 % v/v) and stored at –80ºC.

For transformation, 2 µl of the ligation mixture were added to 50 µl of competent cells and incubated in ice for 30 min. Heat shock was administered at 42ºC for 45 seconds and further incubation in ice for 5 min. The cells were recovered by addition of 100 µl of SOC medium (20 g/l tryptone, 5 g/l yeast extract, 0.5 g/l NaCl, 2.5 mM KCl and 20 mM glucose; pH 7.0) and incubation at 37ºC for 1 hour. Selection of colonies harboring the pCR2.1 ligated vectors was done in LB agar medium (LB with 15 g/l agar) containing 100 µg/ml ampicilin.

Colony PCR was done in 10-20 µl reaction mixtures containing the same conditions mentioned above (see section 2.1.4). The template plasmid DNA used for amplification was obtained by incubating each selected colony in 10 µl of sterile water at 95ºC for 10 min. 2 µl of this cell lysate were used as template DNA. The primers used were pTV-F and pTV-R (IDT; see table A.2). The PCR cycle was as follows:

Temperature	Time	Cycles
94ºC	2 min	1
94ºC	30 sec	34 cycles
50ºC	30 sec	
72ºC	1 min	
72ºC	10 min	

Positive colonies were picked and grown in LB broth supplemented with 100 µg/ml ampicilin. Plasmid DNA was isolated using the Plasmid Mini kit (Qiagen) and quantified with agarose gel electrophoresis. Sequencing was performed using the Nevada Genomics Center facility (University of Nevada, Reno, Nevada).

2.2 MANes;BsasA and MANes;BsasB characterization

2.2.1 In silico characterization of MANes;BsasA and MANes;BsasB

The two different β-CAS cDNA copies isolated were named MANes (*Manihot esculenta*);BsasA and MANes;BsasB and their amino acid sequences were predicted using the Translate tool of ExPASy (Expert Protein Analysis System) proteomics server of the Swiss Institute of Bioinformatics (SIB; URL: http://expasy.org/tools/dna.html). Protein sequence alignments were done using ClustalW2 (Chenna et al., 2003) tool for multiple sequence alignment (see appendix C for accession numbers of the sequences employed). Default parameters for multiple sequence alignment were used, according to the ClustalW2 version available at the server of the European Molecular Biology Laboratory of the European Bioinformatics Institute (EMBL-EBI). Phylogenetic analysis and tree construction was done using *Phylogeny fr*. (Dereeper et al., 2004). Prediction of mitochondrial, plastidic and cytosolic

sub-cellular location was achieved using TargetP (Emanuelsson et al., 2004) and Predotar (Small et al., 2004).

2.2.2 MANes;BsasA, MANes;BsasB and ARAth;Bsas3;1 cloning into pTV118N

Signal peptide length and potential cleavage site was estimated using the MITOPROT tool (Claros and Vincens, 1996). Sequences were also analyzed for restriction sites using the NEB cutter V2.0 tool (New England Biolabs, Ipswich, MA, USA; URL: http://tools.neb.com/NEBcutter2/index.php). Primers were designed to amplify engineered truncated versions of MANes;BsasA, MANes;BsasB and ARAth;Bsas3;1 (*Arabidopsis thaliana* β-CAS), with a start (ATG) codon at amino acid positions 37, 30 and 28 respectively (see Table A.2 of appendix A). These versions of the genes lacked the mitochondrial signal peptide, absent in the mature forms of the proteins. The primers also resembled *Nco*I (forward primer) and *Xba*I (reverse primer) restriction sites for cloning (see Table A.2 of appendix A). The *Arabidopsis thaliana* β-CAS full-length clone (accession AJ010505; Hatzfeld et al., 2000) was kindly provided by Dr. Kazuki Saito (Chiba University, Chiba, Japan) and was used as a control.

MANes;BsasA, MANes;BsasB and ARAth;Bsas3;1 were amplified using the aforementioned cloning primers at the same reaction conditions mentioned in section 2.1.4. The PCR cycle used was the same as indicated in section 2.1.5 for colony PCR. The purified PCR products and the *Escherichia coli* over-expression vector pTV118N (Takara) were digested using the restriction enzymes *Nco*I and *Xba*I (Promega) in 20-50 µl reaction mixtures containing: 0.1 – 0.5 µg DNA, 1X Buffer D (Promega; 6 mM Tris-HCl pH 7.9, 6 mM $MgCl_2$, 150 mM NaCl and 1 mM DTT), 0.1 mg/ml BSA and 2.5 U of each enzyme. Digestions of the plasmid with each enzyme in separate tubes (single digestions) were also done as digestion controls. Reactions were

incubated at 37ºC for 3 hours. After the reaction period, the enzymes were heat inactivated at 65ºC for 10 min. Digestions were checked in 1% agarose gels prepared in 1X TAE buffer.

The purified digested DNA molecules were ligated at a 1:1 vector/insert ratio, in 10-30 µl reactions containing the following: 1X T4 DNA ligase buffer (30 mM Tris-HCl pH 7.8, 10 mM $MgCl_2$, 10 mM DTT and 1 mM ATP), 100 ng of *Nco*I/*Xba*I digested pTV118N and 32 ng of insert (*Nco*I/*Xba*I digested MANes;BsasA, MANes;BsasB and ARAth;Bsas3;1). The reactions were incubated at 4ºC overnight. Transformation conditions were done as indicated above (see section 2.1.5) but *Escherichia coli* NK3 strain was used instead of DH5α, since this strain lacks CAS and CS activities (Hulanicka et al., 1986). This strain was kindly provided by Dr. Kazuki Saito. The expression of the genes cloned in the *Nco*I site of pTV118N are under the control of the *lacZ* promoter for induction with isopropyl β-D-1-thiogalactopyranoside (IPTG).

2.2.3 Bacterial over-expression and preparation of crude protein extracts

NK3 cell growth, induction and crude protein extract preparation was done as previously reported (Maruyama et al., 2001; Hatzfeld et al., 2000). Overnight cultures of NK3 cells harboring the MANes;BsasA, MANes;BsasB and ARAth;Bsas3;1 engineered genes in pTV118N were diluted 100 times in 100 ml LB broth supplemented with 100 µg/ml ampicilin to have an approximate initial OD_{600} of 0.025. After initial growth for 5 hours under agitation at 145 rpm in a InnOva 2000 platform shaker (New Brunswick Scientific) at 37ºC (VWR Incubator Model 1545; VWR, West Chester, PA, USA), the cells were induced by addition of IPTG to a final concentration of 1 mM and grown for nine additional hours at 37ºC under agitation. The cells were harvested by centrifugation at 3,000 g for 10 min at room temperature in a Centrifuge 5810R (Eppendorf). The supernatant was discarded and the cell pellet was resuspended in 10 ml of ice chilled 50 mM Tris-HCl pH 8.5. The cells were lyzed by 1 min sonication at 95% full

power using a Microson Ultrasonic Cell Disruptor (Misonix, Farmingdale, NY, USA). The cell debris was separated by centrifugation at 16,000 g for 20 min at 4ºC and the supernatant was taken as the crude protein extract. Protein concentration was quantified by the Bradford method (Bradford, 1976) using bovine serum albumin as a standard.

2.2.4 β-CAS and CS activity assays

β-CAS activity was assayed as previously reported (Hasegawa et al., 1994; Elias et al., 1997) with slight modifications. The reaction mixture contained 25 mM KCN and 5 mM L-cysteine in 50 mM Tris-HCl pH 8.5 and different quantities of the crude protein extract to a final volume of 1 ml in 1.5 ml microfuge tubes. The reaction was initiated by addition of the crude protein extract and was incubated at 30ºC for 10 min. The reaction was stopped by addition of 200 µl of 0.03 M ferric chloride (prepared in 1.2 N HCl) and 200 µl of 0.02 M N,N-dimethyl-p-phenylenediamine sulfate (DPDS; prepared in 7.2 N HCl). The microtube was closed, shaken and after 20 min of incubation at room temperature, centrifuged at 13,000 x g for 2 min to remove precipitated proteins. The amount of methylene blue formation due to the presence of H_2S, one of the CAS reaction products, was measured spectrophotometrically at 665 nm. The H_2S concentration was measured against a Na_2S standard curve. Activity was defined as µmol of $H_2S \cdot min^{-1} \cdot$ mg protein^{-1} (Hatzfeld et al., 2000; Maruyama et al., 2001).

Cysteine synthase activity was measured according to Maruyama et al., (2001) and Lunn et al., (1990). The reaction mixture contained 2.5 mM Na_2S, 5 mM *O*-acetyl-l-serine and 5 mM dithiothreitol (DTT) in 50 mM sodium phosphate buffer pH 7.5 and varying quantities of the protein extract, to a final volume of 100 µl in 1.5 ml microfuge tubes. The reaction was initiated by addition of the protein extract and incubated at 25ºC for 10 min. The reaction was stopped by addition of 50 µl 20% (w/v) trichloroacetic acid (TCA) and the precipitated protein was

eliminated by centrifugation for two minutes at 13,000 x g. The cysteine formed due to the reaction was measured by the method reported by Gaitonde (1967), which is based on the formation of a red-pink product when ninhydrin reacts with cysteine. The supernatant was transferred to a glass tube containing 100 µl of concentrated acetic acid and 200 µl ninhydrin reagent (25 mg/ml prepared in 60:40 acetic acid/HCl). The tubes were incubated in boiling water for 10 min and rapidly cooled in tap water followed by addition of 550 µl of 95 % ethanol. The solutions were read for absorbance at 560 nm and cysteine concentration was calculated by the molar extinction coefficient for the colorimetric assay (25,000 $M^{-1} \cdot cm^{-1}$).

2.3 Construction of the plant over-expression vectors

2.3.1 MANes;BsasA and ARAth;Bsas3;1 cloning into pKYLX

The pKYLX71:35S^2 binary vector (Schardl et al., 1987) was used for cassava transformation. MANes;BsasA and ARAth;Bsas3;1 were amplified with primers that contained restriction sites as indicated in section 2.2.2 for cloning into pTV118, but instead of an *Nco*I site in the forward primer, a *Sac*I site was included (see Table A.2 of appendix A). The genes were cloned in the *Sac*I (5' end) and *Xba*I (3'end) sites of the pKYLX71:35S^2 multiple cloning site, under the control of the duplicate cauliflower mosaic virus (CaMV) 35S promoter for constitutive expression. Digestions, ligations, DH5α transformation and plasmid isolation were also done as previously stated (see section 2.1.5). Colony selection was done in LB agar plates containing 10 µg/ml tetracycline. Colony PCR was performed as indicated in sections 2.1.4 and 2.1.5, using a forward primer that primes the downstream region of the CaMV 35S promoter and as reverse primer, the cassava β-CAS specific primer 2 for the pKYLX:MANes construct (see Table A.1), or the Arath-Xba primer for the pKYLX:ARAth construct (see Table A.2).

Sequencing was also performed using the Nevada Genomics Center facility (University of Nevada, Reno, Nevada).

2.3.2 *Agrobacterium* transformation

Once the sequence of the constructs was checked for mutations by sequencing, *Agrobacterium tumefaciens* LBA4404 strain ElectroMAX cells (Invitrogen) were transformed by electroporation according to the instructions of the manufacturer. Cells were thawed on ice for 30 min and 20 µl of cells were mixed with 40 ng of plasmid DNA in an ice chilled 1.5 ml microtube. The mixture was transferred to an ice chilled 0.1 cm electroporation cuvette and electroporation was performed at 2.0 kV in a MicroPulser elctroporator (Bio-Rad, Hercules, CA, USA). Immediately after electroporation, 1 ml of YM medium (0.4 g/l yeast extract, 10 g/l mannitol, 0.1 g/l NaCl, 0.8 mM $MgSO_4 \cdot 7H_2O$ and 2.2 mM $K_2HPO_4 \cdot 3H_2O$; pH 7.0) were added to the cuvette, transferred to a 1.5 ml microtube and incubated at 30ºC for 5 hours under slow agitation. After the recovery period, 100 µl of the culture were plated in YM agar medium (YM medium with 15 g/l agar) containing 100 µg/ml streptomycin and 10 µg/ml tetracyclin. Colony PCR was performed as indicated in section 2.1.5.

2.4 Plant transformation

2.4.1 Plant material and media

Plant transformation was done as previously reported by Siritunga and Sayre (2003). Mcol 2215 *in vitro* plants were micro-propagated in 4e medium (1X Murashige and Skoog (MS; 1962) salts supplemented with 2 % (w/v) sucrose, 1 mg/l thiamine, 100 mg/l myo-inositol, 0.04 mg/l N^6-benzilaminopurine (BAP), 0.05 mg/l gibberellic acid (GA_3), 0.02 mg/l naphthalene-1-acetic acid (NAA) and 0.7 % (w/v) agar; pH 5.75). The plants were incubated at 28ºC in a plant

incubator (model I41LL, Percival Scientific, Perry, IA, USA,) under a 12 hour day photoperiod regime (5 μmol photons /m^2/s).

Apical leaves were taken from micro-propagated plants and cultured in MS8 medium (1X MS salts supplemented with 2 % (w/v) sucrose, 10 ml/l 100X Gamborg's B-5 vitamins (Gamborg et al., 1968), 8 mg/l 2,4-dichlorophenoxyacetic acid (2,4-D), 0.5 mg/l $CuSO_4$, 50 mg/l casein and 0.2 % (w/v) phytagel; pH 5.75) for induction of somatic embryos. The cultures were covered with cheesecloth and incubated at the same conditions as the micro-propagated plants for 4-6 weeks.

Embryo clumps were transferred to RM2 medium (1X MS salts supplemented with 2 % (w/v) sucrose, 1 mg/l thiamine, 100 mg/l myo-inositol, 1 mg/l BAP, 0.5 mg/l GA_3 and 0.2 % (w/v) phytagel; pH 5.75) to induce the development of cotyledonary leaves (3-4 weeks).

2.4.2 Plant transformation and selection

Cotyledonary leaves were soaked in 50 ml 1X MS salts devoid of nitrogen, containing a few drops of Tween-20. Leaves were cut into approximately 0.5 cm^2 peaces and placed in MS8 medium containing 100 μM acetosyringone. Two-day-old 15 ml cultures of *Agrobacterium* cells harboring any of the constructs, grown in a rotatory shaker (Model TC-7, New Brunswick Scientific, Edison, NJ, USA) at 30ºC (Fisher Isotemp Incubator, Fisher Scientific, Madison, WI, USA) were centrifuged at room temperature for 7 min at 6,500 x g. The supernatant was discarded and the cells were resuspended in 1X MS salts devoid of nitrogen, to an OD_{600} of 0.45-0.5. The cells were induced by addition of acetosyringone to a final concentration of 200 μM and incubation at 28ºC for 2-4 hours. After this period, the cotyledonary leaves were inoculated with one or two drops of the induced culture. Co-cultivation was done for two days at room temperature in darkness.

Co-cultivated explants were transferred to MS8 medium containing 75 mg/ml paromomycin and 500 mg/l carbenicilin (for selection of transgenic plantlets and killing *Agrobacterium*, respectively), covered with cheesecloth and incubated at the same conditions mentioned above for two months. Clumps of putative transgenic somatic embryos that grew in selective MS8 medium were transferred to RM2 medium containing paromomycin and carbenicilin at the same concentrations and grown for two additional months. After this period, germinated clumps were transferred to RM2 medium devoid of antibiotics and grown for at least one month. Germinated plantlets were then transferred to micro-propagation medium and grown for 3-4 months.

2.5 PCR screening and phenotypical analyses

2.5.1 PCR screening

DNA was extracted from 3-4 month old *in vitro* cassava leaves according to a modified version of the method of Dellaporta et al. (1983). Approximately 100-200 mg of tissue was ground using sand and a micropestle in 1.5 ml microfuge tubes containing 400 µl of extraction buffer (100 mM Tris-HCl, 50 mM EDTA, 500 mM NaCl, 1% (w/v) polyvinyl pyrrolidone (PVPD), 10 mM β-mercaptoethanol and 1% (w/v) sodium dodecyl sulfate) previously warmed at 65ºC. The homogenized tissue was incubated at 65ºC for 15 min with vortexing at 5 min intervals. Samples were cooled to room temperature for 2 min followed by addition of 250 µl of cold 5M potassium acetate and mixed by inverting the tubes several times. The samples were then incubated in ice for 20 min and centrifuged top speed for 15 min at room temperature. The supernatant was transferred to a new 1.5 ml microtube, followed by addition of 700 µl cold isopropanol and mixing by inverting the tubes several times. The samples were then incubated for 2 hours or overnight at -80ºC, centrifuged top speed for 15 min and the DNA pellet was air

dried for 2-3 hours in the bench top. After drying, the pellets were washed with 500 ml cold 70% (v/v) ethanol and air dried for 1 hour in the bench top. The cleaned DNA pellet was dissolved in 50 µl of TE buffer (10 mM Tris-HCl pH 7.5; 1 mM EDTA). The DNA was quantified using a NanoDrop ND 1000 Spectrophotometer (Thermo).

PCR screening of transgenic plants was done using the reaction conditions stated in section 2.1.4, but using 100-1000 ng of genomic DNA and varying cycle programs such as the one stated in section 2.1.5. DNA was analyzed using 1.5 % agarose gels prepared in 1X TAE.

2.5.2 β-CAS activity assays from plant tissue

Approximately 50 mg of *in vitro* leaves were ground in 2 ml of ice chilled extraction buffer (Tris-HCl 50mM pH 8.5, 5 mM DTT, 5 mM phenyl-methyl-sulfonyl fluoride (PMSF), 1 mM EDTA and 2% (w/v) polyvinyl pyrrolidone (PVPD) with an ice chilled mortar and pestle. The ground mixture was transferred to 2 ml microtubes and burst by sonication at 95% full power, followed by centrifugation at 13,000 x g at a temperature of 4ºC. The supernatant was separated and taken as the total leaf protein extract. Enzyme activity was assayed as indicated in section 2.2.5.

RESULTS

3 Results

3.1 *Bsas* cDNAs isolation, cloning and sequencing

3.1.1 3' RACE

Rapid Amplification of cDNA Ends (RACE) PCR was used as the strategy for cDNA isolation, due to availability of a partial internal sequence of the cassava β-CAS gene (figure 3.1). RACE PCR allows isolation of the ends of a cDNA sequence, if at least a partial internal sequence is known. For 3' RACE, an oligo adapter of known sequence was included downstream the oligo-dT primer provided by the GeneRacer kit for RACE PCR (Invitrogen).

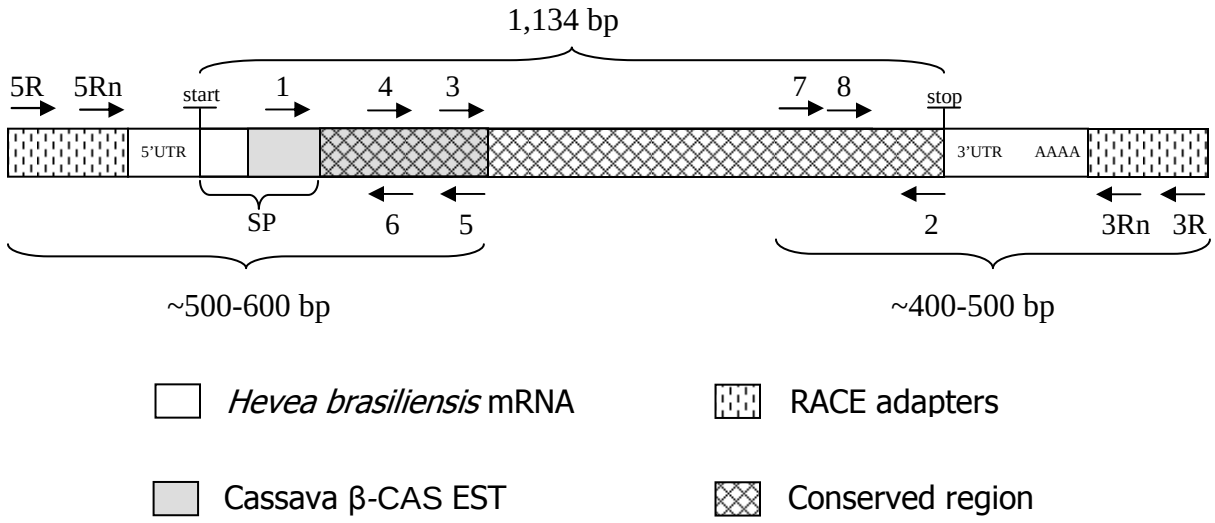

Figure 3.1: Schematic representation of the β-CAS mRNA sequence. The conserved region includes all the coding sequence, except the signal peptide and untranslated regions (UTR). Primers 1, 3, 4, 5 and 6 were designed using the cassava β-CAS EST sequence (accession CK644310). Primers 2, 7 and 8 were designed using the rubber tree (*Hevea brasiliensis*) mRNA sequence (accession AY207389). Primers 5R, 5Rn, 3R and 3Rn were provided by the GeneRacer kit. "SP", signal peptide; "start", start codon; "stop", stop codon; "AAAA", poly-A tail.

The presence of a very long conserved region has been reported to be shared between all the members of the *Bsas* family of enzymes. This region resembles the pyridoxal phosphate (PLP) binding domain, as well as other functional domains involved mainly in substrate binding

and dimer formation, since these enzymes have been reported to be homo-dimers in their active forms in most plants (Maruyama et al., 2001; Hatzfeld et al., 2000). Maruyama et al. (2001) were able to explain with certainty that the conserved region does not include a region of about 20-40 residues in the amino terminal that encodes the targeting sequence for protein sorting into their respective compartments. This non-conserved region which is different between the distinct isoforms (i.e. mitochondrial, plastidic and cytosolic) provides regions of sequence specificity needed for specific primers design. In addition, the non-conserved character of the 5' and 3' untranslated regions (UTR) are useful sequences for specific gene isolation.

Figure 3.1 shows the approximate priming sites of the primers within the β-CAS mRNA molecule. Approximate amplicon sizes are also shown for 5' and 3' ends. Different PCR reactions were done using different template/primers combinations and the reactions were labeled according to these criteria (e.g., reaction **M1/2** means that the template RNA source was from the **M**col 2215 variety, using forward primer **1** and reverse primer **2**; on the other hand **T4/3R** means that the RNA source was the **T**MS 60444 variety, using forward primer **4** and the reverse primer was the **3'R**ACE primer; **Rn** means that the primer used a **R**ACE **n**ested primer).

