LUND UNIVERSITY

MASTER THESIS

Biofortification of Cassava

Author:
Linnea ALMQVIST

Examinator:
Leif BÜLOW
Head Supervisor:
Nick SIRIJOVSKI
Supervisor:
Erik ALEXANDERSSON

A thesis submitted in fulfillment of the requirements for the degree of Master of Science in Biotechnology

in the

Division of Pure and Applied Biochemistry
in cooperation with the
Swedish University of Agriculture

June 7, 2018

Declaration of Authorship

I, Linnea ALMQVIST, declare that this Master thesis titled, "Biofortification of Cassava" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a master degree at this University.
- Where I have consulted the published work of others, this is clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.

Signed: _____

Date: _____

Abstract

Cassava, *Manihot esculenta*, is the third most important staple crop in tropical regions. Cassava is a fibrous root that is tolerant to drought and has a flexible harvest period making it a reliable crop under poorer farming conditions (FAO, 2008). The content of provitamin A, also known as β-carotene, is unfortunately low in cassava even though there are varieties bred for a higher content. Since vitamin A deficiency, VAD, is a large problem of the regions where cassava is of major importance for calorie intake, an increase of the β-carotene content would be of great interest to address malnutrition.

The main idea of this project is to increase the amount of β-carotene accumulated in the cassava root. This is accomplished by using the knowledge of the cassava genome and the carotenoid pathway, in order to block specific enzymes through genome editing techniques . The trials will though take many years and this Master Thesis will cover only a fragment of the entire project. Cassava needs to grow in the field for at least nine months, which will create a long time frame to consider. After creating genetic constructs and transforming the cassava plants, the carotenoid and starch content will be analysed.

Methods for analysing starch and carotenoids for cassava were developed during this Master Thesis by literature research and laboratory trials.

Popular Science Abstract

At the Swedish University of Agricultural Sciences, SLU, in Alnarp, Erik Alexandersson is leading a project that focuses on biofortification of cassava. Cassava is a staple crop grown in tropical regions and serves as the third most important source of calorie intake in these regions (FAO, 2008). Cassava is a root tolerant to drought and can be harvested when needed, which makes it a favorable crop for impoverished families. Cassava can be cooked in several different ways, but one in particular popular method is to dry it and mill to a flour that can be stored and later used for porridge, known as garri. One major issue with cassava, however, is that it produces little β-carotene naturally. β-carotene is active as provitamin A, which is essential in our diets. If the main calorie intake consists of cassava, the lack of β-carotene in the root can cause vitamin A deficiency, leading to blindness and death.

In this project, the β-carotene and starch content in cassava was evaluated in order to understand the genes controlling the biosynthesis. The overall project goal is to enhance the β-carotene content with retained starch levels in cassava (Alexandersson, 2017).

CRISPR/Cas9 is a powerful tool to edit plant genomes which can create new valuable crops. The gene editing is stable and heritable. The method is also user-friendly and not hard-line considered as GMO, which would ease public acceptance. For a growing population and a harsher environment, the method holds promise to be revolutionary (Khatodia *et al*., 2016). In this project, CRISPR/Cas9 was used to deactivate enzymes in the carotenoid pathway hopefully leading to β-carotene accumulation in the mature root.

The carotenoid content varies in plants due to for example species, genotype, ripening stage and environmental conditions. Previously genetic and metabolic engineering has been successful to enhance the carotenoid levels in different plants, for example in the famous so called Golden rice. The carotenoid content can be analysed using high performance liquid chromatography, HPLC, in biological samples. The long conjugated double-bonds will absorb light in at a specific wavelength and the content of specific carotenoids like β-carotene is possible to measure (Saini *et al*., 2015). The starch content can be analysed by using enzymatic methods that will hydrolyse the starch to glucose, which further can be measured by using spectrophotometry.

In this thesis, the β-carotene and starch content was measured for 13 wild cassava genotypes in order to present a recommendation on which of these genotypes that produced the most favorable levels. These recommended genotypes could further be used as targets for the genetic modifications, creating a more revolutionary cassava variety than ever before.

Acknowledgements

I would like to thank my supervisor, Erik Alexandersson, for supporting my work with great passion and knowledge. Erik has shown much enthusiasm and have been presenting countless ideas for new takes and chapters of my thesis. Priscilla Olayide and Livia Stavolone, members of the Biofortification of Cassava project at SLU, have provided me with information, material and feedback of my results and for that I would like to thank them.

I am also grateful to the institution, the Swedish University of Agriculture, for providing me with tools, advice, field expertise and coffee. A special thanks to Karl-Erik Gustavsson, Niklas Olsson, Mia Mogren and Helle Turesson for sharing their competence.

For my head supervisor, Nick Sirijovski, for trusting that I can manage on my own, but still staying only an e-mail away.

My examinator, Leif Bülow, who got me interested in research, genetics and biochemistry in the first place.

Preface

During the first week of this twenty week Master Thesis, in January 2018, the European Court of Justice decided that CRISPR/Cas9 should not be considered as Genetic Modification, GM (Nature, 2018). This relaxation of the rules offers enormous benefits for plant fortifications which is relevant in this thesis, making the subject even more thrilling.

With this CRISPR Master Tool waiting to be used to its full potential to reduce the suffering and starvation in the world, the future appears bright and exciting. I could not have chosen a more accurate and up-to-date subject to write about.
Special thanks to Erik Alexandersson for making this thesis possible and to SLU Alnarp for hosting me.

Without further ado, let us immerse in the biofortification of cassava.

Contents

1 Introduction

The main goal with this Master Thesis is for me to, personally, better understand and gain insight in food and agricultural research work on an applied chemistry level. To accomplish this I assisted members of the project with suitable tasks. Most of my work was conducted at SLU Alnarp, in the laboratory, with data analysis and with computer research work.

Cassava is the third most important staple crop in tropic regions (FAO, 2008). It is a starchy root that can be cooked and consumed in many ways. The problem is that many cassava cultivars do not produce or store accurate amounts of β-carotene, causing tragic human suffering. This is of importance primarily when the diet is not diverse enough and cassava is the main calorie intake. β-carotene acts as a provitamin for vitamin A. Provitamin A can be transformed in our bodies to the active form, vitamin A, which our bodies cannot synthesise nor live without. The effects of vitamin A deficiency, VAD, is blindness, weakened immune system and death.

