Acta Universitatis Agriculturae Sueciae

Doctoral Thesis No. 2022:21
Faculty of Landscape Architecture, Horticulture and Crop Production Science

Transcriptome profiling and gene editing for biofortification of cassava

Priscilla Olukola Olayide

Doctoral Thesis No. 2022:21 • Transcriptome profiling and gene editing for biofortification of cassava • Priscilla Olukola Olayide

Transcriptome profiling and gene editing for biofortification of cassava

Priscilla Olukola Olayide

Faculty of Landscape Architecture, Horticulture and Crop Production Science

Department of Plant Protection Biology

Alnarp

SLU

SWEDISH UNIVERSITY OF AGRICULTURAL SCIENCES

DOCTORAL THESIS

Alnarp 2022

Acta Universitatis Agriculturae Sueciae
2022:21

Cover: Cassava leaves and flowers (Illustration by Franz Eugen Köhler)
ISSN 1652-6880
ISBN (print version) 978-91-7760-917-9
ISBN (electronic version) 978-91-7760-918-6

Swedish University of Agricultural Sciences, Department of Plant Protection Biology, Alnarp, Sweden
Print: SLU Service/Repro, Alnarp 2022

Transcriptome profiling and gene editing for biofortification of cassava

Abstract

Cassava (*Manihot esculenta* Crantz) is an important food crop for millions of people in sub-Saharan Africa. Cassava serves as a food security crop and a good source of energy, but it has relatively low nutritional quality, which has nutritional implications for those who rely on the crop as their main source of calorie intake. Vitamin A deficiency (VAD) is a health problem in populations whose diet constitutes mainly starchy crops like cassava. A possible solution is biofortification by conventional breeding or biotechnology to increase β-carotene content. Unfortunately, some studies report a negative correlation between β-carotene and dry matter content in certain genotypes, which could pose a challenge to these types of cassava biofortification measures. Field-grown cassava landraces were analyzed for agronomic traits and carried out specific and global transcript analyses by real-time quantitative RT-PCR and RNA-sequencing (RNA-seq). This was combined with targeted starch and carotenoids analysis by HPLC and non-targeted metabolite analysis by GC/LC-MS to understand the regulation of key enzymes and intermediate metabolites to identify genes influencing β-carotene accumulation in cassava. Also, using the CRISPR/Cas9-mediated gene editing system in the cassava cultivar TMS60444 we tried to introduce knockout mutations into cassava β-carotene hydroxylase (*MeChyβ*), lycopene-ε-cyclase (*MeLcyε*), and 9-cisepoxycarotenoid dioxygenase 1 (*MeNced1*) which are key genes of the carotenoid pathway. In a separate study, the biosafety regulations and policies in Kenya, Nigeria, Uganda, and the EU represented by Sweden were examined by comparing legislative texts and conducting interviews to determine if policy and regulatory frameworks present problems to perform R&D using new breeding technologies. In the cassava landraces analyzed, we found a weak negative correlation between starch and β-carotene content, whereas there was a strong positive correlation between root yield and carotenoids. Also, cassava landraces with reasonably high content of starch and β-carotene were identified that could be candidates for biofortification by further breeding or plant biotechnological means. Global gene expression profiles grouped cassava landraces into white and yellow landraces; however, at the same time there was no general correlation between the expression profiles of individual genes involved in carotenoid synthesis and accumulation with the storage root color. Gene Ontology (GO) enrichment showed over-representation of upregulated genes involved in protein-related metabolic and catabolic processes in yellow landraces while GO related to photosynthesis and light reactions were enriched in white landraces. Interestingly, we identified a previously reported amino acid change from Alanine to Aspartic acid in Me*Psy1* at position 191 to distinguish the yellow lines from the white lines; however, this change was absent in the pale-yellow lines, confirming that the mutation in *psy* is not solely responsible for carotenoid accumulation in cassava. Non-targeted metabolite analysis revealed higher abundance of several amino acids in white lines, but also higher levels of a few osmolytes indicating differences in stress response. Transformation of cassava FECs with gene targets *Lcyε*, and *Nced1* did not produce transgenic regenerated shoots, whereas *MeChyβ* produced *in vitro* plantlets still under investigation. Finally, our study showed that biosafety regulations on GMO approval in Kenya, Nigeria, and Uganda are not a major hurdle for R&D but might rather be influenced by factors outside of the regulatory framework such as perceptual and financial factors including funding opportunities.

Keywords: Carotenoid biosynthesis, provitamin A, biofortification, CRISPR/Cas9, RNA-seq, transcriptomics, metabolomics, regulatory framework

Author's address: Priscilla Olukola Olayide, Swedish University of Agricultural Sciences, Department of Plant Protection Biology, Alnarp Sweden

Dedication

In loving memory of my beloved grandmother

“To everything there is a season, a time for every purpose under heaven.”

Ecclesiastes 3:1

Contents

List of publications

This thesis is based on the work contained in the listed papers:

I. Priscilla Olayide, Annabel Large, Linnea Stridh, Ismail Rabbi, Susanne Baldermann, Livia Stavolone, and Erik Alexandersson* (2020). Gene Expression and Metabolite Profiling of Thirteen Nigerian Cassava Landraces to Elucidate Starch and Carotenoid Composition. *Agronomy* Vol. 10 (3). pp. 424-441

II. Priscilla Olayide, Erik Alexandersson, Oren Tzfadia, Marit Lenman, Andreas Gisel, and Livia Stavolone. Comparative transcriptome and metabolome analyses identify factors potentially involved in pro-vitamin A accumulation in cassava landraces (manuscript)

III. Priscilla Olayide, Livia Stavolone, Leena Tripathi, Marit Lenman, and Erik Alexandersson. CRISPR-Cas9 directed editing of β-carotenoid hydroxylase in cassava for increased carotenoid composition (manuscript)

IV. Isaac Ongu, Priscilla Olayide, Erik Alexandersson, Barbara Mugwanya Zawedde, Dennis Eriksson*. The biosafety regulatory frameworks in Kenya, Nigeria, Uganda and Sweden and their potential impact on international R&D collaborations (manuscript)

*Corresponding authors

The contribution of Priscilla Olukola Olayide to the papers included in this thesis was as follows:

I. Designed the study together with co-authors, planned and was involved in field work from planting to harvest. Performed some of the laboratory experiments, and prepared tissue samples for laboratory experiments contracted to other laboratories, analyzed the data, and wrote the manuscript with input from co-authors.

II. Contributed to parts of the study design, planned, and performed sample preparations, analyzed parts of the data, and contributed to writing the manuscript with input from co-authors.

III. Designed study together with coauthors. Planned and performed the experimental work together with co-authors. Analyzed parts of the results and wrote the manuscript with input from co-authors.

IV. Contributed to the writing and some background information, editing of the final version.

List of figures

Abbreviations

CRISPR	Clustered regularly interspaced palindromic repeats
Cas9	CRISPR associated proteins
DMC	Dry matter content
GC-MS	Gas chromatography mass spectrometry
GMOs	Genetically modified organism
LC-MS	Liquid chromatography mass spectrometry
SSA	Sub-Saharan Africa
VAD	Vitamin A deficiency
qRT-PCR	Quantitative reverse transcriptase polymerase chain reaction
UHPLC	Ultra-high performance liquid chromatography

1. Introduction

The United Nations (UN) in the year 2015 presented the 2030 Agenda which includes 17 sustainable development goals with 169 associated targets. One of the goals is to "end hunger, achieve food security, improve nutrition, and promote sustainable agriculture" (SDG 2) by the year 2030. Recommendations were presented to reduce food insecurity, hunger, and the various forms of malnutrition; however, in the 2020 follow-up report, the UN emphasized the need to double efforts to achieve food security by 2030 as current food insecurity trends pose various nutritional consequences that could lead to various manifestations of malnutrition (FAO, 2020).

Nutritious food is important for the maintenance of overall body health and for disease prevention; therefore, access to nutritious food is essential in eradicating all forms of malnutrition. The nutritional profiles of foods are important because vitamins and minerals are essential for growth and development, proper functioning of the body and general well-being; therefore, the body requires daily intake of important vitamins and minerals to achieve a healthy state (FAO, 2020). Vitamin A is essential for vision, proper functioning of the immune system, reproduction, growth, and development (Ceballos et al., 2017). Vitamin A deficiency (VAD) is a major nutritional concern globally, resulting from a diet that is chronically insufficient in vitamin A (WHO, 2009). The incidence of VAD is high in preschool children and pregnant women of developing countries and has become the leading cause of visual impairment and blindness (Busie B. Maziya-Dixon, 2006). VAD over a long duration of time can also lead to anemia, a weakened immune response to infection, which increases the severity of infectious diseases and risk of death (WHO, 2009).

In sub-Saharan Africa (SSA), VAD, which is mainly caused by inadequate

intake of foods rich in vitamin A (Low et al., 2017), affects over 40% of children under the age of five. The incidence of VAD is particularly high among preschool children in low-income countries and socioeconomic areas of sub-Saharan African countries (Hombali et al., 2019), due to lack of access to vitamin A rich foods and reliance on crops that are low in vitamin A such as cassava. The high incidence of VAD in these areas could be related to reliance on staple crops that are deficient in provitamin A such as cassava as the main source of dietary calorie intake. Several efforts have been made to tackle VAD through dietary diversification, food fortification and vitamin supplementation; however, VAD remains prevalent in SSA (Ceballos et al., 2017).

Vitamin A can be obtained from both animal and plant sources. In animals, vitamin A is stored as retinol and can be found in liver, egg yolk and dairy products while in plants it is in the form of carotenoids and is present in mangos, papaya, carrots, and sweet potato (Gilbert, 2013). Carotenoids are light harvesting pigments in plants, they are involved in photosynthesis and have antioxidant properties that stabilizes membrane lipids and maintains physiological processes in plants particularly during abiotic stresses (Sun et al., 2018). Carotenoids are also crucial for the regulation of growth and development and the maintenance of normal health in humans (Nisar et al., 2015). Provitamin A carotenoids which include α-carotene, β-carotene, and β-cryptoxanthin; are known for their characteristic bright colors and are converted into vitamin A after ingestion (D'Andrea & Rodriguez-Concepcion, 2019).

Biofortification is a process of increasing the density of vitamins and minerals in crops through plant breeding and transgenic techniques (Bouis & Saltzman, 2017). Approaches to biofortification are either by conventional breeding which is largely based on recurrent natural selection and crosses or by biotechnology through genetic engineering (Khush et al., 2012). Biofortification of staple crops ensures regular intake of important nutrients and vitamins especially in low-income households where dietary diversification is not affordable. In the long term biofortification is a much cost-effective strategy than other methods of nutrient supplementation in addressing global malnutrition because it focuses on enhancing the micronutrient content of edible part of staple crops ensuring bioavailability (Khush et al., 2012). Also, successful biofortification programs in which the needs of both producers and consumers are met promote acceptance and dissemination of seeds and a crop system that is highly sustainable (Penelope Nestel, 2006). Through biofortification, provitamin A can be supplied to

staple crops through plant breeding and biotechnology; thus, increasing the intake of vitamin A in the diet.

This thesis focuses on gaining a better understanding of the molecular mechanisms governing carotenoid biosynthesis and accumulation in cassava through metabolite and gene expression studies. We also examined the legal framework influencing the development and acceptance of gene editing techniques for plant breeding improvement in SSA.

2. Background

2.1 Cassava (*Manihot esculenta* Crantz)

2.1.1 Origin and distribution

Cassava is a perennial shrub thought to have a neotropical origin and thought to be first domesticated in Brazil before it was introduced to West Africa in the sixteenth century (Sayre et al., 2011). It originated from its wild progenitor, *Manihot esculenta ssp. Flabellifolia* (Allem, 1994), and since its domestication, several cassava cultivars have been produced with characteristics such as high biomass and high starch yield; however, it is highly heterozygous due to its outcrossing nature and broad tropical distribution which makes breeding for desirable traits difficult (Wang et al., 2014). Today, cassava is an important staple crop in tropical and subtropical regions of the world which includes Africa, Asia, Latin America, and the Caribbean (Ceballos et al., 2004). Cassava is mostly propagated in the tropics for its starchy roots and dry matter content which can be up to 40% with starch content of 85% (Cock, 1982); its high starch content makes it an important source of dietary energy in tropical countries. Its perennial nature and tolerance to drought increases its capacity to be a food security crop, particularly in drought-prone regions of SSA (El-Sharkawy, 1993). Every part of the cassava plant can be utilized, the starchy roots are a valuable source of energy and can be boiled or prsocessed in different ways for human consumption as fermented starches, dried chips, and meals or used for animal feed (Ceballos et al., 2004). Cassava stems are used as planting material for propagation while the foliage and tender stems are used for animal feeds. Cassava leaves have high nutritive value, they can be eaten as vegetables and are a relish in a number of African countries (Ufuan Achidi et al., 2005).