Figure 3.2 shows results of a 3'RACE PCR reaction using primer combinations 3/3R and 4/3R (long 3'RACE PCR), in which the approximate amplicon sizes expected are 1,050 and 1,170 bp, respectively. The expected bands were obtained in all the lanes (figure 3.2). Furthermore, multiple bands of expected size can be observed in more than one lane (e.g., lane T3/R). However, non specific amplification was observed in all reactions. To reduce non-specific amplification and assure that the multiple bands observed are β-CAS specific, reaction products M3/3R, T3/3R, M4/3R and T4/3R were used as templates for PCR reactions using primers 7, 8 and 3Rn (nested PCR). Results are shown in figure 3.3.

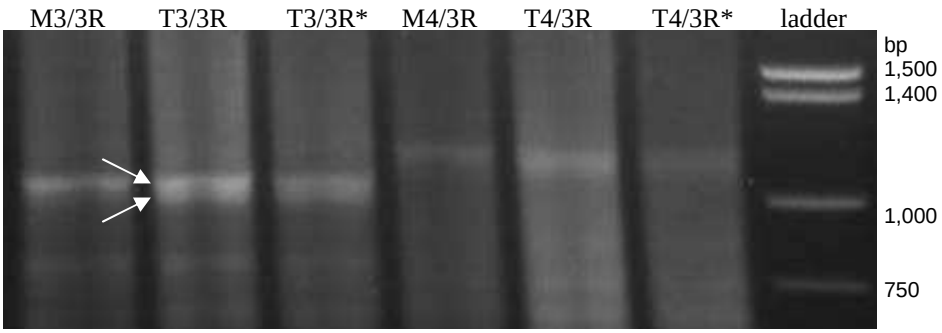

Figure 3.2: Long 3'RACE PCR (1% agarose on 1X TAE). *These lanes contained, by error, cDNA synthesized with oligo dT primer, as well as with GeneRacer oligo dT. This gel was run for 3 hours at 75 mV. It contained approximately 25-35 μl of the total reaction.

At least three visible bands are observed with the expected molecular weight for PCR products 7/3Rn and 8/3Rn (400-500 bp; see figure 3.3), which were named 1, 2 and 3. The elimination of non-specific amplification was achieved by coupling nested PCR (amplification from a PCR product using primers that anneal to internal regions of the amplicon) with a touchdown program in which the annealing temperature was decreased 0.5ºC each cycle, starting with 60ºC until reaching 50ºC (see section 2.1.4). Amplification of more than one band was obtained because the gene specific primers used anneal to a conserved region shared between the different members of the *Bsas* family of genes, thus amplifying all the possible members (see figure 3.1).

To check whether the bands 1, 2 and 3 are β-CAS or cysteine synthase (CS) specific amplicons, bands were cut, purified and sequenced using the Nevada Genomics Center facility (University of Nevada, Reno, Nevada). The sequenced products were checked for homology to β-CAS sequences using the Basic Local Alignment Sequence Tool (BLAST; Altschul et al., 1990). All the sequences showed homology to β-CAS and/or CS sequences (data not shown).

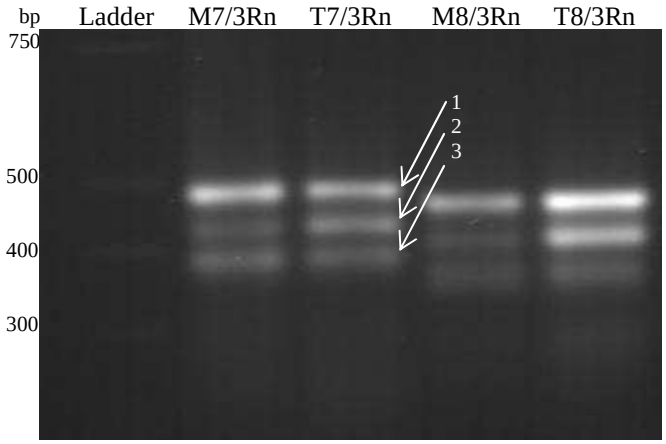

Figure 3.3: Nested PCR of the 3'RACE products M3/3R and T3/3R using primer combinations 7/3Rn and 8/3Rn (1.5 % agarose on 1X TAE). This gel was run for 2 hours at 100 V and then for 3 hours at 75 V.

To check for sequence homology between the three bands, sequence alignments were done with the ClustalW tool using default parameters (Chenna et al., 2003). High homology was found between sequences T7/Rn-1 and T7/Rn-2 (see figure 3.4), except that a gap of about 54 bp was observed upstream the poly-A tail of the T7/Rn-2 sequence, hence explaining the difference in molecular weight observed between these two bands in the agarose gel (see figure 3.3). The third band, T7/Rn-3, exhibited slight differences in its nucleotide sequence, giving the possibility that it belongs to a gene of the same family, but different from the gene encoded by T7/Rn-1 and T7/Rn-2 (see figure 3.4). These results could only be fully understood when a similar result was obtained in 5'RACE (see next section).

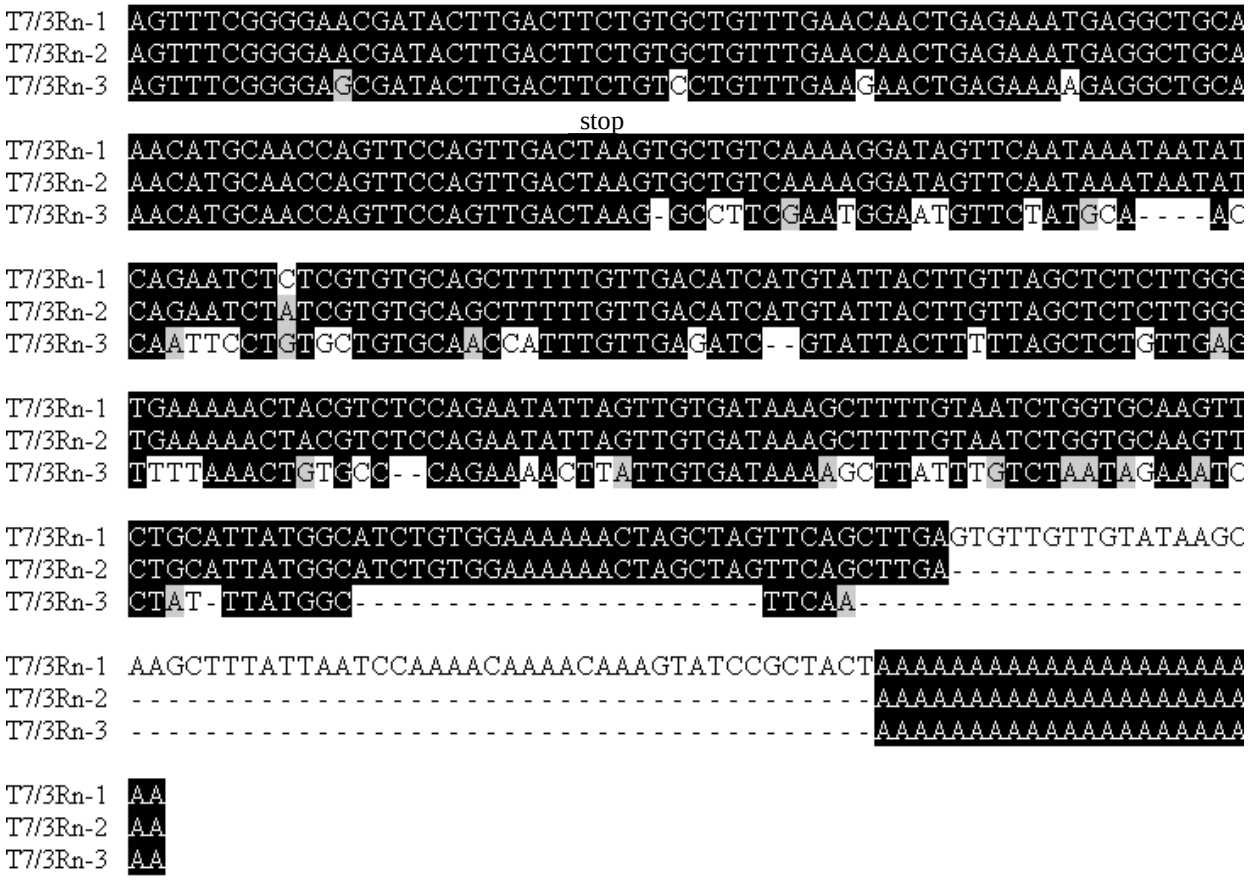

Figure 3.4: ClustalW alignment of bands 1, 2 and 3 of the PCR reaction T7/3Rn (3'RACE). Black backgrounds indicate conserved residues; grey backgrounds indicate similar residues (i.e. purine to purine, or pyrimidine to pyrimidine shifts); white backgrounds indicate completely different residues. The position of the stop codon is indicated for orientation.

3.1.2 5' RACE

For 5'RACE PCR, a RNA molecule of known sequence was ligated to the 5'end of all the mRNA molecules of a total RNA sample from cassava leaves, using the GeneRacer Kit (Invitrogen; see section 2.1.3).

An initial long 5'RACE PCR was done using primer 5R as forward primer and primer 2 as reverse primer. Results are shown in figure 3.5, lanes M5R/2 and T5R/2. The expected PCR product for these two reactions was ~1.2 kb, and for unknown reasons it was only observed in reaction M5R/2. Non-specific amplification was observed and was eliminated by coupling touchdown and nested PCR in a second PCR reaction using products M5R/2 and T5R/2 as

templates, and primer combinations 5Rn/5 and 5Rn/6. Results are shown in the first four lanes of figure 3.5.

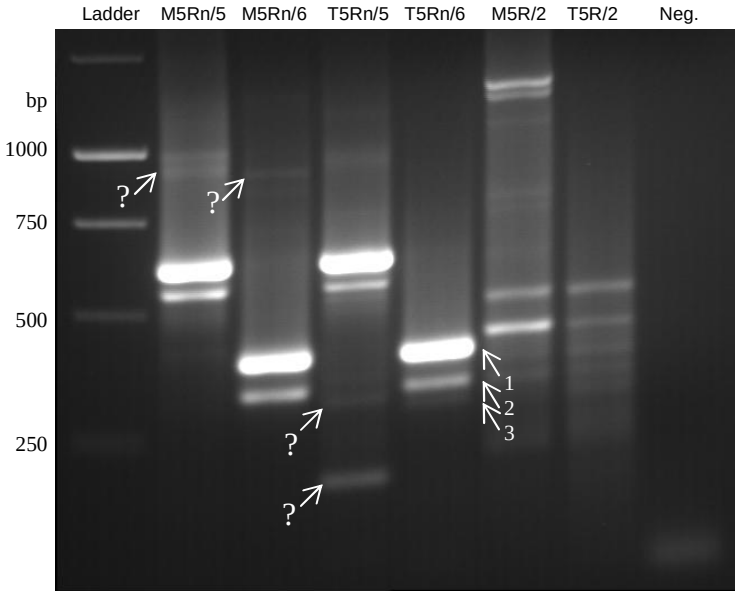

Figure 3.5: 5'RACE PCR. The PCR products from lanes M5R/2 and T5R/2 were used as templates to amplify the contents of the first four lanes. The bands marked as 1, 2 and 3, were isolated and sequenced. The bands indicated with question marks were not isolated even thou many attempts were made to do so.

In figure 3.5, the expected molecular weighs for PCR products using primer pairs 5Rn/5 and 5Rn/6 are approximately 600 bp and 400 bp, respectively (see figure 3.1). The samples belonging to the TMS 60444 variety showed three bands of expected molecular weight (see lane T5Rn/6), while the Mcol 2215 samples amplified only two of the three expected bands, based on the results obtained in 3'RACE. Only the bands labeled as "1", "2" and "3" were successfully isolated and sequenced.

Figure 3.6 illustrates a sequence alignment of bands 1, 2 and 3 from T5Rn/6. High homology was found between sequences T5Rn/6-1 and T5Rn/6-3, and again a difference in sequence length is observed: T5Rn/6-1 is ~100 bp longer than T5Rn/6-3. On the other hand, the sequence T5Rn/6-2 was shown to be more dissimilar in its nucleotide sequence than the first two and showwn to be shorter than T5Rn/6-1 and longer than T5Rn/6-3. This result can be compared

to the result obtained in 3'RACE (see last section), where sequences belonging to the first two bands (T7/3Rn-1 and T7/3Rn-2) were more similar in sequence than the third (T7/3Rn-3), hence raising the possibility that these bands belong to the same genes as bands 1 and 3 from 5'RACE, and that the band named as T7/Rn-3 in 3'RACE makes part of the same gene together with band TRn/6-2 from 5'RACE.

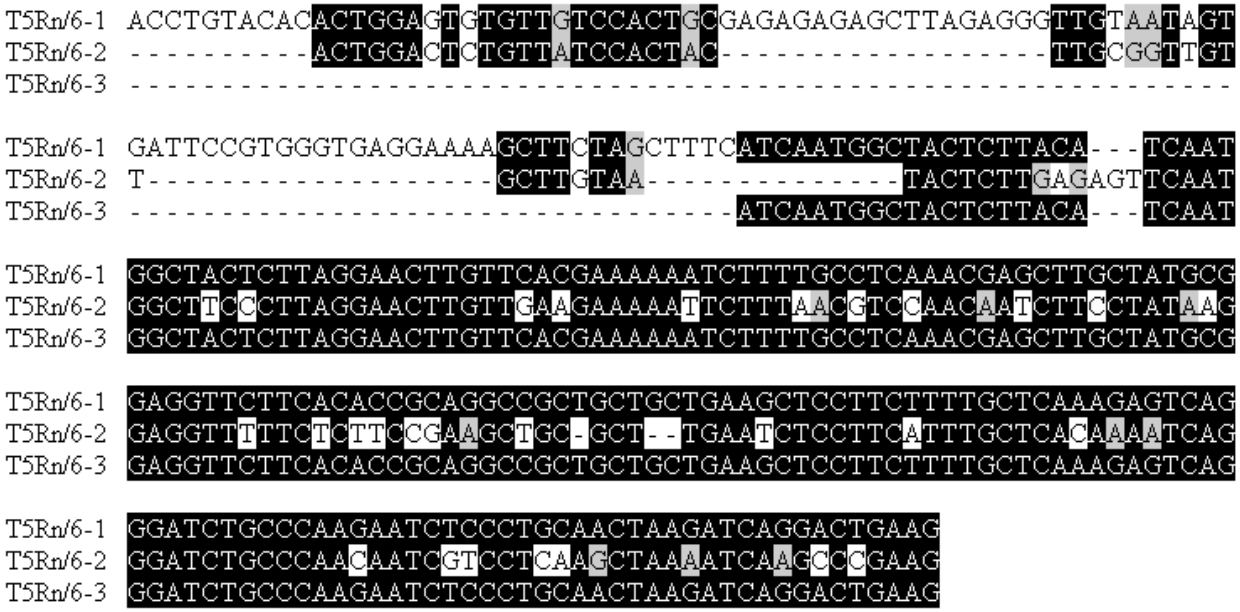

Figure 3.6: ClustalW alignment of bands 1, 2 and 3 from PCR reaction T5Rn/6. Black backgrounds indicate conserved residues; grey backgrounds indicate similar residues (i.e. purine to purine, or pyrimidine to pyrimidine shifts); white backgrounds indicate completely different residues.

These RACE PCR results hint about the presence of at least three genes encoding members of the *Bsas* family of enzymes being expressed in cassava leaves. The two identical sequences found in both 5' and 3' ends tells about the expression of two identical members, whose difference in molecular weight could be due to large and short versions of the same enzyme, if the differences in size are not restricted only to the un-translated regions of the mRNA. However, due to sequence identity, the proteins encoded by these two identical genes

are expected to be functionally identical. The remaining gene is expected to encode a different enzyme.

3.1.3 Isolation of two different full length cDNAs

In order to gather information that could lead to isolation of the full length cDNAs, primers were designed using the sequences obtained from the RACE PCR results, since they show non-conserved regions that provide specificity for gene isolation. Since sequence similarity is an indicator of functionality and this project aimed to identify the β-CAS gene, efforts were directed towards isolation of the genes with differences in their nucleotide sequence, in order to provide insights about their functionality. In addition, the specific isolation of the two identical isoforms cannot be achieved using PCR derived techniques, since primers designed to amplify one sequence will amplify the other as well. Other techniques such as the construction and screening of a cDNA library might be a better method.

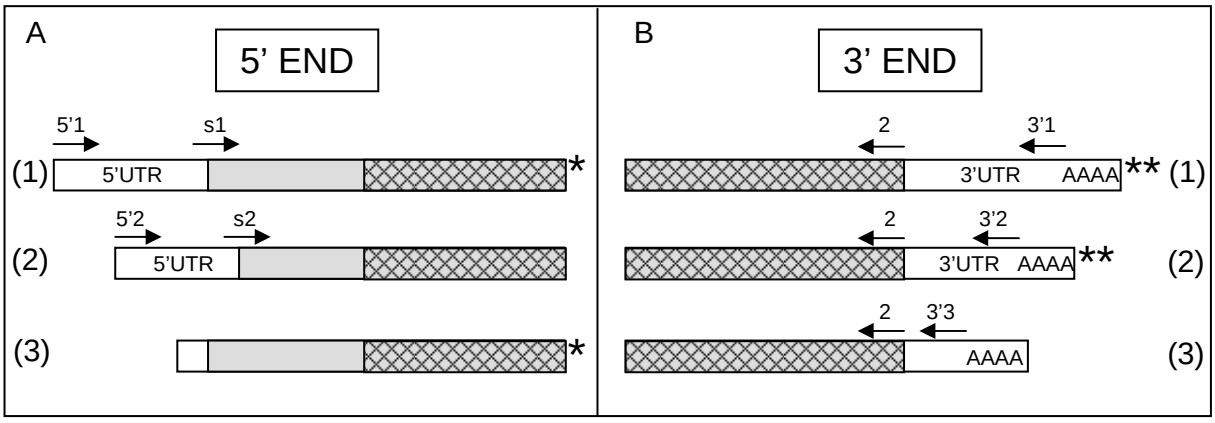

Figure 3.7: Scheme of the different bands obtained in RACE PCR showing the primers designed using these sequences. A, Primers designed using sequences from the 5'end: 5'1 and 5'2 primers anneal to the 5'untranslated region (UTR), while s1 and s2 primers anneal at the start codons; B, Primers designed using sequences from the 3'end. Primers 3'1, 3'2 and 3'3 anneal to the 3'UTR region. Note that primer "2" is the same as the one presented in figure 3.1 and already used for long 5' RACE (see figure 3.5). Legend: white, UTR regions; grey, coding sequence; checked, conserved region; * or **, identical sequences; "AAAA" poly-A tail.

Figure 3.7 shows a schematic representation of the three bands obtained in both 5' RACE and 3' RACE PCR and the approximate localization of the annealing sites of the primers designed. PCR reactions were done using different primer combinations that could trigger the specific isolation of the two different genes. An interesting result can be observed in figure 3.8A. When using primer 5'1, which anneals to one of the two identical sequences of the 5' end, amplification was achieved when using primers 3'1 and 3'2, which anneal to the two identical sequences (potentially the same gene) of the 3'end. On the other hand, when using primer 5'2, which anneal to the second potential isoform, amplification was not achieved using the same cDNA sample. Moreover, when using this forward primer with a primer that could hit the third (non-identical) sequence of the 3'end, amplification of the correct molecular weight was obtained (figure 3.8B, lane T5'2/3'3).

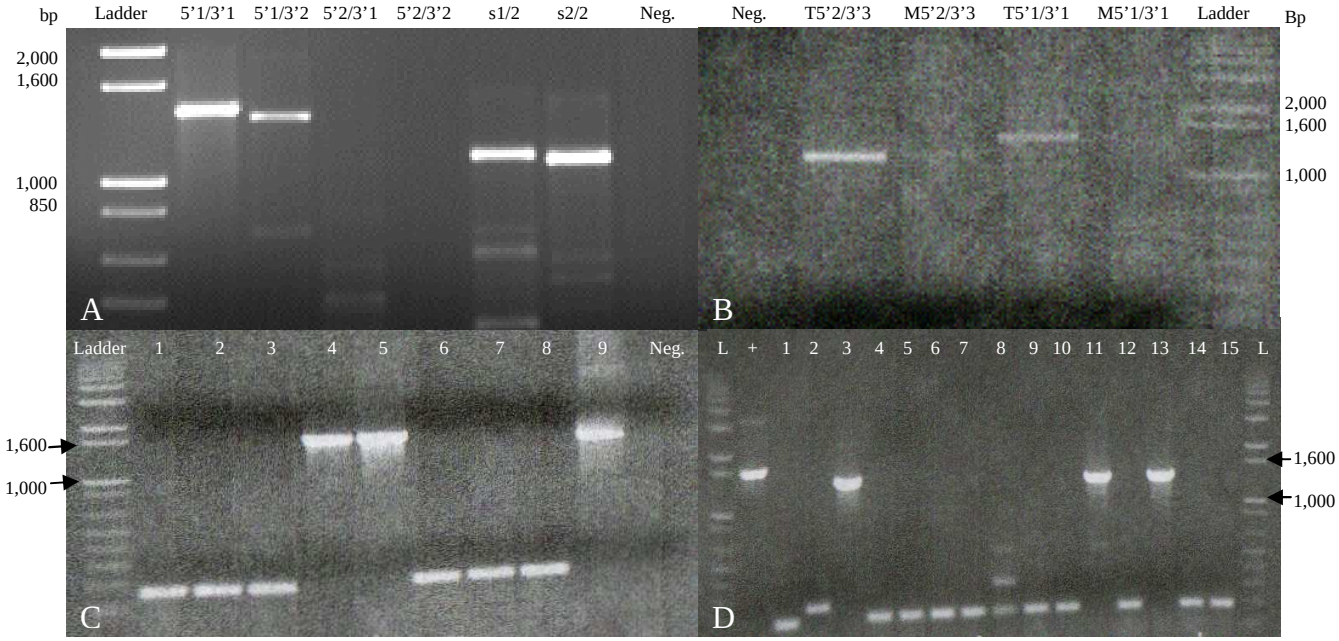

Figure 3.8: PCR amplification and cloning of two *Bsas* full length cDNAs. A, Amplification from TMS 60444 leaf cDNA; B, Amplification from TMS 60444 and Mcol 2215 leaf cDNA; C, colony PCR of amplicon T5'1/3'1 (see B) cloned in pCR2.1 using the TA cloning kit (Invitrogen); D, colony PCR of amplicon T5'2/3'3 (see B panel) cloned in pCR2.1. Both primers used for colony PCR anneal to the vector (pTV-F and pTV-R; see table A.2 of appendix A).

These results support the hypothesis from the last section, stating that the two identical sequences might encode one gene, while the other sequence with slight differences could encode a second gene. The genes were amplified from the start codon to the stop codon using primer combinations s1/2 and s2/2 (see figure 3.8A). The full length cDNAs encoding these two genes were amplified using primer combinations 5'1/3'1 and 5'2/3'3, respectively, from TMS 60444 and Mcol2215 leaf cDNA (figure 3.8B) and the amplicons from TMS 60444 were cloned in pCR2.1 using the Original TA cloning kit (Invitrogen). Figure 3.8 C and D shows the colony PCR screening of the clones resembling the two distinct genes. Positive clones were sequenced using the Nevada Genomics Center facility (University of Nevada, Reno). The two sequences are included in Appendix B.

