The role of genes involved in the biosynthesis of scopoletin in cassava post-harvest physiological deterioration

Lim, Yi-Wen

Award date:
2019

Awarding institution:
University of Bath

The role of genes involved in the biosynthesis of scopoletin in cassava post-harvest physiological deterioration

Yi-Wen Lim

A thesis submitted for the degree of Doctor of Philosophy

University of Bath

Department of Biology and Biochemistry

September 2018

This thesis is dedicated to my family and in memory of my father, Po Oun Lim

Thank you for everything you have done for me,

I would not be here without your support and love.

ACKNOWLEDGEMENTS

To my mother Loo Fong Poh: who never once stopped believing in me, and for the countless prayers over these past four years. I owe it all to you. Thank you!

I would like to acknowledgement the University of Bath for the fee waiver and to CEDIES for funding my PhD. And I would like to thank the Department of Biology and Biochemistry and the Biochemical Society for the travel grant, and the organizing committee for giving me the partial fellowship to attend the 1st World Congress on Root and Tuber Crops (WCRTC) in China, January 2016.

I would like to express my gratitude to my supervisors, Dr John Beeching and Prof Rod Scott, for their help, advice and support throughout my PhD. Especially to John for helping me with the writing of this thesis and guiding me through the different aspects of the PhD.

I am also extremely thankful for the help provided by Dr Herve Vanderschuren from ETH Zurich for the cassava FEC and advice on transformation work; and Dr Nigel Taylor from the Danforth Centre for his advice and help on cassava transformation.

To all the lab members of 4S 1.52; Shi: who kindly allowed me to use his transgenic cassava in this study and for his assistance when I first started in the lab, especially with molecular techniques. To Thoni: for your patience in helping me through so much! I am extremely thankful for your help, especially with the southern blot, transformation work, cassava transplanting, qPCR and biochemical analysis. I would not be able to complete my lab work so swiftly without your advice. To Lidia, Alex, Trupti, Toby and Glenda: thank you for all your support, both academic and non-academic; you have made working in this lab a wonderful experience, to which I will miss being a part of. To Ben: who very kindly helped with looking after my cassava when I hurt my foot, and for believing in Cassoverlord when I did not.

I would also like to thank all the lab members from 3S 1.13 for their help with *Arabidopsis* work, and for providing different perspectives into my work. I am grateful to Rob for looking after my cassava, thank you for tirelessly checking each plant for spider mites and keeping them in such great condition! I am also thankful to the labs 3S 1.09 and 4S 0.53 for the use of the qPCR machine and lab 4S 1.28 for the use of their Nanodrop machine. And to Dr Shaun Reeksting, for helping me with the HPLC biochemical analysis.

Finally, to everyone I have met in my life who has made it possible for me to persevere through the hard times: whether you know it or not, so much of me is made of what I have learned from you. And it may well be that our paths will never cross again, but I am so grateful to have met you; for through knowing you, my life has been changed for the better.

LIST OF ABBREVIATIONS

2ODG	2-oxoglutarate-dependent dioxygenase
4CL	4-coumarate CoA ligase
ABA	Abscisic acid
ABRC	Arabidopsis Biological Resource Center
ACC	1-aminocyclopropane-1-carboxylate
AOX	Alternative oxidase
APX	Ascorbate peroxidase
C3'H	*p*-coumaroyl shikimate 3' hydroxylase
C4H	Cinnamate 4-hydroxylase
CAD	Cinnamyl alcohol dehydrogenase
CAT	Catalase
CBB	Cassava bacterial blight
CBSD	Cassava brown streak disease
CBSV	Cassava brown streak virus
CCoAOMT	Caffeoyl CoA *O*-methyltransferase
CCR	Cinnamoyl-CoA reductase
CMD	Cassava mosaic disease
CMG	Cassava mosaic geminivirus
CoA	Coenzyme A
COMT	Caffeic acid *O*-methyltransferase
C_T	Cycle threshold
DNA	Deoxyribonucleic acid
F5H	Ferulate 5-hydroxylase
F6’H	Feruloyl 6'-hydroxylase
FP	Forward primer

GUS	β-glucuronidase
H_2O_2	Hydrogen peroxide
HCN	Hydrogen cyanide
HCT	Hydroxycinnamoyl-CoA shikimate:quinate hydroxycinnamoyl-transferase
HPLC-MS	High-performance liquid chromatography-mass spectrometry
HPRG	Hydroxyproline-rich glycoproteins
hptII	Hygromycin phosphotransferase II
MAPK	Mitogen-activated protein kinase
NADPH	Nicotinamide adenine dinucleotide phosphate
nptII	Neomycin phosphotransferase II
O_2^-	Superoxide radical
OH	Hydroxyl radical
OPP	Oxidative pentose-phosphate
PAL	Phenylalanine ammonia-lyase
PCD	Programmed cell death
PCR	Polymerase chain reaction
PPD	Post-harvest physiological deterioration
qRT-PCR	Quantitative real-time polymerase chain reaction
RNA	Ribonucleic acid
RNAi	Ribonucleic acid interference
ROS	Reactive oxygen species
RP	Reverse primer
RQ	Relative quantity
RT	Room temperature
RT-PCR	Reverse transcriptase Polymerase chain reaction

S8H	Scopoletin 8-hydroxylase
SAM	S-adenosyl-L-methionine
Scopolin-βG	Scopolin-beta-glucosidase
Scopoletin-GT	Scopoletin-glucosyltransferase
SOD	Superoxide dismutase
TAN	Tropical ataxic neuropathy
THT	Tyramine hydroxycinnamoyl transferase
Ubq10	Polyubiquitin 10
UV	Ultra-violet
WT	Wild-type

ABSTRACT

Cassava is the sixth most important crop in the world, feeding over 500 million people worldwide. The storage root is a staple source of carbohydrates, and its ability to grow in nutrient-poor soil and drought tolerance make it an ideal food security crop. Cassava is also used as animal feed and in industry for starch and biofuel. The potential of cassava is limited by several factors, with post-harvest physiological deterioration (PPD) of storage roots being a major constraint. PPD is a phenomenon triggered upon harvesting and mediated by reactive oxygen species and scopoletin accumulation that ultimately renders the storage roots unpalatable and unmarketable.

Scopoletin is biosynthesised through the phenylpropanoid metabolism, which is highly conserved in plants. Using *Arabidopsis thaliana* as a reference, homologous genes of enzymes involved in the biosynthesis of scopoletin were identified; these were phenylalanine ammonia-lyase (PAL), cinnamate 4-hydroxylase (C4H), 4-coumarate CoA ligase (4CL), Hydroxycinnamoyl-CoA shikimate:quinate hydroxycinnamoyl-transferase (HCT), *p*-coumaroyl shikimate 3' hydroxylase (C3'H), caffeoyl CoA *O*-methyltransferase (CCoAOMT), feruloyl 6'-hydroxylase (F6'H) and caffeic acid *O*-methyltransferase (COMT). Gene expression analysis was performed on said candidate genes and CCoAOMT, F6'H and COMT genes were selected to be knocked-down through RNAi silencing. Two RNAi constructs were generated, pRNAi-CCoAOMT/F6'H (2x), to target CCoAOMT and F6'H simultaneously, and pRNAi-CCoAOMT/F6'H/COMT (3x), to target CCoAOMT, F6'H and COMT simultaneously. Using Agrobacterium-mediated transformation, transgenic cassava lines were generated. Wild-type and pRNAi-F6'H (1x) transgenic cassava were grown alongside the 2x and 3x lines.

1x, 2x and 3x transgenic cassava lines have significantly reduced discolouration and scopoletin accumulation compared to wild-type plants. qRT-PCR results show F6'H having significant reduction in gene expression whereas expressions of CCoAOMT and COMT are not up-regulated during PPD; confirming *de novo* synthesis of scopoletin is mainly through the dominant pathway involving F6'H and not through *O*-methylation of esculetin.

TABLE OF CONTENTS

Acknowledgements 3

List Of Abbreviations 4

Abstract 7

1 Introduction 16

1.1 Cassava 16

1.1.1 Taxonomy and Origins 16

1.1.2 Morphology and Physiology 17

1.1.3 Growth and Cultivation 20

1.1.4 Global Importance of Cassava as a Crop 20

1.1.5 Constraints of Cassava as a Crop 22

1.2 Post-Harvest Physiological Deterioration 26

1.2.1 What is Post-Harvest Physiological Deterioration? 26

1.2.2 Mechanisms of PPD 26

1.2.3 Post-Harvest Responses in Other Root and Tuber Crops 31

1.3 Scopoletin 33

1.3.1 Scopoletin and Other Secondary Metabolites in PPD 33

1.3.2 Scopoletin Biosynthesis 35

1.4 Phenylpropanoid Metabolism 39

1.5 Research Strategy 44

2 Materials and Methods 46

2.1 Instruments 46

2.2 Plant Materials 46

2.3 Bacteria Strains 47

2.4 Plasmid 47

2.5 Software and Statistical Methods 47

2.6 Primers 48
2.7 Bacteria and Plant Manipulation 49
2.7.1 Antibiotics 49
2.7.2 Bacterial Cultures 49
2.7.3 Preparation of Electro-Competent *A. tumefaciens* 50
2.7.4 Transformation of Chemically Competent *E. coli* 50
2.7.5 Transformation of Electro-Competent *A. tumefaciens* 51
2.7.6 Subculture and Growth of Cassava 51
2.7.7 Cassava Transformation 51
2.8 Nucleic Acid Manipulation 52
2.8.1 Cassava Leaf DNA Extraction 52
2.8.2 Southern Blot 52
2.8.3 Cassava RNA Extraction 55
2.8.4 cDNA Preparation 56
2.8.5 *E. coli* Plasmid Extraction 56
2.8.6 Polymerase Chain Reaction (PCR) 56
2.9 Plasmid Manipulation 58
2.9.1 DNA Restriction Endonuclease Digestion 58
2.9.2 DNA Electrophoresis DNA Gel Purification 58
2.9.3 DNA Gel Purification 58
2.9.4 RNAi Construct 58
2.9.5 DNA Sequencing 58
2.10 PPD Assessment 59
2.10.1 Cassava Harvest 59
2.10.2 PPD Induction in Cassava Roots 59
2.10.3 PPD Measurement 60
2.11 Biochemical Analysis 60

2.11.1 Sample Preparation 60
2.11.2 Standard Preparation 60
2.12 Quantitative Gene Expression Analysis 61
2.12.1 Primer Design and Standard Curve 61
2.12.2 Comparative C_t Experiment 61
3 Are Phenylpropanoid Metabolism Genes Involved in Scopoletin Biosynthesis Found in Cassava? 63
3.1 Introduction 63
3.1.1 Background 63
3.1.2 Scopoletin Biosynthesis in Cassava and Arabidopsis 63
3.1.3 Phenylpropanoid Metabolism Genes Involved in Scopoletin Biosynthesis in Arabidopsis 66
3.2 Aim of Research 70
3.3 Results 70
3.3.1 Cassava Candidate Genes 70
3.3.2 Cassava Candidate Gene Expression Profiles 89
3.4 Discussion 93
4 Generating Cassava Knock-Down Lines Through RNAi Silencing 96
4.1 Introduction 96
4.2 Aims of Study and Research Strategies 99
4.3 Results 99
4.3.1 Creation of RNAi Constructs 99
4.3.2 Cassava Transformation and Growth 104
4.4 Discussion 106
5 Do the Transgenic Cassava Plants Differ in Terms of Growth and the PPD Response? 108
5.1 Introduction 108
5.2 Results 111

5.2.1 Do the Transgenic Cassava Plants Have Morphological Differences? 111
5.2.2 Is PPD Affected by the Downregulation Of F6'H, CCoAOMT and COMT? 117
5.2.3 Was the Accumulation of Scopoletin Reduced in Transgenic Cassava Lines? 119
5.2.4 Did the RNAi Constructs Successfully Knock-Down Said Genes of Interest? 121
5.3 Discussion 124
6 General Discussion 131
6.1 The Role of Scopoletin in Plants 132
6.2 The Role of CCoAOMT, F6'H and COMT 133
6.3 Approach to Solving PPD 135
6.4 Future Directions for Cassava PPD 137
References 138
Appendices 158

TABLE OF FIGURES

Figure 1.1 The flowers of cassava 18
Figure 1.2 The roots of (A) domesticated *Manihot esculenta* and (B) wild *M. esculenta flabellifolia* 19
Figure 1.3 A transverse section of a cassava storage root 19
Figure 1.4 PPD development in cassava storage root 26
Figure 1.5 Biological pathways affected the PPD process in cassava storage roots. 27
Figure 1.6 The mode of action of enzymatic ROS scavenging by superoxide dismutase and catalase. 29
Figure 1.7 Scopoletin and its glucoside, scopolin. 33
Figure 1.8 Cassava samples, (A) fresh and (B) 24 hours post-harvest, emit fluorescence under 366nm UV light. 34
Figure 1.9 Three alternative pathways of the biosynthesis of scopoletin in cassava. 37
Figure 1.10 A simplified representation of the phenylpropanoid metabolism to show the various secondary metabolites that are synthesised through it, including the coumarin scopoletin 40
Figure 1.11 The biosynthesis of scopoletin through three alternative pathways in the phenylpropanoid metabolism 45
Figure 3.1 The biosynthesis of scopoletin through three alternative pathways in the phenylpropanoid metabolism 65
Figure 3.2 Neighbour joining consensus phylogenetic analysis of phenylalanine ammonia-lyase (PAL) peptide sequences with 1000 bootstrap replicates 72
Figure 3.3 MUSCLE alignment of the cassava candidate genes (Cassava4.1_002628, Cassava4.1_034377, Cassava4.1_002709, Cassava4.1_002591, Cassava4.1_028434 and Cassava4.1_003117) and the *A. thaliana* PAL genes (PAL1 (At2g37040), PAL2 (At3g53260), PAL3 (At5g04230) and PAL4 (At3g10340)) 73
Figure 3.4 Neighbour joining consensus phylogenetic analysis of cinnamate 4-hydroxylase peptide sequences with 1000 bootstrap replicates 75
Figure 3.5 MUSCLE alignment of the cassava candidate gene's peptide sequence, Cassava4.1_005978, and the *A. thaliana* reference C4H, At2g30490 76
Figure 3.6 Neighbour joining consensus phylogenetic analysis of 4 coumarate CoA ligase (4CL) peptide sequences with 1000 bootstrap replicates 77
Figure 3.7 MUSCLE alignment of the cassava candidate peptide sequences, Cassava4.1_005006, Cassava4.1_005014, Cassava4.1_004658 and