Many trials have been done by crossing and manipulating cassava into producing more carotenoids. For these trials to have greater effect the genetic pathways of the carotene needs to be fully understood.

The focus of this Master Thesis have been on biofortification of cassava through genome editing. Understanding of the genetic code with the enormous amount of information and data we have today is the bottom line. By analysing the genome, a gene target can be identified, evaluated and edited. To analyse the affect of gene editing on the modifications in the crop, a harvesting time of a minimum around 8 months from sowing is needed. Additional harvesting after 12 and 14 months is also needed to achieve a complete picture of the expression.

This thesis work extends over a 20 week period and therefore the final results of gene editing will not be presented. Instead focus will be on developing methods that can be used to both develop new transformations of genes of interest, and analysing carotenes and starch in the harvested roots. Finally the suggested cassava genotype to use in future gene editing in the project will be presented.

2 Theory

2.1 Genetics and Biochemistry

2.1.1 Cassava, *Manihot esculenta*

As a staple crop cassava is popular because of its reliability and high starch content. Its nutritional value is pretty low but the high amount of calories from carbohydrates gives it an important role in fighting hunger in tropical regions (Ceballos *et al.*, 2017). In Figure 2.1 a full grown cassava is shown and the white inside of the root is visible. The flesh colour can vary from yellowish and pinkish to white and is related to the carotenoid content. The colour of the flesh does not notably affect flavour or cooking properties of the root.

FIGURE 2.1: Cassava root, *Manihot esculenta*. Source: Farmaciaorganica (2018).

Cassava can be obtained in Sweden in several vegetable retail stores but is not common to consume traditionally and the prevalence is probably due to the cultural diversity of the Swedish population.

2.1.2 Carotenoids

Carotenoids are structurally related compounds found in many vegetables. They are produced by the plant for several reasons. They help protect the plant from photo-oxidation, take part in the photosynthesis process and give pigmentation. They are also precursors for regulatory systems in for example the synthesis of abscisic acid (Ceballos *et al.*, 2017). Carotenoids are produced in almost all the plasmids in the plant cells and stored in chloroplasts or chromoplasts (Yuan *et al.*, 2015). In Figure 2.2 a map of the plant carotenoid pathway is shown.

2.1.3 Vitamin A

Vitamins are substances required by our bodies to function but the body itself cannot synthesise them *de novo*. In more than half of the countries of the world, mostly

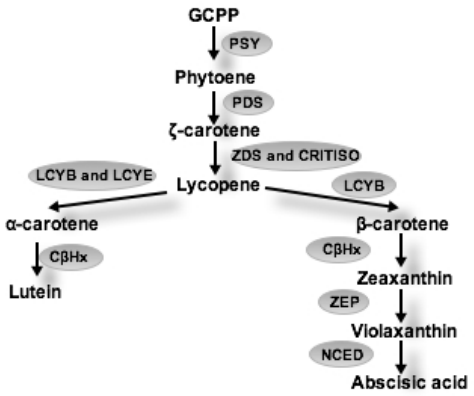

FIGURE 2.2: Schematic diagram showing general carotenoid biosynthesis in plants. The ellipses are presenting the enzymes involved in the biosynthesis.

in Africa and Southeast Asia, vitamin A deficiency is considered a public health problem (Ceballos *et al.*, 2017). Many of the carotenoid compounds have provitamin A activity, but the highest activity is found in the configuration all trans β-carotene, explaining why this specific metabolite is of great importance (Saini *et al.*, 2015). The project's aim is to increase the synthetisation of β-carotene by altering the function of the enzymes involved in the carotenoid pathway.

2.1.4 Expression of genes in the carotenoid pathways

The gene expression in the carotenoid pathway can be analyzed using available tools. The genome expression of cassava has been published and is available online under the name Bart Lab Cassava Atlas. This data can be used to, for example, compare the different amount of gene expression in different biological parts of the plant and further give hints on how the genotype corresponds to the phenotypic expression (Wilson *et al.*, 2017). It must be taken into consideration that the cassava plants analysed to create the database was not grown under natural conditions but in pots. Another important tool in bioinformatics is an online database called Plaza. Plaza compares plant genomics by compiling different genome sequencing projects and initiatives. 55 plant species are available, cassava included. Genomes of these species can be searched by function or homologs and compared through powerful and user friendly tools (PLAZA, 2018). Proteins can be compared between species and further conclusions about their different characteristics can be drawn.

When the genes in the pathway are understood and mapped they can be turned off by using CRISPR/Cas9 techniques and further the biosynthesis and accumulation of carotenoids can be altered. This theory will be explained in later sections.

Phytoene synthase, PSY, for example, is an important enzyme in the carotenoid pathway. It converts GCPP to phytoene as the name reveals. This can be seen in Figure 2.2 above. The prevalence of this enzyme affects the amount of carotenoids synthesized, meaning that the expression of this enzyme is important for the accumulation of carotenoids (Ceballos *et al.*, 2017). This gene was the target in the well known Golden rice project, increasing the β-carotene content drastically in rice by increasing the amount of PSY expressed (Yuan *et al.*, 2015).

This project's aim is not to increase the amount of enzyme, but rather down regulate target enzymes. This method is an advantage since it is easier to cause a loss

of function than a gain or an increase using CRISPR/Cas9. Removing or down regulating genes interferes less with regulations than adding or increasing them. The target enzymes for this project will be: β-carotene hydroxylase, CHY, 9-cis epoxy carotenoid dioxygenase, NCED and lycopene ϵ-cyclase, LCYE.

- CHY is also named CβHx as in Figure 2.2 in section 2.1.3, and is responsible for the conversion of β-carotene to zeaxhanthin. By blocking this way, it can be assumed that β-carotene will accumulate.
- NCED regulates the conversion to abscisic acid at the end of the pathway.
- LCYE catalyses the conversion of lycopene to α-carotene, a branch that is desired to block.

2.1.5 The Orange gene

Another gene of interest is called Orange, or short Or. This gene controls the production of a protein also named orange. This protein is not directly linked to the carotenoid synthesis, but rather to the stability of the carotenoids produced. More precisely it seems to be stabilising phytoene synthase, the enzyme in the beginning of the carotenoid pathway. A linkage of the Or protein and induced chromoplast synthesis is also a published explanation of the accumulation (Yuan *et al*., 2015). Yuan *et al*. (2015) also showed that in *Arabidopsis thaliana* and *Sorghum bicolor*, the carotenoid levels rose from a single amino acid expression by affecting the accumulation. This phenomenon was first observed by a spontaneous mutation resulting in orange cauliflower that was caused by a high accumulation of β-carotene by affecting the orange gene (Crisp *et al*., 1975). By using overexpression vectors for the orange gene in sweet potato, a significant increase of the β-carotene content has been accomplished (Kim *et al*., 2013).