2.1.2 Traits for commercialization

Dry matter and starch content is a major driver of increasing cassava

commercialization and expansion. The global export of cassava starch and flour has increased over the years. In Asia, there has been an increase in the export of cassava starch in recent years, with Thailand emerging as the major exporter of cassava starch to other Asian countries (Fuglie et al., 2006). Starch content and properties are also of more interest to starch industries as cassava starch is also being used in the textile and pharmaceutical industry and within food manufacturing (Karlstrom et al., 2016). Other traits for commercialization include resistance/tolerance to major diseases of cassava such as cassava mosaic disease (CMD) and cassava brown streak disease (CBSD), improved yield, and improved plant architecture and early bulking (Ceballos et al., 2004).

2.1.3 The cassava genome

Cassava has a diploid genome with 18 bivalent chromosomes (2n = 36). For a long time, the lack of a cassava reference genome has hindered exploration and understanding the genetic basis and molecular mechanisms of the cassava crop. The first cassava reference genome was obtained by whole genome shotgun (WGS) of nuclear genomic DNA of the partially inbred line AM560-2 using the 454 GS FLX Titanium platform in 2009. This generated an assembly spanning 532.5 Mb, half of v4 sequences into n = 18 chromosomes which was captured in 487 scaffolds (Prochnik et al., 2012). Several versions of the cassava reference genome have been constructed following the assembly of the first draft genome. Version 6 (v6.0) is an Illumina-based assembly of the same line AM560-2 spanning approximately 582.25 Mb (18% more total contig sequence than v5) arranged on 18 chromosomes and with 33,033 predicted protein-coding genes 96.6% of which are anchored to a chromosomal position (Bredeson et al., 2016). Version 7 (v7.0) incorporates PacBio continuous long read sequences, allowing for a more complete capture of assembled contigs (Bredeson et al., 2016), while the current version, the v8 assembly, incorporates newly generated high-throughput chromatin conformation capture with PacBio data (https://phytozome-next.jgi.doe.gov/info/Mesculenta_v8_1).

Due to the heterozygosity as an outbreeding species, cassava has high allelic variations in form of single nucleotide polymorphisms (SNPs) and structural polymorphisms such as deletions, insertions, and inversions (Bredeson et al., 2016). The genomes of two cassava lines, the wild subspecies Manihot esculenta ssp. flabellifolia (W14) and a variety commonly cultivated in Asia (KU50) were recently assembled using next-generation sequencing technologies. The assembly of W14 spanned 432 Mb while KU50 spanned

495 Mb representing 58.2% and 66.7% of the 742-Mb cassava genome and encoding 34,483 and 38,845 predicted genes, respectively (Wang et al., 2014). Genome comparison between W14, KU50, and the sequenced cassava line AM560-2 revealed a considerable number of single-nucleotide variations, insertions and deletions, and high level of heterozygosity (Wang et al., 2014). There was evidence that some gene sets may have been selected for during evolution and adaptation. For instance, stress response genes were highly expressed in domesticated cassava reflecting adaptation to tropical weather and growth conditions (Wang et al., 2014). Genome comparison between TME7, a farmer-preferred Nigerian line, and AM560-2 revealed differences in genome size and large structural variations (SVs) and thousands of haplotypic SVs in TME7 compared with AM560-2 showing that diversity in cassava goes beyond small nucleotide level variation between these two accessions (Mansfeld et al., 2021). Due to its heterozygosity and allelic variations, it is important to sequence several key cultivars to characterize these genetic variations and get a better understanding of the genetic basis of the cassava crop to accelerate breeding of economically important traits such as disease resistance and nutrition.

2.1.4 Nutritional drawbacks of the cassava crop

Although cassava is a good source of carbohydrate, it is deficient in many essential minerals, vitamins, and proteins, which can have potentially damaging health effects in people who rely on the crop as their sole staple food (Ceballos et al., 2017). Malnutrition is a condition resulting from inadequate dietary or nutritional intake and it is the cause of death of children under five years of age in sub-Saharan Africa (Kramer & Allen, 2015). Consumption of staple crops that are predominantly rich in starch such as cassava increases the risk for vitamin A, zinc, and/or iron deficiency (Gegios et al., 2010). About 761,000 deaths have been linked to vitamin A, iron, and zinc deficiency (Sayre et al., 2011). Most cassava varieties grown and consumed around the world have white storage roots (Figure 1A); however, varieties with yellow storage roots (Figure 1B) have been identified in the South American germplasm. Varieties with white storage roots are generally low in provitamin A carotenoids; hence, consuming cassava as a staple food poses the risk for VAD and other nutritional deficiencies. The global incidence map of VAD shows prevalence in Africa and India (Figure 2).

Figure 1. White and yellow cassava storage roots. Wikipedia commons.

2.1.5 The relationship between starch and β-carotene

Although most cassava around the world produce white storage roots, there is diversity in cassava storage root color (Figure 1B). Cassava root pigmentation varies from white, cream, yellow, orange and pink which can be linked to carotenoid components (Chávez et al., 2005). Reports from previous studies have also confirmed a strong linear relationship between yellow color and carotenoid content (Sanchez et al., 2014; Akinwale et al., 2010; Chávez et al., 2005), suggesting that the higher the intensity of the yellow color the higher the carotenoid content. However, several studies using African germplasm have reported a negative correlation between dry matter and carotenoid content (Njoku et al., 2015; Akinwale et al., 2010). Since dry matter is a desired trait particularly appreciated by the starch industry and farmers, a negative correlation between these two traits may pose a threat to biofortification efforts. Using Genome-Wide Association analysis (GWAS), Rabbi et al. (2016) reported that genomic colocation of the loci for dry matter content and root yellowness suggests that physical linkage between the genes is the reason for the observed negative correlation.

There might also be physiological and mechanistic reasons why high β-carotene and starch is difficult to combine. Since starch and β-carotene are synthesized in plastids, there is a possibility that both substrates compete for the same organelles (Cervantes-Flores et al., 2010; Horner et al., 2007). A negative correlation has also been reported in other crops, which supports the notion that the correlation is physiologically linked. In sweet potato, a negative correlation between dry matter and β-carotene and also between starch and β-carotene have been reported (Cervantes-Flores et al., 2010). However, some studies in cassava have suggested that the negative correlation between dry matter and carotenoid content is negligible and that

the two traits are independent of each other particularly in Latin American germplasm (Sanchez et al., 2014; Ceballos et al., 2013).

In paper I, we observed that some cultivars had high β-carotene and high starch content, and a weak negative correlation was found between β-carotene and starch. In paper II, we analyzed the expression profiles of genes involved in glycolysis and starch metabolism in cassava roots for possible correlation with the β-carotene content. Since glyceraldehyde-3-phosphate (GA3P) is a common precursor metabolite needed for starch and carotenoid biosynthesis, it is important to determine if these pathways compete for this metabolite. Determining the cause of the observed negative correlation in African germplasm between these two traits will help breeding and biofortification efforts to improve these traits in cassava.

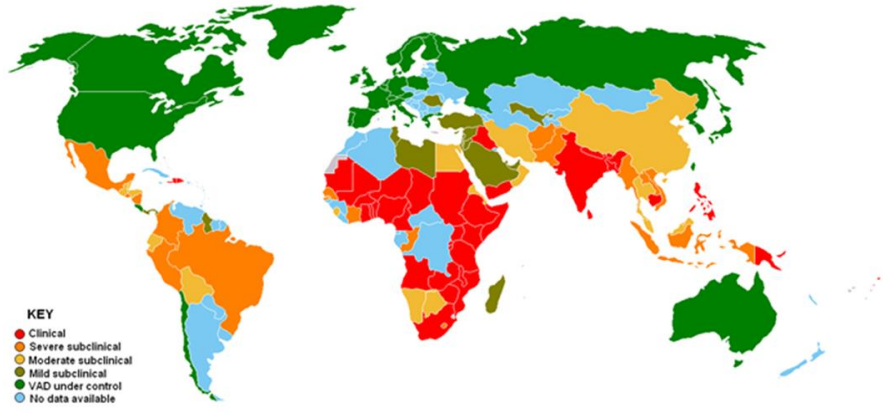

Figure 2. Vitamin A deficiency global incidence map. Wikipedia commons

2.2 Carotenoids

Carotenoids are a class of isoprenoids localized in the plastids of all photosynthetic plants and in some non-photosynthetic bacteria and fungi (Li et al., 2012). Carotenoids are involved in the assembly of photosystems and light harvesting for photosynthesis and photoreception in plants and play a crucial part in plant growth and development (Nisar et al., 2015). In non-photosynthetic tissues such as roots, carotenoids are essential for plant development and the production of strigolactones and abscisic acid (ABA) (Cazzonelli and Pogson, 2010). The biosynthesis of carotenoids occurs

through the methylerythritol 4-phosphate (MEP) pathway (Figure 3). The pathway starts when two molecules of geranylgeranyl pyrophosphate (GGPP) are catalyzed by phytoene synthase (PSY) to produce phytoene which is then converted to lycopene through series of desaturation and isomerization reactions by phytoene desaturase (PDS), 15-cis-ζ-carotene isomerase (ZISO), ζ-carotene desaturase (ZDS) and carotene isomerase (CRTISO) (Luan et al., 2020). Cyclization of lycopene by lycopene-ε-cyclase (Lycε) and/or lycopene-β-cyclase (Lycβ) produces α-carotene and β-carotene respectively. Subsequently, α-carotene and β-carotene undergoes hydroxylation by β-carotene hydroxylases (BCH/CHYβ) to generate lutein in the α-branch and zeaxanthin in the β-branch (Sun et al., 2018). Zeaxanthin is then hydroxylated by zeaxanthin epoxidase (ZEP) to form violaxanthin which can be converted back to zeaxanthin in a reversible reaction by violaxanthin de-epoxidase (VDE) and/or converted to neoxanthin by neoxanthin synthase (NSY). Finally, carotenoids are cleaved by 9-cis-epoxycarotenoid dioxygenases (NCEDs) or carotenoid cleavage dioxygenases (CCDs) to apocarotenoids, abscisic acid (ABA), and strigolactones (Luan et al., 2020).

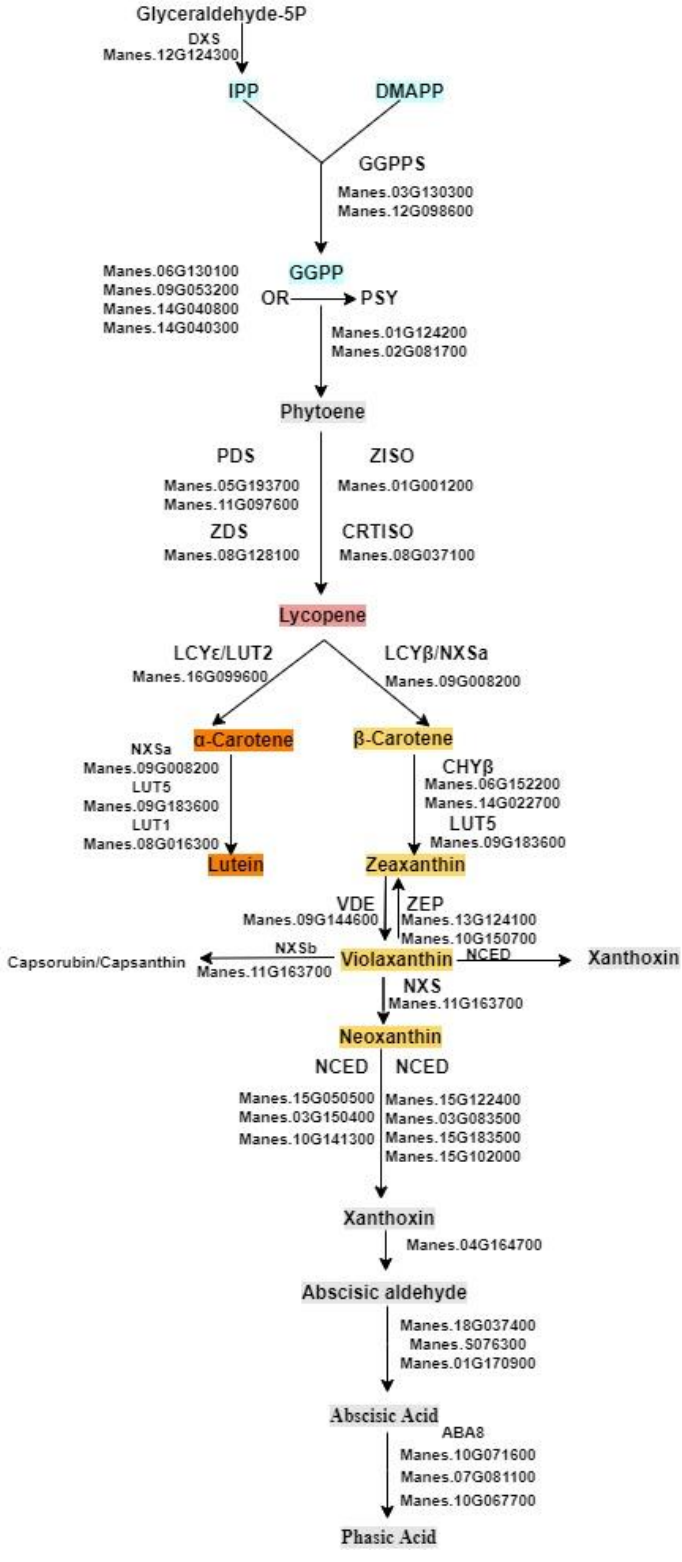
Glyceraldehyde-5P
DXS
Manes.12G124300
IPP
DMAPP
GGPPS
Manes.03G130300
Manes.12G098600
GGPP
Manes.06G130100
Manes.09G053200
Manes.14G040800
Manes.14G040300
OR
PSY
Manes.01G124200
Manes.02G081700
Phytoene
PDS
Manes.05G193700
Manes.11G097600
ZISO
Manes.01G001200
ZDS
Manes.08G128100
CRTISO
Manes.08G037100
Lycopene
LCYε/LUT2
Manes.16G099600
LCYβ/NXSa
Manes.09G008200
α-Carotene
β-Carotene
NXSa
Manes.09G008200
LUT5
Manes.09G183600
LUT1
Manes.08G016300
CHYβ
Manes.06G152200
Manes.14G022700
LUT5
Manes.09G183600
Lutein
Zeaxanthin
VDE
Manes.09G144600
ZEP
Manes.13G124100
Manes.10G150700
NXSb
Manes.11G163700
Capsorubin/Capsanthin
Violaxanthin
NCED
Xanthoxin
NXS
Manes.11G163700
Neoxanthin
NCED
NCED
Manes.15G050500
Manes.03G150400
Manes.10G141300
Manes.15G122400
Manes.03G083500
Manes.15G183500
Manes.15G102000
Xanthoxin
Manes.04G164700
Abscisic aldehyde
Manes.18G037400
Manes.S076300
Manes.01G170900
Abscisic Acid
ABA8
Manes.10G071600
Manes.07G081100
Manes.10G067700
Phasic Acid

