EXPRESSION OF *QQS* GENE IN TRANSGENIC CASSAVA PLANT:

IMPLICATION IN STARCH AND PROTEIN CONTENT

By

MARIEME DIAO

A THESIS

Submitted in Partial Fulfillment of the Requirements for the degree of
Master of Science in the Food Science and Biotechnology
Graduate Program of Delaware State University

DOVER, DELAWARE
December 2017

This thesis is approved by the following members of the Final Oral Committee:

Dr. Bertrand Hankoua, Committee Chairperson, Department of Human Ecology, Delaware State University

Dr. Alberta N.A. Aryee, Co-advisor and Committee Member, Department of Human Ecology, Delaware State University

Dr. Jung-Lim Lee, Committee Member, Department of Human Ecology, Delaware State University

Dr. Sathyanarayana Elavarthi, Committee Member, Department of Agriculture and Natural Resources, Delaware State University

Dr. Vincent N. Fondong, External Committee Member, Department of Biology, Delaware State University.

ACKNOWLEDGMENTS

First, I express my infinite gratitude to the Owner of all blessings for His uncountable provision and guidance. It is with a heart filled with appreciation and humility that I would like to acknowledge the College of Agricultural and Related Sciences (CARS) and the Department of Human Ecology for accepting me into the MS. Program in Food Science and Biotechnology. I would like to express my profound gratitude to the Dean of CARS, Dr Dyremple B. Marsh for his advices and financial support. I would like to thank the Chair department of Human Ecology, Dr Samuel Besong, for giving me the opportunity to integrate the MS. Food Science and Biotechnology program, for offering me a scholarship and for supplying the chemicals needed in my experiments. Thank you also to his blessed family who assisted me in babysitting my kids in moments really needed. It is with an outmost gratitude that I thank my Advisor and Research Mentor Dr Bertrand Hankoua for all the scientific knowledge I learned from him, for his encouragement and support. I could not have a better mentor for my Master's program. It was a blessing working with you, Dr Hankoua. I would like to thank my co-advisor and committee member Dr. Alberta N.A. Aryee, for her input in this work. I express my sincere appreciation and gratitude to all my committee members Dr. Jung-Lim Lee, Dr. Sathyanarayana Elavarthi and Dr. Vincent N. Fondong for their precious comments. Also, I would like to express my sincere appreciation and gratitude to Dr Ayalew Ligaba Osena for his valuable help in this project. I would like to acknowledge and thank my former professor and program adviser Dr Stephen E. Lumor, for his constant encouragement and for making me believe at the beginning of this journey that I will succeed in this program. I would like also to acknowledge the graduate students that supported me along this journey; Immaculate T. Arrey, Bhagya Sri Kaja, Prince G. Boakye, Duchard Louis,

and Cassandra T. Eyong. Moreover, I would like to thank the USDA-NIFA Evans-Allen Grant for funding this research work.

I would like to express my sincere gratitude to my father Diattara El Fecky Agne, my mother Khadidjatou Sow, and my siblings Yaya, Fatimata, Anta, Maimouna and Tidjane for their love, prayers and support. Last, I would like to express my profound gratitude to my husband, Abdoulaye Diao, for this rich journey we took together seventeen years ago, and which brought my spirit at a state which I am very pleased with. Thank you to Natago D. Diao and Baffa Y. Diao, my two blessed children. Thank you all three for the joy to have you in my life, your love and sacrifices you're continuing to make for me.

EXPRESSION OF *QQS* GENE IN TRANSGENIC CASSAVA PLANT:

IMPLICATION IN STARCH AND PROTEIN CONTENT

MARIEME DIAO

Faculty Advisor: Dr. Bertrand B. Hankoua

ABSTRACT

Cassava, *Manihot Esculenta* Crantz, is a dicot woody perennial shrub that belongs to Euphorbiaceae family. Over 800 million persons globally and 250 million sub-Saharan Africans rely on the starchy root crop cassava as their staple and major daily source of calories. It is a good source of calories for human and animals because of its richness in starch (30-60% dry matter). A typical cassava-based diet provides < 30% of the minimum daily requirement (MDR) for protein and 10% –20% for iron, zinc, and pro-vitamin A. Cassava deficiency in important nutrients especially proteins (1% to 2% dry weight basis) is a major cause of death and morbidity of resources poor people in developing countries who rely exclusively on this crop for their daily calorie source. Many attempts to use both traditional and biotechnology approaches to overcome protein deficiency in cassava have been proven to be unsuccessful. Our goal is to employ modern technologies to improve the health of millions of cassava consumers through the development of bioengineered cassava with increased protein levels. *Qui-Quine Starch (QQS),* a novel orphan gene unique to *Arabidopsis thaliana* was used in this study for the development of novel transgenic cassava plants exhibiting higher protein accumulation in their tissues. *QQS* is fully sequenced and available in the Genbank *(***GenBank#: NM_113975)**. This gene acts as a transcription factor and is known to

impart carbon and nitrogen repartition into plant and therefore causing variations in protein, carbohydrate and even oil accumulation when expressed heterologously in transgenic plants. We expressed *QQS* in transgenic cassava plant via *Agrobacterium*-mediated genetic transformation with the goal to increase protein accumulation and study the eventual variations in plant morphology and biochemical profiles of other metabolites such as carbohydrates. A total of 10 transgenic cassava lines stably expressing *QQS* were regenerated and all well established in soil and only two lines (20%) showing malformed phenotype. Line R7 (F) has the best growth vigor. Traditional PCR demonstrated that *QQS* gene was integrated in the genome of all the lines. RT-PCR data showed that the *QQS* gene is differentially expressed and that its transcript levels varied remarkably between shoot and root biomass of each transgenic line. Results of the nutritional profiles have shown that the expression of *QQS* gene into cassava plant increased leaf protein by 1.36% in line R''' (LA) L2 and root protein by 17.02% for the same line compared to their wild-type control plants and non-*QQS* expressers. Moreover, leaf soluble total carbohydrate increased by 51.76% in line R''' (G) L2 and root soluble total carbohydrate increased by 46.75% in line R7 (F). Regression analysis established no correlation between *QQS* transcript levels, carbohydrate, and protein content for transgenic root and shoot samples. Novel functionality of *QQS* to increase starch content in transgenic biomass is demonstrated. No change in the content of specific amino acids was observed among lines and plant parts. In addition, *QQS* expression revealed an increase in biomass, plant vigor, early in-vitro mini-tubers production for line R7 (F). However, it will be very critical to validate these observed changes in nutritional profiles of these young cassava plants expressing *QQS* by biochemically profiling these transgenic cassava plants at maturity once established in the field. Once validated in model cv60444, the transferability of this technology to consumer's preferred cassava plant parts (tuberous roots and leaves) should be attempted.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

and correlations between carbohydrate and protein content among line.

LIST OF ILLUSTRATIONS

CHAPTER 1: INTRODUCTION

1.1 BACKGROUND AND SIGNIFICANCE

To effectively address the issue of agricultural development in developing countries it is critical to recognize which crops are farmer and consumer preferred and most importantly which are the staple crops that provide the majority of calories. In sub-Saharan Africa (SSA), the starchy root-crop cassava ranks first in total productivity and number two behind maize as a source of calories (Chavez *et al*., 2002; FAO/WHO, 2002; Gegios A. *et al*., 2010).

Cassava (*Manihot esculenta* Crantz) is a woody perennial shrub that belongs to Euphorbiaceae family (Hillocks, *et al.*, 2002), grown widely throughout the tropics for its edible, starch-filled, tuberous roots (Chavariaga *et al.*, 2016). Cassava was brought to Africa from Brazil by Portuguese sailors in the 16th century and rapidly spread throughout the continent. Cassava is cultivated in the tropics and subtropics between latitudes 30° north and 30° south of the equator (Nweke, *et al*., 2002). Cassava constitutes a staple food and a source of income for more than 800 million people worldwide (Chavarriaga *et al*, 2016). More than 263 million tons was produced in 2013; in which Africa contributed 54.8% (144.2 million), Asia 33.5% (88.2 million tons) and the Americas 11.6% (30.5 million tons) (FAOSTAT, 2015). Nigeria remains the world's largest cassava producer with 47.4 million tons produced in 2013 (FAO, 2013). Due to its richness in starch (30-60% dry matter) (FAO, 2013), cassava is not only a good source of calories for human and animals but also an excellent raw material for starch-based products.

Although cassava is a good source of calorie for more than 800 million people worldwide and an economic and industrial valuable cash crop for many resources poor farmers and processors throughout the world, cassava has many shortcomings that limits its productivity and utilization. At the agricultural level,

cassava can potentially produce 90 tons/hectare/year when grown under ideal conditions. In Sub-Saharan Africa, average yields (10-15 tons/ha/year) are considerably less, showing a low production yield compare to its estimated potential yield of 90 tons/ha (Chavarriaga *et al*., 2016). This yield depression is because of various biotic and abiotic stresses. Virus, bacteria, and insect diseases are serious biotic constraints which limit the productivity and the large-scale cultivation of cassava (Chavarriaga *et al.*, 2016). At the nutritional level, cassava tuberous roots are rich in starch (85%) and poor in vitamins and proteins (1% to 2% dry weight basis); while, cassava leaves, are known to be a good source of protein and vitamins but poor in starch (Sayre *et al*., 2011). A typical adult-sized cassava meal (500 g) can provide adequate calories but is an insufficient source of iron, zinc, vitamin A, and protein (Table 1). Frequent consumers of cassava are at greater risk for malnutrition—especially deficiencies in protein, vitamin-A, iron, and zinc than consumers of other diets, particularly those that are cereal-based (Graham *et al.*, 1996). Furthermore, the amino acid content of cassava reveals a very low composition in sulfur amino acids which are essential for human health (Omole, 1977; Nassar & Sousa, 2007). Additionally, cassava is known for its high level in cyanogen; which may render the crop poisonous for daily consumption (Chavarriaga *et al.*, 2016). For these reasons, various scientific, technological and economical actions have been undertaken to overcome these multiple cassava drawbacks which impact both human economic and human health.

For example, multiple attempts have been taken by breeding programs to overcome the high level of cyanogen, viral diseases and nutrients deficiency in cassava trying both conventional and transgenic breeding (Chavarriaga *et al*., 2016). Traditional breeding has been proven successful in introducing new agricultural traits to cultivated cassava from its relative wild types (Chavarriaga *et al.*, 2016). However, traditional breeding methods have met little success in increasing the protein content of cassava, especially due to the absence of cultivars with high

protein content. This therefore, makes transgenic technology more attractive in achieving protein enhancement in cassava. Recently, cassava transgenic techniques have been developed in several laboratories (Taylor *et al*., 2004) which have accelerated the use of biotechnology approaches to address traits improvement in cassava. Biotechnology has been proven successful in developing low cyanogen content (Siritunga & Sayre, 2004), insect resistant (Ladino, *et al*., 2002), virus resistant (Chellappan, *et al.*, 2004), and herbicide resistant (Sarria, *et al*., 2000) cassava crop, and many nutritional traits such as vitamin A (Welsch *et al.*, 2010; Failla *et al.*, 2012), iron (Ihemere *et al*. 2012; Narayanan *et al.*, 2015); and vitamin B (Li et *al.*, 2015) have been enhanced. However, the use of biotechnological tools to increase protein content in cassava has not been studied intensively. The three main previous attempts have yielded little success. The first attempt was to create a strong protein sink in roots shown to be sufficient to elevate total root protein. A patatin-driven cassava hydroxynitrile lyase (HNL) construct was used to produce transgenic plants which had a three-fold increase in total root protein and, importantly, an 80% reduction in root linamarin levels (Sayre, 2011). Cassava hydroxynitrile lyase (HNL) is an enzyme that catalyzes the conversion of acetone cyanohydrin to cyanide. Because this enzyme is localized in the apoplastic space around leaf cells, it could presumably be induced to accumulate there by overexpression without turning over. The second attempt consisted of accumulating proteins in roots were the expression of artificial storage proteins *ASP1,* designed to be rich in essential amino acids (Zhang *et al.*, 2003). The third attempt to produce and accumulate protein bodies in transgenic cassava plants expressing a chimeric protein named zeolin was unsuccessful (Abhary *et al.*, 2011). The three-fold increase in protein content obtained by expressing HNL is still below the minimum daily requirement (MDR) for an adult daily cassava starch meal of 500 g (Table 1). These previous attempts suggest that increasing protein content in staple crop such as cassava diet of the

malnourished is still an extremely challenging task (Sayre, 2011). To this date, no cassava cultivar has been developed with high protein content at levels needed to meet the minimum daily requirements (MDR) for millions who consume cassava as a staple crop. We propose to develop a novel bioengineered cassava cultivar and identify the best transgenic plants expressing a high level of protein enhancing trait. Specifically, we will express *QQS* into cassava to increase its leaf and root protein content. Studies have demonstrated that the ectopic expression of *QQ*S (Li *et al.*, 2015) increased soybean protein independently of the genetic background and the original protein content of the cultivar. Furthermore, *QQS* is known to regulate the metabolic processes that affect the partitioning of carbon and nitrogen among proteins and carbohydrates, therefore resulting in a variation in leaf and seed composition (Li *et al*., 2015). Our approach was to leverage these previous successes with *QQS* to increase protein content in cassava via constitutive single expression of *QQS* in transgenic cassava. Furthermore, suitable molecular and biochemical tools were employed to validate the transgenic nature of the cassava lines expressing *QQS* and their nutritional profile to assert protein and soluble total carbohydrate enhancement.

1.2 STATEMENT OF PROBLEM AND HYPOTHESIS

Nowhere in the world is human health more deleteriously impacted by poor nutrition than in SSA (Graham *et al.* 1996). This is the only region where indices of nutritional status have not improved and where life expectancy and child mortality are increasing. Despite the understanding of the importance of micronutrients, micronutrient deficiencies remain a huge problem among young children and women in SSA (Graham *et al.* 996). More than 320 million people suffer from chronic malnutrition and 400 million from micronutrient deficiencies in SSA, where subsistence farmers and their families are most at risk from the implications of under-nutrition (Steven R. *et*

al. 2008). Many of these farmers, 250 million of the poorest people in the world, rely on cassava as their staple food (Gegios A. *et al*., 2010). Children consuming cassava as a staple food are at risk for inadequate zinc, iron, and vitamin A intake (Gegios A. *et al*., 2010).

Among the world's major crops, cassava ranks fifth in direct human consumption and is an important source of calories for over 800 million people (Chavarriaga *et al*., 2016). Particularly, in Sub-Sahara Africa (SSA), more than 263 million tons was produced in 2013; in which Africa contributed 54.8% (144.2 million), Asia 33.5% (88.2 million tons) and the Americas 11.6% (30.5 million tons) (FAOSTAT, 2015). Cassava roots' starch provides over 25% of dietary energy for an estimated 200 million Africans (FAO, 2012; Chavez *et al.*, 2000). Although cassava is highly regarded in many developing countries as a food security crop, it has important deficiencies in essential minerals, vitamins, and protein. For a diet based principally on staple foods such as cassava, damaging health effects may occur unless the diet is supplemented with the deficient nutrients. Most of the people that depend on cassava as their unique source of dietary energy are economically very poor and cannot afford additional foodstuffs to supply an adequate protein diet. Therefore, improving cassava crop to ensure that people relying on it receive their daily nutrient requirements has become a real concern for scientists. Furthermore, cassava roots contain on average of 1-3 percent protein by dry weight, substantially less even than maize. A meal of 500 g cassava provides only 30% of the daily protein requirement, demonstrating that cassava root is a critically poor source of protein (Sayre *et al.*, 2011).

A recent study among children that consume cassava as a staple food, found that 13 percent of Nigerian and 53 percent of Kenyan children had inadequate protein intake; the fraction proportion of dietary energy derived from cassava was negatively correlated with protein intake, protein: energy (P: E) ratio, and dietary diversity. Height-for-age Z score was directly associated

with protein intake and negatively associated with cassava consumption (Stephenson, 2010). In addition, populations that rely on cassava as their major source of calories are at high risk of protein energy malnutrition (Stephenson, 2010), Kwashiorkor (Sreeja *et al*., 2002), and related pathological disorders (Rosling, 1988). Increasing protein content in the diet of the malnourished is a critical problem which needs to be solved. Protein content in human diet should be approximately 15 percent of total caloric intake; however, in many areas of the world, including Africa, and southern Asia, the staple foods contain far lower levels of protein (Li *et al.*, 2012). A recent study showed that among the world's staple crops, cassava has the lowest protein/energy ratio (7.4 mg protein/cal for cassava compared to 26 mg protein/cal for maize) and is deficient in iron, zinc, pro-vitamin A or β-carotene and vitamin E. Thus, the minimal levels of protein intake are not achieved and, consequently, health deteriorates (Sayre *et al*., 2011; Stephenson, 2010). Maximizing protein intake in impoverished areas of the world could dramatically improve the health status of millions of people relying on a single staple crop such as cassava. Production of bio-fortified staple food products can be achieved either through conventional crop breeding or genetic modification. Despite the potential of conventional breeding to improve micronutrient density in cassava and other staple crop varieties (Graham *et al*., 1996; Yan *et al*., 2010; Low *et al*., 2007; Steven *et al.*, 2008), irregular flowering, high degrees of heterozygosity, low seedling number, and genetic variability represent significant drawbacks for cassava breeders (Jennings *et al*., 2002; Reilly *et al*., 2003). Furthermore, the use of conventional breeding will not be an option where genetic variability for a given trait is too low to achieve the target levels (Jennings *et al*., 2002). Additionally, it has been hypothesized that selecting for higher protein content through conventional breeding may eventually lead to a lower yield because protein synthesis requires about twice the primary products of photosynthesis as the synthesis of a similar weight of starch

(Cock, 1985). A traditional breeding strategy had led to the development of elevated pro-vitamin A bio-fortification, but had no improvement in iron, zinc, vitamin E, and protein content (Reilly K. *et al*., 2003; Siritunga *et al*., 2007); therefore, making genetic engineering an attractive tool for crop improvement, particularly bio-fortification. Biotechnology approaches have been proven successful in increasing Vitamins A (Welsch *et al*., 2010; Failla *et al*., 2012), B (Li *et al.* 2015) and iron (Ihemere *et al.* 2012; Narayanan *et al.* 2015) contents into transgenic cassava lines. Transgenic technologies have been used successfully in increasing protein content and overall essential amino-acids in major staples (Li et *al.*, 2015; Chakraborty *et al*. 2010). But for cassava, difficulties encountered in transgenic protein biofortification suggest that these previous attempts to increase protein content in cassava diet of the malnourished is still an extremely challenging task and remains a very difficult task for transgenic breeders (Sayre R. *et al.*, 2011). While recognizing that traditional breeding and biotechnological methods have unique capabilities and constraints, biotechnological approaches have proven to achieve bio-fortification in the most efficient, cost-effective and sustainable manner (Perez-Massot *et al*., 2013).

Recently, a novel *Arabidopsis* orphan gene, *QQS* that alters plant composition was stably expressed into soybean plants. This gene has been shown to increase the amount of protein in soybean seeds by about 30 to 60 percent (Li *et al*., 2015), raising the possibility that this gene could hold great promise to help meet nutritional protein needs of a hungry world. Now that technology concept for transgenic protein bio-fortification of soybean, a protein rich crop expressing *QQS* is proven as well as the field performance of these bio-fortified soybean lines demonstrated in the U.S., the way is now paved in this graduate research thesis to propose for the first time to express *QQS* gene into a starch rich crop, cassava, generate and identify the best transgenic plants expressing a high level of this trait. This step is critically required for cassava consumers to reap

the benefits from advances in molecular science. QQS gene was expressed in the model cassava cultivar (cv 60444) which is not a preferred cultivar for farmers in Africa (Chavarriaga *et al*., 2016), but chosen because it is the only African germplasm for which transgenic capacity had been demonstrated with; time from *Agrobacterium* co-culture to regeneration of whole plants is 4.5–5 months with escape rates of 5% or less (Chavarriaga *et al*., 2016). The subsequent integration of this transgenically imparted protein nutritional traits into genetic backgrounds of cassava favored by farmers in SSA (Yaboga E. *et al*, 2013). After this proof of concept phase is demonstrated in this study using cv 60444, we will guarantee key success for delivering protein biofortified cassava products to farmers.