It has also been shown in *Arabidopsis thaliana* that a single amino acid substitution in the wild type Or gene results in a carotenoid accumulation that is significantly higher without affecting the growth of the plant (Yuan *et al*., 2015). This shows that it is not only the amount of orange protein synthesized but also the structure of the protein itself that is of importance for the carotenoid accumulation. The Or gene in Arabidopsis contains a highly conserved arginine at position 108, and when that single amino acid is substituted to either histidine or alanine, the function drastically changes. Histidine is commonly found in binding or active sites and further the creation of a new active site, or a stronger one, can be the structural explanation of the effect. Alanine may have a similar explanation (Yuan *et al*., 2015). These interactions are preferably studied in 3D-structures of the protein and its active sites, unfortunately this protein is not modelled for plants yet but the future holds promises for new discoveries.

Cho *et al*. demonstrated in 2016 that there is a connection between the Or gene and drought tolerance in plants. Since carotenoids acts as antioxidants, the prevalence of them can protect the plant from abiotic stress. This fact may be part of the explanation of the high preservation of the Or gene in different plants.

2.1.6 Correlation between starch and β-carotene

Research is pointing at a negative correlation between starch and β-carotene (Giuliano, 2017). The reason why this correlation occur is not proven, but two theories can be mentioned. There can be a genetic or a plant physiology linkage between

these two metabolites. Both β-carotene and starch are synthesised inside plastids, leading to a competition for the same organells being one explanation (Cervantes-Flores *et al.*, 2011). If the linkage is genetic, Orange gene modification might not have the same effect as altering the carotenoid pathway.

2.1.7 Plant Breeding

Before the target gene to be altered is chosen, a suitable cassava variety needs to be selected. For commercial crops there are often hundreds of varieties to choose from. When the aim of plant breeding is biofortification a few things needs to be kept in mind. You cannot breed for too many different qualities, but instead pick a few points of great importance and interest. If a plant variety existed that could fulfill all our needs, the world would be perfect. Unfortunately, there is no such plant. As Professor Malachy Akoroda (2018) lectures, "No crop stands alone". We will always need a variation in our diet. Since the final aim is to increase the β-carotene content, the plant variety chosen for the experiments would possibly have a naturally high β-carotene content but not missing other important traits like tolerance for drought and diseases, and keep a high starch content as mentioned earlier. The variety also needs to be suitable for transformation. Since this has been known for a long time, resources and efforts have been put in to develop stocks with a naturally high β-carotene expression. There are varieties with a β-carotene content of as much as 25 μg/g (fresh weight basis) cassava, these have a characteristic yellow root pigmentation (Esuma, 2016).

The main use of the crop also needs to be considered. Since 70 % of the cassava harvested in Nigeria is used for the production of garri (Labuschagne and Akoroda, 2018), a flour that is fried, the garri traits needs to be evaluated and cannot be neglected while breeding.

2.1.8 Genome mapping

When the target genes are decided, bioinformatic tools like plant genome databases and computer programs are used to map the genes and create primers for these to make it possible to amplify them in lab. A primer is an oligonucleotide consisting of around 20 nucleotides that can be designed to anneal prior to the translated protein sequence. By using primers the target genes can be amplified and a construct is created that can further be transferred to the organism.

2.1.9 Construct

In order to genetically modify the genome of cassava the exact genomic construct needs to be determined. The carotenoid pathway is known and presented in Figure 2.2. In genome editing, it is much easier to down regulate a gene than to up regulate it, basically since it is easier to inhibit a genetic sequence from working than programming the gene to produce more copies of proteins. At least for today this is true. To eliminate a gene and further down regulate it, one of the bases can simply be switched creating a non working copy. By looking at the carotenoid pathway,2 several options for genes to down regulate can be discussed. It seems logical to down regulate the metabolite right after β-carotene, CHY1, or block the left path from taking place by removing the LCYE enzyme. When the targeted bases are decided primers can be designed to be used to create a construct. When the form of the

construct is decided, the CRISPR/Cas9 tool can be used to incorporate the desired changes into the genome.

2.1.10 CRISPR/Cas9

Clustered Regularly Interspaced Short Palindromic Repeats, CRISPR, is a revolutionary technique based on proteins found in the bacterial immune system. It is a tool to accurately edit genomes. The technique can be used to both delete and edit genomic sequences. Cas9 is an endonuclease found in bacteria that contains a short RNA or DNA sequence that it uses as a map of where to cut open the DNA or RNA strands. The altered plants resulting from this technique will not be regarded as GMO but GE, gene editing plants, easing the public acceptance (Khatodia *et al.*, 2016). CRISPR/Cas9 is used in this project to alter the genes in the construct.

When the construct of a gene is created, the gene can be transformed back into the plant cell through for example *Agrobacterium rhizogenes*-mediated infiltration (Arola and Narula, 2017).

CRISPR has been used to, for example, enhance flavour, sugar or vitamin content and to increase climate and disease tolerance in different plants and the technique in constantly refined. The first plant genome directed mutations caused by CRISPR were performed in August 2013 (Arola and Narula, 2017).

2.2 Analysis

2.2.1 Starch measurements

The starch content in cassava is of importance to measure since there is a proven correlation between the starch and carotenoid content (Giuliano, 2017). Also the starch content is one major component for making cassava such an important staple crop, and to loose this trait is unfavourable. Starch is the most common carbohydrate in our diets and provides us with energy. The starch content in a sample can, for example, be measured by using a kit from Megazyme (Bray, Ireland) called Total Starch Assay Kit. This is an enzymatic method based on the enzymatic hydrolysis of starch into glucose, which can be detected by using spectrophotometry. By using this method the resistant starch is not measured and it is also assumed that the natural D-glucose content of the samples is negligible.