3.2 MANes;BsasA and MANes;BsasB functional characterization

3.2.1 Sequence cluster analysis

Hatzfeld et al. (2000) proposed for the first time a classification method for this family of enzymes, in which the first three letters of the genus and the two first letters of the specific name of the organism of origin are used. Also the isoform identity based on sub-cellular location and kinetic properties is considered, Bsas1 and Bsas2 being the enzymes showing cysteine synthase (CS) catalytic properties and targeted to the cytosol and the chloroplast, respectively. The enzymes showing β-CAS activity and targeted to the mitochondria have been classified in a third group, Bsas3. For example, the actual names proposed for the Arabidopsis cytosolic CS, plastidic CS and mitochondrial β-CAS are ARAth;Bsas1, ARAth;Bsas2 and ARAth;Bsas3, respectively. In the case where more than one isoform has been isolated from the cytosol, another number will follow: ARAth;Bsas1;1 is the first cytosolic CS isolated from Arabidopsis.

Since no information was available when starting this *in silico* protein characterization, the two cassava proteins were named MANes;BsasA and MANes;BsasB.

The identification of the two isolated sequences was first achieved by analyzing their sequence homology to previously identified protein sequences of the *Bsas* family of enzymes. The full-length cDNA sequences were used for prediction of the amino acid sequences using the Translate tool of ExPASy (Expert Protein Analysis System) proteomics server of the Swiss Institute of Bioinformatics (SIB). These sequences were aligned using the ClustalW tool (Chenna et al., 2003) for multiple sequence alignment (see appendix C for accession numbers of the sequences employed). Default parameters for multiple sequence alignment were used, according to the ClustalW2 version available at the server of the European Molecular Biology Laboratory of the European Bioinformatics Institute (EMBL-EBI).

Figure 3.9 shows the multiple sequence alignment of the two cassava (MANes) proteins with three isoforms (cytosolic, chloroplastic and mitochondrial) from Arabidopsis, spinach and potato. The rubber tree sequence was also included. Both MANes;BsasA and MANes;BsasB align with these sequences, with results compared to previous bioinformatics studies (Maruyama et al., 2001). As mentioned above (section 3.1.1), the *Bsas* proteins share a very long conserved region following the non-conserved signal peptides. It can be observed that the main differences between these sequences lie in the amino-terminal, which is where these signal peptides are encoded. The cytosolic Bsas1 enzymes are the shortest proteins, since a lack of the signal peptide is expected to target these enzymes to the cytosolic compartment. Chloroplastic Bsas2 sequences have the longest signal peptides followed by the mitochondrial isoforms.

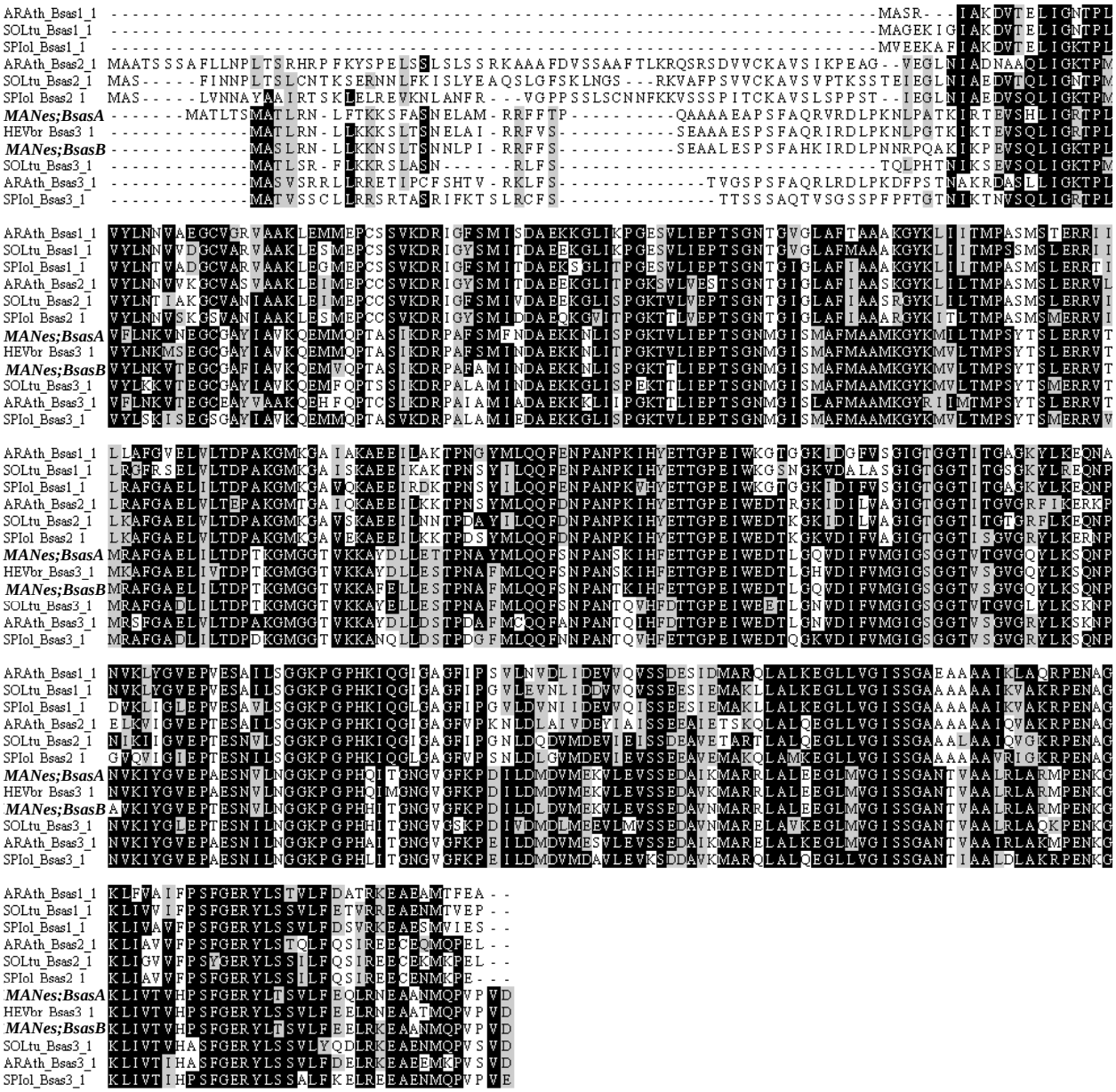

Figure 3.9: Protein sequence alignment of MANes-BsasA and MANes-BsasB with other 10 *Bsas* protein sequences. Three different isoforms (i.e. cytosolic Bsas1, plastidic Bsas2 and mitochondrial Bsas3) were included from three of the plants in which these enzymes have been extensively studied (Hatzfeld et al., 2000; Maruyama et al., 2001), namely *Arabidopsis thaliana* (ARAth), spinach (*Spinacea oleracea*, SPIol) and potato (*Solanum tuberosum*, SOLtu). The β-CAS of rubber tree (*Hevea brasiliensis*, HEVbr) was also included, due to high homology to cassava β-CAS. For a complete list of sequences and accession numbers, see appendix C. Residues are enclosed in a gradient of colored backgrounds ranging from black, dark grey, light grey to white, being black most conserved and white non-conserved residues. The alignment parameters used distribute the sequences from top to bottom based on sequence similarity. The MANes;BsasA and MANes;BsasB sequences are shown in "**Bold**" and "*Italics*" to facilitate observation.

These ten sequences, as well as 14 additional sequences, were used for a robust phylogenetic analysis using the *Phylogeny fr*. tool (Dereeper et al., 2008). This tool provides very simple analysis of ClustalW alignments of biological sequences. This information might provide clearer insights about the identity of MANes;BsasA and MANes;BsasB based on their similarity to previously characterized enzymes. Results of the construction of a phylogenetic tree are shown in figure 3.10.

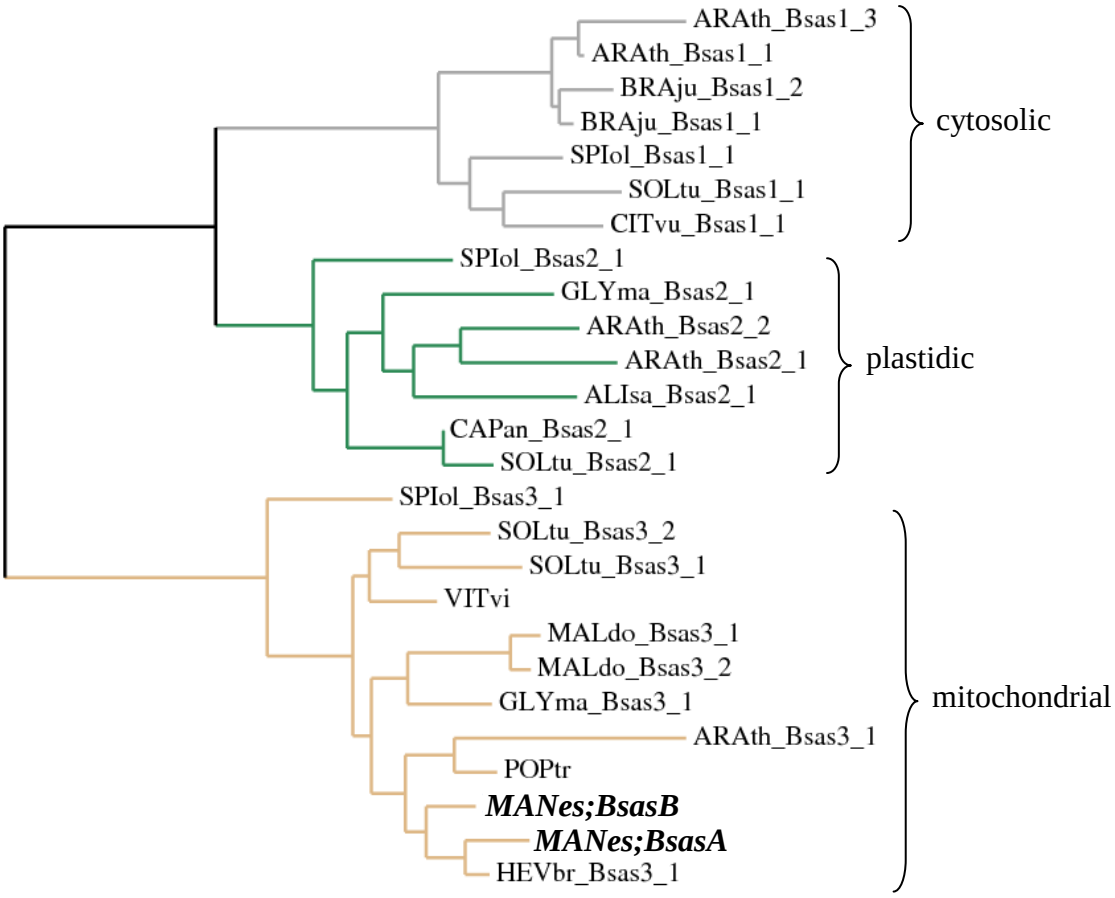

Figure 3.10: Phylogenetic analysis of 24 *Bsas* protein sequences with MANes;BsasA and MANes;BsasB. A multiple sequence alignment using ClustalW was first done, in order to provide the cluster analysis required by the *Phylogeny fr*. tool using default parameters (Dereeper et al., 2008). The accession numbers of the sequences used for this analysis are detailed in Appendix C. Clustered branches are identified by grouped isoforms: cytosolic Bsas1, chloroplastic Bsas2 and mitochondrial Bsas3. The sequences POPtr and VITvi have not been characterized neither *in silico*, nor kinetically**.** The MANes;BsasA and MANes;BsasB sequences are shown in "**Bold**" and "*Italics*" to facilitate observation.

According to the phylogenetic tree in figure 3.10, three main clusters can be observed when using this set of 24 sequences. Each cluster is composed of cytosolic Bsas1, chloroplastic Bsas2 or mitochondrial Bsas3 proteins. This information clarifies that, based on sequence similarity, the *Bsas* enzymes can be classified into these three groups, in agreement with the published work of Hatzfeld et al. (2000). In addition, the two clusters comprised of cysteine synthases are closer together than to the sequences clustered among the mitochondrial cyanoalanine synthases.

Both MANes;BsasA and MANes;BsasB sequences are observed within the mitochondrial proteins grouped in the Bsas3 cluster, thus providing insights about the possibility of both enzymes encoding β-CAS sequences (see figure 3.10). Interestingly, even though MANes;BsasA and MANes;BsasB show short distances within the tree, MANes;BsasA is closer to the rubber tree β-CAS (HEVbr;Bsas3;1), which is also known to be a cyanogenic plant. However, the HEVbr;Bsas3;1 enzyme has not been characterized *in vitro* or *in vivo*; rather, it has been only characterized using sequence similarity.

This information strengthens the hypothesis of both enzymes being mitochondrial. However, clearer information regarding their actual localization must be achieved using tools for this purpose and further kinetic analyses.

3.2.2 *In silico* determination of the sub-cellular location of MANes;BsasA and MANes;BsasB

The development of sequence analysis tools has provided additional ways to analyze biological enzymes, based on comparative approaches to previously studied proteins. In this way, the structural organization of a protein can be analyzed by the comparison of its amino acid

sequence to a library containing protein sequences of a diverse range of protein structures and functions.

Even though signal peptides are not enclosed in conserved protein domains, certain structural features have been identified in plant signal peptides that allow their import to specific compartments (Murcha et al., 2005; Sjöling and Glaser, 1998). Computer algorithms have been designed for the identification of these features in new sequences, given a library of well-characterized molecules. TargetP (Emanuelsson et al., 2000) is one of these programs that has been widely used among scientists aiming the *in silico* characterization regarding the sub-cellular location of newly isolated protein sequences (Subbaiah et al., 2006; Watanabe et al., 2008a). An additional tool used for this purpose is Predotar (Small et al., 2004). These two tools were employed for the determination of the sub-cellular localization of the MANes;BsasA and MANes;BsasB proteins. The set of characterized enzymes previously used for the phylogenetic analysis was also used, in order to provide information leading to the level of certainty to which these programs can estimate N-terminal targeting sequences in plants. Results are shown in Table 3.1.

According to TargetP, all the cytosolic and chloroplastic *Bsas* enzymes were successfully identified as previously reported by each group involved in their characterization. Only the soybean Bsas3 (GLYma;Bsas3;1), which was originally identified and submitted as a β-CAS-encoding enzyme (accession ABQ88337) showed a result compared to chloroplastic enzymes. Moreover, this enzyme was also showed to be chloroplastic using Predotar. Also when using this program, one of the two plastidic isoforms of Arabidopsis was shown to be mitochondrial, but this enzyme has been confirmed experimentally to encode a mitochondrial CS (Watanabe et al., 2008a). This information demonstrates the reliability of these tools in estimating sub-cellular

localization, considering that only together with the phylogenic tree and kinetic studies, a complete characterization of the enzymes will be achieved.

Table 3.1: *In silico* analysis of a set of 24 *Bsas* enzymes together with the isolated MANes;BsasA and MANes;BsasB.

Sequence	TragetP				Predotar			
	cTP	mTP	Other	Location	Plastid	Mitoch.	Elsewhere	Predict.
ARAth;Bsas1;1	0.091	0.073	0.819	-	0.01	0.01	0.98	-
ARAth;Bsas1;3	0.094	0.125	0.731	-	0.02	0.01	0.95	-
BRAju;Bsas1;1	0.105	0.07	0.824	-	0.00	0.01	0.98	-
BRAju;Bsas1;2	0.119	0.074	0.753	-	0.03	0.01	0.97	-
CITvu;Bsas1;1	0.048	0.068	0.886	-	0.00	0.01	0.99	-
SPIol;Bsas1;1	0.059	0.064	0.808	-	0.00	0.01	0.98	-
SOLtu;Bsas1;1	0.053	0.066	0.869	-	0.00	0.01	0.99	-
CAPan;Bsas2;1	0.92	0.027	0.117	C	0.86	0.01	0.14	C
ARAth;Bsas2;1	0.938	0.132	0.026	C	0.93	0.04	0.06	C
ARAth;Bsas2;2	0.744	0.473	0.006	C	0.03	0.33	0.61	~M
SOLtu;Bsas2;1	0.859	0.048	0.238	C	0.90	0.01	0.10	C
SPIol;Bsas2;1	0.747	0.077	0.098	C	0.85	0.03	0.14	C
ALIsa;Bsas2;1	0.836	0.491	0.01	C	0.94	0.32	0.04	C
GLYma;Bsas2;1	0.959	0.139	0.01	C	0.97	0.15	0.02	C
MANes-BsasA	***0.405***	***0.397***	***0.064***	***C***	***0.16***	***0.72***	***0.23***	***M***
MANes-BsasB	***0.261***	***0.648***	***0.042***	***M***	***0.00***	***0.49***	***0.50***	***~M***
HEVbr;Bsas3;1	0.152	0.579	0.082	M	0.00	0.56	0.44	M
SOLtu;Bsas3;1	0.077	0.721	0.087	M	0.00	0.86	0.13	M
SOLtu;Bsas3;2	0.038	0.785	0.123	M	0.00	0.64	0.36	M
GLYma;Bsas3;1	0.365	0.276	0.045	C	0.96	0.38	0.02	C
MALdo;Bsas3;1	0.235	0.828	0.016	M	0.00	0.73	0.26	M
MALdo;Bsas3;2	0.103	0.838	0.015	M	0.00	0.6	0.39	M
ARAth;Bsas3;1	0.095	0.836	0.052	M	0.00	0.51	0.48	M
SPIol;Bsas3;1	0.586	0.784	0.007	M	0.35	0.5	0.30	M
POPtr	0.339	0.354	0.116	M	0.03	0.25	0.72	M
VITvi	0.227	0.534	0.012	M	0.00	0.77	0.23	M

Notes to table 3.1: The analysis was done using TargetP and Predotar tools. The output score values of TargetP do not refer to probabilities, nor do they actually add to 1, but the highest score is the most likely one to tell about the protein's location. Predotar output values are probability estimates as to whether the sequence contains a mitochondrial, plastidic or none targeting sequence. Classification sub-tittles (i.e. "cTP", "mTP", "Plastid", "Mitoch.", etc) in each column were kept as given by each estimating tool. Legend: "cTP", chloroplast targeting peptide; "mTP", mitochondrion targeting peptide; "Other", mostly cytosolic; "C", Chloroplast; "M", Mitochondrion; "~M", probably Mitochondrion. The MANes;BsasA and MANes;BsasB sequences are shown in "**Bold**" and "*Italics*" to facilitate observation.

Using the TargetP program (Table 3.1), MANes;BsasA was shown to be chloroplastic and MANes;BsasA mitochondrial. However, the scores obtained are not as high as in other

enzymes within the group shown to be located in either compartment, such as ARAth;Bsas2;1, SOLtu;Bsas2;1 and SPIol;Bsas2;1 for plastidic enclosed enzymes, and ARAth;Bsas3;1, SPIol;Bsas3;1 and SOLtu;Bsas3;1 for mitochondrial enzymes, which are the most characterized. A different result was observed when the Predotar tool was used. In that case MANes;BsasA showed high probability to be mitochondrial, while no certain result could be obtained for Manes;BsasB. There was a 0.49 probability that MANes;BsasB is targeted to the mitochondrion and a similar probability (0.50) that it is targeted elsewhere (see Table 3.1). This result cannot be taken as a chance of MANes;BsasB to be targeted to the cytoplasm, since cytoplasmic enzymes from the set of sequences used were readily identified as cytosolic by Predotar. Possible difficulties in the accurate prediction of the sub-cellular location of MANes;BsasA and MANes;BsasA by these tools might be due to a possible lack of enough cassava proteins within the databases employed by these tools. Further kinetic evidence will provide deeper insights about the actual identity of these novel proteins from cassava.

3.2.3 *In vitro* kinetic properties of MANes;BsasA and MANes;BsasB

The characterization of any enzyme has been commonly achieved by determination of the kinetic properties of purified protein fractions obtained using methods such as ammonium sulfate precipitation, differential centrifugation and different types of liquid chromatography, being size exclusion, affinity and ion exchange chromatography amongst the most employed. More recently, molecular biology and functional genomics have provided approaches to enzyme characterization in which a DNA construct resembling the gene that encodes the enzyme in question is tested for its ability to compensate for the lost of function in a mutant organism lacking that particular gene. Hulanicka et al. (1986) reported a *Escherichia coli* cysteine auxotroph that was shown to be negative for cysteine synthase activity. This strain was originally

developed and used for cloning and characterization of the *cysAMK* operon required for cysteine synthesis in *Salmonella typhimurium*. Years later Hell et al. (1994) used this mutant strain for the first time in the characterization of *Bsas* enzymes, reporting the actual function of two cDNAs encoding the cytosolic and chloroplastic CS isoforms. A more complete study was achieved by Hatzfeld et al. (2000), which used the NK3 mutant strain for further characterization of other spinach and Arabidopsis *Bsas* isoforms. Maruyama et al. (2001) used it for characterization of two CS and two β-CAS enzymes from potato.

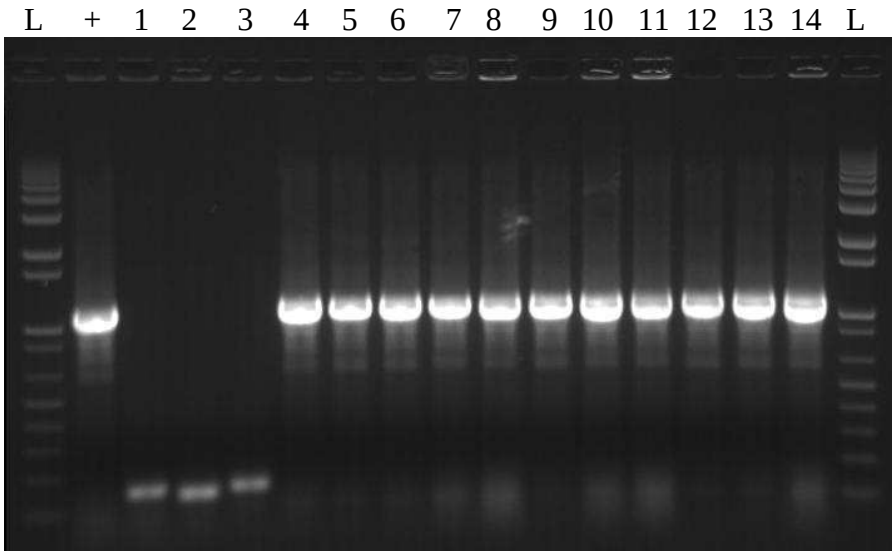

Figure 3.11: Colony PCR of MANes;BsasA and MANes;BsasB in pTV118N. Primers used annealed to regions flanking the multiple cloning site (pTV-F and pTV-R, see appendix A, table A.2). For complete details of addition of restriction sites, digestion and ligations of DNA molecules, as well as bacterial transformation, see section 2.2.2 and appendix A. Lanes 1-5, MANes;BsasA; Lanes 6-14, MANes;BsasB; "+", positive control, ARAth;Bsas3;1.