Cassava4.1_004136, and the *A. thaliana* references, 4CL1 (At1g51680), 4CL2 (At3g21240), 4CL3 (At1g65060) and 4CL4 (At3g21230) 78
Figure 3.8 Neighbour joining consensus phylogenetic analysis of hydroxycinnamoyl-CoA shikimate:quinate hydroxycinnamoyl-transferase (HCT) peptide sequences with 1000 bootstrap replicates..... 80
Figure 3.9 MUSCLE alignment of the cassava candidate peptide sequences, Cassava4.1_008045 and Cassava4.1_008063, and the *A. thaliana* reference, At5g48930 81
Figure 3.10 Neighbour joining consensus phylogenetic analysis of *p*-coumaroyl shikimate 3' hydroxylase (C3'H) peptide sequences with 1000 bootstrap replicates ... 82
Figure 3.11 MUSCLE alignment of the cassava candidate peptide sequence, Cassava4.1_005910, and the *A. thaliana* reference sequence, At2g40890 83
Figure 3.12 Neighbour joining consensus phylogenetic analysis of caffeoyl CoA 3-*O*-methyltransferase (CCoAOMT) peptide sequences with 1000 bootstrap replicates 84
Figure 3.13 MUSCLE alignment of the cassava candidate peptide sequence, Cassava4.1_011832 and Cassava4.1_014783, and the *A. thaliana* reference sequence, At4g34050 85
Figure 3.14 Neighbour joining consensus phylogenetic analysis of caffeic acid *O*-methyltransferase (COMT) peptide sequences with 1000 bootstrap replicates 87
Figure 3.15 MUSCLE alignment of the cassava COMT candidate peptide sequence, Cassava4.1_010187 and Cassava4.1_010203, and the *A. thaliana* reference sequence, At5g54160 and POMT9 from *Populus deltoids* 88
Figure 3.16 Expression of the C4H candidate gene, Cassava4.1_005978 (MeC4H1) 92
Figure 3.17 Expression of the 4CL candidate gene, Cassava4.1_004658 (Me4CL3) 92
Figure 4.1 Biosynthesis of scopoletin through three pathways in the phenylpropanoid metabolism. 98
Figure 4.2 The cassette of pRNAi-GG. 99
Figure 4.3 Nucleotide Muscle alignment of RNAi target regions..... 100
Figure 4.4 ihpRNA construction with the pRNAi-GG vector using Golden Gate technology 101
Figure 4.5 ihpRNA construction with the pRNAi-GG vector through Golden Gate technology for simultaneous silencing of two genes..... 102
Figure 4.6 The pRNAi-GG primers position and restriction sites 103
Figure 4.7 Gel electrophoresis results of recombinant pRNAi-GG with primers P21 and P25 104

Figure 4.8 Confirmation of the RNAi insert by PCR using DNA from plantlets regenerated from transformed FEC with primers P21 and P25. 105
Figure 4.9 Southern blot to identify independent lines and transgene copy number 105
Figure 5.1 The biosynthesis of scopoletin through three alternative pathways in the phenylpropanoid metabolism. 109
Figure 5.2 Height of cassava plants at 5, 7 and 9 months 112
Figure 5.3 Comparison of wild-type and 3x plants at 9 months old 113
Figure 5.4 Wild-type cassava and 3xD transgenic cassava at 9 months old to show difference in leaf morphology 114
Figure 5.5 Comparison of wild-type cassava and 1xA transgenic cassava leaf morphology 115
Figure 5.6 Weight of cassava lines storage root at 9 months 116
Figure 5.7 The discolouration of the wild-type cassava compared with nine RNAi transgenic lines during the PPD time course 118
Figure 5.8 PPD development in the wild-type, 2xA and 3xB cassava storage roots 119
Figure 5.9 Scopoletin concentration (ng/mg) of the 6 lines over the PPD time course 120
Figure 5.10 Relative quantification ($2^{-\Delta\Delta Ct}$) of F6'H transcripts in transgenic and wild-type storage roots over the PPD timescale using qRT-PCR 122
Figure 5.11 Relative quantification ($2^{-\Delta\Delta Ct}$) of CCoAOMT transcripts in transgenic and wild-type storage roots over the PPD timescale using qRT-PCR 123
Figure 5.12 Relative quantification ($2^{-\Delta\Delta Ct}$) of COMT transcripts in transgenic and wild-type storage roots over the PPD timescale using qRT-PCR 124
Figure 6.1 The mode of action of enzymatic ROS scavenging 136

LIST OF TABLES

Table 1.1 Comparison of nutritional content between cassava and other root and tuber crops..... 21
Table 2.1 List of primers. 48
Table 2.2 Antibiotics and their properties. 49
Table 2.3 Composition of reaction for the probe..... 53
Table 2.4 Probe PCR reaction conditions 53
Table 2.5 Master mix composition for the digestion using *Hind*III 54
Table 2.6 Taq DNA polymerase Reaction Mixture 57
Table 2.7 Taq DNA Polymerase Set-up 57
Table 2.8 Q5 High Fidelity DNA Polymerase Reaction Mixture 57
Table 2.9 Reaction set up using Q5 High-Fidelity DNA polymerase..... 57
Table 3.1 RT-PCR results performed on cassava storage roots over the PPD timescale (0 hour, 24 hours, 48 hours and 72 hours) and leaf with gene specific primers excluding CCoAOMT and COMT which used primers that targeted both genes in the family 90
Table 3.2 Names given to cassava candidate genes encoding relevant enzymes. ... 91
Table 4.1 Primers used in the construction of pRNAi-CCoAOMT/F6'H and pRNAi-CCoAOMT/F6'H/COMT. 102
Table 4.2 Primer sequences used to identify recombinants, intron orientation and sequencing for the pRNAi-GG vector..... 103
Table 4.3 Cassava transgenic lines and their transgene number 106

1 INTRODUCTION

1.1 CASSAVA

1.1.1 TAXONOMY AND ORIGINS

Cassava (*Manihot esculenta* Crantz) is a perennial, dicotyledonous shrub belonging to the Euphorbiaceae; a large family that consists of approximately 300 genera and 8000 species and includes other important agronomical species such as castor bean (*Ricinus communis*) and rubber (*Hevea braziliensis*) (Lebot, 2009). Although the Euphorbiaceae group is known for its latex producing capacity (Webster, 1994), cassava is not grown for this purpose but for its storage roots instead.

The genus *Manihot* consists of 98 species which are all found in the Americas (Rogers and Appan, 1973), of which many are a source of food to humans (Olsen and Schaal, 2001). For a period of time, it was thought that *M. esculenta* originated through interbreeding of different species in the *Manihot* genus (Allem, 1994) but the discovery of wild populations led to the reclassification of three cassava subspecies: *M. esculenta* Crantz ssp. *esculenta* (cultivated), *M. esculenta* Crantz ssp. *peruviana* (wild) and *M. esculenta* Crantz ssp. *flabellifolia* (wild) (Allem, 2002). Olsen and Schaal (2001) used microsatellite polymorphism to show that *M. esculenta* was domesticated from the wild subspecies *M. esculenta flabellifolia* since microsatellites of *M. esculenta* are a subset of those of *M. esculenta flabellifolia*.

Cassava may have been domesticated approximately 10,000 years ago or by 7000 BC (Isendahl, 2011; Olsen and Schaal, 1999) but the site of domestication is highly contested with multiple sites in South America being proposed (Allem, 2002). Using a phylogenetic approach by evaluating polymorphism within the single-copy nuclear gene, glyceraldehyde 3-phosphate dehydrogenase (G3pdh), Olsen and Schaal (1999) identified the South Amazon border with Bolivia as the most likely origin of domestication. In addition, with botanical mapping of the current distribution of cassava, it was suggested that the savannas to the south of the Amazon rainforest, the Cerrado of the Brazilian states of Rondonia, and the north-western Mato Grosso are likely to be the geographical origin of cassava (Isendahl, 2011).

Cassava was introduced to Africa and Asia by the Spanish and Portuguese during the sixteenth and seventeenth centuries respectively (Cock, 1982) and has since become a staple food source in many countries. Nowadays, cassava is cultivated for its starchy

storage roots and it is grown in tropical and subtropical regions all around the world. The global yield of cassava is forecasted to be 278.0 million tonnes per annum in 2017 with Nigeria producing 55 million tonnes of the annual yield. Nigeria is followed by Thailand (30.9 million tonnes), Indonesia (20.3 million tonnes) and Brazil (20.1 million tonnes) (FAO, 2017).

1.1.2 MORPHOLOGY AND PHYSIOLOGY

Cassava is considered to be an amphidiploid or sequential allopolyploids and possesses 2n = 36 chromosomes (El-Sharkawy, 2003). Cassava has a height of 1-4 m and has simple leaves with palmated veins, ranging between 3-9 lobes depending on the cultivar, age of the plant and environmental conditions. The leaves are alternate and the length of a petiole from a fully opened leaf usually ranges between 5-30 cm, although up to 40 cm has been recorded. The upper (adaxial) surface of the leaves is covered with a waxy epidermis with most cultivars having stomata located on the undersurface (abaxial). A study by El-Sharkawy and Cock (1990) found that only 2% of 1500 cultivars examined had stomata on the adaxial surface; these stomata are functional and bigger than those on the abaxial surface. The large variability in cassava morphological characteristics suggests a high level of interspecific hybridisation (Alves, 2002).

Cassava is monoecious, bearing both male (pistillate) and female (staminate) flowers on the same inflorescence (Figure 1.1). The male flowers are located on the upper part of the inflorescence and the female flowers on the lower part. There are usually fewer female flowers than male flowers and the female flowers open 1-2 weeks before the males to encourage cross pollination, thus contributing to the high heterozygosity observed in cassava populations (Alves, 2002). Individuals within a population often have staggered flowering time that may be affected by environmental conditions (Ceballos *et al.*, 2004). In addition, some cultivars have never been known to flower (Sayre *et al.*, 2011; Manu-Aduening *et al.*, 2005), therefore making cassava breeding difficult and time consuming. Successful fertilisation will yield three seeds per capsule (Alves, 2002) and a high temperature is required to break dormancy and induce germination (Pujol *et al.*, 2002). The unpredictability of cassava flowering, its highly heterozygous population and germination conditions renders propagation through sexual reproduction difficult (Ceballos *et al.*, 2004).

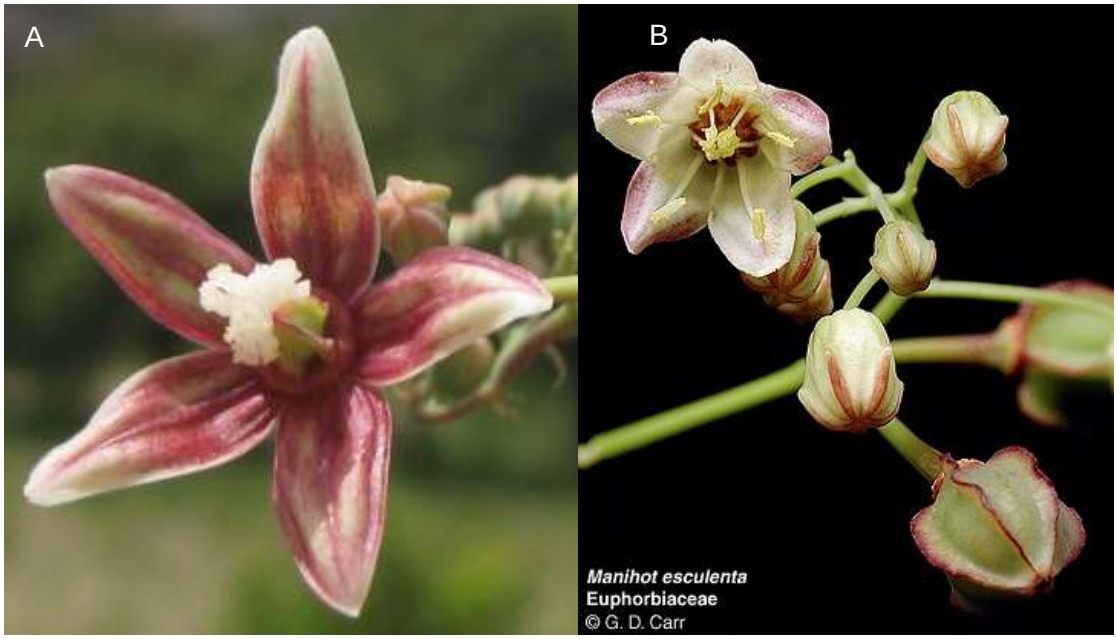


Figure 1.1 **The flowers of cassava.** (A) A female flower with the stigma and ovary present. (B) A male flower with anthers and a fertilised ovary on the bottom right corner. Photos courtesy by Ahmad Fathoni and Gerald D. Carr respectively (Liu, 2016).

The cassava root cannot be considered a true tuber like potato (*Solanum tuberosum*) as tubers are derived from stem tissues whereas cassava storage roots are derived from root tissue entirely (Alves, 2002). As the cassava plant matures, secondary thickening of fibrous roots will lead to the development of storage roots. Zhang *et al.* (2003b) found that during the initiation of storage root formation, fibrous roots that will become storage roots have a rapid upregulation of activity of the cambium, leading to the differentiation of secondary xylem into enlarged parenchyma cells in which starch will accumulate. This is an abnormal but advantageous phenotype that has been selectively bred by man since the domestication of cassava as shown in Figure 1.2 (Jansson *et al.*, 2009). Growth conditions and cultivar variety affects the number of storage root yield per plant; an individual plant will usually develop between 2-3 kg or 3-10 storage roots (van Oirschot *et al.*, 2000). As cassava storage root is not a true tuber, formation of bud primordia will not occur, therefore it cannot be used as propagules. This is why cassava has to be propagated via woody stem cuttings; although this limits the rate of multiplication, it ensures that the economically important part of the crop can be fully utilised and not be reinvested in new planting material (Cock, 1982).