Figure 3. Carotenoid biosynthesis pathway showing genes that were identified cassava cultivars used in this study. 1-deoxy-D-xylulose-5-phosphate synthase (DXS), Geranylgenranyl diphosphosphate synthase (GGPPS), orange protein (OR), Phytoene synthase (PSY), phytoene desaturase (PDS), ζ-carotene desaturase (ZDS), ζ-carotene isomerase (ZISO), carotenoid isomerase (CRTISO), lycopene-ε-cyclase (LCYε/LUT), lycopene β-cyclase (LCYβ) β-carotenoid hydroxylases (CHYβ), zeaxanthin epoxidase (ZEP), violaxanthin de-epoxidase (VDE), 9-cis-expoxycarotenoid dioxygenase, neoxanthin synthase (NXS).

PSY is known to catalyze the first carotenoid committed biosynthetic step and has been identified as the limiting enzyme in plant tissues exhibiting low levels of carotenoid production (Shewmaker et al., 2002;Lindgren et al., 2003;Welsch et al., 2010); thus, overexpression of *Psy* can enhance carotenogenesis as was observed in *Arabidopsis thaliana* and cultivated carrot roots overexpressing *AtPsy* (Maass et al., 2009).

The Orange (*Or*) gene, originally isolated from a cauliflower orange curd mutant is also thought to enhance carotenoid accumulation by triggering the differentiation of non-colored plastids into chromoplasts which in turn acts as a metabolic sink facilitating accumulation and sequestration of carotenoids (Li et al., 2012). The *Or* gene is a DNAJ-like Cysteine rich Domain-containing protein which are a family of molecular chaperones involved in protein folding and assembly; they are known to influence PSY stability and differentiation of non-colored plastids into chromoplasts; thus, enhancing carotenoid synthesis (Lu et al., 2006;Sun et al., 2020). Constitutive expression of the *Or* transgene in potato tubers enhanced accumulation of carotenoid–lipoprotein sequestering structures and levels of lutein and β-carotene (Lopez et al., 2008;Li et al., 2012). Five *Or* genes have been identified in cassava, two of these genes are located on chromosome 14 (Manes.14G040800, Manes.14G040300.1), one on chromosome 6 (Manes.06G130100.1), and one on chromosome 9 (Manes.09G053200.1). Manes.14G040800.1, Manes.14G040300.1, and Manes.06G130100.1 were shown to be closely related to and share sequence similarities with CmOR while Manes.09G053200.1 shares sequence similarities with AtOR-like and CmOR-like (Jaramillo et al., 2022).

Carotenoid biosynthesis occur in the plastids, which are of various types, such as proplastids, etioplasts, amyloplasts, chloroplasts, and chromoplast and they all produce carotenoids except proplastids which are undifferentiated progenitor plastids (Howitt and Pogson, 2006). Etioplasts dark-grown precursors of the chloroplast and chloroplasts are the

photosynthetic plastids in green tissues and are important for photosynthesis and photoprotection (Sun et al., 2018). In chloroplasts, carotenoids accumulate in the form of chlorophyll-carotenoid-protein complexes in the thylakoid membranes associated with light-harvesting antenna and these complexes stabilizes plant light-harvesting complexes (LHCs) and assembles photosystem II (PSII) (Hannoufa and Hossain, 2012). Elaioplasts are lipid-storing plastids which use lipoprotein-sequestering structures to store carotenoids in seeds (Howitt and Pogson, 2006). Amyloplasts are starch-storing plastids found in seeds, roots, and tubers, and they contain relatively low levels of carotenoids (Howitt and Pogson, 2006;Wurtzel et al., 2012;Sun et al., 2018). Chromoplasts are colored plastids known to store significant amounts of carotenoids in membranes, oil bodies or other crystalline structures within the stroma (Cunningham et al., 1998), and are responsible for the massive accumulation of carotenoids in many colored flowers, fruits, and vegetables (Sun et al., 2018). Chromoplasts sequester carotenoids within lipoprotein structures which are classified as globular, crystalline, membranous, fibrillar and tubular; this mechanism of carotenoid sequestration relies on the hydrophobic nature of carotenoids (Howitt and Pogson, 2006;Oleszkiewicz et al., 2018). Mechanism of carotenoid biosynthesis and accumulation in chromoplasts have also been linked to increased transcript abundance of regulatory biosynthetic genes and presence of structures capable of sequestering carotenoids in plastids (Howitt and Pogson, 2006). Types and size of chromoplasts could be related to the levels to which carotenoids accumulate which varies among plant species and between individuals of the same species. For instance, sweet potato accumulates β-carotene while potato accumulates primarily β-cryptoxanthin or lutein (Howitt and Pogson, 2006). Variation in levels of β-carotene in cassava storage roots may also be due to variation in the number of chromoplasts in landraces with contrasting carotenoid profiles (Carvalho et al., 2016).

2.3 Use of biotechnology to address nutritional drawbacks in cassava

The application of biotechnology for cassava improvement is important because of the high heterozygosity, low fertility, and unsynchronized flowering of cassava, which all make improvement by conventional breeding a complicated and tedious process (Zainuddin et al., 2012). Although traits can be improved through phenotype-based recurrent selection, the annual

cropping cycle of 12 to 24 months makes it a lengthy process (Rabbi et al., 2017). Applications of biotechnology to enhance cassava production of transgenic material using Agrobacterium-mediated transformation of somatic embryos and microparticle bombardment of friable embryogenic calli (Bull et al., 2011). Establishing cassava biotechnology platforms in SSA will promote development of a sustainable biotechnology infrastructure for cassava improvement.

For decades cassava improvement programs have focused on characteristics such as root yield, root bulking, tolerance to biotic and abiotic stresses, and increased dry matter content (Okechukwu & Dixon, 2008). Improving these traits is crucial for both food security and poverty alleviation especially in SSA and this has been the primary focus of the International Institute of Tropical Agriculture (IITA) breeding program.

In the past few decades, efforts are being directed towards nutritional improvement of cassava as the incidence of malnutrition in SSA continues to rise. One of such efforts is the BioCassava Plus program which aimed to provide complete nutrition in a typical adult-sized cassava meal by improving zinc, iron, protein, and provitamin A content of cassava while incorporating other economic drivers such as reduced cyanogens, virus resistance, and extended shelf life using transgenic strategies (Sayre et al., 2011). An in silico comparative genomic analysis using arabidopsis, tomato, potato and sweet potato as template plants have identified forty carotenoid genes in cassava thought to be conserved in plants (Sreekumar et al., 2022). Identification of genes of the carotenoid biosynthesis pathway and understanding their roles in the pathway is important in developing cassava varieties with improved carotenoid content.

In a study using transgenic strategies, the bacterial crtB gene and 1-deoxy-D-xylulose-5-phosphate synthase (DXS) gene were co-expressed under the control of patatin promoters to direct flux in the isoprenoid pathway from geranylgeranyl diphosphate (GGDP) towards carotenoid synthesis; thus, enhancing biosynthesis and accumulation of provitamin A in cassava storage roots (Beyene et al., 2017). Welsch et al. (2010) observed a positive correlation between SNPs in psy2 gene and root color, and overexpression of bacterial psy gene (TP-CrtB) resulted in transgenic lines with higher carotenoid content compared to the wild-type control. crtB has also been used to enhance carotenoid accumulation canola seeds (Shewmaker et al., 2002), potato tubers (Diretto et al., 2007), and tomato fruits (Fraser et al., 2002). All-trans-β-carotene is the target for provitamin A biofortification and

efforts to enhance β-carotene can be achieved by directing and enhancing flux into the carotenoid biosynthetic pathway or by downregulating the turnover of β-carotene into downstream metabolites (Sayre et al., 2011).

Aside strategies directed towards manipulating carotenoid biosynthesis or degradation, promoting differentiation of carotenoid sequestration structures such as chromoplasts can improve the sink capacity in plant cells and enhance carotenoid levels (Torres-Montilla & Rodriguez-Concepcion, 2021). Carrot and cauliflower carrying the *Or* mutant displayed large amounts of carotenoid sequestering structures with high levels of carotenoid accumulation (Lopez et al., 2008; Li et al., 2001). Co-transformation of cassava FEC from TMS60444 with ZmPSY1 and MeOr resulted in high levels of β-carotene and total carotenoids that is more than 3-fold higher than in the FEC transformed individually with ZmPSY1 or MeOr (Jaramillo et al., 2022).

Biofortification measures applied to tackle iron and zinc deficiency in cassava include overexpression of Arabidopsis thaliana AtVit1 and co-expression of a mutant A. thaliana iron transporter (AtIrt1) and ferritin (AtFer1) resulting in transgenic cassava lines with elevated levels of iron and zinc in their storage roots (Narayanan et al., 2019).

Table 1. Cassava traits targeted by biotechnology

Traits of interest (genes)	Cassava genotypes	Source
Biofortified β-carotene (crtB)	60444	Welsch et al. 2010
Biofortified β-carotene (crtB, crtI, crtY)	60444	Bonilla 2010
Leaf retention (senescence-inducible *ipt*)	60444	Zhang et al. 2010
Waxy starch (RNAi GBSSI)	60444	Zhao et al. 2011
CBSVD (RNAi FL-CP)	60444	Yadav et al. 2011
Protein content/cyanogenic content (HNL)	60444	Narayanam et al. 2011
RNAi CMD (ACMV/EACMV); CBSD (n.d.)	60444	Taylor et al 2012
Iron biofortification (FEA1)	60444	Ihemere et al. 2012
CMV and CBSV resistance (RNAi-CBSV coat protein)	TME7 (Oko-Iyawo)	Vanderschuren et al. 2012
Waxy starch (RNAi-GBSSI)	Adira4	Koehorst-van Putten et al. 2012
UCBSV resistance (siRNA-UCBSV coat protein)	60444	Ogwok et al. 2012
Biofortified β-carotene (crtB and DXS)	60444	Failla et al. 2012
UCBSV resistance (RNAi-UCBSV coat protein)	60444	Odipio et al. 2014
Resistance to Sri Lankan CMV (AV2 and AV1 coat proteins)	KU50	Ntui et al. 2015
Iron biofortification (AtVIT1)	TME204	Narayanam et al. 2015
Resistance to CBSV and UCBSV, increase carotene content in roots	TME204, TME7, 60444	Chauhan et al. 2015
Biofortified vitamin B6 (AtTDX1.1 and AtTDX2)	60444	Li et al. 2015

2.4 Transcriptome and metabolite analyses

Integrated analysis of metabolites and gene expression can offer insight into several important biological processes in plants. Transcriptomic studies provide a platform for understanding pathways and mechanisms that control cell growth and development, the association between genotype and phenotype and can be used for gene discovery for biotechnological applications (Ruan et al., 2004), which will promote breeding crops that are tolerant to both biotic and abiotic stresses and with promising nutritional traits. To improve nutrition and health, transcriptomic studies can be used for the identification of genes and regulators that can be targets for carotenoid biofortification in major staple crops. Several studies applying biotechnology and/or gene editing technology in cassava to improve specific traits have been recorded (Table 1). The recent elucidation of the cassava genome sequence provides tools for cassava breeding and research and facilitates the acquisition of -omics data. For instance, transcriptome analyses have revealed that cassava roots development is under hormonal regulatory control evidenced by increased expression of genes involved in cell differentiation and proliferation thus promoting growth and starch metabolism (Utsumi et al., 2020). Carotenoid biosynthesis is thought to be a regulated variation in the abundance of carotenogenic gene transcripts in carrots (Clotault et al., 2008). Several transcriptome studies have been conducted on storage root development, response to biotic and abiotic stresses and more recently more studies have been conducted on cassava nutritional traits, particularly with regards to carotenoid accumulation.