Numerous molecular tools now make it possible to express *QQS* in transgenic cassava to advance scientific knowledge and employ modern biotechnologies to improve the health of poor people in developing countries through development and delivery of novel cassava germplasm with increased levels of bioavailable protein. Tools include multiple gene expression systems such as pSAT vectors (Tzfira *et al.*, 2005), discovery of factors critical for optimum heterologous expression of proteins in plants such as root specific promoters (Liu *et al*., 1991), complete cassava genome sequence (Prochnik *et al.*, 2012); codon optimization systems for optimal functional expression *in planta* (Dai *et al.* 2000); ligation-free cloning system (Gibson *et al*., 2009) and an improved, and routine system to recover genetically engineered large numbers of transgenic cassava plants for a given construct (Chavarriaga *et al.*, 2016).

1.3 SPECIFIC OBJECTIVES

This project is focused on 4 specific objectives which are:

1. The design, construction, and characterization of single expression vector with *QQS* gene.

2. The genetic transformation of this construct into cassava embryogenic cells via *Agrobacterium*-mediated techniques and then the regeneration of transgenic cassava transgenic tissues and plants expressing *QQS* gene. Wild-type non-transformed embryogenic cassava tissues will be regenerated alongside and used as a negative control plants.

3. The molecular characterization of putative transgenic lines to confirm their transgenic nature using regular PCR, to evaluate the level of expression of *QQS* in various lines using Real-Time PCR.

4. Finally, the last part of the project will be dedicated to the biochemical characterization of the transgenic lines expressing QQS. We will determine: total soluble protein, total amino acid content and profile as well as soluble total carbohydrate content of root and shoot biomass of these novel bioengineered cassava plants. The chemical profile results obtained from these transgenic lines will be compared to that of the wild type non-transgenic, and transgenic lines only expressing a plant selectable paromomycin marker, *nptII*. This will allow us to validate the role of *QQS* in metabolites partitioning once heterologously expressed in transgenic cassava tissues. We will also monitor any morphological differences between wild-type and transgenic lines.

CHAPTER 2: LITERATURE REVIEW

2.1 CASSAVA AS A STAPLE FOOD AND AN INDUSTRIAL CASH CROP

Cassava, *Manihot Esculenta* Crantz, is a dicot woody perennial shrub that belongs to Euphorbiaceae family (Hillocks, *et al.*, 2002). Cassava is cultivated in the tropics and subtropics between latitudes 30° north and 30° south of the equator (Nweke, *et al*., 2002). Uncertainty shrouds the origin of cassava. However, it is mostly believed that cassava is originated from South America. It has been introduced, later, in West Africa by Portuguese sailors in the sixteenth century. From there, cassava has spread widely into Sub-Saharan Africa, India, the Philippines, and Indonesia (Chavarriaga *et al*, 2016). Today, cassava constitutes a staple food and a source of income for more than 800 million people worldwide (Chavarriaga *et al*, 2016). Ranked as the fourth largest source of calories, global cassava production was more than 263 million tons in 2013; in which Africa contributed 54.8% (144.2 million), Asia 33.5% (88.2 million tons) and the Americas 11.6% (30.5 million tons) (FAOSTAT, 2015). Nigeria remains the world's largest cassava producer with 47.4 million tons produced in 2013 (FAO, 2013). More than 95% of cassava production is used for human consumption; the remaining of the production is used as livestock and raw material for industrial production (Nweke, 2002). Due to its richness in starch (30-60% dry matter) (FAO, 2013), cassava is not only a good source of calories for human and animals but also an excellent raw material for starch producing companies.

Today, the high capacity of cassava roots in accumulating starch attracts bioethanol producers, especially in China and Thailand, who are considering the potential use of cassava roots

in bioethanol production (Chavarriaga *et al*, 2016). This new role of cassava may be a great opportunity for economic growth in Africa and especially for cassava growers.

2.2 SHORTCOMINGS AND CONSTRAINTS TO CASSAVA PRODUCTION AND UTILIZATION

Although cassava is a nutritional and an economic source for more than 800 million people worldwide, cassava has many shortcomings. At the agricultural level, cassava has a low production yield compare to its estimated potential yield. The average yield worldwide was between 12 tons and 13 tons per hectare while its potential yield was estimated as 80 tons per hectare (FAO, 2013). This situation is due to many viral diseases, mainly cassava mosaic disease (CMD) and cassava Brown Streak disease (BSD), affecting the crop especially in Africa where agricultural practices foster contamination, causing important decrease in yield (Legg *et al*., 2014). At the nutritional level, the nutrients are unequally distributed in the edible parts of the plants (leaves and roots). At one hand, cassava roots are rich in starch (85%) and poor in vitamins and proteins (1% to 2% dry weight basis); on the other hand, cassava leaves, are known to be a good source of protein and vitamins while poor in starch (Sayre *et al*., 2011). Moreover, a meal of 500 g cassava provides only 30% of the daily protein requirement, demonstrating that cassava root is a critically poor source of protein (Sayre, 2011). The nutritional qualities of cassava are summarized in the table below (Table 1). Furthermore, the amino acid profile of cassava reveals a very low composition in sulfur amino acids which are essential for human health (Omole, 1977; Nassar & Sousa, 2007). Additionally, cassava is known for its high level in cyanogen; which may render the crop poisonous for daily consumption.

Table 1*: Nutritional qualities of cassava for an adult meal of 500 g*.

Table 1: Nutritional qualities of cassava foods (from FAO: http://faostat.fao.org/site/609/default.aspx)

Cassava meal Recommended Minimum daily (RMD) Intake for an adult	Energy 1700-2400 (Kcal)	Protein 50-80 (g)	Iron 18 (mg)	Zinc 12 (mg)	Vitamin E 8 (mg)	β-carotene, pro-vitamin A 11 (mg)
Boiled (500 gfw) (%RMD)	740 (~36%)	5.5(~9%)	2.0 (11%)	2.0 (17%)	1.0 (13%)	1 (9%)
Dry (500 gdw)	1775 (~87%)	10.5 (~16%)	4.0 (22%)	4.0 (34%)	1.0 (13%)	2 (18%)
Flour (500 gdw)	1710 (~83%)	7.5 (~12%)	4.0 (22%)	3.0 (25%)	1.0 (13%)	0 (0%)
Fresh (500 gfw)	745 (~36%)	6.0 (~9%)	2.0 (11%)	2.0 (17%)	1.0 (13%)	1 (9%)
Roasted (500 gdw)	1360 (~66%)	10 (~15%)	2.5 (13%)	3.0 (25%)	1.0 (13%)	1 (9%)

2.2.1 SOLUTIONS TO ADDRESS CASSAVA SHORTCOMINGS AND CONSTRAINTS TO PRODUCTION AND UTILIZATION

The high consumption of cassava worldwide, particularly in Africa where 200 million people consume cassava on a daily basis (FAO, 2012), and considering its nutritional deficiency, especially in protein and the implications in both economic and human health, it is, therefore, urgent to find solutions to overcome the various drawbacks encountered with this crop. For many several decades, various attempts have been taken by breeding programs to overcome the high level of cyanogen, viral diseases and nutrients deficiency in cassava trying both conventional and transgenic breeding. Conventional breeding uses vegetative propagation techniques to transfer a trait existing in one cultivar into another cultivar lacking that trait. In the case of cassava, traditional breeding is done through stem cuttings. It has been proven successful in introducing new traits to cultivated cassava from its relative wild types with viral and bacterial diseases resistance and improved post-harvest qualities (Nweke, *et al.,* 2002). However, traditional breeding methods have met little success in increasing the protein content of cassava, especially due to the absence of cultivars with high protein content. This therefore, makes transgenic technology more attractive in achieving protein enhancement in cassava. Transgenic breeding is based on improving crop by introducing genes that impart new properties, which the plants and their ancestors never processed through cloning and plant transformation methods. Recently, cassava transgenic techniques have been developed in several laboratories (Taylor *et al*., 2004) which have accelerated the use of biotechnology approaches to address traits improvement in cassava. Biotechnology has been proven successful in developing low cyanogen content (Siritunga & Sayre, 2004), insect resistant (Ladino, et al., 2002), virus resistant (Chellappan, *et al.,* 2004), and herbicide resistant (Sarria, et al., 2000) cassava crop, and many nutritional traits such as vitamin A (Welsch et al., 2010; Failla

et al., 2012), iron (Ihemere *et al.,* 2012; Narayanan *et al.,* 2015); and vitamin B (Li *et al*., 2015) have been enhanced. However, the use of biotechnological tools to increase protein content in cassava has not been studied intensively. The three main previous attempts have yielded little success. The first attempt was to create a strong protein sink in roots shown to be sufficient to elevate total root protein. A patatin-driven cassava hydroxynitrile lyase (HNL) construct was used to produce transgenic plants which had a three-fold increase in total root protein and, importantly, an 80% reduction in root linamarin levels (Sayre, 2011). Cassava hydroxynitrile lyase (HNL) is an enzyme that catalyzes the conversion of acetone cyanohydrin to cyanide. Because this enzyme is localized in the apoplastic space around leaf cells, it could presumably be induced to accumulate there by overexpression without turning over. The second attempt consisted of accumulating proteins in roots were the expression of artificial storage proteins *ASP1,* designed to be rich in essential amino acids (Zhang *et al.* 2003). The third attempt to produce and accumulate protein bodies in transgenic cassava plants expressing a chimeric protein named zeolin was unsuccessful (Abhary, *et al.* 2011). The three-fold increase in protein content by expressing HNL is still below the minimum daily requirement (MDR) for an adult daily cassava starch meal of 500 g. These previous attempts suggest that increasing protein content in staple crop such as cassava diet of the malnourished is still an extremely challenging task (Sayre *et al.,* 2011).

2.2.2 *QUI-QUINE STARCH (QQS),* A TRANSCRIPTION FACTOR THAT MODULATE NITROGEN AND CARBON PARTITIONING WITH MOLECULAR CAPABILITIES TO INCREASE PROTEIN CONTENT *IN PLANTA*

QQS is a novel orphan gene unique to *Arabidopsis thaliana* known for its capability to enhance protein accumulation when expressed in plants. The *QQS* gene encodes a protein of only 59 amino acids (**GenBank#: NM_113975)** whose homolog is not identifiable by primary sequence comparisons to any other sequenced species, not even the closely related Brassica napus (Li *et al*., 2009). Studies have demonstrated that the ectopic expression of *QQ*S increased soybean protein content which is already a very high producing crop. The expression of *QQS* into plant tissues induce a regulation of the metabolic processes that affect the partitioning of carbon and nitrogen among proteins and carbohydrates, therefore resulting in a variation in leaf and seed composition (Li *et al*., 2009).

The processes by which *QQS* affect the partitioning of carbon and nitrogen among proteins and carbohydrates, resulting in a variation in leaf and seed composition demonstrate that this *QQS* ability can be transferred to any other plants independently of the genetic background and the original protein content of the cultivar (Li *et al*., 2015). Hence, we are choosing to express this gene in cassava to generate data which will broaden the concept of *QQS* as a modulator of carbon and nitrogen allocation, and possibly demonstrate that this transcription factor can affect the shoot and root composition of cassava which is an economically important agronomic crop with significant divergence from *Arabidopsis* where *QQS* was originally cloned.

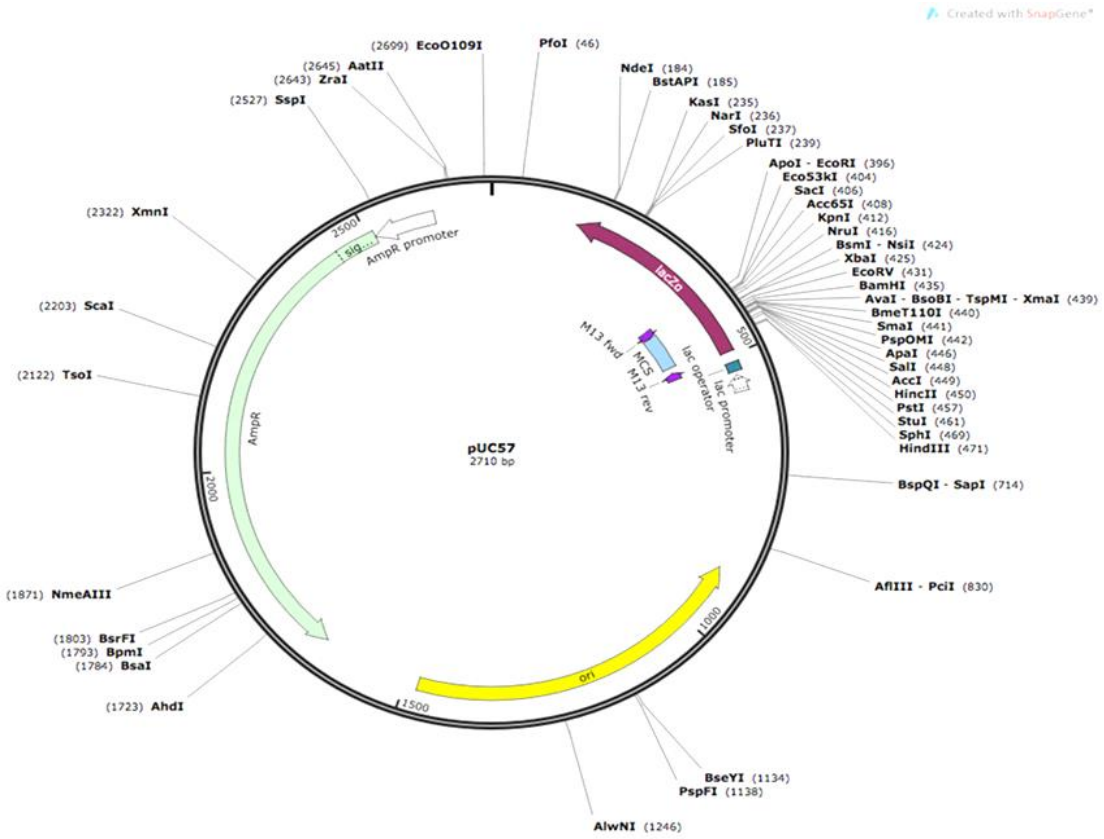
CHAPTER 3: MATERIALS AND METHODS

3.1 MOLECULAR CLONING AND PLANT TRANSFORMATION

3.1.1 *QQS* gene synthesis and characterization of the donor plasmid

QQS sequence *(***GenBank#: NM_113975)** was engineered by adding restriction sites at both 3' and 5' (Sac1 and Kpn1) to facilitate restriction cloning of the synthesized fragment in a shuttle vector of a pSAT series (Tzfira et *al*., 2005) and a 3'-Histag modification for *QQS* protein detection and purification in transgenic cassava tissues and plants. The modified sequence was sent for synthesis and *QQS* codon optimization at the Integrated DNA Technologies, Inc (IDT Coralville, Iowa, USA); pCU57, a 2.7 Kb Genscript vector which has a bacteria ampicillin resistant marker and comporting the optimized and modified *QQS* gene (pCU57-QQS), pSAT1 and pPZP vectors (Tzfira et al., 2005) were all used to create the *QQS* expression construct in pPZP plant transformation vector to facilitate genetic engineering of QQS in cassava plant. A double digestion of the pCU57-*QQS* vector (figure 2) with SacI and KpnI (NEB, MA, USA) restriction enzymes was performed to confirm the presence of the *QQS* insert. Digestion conditions followed manufacturer's recommendations. Followed is the pUC57 *QQS* donor plasmid from Gen script.

Figure 1*: pUC57-QQS donor plasmid from Genscript*



3.1.2 Insertion of *QQS* optimized gene into pSAT1 shuttle vector through Gibson assembly technique.

The purpose of cloning *QQS* optimized sequence into pSAT1 shuttle vector was to assign our gene of interest with genetic elements available in pSAT modular vectors such a promoter (MAS and enhanced 35S in most cases), a multiple cloning site, a homing endonuclease or rare restriction enzyme site, a translation enhancer (5'-UTR from tobacco etch virus), and a terminator sequence. Gibson assembly method was used for this cloning after multiple unsuccessful attempts with ligation technique. Gibson assembly method is a ligation-free cloning technique developed

to increase efficiency and accuracy of DNA assembly (Gibson *et al*., 2009). The Snapgene strategy to generate the assembled pSAT1-*QQS* vector via Gibson assembly is shown in Figure 2.

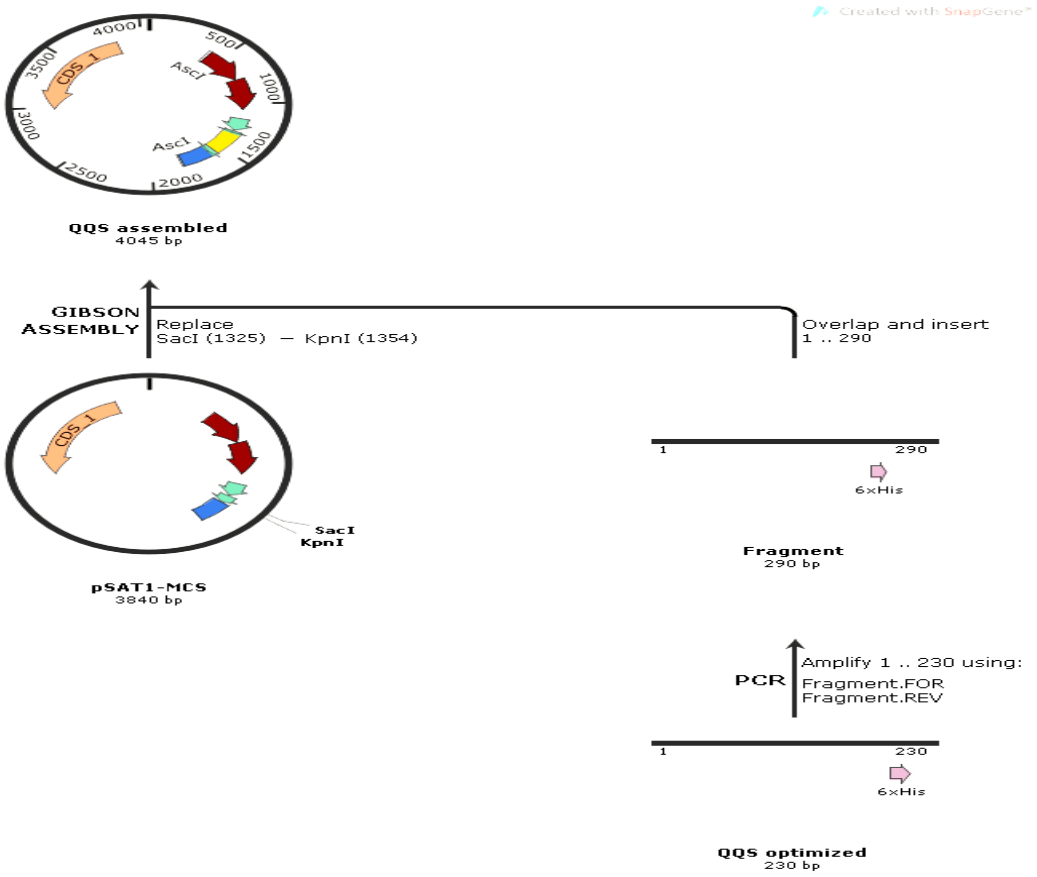


Figure 2*: Assembled pSAT1-QQS vector via Gibson assembly using Snapgene software.*

For this, the optimized *QQS* sequence was redesigned using Snapgene software to assign it with a 5', and 3' overlap sequences (30 nucleotides were added at each 5' and 3' end). The corresponding *QQS* gene specific Gibson primers, *QQS* fragment primers (Table 2), were designed by the software to assign to the *QQS* gene a 5' and 3' overlapping regions via PCR amplification. After synthesis of the *QQS* Gibson primers (IDT, USA) (Table 2), the Gibson *QQS* fragment was generated through a PCR reaction involving the synthesized *QQS* fragment primers with

overlapping regions and the plasmid pCU57 comporting the optimized *QQS* gene sequence as template. Q5 High-fidelity DNA polymerase (NEB, USA) was used for the PCR reaction and following manufacturer's protocol at an annealing temperature of 72 °C (NEB Tm calculator tools). The outcome of this PCR reaction was an amplicon of a linearized *QQS* gene comporting overlapping sequences to its 5' and 3' ends.