The aforementioned studies cloned the *Bsas*-encoding cDNAs in the pTV118N (Takara Bio, Shiga, Japan) vector, in which the cloning of genes with a start codon at exactly 8 bp downstream from the Shine Dalgarno (SD) sequence is guaranteed with the inclusion of the *Nco*I restriction site in the forward primer, exactly at the start codon, since this restriction site provides a start codon (cc**atg**g). This positioning of the gene assures maximum level of expression using

the *lacZ* promoter, whose expression is increased by induction using IPTG. The MANes;BsasA and MANes;BsasB cDNAs were cloned in this vector (see figure 3.11), as well as the ARAth;Bsas3;1 (which together with the NK3 mutant was kindly provided by Dr. Kazuki Saito, Chiba University, Chiba, Japan). The Arabidopsis β-CAS was used as a positive control. The reverse primer contained a restriction site for *Xba*I (for a detailed list of primers used see table 2 of appendix A). A fourth clone resembled the pTV118N empty vector, serving as a negative control.

After growth and induction of NK3 cultures harboring any of the four clones, cells were lyzed by sonication, centrifuged and the supernatant was taken as the crude protein extract (Hatzfeld et al., 2000). CAS activity was measured by a colorimetric method that measures the accumulation of sulfide (Hasegawa et al., 1994), while CS activity was achieved according to a cysteine colorimetric method using nin-hydrin (Gaitonde, 1967). Each reaction is shown below.

$$\text{Cysteine} + \text{CN} \longrightarrow \beta\text{-CAS} + H_2S$$

$$\textit{O}\text{-acetyl-serine} + H_2S \longrightarrow \text{Cysteine} + \text{acetate}$$

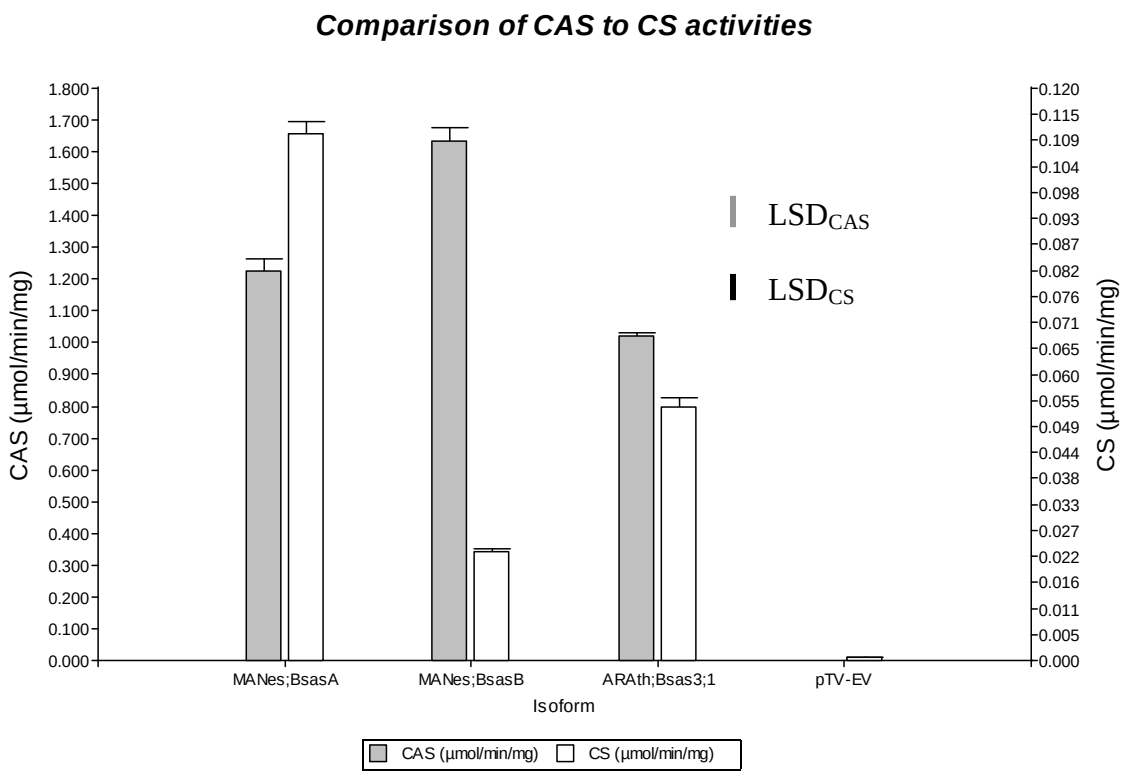

Figure 3.12: Comparison of CAS to CS activity levels of MANes;BsasA and MANes;BsasB proteins over-expressed in the *Escherichia coli* NK3 cysteine auxotroph. Cultures harboring the ARAth;Bsas3;1 and pTV118N empty vector (pTV-EV) were used as controls. CAS activity levels were measured in µmol sulfide/min/mg protein. CS activity levels were measured in µmol cysteine/min/mg protein. Statistical analysis and chart construction were done using InfoStat. All values showed statistical differences using ANOVA test ($\alpha = 0.05$), including Arath;Bsas3;1 and pTV-EV in both assays (see Appendix D for detailed statistical test scores). Bars depict standard error. LSD, Least Significant Difference ($\alpha = 0.05$; $LSD_{CAS} = 0.09504$; $LSD_{CS} = 0.00580$).

Summarized information regarding enzyme activity assays for the MANes;BsasA and MANes;BsasB proteins is shown in figure 3.12. As can be observed, the MANes;BsasB isoform showed the highest β-CAS activity, even than the ARAth;Bsas3;1 protein, which has been previously reported to be capable of mainly catalyzing this reaction (Hatzfeld et al., 2000; Watanabe et al., 2008). Moreover, the MANes;BsasA enzyme also showed higher β-CAS activity than ARAth;Bsas3;1, suggesting both genes encode β-CAS enzymes. All the β-CAS activity values for each enzyme showed to be statistically different (figure 3.12 and appendix D).

The levels of the NK3 mutant strain in catalyzing the β-CAS reaction (pTV-EV in figure 3.12 and table 3.2) were barely detectable and actually a negative value was obtained for the mean value.

Since the *Bsas* enzymes have been shown to be multifunctional, characterized by their ability to catalyze the synthesis of cysteine and as well as of other β-substituted alanines (Ikegami and Murakoshi, 1994), the novel cassava isoforms were tested for their ability to catalyze the CS reaction. In addition, is has been suggested that in any study reporting the isolation of any of the *Bsas* members, a multifunctional characterization should be addressed, at least on the ability of these enzymes in synthesizing cysteine and, in this case, β-cyanoalanine (Warrillow and Hawkesford, 1998). Summarized results about the effectiveness of the MANes;BsasA and MANes;BsasB compared to ARAth;Bsas3;1, in catalyzing the CS reaction are also shown in figure 3.12. Interestingly, MANes;BsasA showed remarkably higher CS activity levels than ARAth;Bsas3;1, but even a more drastic difference can be observed when compared to the levels to which MANes;BsasB can catalyze this reaction. Statistical differences were also observed between all isoforms (see appendix D), including the NK3 strain, which as expected, showed no CS activity (see figure 3.12 and table 3.2).

Table 3.2: Comparison of β-CAS to CS activities of MANes;BsasA and MANes;BsasB proteins over-expressed in the *Escherichia coli* NK3 cysteine auxotroph.

Isoform	CAS	CS	CAS/CS	CS/CAS
	μmol H_2S min^{-1} mg^{-1} protein	μmol Cys min^{-1} mg^{-1} protein		
MANes;BsasA	1.2256	0.1105	11.091	0.090
MANes;BsasB	1.6333	0.0227	72.011	0.014
ARAth;Bsas3;1	1.0233	0.0533	19.207	0.052
pTV-EV	-0.0358	0.0005	-	-

These results suggest that MANes;BsasB might have a main role in cyanide detoxification, or at least higher than the MANes;BsasA enzyme. In addition, CAS/CS and

CS/CAS activity ratios were calculated (see table 3.2), since this have provided a simple way to test the actual roles of *Bsas* isoforms in either catalyzing the CAS or the CS reaction (Hatzfeld et al., 2000; Maruyama et al., 2001). These results show that the MANes;BsasB isoform had a 6.5 fold CAS/CS ratio, compared to MANes;BsasA, while compared to ARAth;Bsas3;1 it had a 3.75 fold CAS/CS ratio. On the other hand, the CS/CAS ratio for MANes;BsasA was 6.4 fold compared to MANes;BsasB, while an almost two fold was observed compared to ARAth;Bsas3,1. These results clearly demonstrate that based on their kinetic properties, MANes;BsasA has a main role in cysteine biosynthesis, while MANes;BsasB is more likely to encode a true β-CAS enzyme.

3.3 *In vivo* testing of the role of *Bsas* enzymes in cyanide metabolism in cassava

The previous report about the presence of β-CAS activity in cassava protein extracts (Elias et al., 1997a) did not address information regarding which genes and/or isoforms are actually contributing to cyanide assimilation. Even though significant differences between MANes;BsasA and MANes;BsasB in cyanide assimilation were found *in vitro* (see previous section), the unclear results concerning their sub-cellular location using software tools (see section 3.2.2) makes the analysis of these genes using biotechnological tools an intriguing approach. In other words, this study also used plant transformation experiments in cassava to aim for information regarding the role of the over-expression of these enzymes in cyanide detoxification within the plant. In this sub-chapter the MANes;BsasA, MANes;BsasB and ARAth;Bsas3;1 were cloned in a *Agrobacterium* binary vector and used for cassava transformation of germinated somatic embryos (GSEs).

3.3.1 Construction of plant transformation vectors and transformation of the LBA4404 *Agrobacterium tumefaciens* strain

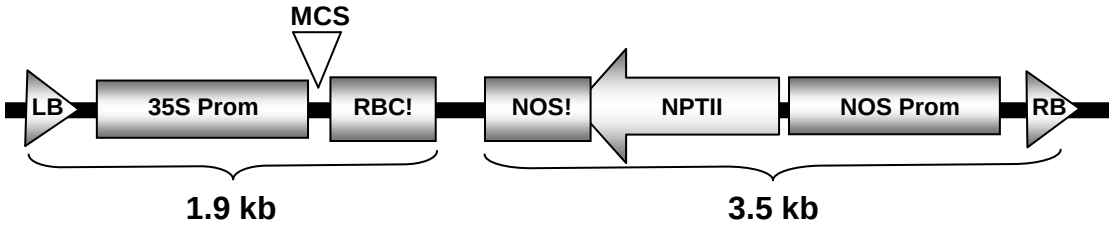

Figure 3.13: Schematic representation of the T-DNA region of pKYLX:35S^2 binary vector. The disarmed Ti-plasmid contains a cassette for cloning genes under the control of the constitutive promoter of the cauliflower mosaic virus 35S ribosomal subunit. Selection of transgenic plantlets can be done using the neomycin phosphotransferase II gene (*nptII*) that confers resistance to amino-glycoside antibiotics. Legend: LB and RB, left and right borders; 35S Prom, double CaMV 35S constitutive promoter; RBC!, terminator of rubisco small subunit; NOS!, terminator of nopaline synthase; NOS Prom, nopaline synthase promoter; MCS, multiple cloning site: *Hin*dIII*, *Bam*HI, *Xho*I*, *Pst*I, *Sac*I*, *Xba*I* ("*", unique sites).

The pKYLX71:35S^2 binary vector (figure 3.13; Schardl et al., 1987) was used, since previous reports have used this vector for cassava transformation (Siritunga et al., 2004). This vector provides constitutive expression of genes under the control of the cauliflower mosaic virus 35S promoter. Primers were designed to amplify the MANes;BsasA, MANes;BsasB and ARAth;Bsas3;1 genes with forward primers resembling *Sac*I sites and start codons at amino acid positions 37, 30 and 28, respectively (see primers ManesA-Sac, ManesB-Sac and Arath-Sac in table A.2 of appendix A). An *Xba*I site was included in the reverse primer, which also had the naturally occurring stop codon of each isoform (see primers Manes-Xba and Arath-Xba in table A.2 of appendix A). These engineered versions of the proteins were expected to assure the translation of the mature forms of the protein, since most N-terminal encoded mitochondrial signal peptides are cleaved after protein import to the mitochondria, hence not contributing to the mature structure of the active enzyme (Sjöling and Glaser, 1998). Besides, a higher reduction of

cyanide levels within the plant might be expected if the over-expression of the protein is achieved, not only in plant mitochondria, but also distributed within all the cytoplasm, which is where the protein will be delivered when lacking a targeting sequence.

Figure 3.14 shows the colony PCRs of the three constructs developed for this study. At least four positive colonies (except from the 35S:ARAth;Bsas3;1 construct, which only gave one positive colony) were picked and grown for plasmid isolation and sequenced using the Nevada Genomics Center Facility (University of Nevada, Reno, Nevada).

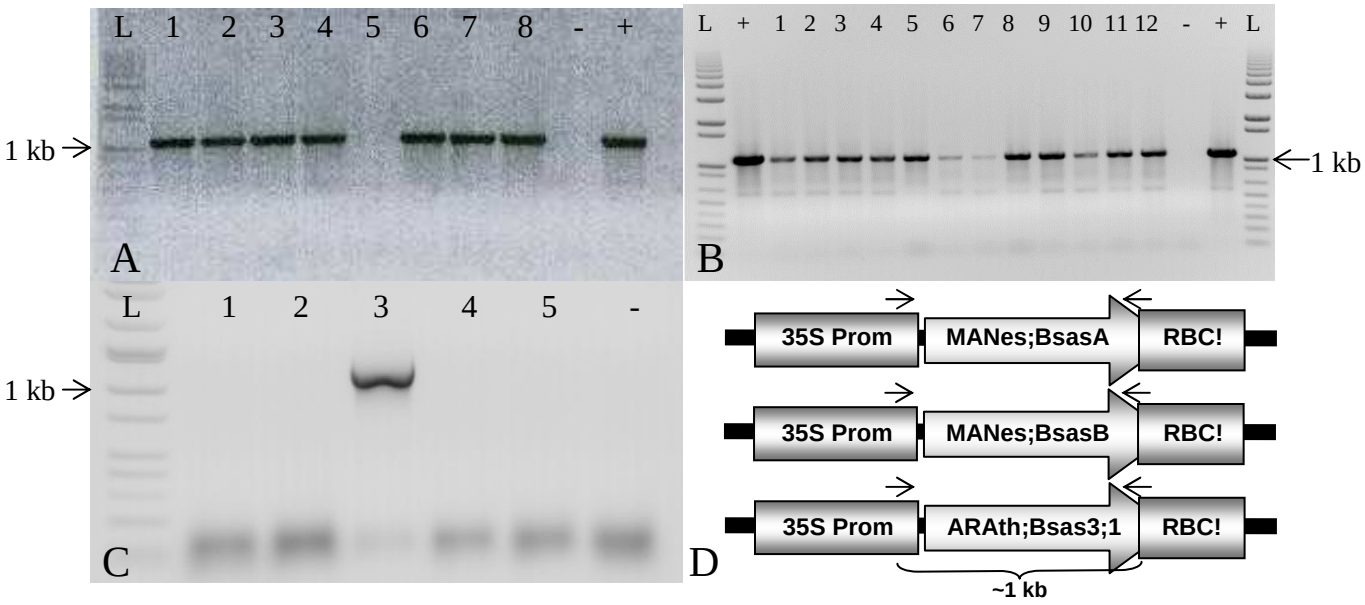

Figure 3.14: Colony PCR of MANes;BsasA, MANes;BsasB and ARAth;Bsas3;1 genes cloned in pKYLX binary vector (1 % agarose on 1X TAE). The screening was done using a forward primer that anneals to the downstream region of the CaMV 35S[2] promoter and a reverse primer specific for each gene (see table A.2 of appendix A). A, screening of 35S:MANes;BsasA colonies; B, screening of 35S:MANes;BsasB colonies; C, screening of 35S:ARAth;Bsas3;1 colonies; D, scheme of each over-expression construct showing approximate annealing sites of primers used and amplified PCR products sizes.

Clones that showed no nucleotide mutations were used for transformation of *Agrobacterium tumefaciens* LBA4404 by electroporation. This strain contains the virulence genes used for T-DNA transfer and integration into the plant's genome in a second plasmid that also contains the gene conferring streptomycin resistance. The pKYLX vector contains the

tetracycline resistance gene, thus selection of *Agrobacterium* colonies harboring the binary system must be achieved using these two antibiotics. Results of transformation of the 35S:MANes;BsasA and 35S:ARAth;Bsas3;1 constructs into LBA 4404 electro-competent cells are shown in figure 3.15. The 35S:MANes;BsasA and pKYLX empty vector constructs were also used for *Agrobacterium* transformation (data not shown).

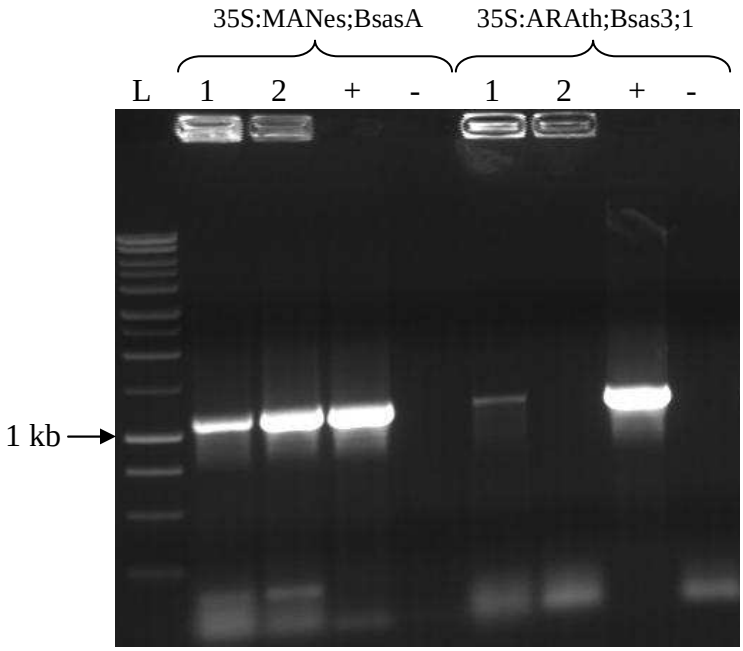

Figure 3.15: Colony PCR screening of *Agrobacterium tumefaciens* LBA 4404 strain harboring either the 35S:MANes;BsasA or the 35S:ARAth;Bsas3;1 constructs (1% agarose in 1X TAE). Primers used were the same as those employed in figure 3.14 for each construct.

3.3.2 Transformation of cassava germinated somatic embryos and plant selection

Different explants have been reported to be suitable for transformation in cassava, in which *de novo* regeneration of completely new plants has been achieved (Taylor et al., 2004; see sections 1.4.2 and 1.4.3 for a detailed discussion). One of these methods employ germinated somatic embryos developed from meristematic tissue in MS salts containing 2,4-D at a concentration of 8 mg/l (MS8 medium). After embryo induction, regeneration medium (RM2) containing cytokinins and gibberellins was used for shoot induction and development of

elongated cotyledons. This tissue, when taken as explants for further acquisition of somatic embryos, shows higher embryogenic properties than leaf apical meristems. In our laboratory, this cyclic embryogenic system was established using the highly embryogenic cultivar Mcol 2215 (see figure 3.16).

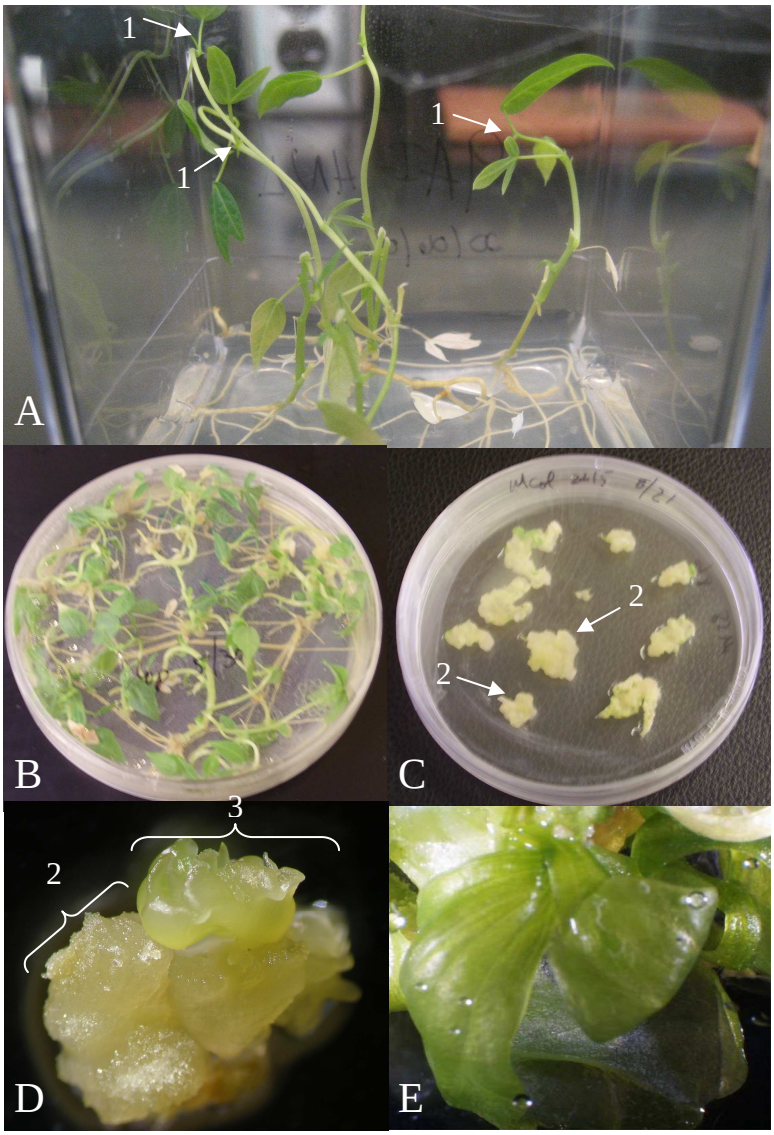

Figure 3.16: Somatic embryogenesis in the Colombian cultivar Mcol2215. A and B, apical leaf meristems (1) are taken from *in vitro* cassava plants propagated in MS (Murashige and Skoog, 1962) salts and vitamins; C and D, after four weeks of cultivation of meristematic tissue in MS salts supplemented with 2,4-D, white hard callus (2) proliferation is induced, from where organized embryogenic structures (3) form; E, somatic embryos are then germinated in MS salts containing cytokinins and gibberellins, which trigger the formation of elongated cotyledonary leaves, suitable for regeneration and/or transformation.