The basis of the starch measurement is to initially hydrolyse starch into D-glucose by the use of α-amylase. The hydrolysis is followed by two reactions, shown as Equation 2.1 and 2.2, using the glucose oxidase and peroxidase that is included in the Megazyme GOPOD reagent.

$$D\text{-}glucose + O_2 + H_2O \xrightarrow{\text{glucose oxidase}} D\text{-}gluconate + H_2O_2 \tag{2.1}$$

$$\begin{aligned} &2H_2O_2 + p\text{-}hydroxybenzoicacid + 4aminoantipyrine \\ &\xrightarrow{\text{peroxidase}} quinoneiminedye + 4H_2O \end{aligned} \tag{2.2}$$

In order to further confirm the results from the starch measurements alternative methods for analysis could be used. An iodine test could for example additionally be used to measure the starch levels in the different genotypes. If the starch level

results from the two different methods overlap, the accuracy of the results will be strengthened.

2.2.2 Spectrophotometric analysis of carotenoids

To allow the target compounds, carotenoids, to be released from the cell and available for detection and quantification, extraction needs to be done. In this step hazardous chemicals are sometimes needed and it is of great importance to handle them with care. In order to determine the most effective extraction solvent for making the carotenoids available for detection, spectrophotometric measurements were done and the solvent that gave the highest absorbance was used in further trials. HPLC analysis of β-carotene is more accurate than a spectrophotometric method since it can be confirmed that the compounds detected are β-carotene by looking at the shape of the absorbance spectrum. This is further explained in the result section. But since a spectrophotometric method is both easier, faster and cheaper, the samples were analyzed using spectrophotometry in parallel to the HPLC measurements.

2.2.3 HPLC analysis of carotenoids

High Performance Liquid Chromatography, HPLC, is a method for separating compounds to analyse the composition in a sample and in what concentration the compounds are present. The method is based on the interaction of compounds with the columns that they pass through. Further, the compounds are detected by using their light absorbance properties. β-carotene is an orange chromophore and absorbs light at a wavelength of 450 nm. An absorbance spectrum is also scanned by a Diode Array Detector, DAD, in order to further analyse the peaks. This light spectrum and the peaks can be used to distinguish the different configurations of carotenes that are present, all-trans β-carotene can be separated from its cis-forms for example. The data used for the calculations was the total amount of β-carotene, including both cis- and trans-forms.

2.2.4 Orange gene

The analysis of the orange gene is done by using bioinformatic tools. Online databases such as Plaza and the Protein Data Bank are used to collect data and a program called CLC Genomics Workbench from Qiagen bioinformatics can be used to further study the genes and compare them between different organisms.

2.2.5 Statistical Analysis

The standard deviation, the correlation, the linear regression and the variance were calculated in Excel for the data sets. ANOVA was used to perform a Tukey test in Minitab. Minitab was used to obtain p-values for the regression fit.

3 Materials and Methods

3.1 Cassava samples

The cassava plants used were grown at IITA in Ibadan, Nigeria and harvested after eight months. At eight months, the root is mature and all necessary metabolites are developed. The roots are usually harvested after twelve to fourteen months but can be harvested as late as 24 months from sowing. Thirteen different genotypes were planted. Three biological replicates that gave three technical replicates each, of roots and leaves were collected, cut in small pieces, lyophilised and shipped with silica gel to Alnarp where they were stored refrigerated at 4°C, dry and dark. The genotypes were chosen based on their starch and β-carotene content since there is a negative correlation between these two as shown previously (Giuliano, 2017). The genotypes have varying degrees of both starch and β-carotene.

3.2 Construct

RNA extractions were done at SLU, but since they gave low RNA concentrations, it was decided that this work should be performed at IITA instead due to the access of fresh material.

3.3 Starch measurements

A kit called Total Starch Assay was ordered from the company Megazyme. The protocol included was followed. The method called "a" was used. This method excludes D-glucose rinsing and measurement of resistant starch of the sample. This choice was decided after consulting experienced people at SLU.

For all analysis methods the freeze dried cassava samples was milled for 8 seconds to a powder using an IKA basic mill. Approximately 50 mg of the powder was transferred, and the exact weight recorded, to 2.0 ml eppendorf tubes and 0.1 ml of aqueous ethanol (80% v/v) was added to aid dispersion. After vortexing, 1.5 ml of thermostable α-amylase was added and the tubes were incubated in boiling water for 12 minutes. The samples were then cooled down in room temperature for a minute before 50 μl of amyloglucosidase was added followed by an incubation at 50°C for 30 minutes. After the incubation, the content of the tubes were transferred to a 50 ml polypropylene tube and distilled water was used to adjust the volume to 50 ml. The 50 ml tubes were centrifuged and 50 μl of the supernatant was mixed with 1.5 ml of GOPOD reagent and incubated for 20 minutes at 50°C. A glucose standard was treated the same way. The absorbance was measured at 510 nm.

The starch values were calculated by using the weight and the absorbance value. A calculation guide was included in the kit and this guide was followed. A maize standard that was included in the kit was used as a reference, and its starch value was considered as true by correcting the samples after the true starch standard value of 93 %.

3.4 Extraction

When measuring carotenoid content, great precautions needs to be considered since these substances are light sensitive and will quickly break down or change configurations from UV-light resulting in less accurate results. Further, the samples and powder were always wrapped in foil or stored in dark tubes, and all treatments was done with the light switched off and the windows covered. The lyophilized raw material was milled into a flour using an IKA 10A basic mill, same as for the starch measurements. The powder was divided into 200 mg, $\pm$2 mg replicates from each sample. The extraction solvent consisted of ethyl acetate/ethanol/BHT 80:20:0.1 (v/v). To the samples, 1 ml of the solvent was added followed by vortexing. The samples were then placed in an ultrasonic bath for 10 minutes and incubated at 60°C for 60 minutes. After the incubation, the samples were centrifuged for ten minutes at 3000 rpm and the supernatant was collected. To further increase the carotenoid content in the samples and increase the resolution, they were evaporated and solved in 100 μl in the same solvent as used before. Hence a ten time concentration was obtained. When the extraction was done the samples were stable in dark at 4°C.

The trial method was developed by consulting an experienced technician at SLU. To conclude the method that would be used in the final measurements, test trials were done, which included a different extraction solvent: Hexane/ethanol 4:3 (v/v) with powder concentration of: 100 mg, 200 mg and 500 mg per ml. HPLC was tried on different concentrations of cassava powder in order to obtain the best resolution. 100 and 200 mg powder/ml solvent gave signals too small to be reliable in comparison to the background noise. The concentration finally used was 2.0 g powder/ml solvent. The method used for the final trials was selected by looking at the resolution of the test trials. This is explained in detail in the section 4.2.