Metabolic profiling is used to quantify the metabolome of biological samples allowing for untargeted screening of the metabolic status for a given sample. Liquid chromatography (LC) coupled with mass-spectrometry (LC-MS) and gas chromatography (GC) coupled with MS (GC-MS) are two prominent analytical approaches for untargeted metabolic profiling covering a wide spectrum of low-weight compounds (Ganna et al., 2015). Metabolite composition of several cassava varieties have been analyzed on UPLC and LC/GC-MS platforms, creating metabolic profiles that can be linked to characteristic traits that are helpful in elucidating the relationship between genotype and phenotype and ultimately useful for cassava improvement (Obata et al., 2020; Drapal et al., 2019).

2.5 Legal framework for gene editing in SSA

Improvements to cassava have mostly been led via conventional plant breeding programs (Bull et al., 2011); however, the limitations identified with the crop may require modern biotechnology tools to hasten the generation of improved farmer-preferred cultivars. As genome-editing tools such as CRISPR/Cas9 tool are becoming more popular, it is now a focus of discussions in plant breeding as talks of its safety and risk concerns from its implementation are on the rise. Also, with breeders embracing genome-edited varieties in their breeding selections, there have been discussions on whether products resulting from using these techniques should be classified as genetically modified organisms (GMOs) and whether similar regulatory policies applied to GMOs should be applied to these products as well (Jenkins et al., 2021). For instance, the European Union classified organisms obtained through genome editing methods such as CRISPR-CAS and TALEN as GMOs and as a result are subject to regulations currently required for transgenic organisms GMO legislation (Bettina Wanner, 2019). However, regulatory perspectives and policies on gene editing techniques remain different from nation to nation and strong growth has been reported in both developing and industrialized countries (Zhang et al., 2020). Application of genetic modification technologies appears limited relative to non-transgenic biotechnology methods in developing countries of SSA, South Africa being an exception, compared to other countries (Takeshima, 2010). As the global area of GM products continues to grow, it is important to note the drivers of the differences in policies and regulations in different countries. For instance, studies have revealed regulatory and social uncertainties, public funding and technological uncertainty, and a strict legal regime governing agricultural biotechnology are among the drivers of development (Einsele, 2007;Eriksson et al., 2018;Lassoued et al., 2018). Other drivers include regulatory assessment, intellectual property and patents, acceptance by consumers, regulators, influence of non-governmental organizations (NGOs), conflicting interests of stakeholders, socio-political legitimacy, where cultural aspects and political influences matter (Einsele, 2007;Eriksson et al., 2018;Lassoued et al., 2018). Ultimately, each country has a different stance on regulatory approaches to agricultural biotechnology and in some countries, techniques are treated on a case-by-case method while in some follow a more restrictive approach. Paper IV compares the biosafety regulatory frameworks and approaches towards GMOs and new breeding technologies in Kenya, Nigeria and Uganda, and Sweden as a representative of an EU member country, and how they influence research and development. Our findings showed that political,

financial, and economic factors and funding opportunities are more likely to influence research and development than biosafety regulatory frameworks in these countries.

3. Thesis aims

The aim of this thesis is to provide a better understanding of the molecular mechanisms governing carotenoid and starch biosynthesis in cassava with the end goal of producing cassava cultivar(-s) that are high in carotenoid composition while retaining their starch and/or dry matter content. Furthermore, through biofortification of cassava using gene editing techniques this project aims to improve the provitamin A content of cassava.

Through metabolite and gene expression studies we aim to identify transcript signatures associated with carotenoid and starch biosynthesis. Linking carotenoid and starch levels to gene expression is required to form the knowledge base leading to nutritional improvement of cassava.

Finally, the study aims to analyse and compare the biosafety regulatory practices and policies in Kenya, Nigeria, Uganda, and Sweden to show how policy and regulatory developments in these countries may present bottlenecks to collaborations involving the use of gene editing technologies.

Specific aims

- To elucidate the carotenoid biosynthesis pathway and identify transcript signatures associated with βcarotene accumulation and regulation in cassava.
- To quantify the carotenoid and starch content of selected Nigerian cassava landraces by targeted and untargeted metabolite analysis to reveal their metabolite composition.
- To identify new genes and mutations based on transcript and metabolite data from cassava and understand the regulation of these genes.
- To integrate gene expression and metabolic data to understand how to modulate the carotenoid and starch pathway in cassava.

- To modify selected genes associated with the carotenoid biosynthesis pathway in cassava using CRISPR/Cas9 targeted gene editing and see if modification of these genes results in increase in yellowness and carotenoid content.

4. Methods

4.1 Field trials

While greenhouse studies are mostly preferred to field studies, a field study provides data from an actual agricultural setting. Also, the laboratory being a controlled environment, prevents insight into the effects of biotic and abiotic stressors on the agronomic properties and traits of plants. Since cassava cultivation is carried out in the fields and agronomic properties/traits are an important part of this research, a field experiment rather than greenhouse studies is a more appropriate approach. In paper I, we carried out a field study from which agronomic parameters and performance were collected and recorded during sampling and finally at harvest. Leaf and root samples (Figure 4) used in all laboratory experiments were collected from four- and eight-months old plants. Plants from this field study were also used in paper II for untargeted metabolite and RNA sequencing studies.

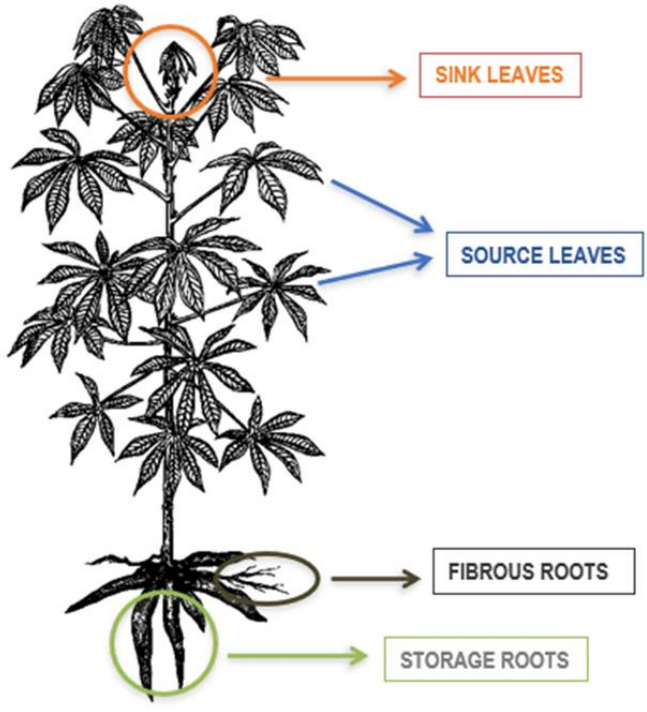

Figure 4. Sketch of the cassava plant, source leaves, fibrous and storage roots. Source leaves and storage roots were collected to obtain agronomic, metabolites, and transcriptomic analyses. John Eifick University of Queensland, AU.

4.2 RT-qPCR and transcriptomics

Gene expression analysis such as reverse transcription quantitative real-time PCR (RT-qPCR) and RNA-seq are used to measure transcript levels of genes in a sample. RT-qPCR is fast, accurate, sensitive, cost-effective, and medium throughput gene expression analysis able to measure expression of a limited number of genes (Derveaux et al., 2010). It is also the golden standard for the detection and quantification of target genes of interest, allowing for the analysis of a wide range of samples with differing target abundance (Bustin and Mueller, 2005). While RT-qPCR can analyze more than one transcript at a time, RNA-seq is a high-throughput technique with the capacity to measure thousands of transcripts per assay and is a preferred technique for whole transcriptome analysis. In paper I, gene expression levels of five carotenoid biosynthesis pathway genes phytoene synthase 1 and 2 (*PSY1* and

2), β-carotenoid hydroxylase (*CHYβ*), lycopene-ε-cyclase (*LCYε*), and 9-cis-epoxycarotenoid dioxygenase (*NCED1*) in white and yellow cassava landraces were determined using RT-qPCR. More importantly, we assessed correlation between carotenoid content and expression levels of selected carotenoid-related genes.

Next-generation sequencing (NGS) technologies through RNA-seq revolutionized transcriptomics by eliminating the many challenges posed by hybridization-based microarray approaches that were previously used for measuring gene expression (Kukurba and Montgomery, 2015). Gene expression profiling by RNA-seq provides a high-resolution view of the global transcriptional landscape. In a typical RNA-seq experiment, isolated RNA is fragmented to meet size specifications and amplified to generate complementary DNA (cDNA) libraries, cDNA is then sequenced on NGS platform, and the resulting reads are aligned to a reference genome (Auer and Doerge, 2010). Alternatively, a transcriptome can be *de novo* assembled where after the reads are mapped. In paper II, we employed RNA-seq to analyze differences in gene expression in leaf and root samples of white and yellow landraces. Sequencing reads were mapped against the reference genome and gene expression and differential gene expression among landraces were calculated.

4.3 Targeted and untargeted metabolite analysis

Targeted and untargeted metabolomics platforms allow for assessment of a wide range of metabolites in plants. As a target for nutritional improvement, concentrations of β-carotenoids in leaf and root samples of cassava landraces were determined using ultra-high performance liquid chromatography coupled with diode array detection and time-of-flight mass spectrometry (UHPLC–DAD–ToF–MS) to determine (paper I).

For untargeted metabolite analysis, Gas Chromatography-mass spectrometry (GC-MS) and Liquid Chromatography-Mass Spectrometry (LC-MS) were employed (paper II). GC-MS is used to analyze and separate compounds in a sample and the mass to spectrum ratio of ions, which is done as the separated gas moves down the column of the device. LC-MS is similar to GC-MS, but it has the capacity to analyze complex natural biological mixtures. Also, GC–MS has a preference for small polar metabolites covering primary metabolism while LC–MS covers large hydrophobic metabolites that are predominant in secondary metabolism (Doerfler et al., 2013). While targeted metabolite analysis is used for the identification and

measurement of specific metabolites, untargeted metabolomic techniques allows for comprehensive profiling of metabolites on a broader scale (Yandeau-Nelson, 2015); however, many of these metabolites may not be identified due to their absence in public mass spectral libraries and/or a lack of appropriate standards for comparison (Kueger et al., 2012). Furthermore, metabolites may be masked by other metabolites or not properly ionized, preventing detection (Moco et al., 2007).

4.4 CRISPR-Cas9 mediated gene editing

CRISPR-Cas9 technology is a site-specific genome engineering technology derived from prokaryotic adaptive immune system that provides precise molecular editing in organisms (Barrangou and Doudna, 2016). CRISPR-Cas9 reagents can create a targeted DNA double-strand break (DSB) in a genome. DNA breaks can be repaired either by nonhomologous end joining (NHEJ), which creates insertion/deletion (indel) mutations at the break site knocking out gene function or by homology dependent repair (HDR), which copies information from a donor DNA template to achieve the desired DNA sequence modification (Cermak et al., 2017).

Using the CRISPR/Cas9 system guide design program CRISPOR by tefor infrastructure (Concordet and Haeussler, 2018), CRISPR guides were designed from the coding sequence of the Manihot esculenta CHYβ gene. Generated 20bp targets containing a protospacer adjacent motif (PAM) at the 3' end was analyzed, and two targets were selected based on specificity score and low number of off targets. Using the editing vector pDIRECT_22C (Addgene plasmid, #91135), primers were designed using the primer design web tool for plant genome engineering by Voytas laboratory (Cermak et al., 2017). CRISPR construct was assembled by "Golden Gate reaction" and transformed into competent E. coli DH5α. The resulting clones were mobilized into Agrobacterium tumefaciens strain LBA4404 by electroporation as described in paper III.

4.5 Agrobacterium-mediated transformation

Agrobacterium-mediated transformation is the use of the bacterial pathogen Agrobacterium tumefaciens to deliver genes of interest into a host plant. This technology has been used to enhance plant tolerance/resistance to biotic/abiotic stresses which ultimately increases crop productivity and to enhance nutritional content of crop plants (Ziemienowicz, 2014).

The process involves the transfer of genetic components carried by *A. tumefaciens* into the host plant. Genetic components required include the T-DNA, the Ti plasmid virulence vir region and the chromosomal virulence loci, chvA, chvB, and pscA which are essential for the transfer process (Ziemienowicz, 2014)). To initiate T-DNA transfer, plant cells are wounded, and Agrobacterium recognizes molecules secreted by wounded cells as a signal to induce the expression of vir genes which activates T-DNA transfer (Ziemienowicz, 2014)).

Agrobacterium-mediated transformation of friable embryogenic calli (FEC) is the most widely used method in cassava transformation to generate transgenic cassava plants (Bull et al., 2009). Paper III describes Agrobacterium-mediated transformation of FECs generated from axillary buds of the cassava cv. TMS60444 with LBA4404-CHYβ following previously defined methods by Zainuddin et al. (2012).

5. Results and discussion

5.1 Agronomic data and targeted metabolic profiling

The understanding of the mechanism of carotenoid accumulation and sequestration in cassava through metabolite and transcriptomic analysis of landraces with varying β-carotene content is central in this thesis. Although the core carotenoid biosynthesis that constitutes the pathway in plants is well established, the mechanism of accumulation in non-photosynthetic tissues such as in storage roots of cassava is not well understood. Also, to identify suitable material for biofortification among locally adapted and farmer-preferred landraces we analyzed agronomic, metabolite and gene expression profiles of a select subset of cassava landraces to increase the understanding of the physiological and genetic basis of carotenoid biosynthesis and regulation.