Even though three constructs were prepared for plant transformation, only two constructs were actually used due to the complexity of cassava transformation, since transformation efficiency in cassava is about 1-3%, considering the amount of transgenic lines obtained per amount of explants used (Ihemere, 2003; Siritunga, 2002). In this way, for obtaining at least 25 lines with 1% transformation efficiency, 2,500 explants must be obtained from somatic embryogenesis and transformed. The establishment of cyclic somatic embryogenesis took a period of acclimation of Mcol 2215 plants to the starting tissue culture facilities employed for this project, which was combined with a period of initial training on tissue culture techniques and cassava micropropagation. Hence, even though multiplication of plants and callus induction in *in vitro* conditions was started in June 2006, it was not until February 2007 that the first transformation was done using the construct resembling the ARAth;Bsas3;1 gene (see table 3.3).

Table 3.3: Transformation of explants from GSEs using either the 35S:MANesA (MANes;BsasA) or the 35S:ARAth (ARAth;Bsas3;1) constructs. The number of explants used for co-cultivation with any of the two constructs is shown. Stars depict transformation events that lead to the production of lines that endured selection in paromomycin.

Date	Construct	Transformation event	Number of explants	Time in selection medium
13-Feb-07	35S:ARAth	1	160	-
15-May-07	35S:ARAth	2	203	-
15-Aug-07	35S:ARAth	3	361	-
16-Aug-07	35S:ARAth	4*	231	4 months
16-Oct-07	35S:ARAth	5*	395	4 months
19-Oct-07	35S:ARAth	6	500	-
24-Oct-07	35S:ARAth	7*	194	4 months
24-Nov-07	35S:ARAth	8*	344	not recorded
26-Nov-07	35S:ARAth	9*	418	not recorded
6-Mar-07	35S:MANesA	1	70	-
2-May-07	35S:MANesA	2	184	-
28-May-07	35S:MANesA	3	143	-
21-Aug-07	35S:MANesA	4	197	-
27-Aug-07	35S:MANesA	5*	367	5 months, 3 weeks
17-Oct-07	35S:MANesA	6*	456	4 months
23-Oct-07	35S:MANesA	7*	515	4 months
16-Jan-08	35S:MANesA	8*	581	5 months, 3 weeks

Table 3.3 shows the amount of explants used for co-cultivation with LBA4404 harboring either the 35S:MANesA construct or the 35S:ARAth construct. As can be appreciated, transformations were done by alternating the two constructs, so that a paired number of putative transgenic lines from both constructs growing in selective medium were obtained through time. After co-cultivation for two days, the explants were transferred to MS8 medium containing paromomycin and carbenicillin for selection of plants and killing *Agrobacterim*, respectively. Organized embryos formed in selective medium were then germinated using RM2 medium containing the aforementioned antibiotics. The selection period extended to not less than four months {see table 3.3).

Table 3.4: Determination of transformation efficiencies for the 35S:MANesA and 35S:ARAth constructs. Detailed information regarding the transformation event number 9 is given. TE, Transformation Efficiency calculated by the following equation: TE (%) = (Lines/Explants)*100.

Construct (transformation event)	Explants	Lines	TE (%)
35S:ARAth (all)	2,806	80	2.9
35S:MANesA (all)	2,513	212	8.4
35S:MANesA (1-8)	1,932	94	4.9
35S:MANesA (9)	581	118	38.2

In table 3.4, data concerning transformation efficiency is presented. The 35S:ARAth construct showed levels of transformation efficiency compared to previous studies in which somatic embryogenesis was used for transformation of Mcol 2215 (Siritunga, 2002; Ihemere, 2003). An interesting result could be appreciated when considering the transformation efficiency for 35S:MANesA construct, since it obtained 8.4%. This value was mainly due to a high amount of lines obtained from the co-cultivation of the last 581 explants, which was abnormally higher, showing to have 38.2% transformation efficiency. Despite values closer to normal efficiency

were obtained in the first 8 transformation events, in which a total of 1,932 explants were transformed, they were still too high, showing 4.9%.

3.3.3 PCR screening of transgenic plants

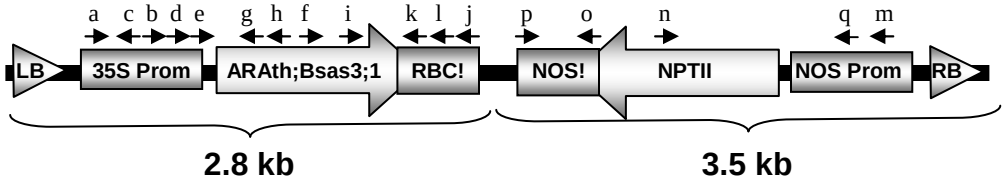

Figure 3.17: Approximate annealing sites of primers designed for PCR screening of putative transgenic 35S:ARAth plants. See table A.3 of appendix A for a detailed description of each primer.

Due to the apparent presence of escapes within the population comprised of putative transgenic plants developed using the 35S:MANes construct, PCR screening of only the plants developed with the 35S:ARAth construct was done. After micropropagation of putative transgenic plants, DNA was extracted from leaves of four month old *in vitro* plants using a method derived from Dellaporta et al. (1983). Primers designed for screening of putative transgenic plants and their approximate priming sites are shown in figure 3.17 (see Table A.3 of appendix A for a detailed description). Summarized results of all the PCR reactions done for the screening are shown in appendix E. Only several particular results will be discussed in detail.

As can be observed in tables E.1, E.2 and E.3 of appendix E, no line was observed to show amplification from both the selection (*nptII*) and the 35S:ARAth cassettes included within the T-DNA region of pKYLX, except several cases that did not show a repetitive pattern of amplification. An interesting result was obtained when amplifying the downstream region of the 35S:ARAth construct: some lines showed to amplify using a primer pair that hit the 35S:ARAth gene (see figure 3.18 and table E.3, reaction i/Arath-Xba).

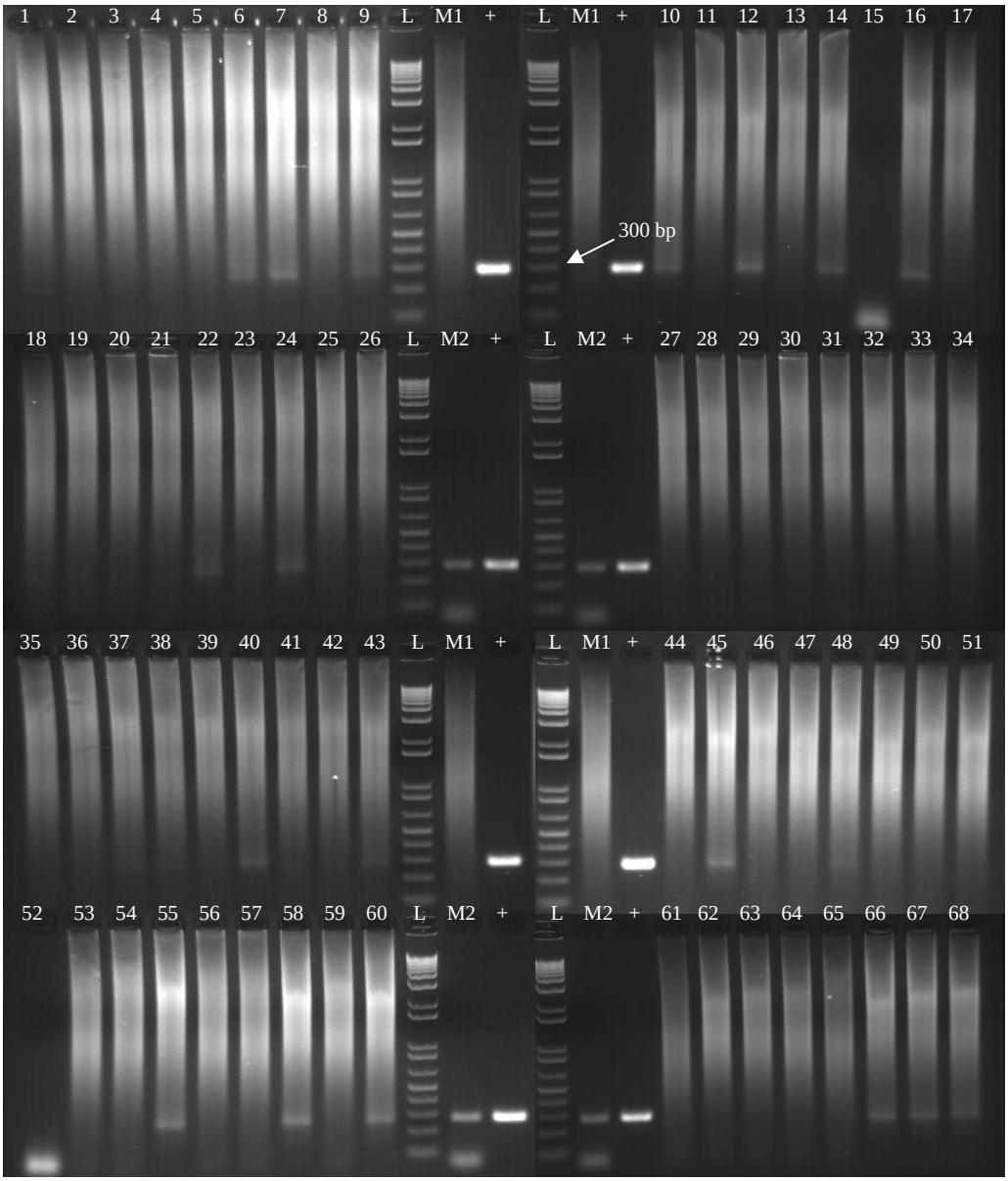

Figure 3.18: PCR screening of putative 35S:ARAth plants using primer "i" as forward primer (figure 3.17; table A.3) and Arath-Xba as reverse primer (table A.2) (1.5% agarose on 1X TAE). The expected amplified PCR product was of about 300 bp long. These results are summarized in table E.3 of appendix E. "M1" and "M2", amplifications from two different DNA samples of Mcol 2215 wild type plants; "+", amplification from pKYLX 35S:ARAth plasmid.

The reverse primer of this PCR reaction was supposed to amplify from a short portion of the Arabidopsis β-CAS un-translated region included within the construct. However, as can be observed in figure 3.18 this primer set amplified the M2 sample, which corresponded to genomic DNA extracted from Mcol 2215 wild type plant.

To check whether this result was or was not a false positive, amplification from five different DNA samples extracted from Mcol 2215 wild type plants was achieved. The same PCR reaction was done, also using a plasmid resembling the cDNA sequence of the endogenous MANes;BsasB gene. Results are shown in figure 3.19. As can be observed, this primer pair showed to amplify the same sample at irregular intervals, since all reactions were done in triplicates from the Mcol 2215 samples and amplification was not always observed. Moreover, the primer pair showed to hit the MANes;BsasB gene. In addition, this reaction was done several times (see tables E.1 and E.2 of appendix E) and samples showing amplification with this primer pair did not amplify with many other different primer sets, thus the discussed result might be a false positive.

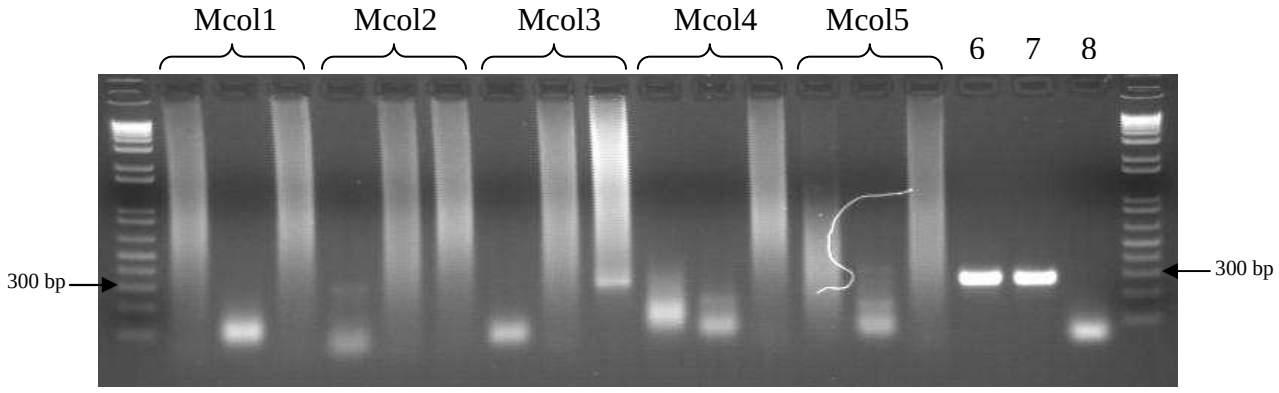

Figure 3.19: Amplification from different Mcol 2215 wild type plants using primer set i/Arath-Xba (1.5 % agarose in 1X TAE). See tables A.2 and A.3 of appendix A for details concerning primers. Each DNA sample (Mcol1, Mcol2, etc.) was amplified in triplicates. Lanes 6, 7 and 8 were amplifications from MANes;BsasB, pKYLX 35S:ARAth plasmid and negative control, respectively.

A second interesting result is shown in figure 3.20. In this case, amplification using primers hitting the NOS promoter and *nptII* gene were employed. This result showed that when using this primer pair, seven lines showed amplification. However, when a different set of primers was used for amplification of the selective cassette, amplification was not obtained (see figure 3.21 and table E.3.)

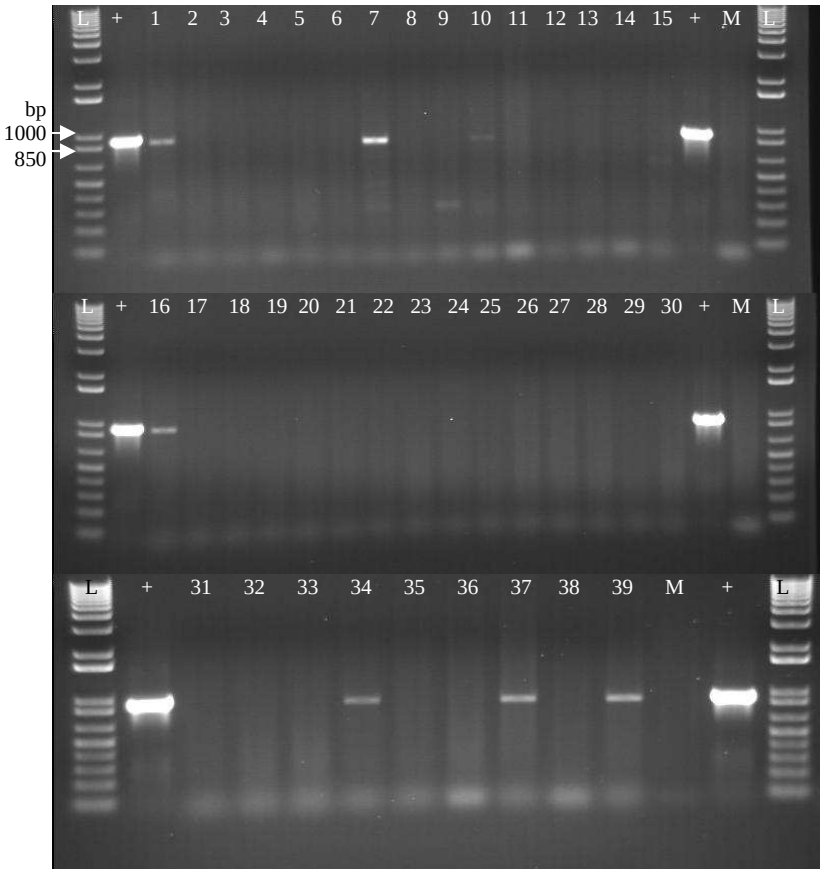

Figure 3.20: PCR screening of putative 35S:ARAth lines with the primer pair m/n, amplifying the *nptII* selective cassette (1.5 % agarose in 1X TAE). The expected amplified PCR product had a molecular weight of about 900 bp. These results are summarized in table E.3 of appendix E. See figure 3.17 and table A.3 for details concerning each primer. "M", Mcol 2215 wild type plant; "+", pKYLX 35S:ARAth plasmid.

The above (figure 3.20) and below (figure 3.21) results, even though contradictory, are in accordance to previous results showing amplification of close-related bands from this particular genetic background. Siritunga (2002) showed that when employing primers to amplify the *nptII* gene from transgenic plants developed from Mcol 2215, a non-identified endogenous DNA amplicon was obtained. Moreover, when employing primers that could hit the NOS terminator sequence, no amplification was obtained from this same set of samples (see reaction o/p from table E.3). When the first attempts of amplification of the NOS terminator were done, two lines showed a positive result (see table E.1). However, no line showed a repetitive pattern of

amplification, thus there was no molecular evidence of the presence of the T-DNA within their genome.

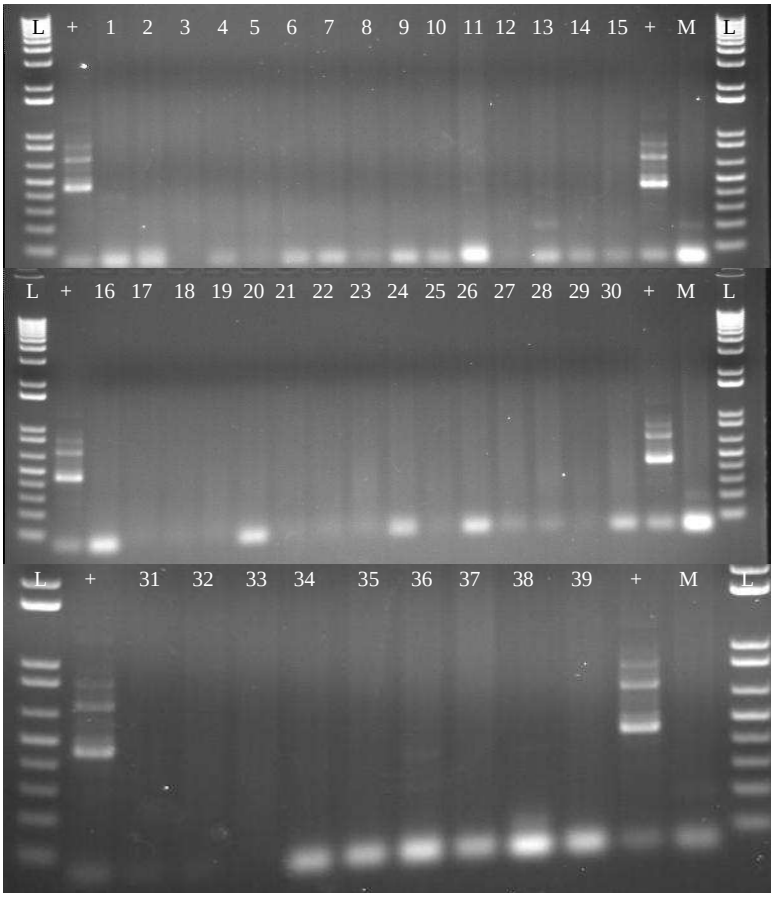

Figure 3.21: PCR screening of putative 35S:ARAth lines with the primer pair q/n, amplifying the *nptII* selective cassette (1.5 % agarose in 1X TAE). The expected amplified PCR product had a molecular weight of about 450 bp. These results are summarized in table E.3 of appendix E. See figure 3.17 and table A.3 for details concerning each primer. “M”, Mcol 2215 wild type plant; “+”, pKYLX 35S:ARAth plasmid.

PCR is not considered a concrete prove of the presence of a transgene in a genome, mainly due to non-specific amplification to which this technique is prompt to. Concrete prove of transgenesis can only be achieved using blotting techniques, such as southern blot, which can also provide information regarding insertion copy number. Quantitative Real Time PCR has been suggested and employed as a tool for estimating copy number, but this technique needs further optimization for its employment in copy number determination experiments. However, PCR has

been extensively used for screening purposes due to its easiness compared to other more reliable tools such as southern blot. Once the PCR screening is achieved, the plants are further tested using southern blots that confirm the stable integration of the T-DNA into the plant's genome. In this experiment, PCR was used only for screening purposes and these results do not provide a reliable evidence of the presence of the pKYLX 35S:ARAth T-DNA region within any of the lines developed in this experiment.

3.3.4 Phenotypical analysis: β-CAS activity assays from putative transgenic plants

In order to check whether these plants have the T-DNA integrated in a region that might not be facilitated by PCR amplification, some of the most healthy plants that showed a positive result were propagated, grown for at least three months and checked for increased levels of β-CAS activity. Results are shown in figure 3.22. As can be observed, only lines 42, 581 and 722-1 had activity levels lower than that of the Mcol 2215 wild type plants. All other lines had an activity level not different from Mcol 2215. This result makes stronger the assumption to which the PCR results lead, showing that no evidence was found of these plants being transgenic.