3.5 HPLC

An injection volume of 10.0 μl was used for all cassava root samples and the flow was 0.5 ml/min. The column used was a Luna 3 μm Silica 150*3 mm from Phenomenex. The extractions were run for 6 minutes in the isocratic mobile phase n-hexane/isopropanol (96:4) at 27°C. The computer program used was LC/MSD Chemstation from Agilent Technologies. The detection wavelength was 450 nm and the light spectrum used for the DAD included 230-550 nm.

3.6 Spectrophotometry

A spectrophotometer was used to measure the absorbance at 455 nm for two samples of each genotype in order to compare with the results from the HPLC. The same extraction method was used.

3.7 Orange gene

When using bioinformatics to analyze genes there are a number of softwares available that was used. These are all mentioned in the theory section. The species chosen to be used for alignment with cassava was Arabidopsis, *Arabidopsis thaliana*, cauliflower, *Brassica oleracea* and sweet potato, *Ipomoea batatas*. Sweet potato was picked since it is known to have a high β-carotene content, cauliflower since that

was where the Orange gene was found and researched primarily and Arabidopsis is commonly used as a model organism, but was also used since the published amino acid substitution of interest was done in Arabidopsis.

3.8 Statistical analysis

Microsoft Excel was used to compile the data values from the starch and β-carotene analysis. Minitab and its functions was used to calculate the variance by using ANOVA-analysis. A Tukey test was used in order to draw conclusions about the different genotypes and how they differ from each other. The Minitab function Regression fit was used to calculate the p-values for the regression lines in order to prove correlations.

4 Results

Since this Master Thesis was shorter than the SLU cassava project itself, the results can be considered as lab training and developing analysis methods. The results can be used to choose what genotypes of cassava that will be used for the CRISPR/Cas9 genetic modifications since the actual field samples were analysed. It is desired to find a genotype that is naturally high in β-carotene but still has a high starch content. The relationship between the starch content and the carotenoid content mentioned in the theory section of the report was analysed, and will be discussed.

4.1 Starch measurements

Enzymatic measuring methods may not be as accurate as for example HPLC analysis. Since enzymatic methods are based on working proteins there is a risk of loss in effectiveness of the enzymes. The method used also includes more steps and further increases the risk of variation between measurements. The starch content in the cassava samples was measured using the Total Starch Assay Kit from Megazyme and stretched from 50 to 85 % of the dry weight. The results are presented in Figure 4.1.

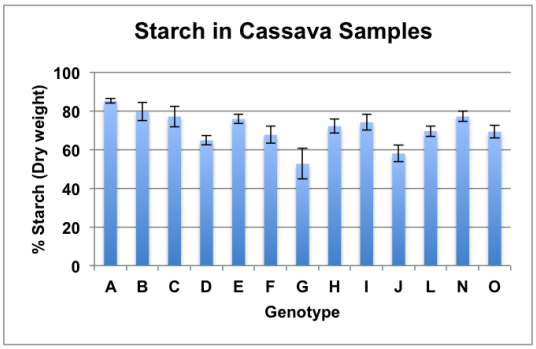

FIGURE 4.1: Starch percentage of dry weight of the different cassava genotypes. Three technical replicates was obtained from three biological replicates of each genotype, giving nine replicates of each genotype. Every replicate was measured once. The error bars represents the sample standard deviation for each genotype.

4.2 Carotenoid measurements

4.2.1 Spectrophotometry

Four different extraction methods was tested on cassava samples to decide which extraction solvent and method that gave the highest absorbance. This method was further assumed to give the best resolution and would be used in the trials with the harvest from the end of February.

The extraction made with ethyl acetate/ethanol/BHT 80:20:0.1 (v/v) as a solvent, followed by a ten minute ultrasonication, an incubation and a centrifugation, clearly gave the highest absorbance and was therefore decided to be used for the HPLC protocol.

Two of the rejected methods used a different extraction solvent: hexane/ethanol 4:3 (v/v) followed by a 30 minute ultrasonication and a centrifugation. In one of the methods, these steps were followed by an over night shaking in 4°C. The third rejected extraction method also used hexane/ethanol as a solvent. After the addition of the extraction solvent, the samples were ultrasonicated for 30 minutes, followed by a 60 minute 60°C incubation. Since all the three methods using hexane/ethanol gave a lower absorbance than the method using ethyl/ethanol/BHT, they were rejected. This was fortunate also, since hexane is a hazardous chemical and should be avoided when possible.

4.2.2 HPLC

The β-carotene content was analyzed by using HPLC DAD. The total area of the detected peaks were used to calculate the concentration in μg β-carotene/g dry weight cassava. A β-carotene standard was used to calculate this conversion. The results are presented in Figure 4.2. The β-carotene content appear to have a larger variation than the starch content.

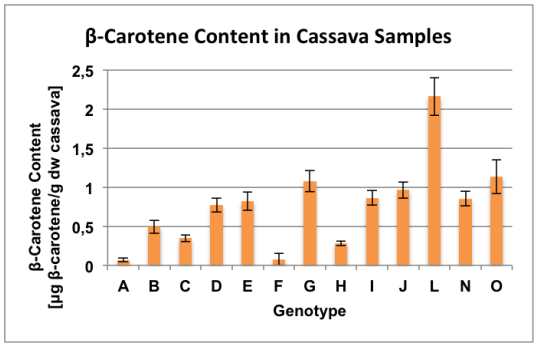

FIGURE 4.2: β-carotene content in the different genotypes presented in μg β-carotene per g dry weight cassava. The error bars represents the sample standard deviation from the nine replicates.

In order to calculate the total area and the total β-carotene concentration, chromatograms from the HPLC measurements are used. A chromatogram of a cassava root is shown in Figure 4.3. The computer program sums up the areas of the two peaks giving a total concentration. Peak I presents all-trans β-carotene and peak II presents the cis configuration forms of β-carotene.

If it is desired to analyse the cis- and trans forms of β-carotene, the HPLC DAD method used offers this as well. When characterising the peaks from the chromatogram in order to decide what compounds they present, a spectrum is used. In Figure 4.4 a spectrum of Peak I from Figure 4.3 is presented to the left, and a spectrum showing an all-trans β-carotene reference to the right. Since the Peak I spectrum on the left well resembles the forms of the reference spectrum on the right, it can be concluded that they are the same substance.