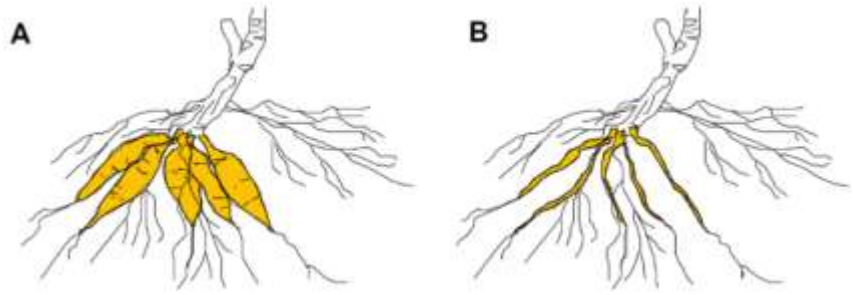


Figure 1.2 The roots of (A) domesticated *Manihot esculenta* and (B) wild *M. esculenta flabellifolia*. The storage roots are coloured yellow. Picture modified from (Jansson *et al.*, 2009).

A mature storage root has a diameter of 3-15 cm and a length of up to 100 cm (Alves, 2002). It is comprised of three distinct tissues: bark (periderm), peel (cortex) and the edible parenchyma (Figure 1.3). During the development of storage roots, the periderm which is made of a thin layer of dead cork cells will be shed and replaced by new cells to accommodate the thickening of the root. The flexible cortex which is only 1 mm thick contains the sclerenchyma, cortical parenchyma and phloem (Alves, 2002). The parenchyma accounts for 85% of the root's weight and the colour varies considerably between cultivars as it correlates with the total carotenoid content of the storage root (Sánchez *et al.*, 2006). The parenchyma is the edible part of the storage root and contains a large amount of starch, accounting for approximately 85-91% of the total weight (Alves, 2002). It also contains radially distributed xylem vessels and a large vascular bundle running through the middle of the storage root (Onwueme, 1978).

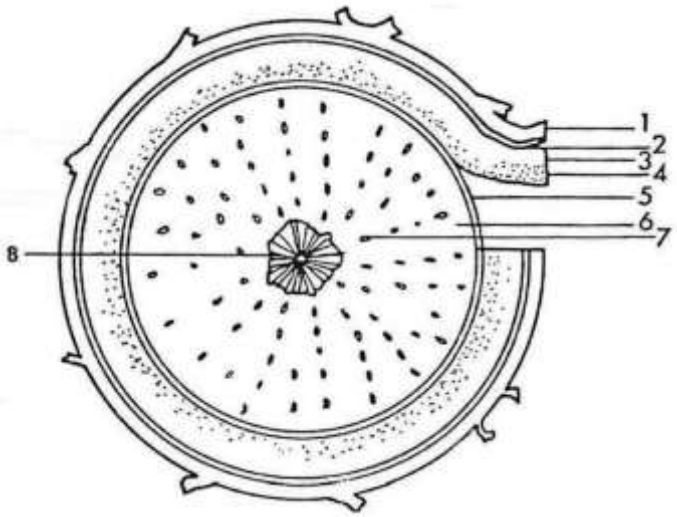


Figure 1.3 A transverse section of a cassava storage root. (1) bark/periderm, (2-4) the peel, (2) sclerenchyma, (3) cortical parenchyma, (4) phloem, (5) cambium, (6) storage parenchyma, (7) scattered vascular bundles, (8) central vascular bundle. (Hunt *et al.*, 1977)

1.1.3 GROWTH AND CULTIVATION

Cassava is a hardy plant that can grow on nutrient-poor soil and withstand up to six months of drought (Lokko *et al.*, 2007). Due to this, it is often grown on marginal soil that is unsuitable for other crops. Depending on the cultivar and growth conditions, cassava can reach a height of up to four metres (Alves, 2002). As cassava is a tropical crop, its optimal growth conditions would be between 25-29°C with an annual rainfall of 1000-1500 mm but it can still grow and yield despite limited water supply. Cassava growth will halt if the temperature is lower than 17°C or higher than 37°C (El-Sharkawy, 2003).

Since the cassava storage root is a true root and not a propagule, it must be propagated via hardwood stem cuttings. Stems that have hardened but not extensively lignified from 8-18 month old cassava are ideal for propagating (Lozano *et al.*, 1977). Stem cuttings of approximately 20-30 cm with 5-7 nodes are used for propagation as it will have sufficient nutrients for the initial establishment of roots and leaves (Onwueme, 1978). Within 5-7 days, adventitious roots will develop from the nodes of axillary buds and a callus will form at the base of the cutting, from which roots will develop (El-Sharkawy, 2003). Leaves will start developing 10-13 days after planting. The cassava plant will produce a well-established root system, stem and leaves within the first two months. By the third month, 3-10 fibrous roots will start thickening and initiating the development into storage roots. After 4-5 months, the plant will have achieved its maximal canopy and will start to produce more photosynthetic products than what is required for growth, leading to carbohydrate being stored as starch in the storage roots. From the sixth month onwards, root bulking accelerates, leaves begin to senesce as carbohydrates are increasingly being transported from the leaves to the root, and the stem becomes lignified. After ten months, plants may go through a period of inactive growth, after which it can resume storage root bulking and vegetative growth (El-Sharkawy, 2003; Alves, 2002).

1.1.4 GLOBAL IMPORTANCE OF CASSAVA AS A CROP

Cassava is the sixth most important crop globally (Mann, 1997) and feeds approximately 500 million people worldwide (Best and Henry, 1992). The starchy storage root is a staple source of carbohydrates and accounts for 50% of total calorific intake in certain countries (Cock, 1985). Besides that, cassava is a rich source of minerals and vitamins; it not only has a high carbohydrate content but is also rich in vitamin C (Table 1.1), with just 450g of cassava meeting the recommended daily dietary intake of vitamin C for an adult male (Chandrasekara and Josheph Kumar, 2016).

Table 1.1 Comparison of nutritional content between cassava and other root and tuber crops (Chandrasekara and Josheph Kumar, 2016).

Nutrient (per 100 g)	Cassava	Potatoes		Sweet potatoes	Yam
		White	Red		
Energy (kcal)	160	69	70	86	118
Carbohydrate (g)	38.1	15.7	15.9	20.1	27.9
Protein (g)	1.4	1.7	1.9	1.6	1.5
Total lipid (g)	0.3	0.1	0.1	0.1	0.2
Dietary fibre (g)	1.8	2.4	1.7	3.0	1.8
Calcium (mg)	16	9	10	30	17
Magnesium (mg)	21	21	22	25	21
Potassium (mg)	271	407	455	337	816
Vitamin C (mg)	20.6	19.7	8.6	2.4	17.1
Vitamin A (IU)	13	8	7	14187	138
Vitamin K (µg)	1.9	1.6	2.9	1.8	2.3
Vitamin E (mg)	0.19	0.01	0.01	0.26	0.35

Due to its ability to grow on nutrient-poor soil and withstand extended periods of drought, cassava plays a key role in food security, especially in African countries. Cassava is an attractive crop for poorer farmers in the tropics as land unfitted for other agronomic purposes can be used. It can also be incorporated into an intercropping strategy since it can be grown at the end of a cropping cycle when the soil nutrient levels are low (Borin and Frankow-Lindberg, 2005).

Nowadays, cassava storage roots are also used as animal feed and as raw material for industrial purposes such as starch, paper, textiles, biofuel and many more (Vlaar *et al.*, 2007). As demands for renewable alternatives to fossil fuels increases, cassava is becoming a favoured alternative to corn starch in the biofuel industry due to its high yield and cheaper price (Zhang *et al.*, 2003a). Cassava is an ideal alternative to corn for bioethanol production as firstly, it can be grown and harvested all year round without seasonal limitations (Nguyen *et al.*, 2007). Secondly, it produces a good yield under sub-optimal growth conditions such as limited water and nutrients, making it a crop of low maintenance and cost. Thirdly, land that is not able to sustain other crops can be used to grow cassava which maximises land use. Finally, the high starch content of the cassava root makes it ideal for ethanol fermentation (Ziska *et al.*, 2009). In Asia, especially Indonesia, China, Thailand and Vietnam, cassava is mainly grown for

industrial purposes. China is currently the biggest importer of cassava chips and the promising cassava market has contributed to the industrialisation and development of rural areas in developing countries. For example, in Vietnam, cassava has become a cash crop where cassava chips and starch are being produced competitively. This has created many jobs, attracted foreign investment, increased exports and contributed to modernisation and industrialisation of rural areas (Kim *et al.*, 2017).

1.1.5 CONSTRAINTS OF CASSAVA AS A CROP

Cassava, like most other crops, has its share of constraints and setbacks that limits its production and utilisation. These include the low nutritional value of the storage root, the accumulation of toxic cyanogenic glucosides, diseases and pests and the short shelf-life of the storage root post-harvest.

1.1.5.1 NUTRITIONAL CONTENT

Despite being a good source of calories due to its reliable yield and high carbohydrate content, cassava is unfortunately lacking in other macronutrients and vitamins (Nassar and Ortiz, 2010). Cassava is severely lacking in protein as it only accounts for 1-5% of the root's total dry weight unlike cereals where 14% of its dry weight is protein, making cassava one of the staple foods with the lowest protein to energy ratio (Stephenson *et al.*, 2010; Young and Pellett, 1994). Besides that, essential sulphur-containing amino acids such as leucine and lysine are also only present at low levels (Stupak *et al.*, 2006). Cassava leaves, on the other hand, at the age of 11-12 months, consists of 20-30% of crude protein, with essential amino acids apart from sulphur-containing ones (Eggum, 1970). A setback to this is the cyanogenic compounds in the leaves require prolonged boiling to be removed thereby destroying the amino acids (Stupak *et al.*, 2006). Since cassava is mainly a staple food in poorer countries and communities, people relying on cassava as a vital food source may suffer from protein energy malnutrition (Stupak *et al.*, 2006).

1.1.5.2 CYANOGENIC COMPOUNDS

In its raw form, cassava is potentially toxic to human consumption due to the presence of cyanogenic glucosides which are present in all cassava tissues, apart from the seeds. The cyanogenic glucosides present are mainly linamarin, which accounts for approximately 95% of total cyanogenic glucosides, the glucoside of acetone cyanohydrin

and lesser quantities of lotaustralin (Siritunga and Sayre, 2003; Santana *et al.*, 2002; Cock, 1985). Linamarin is synthesised in the leaves and then transported to the roots and stored in the cell vacuoles (Sayre *et al.*, 2011; Du *et al.*, 1995). Leaves of cassava and the peel of storage roots contain the highest levels of cyanogenic glucosides. The quantity of cyanogenic glucosides also depends on the cassava cultivar (Wheatley *et al.*, 1992); most cultivars have less than 100 ppm fresh weight of cyanogenic glucosides but there are 'bitter' cultivars which have up to 500–1000 ppm of linamarin (Cardoso *et al.*, 2005). According to WHO, the maximal residual cyanogen concentration is 10 ppm, therefore these cultivars are obviously harmful if not processed appropriately.

The enzyme linamarase, located in the cell membrane on the inner cell wall, is responsible for the release of hydrogen cyanide (HCN). It is found in high concentrations in cassava peel but is also present in the parenchyma (flesh) (Cooke *et al.*, 1978). The toxic cyanogenic glucosides are a defence mechanism against herbivores; when the cells are mechanically damaged, linamarin and linamarase will come in contact with each other and react to produce acetone cyanohydrin. Cyanohydrin will then decompose to release HCN spontaneously if the pH is greater than 5 or catalysed by hydroxynitrile lyase (Sayre *et al.*, 2011). This release of HCN as a gas provides a method to detoxify cassava roots and avoid cyanide poisoning. By mashing, soaking, fermenting and boiling the cassava roots, most of the cyanogenic glucosides will be successfully removed through gaseous HCN unless there is an exceedingly high concentration of cyanogen.

Tropical ataxic neuropathy (TAN) and konzo (irreversible spastic paraparesis) are diseases that arise from acute or chronic exposure to cassava cyanogen and can happen if the cyanogenic glucosides are inadequately removed (Cardoso *et al.*, 2005; Siritunga and Sayre, 2003). In addition, the low protein content of cassava, especially the lack of sulphur-containing amino acids, compounds the condition as they are involved in cyanide detoxification (Cliff *et al.*, 2011; Ernesto *et al.*, 2002).

To address the problem of cyanide toxicity from cassava roots, transgenic lines where linamarin synthesis was blocked were generated. Unfortunately, inhibition of linamarin biosynthesis in the leaves led to the reduction of linamarin in the root and caused transgenic plants to have impaired growth in the field (Sayre *et al.*, 2011).

1.1.5.3 PESTS AND DISEASES

Cassava is susceptible to many pests and diseases resulting in significant yield loss, especially in South America and Africa (Lebot, 2009). Cassava is affected by

approximately 200 different species of pests including mealybugs, whiteflies, mites, thrips, gallmidges, stem borers, root mealy bugs, lace bugs, shoot flies, hornworm, burrower bugs and more (Bellotti *et al.*, 1999). These pests not only cause physical damage to the plant, but some are involved in transferring and spreading diseases. Nematodes can also cause damage to the plant, increasing its vulnerability towards subsequent bacterial and fungal infections. One of the most damaging bacterial diseases is cassava bacterial blight (CBB) caused by the pathogen *Xanthomonas axonopodis* pv. *Manihotis*; CBB significantly impacts cassava yield as leaf biomass, storage root and stem (planting material) are all affected (Wydra *et al.*, 2001). Besides that, there are around 250 species of fungi that affect cassava; *Cercospora henningsii* being the most prominent as it causes brown leaf spot disease in cassava. This disease is widespread in all cassava producing countries and is characterised by the brown spots on both leaf surfaces (Lebot, 2009).

Cassava is also susceptible to a range of viral diseases. At least 17 different viruses that infect cassava have been identified, with cassava mosaic disease (CMD) and cassava brown streak disease (CBSD) being the most damaging to cassava production in Sub-Saharan Africa (Ephraim *et al.*, 2015). The economic losses due to these two diseases are estimated to be more than USD 1 billion per year (Legg *et al.*, 2006b).

CMD is caused by cassava mosaic geminiviruses (CMGs) which are transmitted by the whitefly (*Bemisia tabaci*) and reports show that the CMD pandemic has caused an economic loss of approximately US1.9-2.7 billion annually in East and Central Africa (Patil and Fauquet, 2009). It is one of the most detrimental plant virus diseases in the world, causing famine that has led to the death of thousands of people (Legg *et al.*, 2006a). CMD is currently only present in the African continent and Indian subcontinent but is continuing to expand (Fargette *et al.*, 2006). Even though countries in South-East Asia and South America are cultivating cassava at a large-scale, CMD is absent due to *B. tabaci* biotype being unable to colonise cassava successfully in said countries (Carabali *et al.*, 2005). However, CMD still poses a huge threat and strict quarantine restrictions are necessary to prevent the import of CMGs and cassava-adapted *B, tabaci* biotypes (Patil and Fauquet, 2009). Several cultivars that are resistant to CMD are currently being used to control CMD in Africa but are inadequate to combat the CMD pandemic (Legg and Fauquet, 2004). RNAi silencing mechanisms have been used to target CMG promoters and genes involved in CMD and have produced promising results though a fully resistant cultivar has not yet been successfully engineered (Bull *et al.*, 2011; Sayre *et al.*, 2011; Zhang *et al.*, 2005).