PSAT1 which was our initial entry shuttle vector was also linearized using double restriction enzymatic reaction with Sac1 and Kpn1 to generate a linearized vector to be used in the Gibson reaction. Then, the linearized *QQS* gene comporting overlapping regions and the linearized vector were gel purified then the purity and their concentration determined. The two fragments (linearized *QQS* gene and PSAT1 vector) were then combined in a single Gibson reaction following NEBuilder® HiFi DNA Assembly Master Mix/NEBuilder HiFi DNA Assembly Cloning Kit protocol. The hypothetical assembled pSAT1-*QQS* vector with all its important genetic features obtained through Snapgene software design is shown in Figure 3.

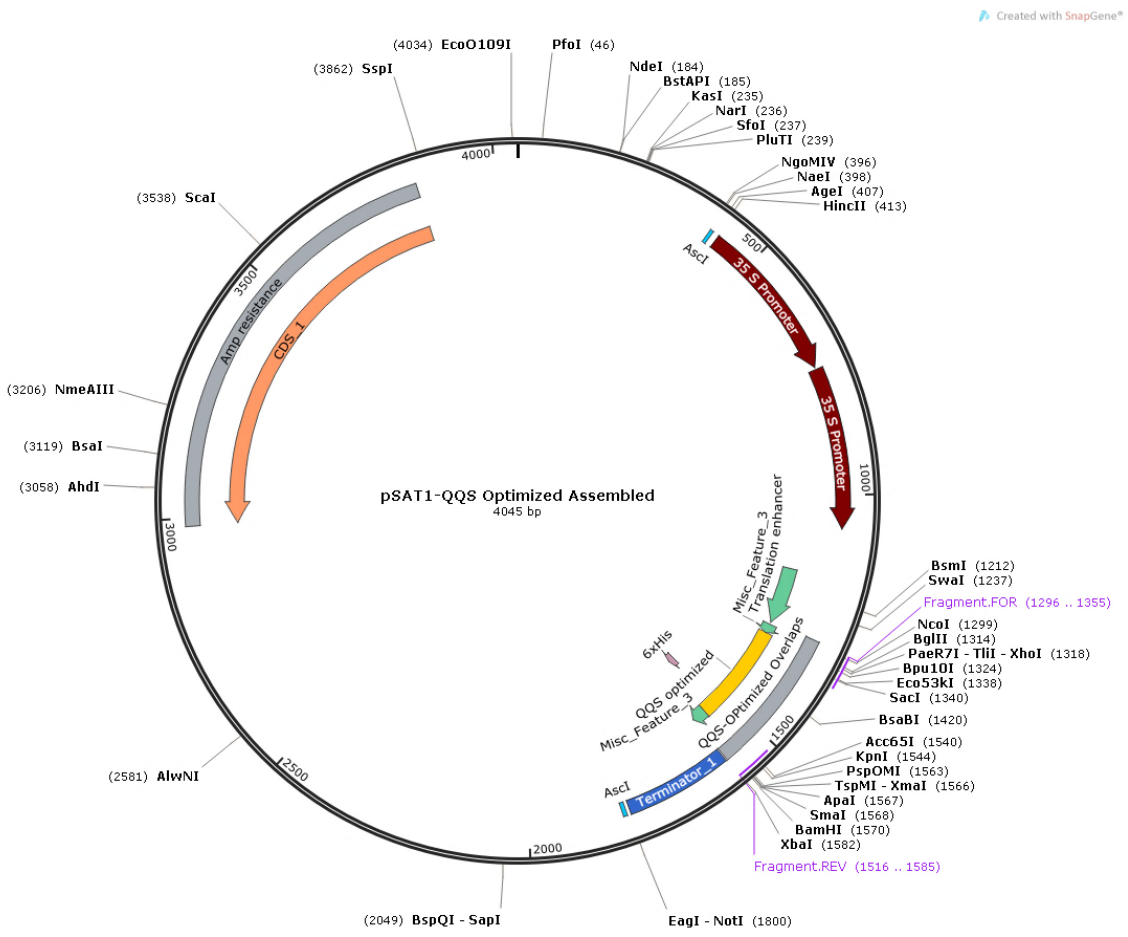


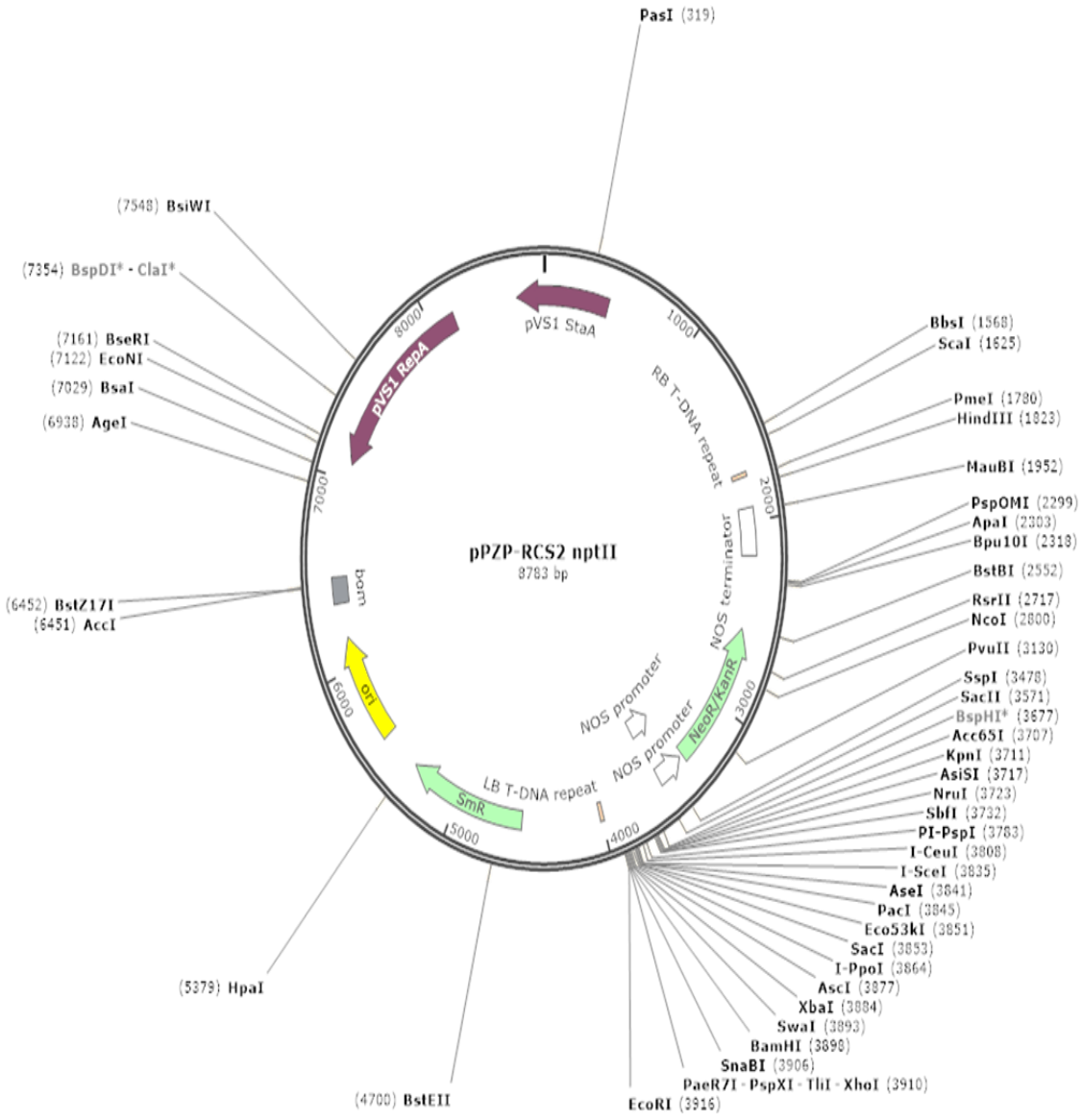
Figure 3*: The hypothetical assembled pSAT1-QQS vector map comporting all its important genetic features obtained through Snapgene software design.*

An aliquot from the Gibson reaction supposed to contain the correct assembled vector (pSAT1-*QQS*) as shown in Figure 3 above was transformed by heat shock into E. coli NEB-5-alpha competent cells. The reaction was cultured on LB plate supplemented with ampicillin (100g/L) for an overnight incubation at 37 °C. To confirm proper assembly of the fragment in pSAT1 vector, a colony PCR was done on 12 wells agar grown selected colonies.

Colony PCR involves the breakdown of the cell wall of the bacteria to liberate the plasmid DNA which is later amplified through PCR reaction. For that, each selected colony was diluted in a 2-mL tube containing 20 µL of lysis buffer (TE+0.1% Triton-X100), then boiled for 20 minutes at 350 °C. The reaction was then spun down at 1300 x g for 20 min. 2 µl of the supernatant was used as template in a PCR reaction (Gotaq green Master Mix 2x, Promega, USA) with the synthesized Gibson *QQS* check primers (Table 2). In addition to the colony PCR to check the success of the Gibson assembly, a double digestion of the assembled pSAT1-*QQS* vector with SacI and KpnI was done to confirm proper insertion of the *QQS* optimized gene in this vector.

3.1.3 Insertion of *QQS* expression cassette from pSAT1-*QQS* assembled vector into Ppzp final transformation vector through ligation dependent technique

pPZP is a plant expression binary vector pSAT/pPZP-RCS as described in Tzfira et al. 2005, 2007; Zeevi et al. 2012 (Figure 4). *QQS* confirmed single-gene expression cassette was obtained from pSAT1-*QQS* assembled vector through restriction digestion with a homing enzyme AscI (NEB, USA) for its insertion into the pPZP final vector. The digestion product, *QQS* expression cassette was ran on gel, then *QQS* cassette was gene purified using QIAGEN mini elute gel extraction kit (QIAGEN, USA) and following manufacturer protocol. pPZP final vector was also linearized through double digestion with AscI homing enzyme (NEB, USA); the product of the reaction, a linearized pPZP was dephosphorylated with the shrimp alkaline phosphatase (NEB, USA) following instruction from the manufacturer's *QQS* cassette was inserted into the linearized and dephosphorylated pPZP vector by ligation using T4 DNA ligase (NEB, USA).



<u>Figure 4</u>*: Plant expression binary vector pSAT/pPZP-RCS with all genetic features.*

For optimal insertion into pPZP final vector, an insert/vector molar ratio of 3:1 was used. The amounts of insert and vector in µl were calculated using the following formula: [ng of vector] * [kb size of insert]/ [kb size of vector] *[insert: vector ratio]. An aliquot of the ligation product was transformed into *E. coli* DH5α competent cells; a 100 µL of the reaction was spread on LB

plate supplemented with spectinomycin and grown overnight at 37 °C for vector characterization. Colony PCR was done on 13 well grown colonies, according to the protocol previously describe in this chapter. In addition, the proper cloning of *QQS* cassette into pPZP final vector was also confirmed through restriction digestion with ASCI and confirmed on 1 % gel electrophoresis. Colony PCR was also performed on these colonies and 5 positive colonies were sequenced at Macrogen, USA. The resulting trans-conjugant was used as vector system for cassava transformation (Hankoua et al. 2006; Taylor *et al*. 2012). Correct *QQS* expression construct (pPZP-*QQS*) was mobilized into disarmed *A. tumefaciens* strain LBA4404 (Lazo, 1991) by heat-shock. Colony PCR was done on 10 well grown *Agrobacterium* colonies to conformity of *QQS* construct into disarmed *A. tumefaciens* strain LBA4404.

3.1.4 **Insertion of *QQS* gene inside cassava genome and regeneration of cassava plant**

To insert *QQS* construct inside the cassava genome, friable embryogenic cassava callus (FEC) (illustration in Figure 7) was used as target tissue in *Agrobacterium* mediated DNA delivery with paromomycin as a plant selectable antibiotic system. The transformation and regeneration of cassava was done following the techniques of cassava transformation and regeneration developed by Hankoua et al. 2006 and Taylor *et al*. 2012. Briefly, LBA4404 cells comporting our trans-conjugant were grown overnight into LB supplemented with 50% Kanamycin, 30 Rifampicin and 30% Streptomycin. This starting culture was used to inoculate 25 ml of Yeas Mold (YM) media broth (composition for 1 L of media: yeast extract 3 g, malt Extract 3 g, peptone 5 g, dextrose 10 g) containing the same antibiotics and grown until an OD600 between 0.75 and 1. Then, 42 ml of that culture was washed twice with GD2 50P (composition in 1L of media: 40 ml of the GD macronutrient solution (25 times stock), 5 ml of the Fe-EDTA solution (200 times stock solution),

1 ml of micronutrients solution (1000 times stock), 1 ml of GD vitamin solution (1000 times stock), 50 ml of picloram solution (1 mM stock), 20 g of sucrose;) liquid media with 200 µM acetosyringone to obtain a final OD600 of 0.5. To infiltrate cassava FEC, 1 ml of GD2 50P containing 200 µm of acetosyringone was placed in each well, then 0.8 cm3 of cassava FEC (about 3 big clusters) were added to each well. The excess of water was removed, and then 3 ml of *Agrobacterium* suspension was used to infiltrate the cassava somatic embryogenic tissues. To optimize cell infiltration, wells with inoculated cassava tissues were placed under vacuum for 30 mins. Then, using an electronic pipette, the suspension of tissues was spread onto a mesh and the excess of liquid drained on filter papers. The mesh was then placed on co-culture plate (GD2 50P+100 µM acetosyringone) and incubation at 22 °C full light for 2 days.

At the end of the co-culture period, tissues were suspended and washed twice into GD2 50P containing 500 mg of carbenicillin then placed on a recovery media plate (GD2 50P+500 mg/l carbenicillin) to initiate cell division from co-cultured tissues. After 7 days on the recovery plate, the tissues were placed onto the FEC selection plate (GD2 50P+500 mg/L carbenicillin+ 27.5 µM paramomycin) for 4 weeks. The 27.5 µM paramomycin is not suitable for cassava cell proliferation and it is used for routine selection of putative transgenic cell growing on a selection plate.

Cassava cells that survived and proliferate on 27.5 µM paramomycin were placed onto embryo regeneration media (MS2 0.5NAA+45 µM paramomycin) for 3 to 6 weeks to encourage embryo emergence. At this stage of the transformation process, maturing embryos (illustration in Figure 7) were monitored and picked regularly and transferred onto embryo germination plate (MS2 2BAP: 4.31 g/l of MS, 1 ml of MS vitamin solution (1000 times), 20 g of sucrose, 2 ml of BAP (1Mm)) for 3 to 4 weeks for putative transgenic shoots emergence. This embryo germination

process was done in a cyclic manner and as new putative transformed maturing embryos emerged from maturation plate.

Finally, shoots were grown on plant maintenance media (Composition in 1L of MS2 phytogel: 4.31 g/l of MS powder, 1 ml of MS vitamin (1000 times), 20 g of sucrose, 4 ml/L of CuSO4 2 Mm, 2.2 g/l of phytogel) for clonal maintenance and propagation of putative transgenic plant lines. After 3 weeks, well grown plantlets were transferred from MS2 media plates to soil according to laboratory protocol for hardening. Briefly, jars were half filled with Fafard 52 soil (Maryland plants & supplies INC., MD, USA) then placed in a larvicide (Gnatrol) mixture (1Tsp of gnatrol/1Gallon of water). After soil treatment, the plantlet was removed from the MS2 plate with precaution to avoid root system damage; then the remaining of gel on roots was washed with running water. The plantlet was then laid on the treated soil and covered with additional soil while leaving the apical bud exposed. Soil established plantlets was watered and the potted plantlets were placed on a tray and covered with a low dome for about a week to facilitate adaptation to new environment. To avoid the effect of nutrient deficiency of young cassava plantlets, soil was treated with fertilizers after every 3 weeks after transfer to soil. Nitrogen-Phosphate-Potassium (NPK) fertilizer, USA (1.3g/L of water) and all-purpose plant fertilizer, Miracle Gro, USA, (1.5 tsp / Gallon) were added, by alternation, ones a week until biomass harvest for downstream characterization. Incubations for *In vitro* cassava plantlets growing either on an agar media plates or in soil, was done in the same growth chamber under long day (LD) conditions at 28°C and 80% relative humidity.

3.2 MOLECULAR CHARACTERIZATION OF REGENERATED CASSAVA PLANT

Preliminary data to assert of the insertion of *QQS* gene into genome of the transgenic lines was obtained by PCR reaction at different stages of cassava plant regeneration and developmental process. Biomass were obtained from putative transgenic: 1)- maturing embryos; 2)- 3 weeks old cassava plantlets growing on MS2 cassava maintenance medium; and 3)-6 weeks old well-established soil grown cassava plants. Transgenic biomass from 6 weeks old well-established soil grown cassava plants were used for quantitative RT-PCR to check the level expression of the *QQS* gene in different transgenic lines as well as perform their biochemical characterization.

3.2.1 QQS determination by PCR

Tissues (leaves, roots and stem separately) from randomly selected young plantlets of transgenic lines were harvested and grinded in liquid nitrogen with a mortar and pestle. 100 mg of the grinded tissues were used to extract DNA using Genejet Plant Genomic DNA purification Mini kit (Fisher, USA) and following manufacturer's protocol. 250 ng of the extracted DNA, for each line, was used as template for the PCR reaction (Gotaq green Master Mix 2x, Promega, USA) using QQS and NPTII genes specific PCR primers. For soil grown cassava plants, DNA extraction was done on newly fully-expanded leave and roots freshly harvested and grinded in liquid nitrogen.

3.2.2 Quantitative Real-Time Reverse Transcription-PCR (qRT-PCR)

To evaluate *QQS* expression or transcript level in root and leave tissues of different transgenic lines, total RNA was extracted from newly fully-expanded leaves and roots of each selected transgenic lines and wild-type non-transformed. After harvested, tissues were grinded into

liquid nitrogen using a mortar and pestle. RNA was extracted using Spectrum Plant Total RNA kit (Sigma, Aldrich, USA). For each line, 4 μg of the extracted RNA was treated with Dnase using Deoxyribonuclease I the kit, amplification grade (Invitrogen life technology, USA). After Dnase treatment, 16 μl of the resulting reaction was used to synthesize cDNA using the kit superscript III first-strand synthesis system for RT-PCR (Invitrogen life technology, USA). cDNA concentration was checked on Nanodrop; then samples were normalized to 1.25 G/μl for their use in real time PCR reaction. RT-PCR *QQS* primers were designed and sent to IDT for synthesis. After reception, RT-PCR was carried out using Power SYBR Green Master Mix (Applied Biosystems, USA). Cassava tubulin gene was used as internal control or housekeeping gene. Three technical replicates were used for each line and the wild-type non-transgenic line was used as the reference sample. Relative gene expression was calculated using $\Delta\Delta C_T$ method available on the 7500 and 7500 fast real-time PCR system version 2.0.1 software. The mean and standard error of the RQ values collected from the real-time PCR data were used to compare gene expression level between lines.

3.3 BIOCHEMICAL CHARACTERIZATION OF WILD-TYPE AND QQS TRANSGENIC CASSAVA LINES

Leaves, stem and roots of 6 weeks *in-vitro* soil grown wild-type, five selected *QQS* stable transgenic lines (R'''(G) L1, R'''(G) L2, R'''(G) L3, R''' (LA) L2, R7 (F)) including NPTII expressers were harvested, grinded in liquid nitrogen then stored at -80°C in 50 ml sterile tubes. These samples were later used for biochemical analysis to evaluate and compare protein and Amino-Acid composition, soluble total carbohydrate content between wild-type non-transformed and stably transgenic lines.