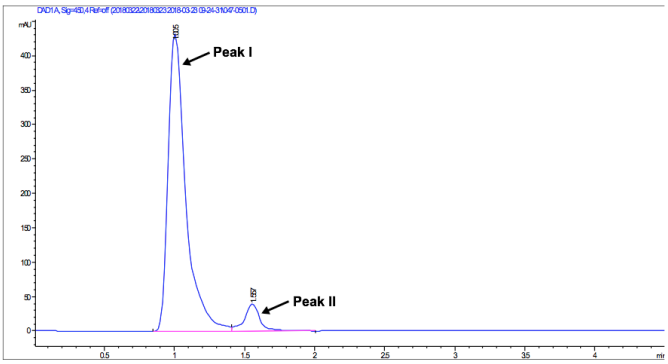

FIGURE 4.3: Chromatogram of cassava root from HPLC analysis. Peak I shows the detection of all-trans β-carotene and Peak II the different cis forms of β-carotene. The area of the peaks represents the quantity of the compounds.

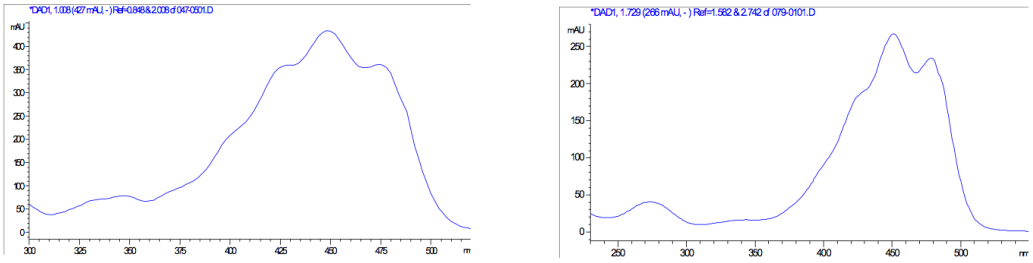

FIGURE 4.4: To the left a spectrum of the all-trans β-carotene Peak I from Figure 4.3 is presented. To the right a spectrum of an all-trans β-carotene reference is shown.

4.3 Combining Starch and β-carotene Content

If the results from the starch and the β-carotene measurements are plotted in the same graph, and sorted after the β-carotene content, a negative correlation is observed in Figure 4.5. If the correlation is calculated using Excel, the result shows a correlation of -0.3893. This number corresponds to a moderate to a low negative correlation.

By looking at the scatterplot in Figure 4.6, a negative correlation can be suspected here as well, although the R-square value is 0.12692, which is far from one, pointing at a low or no correlation. By performing a regression fit in Minitab, a p-value of 0.189 is obtained. Since this number is larger than 0.05, a negative starch and β-carotene correlation cannot be concluded for these genotypes.

4.4 Spectrophotometric method compared with HPLC method

The results of the absorbance of the extractions can be compared with the results from the HPLC. The HPLC is more accurate, but since a spectrophotometric method is faster, cheaper and easier, it can be a method to choose if the results are close enough to the HPLC results. In Figure 4.7, a scatter plot is presented that correlates

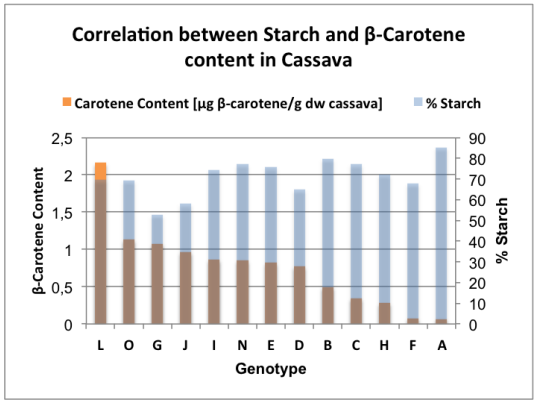

FIGURE 4.5: β-carotene and starch content for the genotypes plotted in the same graph sorted after β-carotene content.

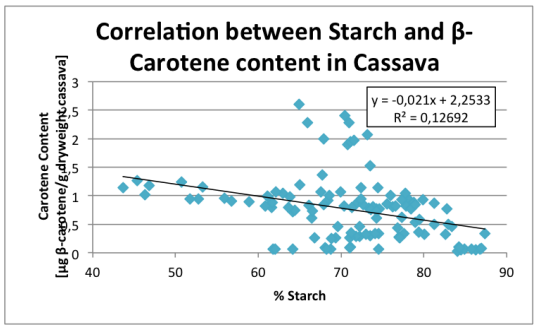

FIGURE 4.6: Scatterplot of β-carotene and starch content of the 117 samples from the 13 genotypes.

the spectrophotometric method with the HPLC method to make it possible to graphically compare the data.

The R-value of 0.99212 is close to 1.0 suggesting that there is a very high correlation between these analysis methods. From a regression fit done in Minitab, a p-value with an exponent of negative twelve is obtained concluding that this correlation is significantly reliable.

4.5 Orange gene

Considering the information presented in section 2.1.5, a modification of the amino acid 108 in the orange gene from arginine to histidine would be of interest for future work. The Orange gene from cassava was aligned with Arabidopsis, cauliflower and sweet potato. The alignment can be seen in Figure 4.8.

The alignment of sweet potato is low, concluding that it cannot be used for further discussion in this case. The alignment for the other species shows that position 108 express arginine, R, in cassava and cauliflower as well as for Arabidopsis. The gene is conserved between the species in this area, indicating that an alteration at position 108 in cassava might cause similar effects as it did in Arabidopsis.

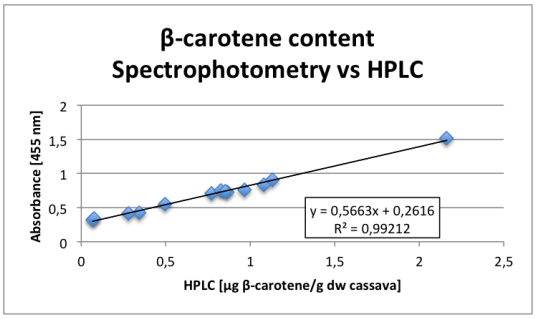

FIGURE 4.7: Correlation results from spectrophotometric and HPLC analysis of β-carotene.