Cassava brown streak disease (CBSD) is caused by two viruses belonging to the *Ipomovirus* genus, cassava brown streak virus (CBSV) and Ugandan cassava brown streak virus (UCBSV) (Revers and García, 2015). This disease is widespread in the East and coastal regions of Africa (Hillocks and Jennings, 2003). CBSD causes severe necrosis and chlorosis of infected leaves, dry brown-black necrotic rot of the storage root which starts as a lesion and may spread to the entire root and may lead to stunted growth of the cassava plant (Bull *et al.*, 2011; Alicai *et al.*, 2007). Tolerant cultivars may be able to restrict the rot in the storage root to a lesion, but susceptible cultivars may suffer the loss of the entire root. Depending on the cultivar, CBSD can cause up to 70% of root weight loss (Maruthi *et al.*, 2005). Current means of controlling the disease are through field sanitation and thorough screening to avoid using diseased stem cuttings. Using CRISPR/Cas9-mediated genome editing, cassava mutants show suppressed disease symptoms but still not a complete resistance to CBSD (Gomez *et al.*, 2018).

1.1.5.4 POST-HARVEST PHYSIOLOGICAL DETERIORATION (PPD)

Additionally, a major problem cassava faces as a crop is its short shelf life after which the storage root will deteriorate to the point of becoming unpalatable and unmarketable. The short shelf-life of cassava is attributable to a response is called post-harvest physiological deterioration (PPD), and is an active physiological response triggered by harvesting. This causes huge economical losses for farmers and severely restricts the potential of cassava as a crop. It is estimated that 5-25% of cassava roots are lost due to PPD worldwide (Saravanan *et al.*, 2014), with Africa suffering up to 29% of losses and Latin America and Asia losing 10% and 8% respectively (Salcedo and Siritunga, 2011). The issue of shelf-life is especially problematic in the case of cassava farming as many farmers operate at small-scale in developing countries and PPD makes it hard for them to sell their products as the market demand mainly exists for roots in their fresh form. A study by (Rudi *et al.*, 2010) has shown that delaying PPD by two weeks in Nigeria could increase profit by approximately USD 2.9 billion over a 20-year period. In the North-East of Thailand alone, it could save around USD 26.5 million per year for the starch industry and USD 8.5 million per year for farmers if PPD is delayed by 45 days (Vlaar *et al.*, 2007).

1.2 POST-HARVEST PHYSIOLOGICAL DETERIORATION

1.2.1 WHAT IS POST-HARVEST PHYSIOLOGICAL DETERIORATION?

Post-harvest physiological deterioration (PPD) can be characterised by the blue-black discolouration of the parenchyma and xylem vessels known as vascular streaking (Figure 1.4). This is caused by the inevitable wounding during harvest as discolouration appears first around the wounds and gradually spreads through the root (Reilly *et al.*, 2007). PPD is intricately connected to the wound response as it is induced by mechanical damage to the root but is further influenced by various environmental factors such as growth and storage conditions (Tumuhimbise *et al.*, 2015). These traits make the mechanism of PPD difficult to understand as it is not caused by a single factor and is more of a network comprised of multiple physiological responses. Early studies found that PPD is a spontaneous endogenous physiological response rather than a microbial or pathological response since neither bacteria nor fungi were isolated from freshly harvested roots and bactericide and fungicide did not influence the deterioration process (Noon and Booth, 1977). Although this is the case initially, PPD is often followed by microbial infection, known as secondary or microbial deterioration, due to its disordered physiology (Buschmann *et al.*, 2000a).

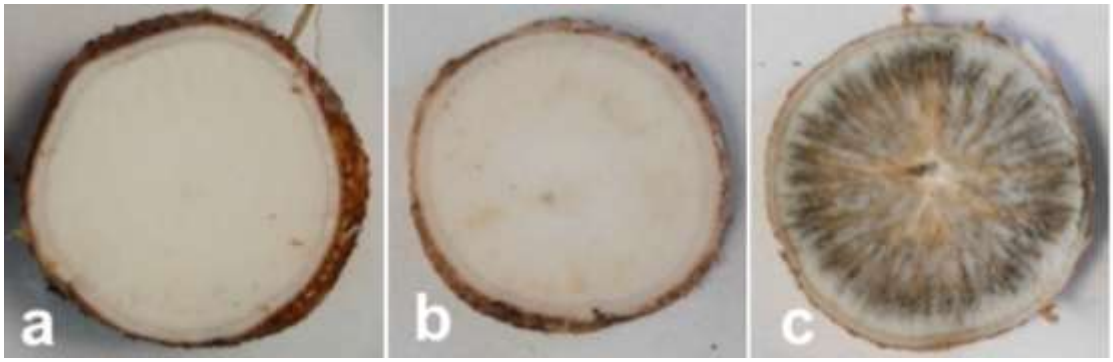


Figure 1.4 PPD development in cassava storage root. (A) Fresh root (B) Root 24 hours post-harvest (C) Root 72 hours post-harvest.

1.2.2 MECHANISMS OF PPD

1.2.2.1 BACKGROUND

PPD is a complicated phenomenon that is related to the wound response and shares many similarities such as gene expression changes, accumulation of secondary metabolites and production of enzymes and molecules involved in defence (Beeching *et al.*, 2000). PPD is triggered by the wounding of the root during harvest which causes an initial oxidative burst of superoxide radical (O_2^-) within 15 minutes. Following this,

reactive oxygen species (ROS) such as hydrogen peroxide (H_2O_2), hydroxyl radical (OH) and secondary metabolites such as scopoletin, which have anti-microbial and/or anti-oxidant activities are produced (Qin *et al.*, 2017; Vanderschuren *et al.*, 2014). The initial wounding stimulates the cells to release signal molecules that lead to the production of compounds in different pathways that are involved with defence, tissue cell protection, restoration and programmed cell death (PCD) (Figure 1.5) (Owiti *et al.*, 2011). Signal molecules such as the phytohormone ethene, which is involved in coordinating wound and senescence response in plants (Ecker and Davis, 1987), is increased during PPD (Hirose *et al.*, 1984). Other signal molecules produced include jasmonic acid, salicylic acid, abscisic acid and H_2O_2 (Beeching *et al.*, 1994). PPD also causes an increase in respiration and activity of acid invertase, which causes the mobilisation of starch to sugars (Tanaka *et al.*, 1983).

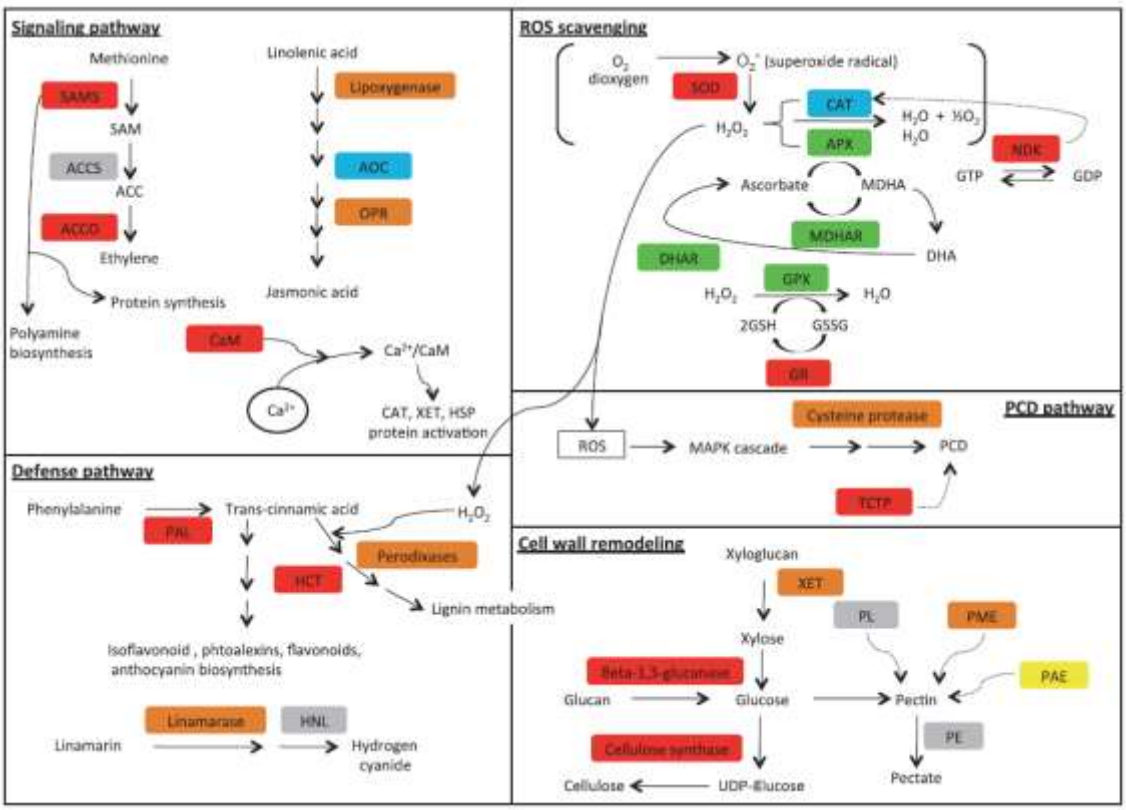


Figure 1.5 **Biological pathways affected the PPD process in cassava storage roots.** Red boxes indicate proteins up-regulated during early stages of PPD, orange/brown boxes indicate proteins up-regulated during late stages of PPD, yellow boxes indicate proteins down-regulated during early PPD, green boxes indicate protein down-regulated during late PPD, and blue boxes indicate proteins with both up- and down-regulation during PPD. SAMS, S-adenosyl methionine synthase; ACCS, 1-aminocyclopropane-1-carboxylate synthase; ACCO, 1-aminocyclopropane-1-carboxylate oxidase; AOC, allene oxide cyclase; OPR, oxophytodienoate reductase; CaM, calmodulin; SOD, superoxide dismutase; CAT, catalase; APX, ascorbate peroxidase; NDK, nucleoside diphosphate

kinase; MDHAR, monodehydroascorbate reductase; DHAR, dehydroascorbate reductase; GPX, glutathione peroxidase; GR, glutathione reductase; PAL, phenylalanine ammonia lyase; HCT, hydroxycinnamoyl transferase; HNL, hydroxynitrile lyase; TCTP, translationally controlled tumor protein; XET, xyloglucan endotransglycosylase; PL, pectate lyase; PME, pectin methyl esterase; PAE, pectin acetylesterase; PE, pectin esterase (Owiti *et al.*, 2011).

1.2.2.2 PPD AND THE OXIDATIVE BURST

PPD is the result of an oxidative process as the visual symptom of PPD, the blue-black discolouration of the xylem vessels and parenchyma, is due to the oxidation of phenolic compounds, in particular, the coumarin scopoletin (Buschmann *et al.*, 2000a). The oxidative burst is one of the earliest observable aspects of a plant's defence strategy where large quantities of reactive oxygen species (ROS) are produced rapidly and transiently (Low and Merida, 1996). In normal conditions, ROS are generated inevitably as a by-product of photosynthesis through successive one-electron reductions of oxygen (O_2). ROS are involved in several plant metabolic pathways such as the cell cycle, growth, hormone signalling, PCD and biotic and abiotic stress responses (Figure 1.5) but they can also cause oxidative damage to cells (Mittler *et al.*, 2004). Plants can normally prevent this potential toxicity with ROS scavenging mechanisms but during the oxidative burst, the rapid production of ROS disrupts the equilibrium between ROS production and scavenging (Apostol *et al.*, 1989).

The oxidative burst is one of the earliest events that lead to PPD development in cassava (Iyer *et al.*, 2010). Within 15 minutes of harvest where inevitable tissue damage occurs, superoxide radical (O_2^-) is detected, followed by hydrogen peroxide (H_2O_2) within 3 hours of injury. The reaction between H_2O_2 and secondary metabolites (mainly scopoletin) results in the blue-black pigmentation observed in PPD (Reilly *et al.*, 2003). Xu *et al.* (2013) showed that increasing the expression of genes encoding ROS scavengers, superoxide dismutase (SOD) and catalase (CAT), delayed PPD development. SOD and CAT are ROS scavengers that work together to remove ROS. Firstly, SOD converts superoxide (O_2^-) to hydrogen peroxide (H_2O_2). Then, CAT converts hydrogen peroxide (H_2O_2) to water (H_2O) and oxygen (O_2) (Figure 1.6) (Apel and Hirt, 2004).

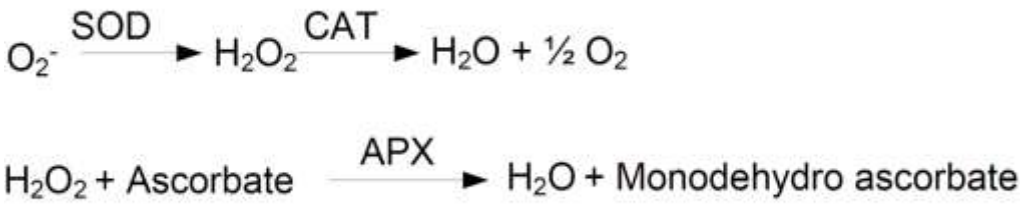


Figure 1.6 The mode of action of enzymatic ROS scavenging by superoxide dismutase and catalase (Apel and Hirt, 2004).