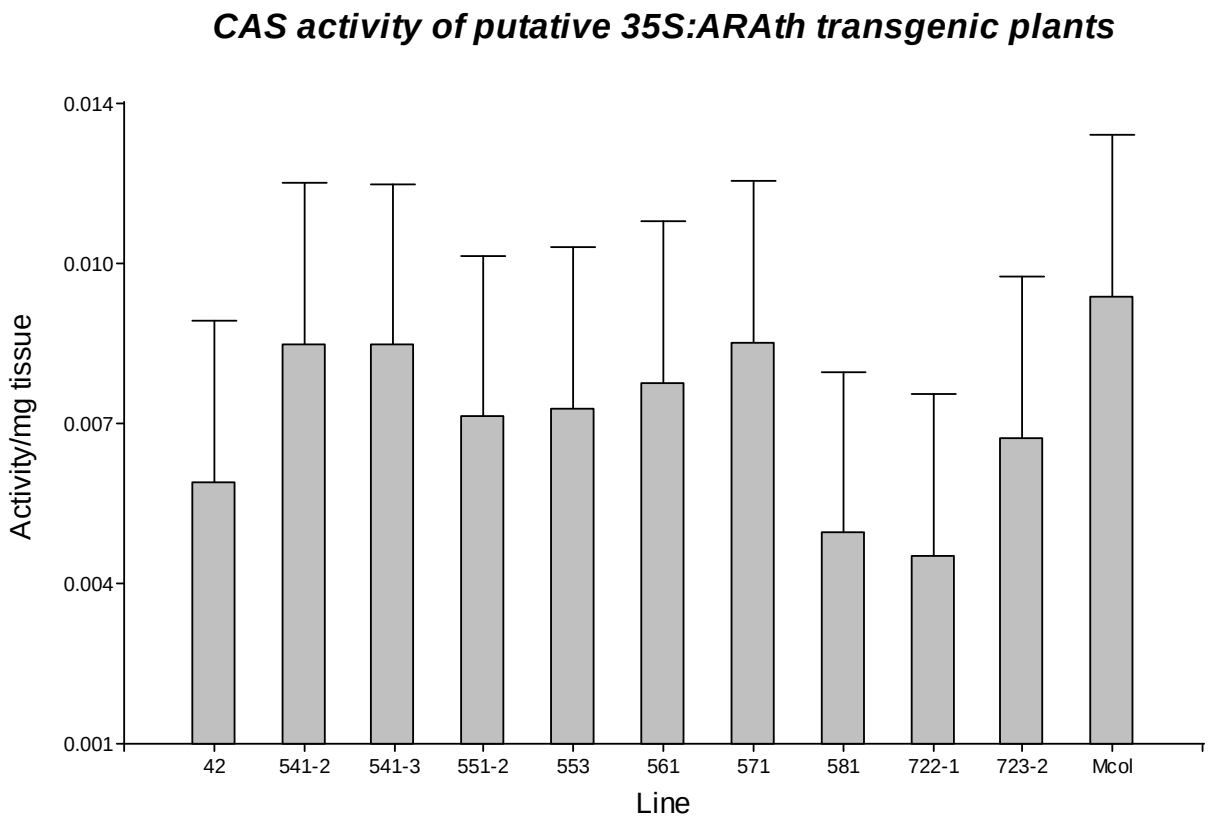

Figure 3.22: β-CAS activity in some putative transgenic 35S:ARAth lines. Activity was measured as μmol/min/mg protein per mg of tissue. "Mcol", Mcol2215 wild type plant. The Analysis of Variace output is shown in Appendix D. Bars depict the size of the least significant difference ($\alpha = 0.05$; LSD = 0.00315).

DISCUSSION AND CONCLUDING REMARKS

4 Discussion and concluding remarks

4.1 Presence of two functional members of the *Bsas* family of enzymes in cassava.

One of the main goals of this project, the cloning and characterization of one or more of the *Bsas* family of genes from cassava in search for the β-CAS gene, has been fully achieved. RACE PCR was used as the PCR-based strategy for gene isolation since it allows amplification of the borders of a cDNA molecule, given the sequence of the internal region. This technique has been previously used for the isolation of different *Bsas* members of potato (Maruyama et al., 2001). Our results show that based on the sizes of the 5' and 3' ends of the cDNA molecules of the *Bsas* enzymes being expressed in cassava leaves, three members could be identified. Two of these are identical in sequence, except for the presence of gaps in both the 5' and 3' un-translated (UTR) regions, which evidence the expression of two versions of the same RNA transcript that differ only in size of the UTR regions. These cDNAs were named MANes;BsasA and both have identical open reading frames. In the other hand, the second functional gene MANes;BsasB had a similar sequence, but was not identical.

Multiple sequence alignment tools were used for clustal analysis that allowed a similarity characterization of the enzymes to previously identified isoforms. These sequence cluster alignments were used for the construction of a phylogenetic tree that evidenced that these two genes are more close together than each one to particular isoforms, except for MANes,BsasA being closer to the rubber tree β-CAS. Phylogenetic analysis show that *Bsas* isoforms can be divided in three groups, depending on sequence similarity (Hatzfeld et al., 2000; this study), these three groups being composed of cytosolic, chloroplastic and mitochondrial enzymes,

respectively. Based on these result both cassava *Bsas* were shown to be within the mitochondrial cluster.

Based on the estimated sub-cellular location, no concrete result was obtained. Two different tools were used for this estimation, though none gave similar results, even though the reliability of the software used was tested by doing the analyses also with other previously characterized enzymes. This result might be due to a lack of cassava enzymes within the data bases used by these programs. The only concrete result that allowed a concrete characterization was the testing of the kinetic properties of these two enzymes when over-expressed in a bacterial background lacking the capability of catalyzing both the β-CAS and CS reactions. These results clearly demonstrate that MANes;BsasB encodes a true β-CAS enzyme, while MANes;BsasA encodes an enzyme that can catalyze better the CS reaction.

These results need to be considered together with the *in silico* analyses results. Sequence alignment shed light about the possible mitochondrial localization of both enzymes. On the other hand, protein sorting using the proper tools is not clear. In Arabidopsis, two mitochondrial enzymes have been identified, each one resembling catalytic properties of β-CAS and a CS (Watanabe et al., 2008a). On the other hand, no true β-CAS has been identified outside the mitochondria; the low cytosolic and chloroplastic β-CAS activities extensively found in plants have been attributed to Bsas1 and Bsas2 cysteine synthases present in these compartments (Warrillow and Hawkesford, 1998; Warrillow and Hawkesford, 2000). Our results regarding the kinetic properties of each isoform indicate that MANes;BsasB is β-CAS and that despite the unclear sub-cellular location estimation, all the published evidence suggests that it should be localized in cassava mitochondria.

As for a concrete characterization of MANes;BsasA, kinetic analyses showed that it is a CS enzyme and any other consideration regarding its sub-cellular location should be taken starting with this proven statement. Moreover, another interesting consideration should be analyzed. In order for a protein to be translated by the ribosomes, a pyridine rich region is needed at 5-15 bp upstream the start codon, for correct positioning of the mRNA transcript (Lewin, 1999). The shorter version of MANes;BsasA lacks this region and, if this transcript is to be translated to a protein, only the third methionine-encoding codon can provide a start codon with the pyridine-rich box upstream the start codon (data not shown). This protein will be 25 amino acids shorter than the longer version of this isoform, providing the cytosolic version of CS which, in Arabidopsis, is the enzyme that mainly contributes to cysteine biosynthesis (Watanabe et al., 2008a).

On the other hand, the longer version of MANes;BsasA could be either targeted to the plastids, serving together with sulfur acetyl transferase in sulfur assimilation into organic molecules, or in the mitochondria, where it could provide a good cysteine supply for cyanide detoxification by MANes;BsasB, which has apparently lost significantly its cysteine synthesis rate, in spite of a higher β-cyanoalanine synthesis rate. An additional option for MANes;BsasA is that somehow, cassava has developed an enzyme that could be targeted to the three discussed compartments: i.e. the long version of the isoform resembles a signal peptide that allows sorting of the enzyme to both mitochondria and plastids, whereas the shorter catalyzes cysteine biosynthesis within the cytosol, as in any other plant. Plant plastidic signal peptides tend to be longer than mitochondrial transit peptides, as can be observed in figure 3.9. Interestingly MANes;BsasA shows a shorter version when compared to plastidic signal peptides, while longer when compared to those of mitochondrial sorting, suggesting maybe that this protein resembles a

targeting sequence for commonly sorting of enzymes to these two compartments. However, this hypothesis needs to be further tested, perhaps using translational fusions of chimeric proteins resembling the signal peptide together with a reporter molecule such as green fluorescent protein.

4.2 Cassava transformation using GSEs needs further optimization.

A possible explanation to have obtained an abnormally high number of putative transgenic plants after selection of the plantlets developed using the 35S:MANes in paromomycin could be the selection methodology. The callus developed in MS8 selective medium was only moved to fresh medium when transferred to regeneration medium for induction of shoots in the organogenesis step, when the developed embryos had at least 8 weeks of culture in the same non-fresh medium in which the callus induction and selection was started. Certain protocols establish that medium containing antibiotics for transgenic plantlet selection must be replaced at least 1-2 weeks (Zhang and Puonti-Kaerlas, 2004), to avoid the overcoming of non-transgenic tissue to these selective conditions. The gene *nptII* encodes the enzyme neophosphotransferase, which was originally isolated from *Escherichia coli*. The enzyme inhibits the lethal action of amino-glycoside antibiotics by addition of a phosphate group, causing inactivation of these molecules which mainly impair protein translation in the ribosomes. It appears unlikely of wild type plants to overcome these conditions, which constitute the functional basis of the utilization of this mechanism for plant selection in transgenic plant regeneration. However, due to the long period of time to which these explants were treated without assuring that fresh antibiotic was administered, it seems plausible that plants might endure such a treatment.

In addition to selection in non fresh medium, plants were regenerated using a treatment that has never been reported. Li et al. (1996) were the first group to report regeneration of

transgenic plants using transformed GSEs, by organogenesis of shoots from callus. This induction of shoots has been reported to be developed in selective medium (Siritunga and Sayre, 2003, 2004; Ihemere et al., 2006). In the protocol employed for the present study, after this period comprised of eight weeks in MS8 selective medium, followed by shoot induction of shoots in RM2 selective medium, the developed embryos were transferred to RM2 medium without the selective agent. This action triggered the formation of shoots in a very marked way. Moreover, this step was administered to the plants developed from the fourth transformation event to the last, which was when plant lines were first observed to develop completely. This action, together with the lack of a proper selection protocol constitutes two of the main reasons for the presence of escapes after apparent selection using paromomycin.

LITERATURE CITED

Allem, A.C. (2002). The origin and taxonomy of cassava. In: Hillocks, R.J., Thresh, J. M., Belloti, A.C. Cassava: Biology, Production and utilization. CAB International. pp. 1-16

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990). Basic local alignment search tool. Journal of Molecular Biology 215: 403-410.

Aregheore, E.M. and Agunbiade, O.O. (1991). The toxic effects of cassava (*Manihot esculenta* Crantz) diets on humans: a review. Veterinary and Human Toxicology 33: 274-275.

Belloti, A.C., Smith, L. and Lapointe, S.L. (1999). Recent advances in cassava pest management. Annual Reviews of Entomology 44: 343-370.

Berkowitz, O., Wirtz, M., Wolf, A., Kuhlmann, J. and Hell, R. (2002). Use of biomolecular interaction analysis to elucidate the regulatory mechanism of the cysteine synthase complex in *Arabidopsis thaliana*. The Journal of Biological Chemistry 277: 30629-30634.

Blumenthal, S., Butler, G.W. and Conn, E.E. (1963). Incorporation of hydrocyanic acid labeled with carbon-14 into asparagine in seedlings. Nature 197: 718-719.

Blumenthal, S.G., Hendrickson, H.R. and Conn, E.E. (1968). Cyanide metabolism in higher plants. III. The biosynthesis of β-cyanoalanine. The Journal of Biological Chemistry 243: 5302-5307.

Bokanga, M. (1994). Processing of cassava leaves for human consumption. Acta Horticulturae (ISHS) 375: 203-207.

Bradford, M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical Biochemistry 72: 248-254.

Buitrago, J.A. (1990). La yuca en la alimentación animal. Centro Internacional de Agricultura Tropical (CIAT). 450 pp.

Calatayud, P.A. and Múnera, D.F. (2002). Defensas naturales de la yuca a las plagas de artrópodos. In: Ospina, B. and Ceballos, H. La yuca en el tercer milenio. CIAT, Colombia.

Calatayud, P.A., Rouland, C. and Le Ru, B. (1996). Interactions between cassava mealybug and cassava: Cytochemical aspects of plant cell wall modifications. Entomologia Experimentalis et Applicata 80: 242-245.

Castric, P.A., Farnden, K.J.F. and Conn, E.E. (1973). Cyanide metabolism in higher plants: V. The formation of asparagine from β-cyanoalanine. Archives of Biochemistry and Biophysics 152 : 62-69.

Chellapan, P., Masona, M.V., Vanitharani, R., Taylor, N.J. and Fauquet, C.M. (2004). Broad spectrum resistance to ssDNA viruses associated with transgene-induced gene silencing in cassava. Plant Molecular Biology 56: 601-611.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G. and Thompson, J.D. (2003). Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Research 31 (13): 3497-3500.

Chiwona-Karktun, L., Tylleskar, T., Mkumbira, J., Gebre-Medhin, M. and Rosling, H. (2000). Low dietary cyanogen exposure from frequent consumption of potentially toxic cassava in Malawi. International Journal of Food Science and Nutrition 51: 33-43.

Ceballos, H. (2002). La yuca en Colombia y el mundo: nuevas perspectivas para un cultivo milenario. In: Ospina, B. and Ceballos, H. La yuca en el tercer milenio. CIAT, Colombia.

Ceballos, H. and De la Cruz, G.A. (2002). Taxonomía y morfología de la yuca. In: Ospina, B. and Ceballos, H. La yuca en el tercer milenio. CIAT, Colombia.

Ceballos, H., Iglesias, C.A., Pérez, J.C. and Dixon, A.D.O. (2004). Cassava breeding: opportunities and challenges. Plant Molecular Biology 56: 503-516.

Claros, M.G. and Vincens, P. (1996). Computational method to predict mitochondrially imported proteins and their targeting sequences. European Journal of Biochemistry 241: 770-786.

Cock, J.H. (1985). Cassava: New potential for a neglected crop. Westview Press, Boulder and London, 191 pp.

Dellaporta, S., Wood, J. and Hicks, J. (1983). A plant DNA minipreparation: Version II. Plant Molecular Biology Reporter 1: 19-21.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.F., Guindon, S., Lefort, V., Lescot, M., Claverie, J.M. and Gascuel, O. (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Research 36(Web Server issue): 465-469.

Dunhill, P.M. and Fowden, L. (1965). Enzymatic formation of β-cyanoalanine from cyanide by *Escherichia coli* extracts. Nature 208: 1206-1207.

Elias, M., Sudhakaran, P. and Nambisan, B. (1997a). Purification and characterization of β-cyanoalanine synthase from cassava tissues. Phytochemistry 46: 469-472.

Elias, M., Nambisan, B. and Sudhakaran, P. (1997b). Catabolism of linamarin in cassava (*Manihot esculenta* Crantz). Plant Science 126: 155-162.

Emanuelsson, O., Nielsen, H., Brunak, S. and von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. Journal of Molecular Biology 300: 1005-1016.

FAO/IFAD - Food and Agriculture Organization of the United Nations/International Fund for Agricultural Development (2000). The world cassava economy. Facts, trends and outlooks. Rome, Italy.

Fregene, M. and Puonti-Kaerlas, J. (2002). Cassava biotechnology. In: Hillocks, R.J., Thresh, J. M., Belloti, A.C. Cassava: Biology, Production and utilization. CAB International. pp. 179-207.

Floss, H.G., Hadwiger, L. and Conn, E.E. (1965). Enzymic formation of beta-cyanoalanine from cyanide. Nature 208: 1207-1208.

Gaitonde, M.K. (1967) A spectrophotometric method for the direct determination of cysteine in the presence of other naturally occurring amino acids. Biochemistry Journal 104: 627-633.

Gamborg, O., Miller, R. and Ojima, K. (1968). Nutrient requirement of suspension cultures of soybean and carrot cells. Experimental Cell Research 50: 151-158.

Gresshoff, P. and Doy, C. (1974). Derivation of a haploid cell line from *Vitis vinifera* and the importance of the stage of meiotic development of anthers for haploid culture of this and other genera. Z. Pflanzenphysiol. 73: 132-141.

Hankoua, B.B., Taylor, N., Ng, S.Y.C., Fawole, I., Puonti-Kaerlas, J., Padmanabhan, C., Yadav, J.S., Fauquet, C.M., Dixon, A.G.O. and Fondong, V.N. (2006). Production of the first transgenic cassava in Africa via direct shoot organogenesis from friable embryogenic calli and germination of maturing somatic embryos. African Journal of Biotechnology 5: 1700-1712.

Hasegawa, R., Tomoko, T., Yuichiro, T. and Yohji, E. (1994). Presence of β-cyanoalanine synthase in unimbibed dry seeds and its activation by ethylene during pregermination. Physiologia Plantarum 91: 141-146.

Hatzfeld, Y., Maruyama, A., Schmidt, A., Noji, M., Ishizawa, K. and Saito, K. (2000). β-cyanoalanine synthase is a mitochondrial cysteine synthase-like protein in spinach and *Arabidopsis*. Plant Physiology 123: 1163-1171.

Hell, R., Bork, C., Bogdanova, N., Frolov, I. and Hauschild, R. (1994). Isolation and characterization of two cDNAs encoding for compartment specific isoforms of *O*-acetylserine(thiol) lyase from *Arabidopsis thaliana*. Federation of European Biochemical Societies (FEBS) Letters 351: 257-262.

Hendrickson, H.R. and Conn, E.E. (1969). Cyanide metabolism in higher plants. IV. Purification and properties of the β-cyanoalanine synthase of blue lupine. The Journal of Biological Chemistry 244: 2632-2640.

Howlett, W.P. (1994). Konzo: a new human disease entity. Acta Horticulturae (ISHS) 375: 323-329.

Hulanicka, M.D, Garret, C., Jagura-Burdzy, G. and Kredich, N.M. (1986). Cloning and characterization of the *cysAMK* region of *Salmonella typhimurium*. Journal of Bacteriology 168: 322-327.

Ibrahim, A.B., Heredia, F.F., Pinheiro, C.B., Aragao, F.J.L. and Campos, F.A.P. (2008). Optimization of somatic embryogenesis and selection regimes for particle bombardment of friable embryogenic callus and somatic cotyledons of cassava (*Manihot esculenta* Crantz). African Journal of Biotechnology 7: 2790-2797.

Iglesias, C.A., Sanchez, T. and Yeoh, H.-H. (2002). Cyanogens and linamarase activities in storage roots of cassava plants from breeding program. Journal of Food Composition and Analysis 15: 379-387.

Ihemere, U.E. (2003). Somatic embryogenesis and transformation of cassava for enhanced starch production. Ph.D. thesis, The Ohio State University, Columbus, OH.

Ihemere, U., Arias-Garzón, D., Lawrence, S. and Sayre, R. (2006). Genetic modification of cassava for enhanced starch production. Plant Biotechnology Journal 4: 453-465.

Ikegami, F., Kaneko, M., Kamiyama, H. and Murakoshi, I. (1988). Purification and characterization of cysteine synthases from *Citrullus vulgaris*. Phytochemistry 27: 697-701.

Ikegami, F. and Murakoshi, I. (1994). Enzymic synthesis of non-protein β-substituted alanines and some higher homologues in plants. Phytochemistry 35: 1089-1104.

Ikegami, F., Takayama, K. and Murakoshi, I. (1989a). Purification and properties of β-cyano-L-alanine synthase from *Lathyrus latifolius*. Phytochemistry 27: 3385-3389.

Ikegami, F., Takayama, K., Kurihara, T., Horiuchi, S., Tajima, C., Shira, R. and Murakoshi, I. (1989b). Purification and properties of β-cyano-L-alanine synthase from *Vicia angustifolia*. Phytochemistry 28: 2285-2291.

Jennings, D.L. and Iglesias, C. (2002). Breeding for crop improvement. In: Hillocks, R.J., Thresh, J. M., Belloti, A.C. Cassava: Biology, Production and utilization. CAB International. pp. 149-166.

Jørgensen, K., Bak, S., Busk, P.K., Sørensen, C., Olsen, C.E., Puonti-Kaerlas, J. and Møller, B.L. (2005). Cassava plants with a depleted cyanogenic glycoside content in leaves and tubers. Distribution of cyanogenic glucosides, their site of synthesis and transport, and blockage of the biosynthesis by RNA interference technology. Plant Physiology 139: 363-374.

Kakes, P. (1994). The function of cyanogenesis in cassava. Acta Horticulturae (ISHS) 375: 79-86.

Kibbe, W.A. (2007). OligoCalc: An online oligonucleotide properties calculator. Nucleic Acids Research 35 (webserver issue): May 25.

Koch, B., Nielsen, V.S., Halkier, B.A., Olsen, C.E. and Møller, B.L. (1992). The biosynthesis of cyanogenic glycosides in seedlings of cassava (*Manihot esculenta* Crantz). Archives of Biochemistry and Biophysics 292: 141-150.

Koch, B., Sibbesen, O., Swain, E., Kahn, R., Liangcheng, D., Bak, S., Halkier, B. and Møller, B.L. (1994). Possible use of a biotechnological approach to optimize and regulate the content and distribution of cyanogenic glucosides in cassava to increase food safety. Acta Horticulturae 375: 45-6.

Lewin, B. (1999). Genes VII. Oxford University Press, 990 pp.

Li, H.-Q., Sautter, C., Potrykus, I. and Puonti-Kaerlas, J. (1996). Genetic transformation of cassava (*Manihot esculenta* Crantz). Nature Biotechnology 14: 736-740.

López, J. (2002). Semilla vegetativa de yuca. In: Ospina, B. and Ceballos, H. La yuca en el tercer milenio. CIAT, Colombia.

López, C., Jorge, V., Piégu, B., Mba, C., Cortes, D., Restrepo, S., Soto, M., Laudié, M., Berger, C., Cooke, R., Delseny, M., Tohme, J. and Verdier, V. (2004). A unigene catalogue of 5700 expressed genes in cassava. Plant Molecular Biology 56: 541-554.

Lunn, J.E., Droux, M., Martin, J. and Douce, R. (1990) Localization of ATP sulfurylase and *O*-acetylserine(thiol)lyase in spinach leaves. Plant Physiology 94: 1345-1352.

Maruyama, A., Ishizawa, K. and Takaqi, T. (2000). Purification and characterization of beta-cyanoalanine synthase and cysteine synthases from potato tubers: are beta-cyanoalanine synthase and mitochondrial cysteine synthase same enzyme? Plant Cell Physiology 41: 200-208.

Maruyama, A., Saito, K. and Ishizawa, K. (2001). β-cyanoalanine synthase and cysteine synthase from potato: molecular cloning, biochemical characterization, and spatial and hormonal regulation. Plant Molecular Biology 46: 749-760.

McMahon, J., White, W. and Sayre, R. (1995). Cyanogenesis in cassava (*Manihot esculenta* Crantz). Journal of Experimental Botany 46: 731-741.

Mejía de Tafur, M.S. (2002). Fisiología de la yuca. In: Ospina, B. and Ceballos, H. La yuca en el tercer milenio. CIAT, Colombia.

Miller, J. and Conn, E. (1980). Metabolism of hydrogen cyanide in higher plants. Plant Physiology 65: 1199-1202.

Mkpong, O.E., Yan, H., Chism, G. and Sayre, R.T. (1990). Purification, characterization and localization of linamarase in cassava. Plant Physiology 93: 176-181.

Morant, A.V., Jørgensen, K., Jørgensen, B., Dam, W., Olsen, C.E., Møller, B.L. and Bak, S. (2007). Lessons learned from metabolic engineering of cyanogenic glucosides. Metabolomics 3: 383-398.