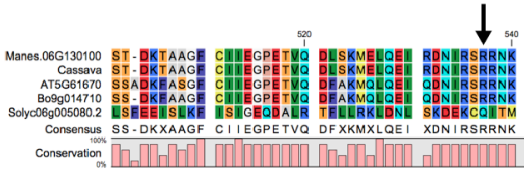

FIGURE 4.8: Alignment of the Orange gene for Cassava, Arabidopsis (*Arabidopsis thaliana*), Cauliflower (*Brassica oleracea*) and Sweet potato (*Ipomoea batatas*). The arrow points out position 108 for Arabidopsis.

4.6 Statistical analysis

The graphs for starch and carotenoid content presented in the result section as Figure 4.1 and 4.2 include error bars showing the SD for the different data sets. These error bars are all in an acceptable range within 95 % of the normal distribution of the mean, except for one single genotype measurement. The starch measurement for genotype G has a higher standard deviation than 95 % of the normal distribution and the reliability of this measurement is therefore lower than for the rest of the results.

To analyse the statistical differences between the data sets, ANOVA methods were used, followed by a Tukey test in Minitab on both the starch and the β-carotene content. For the starch measurements, the Tukey test showed that genotype A is significantly higher in starch than all the other genotypes. The Tukey test for both the starch and the β-carotene is presented in Appendix A. For the β-carotene, genotype L has a significantly higher content than the other genotypes.

In order to analyse the correlation between the starch and β-carotene content, and for the different β-carotene analysis methods, regression fit can be used. The R-square values give information on how significant the correlation is. The correlations can be viewed in Figure 4.6 and 4.7.

5 Discussion

The RNA isolation done at SLU gave no results. This might be due to the lack of fresh material, since RNA is sensitive and breaks down easily. It might also be due to the starch content being high in the samples, obstructing the RNA isolation. Due to the lack of results, all the genetic work was done at IITA.

The starch measurements were differing with more than 30 % for the genotypes tested. This suggests that the starch content can vary considerably between varieties, making it an important factor to consider when selecting the genotype to use in the future. A high starch content is of importance for the quality of the food products derived from cassava. A cassava flour with high provitamin A content would not be favoured among the consumers if the starch content was lower than their regular cassava flour. To improve the accuracy of the results, other methods for starch measurements needs to be added to gain credibility.

The carotene content for the genotypes showed larger differences than the starch content. This suggests that an accurate measurement of the β-carotene content is of great importance when deciding on a suitable genotype. The HPLC analysis did not show significant levels of other carotenoids than β-carotene, confirming that focusing on β-carotene is reasonable. To confirm that all carotenoids are extracted from the cells and detected in the samples, additional measurements needs to be done with samples of known contents. The levels of β-carotene detected in these measurements are somewhat lower than literature values, indicating that the extraction does not fully release the metabolites.

Since there is a moderate correlation of the starch and β-carotene content suggested in this thesis, and a proven correlation published, it would be of interest to use bioinformatic approaches to investigate transcriptional differences in the genes controlling these two components and how they are working together. It is not known how the synthesises of starch and carotenoids is linked and why this occurs, but it could be of great interest to find out more about this correlation.

The correlation between β-carotene analysis from HPLC and spectrophotometry was high. This suggests that in the future, spectrophotometric methods could be sufficient for the analysis of β-carotene content, especially when the cis- and trans forms are not of interest. This does at least seem to be true for the samples used in this report. Time is saved because no technician is needed for HPLC support, and approximately eight minutes is required for HPLC analysis of each sample. For further β-carotene measurements in this cassava project, it would be recommended to use a spectrophotometric method since the advantages of this method stand out and compensate for the reduction of accuracy.

5.1 Suggested genotype for further research

The optimal genotype would give the highest result in both starch and β-carotene content. Unfortunately, the negative correlation or other unknown reasons lead to no such genotype being found. When considering the genotype that is the best for future research, genotype L or N are my suggestions. Genotype L is a candidate since it has a very high β-carotene content as well as a high starch content. Despite the fact that genotype N has a lower β-carotene content compared to genotype L, it is included as a candidate here because its starch content is very high.

The candidates presented only takes into consideration the starch and β-carotene content. There are additional traits of great importance. The genotype needs to be climate tolerable, disease resistant and susceptible for transformation.

When working with biofortification, it is of importance to keep in mind that surviving on a one crop diet is impossible. Trying to reduce human suffering as much as possible by biofortification is of course fundamental, but the best diet is varied.

6 Conclusion

During the process of writing this Master Thesis I learned much more than can be presented in a report. I have gained a lot of lab experience. I learned that even though the internet is great and much information can be obtained from the web, the most important source of information is found with experienced colleagues working with specific tasks.

My conclusion on the genotypes to use for further research is N or L, but the exact decision process lies beyond my expertise. No significant negative correlation between the starch and the β-carotene content for the 13 cassava genotypes analysed could be proven.

CRISPR/Cas9 editing of the Orange gene would be an interesting addition in this field to explore further. This is not yet done but could possible increase the β-carotene content in cassava.

7 Future Work

Considering the results, the suggested genotype to be used in the future work in this project will be genotype N and/or L. It would be of interest to compare genotypes and expression patterns for variety A and E, which are each others opposites in regards to starch and β-carotene. This could possibly tell us which genes are involved and which are of importance in controlling starch and β-carotene content. Leaves from the thirteen genotypes could be used to study the relation in carotenoid content between root and leaf in order to better understand the genomic pathways.

The work of this thesis included the collection of data presenting the prevalance of cis- and trans form of β-carotene. The relation between the configuration forms was not evaluated in this thesis, but it could be of interest to comprehend the data already available.

It would be of great interest to involve CRISPR/Cas9 in cassava trials aiming at substituting the Orange gene amino acid 108 with histidine in order to enhance the β-carotene content.

It is also of importance to research if the improved qualities of the cassava plants will stay among the population and not be lost after a few generations. A stable trait is always desired.

8 References

Alexandersson, E. (2017). MSc or research project in plant nutrition: Carotenoid content in Cassava, Plantlink. Available at: $< http://www.plantlink.se/images/stories/plant_link/MSc_projects/MSc\%20projects_Cassava.pdf > 2017-08-12$.

Arora, L. and Nurula, A. (2017). Gene Editing and Crop Improvement Using CRISPR-Cas9 System. Frontiers. Plant Sci., 08 November 2017.

Ceballos, H., Davrieux, F., Talsma, E. F., Belalcazar, J., Chavarriaga, P. and Andersson, M. S. (2017). Carotenoids in Cassava Roots. Intech. Chapter 12, pp. 189-221.