ROS are signalling compounds that manage various conditions that arise due to stress such as production of restorative or defensive compounds, and PCD. Interaction between ROS and the cells can cause oxidative damage to DNA, proteins and lipids (Apel and Hirt, 2004). Therefore, another approach to prevent PPD could be to stop the cells from producing excessive amounts of ROS. Cyanogenesis plays an important role in the oxidative burst that results in PPD (Zidenga *et al.*, 2012). The disrupted tissues release hydrogen cyanide (HCN) through cyanogenesis from linamarin and linamarase. HCN stops mitochondrial respiration by inhibiting complex IV in the mitochondrial electron transfer chain, leading to a rapid and excessive production of ROS at complex I and III. This suggests that PPD could be cyanide dependent. A study conducted by Zidenga *et al.* (2012) showed that the overexpression of the *Arabidopsis thaliana* mitochondrial alternative oxidase gene (AOX1A) in cassava resulted in a 10-fold reduction in ROS accumulation and delayed PPD for up to two weeks. Unfortunately, this had adverse effects on the root yield under field conditions. In addition, recent studies showed that melatonin can control homeostasis of cellular ROS in the cells and by applying exogeneous melatonin, PPD has been significantly delayed (Hu *et al.*, 2016; Ma *et al.*, 2016). Melatonin activates ROS-scavenging and ROS signal transduction pathways such as calcium signalling, mitogen-activated protein kinases (MAPK) and transcription factors at the early stage of PPD. It also plays a role in the reduction of H_2O_2 by increasing peroxidase and catalase activity during middle and late stages of PPD. Therefore, manipulating the biosynthesis of melatonin could potentially delay the onset of PPD (Hu *et al.*, 2016).

1.2.2.3 PPD AND THE WOUND RESPONSE

Cassava PPD resembles wound responses found in other plant species (Rickard, 1985). During PPD development, changes in total phospholipids, glyceroglylipids and sterol-containing lipids can be observed in damaged tissues; these membrane lipids have

presumably been broken down (Beeching *et al.*, 1999). The interaction of ROS such as H_2O_2 with membrane lipids leads to the production of long chain fatty acids such as linolenic acid which is a precursor in jasmonic acid biosynthesis, a signalling molecule involved in the wound response (Farmer and Ryan, 1992).

Starch degradation is a part of the wound response, and during the cassava PPD development, an increase in total soluble sugars, such as fructose and glucose, is observed (Uarrota *et al.*, 2015). Soluble sugars are involved in stress responses and act as nutrient and metabolite signalling molecules which activate specific or hormone-crosstalk transduction pathways, leading to changes in gene expression and protein synthesis (Couée *et al.*, 2006). High glucose levels can activate NADPH oxidase in the oxidative pentose-phosphate (OPP), which is a major cofactor of pathways involved in ROS scavenging such as ascorbate-glutathione cycles (Couée *et al.*, 2006). Glutathione levels have been shown to correlate with the plant's adaptation to stress conditions (May *et al.*, 1998). Besides that, different sugars may have different effects in oxidative stress responses. For example, ascorbate levels in harvested broccoli florets increased after being fed with sucrose (Nishikawa *et al.*, 2005). Under the action of the ascorbate peroxidase (APX) enzyme, ascorbate converts H_2O_2 to water thus preventing ROS accumulation (Figure 1.6). Sánchez *et al.* (2013) found that storage roots with a higher sugar content correlated with delayed PPD onset in cassava when stored in ambient conditions. The trend of higher sugar level and delayed PPD was also observed in cassava roots when the plant was pruned before harvest (van Oirschot *et al.*, 2000). Although studies have shown that sugars seem to play a role in delayed PPD in cassava storage roots, the exact role that sugars play during PPD development is still unknown.

In addition to the production of signalling and defensive compounds, cassava storage roots also produce reparative compounds such as suberin and lignin to heal damaged tissues and to prevent microorganism invasion (Beeching *et al.*, 1999). Wound repair, or curing, occurs upon tissue disruption. The activation of PAL enzymes and up-regulation of hydroxyproline-rich glycoproteins (HPRGs) during PPD lead to the production of various compounds which include suberin and lignin (Reilly *et al.*, 2007). Cassava roots are capable of repairing wounded tissues independent of PPD when they are still attached to the main stem. Unfortunately, once the storage root is detached from the stem, the ability to repair becomes inadequate and thus PPD is observed (Beeching *et al.*, 2000). The process of sealing and healing of the wound sites does occur in the storage root, but it is too slow under normal conditions to prevent deterioration. If the storage roots are kept under optimum storage conditions of 30-35°C and 80-85% relative humidity, a wound cork can form and delay the progress of PPD (Beeching *et al.*, 1994).

However, the curing method is uneconomical due to the specificity of the storage conditions and if the cured storage roots are wounded again, rapid PPD will develop regardless of the previous treatment (Beeching *et al.*, 1994).

1.2.3 POST-HARVEST RESPONSES IN OTHER ROOT AND TUBER CROPS

Like cassava, other root and tuber crops are often inevitably damaged during harvest, transportation and storage. Wound responses in other root and tuber crops such as potato, sweet potato, taro and yam are similar to cassava, but they recover better from wounding and stresses in comparison to cassava, making their storage easier and shelf-life longer. Understanding and comparing post-harvest responses of other root and tuber crops may provide beneficial insights into cassava PPD.

1.2.3.1 POTATO

Potato (*Solanum tuberosum* L.) is cultivated for its starchy tuber and is the fourth highest produced food crop in the world after maize, wheat and rice. The inevitable wounding of the potato tubers during harvest leads to the production of polyphenolic and polyaliphatic compounds which are involved in the suberisation of wounds thus preventing microbial infection and desiccation of the tuber (Lulai *et al.*, 2008). Within 15 minutes of wounding, callose is produced, leading to the biosynthesis of ethene. Enzymes such as lipoxygenase and plant defence genes such as phenylalanine ammonia-lyase (PAL) and extensin are also activated in response to wounding (Lulai and Corsini, 1998). There was also an increase in the expression of the enzyme, tyramine hydroxycinnamoyl transferase (THT), within 3 to 4 hours of wounding. THT is responsible for the biosynthesis of hydroxycinnamoyl tyramines which are involved in suberisation (Negrel *et al.*, 1993). Like cassava, potato tubers also experience starch degradation; this happens in the first 12 hours after wounding because of the increase in the activity of alpha and beta-amylases (Kato and Uritani, 1976).

1.2.3.2 SWEET POTATO

Sweet potato (*Ipomoea batatas* (L.) Lam) is grown for its storage root like cassava. It is cultivated across Asia, tropical Africa and America for both food and industrial uses (Lebot, 2009). Unlike cassava, sweet potato has a relatively long shelf life, depending on the storage conditions and cultivars, it can last for weeks up to 10 months if it is stored

in optimal conditions (Ray and Ravi, 2005; Holmes and Stange, 2002). Sweet potato has a thin skin that is vulnerable to injury during harvest and handling. When the sweet potato is wounded, the site of injury will undergo curing where the exposed parenchyma suberises and forms a periderm to seal the site of wounding (Ray and Ravi, 2005). Phenolic compounds will also accumulate underneath the periderm after wounding (Schadel and Walter Jr, 1981). Besides that, wounding of the sweet potato roots also triggers ethene synthesis which will acts as an activator of various metabolic responses such as the upregulation of PAL and peroxidase activity (Imaseki *et al.*, 1968). The increased activity of peroxidase indicates that ROS is involved in the wound response mechanism in sweet potato. Ethene also induces responses that increase resistance to fungal resistance (Stahmann *et al.*, 1966).

1.2.3.3 YAM

Yams (*Dioscorea spp.*) consist of multiple species from the Dioscoreaceae family and are cultivated for their starchy tubers. They are also a major food source in Africa, Asia, South America and the Pacific. The deterioration seen in yams is very similar to sweet potato. Post-harvest loss usually stems from an increase in respiration which leads to a reduced weight and a loss of marketability (Ravindran and Wanasundera, 1992). The cortical periderm of yam is vulnerable due to its thinness, which can easily be damaged during harvest (Ejechi and Souzey, 1999). Unlike cassava, yams are stem tubers and not a true root which may be why they can fully cure wounds despite being separated from the main stem (Ravi *et al.*, 1996). When the tuber is wounded, a response is triggered to form a cork periderm underneath the wounds within five days to seal to suppress respiration rate and water loss, thus allowing storage for several months. Abscisic acid (ABA) induces the expression of PAL which leads to the accumulation of phenolic compounds (Cottle and Kolattukudy, 1982). The amount of polyphenols present will be 5-6 times higher five days after wounding than in fresh tubers (Ikediobi *et al.*, 1989). Besides that, oxidation-related enzymes, such as peroxidase, polyphenol oxidase and lipoxygenase, are also activated following wounding; polyphenol oxidase is shown to be greatly up-regulated, suggesting polyphenols are involved in wound suberisation (Rhodes and Wooltorton, 1978) as well as stress resistance and antioxidation (Ikediobi *et al.*, 1989).

In addition, sprouting of the tubers is another major source of losses for yams. After harvest, yams resume a state of dormancy where metabolic activity has been reduced to a minimum (Passam and Noon, 1977). Sprouting can be delayed in temperatures of

around 15°C, regardless of relative humidity. Unfortunately, such a temperature is not ideal for curing and hard to be achieved in practice in tropical regions. Batatasins are endogenous phenolic growth inhibitors that are responsible for inducing and maintaining dormancy; yams only start sprouting when the levels of batatasin fall. Once the yam has sprouted, available nutrients will be used for growth and the tuber will rapidly start to senesce, ultimately leading to tuber deterioration (Passam and Noon, 1977).

1.3 SCOPOLETIN

1.3.1 SCOPOLETIN AND OTHER SECONDARY METABOLITES IN PPD

As mentioned in section 1.2.2.2, during the development of PPD in cassava, secondary metabolites involved in the stress response such as coumarins, phenolic compounds, phytosterols and fatty acids accumulate. Although it is unusual in plant stress responses, a large family of 22 different diterpenic compounds that accumulated in wounded cassava storage roots has been identified (Sakai and Nakagawa, 1988). Other metabolites that are involved in the PPD response include flavanols, leucoanthocyanins, catechin, gallocatechin and delphinidin (Sakai *et al.*, 1986). However, out of all the metabolites involved in PPD, hydroxycoumarins show the most significant accumulation (Bayoumi *et al.*, 2010), especially scopoletin and its glucoside scopolin (Figure 1.7). Esculetin and its glucoside esculin are also present, although to a lesser extent. These secondary metabolites may have antimicrobial abilities or act as anti-oxidants (Buschmann *et al.*, 2000b).

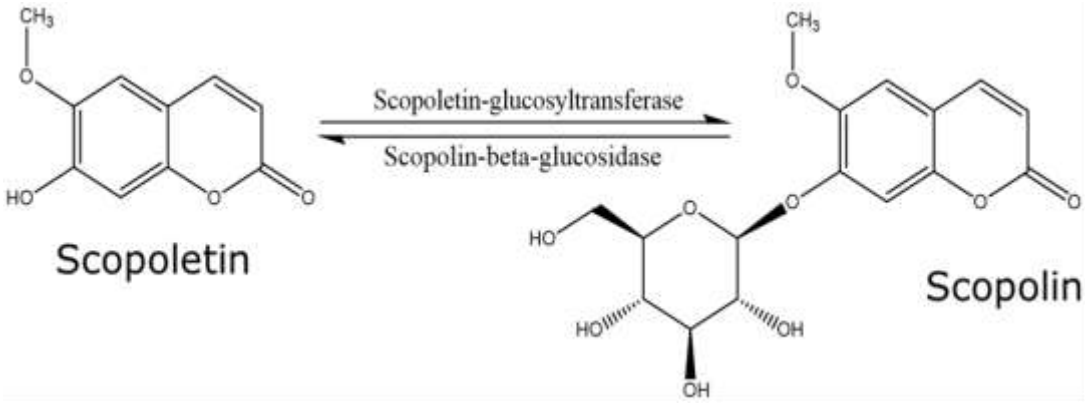


Figure 1.7 Scopoletin and its glucoside, scopolin. Scopoletin can be converted to scopolin by the enzyme scopoletin-glucosyltransferase and vice versa by scopolin-beta-glucosidase.

The hydroxycoumarin, scopoletin (7-hydroxy-6-methoxychromen-2-one), is of interest in understanding cassava PPD as it shows a significant accumulation over the PPD process. Harvested cassava roots show a rapid increase in the concentration of scopoletin from <1 ng/mg in fresh roots to 60-80 ng/mg two days post-harvest, and remains at a high concentration throughout the whole PPD process (Buschmann *et al.*, 2000b). This accumulation of scopoletin observed in PPD suggests that it is involved in storage root deterioration. Besides that, peroxidase has a high affinity to scopoletin, which oxidises it to a blue-black insoluble metabolite. This correlates with the blue-black discolouration known as vascular streaking observed in the roots. When exogenous scopoletin is applied to fresh cassava root, the roots show a faster rate of discolouration and more intense fluorescence under UV light (Figure 1.8) (Wheatley and Schwabe, 1985). In addition, a study performed in sunflower (*Helianthus annuus*), where peroxidase was extracted from a detached leaf was incubated with scopoletin and H_2O_2 produced an insoluble blue product. But, if H_2O_2 was absent or if catalase was added, no blue product was formed, suggesting that it is the oxidation of scopoletin that forms the observed blue product (Edwards *et al.*, 1997). Due to the bioactive nature of scopoletin, under normal conditions, it is stored as the inactive glucoside known as scopolin. During PPD or stress conditions, scopoletin is both synthesised *de novo* and released from its inactive glycone, scopolin, via deglycosylation through a specific beta-glucosidase.

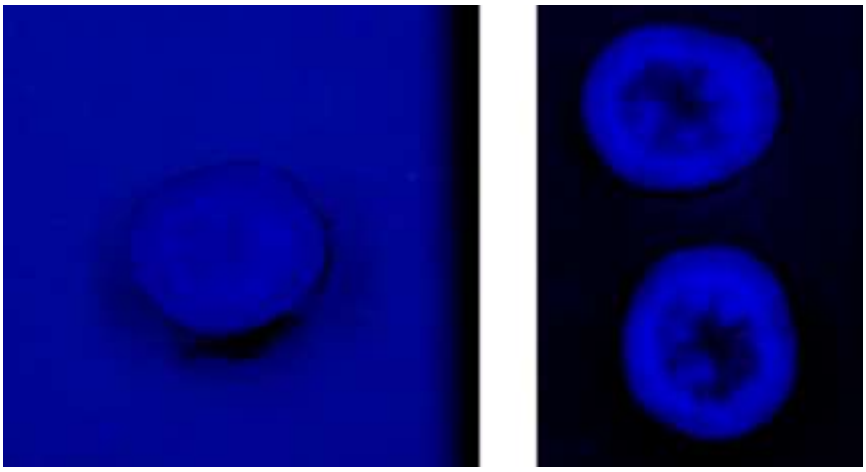


Figure 1.8 Cassava samples, (A) fresh and (B) 24 hours post-harvest, emit fluorescence under 366 nm UV light. (A) The fluorescence of the root is not significantly different compared to the filter paper beneath; (B) Cassava roots 24 hours post-harvest developed significant fluorescence in comparison to the filter paper beneath. Picture taken from Liu (2016).