3.3.1 Soluble total Protein extraction and quantification

Soluble total protein extraction and quantification was performed on leave, stem and root tissues of wild-type non-transformed and selected transgenic lines. For each line, 100 mg of grinded tissues were used, and each extraction was done in triplicates. Protein extraction was done by using 1 ml of cold extraction buffer (200 mM NaCL, 1 Mm Tris-HCl Ph 7.8, 4% of 2-mercaptoethanol and 1% of complete protease inhibitor). After extraction, the samples were vortexed and centrifuged to obtain protein pellet. The pellet was suspended into 100 µl buffer and the resulting solution used for protein quantification. Protein content was determined using CB-X kit (G Biosciences, Maryland Heights, MO), following the manufacturer protocol. The absorbance of the samples at 595 nm was used to determine the protein concentration (µg/µ) according to an equation obtained by generating a standard curve using Bovine Serum Albumin (BSA) 2 mg/ml and according to the CB-X kit protocol.

3.3.2 Soluble total carbohydrate extraction and quantification

Soluble total carbohydrate extraction and quantification was performed on the leave and root tissues of wild-type non-transformed and selected transgenic lines. For each line, 50 mg of grinded tissues were used with 2 replicates per line. Soluble total carbohydrate was extracted using 200 µl of an ice-cold Assay buffer supplied by the soluble total carbohydrate Assay kit (Sigma-Aldrich) used for the analysis. After extraction, samples were centrifuged (13,000g/5min) and the pellet were re-suspended in 30 µl buffer and the resulting solution used for soluble total carbohydrate quantification following manufacturer's protocol. For samples that had less than 30 µl of pellet resuspension solution, water was used to adjust the volume to a final volume of 30 µl as required by the protocol. The absorbance of the samples at 490 nm was used to determine the

soluble total carbohydrate concentration (μg/50 mg) according to an equation obtained by generating a standard curve using a D-glucose standard 2 mg/ml and according to the soluble total carbohydrate assay kit protocol.

3.3.3 Amino-acid analysis

The grinded samples were dried overnight in an oven at 100 °C until complete drying. The Amino-Acid analysis was done at the USDA, Wyndmoor, PA laboratory. The method followed in the analysis is described in Cohen and De Antonis, 1994. To describe briefly, dehydrated transgenic plant samples were subjected to hydrolysis using 6.1 N HCl containing a small amount of phenol. The hydrolysis flasks were extensively purged of oxygen using a PicoTag workstation (Waters Corp., Milford, MA), and then incubated at 110°C for 20h. Hydrolyzed samples were dried under vacuum, and derivatized with AccQFluor reagent (Waters) according to the manufacturer's directions. Chromatography was performed using procedures described as 'system 1' in Cohen and De Antonis (1994), with α-aminobutyric acid as an internal standard. Separation was achieved using an AccQTag C18 reverse phase column (Waters); detection by fluorescence using excitation with 250 nm light and measured emission at 395 nm. Hydrolysis, derivitization and analysis of each sample were performed in triplicate.

3.4 STATISTICAL ANALYSES

Statistical Analysis System (SAS) Enterprise Guide 5.1. Experiments were done in duplicate for soluble total carbohydrate content and in triplicates for soluble total protein content. ANOVA was used to check differences between the mean values. Significant probability was set at $p \leq 0.05$. In addition, excel Microsoft version 2016 was used to check correlations between *QQS*

transcript level in root and leave and soluble total carbohydrate or soluble total protein content of these plant parts. Correlations analysis were also performed to determine any relationship between carbohydrate and protein content in leaves and roots.

CHAPTER 4: RESULTS AND DISCUSSION

Our project concept took advantage of the availability of the full length sequence of *Arabidopsis thaliana* qua-quine starch (*QQS*) sequence (GenBank: NM_113975) is 654 bp and data demonstrating that the expression of this transcription factor in a high protein staple crop, effectively induces a regulation of the metabolic processes affecting partitioning of carbon and nitrogen leading to increase in transgenic seed proteins and decrease of carbohydrates of field grown plants (Li *et al.,* 2009; Li *et al.* 2015)**.** For this project, a sequence of 230 base pairs of the *QQS* gene was used to generate a chimeric construct to generate the various stably transgenic cassava lines.

The availability of the complete nucleotide sequence of the cassava genome (Prochnik *et al*., 2012). The Cassava Genome: Current Progress, Future Directions Trop Plant Biol. 5(1): 88–94) prompt us to optimize the sequence of the *QQS* gene for better expression in cassava by taking advantage of the preferred codon usage of the gene expression machinery of cassava. The sequence alignment of the original and optimized *QQS* sequences are shown in Figure 5. In this Figure 5, nucleotides in "red" in the optimized sequence are codons that are preferred by the cassava expression machinery. Expression of this optimized *QQS* codon usage in transgenic cassava plants could increase its expression.

Optimized	17	ATGAAAACTAACAGAGAACAGGAAATCTACGTTGAGAGGTCTTTTAAGCCAAACAATAGC
Original	17	ATGAAGACCAATAGAGAGCAGGAAATTTACGTTGAAAGAAGCTTCAAACCAAACAATTCA
Optimized	77	ACAATTCAGAACTTGATGGATATCGAGAGGTTTATTCTTCCACATACTTCTACATCAGGA
Original	77	ACAATTCAGAATTTGATGGACATTGAAAGGTTCATTTTGCCTCACACTTCTACATCAGGT
Optimized	137	GTTGCTAGGTTGAAGATGAGAGTTATTAGCTGGGTGGGTCTTCAGTTCTACAATTATCAT
Original	137	GTCGCAAGGCTCAAAATGAGGGTCATATCATGGGTCGGGCTTCAGTTCTACAACTACCAC
Optimized	197	CATCATCATCATCATTAA
Original	197	CATCACCATCACCATTGA

Figure 5*: Sequence alignment of the original and optimized QQS sequence. In red are preferred codon usage for efficient gene expression in cassava.*

4.1 MOLECULAR CLONING RESULTS

4.1.1 Characterization of the insertion of the *QQS* optimized sequence in pUC57-*QQS*, pSAT1-*QQS*, and pPZP-*QQS*

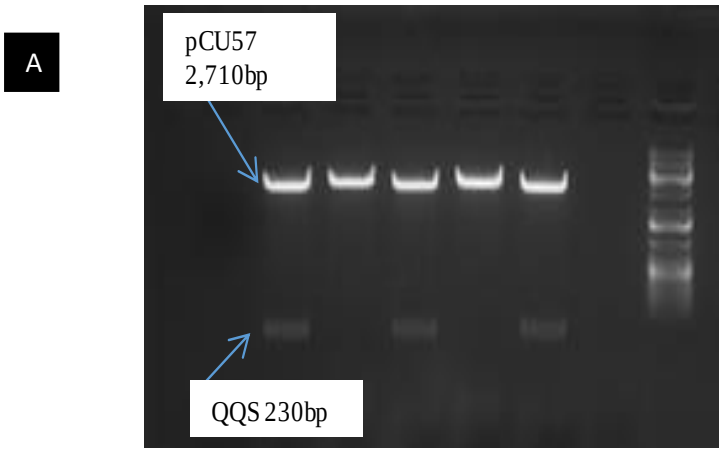
Characterization of *QQS* optimized sequence from pCU57-*QQS* (see Figure 1 showing the map of pCU57) was done by double restriction digestion with KpnI and SacI restriction enzymes of pUC57 vector carrying our gene of interest. Gel electrophoresis of the restricted fragments produced two bands of different sizes, a higher band of 2,710 bps and the lower band of 230bp both representing pUC57 backbone vector and *QQS* optimized sequence (Figure 6 section A). Gel electrophoresis results in Figure 6 Section B1 showed the PCR amplicons (720 bps) using primers (*QQS* check) designed to check the insertion of *QQS* in pSAT1-*QQS* (See figure 3 showing the map of pSAT1-*QQS*) after the Gibson assembly reaction.

Figure 6 Section B2 showed a gel electrophoresis result of double digestion of the pSAT1-QQS with SacI and KpnI enzymes to check proper insertion of *QQS* into pSAT1. Gel electrophoresis of these restricted fragments from pSAT1-*QQS* produced two bands of different

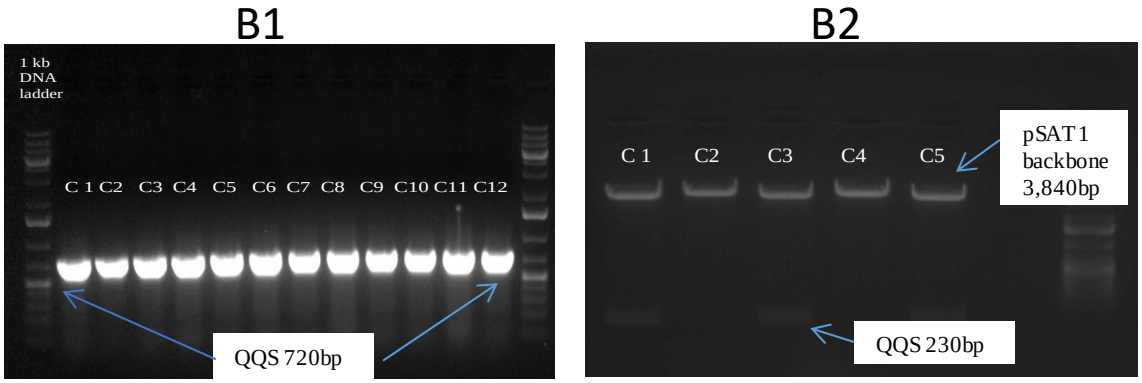
sizes, a higher band of 3,840 bp and the lower band of 230 bp both representing pSAT1 backbone vector and *QQS* optimized sequence respectively (Figure 6 section A). Figure 6 Section C1 showed a gel electrophoresis result of digestion of the pSAT1-*QQS* with the homing endonuclease AscI to extract the *QQS* expression cassette (promoter sequence, *QQS* optimized gene, enhancer, Histag, terminator sequence) from pSAT1-*QQS* vector as well as the linearization with AscI of the final destination vector pPZP of the *QQS* expression cassette. Gel electrophoresis of these restricted AscI fragments of the pSAT1-*QQS* plasmid produced two bands of different sizes, a higher band of 2637 bp and the lower band of 1433 bp both representing pSAT1 backbone vector and *QQS* expression cassette respectively. The *QQS* expression cassette (1433bp) resolved on the gel has all genetic elements (promoter sequence, *QQS* optimized gene, translation enhancer, Histag, terminator sequence) required to drive the expression of the *QQS* optimized gene in stably transgenic cassava plants. In Figure 6 Section C1 a gel electrophoresis of the AscI restriction digestion of the pSAT1-*QQS* showing a higher and linear band of very high molecular weight, approximately 8,475 bp. The size of this linear band is near the size of the circular pPZP vector. pPZP-RCS binary vector contains T-DNA left and right borders and carries the spectinomycin gene for bacterial resistance and the gene encoding the neomycin phosphotransferase II (*nptII*) as a plant selectable marker within its T-DNA borders (Figure 4 that showed map of the pPZP). The AscI restriction QQS expression cassette (1433bp) fragment were ligated with the linearized pPZP (8,475 bp**)** vector fragment to generate the assembled pPZP plant transformation vector (pPZP-*QQS*). Figure 6 Section C2 showed positive colony PCR of thirteen isolated E. coli colonies transformed with the assembled pPZP-*QQS* vector. Gel result of the colony PCR using *QQS* specific primers check showed that *QQS* optimized sequenced is properly inserted in the pPZP-*QQS* vector. Figure 6 Section D showed colony PCR results of ten isolated *Agrobacterium*

colonies transformed with pPZP-*QQS* vector. Gel data showed that *QQS* optimized sequence can be amplified from the pPZP-QQS plasmid isolated from transformed *Agrobacterium* strain LBA4404. These latter data confirmed that a novel LBA4404 trans-conjugant comporting a fully characterized pPZP-*QQS* vector is generated will be used as vector system for inserting optimized *QQS* in transgenic cassava via *Agrobacterium*-mediated transformation. Figure 6 Section E depict the fully characterized pPZP-*QQS* expression cassette obtained through various molecular cloning manipulations described in this section.

Figure 6: ***Molecular cloning results***

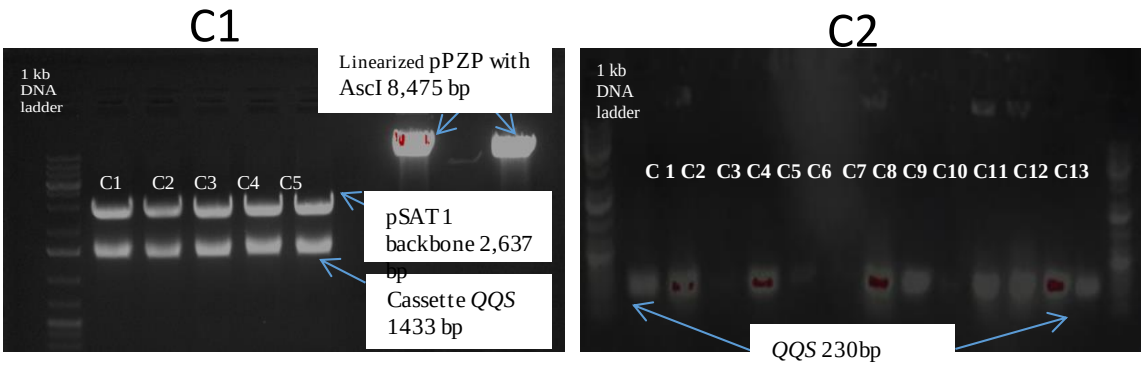


Section A : ***Characterization of QQS optimized gene sequence through double digestion with SacI and KpnI and gel electrophoresis.***



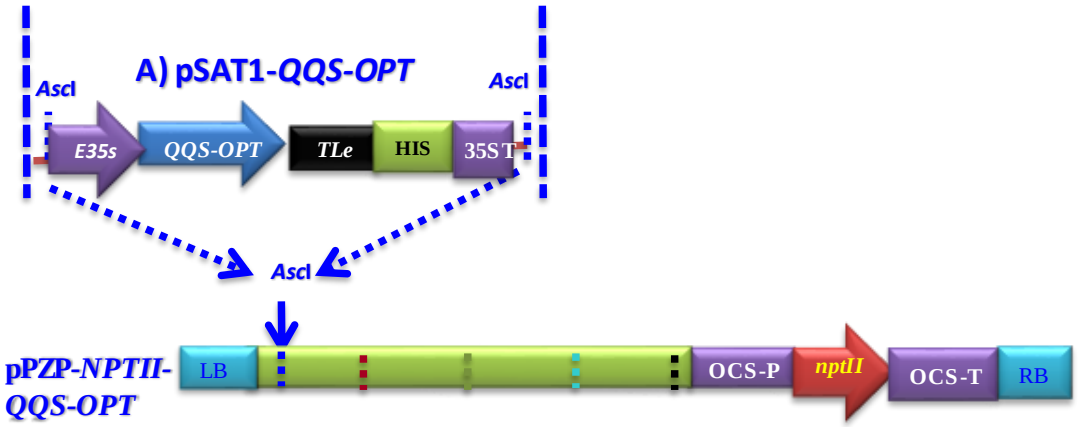
Section B: ***QQS check into pSAT1 using QQS check primers and double digestion with SacI and KpnI to check proper insertion of QQS into pSAT1.***

C



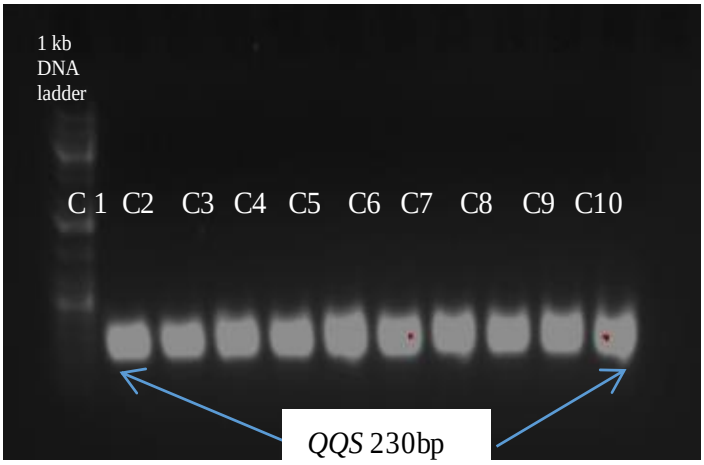
Section C: ***Extraction of QQS cassette from pSAT1 vector and linearization of pPZP vector with AscI (C1). Colony PCR of thirteen isolated colonies to check proper insertion of QQS cassette into Ppzp (C2).***

D



Section E: ***Schematic representation of the QQS-pPZP expression construct for constitutive expression of optimized QQS in transgenic cassava.***

E



Section D: ***ten isolated Agrobacterium colonies to check the proper insertion of QQS-pPZP construct into LBA4404 used to transform the cassava Friable Embryogenic tissues shown on figure 7 section B.***

***Table 2*: List of specific primers designed for molecular cloning**

Gene name	Forward primer	Reverse primer
QQS gene specific primers	aaggaaaggagagctcatga	tcctttccttggtaccttaa
Gblock *QQS* fragment primers	tagccatggtccggactcagatctcgagctaaggaaaggagagctcatgaaaactaacag	tagactaggtggatcccgggcccgcggtactcctttccttggtaccttaatgatgatgatgatgatgata
Gblock *QQS* check primers	aagggatgacgcacaatccc	gcgtcactggattttggttt
NPTII primers	Atggggattgaacaagatggattgc	gaagaactcgtcaagaaggcgatag

4.2 PLANT TRANSFORMATION RESULTS

Figure 7 summarizes the steps taken for the transformation of pPZP-NPTII-*QQS* construct into cassava friable embryogenic tissues via Agrobacterium tumefaciens strain LBA4404.

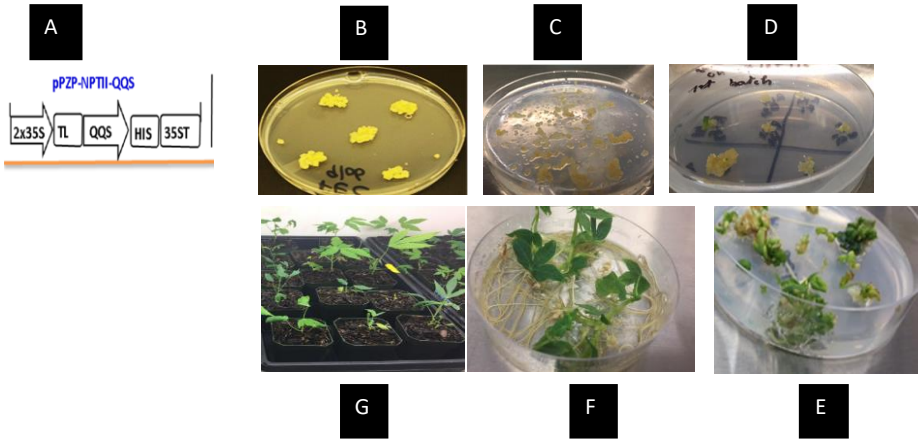


Figure 7-A-G: ***Steps taken in cassava plant transformation and hardening of transgenic plants.***

The *QQS* construct (A) was inserted into cassava embryogenic friable (B) generated from cassava immature leaf lobes. After infiltration of the *Agrobacterium* carrying the *QQS* cassette, transformed cassava tissues were placed in co-culture plate (C) for 6 days under full light. Growing clusters from co-culture plate were placed on 1st embryo regeneration (D) then matured embryo were transferred on germination plate (E). Plantlets were maintained on MS2 phytogel media (F). Cassava plantlets, transgenic and wild-type, were transferred to soil after 3 weeks on MS2 phytogel plate (G).

4.2.1 Morphological comparison between transgenic plants and Wild-Type Plant

The *Agrobacterium*-co-culture and plant recovery generated 16 independent putative transgenic *in vitro* cassava plants. Seven of these transgenic plants were lost due to recurrent *in vitro* microbial contaminations. *In vitro* transgenic plantlets look normal with no detectable morphological anomalies as compared with their non-transformed counterparts (WT) and NPTII (empty vector) expressers (illustration 1, for some *in vitro* transformed plants). *In vitro* plantlets showed a 100% survival rate after transplantation in soil (illustration 1, for some soil grown plants). Symptoms of fasciated leaf, tiny stem, and slow growth were observed in two soil established lines (R7 (E), R7 (D) illustration 3 section B and Section A). Visual comparison of plant height and biomass production at 6-7 weeks after soil transplantation between wild-type, NPTII and each transgenic cassava is shown on illustration 2. No marked variation in these morphological parameters was observed between these lines except line R7 (B), showing slow growth with poor biomass formation compared with the wild type non-transgenic plant and NPTII expresser (illustration 2). Among all these transgenic cassava lines, line R7 (F) exhibit a particular

morphological phenotype in relation to its growth habit and capability to make huge among of biomass (illustration 3 Section A).