Murashige, T. and Skoog, F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiologia Plantarum 15: 473-497.

Murcha, M.W., Rudge, C., Elhafez, D., Adams, K.L., Daley, D.O. and Whelan, J. (2005). Adaptations required for mitochondrial import following mitochondrial to nucleus gene transfer of ribosomal protein S10. Plant Physiology 138: 2134-2144.

Nartey, F. (1969). Studies on cassava, *Manihot utilissima*. II. Biosynthesis of asparagine-^{14}C from ^{14}C-labelled hydrogen cyanide and its relations with cyanogenesis. Physiologia Plantarum 22: 1085-1096.

Nhassico, D., Muquingue, H., Cliff, J., Cumbana, A. and Bradbury, J.H. (2008). Rising African cassava production, diseases due to high cyanide intake and control measures. Journal of the Science of Food and Agriculture 88: 2043-2049.

Nweke, F., Spencer, D. and Lynam, J. (2002). The cassava transformation: Africa's best kept secret. Michigan State University Press, East Lansing, USA.

Oluwole, O., Onabolu, A., Link, H. and Roslin, H. (2000). Persistence of tropical ataxic neuropathy in a Nigerian community. Journal of Neurology, Neurosurgery and Psychiatry 69: 96-101.

Oluwole, O., Onabolu A. and Sowunmi, A. (2002). Exposure to cyanide following a meal of cassava food. Toxicology Letters 135: 19-23.

Padmaja, G. (1995). Cyanide detoxification in cassava for food and feed uses. CRC Critical Reviews in Food Science and Nutrition 35: 299-339.

Poulton, J. (1990). Cyanogenesis in plants. Plant Physiology 94: 401-405.

Raemakers, C.J.J.M., Amati, M., Staritsky, G., Jacobsen, E. and Visser, R.G.F. (1993). Cyclic somatic embryogenesis and plant regeneration in cassava. Annals of Botany 71: 289-294.

Raemakers, C.J.J.M., Sofiari, E., Taylor, N., Hanshaw, G., Jacobsen, E. and Visser, R.G.F. (1996). Production of transgenic cassava (*Manihot esculenta* Crantz) plants by particle bombardment using luciferase activity as selection marker. Molecular Breeding 2: 339-349.

Raemakers, C.J.J.M., Schreuder, M., Pereira, I, Munyikwa, T., Jacobsen, E. and Visser, R.G.F. (2001). Progress made in FEC transformation of cassava. Euphytica 120: 15-24.

Rozen, S. and Skaletsky, H. (1998). Primer3. Code available at http://www-genome.wi.mit.edu/genome_software/other/primer3.html.

Sambrook, J. and Russell, D. (2001). Molecular Cloning: A Laboratory Manual. Vol. 1, 3rd Ed., Cold Spring Harbor Laboratories Press, Cold Spring Harbor, NY.

Sarria, R., Torres, E., Angel, F., Chavarriaga, P. and Roca, W.M. (2000). Transgenic plants of cassava (*Manihot esculenta*) with resistance to Basta obtained by *Agrobacterium*-mediated transformation. Plant Cell Reports 19: 339-344.

Schardl, C.L., Byrd, A.D., Benzion, G., Altschuler, M.A., Hildebrand, D.F. and Hunt, A.G. (1987). Design and construction of a versatile system for the expression of foreign genes in plants. Gene 61(1): 1-11.

Schöpke, C., Taylor, N., Cárcamo, R., Konan, N.K., Marmey, P., Henshaw, G.G., Beachy, R.N. and Fauquet, C. (1996). Regeneration of transgenic cassava plants (*Manihot esculenta* Crantz) from microbombarded embryogenic suspension cultures. Nature Biotechnology 14: 731-735.

Schreuder, M.M, Raemakers, C.J.J.M., Jacobsen, E. and Visser, R.G.F. (2001). Efficient production of transgenic plants by *Agrobacterium*-mediated transformation of cassava (*Manihot esculenta* Crantz). Euphytica 120: 35-42.

Selmar, D. (1994). Translocation of cyanogenic glucosides in cassava. Acta Horticulturae (ISHS) 375: 61-68.

Selmar, D., Lieberei, R. and Biehl, B. (1988). Mobilization and utilization of cyanogenic glycosides. Plant Physiology 86: 711-716.

Siritunga, D. (2002). Generation of acyanogenic cassava (Manihot esculenta Crantz): Transgenic approaches. Ph.D. thesis, The Ohio State University, Columbus, OH.

Siritunga, D., Arias-Garzón, D., White, W. and Sayre, R. (2004). Over-expression of hydroxynitrile lyase in transgenic cassava roots accelerates cyanogenesis and food detoxification. Plant Biotechnology Journal 2: 37-43.

Siritunga, D. and Sayre, R. (2003). Generation of cyanogen-free transgenic cassava. Planta 217: 367-373.

Siritunga, D. and Sayre, R. (2004). Engineering cyanogen synthesis and turnover in cassava (*Manihot esculenta*). Plant Molecular Biology 56: 661-669.

Siritunga, D. and Sayre, R. (2007). Transgenic approaches for cyanogen reduction in cassava. Journal of AOAC International 90: 1450-1455.

Sjöling, S. and Glaser, E. (1998). Mitochondrial targeting peptides in plants. Trends in Plant Science 3: 136-140.

Small, I., Peeters, N., Legeai. F. and Lurin, C. (2004). Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. Proteomics 4: 1581-1590.

Sreeja, V., Nagahara, N., Li, Q. and Minami, M. (2003). New aspects in pathogenesis of konzo: Neural cell damage directly caused by linamarin contained in cassava (*Manihot esculenta* Cranz). The British Journal of Nutrition 90: 467-472.

Stamp, J.A. and Henshaw, G.G. (1987). Somatic embryogenesis from clonal leaf tissue of cassava. Annals of Botany 59: 445-450.

Subbaiah, C.C., Palaniappan, A., Duncan, K., Rhoads, D.M., Huber, S.C. and Sachs, M.M. (2006). Mitochondrial localization and putative signaling function of sucrose synthase in maize. The Journal of Biological Chemistry 281: 15625-15635.

Taiz, L. and Zeiger, E. (2002a). Plant physiology. 3rd Edition. Sinauer Associates, Inc. Sunderland, MA, USA. pp 264-266.

Taiz, L. and Zeiger, E. (2002b). Plant physiology. 3rd Edition. Sinauer Associates, Inc. Sunderland, MA, USA. pp 272-275.

Tattersall, D.B., Bak, S., Jones, P.R., Olsen, C.E., Nielsen, J.K., Hansen, M.L., Høj, P.B. and Møller, B.L. (2001). Resistance to an herbivore through engineered cyanogenic glucoside synthesis. Science 293: 1826-1828.

Takahashi, H. and Saito, K. (1996). Subcellular localization of spinach cysteine synthase isoforms and regulation of their gene expression by nitrogen and sulfur. Plant Physiology 112: 273-280.

Taylor, N., Edwards, M., Kiernan, R.J., Davey, C.D.M., Blakesley, D. and Henshaw, G.G. (1996). Development of friable embryogenic callus and embryogenic suspension culture systems in cassava (*Manihot esculenta* Crantz). Nature Biotechnology 14: 726-730.

Taylor, N., Masona, M.V., Carcamo, R., Ho, T., Schöpke, C. and Fauquet, C.M. (2001). Production of embryogenic tissues and regeneration of transgenic plants in cassava (*Manihot esculenta* Crantz). Euphytica 120: 25-34.

Taylor, N., Chavarriaga, P., Raemakers, K., Siritunga, D. and Zhang, P. (2004a). Development and application of transgenic technologies in cassava. Plant Molecular Biology 56: 671-688.

Taylor, N., Kent, L. and Fauquet, C. (2004b). Progress and challenges for the development of transgenic technologies in cassava. AgBioForum 7: 51-56.

Tylleskar, T., Cooke, R.D., Banea, M., Poulter, N.H., Bikangi, N. and Rosling, H. (1992). Cassava cyanogens and konzo, an upper motoneuron disease found in Africa. Lancet 339: 208-211.

Vetter, J. (2000). Plant cyanogenic glycosides. Toxicon 38: 11-36.

Warrillow, A. and Hawkesford, M. (1998). Separation, subcellular location and influence of sulphur nutrition on isoforms of cysteine synthase in spinach. Journal of Experimental Botany 49: 1625-1636.

Warrillow, A. and Hawkesford, M. (2000). Cysteine synthase (*O*-acetylserine (thiol) lyase) substrate specificities classify the mitochondrial isoform as a cyanoalanine synthase. Journal of Experimental Botany 51: 985-993.

Watanabe, M., Kusano, M., Oikawa, A., Kukushima, A., Noji, M. and Saito, K. (2008a). Physiological roles of the β-substituted alanine synthase gene family in Arabidopsis. Plant Physiology 146: 310-320.

Watanabe, M., Mochida, K., Kato, T., Tabata, S., Yoshimoto, N., Noji, M. and Saito, K. (2008b). Comparative genomics and reverse genetics analysis reveal indispensable functions of the serine acetyltransferase gene family in Arabidopsis. The Plant Cell 20: 2484-2496.

Wenck, A. and Hansen, G. (2004). Positive selection. In: Peña, L. Methods in Molecular Biology, vol. 286: Transgenic Plants: Methods and Protocols. Humana Press Inc., Totowa, N.J.

White, W.L.B., Arias-Garzón, D.I., McMahon, J.M. and Sayre, R.T. (1998). Cyanogenesis in cassava. The role of hydroxynitrile lyase in root cyanide production. Plant Physiology 116: 1219-1225.

White, W.L.B., McMahon, J.M. and Sayre, R.T. (1994). Regulation of cyanogenesis in cassava. Acta Horticulturae (ISHS) 375: 69-78.

Winkel, B.S.J. (2004). Metabolic channeling in plants. Annual Review of Plant Biology 55: 85-107.

Wurtele, E.S., Nikolau, B.J. and Conn, E. (1984). Tissue distribution of β-cyanoalanine synthase in leaves. Plant Physiology 75: 979-982.

Wurtele, E.S., Nikolau, B.J. and Conn, E. (1985). Subcellular and developmental distribution of β-cyanoalanine synthase in barley leaves. Plant Physiology 78: 285-290.

Yamaguchi, Y., Nakamura, T., Kusano, T. and Sano, H. (2000). Three Arabidopsis genes encoding proteins with differential activities for cysteine synthase and beta-cyanoalanine synthase. Plant Cell Physiology 41: 465-476.

Zagrobelny, M., Bak, S., Rasmussen, A.V., Jørgensen, B., Naumann, C.M. and Møller, B.L. (2004) Cyanogenic glycosides and plant-insect interactions. Phytochemistry 65: 293-306.

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

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

Zhang, P., Jaynes, J.M., Potrykus, I., Gruissem, W. and Puonti-Kaerlas, J. (2003). Transfer and expression of an artificial storage protein (ASP1) gene in cassava (*Manihot esculenta* Crantz). Transgenic Research 12: 243-250.

Zhang, P. and Puonti-Kaerlas, J. (2004). Regeneration of transgenic cassava from transformed embryogenic tissues. In: Peña, L. Methods in Molecular Biology, vol. 286: Transgenic Plants: Methods and Protocols. Humana Press Inc., Totowa, N.J.

APPENDIXES

APPENDIX A: Lists of primers used in this study.

Table A.1: Primers designed for the isolation of the cassava β-CAS cDNA

Primer name	Abbreviation	Sequence used for design and accession #	Primer sequence
Manes-CDS FW	1	*M. esculenta* (CK644310)	Cttgctatgcggaggttcttc
Hevea-CDS RV	2	*H. brasiliensis* (AY207389)	Ttagtcaactggaactgg
MANes-Lpz-FW1	3	*M. esculenta* (CK644310)	Gacggtcttgatagagccaacttc
MANes-Lpz-FW2	4	*M. esculenta* (CK644310)	Gtggagcttacatagccgtcaag
MANes-Lpz-RV1	5	*M. esculenta* (CK644310)	Ccctcatagtcacccttctttcc
MANes-Lpz-RV2	6	*M. esculenta* (CK644310)	Cttgacggctatgtaagctccac
HEVbr-FW1	7	*H. brasiliensis* (AY207389)	Gtgggaatatcatctggagccaac
HEVbr-FW2	8	*H. brasiliensis* (AY207389)	Cacagttgctgcacttagacttgc
5UTR-1	5'1	*M. esculenta* (EU350583)	Tgtgttgtccactgcgagag
5UTR-2	5'2	*M. esculenta* (MANes;BsasB)	Actctgttatccactacttgcgg
CDS-A	s1	*M. esculenta* (EU350583)	gcttctagctttcatcaatggc
CDS-B	s2	*M. esculenta* (MANes;BsasB)	gttcaatggcttcccttagg
3UTR-A	3'1	*M. esculenta* (EU350583)	Agtagcggatactttgttttg
3UTR-B	3'2	*M. esculenta* (EU350583)	tcaagctgaactagctag
3utr-CASB	3'3	*M. esculenta* (MANes;BsasB)	Gcacaggaattggttgcatag

Table A.2: β-CAS specific cloning primers and other primers used for this study

Primer name	Abbreviation	Sequence used for design and accession	Primer sequence
MAN-FCP-SacI	ManesA-Sac	*M. esculenta* (EU350583)	ccgcagggagctcctatggaagctc
MAN-RCP-XbaI	Manes-Xba	*M. esculenta* (EU350583)	ggtctagattagtcaactggaactgg
ManesA-H	ManesA-Nco	*M. esculenta* (EU350583)	ctgcccatggatctccctgc
Manes-CASB NcoI-SacI	ManesB-Sac	*M. esculenta* (MANes;BsasB)	ccggagctcccatggaatctc
ManesB-H	ManesB-Nco	*M. esculenta* (MANes;BsasB)	tgcccatggatcgtcctcaag
Arath-H2	Arath-Nco	*A. thaliana* (AJ010505)	ccccatggacttcccctctaccaac
ARAth-FCP-SacI	Arath-Sac	*A. thaliana* (AJ010505)	ctgagctccaccatgggatct
ARAth-RCP-XbaI	Arath-Xba	*A. thaliana* (AJ010505)	tctagagctcccgagatttttggt
pTV-F	pTV-F	pTV118N (U07649)	tgtgtggaattgtgagcgg
pTV-R	pTV-R	pTV118N (U07649)	tttcccagtcacgacgttg

Table A.3: primers used for PCR screening of putative transgenic plants

Primer name	Abbreviation (see figure below)	Sequence used for design	Primer sequence
35sFW1	a	pKYLX (appendix B)	aagttcgacggagaaggtgacg
35sFWmod	b	pKYLX (appendix B)	ccactatccttcgcaagacccttc

35sRV2	c	pKYLX (appendix B)	Tccatggcctttgattcagtggg
35sFW5.1	d	pKYLX (appendix B)	tcaccagtctctctctaagc
35sFW5.2	e	pKYLX (appendix B)	ggatcctcgagctgcagg
Arath-FW	f	ARAth;Bsas3;1(AJ010505)	tgacgcctctctgcttattgggaa
Arath-Rv5.1	g	ARAth;Bsas3;1(AJ010505)	tgaggaacacaagaggcg
Arath-Rv5.2	h	ARAth;Bsas3;1(AJ010505)	Cccatgtttcctgaagtaggctc
Arath-Fw3	i	ARAth;Bsas3;1(AJ010505)	Gaggttagtagtgaggatgc
RBC	j	pKYLX (see below)	Gttgtcgaaaccgatgatacgaacg
RBC-Rv3	k	pKYLX (see below)	Gattctggtgtgtgcgcaatg
RBC-RV3.2	l	pKYLX (see below)	Acttcggtcattagaggccacgat
nospromFW1	m	pBI121 (AY781296)	Ccgcaacgattgaaggagccac
nospromFW2	q	pBI121 (AY781296)	Caaaaatgctccactgacgttcc
nptIIRv1	n	pBI121 (AY781296)	Tatcacgggtagccaacgctatg
NOST Fw	o	pBI121 (AY781296)	Gaatcctgttgccggtcttg
NOST Rv	p	pBI121 (AY781296)	Ttatcctagtttgcgcgcta
HNL-F-B	-	HNL gene (Z29091)	Gcttttggagtcgtttcctg
HNL-R-B	-	HNL gene (Z29091)	Ccgtaacctttttcggtgaa

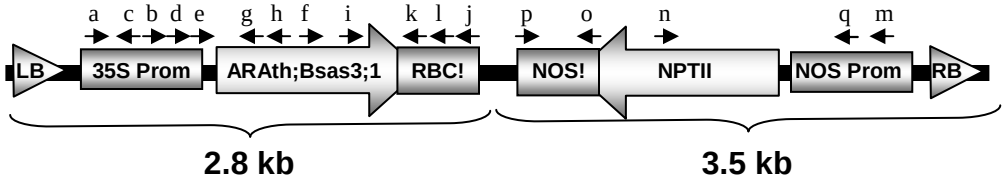

Figure 3.17: Approximate annealing sites of primers designed for PCR screening of putative transgenic 35S:ARAth plants. See table A.3 for detailed information regarding each primer.

Sequence of the empty expression cassette (35s promoter, MCS and RBC!) of pKYLX (available at URL: http://www.uky.edu/~aghunt00/kylx.html**}**) was used for the design of primers employed in the PCR screening of transgenic plants. The sequence is shown below in FASTA format.

>pKYLX

GAATTCGCCCGGGGATCTCCTTTGCCCCAGAGATCACAATGGACGACTTCCTATATC
TCTACGATCTAGTCAGGAAGTTCGACGGAGAAGGTGACGATACCATGTTCACCACTG
ATAATGAGAAGATTAGCCTTTTCAATTTCAGAAAGAATCCTAACCCACAGATGGTTA
GAGACGCTTACGCAGCAGGTCTCATCAAGACGATCTACCCGAGCAATAATCTCCAG
GAGATCAAATACCTTCCCAAGAAGGTTAAAGATGCAGTCAAAAGATTCAGGACTAA
CTGCATCAAGAACACAGAGAAAGATATATTTCTCAAGATCAGAAGTACTATTCCAGT
ATGGACGATTCAAGGCTTGCTTCACAAACCAAGGCAAGTAATAGAGATTGGAGTCT
CTAAAAAGGTAGTTCCCACTGAATCAAAGGCCATGGAGTCAAAGATTCAAATAGAG
GACCTAACAGAACTCGCCGTAAAGACTGGCGAACAGTTCATACAGAGTCTCTTACG
ACTCAATGACAAGAAGAAAATCTTCGTC{AACATGGTGGAGCACGACACGCTTGTCT
ACCTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGAATTGAGACTTTT
CAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCAC
TTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGAT
AAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACC
CCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGC
AAGTGGATTGATGTGAT}AACATGGTGGAGCACGACACGCTTGTCTACCTCCAAAAA
TATCAAAGATACAGTCTCAGAAGACCAAAGGGAATTGAGACTTTTCAACAAAGGGT
AATATCCGGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTTATTGTGAAG
ATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGC
CATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGA
GGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGA
TGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGAC
CCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGACACGCTGAAATCACCAG
TCTCTCTCTAAGCTTGGATCCTCGAGCTGCAGGAGCTCGAATTGATCCTCTAGAGCTT
TCGTTCGTATCATCGGTTTCGACAACGTTCGTCAAGTTCAATGCATCAGTTTCATTGC
GCACACACCAGAATCCTACTGAGTTCGAGTATTATGGCATTGGGAAAACTGTTTTTC
TTGTACCATTTGTTGTGCTTGTAATTTACTGTGTTTTTTATTCGGTTTTCGCTATCGAA
CTGTGAAATGGAAATGGATGGAGAAGAGTTAATGAATGATATGGTCCTTTTGTTCAT
TCTCAAATTAATATTATTTGTTTTTTCTCTTATTTGTTGTGTGTTGAATTTGAAATTAT
AAGAGATATGCAAACATTTTGTTTTGAGTAAAAATGTGTCAAATCGTGGCCTCTAAT
GACCGAAGTTAATATGAGGAGTAAAACACTTGTAGTTGTACCATTATGCTTATTCAC
TAGGCAACAAATATATTTTCAGACCTAGAAAAGCTGCAAATGTTACTGAATACAAGT
ATGTCCTCTTGTGTTTTAGACATTTATGAACTTTCCTTTATGTAATTTTCCAGAATCCT
TGTCAGATTCTAATCATTGCTTTATAATTATAGTTATACTCATGGATTTGTAGTTGAG
TATGAAAATATTTTTTAATGCATTTTATGACTTGCCAATTGATTGACAACATGCATCA
ATCGAT

APPENDIX B: Sequences determined in this study

Sequence of MANes;BsasA (accession number EU350583)

> MANes;BsasA

CACTGGAGTGTGTTGTCCACTGCGAGAGAGAGCCTAGAGGGTTGTAATAGTGATAC
CGTGGGTGAGGAAAAGCTTCTAGCTTTCATCAATGGCTACTCTTACATCAATGGCTA
CTCTTAGGAACTTGTTCACGAAAAAATCTTTTGCCTCAAACGAGCTTGCTATGCGGA
GGTTCTTCACACCGCAGGCCGCTGCTGCTGAAGCTCCTTCTTTTGCTCAAAGAGTCA
GGGATCTGCCCAAGAATCTCCCTGCAACTAAGATCAGGACTGAAGTTTCTCATCTCA
TTGGGAGAACTCCCCTTGTTTTTCTTAACAAAGTGAATGAAGGATGTGGAGCTTACA
TAGCCGTCAAGCAAGAGATGATGCAACCAACTGCCAGCATCAAAGACAGGCCGGCT
TTTTCAATGTTCAATGATGCAGAAAAGAAGAATTTAATCTCTCCTGGGAAGACGGTC
TTGATAGAGCCAACTTCTGGTAATATGGGGATTAGTATGGCTTTTATGGCAGCCATG
AAAGGATACAAGATGATTTTAACCATGCCCTCTTACACAAGCTTGGAAAGAAGGGT
GACTATGAGGGCATTTGGAGCTGAGCTAATTCTCACTGATCCCACCAAGGGGATGG
GAGGAACAGTTAAGAAGGCCTATGATCTTTTGGAAACCACACCAAATGCTTACATG
CTACAACAGTTTTCGAATCCTGCAAATAGTAAGATCCATTTTGAAACGACAGGGCCT
GAAATATGGGAGGATACTCTTGGACAAGTTGATATCTTTGTAATGGGAATAGGCAGT
GGAGGAACAGTCACTGGCGTTGGACAGTACCTTAAATCCCAAAATCCTAATGTTAA
GATATATGGAGTGGAGCCTGCTGAAAGTAATGTGCTGAATGGTGGTAAACCAGGTC
CTCATCAAATTACTGGAAATGGAGTAGGATTTAAACCAGACATATTGGACATGGAC
GTTATGGAAAAAGTTCTTGAGGTTAGCAGTGAAGATGCGATAAAAATGGCTAGGAG
GTTGGCATTGGAGGAGGGACTTATGGTGGGAATTTCATCTGGAGCCAACACAGTTGC
TGCACTTAGACTTGCAAGGATGCCAGAGAACAAAGGCAAACTCATCGTGACTGTTC
ATCCAAGTTTCGGGGAGCGATACTTGACTTCTGTGCTGTTTGAACAACTGAGAAATG
AGGCTGCAAACATGCAACCAGTTCCAGTTGACTAAGTGCTGTCAAAAGGATAGTTC
AATAAATAATATCAGAATCTATCGTGTGCAGCTTTTTGTTGACATCATGTATTACTTG
TTAGCTCTCTTGGGTGAAAAACTACGTCTCCAGAATATTAGTTGTGATAAAGCTTTT
GTAATCTGGTGCAAGTTCTGCATTATGGCATCTGTGGAAAAAACTAGCTAGTTCAGC
TTGAGTGTTGTTGTATAAGCAAGCTTTATTAATCCAAAACAAAACAAAGTATCCG
CTACTAAAAAAAAAAAAAAAAAAAAA