Scopoletin also plays a role in plant defence as it has antifungal and antimicrobial properties (Rodríguez *et al.*, 2000). The fungitoxicity of cassava was investigated on

Microcyclus ulei (P. Henn.) V. Arx that infects *Hevea brasiliensis* (rubber tree) which is also in the Euphorbiaceae family. Following inoculation, the leaflets showed scopoletin accumulation and the *in vitro* tests confirm the inhibitory action of scopoletin on *M. ulei* conidium germination and germ tube elongation (Garcia *et al.*, 1995). Besides that, *H. brasiliensis* infected with *Phytophthora palmivora* also showed a high level of scopoletin accumulation and that scopoletin inhibited *P. palmivora* mycelial growth (Churngchow and Rattarasarn, 2001). In addition, the study by Valle *et al.* (1997) showed a rapid accumulation of scopoletin following inoculation of *Ophiostoma ulmi* on cell suspension cultures of *Ulmus campestris* Linn. and *U. pumila* L.; antifungal activity was also directly observed in the *in vitro* bioassays performed.

Scopoletin may also have some health implications as it can trigger apoptosis in mammalian cells. A study by Kim *et al.* (2005) found that scopoletin induced apoptosis in human premyeloleukemic cells and activated caspase-3 as shown by the proteolytic cleavage of the proenzyme and protease activity. Liu *et al.* (2001) also found that at doses of 100, 200 and 400 ug/ml, scopoletin induces apoptosis and cell cycle arrest thereby inhibiting the proliferation of human prostate adeno-carcinoma cells.

1.3.2 SCOPOLETIN BIOSYNTHESIS

Since scopoletin plays a vital role in the development of PPD in cassava storage roots, understanding its biosynthesis will provide an important insight into extending the shelf life and delaying PPD in cassava. Scopoletin is synthesised through the phenylpropanoid metabolism which is also responsible for many other secondary metabolites that are essential for the survival of the plant. These secondary metabolites include lignin which is needed for structural support and vascular integrity, flavonoids which are involved in UV protection and in attracting pollinators, coumarins and phytoalexins which are involved in defence mechanisms and many more (Vogt, 2010; Dixon *et al.*, 2002). Within the phenylpropanoid metabolism, there are three likely pathways leading to the biosynthesis of scopoletin (Figure 1.9). These are via (1) hydroxycinnamate, (2) hydroxycaffeate and (3) ferulate intermediates; the dominant pathway is via the ferulate intermediate where 90% of total scopoletin in cassava is produced (Bayoumi *et al.*, 2008).

As shown in Figure 1.9, L-phenylalanine (L-Phe) is processed to *trans*-cinnamate ((*E*)-cinnamate) by phenylalanine ammonia-lyase (PAL) and then (*E*)-*p*-coumarate through cinnamate 4-hydroxylase (C4H). It is then linked to a coenzyme A (CoA) through a thioester bond to its carboxyl group by 4-coumarate CoA ligase (4CL). *p*-coumaroyl

shikimate 3' hydroxylase (C3'H) then adds a hydroxyl group to the 3'-carbon of the activated (*E*)-*p*-coumaroyl CoA to convert it to (*E*)-caffeoyl CoA. The 3'-hydroxyl group is then methylated to (*E*)-feruloyl CoA (the CoA-linked form of ferulate) by caffeoyl CoA 3-*O*-methyltransferase (CCoAOMT) (Ferrer *et al.*, 2008).

The last enzyme involved in the biosynthesis of scopoletin in the major pathway is a 2-oxoglutarate-dependent dioxygenase (2ODG) known as feruloyl CoA 6'-hydroxylase 1 (F6'H1) in *A. thaliana* (Kai *et al.*, 2008). This enzyme adds a 6'-hydroxyl group to the benzene circle of (*E*)-feruloyl CoA to form (*E*)-6'-hydroxylferuloyl CoA. For lactonization to occur, the 6'-hydroxyl and carboxyl group needs to be converted via *E*-*Z* isomerisation to the Z-isomer. Once (*E*)-6'-hydroxylferuloyl CoA is converted by an isomerase to (*Z*)-6'-hydroxylferuloyl CoA, the lactone structure can form spontaneously thus generating a scopoletin molecule (Kai *et al.*, 2008). Scopoletin can then be glucosylated to its glucoside, scopolin, through scopoletin-glucosyltransferase (scopoletin-GT); and deglucosylated from scopolin back into scopoletin through scopolin-β-glucosidase (scopolin-βG) (Ahn *et al.*, 2010; Gachon *et al.*, 2004).

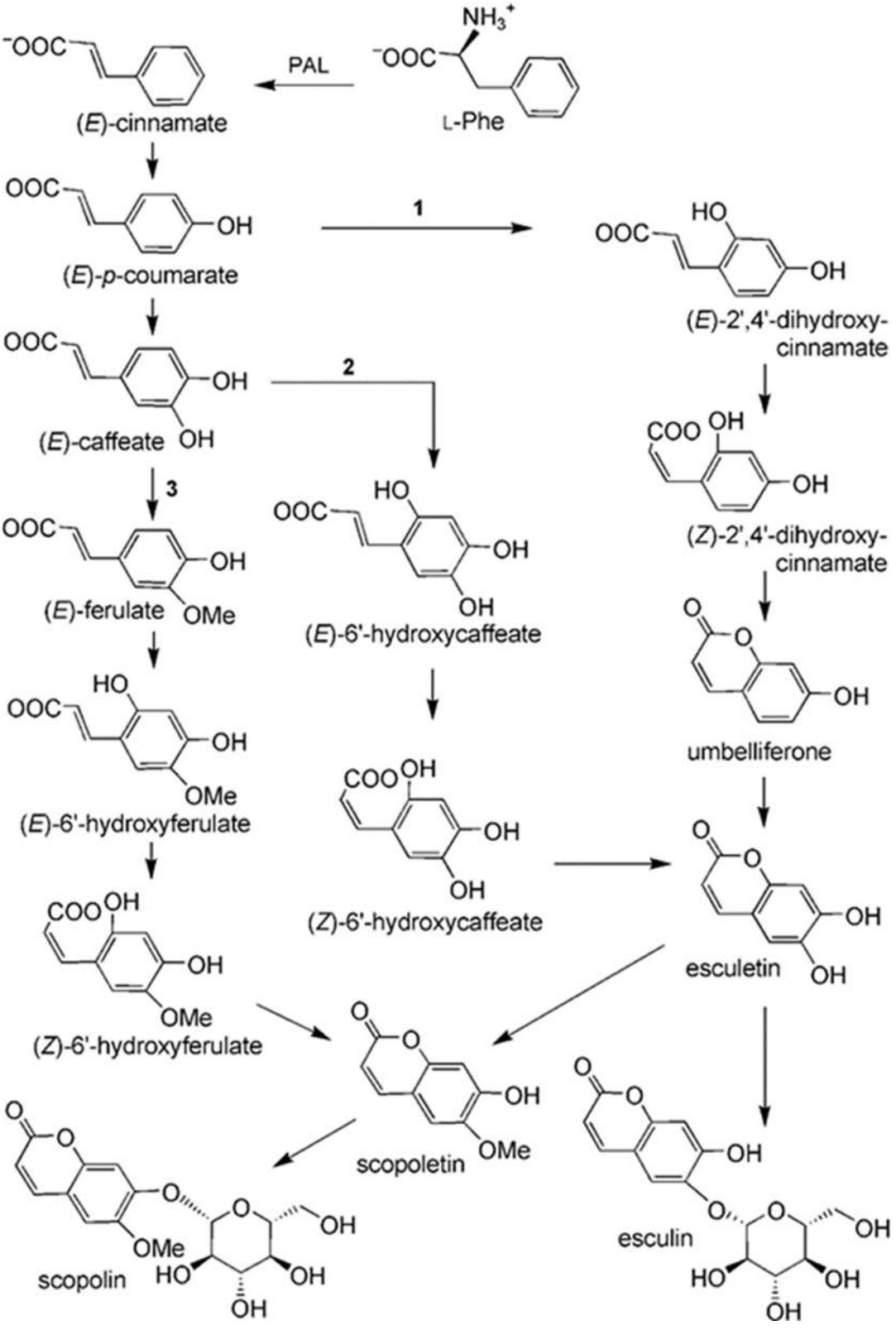


Figure 1.9 Three alternative pathways of the biosynthesis of scopoletin in cassava. Entry points of the three possible pathways are labelled 1, 2 and 3 (Bayoumi *et al.*, 2008).

Alternatively, scopoletin can be synthesised through esculetin (pathways 1 and 2). In pathway 1, (*E*)-*p*-coumarate is 2'-hydroxylated to produce 2', 4'-dihydrocinnamate which then goes through several reactions to produce esculetin. Pathway 2 branches from (*E*)-caffeoyl CoA which is 6'-hydroxylated into (*E*)-6'-hydroxycaffeoyl CoA and converted to (*Z*)-6'-hydroxycaffeoyl CoA through *E*-*Z* isomerisation and then processed into esculetin. Esculetin is then methoxylated into scopoletin through an *O*-methyltransferase enzyme. This was shown by Kim *et al.* (2006) where *Escherichia coli* (*E. coli*) expressing *O*-methyltransferase (*POMT-9*) from poplar (*Populus deltoids* Marsh) was able to generate scopoletin, isoscopoletin and scoparone from esculetin through the methoxylation of the C6 hydroxyl group. As mentioned previously, the two alternative pathways only account for less than 10% of the total amount of scopoletin produced in cassava roots undergoing PPD (Bayoumi *et al.*, 2008).

Through the understanding of the biosynthesis of scopoletin in cassava, studies have been done to reduce scopoletin accumulation with the aim of delaying PPD and extending the shelf-life of cassava. Liu *et al.* (2017) managed to significantly reduce scopoletin accumulation and deterioration in transgenic cassava through RNAi inhibition of cassava homologs of the *A. thaliana* F6'H1 gene. Unfortunately, scopoletin production was not completely inhibited, as there are seven members in the F6'H gene family in cassava for which the RNAi construct may not be able to target them all, despite the fact that most of the genes within the family share a high sequence similarity. Another reason for the accumulation of scopoletin in the transgenic lines may be due to the production of scopoletin through the alternative pathways. Although the alternative pathways normally account for less than 10% of the scopoletin produced during PPD, there might be a feedback mechanism that causes an increase in scopoletin production due to the inhibition of the major pathway. Besides that, transgenic cassava lines where cassava scopoletin-GT was knocked-down through RNAi inhibition and where *A. thaliana* scopolin-βG was overexpressed also showed reduction in scopoletin accumulation and delayed PPD. Interestingly, both transgenic lines would initially lead to increased accumulation of scopoletin in cells under normal conditions. However, due to its reactive nature, it is possible that enhanced peroxidase activity occurring as a direct consequence of increased scopoletin levels resulted in a higher oxidation rate of scopoletin, thereby reducing its overall accumulation (Chong *et al.*, 2002). It could also be due to a feedback mechanism where upstream enzyme activity is downregulated because of the accumulation of scopoletin (Fathoni, 2017). These findings further confirm the importance of scopoletin in PPD development and provide a platform for further to understand the deterioration process in cassava storage root.

1.4 PHENYLPROPANOID METABOLISM

As discussed in section 1.3, scopoletin plays a major role in the development of PPD in cassava storage roots and by manipulating genes involved in its biosynthesis, PPD can be delayed. Scopoletin is one of many secondary metabolites synthesised through the phenylpropanoid metabolism. Phenylpropanoids are a diverse group of compounds that are derived from the carbon skeleton of phenylalanine (the end product of the shikimate pathway) which may be involved in plant structural support, defence and survival (Figure 1.10) (Vogt, 2010). Compounds derived from the phenylpropanoid metabolism are referred to as 'secondary metabolites' because they are not directly involved in basal cellular processes such as nucleic acid and protein synthesis, respiration and photosynthesis. Due to the lack of understanding of the roles these secondary metabolites have in plants, they can appear to be superfluous to plant survival. Although not every compound's function is understood still, the importance of phenylpropanoids to the survival of the plant is undeniable. Specifically, the production of hydroxycinnamyl alcohols, also known as monolignols, which serve as building blocks of lignin (lignin monomers) and confer structural support, vascular integrity and defence in plants (Boerjan *et al.*, 2003). Lignin is formed through the dehydrogenative polymerisation of three monolignols, *p*-coumaroyl alcohol, coniferyl alcohol and sinapyl alcohol, which respectively give rise to *p*-hydroxyphenyl (H), guiacyl (G) and syringyl (S) units of the lignin polymer.

The entry point of the phenylpropanoid pathway is the amino acid L-phenylalanine (L-Phe) which is the end product of the shikimate pathway. Phenylalanine ammonia-lyase (PAL) catalyses the first step where L-phenylalanine is deaminated to produce *trans* (*E*)-cinnamate and ammonia (Figure 1.10). In *A. thaliana*, the active PAL isoforms are encoded by a gene family, as is the case in many other plant species; the 4 genes are PAL1-PAL4 (Raes *et al.*, 2003). According to a study by Raes *et al.* (2003), PAL1 and PAL2 encode the main PAL enzymes of *A. thaliana* as they have conserved specific promoter elements. In addition, *A. thaliana pal1pal2* double mutant over-accumulates phenylalanine and has reduced lignin content with a higher syringyl (S) to guaiacyl (G) ratio (Rohde *et al.*, 2004), and is also deficient in anthocyanin and tannin biosynthesis (Huang *et al.*, 2010).