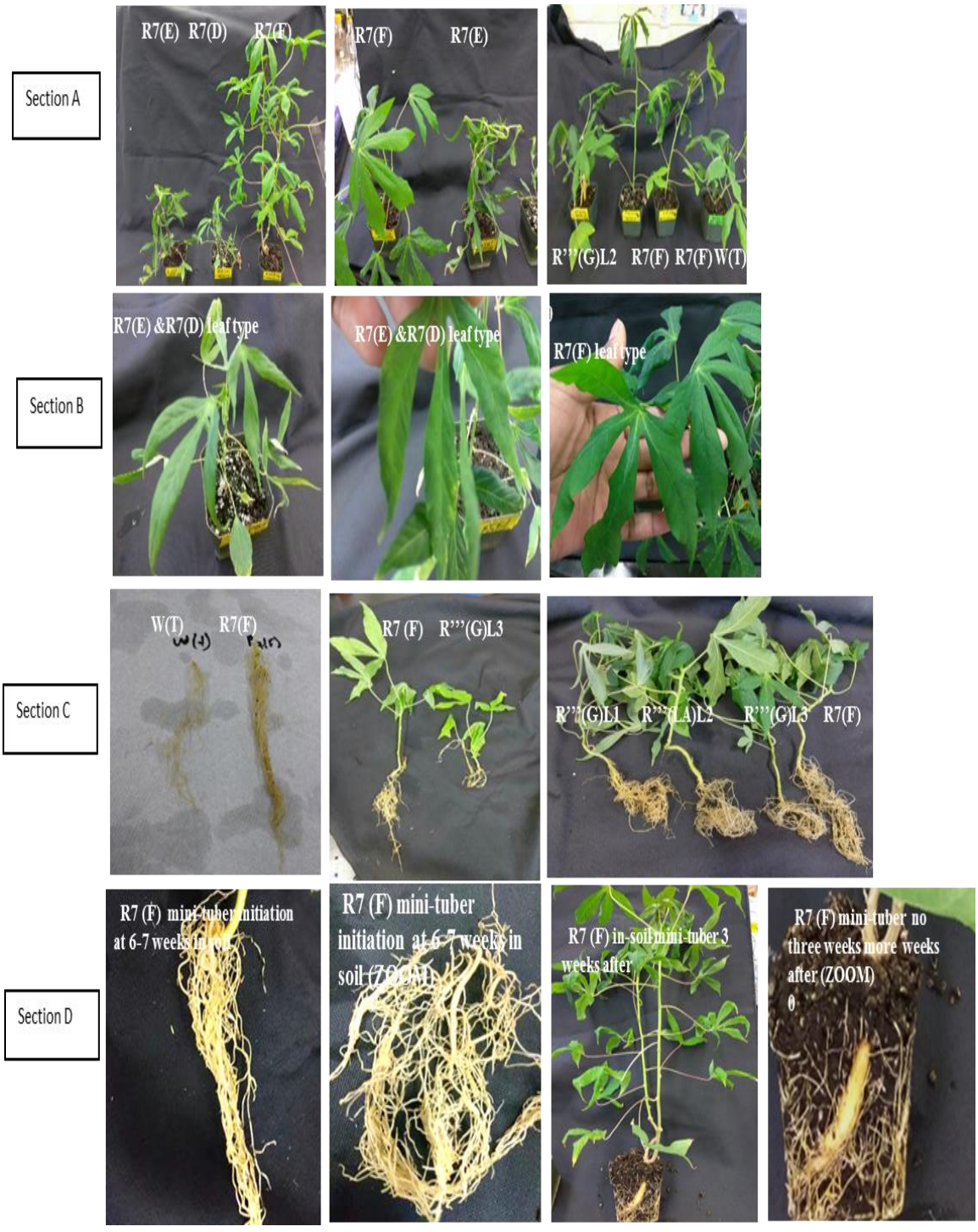
Section A illustrates the capacity of R7 (F) to grow profusely (higher and producing more biomass) compare to other transgenic lines such as R''' (G) L2 and W (T) non-transgenic line. In the same section, pictures illustrate also morphological differences between lines R7 (F), R7 (D) and R7 (E). R7 (D) and R7 (E) have tiny leaves and stem, and slow growth rate compare to R7 (F). Leaf morphology of these lines are also illustrated in section B which offer a better visualization of the disparity observed.

Moreover, illustration 3 section C shows the disparities in plant height, leaf production and root size. R7 (F), root size compared to the wild-type root size; then its plant and root size was compared with plant and root size of line R''' (G) L3. From this picture, it is clear that R7 (F) has the capability to make more shoot and root biomass compared with line R''' (G) L3 and wild-type non-transformed plants. Finally, from the same section R7 (F) plant size and root size were compared with the ones obtained with R''' (LA) L2, R''' (G) L3 and R''' (G) L1. It is important to mention that this excellent growth ability of line R7 (F) is unique in comparison with all other transgenic lines and control plants. All plant was transferred from *in-vitro* to soil, under the same growing conditions and they were evaluated at the same age (6-7 weeks). The final section D, illustrates the capability of R7 (F) to induce and produce harvestable mini-tubers early after *in-vitro* plantlets were transferred to soil between 6-7 weeks; while the other transgenic lines and wild type non-transformed control did not exhibit this capability after more weeks of growth in soil.

Illustration 1: Picture showing in-vitro soil established transgenic lines compared with WT and NPTII transgenic lines.

Illustration 2: *6 weeks in-vitro soil grown cassava plants regenerated from cassava transformation (8 lines), NPTII cassava transformation and the wild type non- transformed line.*

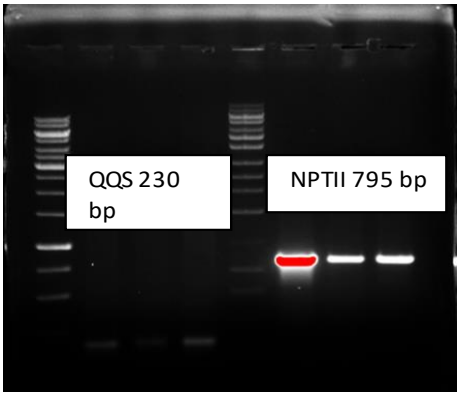
<u>Illustration 3</u>: Particularity of R7 (F) in making more biomass and in initiating mini in-vitro tubers at very early age (6-7 weeks).



4.3 MOLECULAR ANALYSIS OF PUTATIVE TRANSGENIC CASSAVA LINES

4.3.1 Confirmation of transgenicity of tissues through PCR reaction

Gene specific PCR was performed on putative transgenic tissues as described in the material and methods section to check the proper insertion *QQS* optimized gene into the cassava genome. Gene specific PCR was initially performed on DNA extracted from three randomly chosen transgenic lines using *QQS* gene specific primers and NPTII, kanamycin resistance gene primers, to amplify the inserted QQS gene (230 bp) and the NPTII gene (790 bp) used as plant selectable marker to recover putative transformed tissue. Gel pictures in Figure 8-A, Figure 8-B, and Figure 8-C, show a clear amplification of the QQS sequence as well as the NPTII gene for all the transgenic lines and all plant parts screened. PCR amplification of QQS and NPTII in all plant parts (leaves-petioles-roots) of all the transgenic lines generated demonstrate the efficiency of the enhanced 35S and the octopine synthase promoter in driving the expression of these genes constitutively in all these transgenic tissues.



<u>Figure 8-A</u>: ***Gel illustration of PCR amplification of QQS and NPTII genes from putative transgenic tissues on 1st germination plate; 3 months after transformation. Lines checked are from left to right R13, R7, and QR17.***

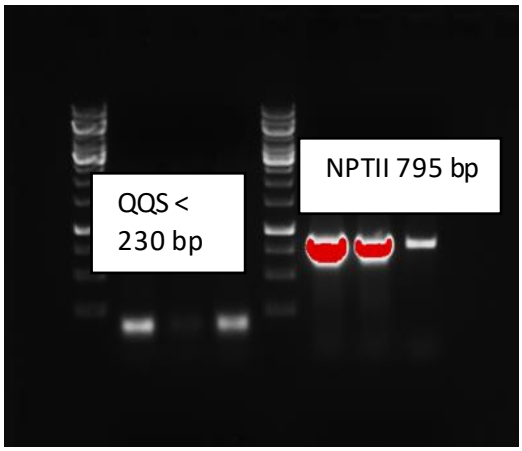


Figure 8-B: ***Gel picture of PCR amplification of QQS and NPTII genes from putative transgenic on 3 randomly chosen lines, from left to right R7(B), R'''G and Old FEC2 about 5 months, after 3 weeks on MS2 phytogel maintenance plate.***

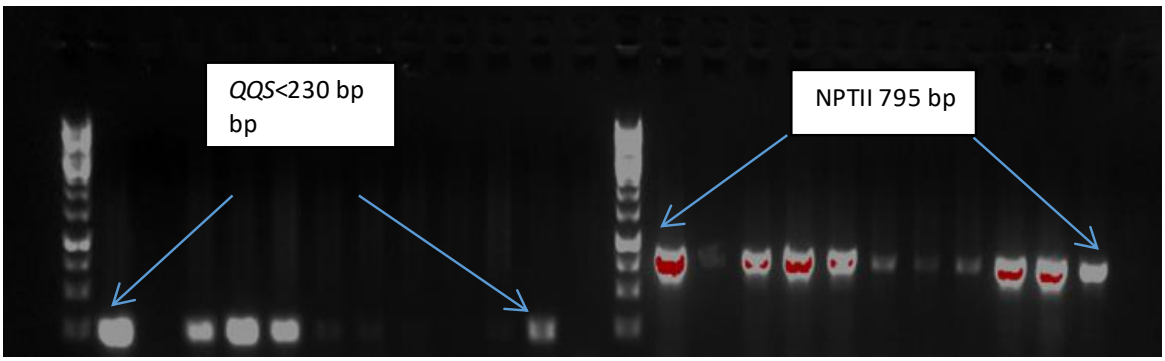


Figure 8-C: ***Gel picture of PCR amplification of QQS and NPTII genes from putative transgenic tissues***.

Gene specific PCR done on transgenic plantlets after about 6 months from 3rd subculturing on MS2 phytogel maintenance media. To check efficiency of the constitutive promoteur DNA extraction was done, for some lines, on different parts of plant (leaves-petioles-roots). The lines checked and their corresponding bands are: 1-R'''LA leaves, 2-R'''LA petiole, 3-R'''LA roots, 4-R'''LB leaves, 5-R'''LB roots, 6-R'''(G) leaves, 7-R'''(G) petiole, 8-R'''(G) roots, 9-R'''(A)

mix of leaves, petiole and roots, 10-R’’’(D), 11-R7 (G), 12-wild-type non-transgenic used as negative control.

4.3.2 Confirmation of *QQS* expression in transgenic lines through Real-Time PCR

The level of expression of the QQS gene in the leaf and root biomass of the wild-type and five transgenic lines was performed using SYBR GREEN master mix and *QQS* RT-PCR primers (designed for the analysis, Table 3) was used as target gene and Tubulin primers was used as internal control. The results obtained from the expression analysis of *QQS* in cassava leaves and roots are summarized in tables 4 for cassava leaves and table 5 for cassava roots. The RQ value of 1 was assigned to the wild-type non-transgenic line, used as reference sample, during expression analysis. The mean and standard error from the relative quantification values (RQ) were used to generate graphs as shown in figure 9A-B. qTR-PCR data in Figure 9 show that the QQS gene is differentially expressed in leaf and root tissues of all the transgenic lines tested. Transcript levels varied widely between lines and tissue types. Line QQSL2 produce the highest transcript level for leaf ad root tissues, while lines QQSL3 and QQS L4 generated the highest compared to the WT reference control. Line QQSL5 generate the least QQS transcript in both leaf and root tissue. The expression of QQS showed a very sharp decrease in the root compared to the leaf tissue in line QQSL2. For root tissue, line QQSL5 and QQSL2 generated lowest level of QQS transcript even below the transcript of the reference wild type non-transformed plants. For root biomass, line QQSL3 and QQSL4 produce the highest QQS transcript levels compared to QQSL1, QQSL2, and QQSL5.

Table 3*: List of specific primers used for Real-time PCR.*

Gene name	Forward primer	Reverse primer
QQS RT-PCR primers	GAACAGGAAATCTACGTTGA	CCACCCAGCTAATAACTCTC
Tubulin primers (TC3055)	CAAGTGCGATCCTCGACATG	GATACCGCACTTGAACCCAG

The real-time PCR for cassava leaves were run twice. The RQ values of both essays were summarized in the table below. For each essay, a technical replicate (TR) of each sample was generated. Wild-type leaf sample (WT) was used as reference sample; therefore, its RQ value is 1.

Table 4*: Results from QQS expression into cassava leaves from in-vitro six weeks' old soil grown cassava plant.*

Line	RQ of 2 essays	Mean RQ	RQ L/RQ WT	Mean RQ	SE
WT1	1	1.209482	1	1	0
WT2	2.247991562				
WT3	1				
WT4	0.589936495				
QQSL1-1	5.538040638	4.5788532	4.57885324	1.8759997	1.343804
QQSL1-2	0.213516623	0.1765356	0.17653559		
QQSL1-3	1.979752064	1.6368594	1.63685945		
QQSL1-4	1.344642162	1.1117504	1.11175044		
QQSL2-1	4.019497871	3.3233217	3.32332174	20.347245	21.26329
QQSL2-2	78.78059387	65.135813	65.1358126		
QQSL2-3	12.61906433	10.433445	10.4334452		
QQSL2-4	3.019353151	2.4964019	2.49640186		
QQSL3-1	3.994669437	3.3027936	3.30279358	1.7787075	1.298468
QQSL3-2	0.165841863	0.1371181	0.13711809		
QQSL3-3	4.152162552	3.4330089	3.43300893		
QQSL3-4	0.292585194	0.2419095	0.2419095		
QQSL4-1	5.050534248	4.1757828	4.17578285	1.8621351	1.475981
QQSL4-2	0.074504897	0.0616007	0.06160067		
QQSL4-3	3.724536896	3.0794479	3.07944794		
QQSL4-4	0.159299403	0.1317088	0.13170878		
QQSL5-1	1.048328638	0.8667584	0.86675835	0.5970746	0.350306
QQSL5-2	0.412582576	0.3411234	0.34112337		
QQSL5-3	1.375174284	1.1369944	1.13699441		
QQSL5-4	0.05251861	0.0434224	0.0434224		

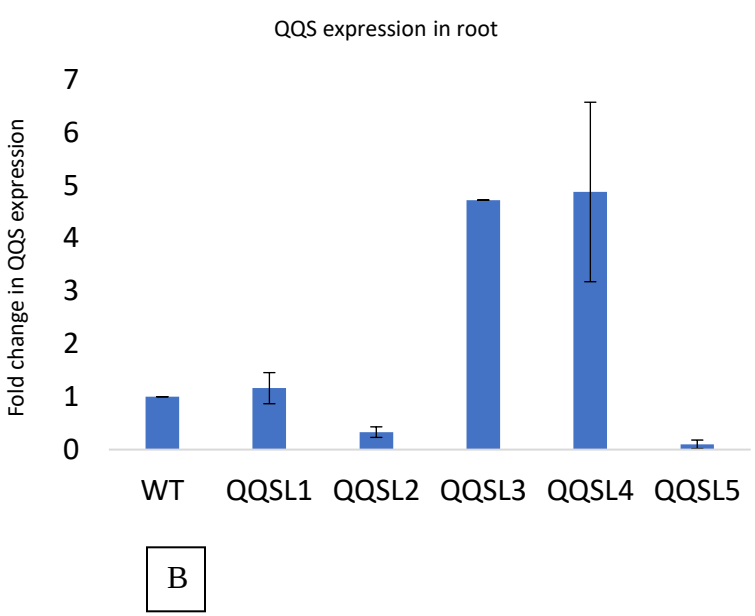
Table 5*: Results from QQS expression into cassava roots from in vitro six weeks' old soil grown cassava plant.*

The real-time PCR data for cassava roots was collected from one RT-PCR run. The RQ values are summarized in the table below. For each sample, a technical replicate (TR) was generated. Wild-type root sample (WT) was used as reference sample; therefore, its RQ value is 1. The missing values for L3-2, technical replicate of L3 were not included in the analysis since they were shown as ''flaws'' on the real-time data result.

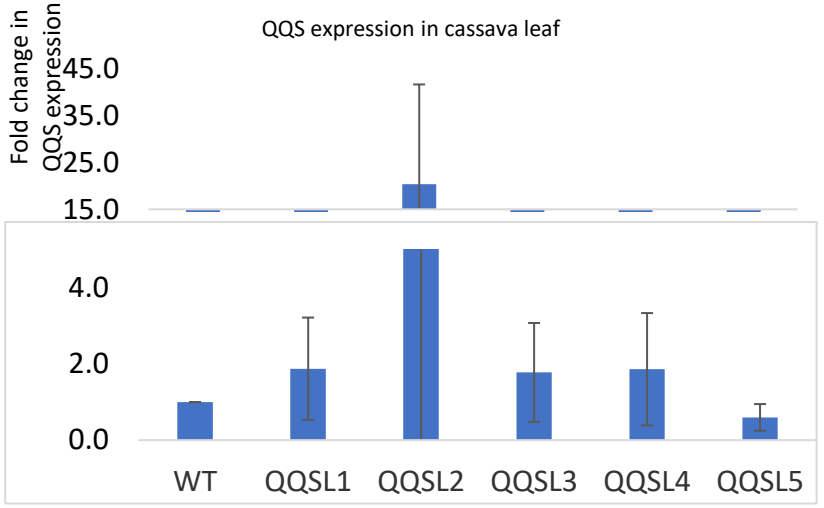
	RQ	Mean RQ	RQL/RQ WT	SE
WT1	1	0.776877254	1	0
WT2	0.55375451			
QQSL1-1	0.67391741	0.867469615	1.161872856	0.2944032
QQSL1-2	1.13134778	1.456276097		
QQSL2-1	0.33411714	0.430077136	0.330299288	0.0997779
QQSL2-2	0.17908686	0.230521441		
QQSL3-1	3.66780615	4.721217063	4.721217063	0
QQSL3-2	.	.		
QQSL4-1	2.46733952	3.175970853	4.874417991	1.6984471
QQSL4-2	5.10630941	6.572865129		
QQSL5-1	0.01452351	0.018694727	0.099943207	0.0812485
QQSL5-2	0.1407637	0.181191687		

<u>Figure 9</u>: ***Quantitative real-time PCR analysis of QQS from leaf (A) and root tissues (B) of six weeks old soil grown transgenic cassava plants.***

The graphs were plotted using the mean RQ of samples from each line and the standard error (SE) obtain after the real time PCR.QQSL1: R''' (G) L1 QQSL2: R''' (G) L2 QQSL3: R''' (G) L3 QQSL4: R''' (LA) L1 QQSL5: R7 (F.)



B



A

4.4 BIOCHEMICAL RESULTS

Biochemical analysis was done on wild-type line, NPTII line (cassava transformed with the pPZP vector exempt of the *QQS* insert) and five confirmed transgenic lines R''' (G) L1, R''' (G) L2, R''' (G) L3, R''' (LA) L2 and R7 (F).

4.4.1 Soluble total protein content from leaves, stem, and roots biomass of stably transformed cassava plants

The Absorbance at 595 nm of each sample was read using Evolution 260 Bio UV-visible spectrophotometer (Thermo scientific, USA), then expressed in protein content (µg/µl) using the following equation: **X= (OD at 595 nm-0.013)/1.2325.** This equation was obtained through generating a standard curve using BSA standard provided by the kit used for analysis and following manufacture's protocol (GBiosciences, USA).