Paper I detailed the outcome of the targeted metabolite and RT-qPCR analysis and agronomic properties of cassava landraces. Agronomic data analyzed over a period of twenty-six months representing two cropping cycles showed that starch and dry matter content are positively correlated as could be expected. Percentage starch ranged from 50% to 80% dry weight with Kaleso a white cultivar having the highest percent starch and IBA083565 a yellow cultivar having the lowest starch content; however, starch assay and dry matter analysis also showed that some landraces with yellow storage roots have reasonably high percentage of starch and dry matter as well (paper I: figure 4).

Metabolite composition of thirteen landraces with white and varying intensity of yellow roots were determined by UHPLC–DAD–ToF–MS (Figure 5a) showed that leaf samples contain Lutein, Chlorophyll A and B,

β-carotene, 9-cis-β-carotene, 15-cis-β-carotene, Zeaxanthin, Neoxanthin 1 and 2, and Violaxanthin 1 and 2. Surprisingly, the chlorophyll A:B ratio ranged from 6.7 to 8 which is particularly high compared to previous reports and other crops(Zhang et al., 2017; Anand & Byju, 2008). β-carotene, 9-cis-β-carotene, and 15-cis-β-carotene were the only carotenoid type detected in root samples and their content was lower compared to the content in leaf samples. In the root samples, total carotenoid content increases with increasing intensity of the yellow color with values ranging from 0.3μg g−1 dry weight in white landrace to Kaleso to 36.1μg g−1 dry weight in deep yellow landrace IBA141092. Total carotenoid content of up to 25.8 μg g-1 was previously reported in cassava and increasing total carotenoid was found to be accompanied by increasing total β-carotenoids (Ceballos et al., 2013).

Correlation analysis of agronomic and metabolite data were conducted to check for negative and/or positive correlations that may exist between these traits. It has been previously reported that dry matter content (DMC) and carotenoid content are negatively correlated traits in cassava, particularly in African germplasm (Rabbi et al., 2017; Esuma et al., 2016b; Esuma et al., 2016a; Esuma et al., 2012; Akinwale et al., 2010), which may hamper biofortification efforts that aim to combine these two traits since dry matter and starch content are important traits for farmers and industry. Transgenic cassava and potato lines co-expressing transgene crtB and DXS also showed increased β-carotone content and a decrease in DMC and starch (Beyene et al., 2017). As discussed in paper I, Pearson and Spearman correlation showed a weak negative correlation of −0.3 and −0.2, respectively between β-carotenoid content and dry matter/starch content, suggesting that the negative correlation is not so strong to prevent combining these traits. However, the number of cassava genotypes used to calculate these correlations were few, which hampers this analysis. Large scale studies using cassava genotypes of South American variants reported low regression coefficients close to zero between DMC and total carotenoid content, suggesting that dry matter and carotenoid contents are independent variables and there is possibility of developing clones combining high carotene and high dry matter levels (Sanchez et al., 2014; Ceballos et al., 2013). Also, many of the high carotene content cassava cultivars used in the African breeding programs were introduced from South America through the collaborations between the International Center for Tropical Agriculture (CIAT), the International Institute of Tropical Agriculture (IITA), and Nigeria's National Root Crops Research Institute (Ceballos et al., 2013; Njoku et al., 2011). Hence, there is a possibility that the negative correlation between DMC and total carotenoid content observed in African germplasm

is linked to location and environmental conditions and there is possibility that these two traits can be combined in African germplasm as well (Ceballos et al., 2013).

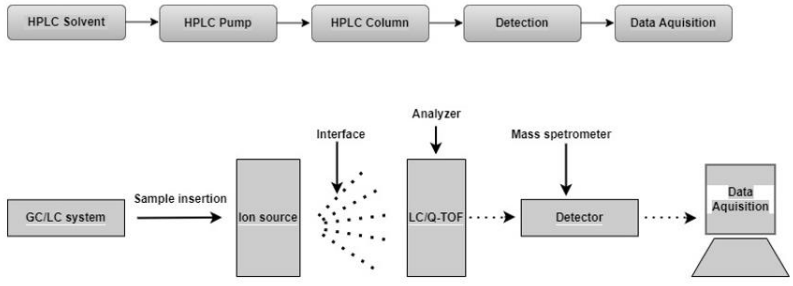

Figure 5. Schematic representation of targeted and untargeted metabolite analyses by HPLC and GC/LC-MS systems.

5.2 Gene expression by qRT-PCR

Gene expression analysis by qRT-PCR was done for five carotenoid biosynthesis genes *psy1*, *psy2*, *chyβ*, *lcyε*, and *nced1* across six cassava genotypes. Gene expression levels were determined in the storage roots of six cultivars: Kaleso, IBA011231, IBA083724, IBA011663, IBA141092, and TMS60444. Expression analysis showed Kaleso, a white landrace, had the lowest expression of all five genes. The four yellow landraces with varying color intensity expressed the select genes differently; however, IBA011663 a yellow cultivar had the highest expression of *psy2*, *chyβ* and *lcyε* compared IBA141092 a deep yellow cultivar which had the highest total carotene and β-carotene content (paper I; figure 6). In line with the report by Welsch et al. (2010), there seem to be no correlation between β-carotene content and levels of expression of carotenoid biosynthesis genes.

5.3 Sequence analysis

Sequence alignment of selected genes based on RNA-seq data revealed three sequence changes between white cultivars A, P, R and yellow cultivars J, O, L, D, I, G (paper II, S1). PSY protein sequence alignment showed an amino

acid change at position 191 from Alanine in the white cultivars to Aspartic acid in the yellow cultivars; however, this mutation is not found in pale-yellow cultivars B and C (paper II; supplementary figure 1) suggesting that the mutation does not fully explain the differences in carotenoid accumulation in cassava. Also, PSY protein modeling in white and yellow cultivars, Kaleso (A) and IBA102103 (J) showed a unique difference in structure exactly upstream of position 191. In this position, model A's loop looks outwards whereas model J has a loop looking inwards (Figure 6). In a previous study by Welsch et al. (2010) single nucleotide polymorphism in *psy2* carrying a C572A nucleotide substitution leading to an A191D amino was found in yellow cultivars. Due to the correlation between color and this SNP was hypothesized to be crucial for root color formation (Welsch et al., 2010).

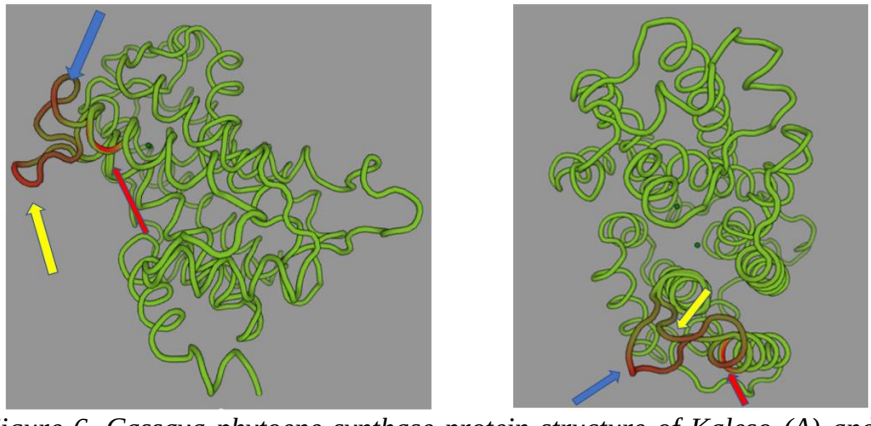

Figure 6. Cassava phytoene synthase protein structure of Kaleso (A) and IBA102103 (J). Overlaying the two models A and J shows the unique difference in structure exactly upstream of pos 191 (red arrow). Model A has a loop looking outwards (yellow arrow) whereas model J has a loop looking inwards (blue arrow)

5.4 Transcriptomic and untargeted metabolomic data

To determine molecular differences between white and yellow cassava cultivars, we analyzed untargeted metabolic profiling and comparative gene expression analysis from RNA-seq data (Figure 7) of leaf and root samples. In the expression matrix generated, leaf and root samples had 13,383 and 14,136 genes with an FPKM value greater than one (FPKM > 1) in at least one sample respectively. Principal Components Analysis (PCA) of the genes from both leaf and root samples separated landraces into three

groups/clusters (paper II; figure 1). We observed partial separation of white and yellow cultivars in the two in PC1 and PC2 and clear separation of the three white cultivars A, P, and R from the yellow cultivars in the PC2 and PC3. Cultivars I, J, and L have the highest β-carotene content in roots (paper II; Table 1) and they can be seen clustering together in PC1 and PC2 (paper II; figure 1). In a different analytical approach, hierarchical clustering was performed on significantly expressed genes (FPKM>3) in roots using EXPANDER (Hait et al., 2019) and clusters that were upregulated in white cultivars, but downregulated in yellow and vice versa were selected for functional analysis (paper II, figure 6). Functional analysis by Gene Ontology analysis (paper II; figure 5) revealed that amino sugar biosynthetic processes, photosynthetic processes and light reactions were enriched in white landraces compared to yellow landraces while protein catabolic processes, organonitrogen metabolic processes, protein/ribonucleotide binding and ncRNA processing were most enriched in yellow landraces. These findings suggest protein-related catabolic processes may influence carotenoid biosynthesis and accumulation in yellow cultivars cassava. Ion transporter-related transcripts which may play a part in carotenoid transport were also overrepresented in yellow landraces. Among other regulatory mechanisms, carotenogenesis is thought to be regulated post-translationally via protein-protein interactions leading to carotenoid sequestration (Sun and Li, 2020). For instance, PSY, a major enzyme in the carotenoid biosynthesis pathway is known to be under the post-translational regulation of molecular chaperone OR protein, to maintain PSY activity in both chloroplasts and chromoplasts (Li et al., 2012;Zhou et al., 2015;Park et al., 2016).

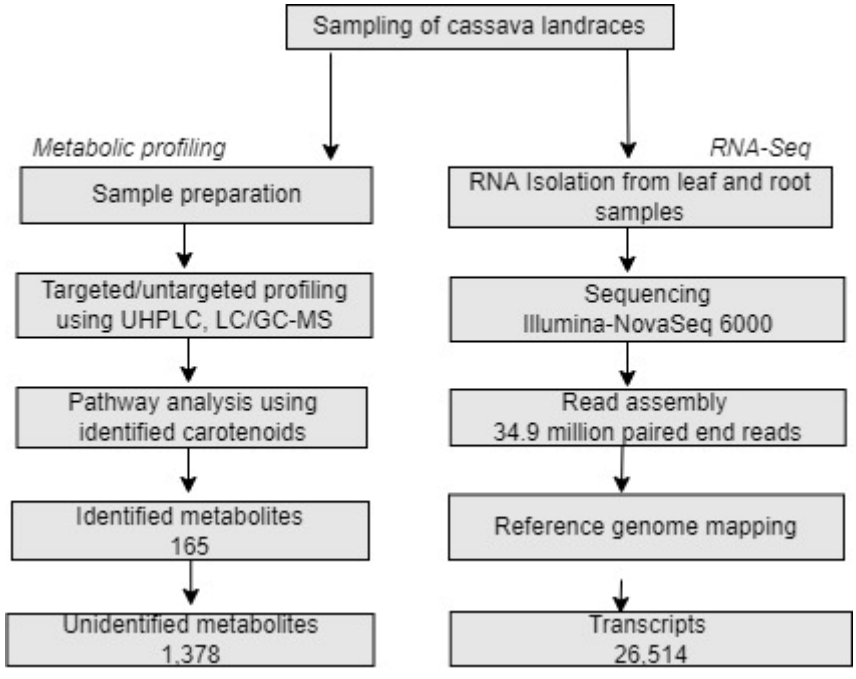

Figure 7. Workflow of metabolite and transcriptomic analyses

Genes of the carotenoid biosynthesis pathway are conserved in plants. In this study, 46 carotenoid biosynthesis genes were identified and their expression levels between genotypes were compared. Between leaf and root samples, 26 genes showed higher expression in leaves than in roots while 20 genes showed lower expression in leaves than in roots (paper II; figure 3). Consistent with the previous report by Welsch et al. (2010), there was no general correlation between the expression of carotenoid biosynthesis genes and root color and/or β-carotene content in root samples. Nevertheless, we observed a higher expression of carotenoid cleavage dioxygenases (CCDs) in white cultivars compared to yellow cultivars. To identify potential candidate genes that may regulate carotenoid accumulation, we analyzed significant differential gene expression (p-value < 0.05, Fold Change +/- 1.5) between yellow and white genotypes. Results showed that members of the RING/U-box protein superfamily mediating protein–protein interactions and ubiquitination (Trujillo, 2018), C2H2-type domain-containing protein (Manes.04G155200) a Zinc-finger protein known for its important roles in plant growth and development, and responses to biotic and abiotic stresses (Ming et al., 2019), and Cystatin B, a known cysteine proteinase inhibitor (Martinez et al., 2012) were much more expressed in yellow genotypes

compared to white genotypes.