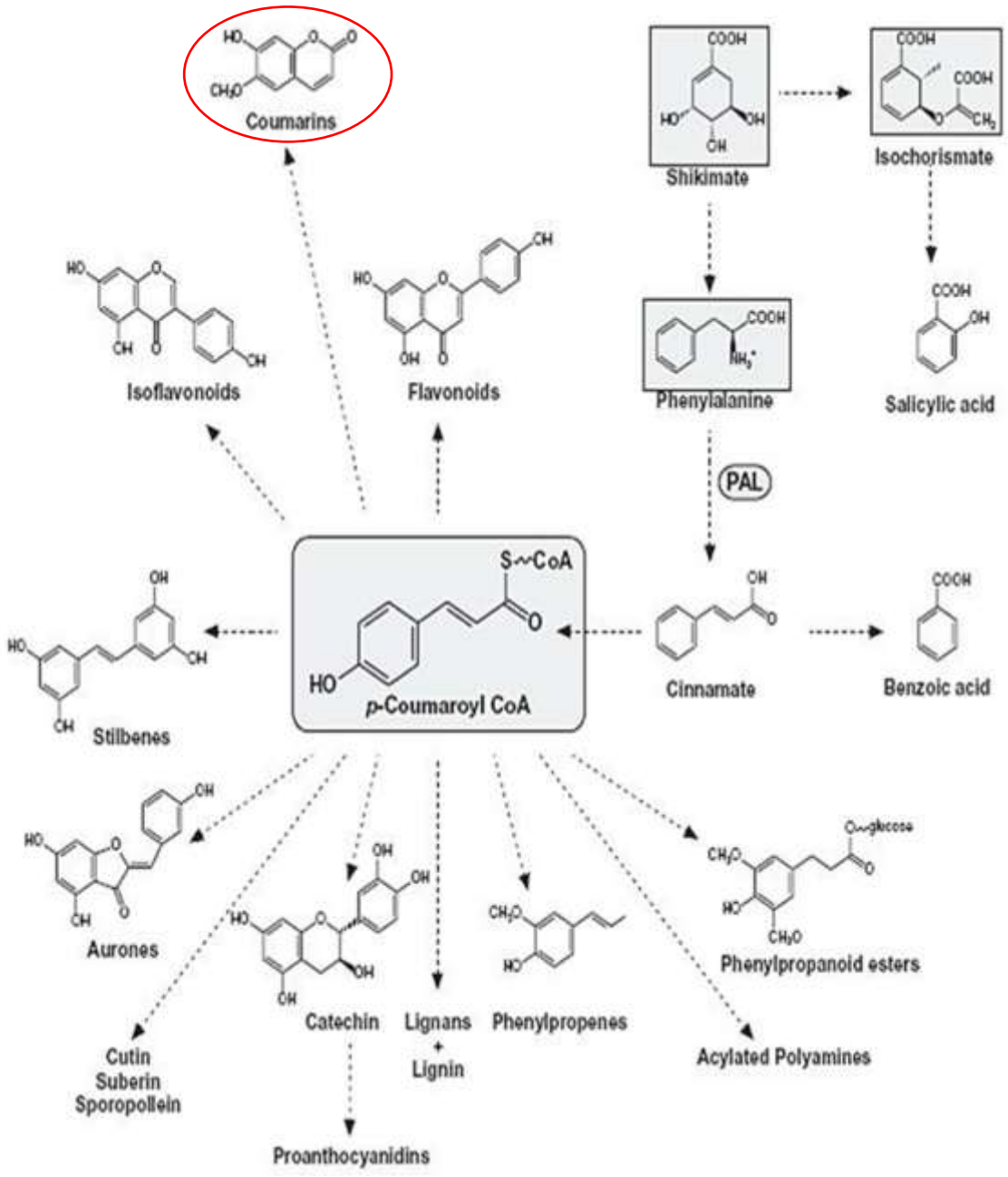


Figure 1.10 A simplified representation of the phenylpropanoid metabolism to show the various secondary metabolites that are synthesised through it, including scopoletin which is within the coumarin group (Vogt, 2010).

(*E*)-cinnamate is then hydrolysed by cinnamate 4-hydroxylase (C4H) which is a cytochrome P450-dependent monooxygenase (P450, CYP) to yield (*E*)-*p*-coumarate (also known as 4-coumarate). The P450 group is divided into families which are at least 40% identical to one another and sub-families of proteins which are at least 55% similar to one another, although there are cases where sequence identity among P450s can be less than 20% (Werck-Reichhart *et al*., 2002). The *A. thaliana* genome comprises of 272 cytochrome P450 genes and is one of the largest protein families in higher plants. C4H

is the first of three P450s in the phenylpropanoid metabolism that is involved in lignin biosynthesis and CYP73A5 is the only CYP73 family member in *A. thaliana* (Werck-Reichhart *et al*., 2002) which is why the knock-out mutant of that gene is inviable (Bell-Lelong *et al.*, 1997).

p-coumarate is then catalysed by 4-coumarate:CoA ligase (4CL) in an ATP dependent formation of the thioester bond between coenzyme A (CoA) and its carboxyl group to form *p*-coumaroyl CoA. There are four 4CL proteins, 4CL1-4CL4, in *A. thaliana* (Costa *et al.*, 2005). Studies show that the isozymes have different expression patterns and activities. 4CL1 and 4CL2 are 83% identical and they have similar expression profiles (Hamberger and Hahlbrock, 2004). 4CL3 is only 61% identical to 4CL1, shows higher expression in siliques and is induced by UV radiation and not pathogen attack like 4CL1 and 4CL2 (Ehlting *et al.*, 1999). Expression profiles and data from 4CL homologues in other plant species suggest that 4CL1 and 4CL2 are expected to be involved in lignin biosynthesis and 4CL3 is likely to be involved in flavonoid biosynthesis (Ehlting *et al.*, 1999). On the other hand, 4CL4 likely has a different metabolic function from the other three as it shows preferential activity towards ferulate and sinapate instead of *p*-coumarate (Hamberger and Hahlbrock, 2004).

Hydroxycinnamoyl-CoA shikimate:quinate hydroxycinnamoyl-transferase (HCT) is involved in two reactions in the phenylpropanoid metabolism. Firstly, it catalyses the transfer of the *p*-coumaroyl group in *p*-coumaroyl CoA to shikimate to produce *p*-coumaroyl shikimate (Hoffmann *et al.*, 2003). *p*-coumaroyl shikimate 3' hydroxylase (C3'H) then catalyses the hydroxylation of *p*-coumaroyl shikimate to produce caffeoyl shikimate (Franke *et al.*, 2002a). Caffeoyl shikimate is then catalysed by HCT again to transfer the caffeoyl moiety back to CoA to form caffeoyl CoA (Hoffmann *et al.*, 2003). HCT can use quinate instead of shikimate but not very well. *A. thaliana* mutants where the HCT gene has been silenced show a dwarf phenotype with accumulation of flavonoids and reduced S lignin in comparison to wild-type plants (Hoffmann *et al.*, 2004). C3'H is the second out of three P450s that is involved in lignin biosynthesis in the phenylpropanoid metabolism. It is also one of the three belonging in the CYP98 family (CYP98A3) in *A. thaliana* (Franke *et al.*, 2002a; Schoch *et al.*, 2001). C3'H is able to act on both the shikimate ester and the quinate ester of *p*-coumarate but the conversion rate for the shikimate ester is four times higher than the latter (Schoch *et al.*, 2001). *ref8* is a CYP98A3 mutant where a single transition mutation resulting in a substitution that greatly reduces C3'H activity (Franke *et al.*, 2002a). Like other *ref* mutants, *ref8* plants have lower levels of sinapoylmalate accumulation but in addition to that, *ref8* plants also

deposit H lignin instead of G and S lignin and are dwarfed and hyperaccumulate flavonoids (Franke *et al.*, 2002b).

Caffeoyl CoA 3-*O*-methyltransferase (CCoAOMT) catalyses the methyl-transfer of the 3'-hydroxyl group of caffeoyl CoA to yield feruloyl CoA. Although there are seven CCoAOMT-like genes in *A. thaliana*, only one is a functionally confirmed CCoAOMT gene (CCoAOMT1). T-DNA insertion mutants have lower stem lignin content and collapsed xylem vessels but no visible morphological differences compared to the wild-type (Do *et al.*, 2007). Coumarin biosynthesis is also significantly affected; the levels of scopoletin and scopolin compared to wild-type roots are approximately 30% and 15% lower respectively (Kai *et al.*, 2008).

Feruloyl CoA is then catalysed by cinnamoyl-CoA reductase (CCR) to produce hydroxycinnamaldehydes. CCR can act on caffeoyl CoA, sinapoyl CoA and feruloyl CoA but has the highest affinity for the latter. Two CCR proteins have been identified in *A. thaliana*, AtCCR1 and AtCCR2, and they are 80% identical (Baltas *et al.*, 2005). Although they are very similar, they have different expression profiles; AtCCR1 expression is detected in flowers and leaves and highly expressed in stems under normal conditions, but AtCCR2 is not detected in any tissue type until exposure to bacterial pathogens. When leaves are exposed to bacterial pathogens, AtCCR2 transcripts begin to steadily increase over a 12-hour period whereas AtCCR1 transcripts decrease. These findings suggest that AtCCR1 is involved in normal lignin biosynthesis and AtCCR2 is involved in pathogen-induced lignification (Lauvergeat *et al.*, 2001). AtCCR1 antisense transgenic plants are severely dwarfed, have approximately 50% less lignin than wild-type plants and have collapsed xylem vessels. These mutants also exhibit reduced levels of sinapoylmalate because of reduced levels of hydroxycinnamaldehydes and hydroxycinnamoyl alcohols and accumulate feruloylmalate due to the redirecting of feruloyl CoA into a potential thioester-based pathway (Derikvand *et al.*, 2008).

The third P450 involved in lignin biosynthesis in the phenylpropanoid metabolism is ferulate 5-hydroxylase (F5H) and is from the CYP84 family (Meyer *et al.*, 1996). F5H catalyses the 5-hydroxylation of coniferaldehyde and coniferyl alcohol to form 5-hydroxyconiferaldehyde and 5-hydroxyconiferyl alcohol respectively. F5H mutants fail to accumulate sinapoylmalate and deposit S lignin (Chapple *et al.*, 1992).

Caffeic acid *O*-methyltransferase (COMT) methoxylates the C5 position of the phenolic ring of aldehyde and alcohol precursors of S lignin such as 5-hydroxyconiferaldehyde and 5-hydroxyconiferyl alcohol to produce sinapaldehyde and sinapyl alcohol, respectively. The T-DNA knockout-mutant for the COMT gene (COMT1) in *A. thaliana*,

COMT1, contains lignin lacking in S units and reduced sinapoylmalate levels in stem, leaves and seedlings (Goujon *et al.*, 2003). COMT also plays a role in the biosynthesis of scopoletin through the alternative pathways in which esculetin is methoxylated to yield scopoletin (Alhalaseh, 2017).

Cinnamyl alcohol dehydrogenase (CAD) is responsible for the last step in lignin biosynthesis of G and S subunits. CAD catalyses the NADPH-dependent reduction of coniferaldehyde and sinapaldehyde to the monolignols coniferyl alcohol (G lignin) and sinapyl alcohol (S lignin) respectively. The primary CADs that are involved in lignin biosynthesis in *A. thaliana* are AtCAD4 and AtCAD5 which are both highly expressed in vascular bundles and several other tissue types (Kim *et al.*, 2007). The double mutant for AtCAD4 and AtCAD5 has a 94% reduction in G and S lignin monomers in comparison to wild-type plants and deposits lignin composed primarily of coniferaldehyde and sinapaldehyde subunits (Sibout *et al.*, 2003).

These findings show how significant the phenylpropanoid metabolism is to the survival of the plant as it is responsible for the biosynthesis of lignin, flavonoid, coumarins and many other important secondary metabolites.

1.5 RESEARCH STRATEGY

Cassava PPD is a complex phenomenon that is yet to be fully understood, but there is sufficient evidence suggesting that scopoletin plays an important and significant role in the development of cassava root deterioration. The understanding of the biosynthesis of scopoletin through the phenylpropanoid metabolism allowed Liu *et al.* (2017) to successfully manipulate genes to reduce scopoletin biosynthesis, though not fully inhibited, by silencing F6'H genes using RNAi constructs. This study, however, was insufficient to tackle the entire problem as only the major biosynthetic pathway of scopoletin was addressed and the alternative pathways may have played a compensatory role. Besides that, due to having seven members in the F6'H gene family in cassava, the RNAi construct may not be able to target all the F6'H genes leading to some scopoletin being synthesised through the major pathway.

To fully understand the biosynthesis of scopoletin and block its production with the aim of extending the shelf-life of cassava, this study intends to further the understanding into the function and gene expression patterns of the enzymes responsible for the biosynthesis of scopoletin (Figure 1.11).

1. The initial phase of the investigation will be to identify homologous genes of PAL, C4H, 4CL, HCT, C3'H, CCoAOMT and COMT in the cassava genome using functionally confirmed reference genes from *A. thaliana*. By understanding the gene expression of these genes, candidate genes for CCoAOMT and COMT will be taken forward for RNAi silencing along with F6'H.

2. As the F6'H single knock-down transgenic lines have already been generated by Liu *et al.* (2017), transgenic cassava with a double RNAi construct for CCoAOMT and F6'H and a triple RNAi construct for CCoAOMT, F6'H and COMT for simultaneous gene silencing will be generated.

3. The storage roots of all three transgenic lines will be harvested and PPD assessment will be performed to analyse the rate of PPD development. Lines from each RNAi construct that show the slowest rate of PPD development will be analysed for gene expression and changes in metabolites.