Figure 11 below shows the standard curve generated by using the BSA standard and following manufacturer's protocol. The graph was obtained by plotting the BSA standard concentrations against the OD read at 595 nm Y= 1.23 X+ 0.013; this equation will be used to express the OD values read at 595 nm into soluble total protein content in µg/µl. Figure 12 shows that soluble total protein content of all the lines varied depending of the tissue types analyzed. In general, all the transgenic lines including the two controls and non-*QQS* expressers (WT and NPTII) make more protein compared with the root and stem biomass, with root biomass making less protein from all the lines. For leaf tissues, no significant differences were observed in protein content between all the lines except biomass of line R7 (F) showing slight reduction in soluble total protein content (1.1µg/µl) compared to other lines and the two controls non-*QQS* expressers

(1.3 µg/µl for WT and NPTII). More variabilities were observed in protein content of stem and root biomass, with lines R''' (G) L1 (0.6 µg/µl) and R''' (G) L3 (0.5µg/µl) show a significant protein decrease compared to other lines, especially WT (0.75 µg/µl) and NPTII (0.70 µg/µl) biomass. Stem biomass from other lines exhibit protein level almost similar to the control plants. Variabilities in protein content of root biomass were observed in all the lines, with lines R''' (G) (LA) L2 (0.53 µg/µl) and R7 (F) (0.5µg/µl) show a significant protein increase compared to other lines, especially the WT (0.42 µg/µl) and NPTII (0.40 µg/µl) biomass. Root biomass from lines R''' (G) L1 ((0.35 µg/µl) and R''' (G) L2 (0.55µg/µl) showed lower protein content.

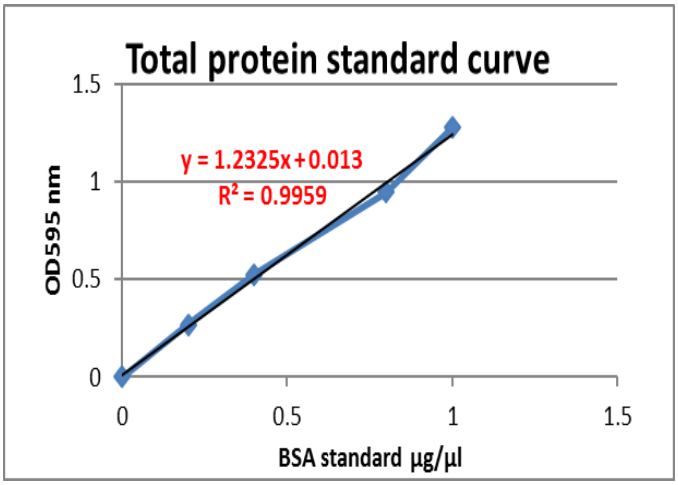
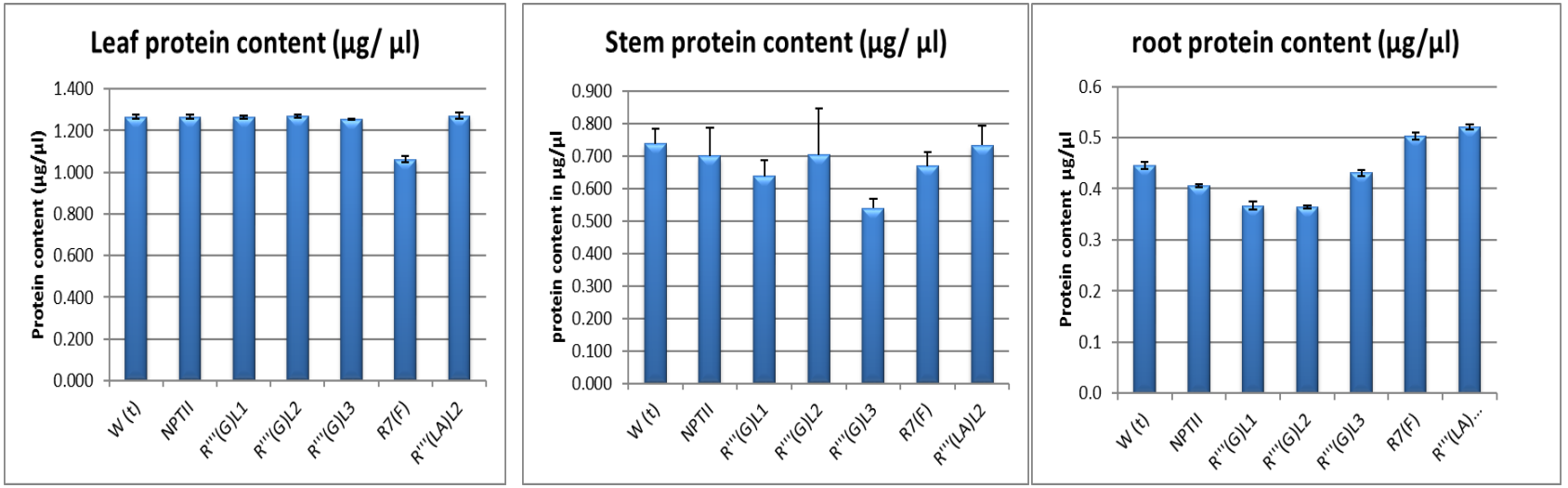


Figure 11: ***standard curve generated by using the BSA standard.***

Table 6*: Soluble total protein extraction and quantification results for cassava leaves, stem and roots from in vitro six* **weeks' soil grown plants.**

Line	Absorbance at 595 nm	Leaf protein content (µg/ µl)	Mean	SE	Absorbance at 595 nm	Root protein content (µg/ µl)	Mean	SE	Absorbance at 595 nm	stem protein content (µg/ µl)	Mean	SE
WT	1.592	1.281	1.264	0.01	0.537	0.425	0.445	0.007	0.859	0.686	0.736	0.027
	1.571	1.264			0.559	0.443			0.973	0.779		
	1.55	1.247			0.588	0.467			0.929	0.743		
NPTII	1.552	1.249	1.266	0.011	0.527	0.417	0.407	0.003	0.994	0.796	0.7	0.05
	1.599	1.287			0.51	0.403			0.844	0.674		
	1.568	1.262			0.506	0.4			0.789	0.63		
R'"(G)L1	1.572	1.265	1.262	0.008	0.501	0.396	0.367	0.008	0.73	0.582	0.637	0.028
	1.551	1.248			0.45	0.355			0.833	0.665		
	1.584	1.275			0.446	0.351			0.831	0.664		
R'"(G)L2	1.585	1.275	1.268	0.006	0.469	0.37	0.365	0.003	1.067	0.855	0.703	0.083
	1.567	1.261			0.448	0.353			0.715	0.57		
	.	.			0.47	0.371			0.856	0.684		
R'"(G)L3	1.553	1.249	1.254	0.004	0.524	0.415	0.431	0.006	0.646	0.514	0.538	0.017
	1.564	1.258			0.54	0.428			0.715	0.57		
	.	.			0.568	0.45			0.65	0.53		
R7(F)	1.284	1.031	1.063	0.016	0.629	0.5	0.503	0.006	0.782	0.624	0.668	0.025
	1.349	1.084			0.612	0.486			0.888	0.71		
	1.335	1.073			0.659	0.524			0.839	0.67		
R'"(LA)L2	1.56	1.255	1.271	0.013	0.674	0.536	0.521	0.005	0.988	0.791	0.733	0.035
	1.612	1.297			0.654	0.52			0.92	0.736		
	1.568	1.262			0.636	0.505			0.84	0.671		

Figure 12: *Quantitative evaluation of protein in leaves, stem and root tissues of six weeks in-vitro soil grown cassava plants. The graphs were plotted using the mean protein content of samples from each line and the standard error (SE) obtained after protein analysis.*



4.4.2 Amino-Acid composition from leaves, stem, and roots biomass of stably transformed cassava plants

Figure 13 shows data of the amino acid composition of biomass harvested from leaves, stem, and roots biomass of stably transformed cassava plants. A total of 18 amino acids were detected in all the samples. Based on these data, the level of each amino acid did not vary significantly between leaves, stem, and roots biomass of stably transformed cassava plants compared to two controls non-QQS expressers (WT and NPTII). Glycine, glutamic acid, aspartame, asparagine, and leucine appeared to be amino acids found in abundance in all the lines and plant parts, but glycine and glutamic acid are mostly abundant in stem biomass for all the lines. The least abundant amino acids in all the lines are histidine and methionine, with methionine showing more abundance in root biomass of all the transgenic lines tested. An important amino acid, cysteine was not detected in any of the analyzed pectin samples. It is possible that some cysteine is present, but undetectable due to technical limitations of the method employed.

Figure 13: ***Amino-acid compositions of biomass harvested from leaves, stem, and roots biomasses of stably transformed cassava plants***

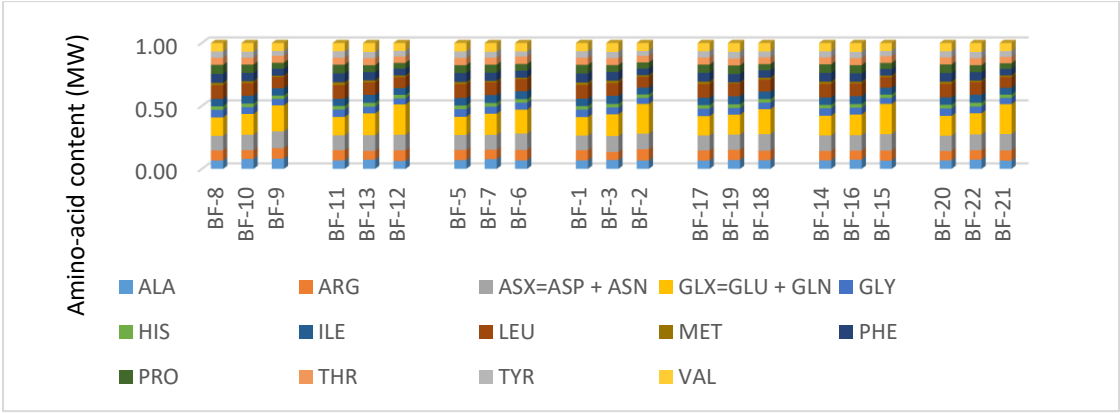


Table 7 A-B-C*: summarizes the results of Amino-Acid composition of each line for leaf, root, and stem*

For each Amino-Acid the mean value of each sample was plotted to obtain the figure above. BF-1 R'"(LA)L2 Leaves, BF-3 R'"(LA)L2 Roots, BF-2 R'"(LA)L2 Stem, BF-5 NPTII Leaves, BF-7 NPTII Roots, BF-6 NPTII Stem, BF-8 W(t) Leaves, BF-10 W(t) Roots, BF-W(t) Stem, BF-14 R'"(G)L1 Leaves, BF-16 R'"(G)L1 Roots, BF-15 R'"(G)L1 Stem, BF-17 R7 (F) Leaves, BF-19 R7 (F) Roots; BF-18 R7 (F) Stem, BF-20 R'"(G) L2 Leaves; BF-22 R'"(G) L2 Roots; BF-21 R'"(G)L2.

Sample type	BF-1		BF-2		BF-3		BF-4		BF-5		BF-6		BF-7	
AA	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
ALA	6.85	0.29	2.28	0.05	2.77	0.13	7.61	0.23	7.34	0.49	2.03	0.05	3.00	0.73
ARG	8.18	0.44	3.04	0.04	2.47	0.24	9.00	0.35	8.57	0.57	2.46	0.04	2.91	0.90
AASX	11.75	0.67	4.24	0.09	4.92	0.31	13.41	0.52	12.63	0.90	3.94	0.06	4.61	0.94
GLX	14.79	0.87	8.09	0.35	6.74	0.48	16.94	0.64	15.36	1.16	5.67	0.08	6.54	1.44
GLY	5.95	0.47	1.62	0.03	2.10	0.11	6.59	0.21	6.37	0.55	1.64	0.04	2.19	0.50
HIS	2.92	0.17	0.93	0.03	1.12	0.05	3.23	0.06	3.09	0.18	0.84	0.03	1.12	0.17
ILE	5.88	0.29	1.86	0.06	2.49	0.11	6.57	0.13	6.34	0.36	1.89	0.07	2.43	0.34
LEU	10.91	0.55	2.88	0.11	3.92	0.18	12.09	0.26	11.54	0.64	2.93	0.11	3.93	0.55
LYS	5.67	0.47	2.41	0.09	2.85	0.16	7.34	0.05	6.43	0.30	2.65	0.07	2.91	0.43
MET	2.20	0.08	0.49	0.04	0.80	0.09	2.31	0.04	1.91	0.57	0.29	0.04	0.46	0.40
PHE	7.11	0.32	1.80	0.09	2.66	0.11	7.84	0.14	7.55	0.42	1.69	0.08	2.55	0.36
PRO	7.04	0.22	1.67	0.05	2.17	0.10	7.67	0.42	6.66	0.92	1.57	0.06	2.46	0.26
SER	5.40	0.32	1.98	0.04	2.63	0.18	6.12	0.25	5.76	0.44	1.80	0.06	2.59	0.52
THR	5.74	0.29	1.79	0.05	2.13	0.12	6.36	0.21	6.11	0.40	1.80	0.03	2.18	0.49
TYR	5.16	0.24	1.33	0.08	1.91	0.09	5.68	0.17	5.39	0.31	1.19	0.08	1.81	0.23
VAL	6.64	0.32	2.17	0.04	2.79	0.11	7.51	0.16	7.20	0.44	2.02	0.06	2.74	0.45

A

Sample type	BF-8		BF-9		BF-10		BF-11		BF-12		BF-13		BF-14	
AA	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
ALA	7.72	0.04	3.27	0.88	2.81	0.40	6.88	0.73	2.09	0.52	2.90	0.13	5.53	1.25
ARG	9.10	0.28	3.39	1.06	2.46	0.17	8.22	0.83	2.77	1.11	2.77	0.04	6.14	1.22
ASX	13.51	0.10	5.43	1.33	4.40	0.32	12.06	1.32	4.09	1.33	4.97	0.22	10.06	2.28
GLX	16.80	0.20	8.36	2.18	5.85	0.57	14.87	1.57	7.97	3.88	6.92	0.39	12.76	2.61
GLY	6.80	0.04	2.06	0.43	1.93	0.11	5.91	0.63	1.57	0.43	2.09	0.08	4.77	1.05
HIS	3.33	0.08	1.09	0.24	1.01	0.08	2.90	0.30	0.92	0.26	1.21	0.06	2.27	0.50
ILE	6.78	0.12	2.33	0.51	2.22	0.16	5.83	0.63	1.79	0.41	2.53	0.10	4.77	1.03
LEU	12.39	0.24	3.67	0.73	3.52	0.26	10.91	1.17	2.85	0.68	3.95	0.16	8.65	1.82
LYS	7.00	0.28	3.24	0.86	2.60	0.21	5.47	0.54	2.23	0.60	2.98	0.12	4.68	1.01
MET	2.32	0.06	0.42	0.04	0.59	0.05	2.46	0.25	0.47	0.11	0.65	0.02	1.64	0.15
PHE	8.11	0.21	2.21	0.38	2.31	0.14	7.17	0.77	1.80	0.38	2.59	0.09	5.65	1.16
PRO	8.31	0.23	1.94	0.40	2.37	0.59	7.00	0.85	1.50	0.31	2.21	0.10	5.56	1.49
SER	6.15	0.02	2.37	0.53	2.30	0.21	5.27	0.56	1.86	0.45	2.60	0.11	4.46	0.89
THR	6.43	0.04	2.26	0.52	1.98	0.19	5.75	0.63	1.70	0.47	2.22	0.07	4.59	0.93
TYR	5.86	0.11	1.52	0.35	1.66	0.40	5.20	0.56	1.45	0.31	1.97	0.07	3.87	0.74
VAL	7.73	0.08	2.59	0.52	2.48	0.15	6.64	0.71	2.08	0.52	2.84	0.12	5.42	1.22

C

Sample type	BF-15		BF-16		BF-17		BF-18		BF-19		BF-20		BF-21		BF-22	
AA	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
ALA	2.17	0.12	2.38	0.55	6.09	1.87	1.68	0.17	2.28	0.19	6.80	0.51	2.42	0.09	2.53	0.19
ARG	2.63	0.11	2.30	0.54	7.17	2.35	1.89	0.20	2.63	0.36	7.46	1.12	2.97	0.22	2.60	0.21
ASX	4.30	0.39	3.97	0.95	10.91	3.30	3.25	0.33	3.90	0.34	12.29	0.87	4.78	0.28	4.34	0.45
GLX	7.88	1.04	5.29	1.19	13.92	3.67	4.89	0.47	5.14	0.37	15.73	0.90	8.70	0.46	5.77	0.83
GLY	1.58	0.09	1.74	0.44	5.31	1.61	1.27	0.13	1.70	0.15	5.95	0.45	1.79	0.08	1.84	0.16
HIS	0.88	0.04	0.93	0.24	2.63	0.82	0.70	0.07	0.96	0.10	2.94	0.18	0.97	0.04	1.04	0.09
ILE	1.78	0.08	2.04	0.52	5.31	1.56	1.47	0.15	2.07	0.21	5.84	0.50	2.00	0.07	2.20	0.18
LEU	2.75	0.11	3.22	0.82	9.74	2.95	2.31	0.25	3.30	0.31	10.60	0.94	3.11	0.09	3.44	0.29
LYS	2.32	0.12	2.35	0.61	5.24	1.86	1.67	0.17	2.20	0.25	5.62	0.63	2.63	0.08	2.66	0.24
MET	0.44	0.03	0.49	0.10	1.91	0.35	0.42	0.06	0.29	0.27	1.97	0.22	0.50	0.02	0.46	0.23
PHE	1.71	0.06	2.15	0.57	6.36	1.96	1.44	0.16	2.10	0.22	6.90	0.64	1.90	0.07	2.26	0.19
PRO	1.60	0.26	2.11	0.80	5.95	2.15	1.22	0.13	2.24	0.17	6.89	0.49	1.65	0.08	1.93	0.17
SER	1.89	0.11	2.17	0.53	5.00	1.45	1.42	0.15	2.15	0.20	5.52	0.42	2.09	0.11	2.29	0.22
THR	1.75	0.11	1.82	0.44	5.09	1.50	1.32	0.12	1.80	0.16	5.64	0.43	1.88	0.07	1.94	0.15
TYR	1.38	0.19	1.56	0.52	4.50	1.46	1.16	0.13	1.70	0.18	4.76	0.65	1.67	0.11	1.68	0.15
VAL	2.09	0.14	2.31	0.59	5.98	1.77	1.66	0.18	2.28	0.22	6.72	0.54	2.36	0.08	2.47	0.20

4.4.3 Soluble total carbohydrate composition from leaves and roots biomass of stably transformed cassava plants

Figure 14 below shows the standard curve generated by using the D-glucose standard provided by the soluble total carbohydrate kit (sigma,USA). The graph was obtained by plotting the concentration of the D-glucose standard against the OD read at 490 nm. The Absorbance at 490 nm of each sample (Table 8) was read using a plate reader (Thermo scientific, USA), then expressed in carbohydrate (µg/µl) using the following equation: X= (OD at 490 nm- Y= 0.1666x+0.16. This equation was obtained through generating a standard curve using D-glucose standard provided by the kit used for analysis and following manufacture's protocol (GBiosciences, USA). The graph below was obtained by plotting the D-glucose standard concentrations against the OD read at 490 nm. Data from Figure 15 showed that there are large variabilities in leaf and root biomass for all the lines analyzed. For leaf biomass, line R'"(G)L1, and R'"(G) L2 produced the highest carbohydrate content reaching approximately 13µg/50mg compared to the controls non-QQS expressers WT (9µg /50mg) and NPTII (11 µg /50 mg). For root sample, line R'"(G) L1 (6 µg /50mg) and R'"(G) L2 (5µg/50mg) produced the lowest carbohydrate content reaching approximately 13µg /50mg, and R7 (F) the highest (18µg /50mg) compared to the controls non-QQS expressers WT (11µg /50mg) and NPTII (10 µg /50mg). Root biomass of line R(7)(F) produce the highest carbohydrate content among all biomasses and lines tested.