Sequence of MANes;BsasB

> MANes;BsasB

ACTGGACTCTGTTATCCACTACTTGCGGTTGTTGCTTGTAATACTCTTGAGAGTTCAA
TGGCTTCCCTTAGGAACTTGTTGAAGAAAAATTCTTTAACGTCCAACAATCTTCCTAT
AAGGAGGTTTTTCTCTTCCGAAGCTGCGCTTGAATCTCCTTCATTTGCTCACAAAATC
AGGGATCTGCCCAACAATCGTCCTCAAGCTAAAATCAAGCCTGAAGTTTCTCAGCTT
ATTGGGAAAACTCCTCTTGTTTACCTTAACAAAGTGACTGAAGGATGTGGAGCTTTC
ATTGCAGTCAAGCAAGAGATGGTGCAGCCAACTGCCAGTATCAAGGACAGGCCGGC
TTTTGCAATGATCAATGATGCAGAAAAGAAGAATTTAATTTCTCCTGGAAAGACGAC
CTTGATAGAGCCAACTTCTGGAAATATGGGGATCAGTATGGCTTTTATGGCAGCCAT
GAAAGGATACAAGATGGTTCTAACCATGCCTTCTTACACCAGCTTGGAGAGGAGGG
TAACTATGAGGGCATTTGGAGCTGAGCTAATTCTCACTGATCCCACCAAGGGAATGG
GAGGAACAGTTAAGAAGGCTTTTGAGCTTTTGGAATCAACACCAAATGCTTTCATGC
TGCAACAATTTTCAAATCCTGCAAATACTAAGATCCATTTTGAAACTACGGGTCCTG
AAATTTGGGAGGATACACTTGGACAAGTTGACATCTTTGTAATGGGAATAGGCAGC
GGAGGAACAGTCTCTGGTGTTGGACAGTACCTTAAATCTCAAAATCCTGCTGTTAAG
ATATATGGAGTGGAGCCTACTGAAAGTAATGTGCTGAATGGTGGTAAACCAGGACC
TCACCATATTACTGGAAACGGAGTTGGATTTAAACCAGATATATTGGACTTGGATGT
AATGGAAAAAGTTCTTGAGGTTAGCAGCGAAGATGCAGTAAACATGGCTCGGAGGT
TGGCGTTGGAGGAGGGGCTTATGGTGGGAATATCATCCGGAGCCAACACAGTTGCT
GCGCTTAGACTTGCTAGAATGCCAGAGAACAAAGGCAAACTTATTGTGACTGTTCAT
CCAAGTTTCGGGGAGCGATACTTGACTTCTGTCCTGTTTGAAGAACTGAGAAAAGAG
GCTGCAAACATGCAACCAGTTCCAGTTGACTAAGGCCTTCGAATGGAATGTTCTATG
CAACCAATTCCTGTGCTGTGCAACCATTTGTTGAGATCGTATTACTTTTTAGCTCTGT
TGAGTTTTAAACTGTGCCCAGAAAACTTATTGTGATAAAAGCTTATTTGTCTAATAG
AAATCCTATTTATGGCTTCAAAAAAAAAAAAAAAAAAAAAA

APPENDIX C: List of sequences used for homology, phylogenetic and sub-cellular location analyses.

Sequence name	Organism	Common name	Accession number
ALIsa;Bsas2;1	*Allium sativum*	garlic	AAX07221
ARAth;Bsas1;1	*Arabidopsis thaliana*	Arabidopsis	X84097
ARAth;Bsas1;3	*Arabidopsis thaliana*	Arabidopsis	X81697
ARAth;Bsas2;1	*Arabidopsis thaliana*	Arabidopsis	X80377
ARAth;Bsas2;2	*Arabidopsis thaliana*	Arabidopsis	X81973
ARAth;Bsas3;1	*Arabidopsis thaliana*	Arabidopsis	AJ010505
BRAju;Bsas1;1	*Brassica juncea*	Indian mustard	Y10845
BRAju;Bsas1;2	*Brassica juncea*	Indian mustard	Y10847
CAPan;Bsas2;1	*Capsocum annum*	cayenne pepper	X64874
CITvu;Bsas1;1	*Citrulus vulgaris*	water melon	D28777
GLYma;Bsas2;1	*Glycine max*	soybean	ABQ88339
GLYma;Bsas3;1	*Glycine max*	soybean	ABQ88337
HEVbr;Bsas3;1	*Hevea brasiliensis*	rubber tree	AAP41852
MALdo;Bsas3;1	*Malus x domestica*	apple	ABF13209
MALdo;Bsas3;2	*Malus x domestica*	apple	ABF13210
POPtr	*Populus trichocarpa*	poplar, cottonwood	ABK96537
SOLtu;Bsas1;1	*Solanum tuberosum*	potato	AF044172
SOLtu;Bsas2;1	*Solanum tuberosum*	potato	AF044173
SOLtu;Bsas3;1	*Solanum tuberosum*	potato	AB027000
SOLtu;Bsas3;2	*Solanum tuberosum*	potato	AB029338
SPIol;Bsas1;1	*Spinacea oleracea*	spinach	D10476
SPIol;Bsas2;1	*Spinacea oleracea*	spinach	D14722
SPIol;Bsas3;1	*Spinacea oleracea*	spinach	D37963
VITvi	*Vitis vinifera*	wine grape	CAO39634

APPENDIX D: Statistical analyses. Analyses were done using InfoStat.

ANOVA test for β-CAS and CS activities determined from MANes;BsasA, MANes;BsasB, ARAth;Bsas3;1 and pTV-EV clones grown in *Escherichia coli* NK3 cysteine auxotroph.

CAS activity

Análisis de la varianza

Variable	N	R²	R² Aj	CV
CAS (µmol/min/mg)	36	0.98	0.98	8.63

Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor
Modelo	13.67	3	4.56	661.16	<0.0001
Isoform	13.67	3	4.56	661.16	<0.0001
Error	0.22	32	0.01		
Total	13.89	35			

Test:Tukey Alfa=0.05 DMS=0.09504

Error: 0.0069 gl: 32

Isoform	Medias	n				
pTV-EV	-0.04	9	A			
ARAth;Bsas3;1	1.02	9		B		
MANes;BsasA	1.23	9			C	
MANes;BsasB	1.63	9				D

Letras distintas indican diferencias significativas(p<= 0.05)

CS activity

Análisis de la varianza

Variable	N	R²	R² Aj	CV
CS (µmol/min/mg)	36	0.99	0.99	10.83

Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor
Modelo	0.06	3	0.02	798.14	<0.0001
Isoform	0.06	3	0.02	798.14	<0.0001
Error	8.2E-04	32	2.6E-05		
Total	0.06	35			

Test:Tukey Alfa=0.05 DMS=0.00580

Error: 0.0000 gl: 32

Isoform	Medias	n				
pTV-EV	5.4E-04	9	A			
MANes;BsasB	0.02	9		B		
ARAth;Bsas3;1	0.05	9			C	
MANes;BsasA	0.11	9				D

Letras distintas indican diferencias significativas(p<= 0.05)

ANOVA test for β-CAS activity measured from putative transgenic 35S:ARAth lines.

Nueva tabla: 13/02/09 - 3:51:11 PM

Análisis de la varianza

Variable	N	R^2	R^2 Aj	CV
Activity/mg tissue	72	0.52	0.45	22.20

Cuadro de Análisis de la Varianza (SC tipo III)

F.V.	SC	gl	CM	F	p-valor
Modelo	1.9E-04	10	1.9E-05	6.72	<0.0001
Line	1.9E-04	10	1.9E-05	6.72	<0.0001
Error	1.7E-04	61	2.8E-06		
Total	3.5E-04	71			

Test:Tukey Alfa=0.05 DMS=0.00315

Error: 0.0000 gl: 61

Line	Medias	n				
722-1	4.8E-03	9	A			
581	0.01	6	A	B		
42	0.01	6	A	B	C	
723-2	0.01	6	A	B	C	D
551-2	0.01	6	A	B	C	D
553	0.01	6	A	B	C	D
561	0.01	6		B	C	D
541-3	0.01	12			C	D
541-2	0.01	3			C	D
571	0.01	6			C	D
Mcol	0.01	6				D

Letras distintas indican diferencias significativas(p<= 0.05)

APPENDIX E: Results of PCR screening of the putative transgenic plants developed using the 35S:ARAth construct.

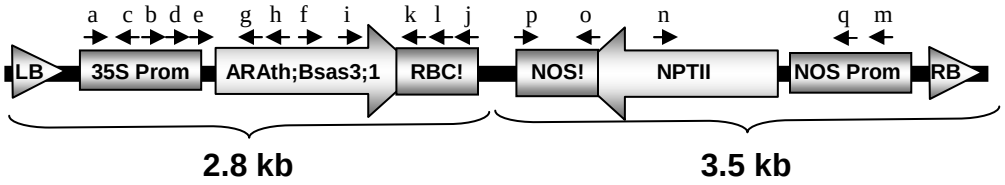

Figure 3.17: Approximate annealing sites of primers designed for PCR screening of putative transgenic 35S:ARAth plants. See table A.3 for detailed information regarding each primer.

Table E.1: PCR screening of the first 35S:ARAth lines from which DNA was extracted. Each primer pair is depicted separating the forward and reverse primers using "/". See tables A.1, A.2 and A.3 of appendix A for details concerning primer sequence, as well as figure 3.17 for approximate annealing sites of primers used. 100-250 ng of genomic DNA were used for each reaction, unless otherwise indicated in parenthesis. A primer pair indicated in parenthesis following a first primer pair means that template DNA used for amplification was taken from a previous PCR reaction attempted with the primer pair specified in parenthesis. Legend: "+", amplified; "-", did not amplify; "~", faint band; "NA", not attempted; "HNL" and "3/2", amplification of endogenous HNL and β-CAS genes, respectively, checking for DNA stability.

Line	b/5	3/2	b/Arath-Xba	HNL	7/j	o/p	o/p(o/p)	i/k	i/Arath-Xba(i/k)
522	-	+	-	+	-	-	-	-	-
511-2	-	+	-	+	-	-	-	-	-
722-2	-	+	-	+	-	-	-	-	-
571	-	+	-	+	-	-	-	-	+
721-3	-	+	-	+	-	-	~	-	-
541-2	-	+	-	+	-	~	+	-	-
532	-	+	-	+	-	-	-	-	-
541-3	-	+	-	+	-	-	-	-	-
511-1	-	+	-	+	-	-	~	-	-
741	-	+	-	+	-	-	-	-	-
531-2	-	+	-	+	-	-	-	-	-
551-2	-	+	-	+	-	-	-	-	-
723-2	-	+	-	+	-	~	~	-	-
711-2	-	+	-	+	-	-	-	-	-
42	-	+	-	+	-	+	+	-	-
572-2	-	+	-	+	-	-	-	-	-
553	-	+	-	+	-	+	+	-	-
722-1	-	+	-	+	-	-	NA	-	-
552-2	NA	NA	NA	NA	NA	-	-	-	-
521-1	NA	NA	NA	NA	NA	-	-	-	-
541-2	NA	NA	NA	NA	NA	-	-	-	-
512-2	NA	NA	NA	NA	NA	-	-	-	-
732-1	NA	NA	NA	NA	NA	-	-	-	-
732-2	NA	NA	NA	NA	NA	-	-	-	-
581	NA	NA	NA	NA	NA	-	-	-	-
552-1	NA	NA	NA	NA	NA	-	-	-	+
512-1	NA	NA	NA	NA	NA	-	-	-	-
733	NA	NA	NA	NA	NA	-	-	-	-

721-2	NA	NA	NA	NA	NA	-	~	-	-
731-2	NA	NA	NA	NA	NA	-	-	-	-
583-2	NA	NA	NA	NA	NA	-	-	-	-
583-1	NA	NA	NA	NA	NA	-	-	-	-
582	NA	NA	NA	NA	NA	-	-	-	-
552-3	NA	NA	NA	NA	NA	-	-	-	-
712	NA	NA	NA	NA	NA	-	-	-	-
723-1	NA	NA	NA	NA	NA	-	-	-	-
559	NA	NA	NA	NA	NA	-	-	-	-
551-1	NA	NA	NA	NA	NA	-	-	-	-
561	NA	NA	NA	NA	NA	-	-	-	+
plasmid	+	+	+	NA	+	+	+	+	+
Mcol	NA	+	NA	+	-	NA	-	-	-
Mcol2	NA	NA	NA	NA	NA	NA	NA	NA	+

Table E.2: PCR screening of the second 35S:ARAth lines from which DNA was extracted. Each primer pair is depicted separating the forward and reverse primers using "/". See tables A.1, A.2 and A.3 of appendix A for details concerning primer sequence, as well as figure 3.17 for approximate annealing sites of primers used. 100-250 ng of genomic DNA were used for each reaction, unless otherwise indicated in parenthesis. A primer pair indicated in parenthesis following a first primer pair means that template DNA used for amplification was taken from a previous PCR reaction attempted with the primer pair specified in parenthesis. Legend: "+", amplified; "-", did not amplify; "~", faint band; "NA", not attempted; "HNL", amplification of endogenous gene, checking for DNA stability.

Line	d/g	d/g(d/g)	d/h	d/h(d/h)	e/h	i/k	i/k(i/k)	HNL(750ng)	i/k(750ng)	i/k(i/k)	i/Arath-Xba(i/k)
721-3	-	-	-	-	-	-	-	+	-	-	-
553	-	-	-	-	-	-	-	+	-	-	-
722-2	-	-	-	-	-	-	-	-	-	-	-
572-2	-	-	-	-	-	-	-	-	-	-	-
532	-	-	-	-	-	-	-	+	-	-	-
551-2	-	-	-	-	-	-	-	+	-	-	-
42	-	-	-	-	-	-	-	+	-	-	-
541-3	-	-	-	-	-	-	-	+	-	-	+
571	-	-	-	-	-	-	-	+	-	-	-
511-1	-	-	-	-	-	-	-	+	-	-	-
541-2	-	-	-	-	-	-	-	+	-	-	-
711-2	-	-	-	-	-	-	-	+	-	-	-
511-2	-	-	-	-	-	-	-	+	-	-	-
513	-	-	-	-	-	-	-	-	-	-	-
584-3	-	-	-	-	-	-	-	-	-	-	-
741	-	-	-	-	-	-	-	+	-	-	-
722-1	-	-	-	-	-	-	-	+	-	-	-
723-2	-	-	-	-	-	-	-	-	-	-	-
531-2	-	-	-	-	-	-	-	-	-	-	-
552-1	-	-	-	-	-	-	-	+	-	-	-
512-2	-	-	-	-	-	-	-	+	-	-	-
712	-	-	-	-	-	-	-	-	-	-	-
551-1	-	-	-	-	-	-	-	-	-	-	-
521-1	-	-	-	-	-	-	-	-	-	-	-
732-2	-	-	-	-	-	-	-	+	-	-	-
732-1	-	-	-	-	-	-	-	+	-	-	-

554	-	-	-	-	-	-	-	+	-	-	-
552-2	-	-	-	-	-	-	-	-	-	-	-
583-2	-	-	-	-	-	-	-	+	-	-	-
541-2	-	-	-	-	-	-	-	+	-	-	-
721-2	-	-	-	-	-	-	-	+	-	-	-
581	-	-	-	-	-	-	-	+	-	-	-
561	-	-	-	-	-	-	-	+	-	-	-
552-3	-	-	-	-	-	-	-	-	-	-	-
512-1	-	-	-	-	-	-	-	+	-	-	-
731-2	-	-	-	-	-	-	-	-	-	-	-
733	-	-	-	-	-	-	-	+	-	-	-
583-1	-	-	-	-	-	-	-	-	-	-	-
Mcol	-	-	-	-	-	-	-	+	-	-	-
Plasmid	+	+	+	+	~	+	+	NA	+	+	+

Table E.3: PCR screening of the third 35S:ARAth lines from which DNA was extracted. Each primer pair is depicted separating the forward and reverse primers using "/". See tables A.1, A.2 and A.3 of appendix A for details concerning primer sequence, as well as figure 3.17 for approximate annealing sites of primers used. 100-250 ng of genomic DNA were used for each reaction, unless otherwise indicated in parenthesis. A primer pair indicated in parenthesis following a first primer pair means that template DNA used for amplification was taken from a previous PCR reaction attempted with the primer pair specified in parenthesis. Legend: "+", amplified; "-", did not amplify; "~", faint band; "NA", not attempted; "HNL", amplification of endogenous gene, checking for DNA stability.

#	Line	HNL	a/m	a/c	i/k	i/k (1µg)	i/Arath-Xba	i/k(i/k)	d/l	d/g(d/l)	d/h(d/l)	f/k(d/l)	o/p	m/n	q/n
1	921	+	-	-	-	-	-	-	-	-	-	-	-	+	-
2	723-1	+	-	-	-	-	-	-	-	-	-	-	-	-	-
3	971-2	+	-	-	-	-	-	-	-	-	-	-	-	-	-
4	964-4	+	-	-	-	-	-	-	-	-	-	-	-	-	-
5	963-2	+	-	-	-	-	-	-	-	-	-	-	-	-	-
6	962	+	-	-	-	-	+	-	-	-	-	-	-	-	-
7	923	+	-	-	-	-	+	-	-	-	-	-	-	+	-
8	976	+	-	-	-	-	-	-	-	-	-	-	-	-	-
9	927	+	-	-	-	-	+	-	-	-	-	-	-	-	-
10	941-3	+	-	-	-	-	+	-	-	-	-	-	-	+	-
11	582	+	-	-	-	-	-	-	-	-	-	-	-	-	-
12	964-1	+	-	-	-	-	+	-	-	-	-	-	-	-	-
13	522	+	-	-	-	-	-	-	-	-	-	-	-	-	-
14	8_2	+	-	-	-	-	+	-	-	-	-	-	-	-	-
15	811-2	+	-	-	-	-	-	-	-	-	-	-	-	-	-
16	915	+	-	-	-	-	+	-	-	-	-	-	-	+	-
17	931-4	+	-	-	-	-	-	-	-	-	-	-	-	-	-
18	922-4	+	-	-	-	-	-	-	-	-	+	-	-	-	-
19	9711	+	-	-	-	-	-	-	-	-	-	-	-	-	-
20	972-2	+	-	-	-	-	-	-	-	-	-	-	-	-	-
21	813	+	-	-	-	-	-	-	-	-	-	-	-	-	-
22	9712	+	-	-	-	-	+	-	-	-	-	-	-	-	-
23	974	+	-	-	-	-	-	-	-	-	-	-	-	-	-
24	852	+	-	-	-	-	+	-	-	-	-	-	-	-	-
25	978	+	-	-	-	-	-	-	-	-	-	-	-	-	-
26	924	+	-	-	-	-	-	-	-	-	-	-	-	-	-

27	914	+	-	-	-	-	-	-	-	-	-	-	-	-	-
28	963-3	+	-	-	-	-	-	-	-	-	-	-	-	-	-
29	913	+	-	-	-	-	-	-	-	-	-	-	-	-	-
30	922-5	+	-	-	-	-	-	-	-	-	-	-	-	-	-
31	821	+	-	-	-	-	-	-	-	-	-	-	-	-	-
32	841	+	-	-	-	-	-	-	-	-	-	-	-	-	-
33	531-2	+	-	-	-	-	-	-	-	-	-	-	-	-	-
34	732-1	+	-	-	-	-	-	-	-	-	-	-	-	+	-
35	541-2	+	-	-	-	-	-	-	-	-	-	-	-	-	-
36	721-2	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	-	-
37	583-1	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	+	-
38	731-2	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	-	-
39	583-2	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	+	-
40	584-3	+	-	-	-	-	+	-	NA	NA	NA	NA	NA	NA	NA
41	552-2	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	NA	NA
42	721-3	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	NA	NA
43	552-3	+	-	-	-	-	+	-	NA	NA	NA	NA	NA	NA	NA
44	521-1	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	NA	NA
45	581	+	-	-	-	-	+	-	NA	NA	NA	NA	NA	NA	NA
46	723-2	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	NA	NA
47	572-2	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	NA	NA
48	512-3	+	-	-	-	-	+	-	NA	NA	NA	NA	NA	NA	NA
49	554	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	NA	NA
50	511-1	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	NA	NA
51	513	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	NA	NA
52	951	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	NA	NA
53	532	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	NA	NA
54	571	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	NA	NA
55	741	+	-	-	-	-	+	-	NA	NA	NA	NA	NA	NA	NA
56	552-1	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	NA	NA
57	853	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	NA	NA
58	551-1	+	-	-	-	-	+	-	NA	NA	NA	NA	NA	NA	NA
59	722-2	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	NA	NA
60	931-2	+	-	-	-	-	+	-	NA	NA	NA	NA	NA	NA	NA
61	553	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	NA	NA
62	963-1	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	NA	NA
63	926	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	NA	NA
64	551-2	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	NA	NA
65	733	+	-	-	-	-	-	-	NA	NA	NA	NA	NA	NA	NA
66	912	+	-	-	-	-	+	-	NA	NA	NA	NA	NA	NA	NA
67	512-2	+	-	-	-	-	+	-	NA	NA	NA	NA	NA	NA	NA
68	512-1	+	-	-	-	-	+	-	NA	NA	NA	NA	NA	NA	NA
+	plasmid	NA	+	+	+	+	+	+	+	+	+	+	+	+	+
M1	Mcol	+	-	-	-	-	-	-	-	-	-	-	-	-	-
M2	Mcol2		-	-	-	-	+	-	-	-	-	-	-	-	-
	CASb	NA	NA	NA	NA	NA	+	NA	NA	NA	NA	NA	NA	NA	NA