Untargeted metabolic profiling of white and yellow cassava cultivars via GC/LC-MS systems identified 154 metabolites. PCA showed white and pale-yellow cultivars clustered together but with high variance between biological replicates (paper II, figure 7). We also observed that amino acids such as L-asparagine and L-serine, organic acids fumaric and malic, and signaling molecule pipecolic acid were up-regulated in the white cassava cultivars compared to yellow cultivars. This observation is in line with a previous study which showed that white flesh cassava contained higher contents of amino acids and their derivatives compared to the yellow varieties (Xiao et al., 2021). We also observed upregulation of amino acid proline and GABA pale-yellow cultivars with a negative correlation between GABA and β-carotene. Overall, the distinguishing factors between white and yellow cassava cultivars are the amino acid, organic acid, osmolyte and primary sugar content. While amino acids, organic acid and osmolyte content are higher in white and/or pale-yellow cultivars compared to yellow cultivars, sugars such as sucrose, fructose, and glucose are less abundant in white lines compared to yellow cultivars as revealed in the GC-MS data. Finally, 1378 features were unmatched in the LC-MS data, which limits the exploration of metabolic diversity of these cultivars. Since elucidating metabolic profiles and genetic regulation of metabolites are important for nutritional improvements of crops (Fang et al., 2019), identifying unmatched features may provide a better understanding of how gene expression or regulation influences metabolism and ultimately phenotype.

5.5 CRISPR-Cas9 Gene editing

The CRISPR-Cas9 system is an RNA-directed DNA endonuclease system is currently considered the most powerful gene editing tool (Mao et al., 2019), and is speedily replacing other programmable nucleases, such as zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) because they are easy to use and less time consuming (Woo et al., 2015). In the CRISPR-Cas9 system, gRNA which consists of CRISPR RNA (crRNA) recognizes specific targets in the genomic DNA and a tracr RNA scaffold for Cas9 binding while Cas9 nuclease cleaves DNA (Mao et al., 2019). CRISPR-induced mutations are small insertions and deletions (indels) rather than large fragment insertions or rearrangements common to transgenesis (Mao et al., 2019). B-carotenoid hydroxylase (CHYB),

lycopene-ε-cyclase (LCYƐ), and 9-cis-epoxycarotenoid dioxygenase (NCED1) are key enzymes of the carotenoid biosynthesis pathway. A-carotene and β-carotene both undergo hydroxylation by LCYƐ and CHYB to generate lutein in the α-branch and zeaxanthin in the β-branch of the carotenoid pathway respectively, while NCED1 like other NCEDs/CCDs, cleaves β-carotenoids violaxanthin/neoxanthin into apocarotenoids and abscisic acid (Luan et al., 2020). Thus, LCYƐ was selected as a candidate for gene editing to block the α-branch of the pathway to increase flux of lycopene to the β-branch of the pathway while CHYB was selected to increase the accumulation of β-carotene by inhibiting or downregulating its hydroxylation activity/reaction. Silencing NCED1 should also increase the accumulation of downstream β-carotenoids such as violaxanthin and neoxanthin. Silencing of CHYB in sweet potato callus using RNAi resulted in a 38-fold increase in β-carotene (Kim et al., 2012). Also, CRISPR/Cas9 directed editing of lycopene epsilon-cyclase in bananas resulted in a significant reduction of α-carotene and an increase in β-carotene content ranging from 46% to 168% (Kaur et al., 2020). Several proof-of-concept studies have been carried out to confirm the efficiency of the CRISPR/Cas9 system and the advantages of its implementation in plant crops. Thus, we employed the CRISPR/Cas9 gene editing system to modify *chyβ*, *lcyε*, and *nced1* genes in cassava. pDIRECT_22C vector (Cermak et al., 2017), encoding two guide RNAs designed from *chyβ, lcyε, and nced1* gene sequences were assembled and cloned into *E. coli* cells. Successful transformation was confirmed by restriction enzyme digest of isolated plasmid and Sanger sequencing (paper III, figure 3A & 3B). Following confirmation, pDIRECT_22C vectors carrying two gRNAs for each gene were introduced to TMS60444 Calli via agrobacterium-mediated transformation.

5.6 Agrobacterium transformation

Genetic transformation is a technique used to introduce desirable traits and characteristics into an existing plant genomes and *Agrobacterium*-mediated transformation is a well-established and most reported method used to achieve genetic modifications in plants (Ziemienowicz, 2014). *Agrobacterium tumefaciens* strains generate transgenic plants by transferring its DNA (T-DNA) into the host plant genome. The molecular mechanism of

gene transfer involves: (1) initial attachment of *A. tumefaciens* to the cell surface, (2) the transfer of T-DNA and Vir effector proteins into plant plasma membrane, (3) cytoplasmic trafficking of VirD2/T-strands to the nucleus, (4) integration of T-DNA into plant genome, (5) expression of transgene (Gelvin, 2010;Pitzschke and Hirt, 2010). In this study, FEC from cv. TMS60444 were transformed with pDIRECT_22C vector (Cermak et al., 2017) carrying two guide RNAs each from *CHYβ,* LCYε, and NCED1 via *Agrobacterium*-mediated transformation. A deeper yellow coloration observed in transformed Calli (paper III, figure 2A) suggests a higher beta-carotene content than untransformed control Calli; however, no shoots were recovered from LBA4404_LCYε and LBA4404_NCED1 transformation experiments. Shoots from LBA4404_CHYβ rooting on selection media (paper III, figure 2B) (basic cassava media supplemented with kanamycin) is also an indication of possible successful transformation (paper III, figure 3). Transformation experiment gave five putative transgenic shoots from the LBA4404_CHYβ transformed Calli and were maintained in basic cassava media with antibiotic selection for subsequent analysis to confirm that shoots are transformants.

5.7 PCR Confirmation and sequencing

PCR to confirm T-DNA insertion using primers specific for the CmYLCV and Cas9 promoters gave mixed and inconclusive results (paper III, figure 5B). Also, PCR amplification of gDNA from regenerated LBA4404_CHYβ shoots using primers specific for CHYβ gene showed no reduction in product size suggesting there was no deletion event. A fragment from the CHYβ gene spanning the putative deletion site was amplified from regenerated plantlets gDNA and sequenced. Sequence alignment and analysis showed no deletion or mutation in CHYβ sequences of all five shoots and all sequences generated were similar to untransformed control (paper III, figure 6).

6. Conclusions

The steady state levels of carotenoids in plant tissues are maintained by continuous biosynthesis, sequestration, degradation, among other processes governed by complex regulatory networks. Although the knowledge of carotenoid biosynthesis and retention in plants have increased, the mechanism of carotenoid accumulation and regulation in cassava is not fully understood. In this study we determined that carotenoid biosynthesis and accumulation in cassava is not only a factor of mutation in *psy* gene sequence but is also influenced by the levels of expression of genes which are not previously characterized as carotenoid related. Furthermore, in this study, the SNP in the *psy* gene attributed to the increased β-carotene accumulation in yellow cassava varieties was absent in both white and pale-yellow cassava cultivars suggesting that other factors influence carotenoid accumulation in cassava. Therefore, more knowledge on the molecular mechanisms involved in carotenoid accumulation is needed in order to take full advantage of their benefits to health and nutrition.

While increased accumulation of carotenoids seems to be directly linked to the expression of carotenoid biosynthesis genes in other plants like tomato and carrot, cassava cultivars with yellow roots have similar expression profile of β-carotene biosynthesis may also be regulated by genes and factors both within and outside of the carotenoid biosynthesis pathway. There was no clear difference in expression of carotenoid biosynthesis genes between white and yellow rooted cultivars. However, we identified the carotenoid cleavage dioxygenase gene *nced*, as a potential distinguishing factor between white and yellow cassava cultivars and speculate that ABA synthesis may play a role in β-carotene regulation. However, we identified the carotenoid cleavage dioxygenase gene *nced*, as a potential distinguishing factor between white and yellow cassava cultivars and speculate that ABA synthesis may play a role in β-carotene regulation. Additionally, we determined that the expression of several genes outside of the described carotenoid biosynthesis

pathway correlated with high β-carotenoid accumulation in yellow cassava cultivars. Our findings suggest that several protein-related catabolic processes were enriched in yellow cultivars; therefore, their roles in carotenoid accumulation should be further investigated and validated. This study emphasizes the need to study the mechanism of carotenoid accumulation in cassava beyond gene expression and the carotenoid pathway.

Contrary to some reports that carotenoid/β-carotene and dry matter content are negatively correlated in African cassava germplasm, in paper I we identified some cassava landraces with high β-carotene content and relatively high dry matter content and observed a weak negative correlation between the two traits. Also, low regression coefficients were observed between both traits in South American cassava germplasm and therefore deemed insignificant. Based on these findings we are of the opinion that β-carotene and dry matter can be combined in using a suitable cultivar. Therefore, biofortification is still a promising measure to increase the provitamin A content in the cassava. It is also worth mentioning that β-carotene was the only carotenoid type identified in the landraces used in this study.

Agrobacterium-mediated transformation of cv. 60444 FEC with LBA4404_chyβ produced a yellow coloration that was absent in untransformed control FEC. Regenerated shoots from transformed calli showed insertion of a truncated T-DNA while gDNA sequencing analysis showed intact gRNA sequence. DNA sequences of shoots regenerated from transformed FEC were similar to that of untransformed shoots. There was no evidence of insertion or deletion in the sequence generated. We were also not able to generate independent transgenic lines evidenced by the lack of mutation (insertion/deletion) in the regenerated shoots. In summary, we determined the need to continue to investigate the regenerated shoots and repeat the transformation experiments with adjustments in experimental design and process.

Finally, in investigating the impacts of EU regulatory frameworks on the use of new breeding techniques in plant breeding we determined that the European union regulatory framework on genetically modified organisms may not necessarily affect research and development in Africa. We suggest that political influence or decisions, finance and economics should be further studied as these factors might finally affect the decision of African countries on adoption of GMOs and NBTs. Evidence gathered through research and direct interviews with stakeholders in four African countries showed that

each of these countries have their own regulatory frameworks on NBTs which is not reflective of the EU's stand on NBTs and do not apply similar strict rules placed on GMOs on gene edited plants in which there is no foreign DNA in final product.

7. Future perspectives

Our research showed that carotenoid synthesis is regulated by genes and/or factors within and outside of the carotenoid biosynthesis pathway, further investigations into the molecular mechanisms that influence carotenoid biosynthesis and accumulation is important for African cassava germplasm. For instance, plastid differentiation, plastid type and size, carotenoid pathway intermediate products, particularly ABA and downstream products that feedback into β-carotene synthesis should be further studied.

The *nced* gene was significantly differentially expressed in white cultivars compared to yellow cultivars; therefore, we suggest a *nced* and other *ccd* gene knock-out in cassava lines with varying content of β-carotene or transformation with a mutant *Or* gene. Also, we identified candidate genes that could be targets for biofortification that could potentially influence carotenoid biosynthesis and accumulation such as genes encoding the Zinc-finger protein family that are involved in protein-protein interaction. Also, based on the outcome of the transformation experiment we propose the use of multiple guide RNA targets in future experiments.

A major challenge faced in cassava transformation experiments is the duration of time it takes from the transformation events to regeneration and production of cassava shoots. In this study, it took approximately 1-1.5 years, from the time of transformation to the regeneration of shoots; therefore, screening for putative transformants in calli is proposed for future experiments. Also, the low rate of regeneration of shoots from transformed FEC is a challenge. Low regeneration efficiency and generation of few independent transgenic lines have been reported in cassava transformations as seen in the outcome of this study, suggesting the need to develop a better

transformation protocol for cassava. Therefore, the need for technical innovations for an effective delivery of the CRISPR/Cas9 editing system in cassava transformation as an additional tool to conventional plant breeding for the improvement of cassava cannot be overemphasized.

Since negative correlation between carotenoid and dry matter content is still a problem in some varieties, generating transgenes of several landraces with diverse backgrounds would identify cultivars that can combine both traits. Landraces are locally adapted cultivars generated from years of selection for desired traits and are known to carry unique alleles that are important for their genetic diversity; therefore, generating FEC from these landraces and including these landraces in genetic resources will increase the chances of producing farmer preferred cultivars with economically and nutritionally important traits. Genetic diversity is important for crop improvement as it allows for diverse selections or combinations which result in genetic gain. The cassava cultivar TMS60444 used in the transformation experiment detailed in paper III is susceptible to cassava mosaic virus infection and therefore not ideal for cultivation. It is therefore important to apply biofortification to locally adapted landraces to improve their nutritional qualities. Our attempt to generate FEC from two cassava landraces was not successful using both similar and modified protocols used to generate FEC in TMS60444 and other elite accessions; therefore, more efforts and knowledge is needed in bringing landraces into tissue culture.

References

References

Allem, A.C. (1994). The origin of Manihot Esculenta Crantz. *Genetic resources and crop evolution* 41**,** 133-150.

Auer, P.L., and Doerge, R.W. (2010). Statistical design and analysis of RNA sequencing data. *Genetics* 185**,** 405-416.