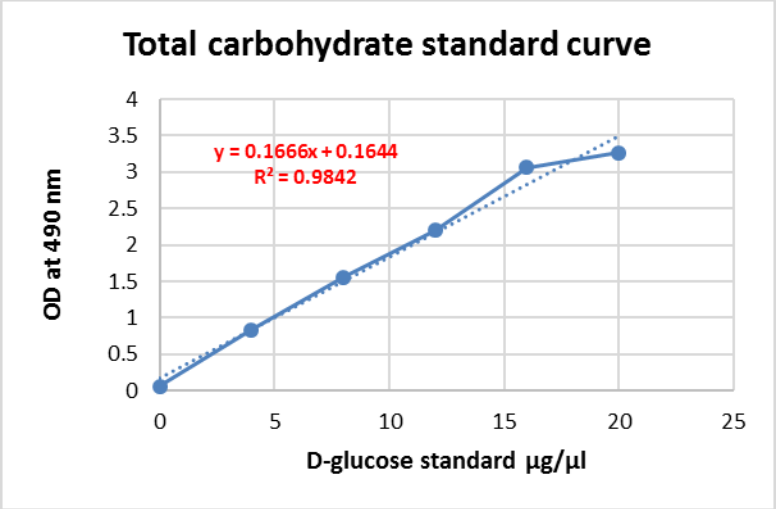


Figure 14*: Standard curve generate using D-glucose standard.*

Y= 0.1666x+0.16; this equation will be used to express the OD values read at 490 nm into soluble total carbohydrate content in 50 mg.

Table 8*: Soluble total carbohydrate from leaf and root tissues of wild-type, NPTII and QQS expressing lines*

Evaluation was done in duplicate, the mean of the samples and the standard errors obtained after soluble total carbohydrate analysis are summarized in this table.

Line	Leaf carbohydrate	SE	Root carbohydrate	SE
WT	9.385	0.738	12.920	2.977
NPTII	11.798	0.378	9.906	1.477
R'"(G)L1	13.707	2.863	7.025	0.276
R'"(G)L2	14.247	2.611	5.753	0.732
R'"(G)L3	11.852	1.369	12.619	3.637
R'"(LA)L2	12.429	3.277	14.816	0.936
R7(F)	12.609	1.188	18.958	10.480

Figure 15: ***Soluble total carbohydrate in leaf and roots root tissues of wild-type, NPTII and, QQS expressing lines.***

A B

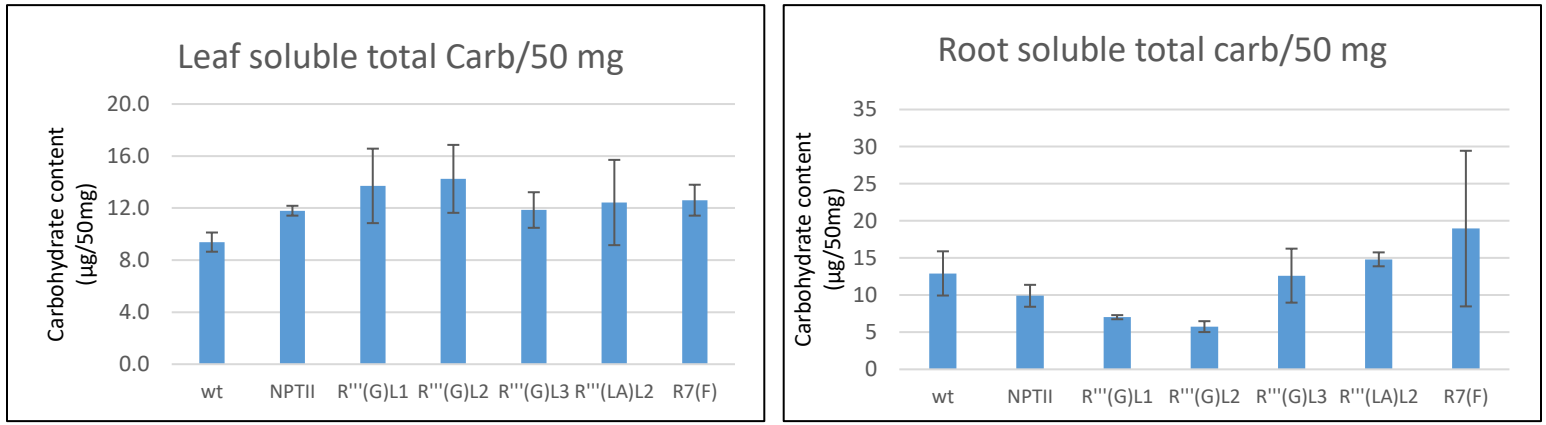


Figure 15-A-B: ***Quantitative evaluation of carbohydrate in leaves (A) and root tissues of wild-type, NPTII and QQS expressing lines (B) of six weeks in-vitro soil grown cassava plants.***

The graphs were plotted using the mean protein content of samples from each line and the standard error (SE) obtained after carbohydrate analysis.

4.4.4 STATISTICAL ANALYSIS OF CORRELATIONS BETWEEN VARIABLES

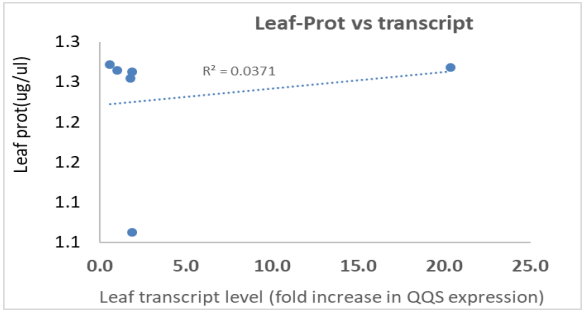
Results of mean comparison using ANAVO SAS enterprise guide 5.1 are shown in Table 9 below. Line were used as independent variable and protein content or carbohydrate content in leaf, root and stem as dependent variables with a significant level of 5%.

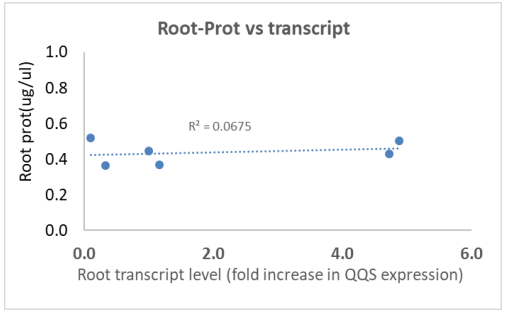
Dependant variable	P value obtained
Leaf protein	0.27
Stem protein	**0.046**
Root protein	**0.0001**
Leaf carbohydrate	0.71
Root carbohydrate	0.45

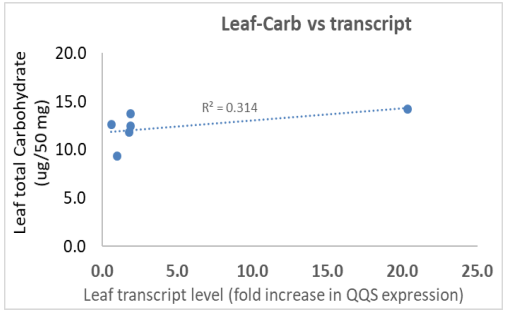
Table 9*: Results of mean comparison using ANOVA* SAS enterprise guide 5.1

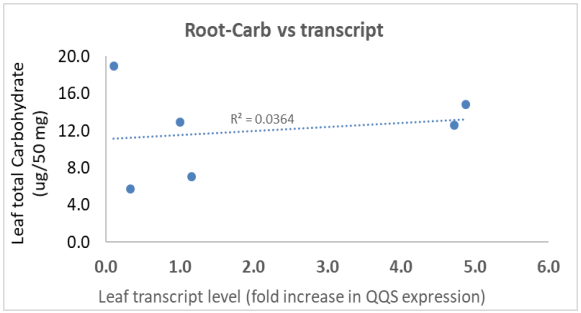
Correlation analysis in Figure 16 show relationship results between QQS transcript levels and carbohydrate or protein content as well as relation between carbohydrate and protein content between and within leaf and root biomass for all transgenic lines tested. Based on the various equations of best fit and the value of the coefficient of determination R^2 of all the correlation data, it is obvious that no strong correlation is observed between protein and carbohydrate content of the leaf and root biomasses and the level of QQS transcripts produced from these plant parts in all the lines. No relationship was also depicted between the leaf carbohydrate and leaf protein for all the lines. The only positive correlation observed was between root carbohydrate and root protein. Very poor correlation was observed between the content of leaf carbohydrate and root carbohydrate as well as content of leaf protein and root protein.

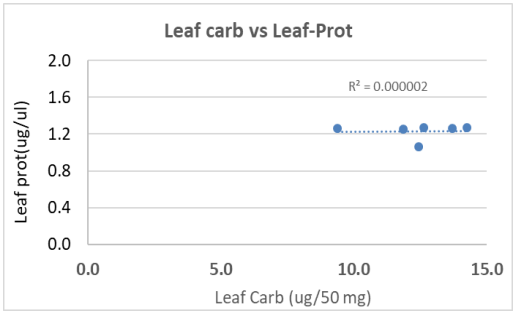
<u>Figure 16</u>: ***Results of correlations between expression level and carbohydrate or protein content and correlations between carbohydrate and protein content among line. Correlations were generated using excel Microsoft version 2016.***

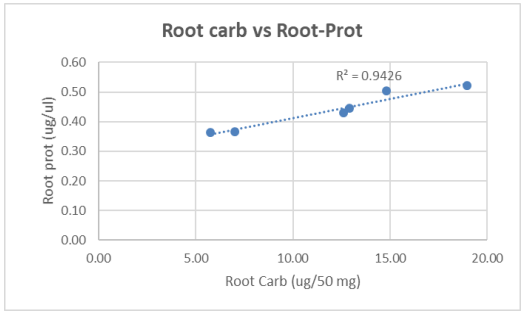


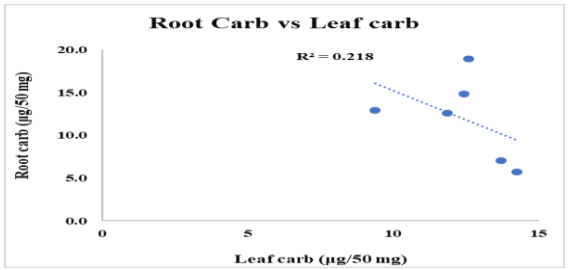


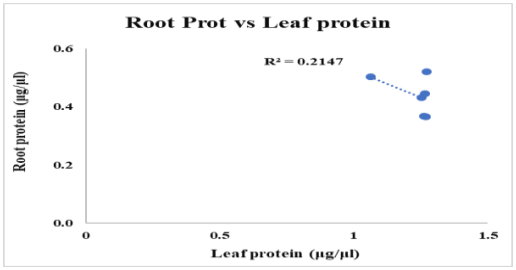












GENERAL DISCUSSION

A successful cloning of the *QQS* optimized sequence in the plant transformation vector pPZP-RCS2-ocs-nptII binary vector (Goderis *et al.*, 2002), a set of modular plant transformation vectors allowing flexible insertion of up to six expression units was achieved through the combination of ligation-free (Gibson et al. 2009) and ligation dependent cloning system. We utilize Gibson assembly because it is scarless (i.e., leaves no unwanted sequences between the assembled fragments). This ligation free-cloning system is widely utilized in various cloning project because it simplifies construction of complex DNA molecules by breaking down a long stretch of DNA sequence into shorter overlapping "brick" fragments, which are amplified by PCR (or directly synthesized) and combined via overlapping sticky ends (Gibson *et al.*, 2009). The pSAT modular entry (Tzfira *et al*., 2005) is also an excellent cloning system because of its flexibility of assembling expression cassettes in various modular cloning units with homing restriction enzymes such as (*Asc*I) (pSAT1) (this thesis), I-*Ppo*I/*Asc*I (pSAT2), I-*Ppo*I (pSAT3), I-*Sce*I (pSAT4), I-*Ceu*I (pSAT5), PI-*Psp*I (pSAT6) and PI-*Tli*I (pSAT7) to facilitate assembly of many single expression vectors in a the multicloning site of pPZP-RCS2-ocs-nptII. We employed Gblock technique after several unsuccessful trials with traditional ligation technique. Difficulties were experienced in correctly amplifying *QQS* after its insertion in pSAT1 modular vector and the successful extraction and characterization of the t*QQS* expression cassette after ligation. After successful insertion of *QQS* gene into cassava genome through *Agrobacterium*-mediated transformation using established protocol (Hankoua *et al.*, 2006 & Taylor *et al*., 2012), at least 13 individual transgenic lines were

regenerated among them 4 were lost due to microbial contaminations. Microbial contamination is a routine problem encountered in tissue culture experiments in general and particularly in *in vitro* manipulations of cassava for micropropagation and gene insertion purposes (Hankoua *et al.*, 2006). Out of the remaining 8 stably transgenic lines, only 5 lines were selected for molecular and biochemical analysis because of the availability of root, leaf, and stem biomasses. Morphological characterization of the established and soil grown transgenic cassava lines identified 4 lines (R7(D), R7 (B), R''' (LA)L1, R7 (E) out of the 8 successfully regenerated transgenic lines (50%) presented some form of morphological abnormalities, but R7 (B), R7 (E) showed a pronounced abnormality characterized with very small and fasciated leaves, very tiny stem and an overall slow growth rate. Because of these physiological abnormalities, lines R7 (B), R7 (E) were not used for the downstream biochemical analysis because of their inability to produce enough biomass from all plant part. The low growth rate can be related to low photosynthetic capability due to the small leaf surface area of the malformed plants. The cassava target tissues, friable embryogenic callus (FEC) used for inserting QQS optimized sequence into cassava genome were 24 months old and was established and maintained from organized embryogenic tissues (Hankoua *et al*., 2005). The observed growth abnormality of these malformed lines could be due to the age of the FEC used in this study which was about twenty-four-months old. This observation is consistent with Taylor *et al.,* (2012) who suggested that transformation of FEC at an optimum stage (about 9 weeks after induction from organized embryogenic tissue), and timely subculturing of transgenic tissues could reduce the formation of abnormal plants. The observed growth defect could be due to genetic or epigenetic variation (Machczynska *et al*., 2015). The growth variability may also be due to high-level expression of the optimized *QQS* gene, but this will be contrary to the observation that expression of QQS in soybean (Li *et al*., 2015) has no impart in the morphology of the field grown

transgenic soybean. However, 69.2% of the regenerated line presented enhanced morphological characteristics compare to their wild-type sibling plants with large leaves, very stick stem, high biomass production and an early mini-tuber induction particularly for line R7 (F) at 6-7 weeks after its transfer in soil. In opposition to the expression of *QQS* into *Arabidopsis Thaliana* and soybean, the regenerated lines were found indistinguishable in morphology from their wild-type sibling plants throughout development (Li *et al*., 2009; Li *et al.*, 2015). In the case of *QQS* expression into cassava plant, this general tendency observed in line R7 (F) is an enhancement of the physiological characteristics of the plant by overall increase of plant vigor is a novel characteristic which could be related to the expression of the *QQS* in cassava. *QQS* is an orphan gene from *Arabidopsis* with initial function was elucidated as transcription factor modulating metabolites allocation in soybean seed (Li *et al*., 2015). The expression of this transcription factor in cassava, which is a high starch producing crop compared to high protein soybean might unveil addition interaction of *QQS* with other transcription factors such as microRNA that play important role in plant development such as enhanced plant biomass production (Zhang *et al., 2015).* Insertion of *QQS* in cassava genome checked through gene specific PCR of extracted DNA from different lines at early stage of regeneration (3 months after transformation) amplified the QQS modified sequence as well as the NPTII binary vector. Real-time PCR results revealed differentially expressed *QQS* between different lines first, within and between leaves and roots of these lines. The variability in the transcript level of a *QQS* sequence may be due to the combination of various important genetic factors shown intensively to affect transcript level, such as tissue type, promoter sequence, enhancer, and codon usage, and transgene insertion location within the chromosome (Streatfield SJ., 2007). We used a very strong constitutive promoter, the enhanced 35S promoter (Zhou J *et al.*, 2013) to drive *QQS* expression in all tissues of these lines, but many

studies have demonstrated that large variability in the transcript levels of a transgene in various tissues is achievable even with a strong constitutive promoter and this could be due to transgenic position effect (Liu C, 2013) and the physiological environment of the cell and tissue types where transgene is expressed. This position effect could explain the reason of higher *QQS* expression was noticed in leaves, especially for R''' (G) L2 line compared with other transgenic lines and the wild type control. Furthermore, *QQS* expression in R''' (LA) L2 was low in leaves while very high in roots. In addition, *QQS* is highly expressed in R''' (G) L2 leaves while very low in the roots. Moreover, *QQS* was expressed at low level in R7(F) both in leaves and roots, and transgene position effect could account for these observed variabilities in QQS expression.

Biochemical analysis revealed either an increase or a decrease in protein content depending on lines and a general increase in leaf carbohydrate while a decrease in root carbohydrate was recorded only for R'''(G) L1 and R''' (G) L2 putative transgenic lines. In general, 4 tendencies have been identified in the biochemical results. For some cases, where the *QQS* was expressed at low level in leaves (R7 (F)), protein content in leaves decreased compare to the wild-type non-transgenic while soluble total carbohydrate content is very high. This tendency of plants expressing *QQS* to have a decrease in protein content while increasing in carbohydrate production, when expressed at very low level, was also found in some transgenic *Arabidopsis thaliana* and soybean lines expressing *QQS* (Li *et al*., 2009).

Whereas, a high expression of *QQS* in the leaves of line (R(G) L2) reveals no variation in the protein content compare to the wild-type while carbohydrate accumulation in this line (R(G) L2 increased considerably. Li *et al.,* (2009) demonstrated that a high expression level of *QQS* in transgenic soybean lead to an increased protein accumulation while carbohydrate accumulation

was decreased. This observation is no aligning with the observed effect of high expression of *QQS* in cassava leaves of line (R(G) L2) possibly pointing to another molecular mechanism of this transcription factor in plant development in a related soybean plant, cassava. Moreover, in cassava roots, a very low expression of *QQS* in line R7 (F) increased both protein and soluble total carbohydrate; while, a very high expression (R (LA) L2) revealed a high accumulation in both protein and soluble total carbohydrate. However, a high expression of *QQS* in R'''(G) L3 roots revealed any noticeable variation in both protein and soluble total carbohydrate accumulation.

No correlations were observed between the *QQS* transcript levels, and carbohydrate, and protein content for leaf and root biomass of all the transgenic lines. This lack of relationship demonstrates that the level of *QQS* transcripts in a particular plant part of a transgenic line does not influence *QQS* to exert the ability to partition the allocation of metabolites during plant development. This phenomenon was also observed in soybean lines expressing *QQS* (Li *et al.,* 2015). It was also demonstrated that the effect of transcription factor such as *QQS* on the starch and protein accumulation in lines from the same transformation event may vary as the expression level of the transgene in different lines of the same transformation event may be different (Shou *et al.,* 2004). *QQS* is a regulatory protein, and it is not unusual for regulatory proteins to have very low levels of expression (Nagaraj *et al.*, 2011). Thus, it may be that only a very small concentration of *QQS* saturates the *QQS* receptor, and any increases over this concentration do not affect protein and starch contents.