Cervantes-Flores, J., C., Sosinski, B., Pecota, K., V., Mwanga, R., O., M., Catignani, G., L., Truong, V., D., Watkins, R., H., Ulmer, M., R. and Yencho, G., C. (2011). Identification of quantitative trait loci for dry-matter, starch, and b-carotene content in sweetpotato. Mol Breeding. Chapter 28, pp. 201–216.

Cho, K-S., Han, E-H., Kwak, S-S., Cho, J-H., Im, J-S., Hong, S-Y., Sohn, H-B., Kim, Y-H. and Lee, S-W. (2016). Expressing the sweet potato orange gene in transgenic potato improves drought tolerance and marketable tuber production. Comptes Rendus Biologies, vol. 339, Issues 5-6. pp. 207-213.

Crisp, P., Walkey, DGA., Bellman, E. and Roberts, E. (1975). A mutation affecting curd colour in cauliflower (Brassica oleracea L. var. Botrytis). Euphytica 24: 173–176.

Esuma, Williams. (2016). Genetic Analysis and Genome-wide Association Mapping of Carotenoid and Dry Matter Content in Cassava. Department of Plant Sciences. Faculty of Natural and Agricultural Sciences. University of the Free State, Bloemfontein, South Africa.

FAO, Food and Agriculture Organization of the United Nations. (2017). Available at: <http://www.fao.org/ag/agp/agpc/gcds/> 2017-08-17.

Farmaciaorganica. (2018). Cassava/Yuca. Available at:
<http://farmaciaorganica.com/product/cassavayuca/> 2018-01-23.

Giuliano, G. (2017). Provitamin A biofortification of crop plants: a gold rush with many miners. Current opinion in biotechnology, vol. 44, pp. 169-180.

Khatodia, S., Bhatotia, K., Passricha, N., Khurana, S. M. P., and Tuteja, N. (2016). The CRISPR/Cas Genome-Editing Tool: Application in Improvement of Crops, Frontiers in Plant Science, vol. 7, article 506.

Kim, S. H., Ahn, Y. O., Ahn, M-J., Jeong, J. C., Lee, H-S. and Kwak, S-S. (2013).

Cloning and characterization of an Orange gene that increases carotenoid accumulation and salt stress tolerance in transgenic sweet potato cultures. Plant Physiology and Biochemistry, vol. 70, pp. 445-454.

Labuschagne, M. and Akoroda, M. (2018). Principal value chains in Nigeria: the case of Cassava. Seminar. Swedish University of Agricultural Sciences. Delivered 2018-02-01.

Nature, International Journal of Science. (2018). Europe court suggests relaxed gene-editing rules. Available at: $< https : //www.nature.com/articles/d41586 - 018 - 01013 - 5?WT.feed_name = subjects_climate - sciences > 2018 - 01 - 22$.

PLAZA 4.0: an integrative resource for functional, evolutionary and comparative plant genomics. (2018). Nucleic Acids Research. (online access). Van Bel, M., Diels, T., Vancaester, E., Kreft, L., Botzki, A., Van de Peer, Y., Coppens, F. and Vandepoele, K. vol. 46:D1, 4 January 2018, pp D1190–D1196, Published 23 October 2017.

Saini, R. K., Nile, S. H., and Park, S. W. (2015). Carotenoids from fruits and vegetables: Chemistry, analysis, occurrence, bioavailability and biological activities. Food Research International, vol. 76, pp 735-750.

Wilson, M., Mutka, A., Hummel, A., Berry, J., Chauhan, R., Vijayaraghavan, A., Taylor, N., Voytas, D., Chitwood, D. and Bart, R. (2017). Gene expression atlas for the food security crop cassava. New Phytologist. vol. 213:4, pp 1632–1641. Published March 2017.

Yuan, H., Owsiany, K., Sheeja, T.E., Zhou, X., Rodriguez, C., Li, Y., Welsch, R., Chayut, N., Yang, Y., Thannhauser, T., Parthasarathy, M., Xu, Q., Deng, X., Fei, Z., Schaffer, A., Katzir, N., Burger, J., Tadmor, Y. and Li, L. (2015). A Single Amino Acid Substitution in an ORANGE Protein Promotes Carotenoid Overaccumulation in Arabidopsis. Plant Physiology. vol. 169:1. Published Sep 2015.

A Tukey Test for Starch and β-Carotene Measurements

A.1 Starch

Starch

Grouping Information Using the Tukey Method and 95% Confidence

Genotyp	N	Mean	Grouping
A	9	85,288	A
B	9	78,02	B
N	9	77,162	B
C	9	77,15	B
E	9	75,869	B C
I	9	74,19	B C D
H	9	72,19	B C D
L	9	69,621	C D E
O	9	69,29	C D E
F	9	67,74	D E
D	7	64,851	E F
J	9	58,02	F G
G	9	52,72	G

Means that do not share a letter are significantly different.

FIGURE A.1: Table presenting the Tukey test results for the starch measurements.

A.2 β-Carotene

Carotene

Grouping Information Using the Tukey Method and 95% Confidence

Genotyp	N	Mean	Grouping
L	9	2,1608	A
O	9	1,1308	B
G	9	1,0790	B
J	9	0,9640	B C
I	9	0,8632	C
N	9	0,8516	C
E	9	0,8242	C
D	8	0,7718	C
B	9	0,4936	D
C	9	0,3443	D E
H	9	0,28009	E
F	9	0,07490	F
A	9	0,06725	F

Means that do not share a letter are significantly different.

FIGURE A.2: Table presenting the Tukey test results for the β-carotene measurements.

B Calculations

To convert the HPLC [mAU*s] signal to μg β-carotene in the sample, equation B.1 was used. A β-carotene reference containing 0.024 μg of β-carotene which gave a signal of 977.45 mAU*s was used for the conversion.

$$SampleSignal[mAU*s]*\frac{(0.024[\mu g]}{977.45[mAU*s]}=\beta-carotene[\mu g] \tag{B.1}$$

When the amount of β-carotene in μg in the sample was known, equation B.2 was used to convert this number into the unit μg β-carotene/mg dry weight cassava. The injection volume was 5 μl, the concentration factor 10 and the starting volume 100 μl.

$$\begin{aligned}&\frac{[\mu g]insample}{5[\mu l]}/10*100/powderdryweight[mg]*1000\\&=\frac{[\mu g]\beta-carotene}{[mg]dryweightcassava}\end{aligned} \tag{B.2}$$