Barrangou, R., and Doudna, J.A. (2016). Applications of CRISPR technologies in research and beyond. *Nat Biotechnol* 34**,** 933-941.

Bettina Wanner, H.M., Andrea Mertens, Jörg Thomaier (2019). CJEU renders decision on the interpretation of the GMO Directive. *Journal of Intellectual Property Law & Practice* 14**,** 88-90.

Bull, S.E., Ndunguru, J., Gruissem, W., Beeching, J.R., and Vanderschuren, H. (2011). Cassava: constraints to production and the transfer of biotechnology to African laboratories. *Plant Cell Rep* 30**,** 779-787.

Bull, S.E., Owiti, J.A., Niklaus, M., Beeching, J.R., Gruissem, W., and Vanderschuren, H. (2009). Agrobacterium-mediated transformation of friable embryogenic calli and regeneration of transgenic cassava. *Nat Protoc* 4**,** 1845-1854.

Bustin, S.A., and Mueller, R. (2005). Real-time reverse transcription PCR (qRT-PCR) and its potential use in clinical diagnosis. *Clin Sci (Lond)* 109**,** 365-379.

Carvalho, L.J.C.B., Agustini, M.a.V., Anderson, J.V., Vieira, E.A., De Souza, C.R.B., Chen, S.B., Schaal, B.A., and Silva, J.P. (2016). Natural variation in expression of genes associated with carotenoid biosynthesis and accumulation in cassava (Manihot esculenta Crantz) storage root. *Bmc Plant Biology* 16.

Cazzonelli, C.I., and Pogson, B.J. (2010). Source to sink: regulation of carotenoid biosynthesis in plants. *Trends Plant Sci* 15**,** 266-274.

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

Cermak, T., Curtin, S.J., Gil-Humanes, J., Cegan, R., Kono, T.J.Y., Konecna, E., Belanto, J.J., Starker, C.G., Mathre, J.W., Greenstein, R.L., and Voytas, D.F. (2017). A Multipurpose Toolkit to Enable Advanced Genome Engineering in Plants

The Plant Cell 29**,** 1196-1217.
Cock, J.H. (1982). Cassava: A fBasic Energy Source in the Tropics. *SCIENCE* 218.
Concordet, J.P., and Haeussler, M. (2018). CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res* 46**,** W242-W245.
Cunningham, F.X., Gantt, J., and Gantt, E. (1998). GENES AND ENZYMES OF CAROTENOID BIOSYNTHESIS IN PLANT. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 49**,** 557-583.
Derveaux, S., Vandesompele, J., and Hellemans, J. (2010). How to do successful gene expression analysis using real-time PCR. *Methods* 50**,** 227-230.
Einsele, A. (2007). The gap between science and perception: the case of plant biotechnology in Europe. *Adv Biochem Eng Biotechnol* 107**,** 1-11.
El-Sharkawy, M.A. (1993). Drought-Tolerant Cassava for Africa, Asia, and Latin America. *BioScience* 43**,** 441-451.
Eriksson, D., Brinch-Pedersen, H., Chawade, A., Holme, I.B., Hvoslef-Eide, T.a.K., Ritala, A., Teeri, T.H., and Thorstensen, T. (2018). Scandinavian perspectives on plant gene technology: applications, policies and progress. *Physiol Plant* 162**,** 219-238.
Fang, C., Fernie, A.R., and Luo, J. (2019). Exploring the Diversity of Plant Metabolism. *Trends Plant Sci* 24**,** 83-98.
Gelvin, S.B. (2010). Plant proteins involved in Agrobacterium-mediated genetic transformation. *Annu Rev Phytopathol* 48**,** 45-68.
Hait, T.A., Maron-Katz, A., Sagir, D., Amar, D., Ulitsky, I., Linhart, C., Tanay, A., Sharan, R., Shiloh, Y., Elkon, R., and Shamir, R. (2019). The EXPANDER Integrated Platform for Transcriptome Analysis. *J Mol Biol* 431**,** 2398-2406.
Hannoufa, A., and Hossain, Z. (2012). Regulation of carotenoid accumulation in plants. *Biocatalysis and Agricultural Biotechnology* 1**,** 198-202.
Howitt, C.A., and Pogson, B.J. (2006). Carotenoid accumulation and function in seeds and non-green tissues. *Plant Cell Environ* 29**,** 435-445.
Jaramillo, A.M., Sierra, S., Chavarriaga-Aguirre, P., Castillo, D.K., Gkanogiannis, A., Lopez-Lavalle, L.a.B., Arciniegas, J.P., Sun, T., Li, L., Welsch, R., Boy, E., and Alvarez, D. (2022). Characterization of cassava ORANGE proteins and their capability to increase provitamin A carotenoids accumulation. *PLoS One* 17**,** e0262412.

Jenkins, D., Dobert, R., Atanassova, A., and Pavely, C. (2021). Impacts of the regulatory environment for gene editing on delivering beneficial products. *In Vitro Cell Dev Biol Plant*, 1-18.

Kaur, N., Alok, A., Shivani, Kumar, P., Kaur, N., Awasthi, P., Chaturvedi, S., Pandey, P., Pandey, A., Pandey, A.K., and Tiwari, S. (2020). CRISPR/Cas9 directed editing of lycopene epsilon-cyclase modulates metabolic flux for beta-carotene biosynthesis in banana fruit. *Metab Eng* 59, 76-86.

Kim, S.H., Ahn, Y.O., Ahn, M.J., Lee, H.S., and Kwak, S.S. (2012). Down-regulation of beta-carotene hydroxylase increases beta-carotene and total carotenoids enhancing salt stress tolerance in transgenic cultured cells of sweetpotato. *Phytochemistry* 74, 69-78.

Kukurba, K.R., and Montgomery, S.B. (2015). RNA Sequencing and Analysis. *Cold Spring Harb Protoc* 2015, 951-969.

Lassoued, R., Smyth, S.J., Phillips, P.W.B., and Hesseln, H. (2018). Regulatory Uncertainty Around New Breeding Techniques. *Front Plant Sci* 9, 1291.

Li, L., Yang, Y., Xu, Q., Owsiany, K., Welsch, R., Chitchumroonchokchai, C., Lu, S., Van Eck, J., Deng, X.X., Failla, M., and Thannhauser, T.W. (2012). The Or gene enhances carotenoid accumulation and stability during post-harvest storage of potato tubers. *Mol Plant* 5, 339-352.

Lindgren, L.O., Stalberg, K.G., and Hoglund, A.S. (2003). Seed-specific overexpression of an endogenous Arabidopsis phytoene synthase gene results in delayed germination and increased levels of carotenoids, chlorophyll, and abscisic acid. *Plant Physiol* 132, 779-785.

Lopez, A.B., Van Eck, J., Conlin, B.J., Paolillo, D.J., O'neill, J., and Li, L. (2008). Effect of the cauliflower Or transgene on carotenoid accumulation and chromoplast formation in transgenic potato tubers. *J Exp Bot* 59, 213-223.

Lu, S., Van Eck, J., Zhou, X., Lopez, A.B., O'halloran, D.M., Cosman, K.M., Conlin, B.J., Paolillo, D.J., Garvin, D.F., Vrebalov, J., Kochian, L.V., Kupper, H., Earle, E.D., Cao, J., and Li, L. (2006). The cauliflower Or gene encodes a DnaJ cysteine-rich domain-containing protein that mediates high levels of beta-carotene accumulation. *Plant Cell* 18, 3594-3605.

Luan, Y., Fu, X., Lu, P., Grierson, D., and Xu, C. (2020). Molecular Mechanisms Determining the Differential Accumulation of Carotenoids in Plant Species and Varieties. *Critical Reviews in Plant Sciences* 39, 125-139.

Maass, D., Arango, J., Wust, F., Beyer, P., and Welsch, R. (2009). Carotenoid crystal formation in Arabidopsis and carrot roots caused by increased phytoene synthase protein levels. *PLoS One* 4**,** e6373.

Mao, Y., Botella, J.R., Liu, Y., and Zhu, J.-K. (2019). Gene editing in plants: progress and challenges. *National Science Review* 6**,** 421-437.

Martinez, M., Cambra, I., Gonzalez-Melendi, P., Santamaria, M.E., and Diaz, I. (2012). C1A cysteine-proteases and their inhibitors in plants. *Physiol Plant* 145**,** 85-94.

Ming, N., Ma, N., Jiao, B., Lv, W., and Meng, Q. (2019). Genome Wide Identification of C2H2-Type Zinc Finger Proteins of Tomato and Expression Analysis Under Different Abiotic Stresses. *Plant Molecular Biology Reporter* 38**,** 75-94.

Nisar, N., Li, L., Lu, S., Khin, N.C., and Pogson, B.J. (2015). Carotenoid metabolism in plants. *Mol Plant* 8**,** 68-82.

Oleszkiewicz, T., Klimek-Chodacka, M., Milewska-Hendel, A., Zubko, M., Stroz, D., Kurczynska, E., Boba, A., Szopa, J., and Baranski, R. (2018). Unique chromoplast organisation and carotenoid gene expression in carotenoid-rich carrot callus. *Planta* 248**,** 1455-1471.

Park, S., Kim, H.S., Jung, Y.J., Kim, S.H., Ji, C.Y., Wang, Z., Jeong, J.C., Lee, H.S., Lee, S.Y., and Kwak, S.S. (2016). Orange protein has a role in phytoene synthase stabilization in sweetpotato. *Sci Rep* 6**,** 33563.

Pitzschke, A., and Hirt, H. (2010). New insights into an old story: Agrobacterium-induced tumour formation in plants by plant transformation. *EMBO J* 29**,** 1021-1032.

Sayre, R., Beeching, J.R., Cahoon, E.B., Egesi, C., Fauquet, C., Fellman, J., Fregene, M., Gruissem, W., Mallowa, S., Manary, M., Maziya-Dixon, B., Mbanaso, A., Schachtman, D.P., Siritunga, D., Taylor, N., Vanderschuren, H., and Zhang, P. (2011). The BioCassava plus program: biofortification of cassava for sub-Saharan Africa. *Annu Rev Plant Biol* 62**,** 251-272.

Shewmaker, C.K., Sheehy, J.A., Daley, M., Colburn, S., and Ke, D.Y. (2002). Seed-specific overexpression of phytoene synthase increase in carotenoids and other metabolic effects. *The Plant Journal* 20**,** 401-412.

Sun, T., Yuan, H., Cao, H., Yazdani, M., Tadmor, Y., and Li, L. (2018). Carotenoid Metabolism in Plants: The Role of Plastids. *Mol Plant* 11**,** 58-74.

Sun, T., Yuan, H., Chen, C., Kadirjan-Kalbach, D.K., Mazourek, M., Osteryoung, K.W., and Li, L. (2020). ORHis, a Natural Variant of OR, Specifically Interacts with Plastid Division Factor ARC3 to

Regulate Chromoplast Number and Carotenoid Accumulation. *Molecular Plant* 13**,** 864-878.

Takeshima, H. (2010). Prospects for Development of Genetically Modified Cassava in SubSaharan Africa. *AgBioForum* 13**,** 63-75.

Trujillo, M. (2018). News from the PUB: plant U-box type E3 ubiquitin ligases. *J Exp Bot* 69**,** 371-384.

Ufuan Achidi, A., Ajayi, O.A., Bokanga, M., and Maziya-Dixon, B. (2005). The Use of Cassava Leaves as Food in Africa. *Ecology of Food and Nutrition* 44**,** 423-435.

Wang, W., Feng, B., Xiao, J., Xia, Z., Zhou, X., Li, P., Zhang, W., Wang, Y., Moller, B.L., Zhang, P., Luo, M.C., Xiao, G., Liu, J., Yang, J., Chen, S., Rabinowicz, P.D., Chen, X., Zhang, H.B., Ceballos, H., Lou, Q., Zou, M., Carvalho, L.J., Zeng, C., Xia, J., Sun, S., Fu, Y., Wang, H., Lu, C., Ruan, M., Zhou, S., Wu, Z., Liu, H., Kannangara, R.M., Jorgensen, K., Neale, R.L., Bonde, M., Heinz, N., Zhu, W., Wang, S., Zhang, Y., Pan, K., Wen, M., Ma, P.A., Li, Z., Hu, M., Liao, W., Hu, W., Zhang, S., Pei, J., Guo, A., Guo, J., Zhang, J., Zhang, Z., Ye, J., Ou, W., Ma, Y., Liu, X., Tallon, L.J., Galens, K., Ott, S., Huang, J., Xue, J., An, F., Yao, Q., Lu, X., Fregene, M., Lopez-Lavalle, L.A., Wu, J., You, F.M., Chen, M., Hu, S., Wu, G., Zhong, S., Ling, P., Chen, Y., Wang, Q., Liu, G., Liu, B., Li, K., and Peng, M. (2014). Cassava genome from a wild ancestor to cultivated varieties. *Nat Commun* 5**,** 5110.

Welsch, R., Arango, J., Bar, C., Salazar, B., Al-Babili, S., Beltran, 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. *Plant Cell* 22**,** 3348-3356.

Woo, J.W., Kim, J., Kwon, S.I., Corvalan, C., Cho, S.W., Kim, H., Kim, S.G., Kim, S.T., Choe, S., and Kim, J.S. (2015). DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. *Nat Biotechnol* 33**,** 1162-1164.