No correlation was found between the content of leaf carbohydrate and leaf protein, leaf carbohydrate and root protein, leaf carbohydrate and root carbohydrate but root carbohydrate and root protein content were found to be highly correlated (R^2=O.9). It will be of great interest to see if this ability to accumulate at the same time starch and protein in young roots is translated to

mature tuberous root which is known to accumulate high starch and low protein (Chavarriaga *et al*., 2016). The lack of linear relationship between the content of carbohydrate and protein from leaf, root, stem biomass of transgenic lines versus *QQS* transcript was also observed in *QQS* transcript accumulation in *Arabidopsis* and in soybean (Li *et al*., 2015), but in the case of *Arabidopsis* and in soybean, elevation of *QQS* RNA accumulation was related to an increase in protein accumulation and a decrease in starch accumulation (Li *et al*., 2015), therefore validating one of the molecular mechanism of *QQS* (Li *et al*., 2009). For cassava line R"' (G) L1 and line R"" (G) L2 leaf protein did not increase significantly compare to the control plants, but have significantly reduced carbohydrate content, and this could be attributed to the QQS expression and modulation effect in metabolites distribution as demonstrated in *Arabidopsis* and soybean *QQS* expressers. Transgenic cassava line R7(F) produced the lowest level of *QQS* transcript in both root and leave tissues but produced the highest level of protein increase in the root tissue compared to all the other root and leaf samples for all the lines tested. This phenomenon was also observed in some soybean *QQS* expressers. This phenomenon is not unusual for regulatory proteins to correlate very low levels expression (Nagaraj *et al.*, 2011). Therefore, only a very small concentration of *QQS* saturates the *QQS* receptor (Li *et al.*, 2009), and any increases over this concentration do not affect metabolite content. They might be other possible explanations for a lack of strong linear correlation between levels of *QQS* transcript and metabolite composition such as *QQS* translational efficiency or stability or the effectiveness of the *QQS* protein to biochemically express its function or post-translational modification is limiting. Indeed, a variety of post-translational regulatory mechanisms can come into play, as have been described for other transgenes (Lillo *et al*., 2004).

CHAPTER 5: CONCLUSIONS & RECOMMENDATION

The expression of QQS gene into cassava plant increased leaf protein until 1.36% in line R'''(LA) L2 and root protein until 17.02% for the same line. Moreover, leaf soluble total carbohydrate increased until 51.76% in line R'''(G) L2 and root soluble total carbohydrate increased until 46.75% in line R(F). For line R7 (F), a low expression in leaves caused a decrease in leaf protein content by 9% while root protein, leaf and root carbohydrate increased respectively by 13.03%, 34.29% and 46.75%. A strong correlation (0.9) was found between root protein and total root carbohydrate accumulation. There was no correlation found between level of expression and chemical compound accumulation. But the general tendency seems to be that a low expression in leaves, which is the case for R7 (F) and R'''(LA) L2 offer better results in term of protein and carbohydrate accumulation in leaves and roots. Morphological observation has revealed a positive impact of *QQS* expression in cassava giving a stronger plant with larger leaves, a stick stem and an increased biomass production. These data will represent an interesting new discovery if the biochemical profile results are confirmed and recapitulated in tuberous root, stem and leaves of matured transgenic cassava grown in the field. Therefore, the novel cassava developed in this project will contribute in ensuring a well-balanced protein diet to the 800 million people in Asia, Africa and Latin America who rely on the starch crop cassava as their staple food. The capability of *QQS* to increase the soluble total carbohydrate content consequently starch accumulation if validated in these transgenic mature plants can be explored by starch and bioethanol producers. *QQS* has been proven to decrease also plant oil accumulation. Cassava is not an oil crop, and most

of the food preparations involve an addition of oil rich food; in some occasions through addition of grinded peanut. For scientific curiosity, it will be interesting to compare crude oil content between wild-type leaves and roots and the transgenic ones. To validate the molecular mechanism of *QQS* in transgenic cassava plant, the two-yeast hybrid system could be used in further experimentation to see if *QQS* will bind to the conserved transcription factor, nuclear factor Y, subunit C4 (NF-YC4) which is an ubiquitous factor modulating plant development. In terms of safety regarding genetic modified food (GMO), at this day, any concrete identification of toxicity was found in GMO. Nevertheless, investigations should be done on transgenic *QQS* plants to ensure safety of crop. It will be interesting to study how *QQS* impact carbon reparation in plant and therefore causing variations in plant chemical composition and vigor. Western blot should be done for protein detection and Southern blot to check the integration of *QQS* into the cassava genome and to evaluate transcript copy number. In addition, the model cassava cultivar (cv 60444) which is not a preferred cultivar for farmers in Africa (Chavarriaga *et al*., 2016) was used in this study; if the results observed are confirmed on tissues from matured plants, the new traits should be transferred to the consumers and farmers preferred cultivar. Finally, other protein genes such as *AMA1* can also be expressed in cassava to compare their effects with the ones discovered with the expression of *QQS* gene.

REFERENCES

Abhary, M., Siritunga, D., Stevens, G., Taylor, N.J. and Fauquet, C.M. (2011). Transgenic biofortification of the starchy staple cassava (*Manihot esculenta*) generates a novel sink for protein. PLoS ONE 6(1): e16256). (Retracted).

Chavarriaga-Aguirre, P., Brand, A., Medina, A., Prías, M., Escobar, R., Martinez, J., Díaz, P., López, C., Roca, W.M. and Tohme, J. (2016). The potential of using biotechnology to improve cassava: a review. In Vitro Cell. Dev. Biol. – Plt. 52(5): 461-478.

Chavez AL, Bedoya JM, Sanchez T, Iglesias C, Ceballos H, Roca W. (2000). Iron, carotene, and ascorbic acid in cassava roots and leaves. Food Nutr. Bull. 21:410–13.

Chellappan, 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. Plt. Mol. Biol. 56(4): 601 Chavez A

Cock, J.H. (1985). Cassava: New potential for a neglected crop. Westfield Press, Boulder; 1985.

Cohen, S. A. and K. M. De Antonis. (1994). Applications of amino acid derivatization with 6-aminoquinolyl-n-hydroxysuccinimidyl carbamate. Analysis of feed grains, intravenous solutions and glycoproteins. Journal of Chromatography A 661(1-2): 25-34.

Dai Z, Hooker BS, Daniel B, Anderson DB and Thomas SR. (2000). Improved plant-based production of E1 endoglucanase using potato: expression optimization and tissue targeting. Molecular Breeding 6 (3). 277-285).

Failla, M.L., Chitchumroonchokchai, C., Siritunga, D., De Moura, F.F., Fregene, M., Manary, M.J. and Sayre, R.T. (2012). Retention during processing and bioaccessibility of β-carotene in high β-carotene transgenic cassava root. J. Agric. Food Chem. 60: 3861-3866.

FAO. (2013). Save and grow cassava: a guide to sustainable production and identification. E-ISBN 978-92-5-107642-2.

FAO/WHO. (2002). Chapter 7: vitamin A in FAO/WHO Expert Consultation on Human Vitamin and Mineral Requirements, pp. 87–107. Rome: World Health Org. /Food Agric. Org. United Nations.

FAOSTAT. (2015). Food and Agriculture Organization of the United Nations. http://faostat3.fao.org/browse/Q/QC/E. Accessed 28 November 2017.

Gegios, A., Amthor, R., Maziva-Dixon, B., Egesi, C., Mallowa, S., Nungo, R., Gichuki, S., Mbanaso, A. and Manary, M. J. (2010). Children consuming cassava as a staple food are at risk for inadequate zinc, iron, and vitamin A intake. Plant Foods for Human Nutrition 65: 64-70.

Gibson, DG., Young, L., Chuang, RY., Venter, C., Hutchison, CA., Smith, HO. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods**.6** (5):343–5).

Goderis, I.J.W.M., Miguel, F.C., Bolle, D., Francois, I.E.J.A., Wouters, P.F.J.W., Broekaert, W.F. and Cammue, B.P.A. (2002) *Plant Mol. Biol.* **50,** 17–27).

Graham RD, Welch RM (1996): Breeding for staple-food crops with high micronutrient density: Agricultural Strategies for Micronutrients. Washington, D.C: Working Paper No 3; International Food Policy Research Institute.

Hamilton, R.J. (1992). In: Lipid Analysis. A practical approach; Hamilton, S. (ed.) Oxford University Press: Oxford, U.K., pp 163-166.

Hankoua, B.B., Taylor, N.J., Ng, S.Y.C., Fawole, I., Puonti-Kaerlas, J., Chellappan, P., Dixon, A.G.O., Fauquet, C.M. and Fondong, V. (2006). Production of the first transgenic cassava (*Manihot esculenta Crantz*) plants in Africa via direct shoot organogenesis of friable embryoids and germination of maturing somatic embryos. Afr. J. Biotechnol. 5(19): 1700-1712.

Hankoua, B.B.; Ng, S.Y.C.; Fawole, I.; Puonti-Kaerlas, J.; Pillay, M.; Dixon, A.G.O. (2005).Regeneration of a wide range of African cassava genotypes via shoot organogenesis from cotyledons of maturing somatic embryos and conformity of the field-established regenerants**.** *Plant cell, tissue, and organ culture 2005 v.82 no.2* pp. 221-231.

Hillocks RJ, Thresh JM, Bellotti AC, eds. (2002). Cassava Biology, Production and Utilization. New York:CABI. 332 pp.

Howeler, Reinhardt H. (2012). The cassava handbook: A reference manual based on the Asian Regional Cassava Training Course, held in Thailand. Centro International de Agriculture Tropical (CIAT), Bangkok, TH. 801 p.

Ihemere, U.E., Narayanan, N.N. and Sayre, R.T. (2012). Iron biofortification and homeostasis in transgenic cassava roots expressing the algal iron assimilatory gene, FEA1. Front. Plant Sci.13 September 2012 https://doi.org/10.3389/fpls.2012.00171.

Jennings, D.L. and Iglesias, C. (2002). Breeding for crop improvement. In Cassava: Biology. Edited by Hillocks Ryan, Thresh JM, Bellotti AC. CAB International: Production and Utilization; pp.149–166.

Ladino, J.J., Echeverry, M., Mancilla, L.I., Lopez, D., Chavarriaga, P., Tohme, J. and Roca, W. (2002). Genetic transformation of cassava: confirmation of transgenesis in clone 60444 and analysis of CRY1Ab protein in transgenic lines. Preliminary data on transformation of farmer-preferred cultivars SM1219-9 and CM3306-4. Annual Report, CIAT, Cali, Colombia.

Legg, J., Lava-Kumar, P., Makeshkumar, T., Triatic, L., Ferguson, M., Kanju, E., Ntawuruhunga, P. and Cuellar, W. (2014). Cassava virus diseases: Biology, epidemiology, and management. In: Gad L, Nikolaos IK (eds) Advances in Virus Research 91, pp 85-142. Li et al., Biotechnologically-modified cassava: Protein absorption relative to casein.

Li, K.T., Moulin, M., Mangel, N., Albersen, M., Verhoeven-Duif, N.M., Ma, Q., Zhang, P., Fitzpatrick, T.B., Gruissem, W. and Vanderschuren, H. (2015). Increased bioavailable vitamin B6 in field-grown transgenic cassava for dietary sufficiency. Nat. Biotechnol. 33: 1029-1032.

Li, L. and Wurtele, E. S. (2015), The *QQS* orphan gene of Arabidopsis modulates carbon and nitrogen allocation in soybean. Plant Biotechnol J, 13: 177–187. doi:10.1111/pbi.12238.

Li, L., Foster, C.M., Gan, Q., Nettleton, D., James, M.G., Myers, A.M. and Wurtele, E.S. (2009) Identification of the novel protein QQS as a component of the starch metabolic network in Arabidopsis leaves. Plant J. 58, 485–498.

Li, L., Ilarslan, H., James, M.G, Myers, A.M. and Wurtele, E.S. (2007) Genome wide co-expression among the starch debranching enzyme genes AtISA1, AtISA2 and AtISA3 in *Arabidopsis thaliana*. J. Exp. Bot. 58: 3323-3342.

Li, Xiangkai and Yang, Jian. (2012). Biotechnologically-Modified Cassava: Protein Absorption Relative to Casein. Journal of Bioequivalence & Bioavailability. 04. 10.4172/jbb.10000109.

Lillo, C., Meyer, C., Lea, U.S., Provan, F. and Oltedal, S. (2004) Mechanism and importance of post-translational regulation of nitrate reductase. J. Exp. Bot. 55, 1275–1282).

Liu C. (2013). Strategies for Designing Transgenic DNA Constructs. *Methods in molecular biology (Clifton, NJ).* 2013; 1027:10.1007/978-1-60327-369-5_8. doi:10.1007/978-1-) 60327-369-5_8.)

Liu XY, Rocha-Sosa M, Hummel S, Willmitzer L, Frommer WB (1991) A detailed study of the regulation and evolution of the two classes of patatin genes in Solanum tuberosum L. Plant Mol Biol 17: 1139–1154) *Frontiers in Plant Science*. 2013; 4:526. doi:10.3389/fpls.2013.00526.

Low, J.W., Arimond, M., Osman, N., Cunguara, B., Zano, F. and Tschirley, D. (2007) A food-based approach introducing orange-fleshed sweet potatoes increased vitamin A intake and serum retinol concentrations in young children in rural Mozambique. J. Nutr. 137(5):1320-1327. M., Kelso, J., Paabo, S. and Mann, M. (2011) Deep proteome and transcriptome mapping of a human cancer cell line. Mol. Syst. Biol. 7, 548)

Machczynska Mutisya, DL., El-Banhawy, E.M., Khamala, C.P.M., and Kariuki, C.W. (2015). Management of cassava green mite *Mononychellus progresivus* (Acari: Tetranychidae) in different agro-ecological zones of Kenya. *System Appl. Acarol.* 20**,** 39–50).

Nagaraj, N., Wisniewski, J.R., Geiger, T., Cox, J., Kircher, M., Kelso, J., Paabo, S. and Mann, M. (2011). Deep proteome and transcriptome mapping of a human cancer cell line. Mol. Syst. Biol. 7, 548

Narayanan, N., Beyene, G., Chauhan, R.D., Gaitán-Solis, E., Grusak, M.A., Taylor, N. and Anderson, P. (2015). Overexpression of Arabidopsis VIT1 increases accumulation of iron in cassava roots and stems. Plt. Sci. 240: 170-181.

Nassar, N., Vizzotto, C.S., Schwartz, C.A. and Pires, O.R. Jr. (2007). Cassava diversity in Brazil: the case of carotenoid-rich landraces. Genet. Mol. Res. 6:116-121.

Nweke, F., Spencer, D. and Lynam, J. (2002). The cassava transformation: Africa's best kept secret. Lansing, Mich., USA: Michigan State University Press. http://www.jstor.org/stable/10.14321/j.ctt7ztc0t (Retrieved: 05/15/2017).

Omole, J.A. (1977). Cassava in the Nutrition of layer, In: Cassava as animal feed. Nestel B, Grahams M (Eds). Process Workshop (IITA), Ibadan.

Pérez-Massot, E., Banakar, R., Gómez-Galera, S., Zorrilla-López, U., Sanahuja, G., Arjó, G., Miralpeix, B., Vamvaka, E., Farre´, G., Rivera, S.M., Dashevskaya, S., Berman, J., Sabalza, M., Yuan, D., Bai, C., Bassie, L., Twyman, R.M., Capell, T., Christou, P. and Zhu, C. (2013). The contribution of transgenic plants to better health through improved nutrition: opportunities and constraints. Genes Nutr. 8: 29-41

Prochnik, S., Marri, P.R., Desany, B., Rabinowicz, P.D., Kodira, C., Mohiuddin, M., Rodriguez, F., Fauquet, C., Tohme, J., Harkins, T., Rokhsar, D.S. and Rounsley, S. (2012). The Cassava genome: current progress, future directions. Trop. Plt. Biol. 5(1): 88-94

Reilly, K., Gómez-Vásquez, R., Buschmann, H., Tohme, J. and Beeching, J.R. (2003). Oxidative stress responses during cassava post-harvest physiological deterioration. Plt. Mol. Biol. 53: 669-685.

Rosling, H. (1988). Cassava toxicity and food security: A review of health effects of cyanide exposure from cassava and of ways to prevent these effects (UNICEF African household food security program, Uppsala, Sweden).

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(4): 339-344.

Sayre R, Beeching JR, Cahoon EB, Egesi C, Fauquet C, et al. (2011) The BioCassava plus program: biofortification of cassava for sub-Saharan Africa. Annu Rev Plant Biol 62: 251-272.

Shou, H., Frame, B.R., Whitham, S.A. and Wang, K. (2004) Assessment of transgenic maize events produced by particle bombardment or *Agrobacterium*-mediated transformation. Mol. Breed. 13, 201–208.

Siritunga, D. and Sayre, R.T. (2007). Transgenic approaches for cyanogen reduction in cassava. J. Assoc. Off. Anal. Chem. 90: 1450-1455.

Siritunga, D., Arias-Garzon, D., White, W. and Sayre, R.T. (2004). Over-expression of hydroxynitrile lyase in transgenic cassava roots accelerates cyanogenesis and food detoxification. Plt. Biotechnol. J. 2: 37-43.

Sreeja, V.G. and Leelamma, S. (2002). Cassava diet-a cause for muco polysaccharidosis. Plant Hum. Nutr. 57: 141-150.

Stephenson, K., Amthor, R., Mallowa, S., Nungo, R., Maziya-Dixon B.et al. (2010). Consuming cassava as staple food places children 2-5 years old at risk for inadequate protein intake, an observational study in Kenya and Nigeria. Nutr. J. 9: 9.

Steven R. and Winter-Nelson, A. (2008). Consumer acceptance of provitamin A biofortified maize in Maputo. Mozambique. Food Policy 33: 341-351.

Streatfield SJ (2007). Approaches to achieve high-level heterologous protein production in plants**.** Plant Biotechnol J. 2007 Jan;5(1):2-5.

Taylor, N., Gaitán-Solís, E., Moll, T., Trauterman, B., Jones, T., Pranjal, A., Trembley, C., Abernathy, V., Corbin, D. and Fauquet, C.M. (2012). A high-throughput platform for the production and analysis of transgenic cassava (*Manihot esculenta*) plants. Tropical Plant Biol. 5:127-139.

Taylor, N.J., Chavarriaga, P., Raemakers, K., Siritunga, D. and Zhang, P. (2004). Development and application of transgenic technologies in cassava. Plt. Mol. Biol. 56: 671-678.

Tzfira, T., Kozlovsky, S.V. and Citovsky, V. (2007). Advanced expression vector systems: new weapons for plant research and biotechnology. Plant Physiol. 145: 1087-1089.

Tzfira, T., Tian, G.W., L.A., Croix, B., Vyas, S., Li, J., Leitner-Dagan, Y., Krichevsky, A., Taylor, T., Vainstein, A. and Citovsky, V. (2005). pSAT vectors: a modular series of plasmids for auto fluorescent protein tagging and expression of multiple genes in plants. Plt. Mol. Biol. 57: 503-516.

Udvardi, M.K., Czechowski, T. and Scheible, W.R. (2008). Eleven golden rules of quantitative RT-PCR. The Plant Cell. 7: 1736-1737.

Welsch, R., Arango, J., Bär, C., Salazar, B., Al-Babili, S., Beltrán, J., Chavarriaga, P., Ceballos, H., Tohme, J. and Beyer, P. (2010). Provitamin A accumulation in cassava (*Manihot*

esculenta) roots driven by a single nucleotide polymorphism in a phytoene synthase gene. Plt. Cell. 22: 3348-3356.

yaboga E, Njiru J, Nguu E, Gruissem W, Vanderschuren H, Tripathi L (2013). Unlocking the potential of tropical root crop biotechnology in east Africa by establishing a genetic transformation platform for local farmer-preferred cassava cultivars. *Frontiers in Plant Science*. 2013; 4:526. doi:10.3389/fpls.2013.00526.)

Yan, J., Kandianis, C.B., Harjes, C.E., Bai, L., Kim, E.H., Yang, X., et al (2010): Rare genetic variation at Zea mays crtRB1 increases β-carotene in maize grain. Nat Genet. 42: 322-327.

Zeevi, V., Liang, Z., Arieli, U. and Tzfira, T. (2012). Zinc finger nuclease and homing endonuclease-mediated assembly of multigene plant transformation vectors. Plt. Physiol. 158: 132-144.

Zhang, B. and Wang, Q. (2015) MicroRNA-based biotechnology for plant improvement. J. Cell. Physiol. 230, 1, 1–15. In Pres).

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

Zhang, X., Myers, A.M. and James, M.G. (2005). Mutations affecting starch synthase III in Arabidopsis alter leaf starch structure and increase the rate of starch synthesis. Plant Physiol. 138: 663-674.

Zhou J, Yang Y, Wang X, Enhanced transgene expression in rice following selection controlled by weak promoters. *BMC Biotechnology*. 2013; 13:29. doi:10.1186/1472-6750-13-29.