Wurtzel, E.T., Cuttriss, A., and Vallabhaneni, R. (2012). Maize provitamin a carotenoids, current resources, and future metabolic engineering challenges. *Front Plant Sci* 3**,** 29.

Xiao, L., Cao, S., Shang, X., Xie, X., Zeng, W., Lu, L., Kong, Q., and Yan, H. (2021). Metabolomic and transcriptomic profiling reveals distinct nutritional properties of cassavas with different flesh colors. *Food Chemistry: Molecular Sciences* 2.

Zhang, D., Hussain, A., Manghwar, H., Xie, K., Xie, S., Zhao, S., Larkin, R.M., Qing, P., Jin, S., and Ding, F. (2020). Genome editing with

the CRISPR-Cas system: an art, ethics and global regulatory perspective. *Plant Biotechnol J* 18**,** 1651-1669.

Zhou, X., Welsch, R., Yang, Y., Alvarez, D., Riediger, M., Yuan, H., Fish, T., Liu, J., Thannhauser, T.W., and Li, L. (2015). Arabidopsis OR proteins are the major posttranscriptional regulators of phytoene synthase in controlling carotenoid biosynthesis. *Proc Natl Acad Sci U S A* 112**,** 3558-3563.

Ziemienowicz, A. (2014). Agrobacterium-mediated plant transformation: Factors, applications and recent advances. *Biocatalysis and Agricultural Biotechnology* 3**,** 95-102.

Popular science summary

The Agenda for Sustainable Development adopted by the United Nations listed 17 sustainable development goals (SDGs) to be achieved by the year 2030. Goal 2 listed as "Zero Hunger" aims to end hunger and starvation, ensure food security, and improve nutrition while promoting sustainable agriculture. However, food insecurity and malnutrition continue to be a major problem globally, more so in developing nations of the world. Food insecurity is a threat to nutrition intake and pose consequences that could lead to various forms of malnutrition. Since food security alone is not enough to ensure nutrition and prevent malnutrition the importance of production of foods that meets the nutritional needs of people cannot be overemphasized.

Vitamin A deficiency (VAD) is a form of malnutrition resulting from inadequate intake of vitamin A in the diet and could result from relying on staple crops that have little to no provitamin A content as the main source of food. Vitamin A is important for vision and proper function of the immune system; therefore, someone with VAD is at risk of visual impairment and prone to infections because the immune system is not able to fight off infections. VAD is prevalent in sub-Saharan Africa (SSA) and especially in low-income countries where people cannot afford nutritionally rich food, so they rely on staple crops that are deficient in provitamin A.

Cassava is one of the most important staple crops in the world and is widely grown for its starchy roots and the ability to give reasonable yield even with minimal input and in harsh weather conditions. As a result, cassava is popular among small-scale farmers who rely on it for food and income generation. However, commercial cassava cultivars are deficient in provitamin A; therefore, making it the main source of dietary calorie intake predisposes to VAD. Unfortunately, a lot of people do rely of cassava as their main food source.

Biofortification of cassava either by conventional breeding or by applications of biotechnology are attempts to increase the production and accumulation of provitamin A in cassava and consequently the intake of vitamin A in the diet and fight VAD. Unfortunately, cassava has a lengthy cropping cycle which makes increasing its provitamin A production and accumulation through conventional breeding a time-consuming venture. Biofortification by genetic modification can shorten the time it takes to produce cassava varieties that have high provitamin A content. Another problem is that many have observed that increasing provitamin A content in cassava causes a reduction in dry matter and starch content which are commercially important traits for farmers and industry. Since high dry matter and starch content influences the market value of cassava, varieties with low concentrations may not be readily accepted by farmers which may defeat the biofortification purpose.

In this study, we analyzed the dry matter, starch, and provitamin A content of 13 cassava landraces with root color ranging from white to pale-yellow and deep yellow. Our results showed that provitamin A content is higher in cassava landraces with yellow roots than those with white roots. We also discovered that some cassava landraces that have high provitamin A content also have relatively high dry matter and starch content. This confirms that high provitamin A content does not necessarily affect dry matter and starch content negatively in some cultivars.

Provitamin A is a type of carotenoid. Therefore, we also analyzed the expression of genes that are related to carotenoid biosynthesis and accumulation and that are thought to influence the level of carotenoids in plants. Our result showed that both white and yellow landraces have similar expression of these genes, and thus do not seem to influence the levels; however, some genes that promote break down of carotenoids were more highly expressed in white landraces. Still, through a global gene expression study we could find types of genes which previously have not been identified in carotenoid production and storage, which might increase the provitamin A content. These should be studied further. Also, we employed a gene editing technique, CRISPR-Cas9 or the so called “gene scissor”, to increase the carotenoid content of a cassava variety by inhibiting genes that breaks down provitamin A. The result of this study showed that the approaches for gene editing in cassava needs to be refined.

Finally, in light of the European Union Court of Justice stand on products derived from new breeding techniques, we researched how regulatory

frameworks and policies may influence research and development collaboration and the acceptance of new breeding technologies in sub-Saharan Africa. Using, Kenya, Nigeria, Uganda as case studies we determined that rather than regulatory frameworks and policies, funding and economics are more likely to affect research and development collaborations.

Populärvetenskaplig sammanfattning

FNs ”Utvecklingsagenda och globala mål för hållbar utveckling” listar 17 mål som ska uppnås till år 2030. Mål 2, "Ingen hunger", syftar till att stoppa hunger och svält, säkerställa livsmedelsförsörjningen, förbättra kostintag och samtidigt främja hållbart jordbruk. Men matförsörjning och undernäring är fortsatt stora problem globalt, särskilt i världens utvecklingsländer. Brister i matförsörjningen äventyrar nödvändigt näringsintag och riskerar att leda till olika former av undernäring. Men tillräckligt med mat motverkar inte i sig undernäring och man kan inte nog betona vikten av en livsmedelsproduktion som även möjliggör en varierad och näringsrik kost.

Vitamin A-brist är ofta ett resultat av ett stort beroende av basgrödor med litet eller inget provitamin A-innehåll. Vitamin A är viktigt för synen och immunförsvarets funktion; därför riskerar någon med vitamin A-brist att drabbas av synnedsättning och är mer benägen att få infektioner. Vitamin A-brist är särskilt utbrett i Afrika söder om Sahara och i låginkomstländer där människor inte har råd med näringsrik mat och därför måste förlita sig på basgrödor med lågt provitamin A.

Kassava är en av de viktigaste basgrödorna i världen och odlas allmänt för sina stärkelsehaltiga rötter och förmågan att ge bra avkastning, även med minimal insats och under dåliga odlingsförhållanden. Som ett resultat är kassavan populär bland småskaliga bönder som är beroende av den för mat och inkomst. Kommersiella kassavasorter har dock brist på provitamin A; därför är det riskabelt att ha den som huvudsaklig källa för kaloriintag. Tyvärr är det dock ändå många som har kassava som huvudsaklig näringskälla.

Biofortifiering av kassava, antingen genom konventionell växtförädling eller genom tillämpad bioteknik, är ett sätt att öka produktionen och ackumuleringen av provitamin A kans således förhindra vitamin A-brist.

Tyvärr har kassava en lång odlingscykel vilket gör att det är en tidskrävande insats att öka halten av provitamin A genom konventionell förädling. Biofortifiering genom genetisk modifiering kan korta tiden det tar att producera kassavavarianter som har hög provitamin A-halt. Ett problem är dock enligt vissa observationer att ökad provitamin A-halt i kassava orsakar en minskning av torrsubstans och stärkelsehalten, vilket är kommersiellt viktiga egenskaper för jordbrukare och industri. Eftersom högt torrsubstans- och stärkelseinnehåll påverkar kassavas marknadsvärde, är det inte säkert att sorter med låg halt accepteras av lantbrukare, vilket kan motverka syftet med biofortifiering.

I denna studie analyserade vi torrsubstansen, stärkelsehalten och provitamin A-innehållet i 13 lokala kassavasorter vars rotfärger varierade från vit, via blekgul, till djupgul. Våra resultat visade att provitamin A-halten är högre i kassavasorter med gula rötter än de med vita rötter. Vi upptäckte också att vissa lokala sorter av kassava med hög provitamin A-halt också har relativt hög torrsubstans- och stärkelsehalt. Detta bekräftar att hög provitamin A-halt inte nödvändigtvis påverkar torrsubstans och stärkelsehalt negativt i vissa sorter.

Provitamin A är en typ av karotenoid. Därför analyserade vi också uttrycket av gener som är relaterade till karotenoidbiosyntesen och –ackumuleringen, och som tros påverka mängden av karotenoider i växter. Vårt resultat visade att både vita och gula sorter har liknande uttryck av dessa gener och alltså inte verkar inverka på nivåerna; dock var uttrycket av vissa gener som främjar nedbrytning av karotenoider högre i vita sorter. Men med en global genuttrycksstudie kunde vi hitta typer av gener som tidigare inte kopplats till karotenoidproduktion eller -lagring, och som kanske kan öka innehållet av provitamin A. Dessa bör studeras närmre. Vi använde också en genredigeringsteknik, CRISPR-Cas9 ibland kallad "gensaxen", för att öka karotenoidhalten i en kassavavariant genom att hämma gener som bryter ner provitamin A. Resultatet av denna studie visade att metoderna för genredigering i kassava behöver förfinas.

Slutligen, i ljuset av EU-domstolens hållning till produkter som härrör från växtbioteknik, undersökte vi hur regelverk och policyer påverkar forsknings- och utvecklingssamarbeten samt acceptansen av ny bioteknik i Afrika söder om Sahara. Genom att använda Kenya, Nigeria, Uganda som exempelländer fastställde vi att det snarare är finansiering och ekonomi som inverkar på forsknings- och utvecklingssamarbeten än regelverk och policyer.

Acknowledgements

This project would not have been possible without the support of many people.

I want to specially thank my supervisor, Dr. Erik Alexandersson for your tremendous support, patience, and guidance through the years. Thank you for your help and for always being one WhatsApp message away. Thank you for supporting me even when things did not go as expected.

I express my sincere gratitude to my co-supervisor Marit Lenman who offered guidance and support. Many thanks for your efforts and help with the gene editing and transformation part of this project. I sincerely appreciate your help.

My sincere gratitude to my co-supervisors Dr. Livia Stavolone and Dr. Andreas Gisel. Thank you for your support and guidance throughout this project. I thank you especially for your efforts on paper II. Dr. Livia, thank you for always pushing me. Finally thank you for always being available to help.

Many thanks to my co-supervisor Dr. Leena Tripathi for the guidance and support rendered to this project in Nairobi.

I also want to a big thank you to members of the Resistance Biology research unit: Erik Andreasson, Erik Alexandersson, Marit, Svante, Erland, Salla, Linnea, Awais, Nam, Sajeevan, Oliyad, Naga, Rizwan, Milla, for creating a friendly working environment.

Many thanks to the co-authors for their contributions to the paper and project. I thank all who have contributed to this project in one way or the other. I

especially thank IITA cassava breeding unit under the leadership of Mr peter Iluebbey for the help and support with the field studies.

I specially thank my colleagues at IITA bioscience centre Ibadan. Many thanks to my colleagues who became good friends: Seun, Chisom, Joseph, Ifeanyi, Anetor, and Damilola. Thank you all for your friendship and encouragement.

Many thanks to Beena, Valentine, and Sam for the laboratory support in Nairobi.

I thank my friend Tope Jekayinoluwa. Thank you for always being available to talk, even at odd hours. I am glad we met.

Many thanks to my friend Lekan Jolayemi, for your positive attitude and disposition. Thank you for your words of encouragement and for always being available to listen. Thank you for allowing me to trouble you and for not complaining.

Finally, I say a big thank you to my wonderful family for the love and support you have shown me during this journey. Thank you for the many words of encouragement and for always checking on me and for always being there despite the distance in the past few years.

Acta Universitatis agriculturae Sueciae
Doctoral Thesis No. 2022:21

Carotenoids are indispensable to plants and are critical in human diets and health. The studies in this thesis were designed to improve the understanding of carotenoid biosynthesis and accumulation in cassava storage roots. Transcriptome profiling and metabolite analysis of cassava landraces identified genes and metabolites that are associated with carotenoid accumulation in cassava roots. It is important to study how these genes can be used to modulate carotenoid accumulation in cassava roots. Interviews and comparison of regulatory documents revealed that EU regulatory framework and policy on GMOs and new breeding techniques do not clearly influence African R&D collaborations

Priscilla Olukola Olayide received her doctoral education at the Department of Plant Protection Biology, SLU, Alnarp; and obtained her MSc in Biomedical Sciences and BSc in Molecular Biology at Liberty University, Virginia, The United States of America.

Acta Universitatis agriculturae Sueciae presents doctoral theses from the Swedish University of Agricultural Sciences (SLU).

SLU generates knowledge for the sustainable use of biological natural resources. Research, education, extension, as well as environmental monitoring and assessment are used to achieve this goal.

Online publication of thesis summary: https://pub.epsilon.slu.se

ISSN 1652-6880
ISBN (print version) 978-91-7760-917-9
ISBN (electronic version) 978-91-7760-918-6