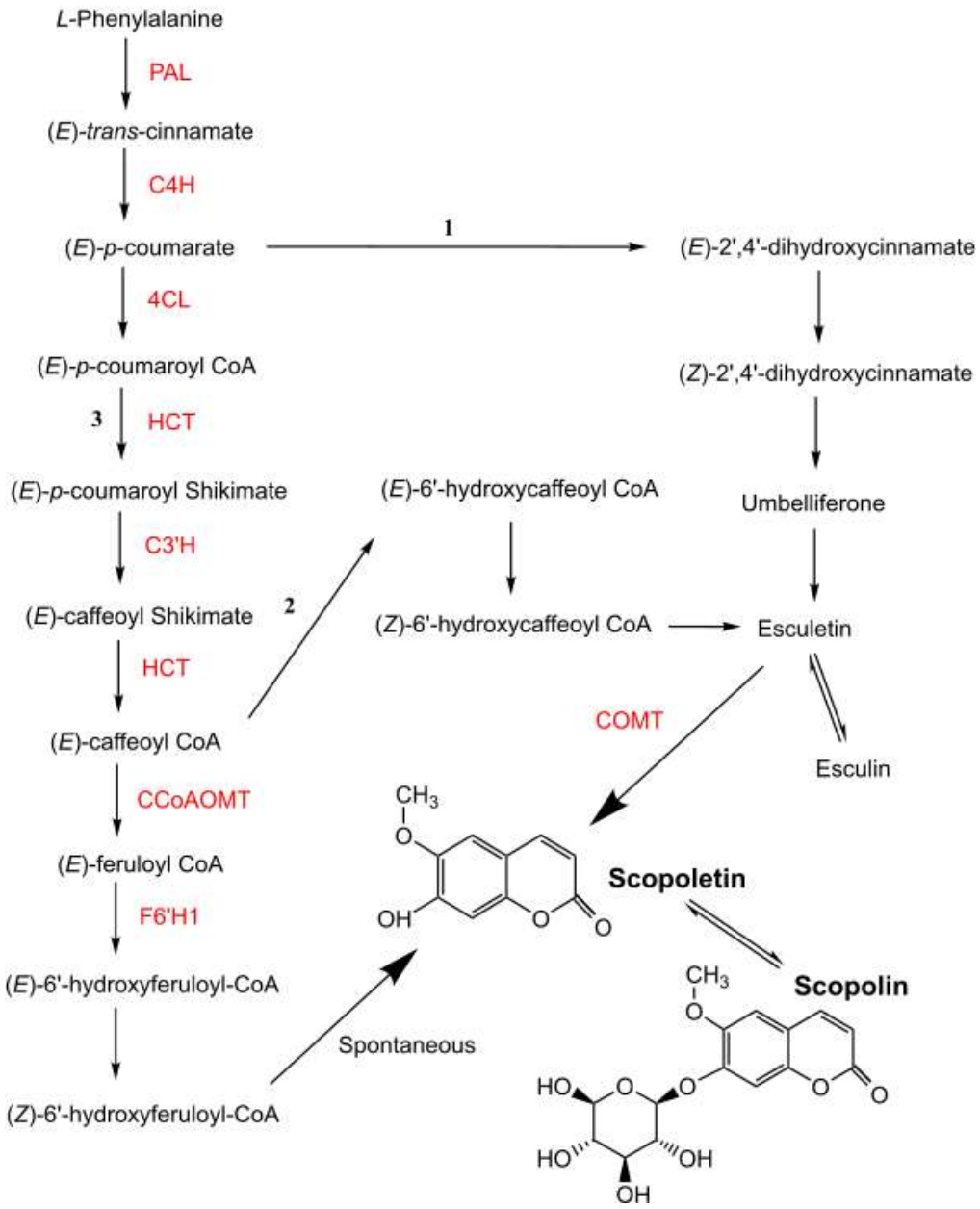


Figure 1.11 **The biosynthesis of scopoletin through three alternative pathways in the phenylpropanoid metabolism.**

2 MATERIALS AND METHODS

2.1 INSTRUMENTS

Unless stated otherwise, all reagents and chemicals used were obtained from Sigma-Aldrich and all plastic consumables were obtained from Fisher Scientific. A Thermo Cycler PTC-200 (MJ Research) was used for PCR and other reactions which required incubation. The PowerPac 200 (Bio-Rad) was used for gel electrophoresis of DNA and RNA samples and the GDS 7500 UV transilluminator (UVP) with a Grab-IT 2.0 (Synoptics Ltd.) imaging system was used to visualise the gel and capture images. Different centrifuges were used depending on the requirement of the sample, e.g. speed and volume. A MiniSpin microfuge (Eppendorf) was used to centrifuge samples with a volume less than 2 mL; an Allegra 25R centrifuge (Beckman Coulter) was used for high speed centrifugation; and a Universal 32R centrifuge (Hettich) was used for low-speed centrifugation. A Helios spectrophotometer (ThermoFisher Scientific) was used to measure the absorbance of cell cultures and a Nanovue spectrophotometer (GE Healthcare Life Sciences) was used to quantify nucleic acid samples. *Arabidopsis thaliana* plants and *in vitro* cassava were grown in a Fitotron plant growth chamber (Weiss Gallenkamp). The electro-transformation of *E.* coli and *Agrobacterium* was carried out through a MicroPulser electroporator (Bio-Rad). Petri dishes were incubated in a Function Line microbiological incubator (Heraeus) and bacterial cultures that required shaking were incubated in an Orbital Incubator (Sanyo). The film developer OPTIMAX X-Ray Film Processor Model 1170-1-0000 (PROTEC GmbH & Co KG, Germany) was used to develop southern blot films.

2.2 PLANT MATERIALS

Wild-type cassava cultivar TMS60444 (tropical *Manihot* series 60444) was obtained from the glasshouse of Department of Biology and Biochemistry, University of Bath. Friable embryogenic calli (FEC) cultivar TMS60444 were obtained from the cassava research group from the Institute of Molecular Plant Biology in ETH Zürich.

2.3 BACTERIA STRAINS

Library Efficiency™ DH5α™ Competent Cells (Invitrogen) were used in *Escherichia coli* transformation. *Agrobacterium tumefaciens* line LBA4404 (disarmed pAL4404 Ti plasmid; genomic antibiotic: Rifampicin; Ti plasmid antibiotic: Tetracycline) (Hoekema *et al.*, 1983; Negrotto *et al.*, 2000) was obtained from ETH-Zürich and electro-competent cells were made in the lab.

2.4 PLASMID

The RNAi vector pRNAi-GG was made by Yan *et al.* (2012) and was purchased from the Arabidopsis Biological Resource Center (ABRC) at the Ohio State University. The vector was made using the binary vector pBI121 as the skeleton. It contains the binary borders for Agrobacterium-mediated transformation, a kanamycin resistant gene to select for bacteria that carry it and neomycin phosphotransferase (nptII) gene to select transformed plants. A constitutive cauliflower mosaic 35S promoter (CaMV 35S) regulates the expression of the T-DNA region.

2.5 SOFTWARE AND STATISTICAL METHODS

The sequences of cassava and *Arabidopsis thaliana* genes were obtained from Phytozome (http://phytozome.jgi.doe.gov/pz/portal.html) and National Center for Biotechnology Information (NCBI). BLAST analyses to compare similarities between the gene sequences were carried out in Phytozome (Goodstein *et al.*, 2012). PROSITE (https://prosite.expasy.org) was used to identify protein domains and functional sites for cassava homologous genes. Primers for PCR, qPCR and sequencing were designed with NCBI primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and IDT oligo-Analyzer (https://www.idtDNA.com/calc/analyzer) (Ye *et al.*, 2012). The alignments, comparison and phylogenetic tree reconstruction of sequences were carried out using Geneious R8 (Kearse *et al.*, 2012). The PPD discolouration of cassava roots was quantified and analysed with a MatLab-based programme designed at ETH-Zurich (Vanderschuren *et al.*, 2014). StepOne Software v2.3 by Applied Biosystems (UK) was used in qRT-PCR for gene expression quantification.

All statistical analyses were carried out using R studio Version 1.1.442 and graphical visualisations were made with SigmaPlot v14 (Systat Software Inc). Statistical methods used were Shapiro-Wilk test to check for normality and Levene's test to check for equal

variance. A one-way ANOVA analysis followed by a Tukey's comparison was performed for data that met the assumptions.

2.6 PRIMERS

Primers listed in Table 2.1 were used in various experiments such as gene expression analysis for specific genes, quantitative gene expression analysis through qRT-PCR, southern blots and sequencing.

Table 2.1 **List of primers**.

Primer name	Primer sequence (5' to 3')
Gene specific primers for RT-PCR	
Cassava4.1_034377_Fwd	ACTTCACCTGGATTTTCCTCTGC
Cassava4.1_034377_Rev	TTGACCCCAGCTCGAGCCT
Cassava4.1_002628_Fwd	GGCAACAATCTCTCAGAATGGTC
Cassava4.1_002628_Rev	AAGGTCTCGCCGCCTAACTTG
Cassava4.1_002709_Fwd	CACTTTCCTTCCACCTTTCCCA
Cassava4.1_002709_Rev	GCCAAAGCCAGTGGTGATG
Cassava4.1_028434_Fwd	ACGGTGAGAATGAGAAGAATCCTAGTAG
Cassava4.1_028434_Rev	AGCAACCCAGTGCCCACC
Cassava4.1_002591_Fwd	AGTGGGAATCCTGCAGTC
Cassava4.1_002591_Rev	CTACTGCAATGGAACCAAAGTG
Cassava4.1_003117_Fwd	CAAGAGGTTCTCGTCAAAAGG
Cassava4.1_003117_Rev	AGCACTTCAACAAAACATCTCC
Cassava4.1_005978_Fwd	AGCTTGATACAGTGCTTGGACC
Cassava4.1_005978_Rev	TTCTGCACCAAACGTCCCAAAG
Cassava4.1_005006_Fwd	GAGTCAGAGCTCCATAATCACTG
Cassava4.1_005006_Rev	GTGATTATGAGTTTCGTGTTAGACC
Cassava4.1_005014_Fwd	CAGAGCTCCCTCATCATCTT
Cassava4.1_005014_Rev	CCAAGAAATGAAAGAACGAATTGG
Cassava4.1_004658_Fwd	ACAAGCTGCTTCAGACACAGGA
Cassava4.1_004658_Rev	TGGACAACCCAGCAGCACA
Cassava4.1_004136_Fwd	TGACATCTATAGCCTCTGAAACC
Cassava4.1_004136_Rev	GGTTGGATAATCCAGCAGCAGT
Cassava4.1_027178_Fwd	AGTTTACGCTGAGAAAGTGAAGCAA
Cassava4.1_027178_Rev	CAGAACGATCGGAGGCACGAA
Cassava4.1_008045_Fwd	GCTCGCGGTCTTGACATTACTG
Cassava4.1_008045_Rev	CCAGGTTGGCTTTGATTGCAGG
Cassava4.1_008063_Fwd	TCCTCCATCCGATCTCCCACTC
Cassava4.1_008063_Rev	TGCACCAGGGCCTTACTAAGCC
Cassava4.1_005910_Fwd	TAGATTCCGATGCTACGCCGAG
Cassava4.1_005910_Rev	GTTCACGAACCGCTTGCCAAATG
Cassava4.1_014783_Fwd	GAATGGTTCTGTGGTGGCAG
Cassava4.1_014783_Rev	AGAGGTCCAATTACACACGTGG

qRT-PCR primers	
qPCR_CCoAOMT_Fwd	
qPCR_CCoAOMT_Rev	
qPCR_F6'H_Fwd	TGGGTTCATGTTCCTCCCAT
qPCR_F6'H_Rev	CTCCTATATCGACCATTGCTGAGTA
qPCR_COMT_Fwd	
qPCR_COMT_Rev	
qPCR_ubq10_Fwd	CACCGGATCAGCAAAGGCTTA
qPCR_ubq10_Rev	CAGACACACAGATCAAAGCAGC
Southern Blot Probes	
SB_Probe_Fwd	ACTAGATCGGGCCTCCTGTC
SB_Probe_Rev	CAAAAGGGCGACATTCAACCG
Other Primers	
PP2A_Fwd	TGCAAGGCTCACACTTTCACT
PP2A_Rev	CTGAGCGTAAAGCAGGGAAG
M13_Fwd	GTAAAACGACGGCCAGT
M13_Rev	AACAGCTATGACCATG

2.7 BACTERIA AND PLANT MANIPULATION

2.7.1 ANTIBIOTICS

Table 2.2 contains all relevant information about the working conditions of the antibiotics.

Table 2.2 **Antibiotics and their properties.**

Antibiotic	Solvent	Sterilisation	Stock concentration
Ampicillin	MilliQ water	Filter sterilisation	100 µg/mL
Carbenicillin	EtOH	N/A	50 µg/mL
Chloramphenicol	MilliQ water	Filter sterilisation	30 µg/mL
Geneticin	MilliQ water	Filter sterilisation	25 ug/mL
Kanamycin	MilliQ water	Filter sterilisation	100 µg/mL
Streptomycin	MilliQ water	Filter sterilisation	100 µg/mL
Rifampicin	DMSO	N/A	50 µg/mL

2.7.2 BACTERIAL CULTURES

E. coli colonies were grown on LB agar plates (5 g/L yeast extract, 10 g/L tryptone, 10 g/L NaCl and 16 g/L agar, pH = 7.0) with appropriate antibiotics at 37°C overnight (16 hours). Single colonies were picked with a sterile 10 uL pipette tip and inoculated in liquid LB medium (5 g/L yeast extract, 10 g/L tryptone and 10 g/L NaCl, pH = 7.0) with

appropriate antibiotics. The culture was kept overnight (16 hours) at 37°C in a shaking incubator (200 rpm).

A. tumefaciens colonies were grown on LB agar plates with the same composition as above with appropriate antibiotics at 28°C for 2 days in the dark. Single colonies were picked with a sterile 10 uL pipette tip and inoculated in liquid LB medium with the same composition as above with appropriate antibiotics. The culture was kept in the dark for 2 days at 28°C in a shaking incubator (200 rpm).

2.7.3 PREPARATION OF ELECTRO-COMPETENT *A. TUMEFACIENS*

A single colony of *A. tumefaciens* LBA4404 was picked with a sterile 10 uL pipette tip and inoculated in 30 mL of liquid LB medium with Streptomycin (100 ug/mL) and Rifampicin (50 ug/mL). The culture was incubated overnight in the dark at 28°C, 200 rpm. The following day, 1 mL of the overnight culture was inoculated in 75 mL of liquid LB medium with the same antibiotics. The culture was incubated overnight (18 hours) in the dark at 28°C, 200 rpm. The following morning, the cell density was monitored by reading the absorption at 600 nm using a spectrophotometer every 30 min. Once OD600 was reached 0.7 – 1.0, 50 mL of the culture was transfered to a 50 mL conical tube and centrifuged in a Universal 32R centrifuge (Hettich) at 4000 rpm at 4°C for 10 min. The supernatant was discarded, and the cell pellet was re-suspended in 25 mL of ice-cold sterile MilliQ water to wash it. The cells were centrifuged (4°C, 4000 rpm, 10 min) again and the supernatant was discarded. The re-suspension and precipitation were repeated twice, and the cells were collected by centrifuging under the same conditions. 10 mL of ice-cold sterile 10% glycerol solution was added the re-suspend the cell pellet. The glycerol suspension was pelleted by centrifugation under the same conditions and the supernatant was discarded. The cell pellet was then re-suspended in 300 uL of ice-cold sterile 10% glycerol solution and was mixed by pipetting gently. 50 uL of suspended cell solution was then aliquoted to sterile pre-chilled 1.5 mL tubes and snap-frozen in LN_2 and stored at -80°C.

2.7.4 TRANSFORMATION OF CHEMICALLY COMPETENT *E. COLI*

Commercial chemically competent DH5α cells (Library Efficiency™ DH5α™ Competent Cells) (Invitrogen) were used. The protocol used was provided by Invitrogen, nothing was amended.

2.7.5 TRANSFORMATION OF ELECTRO-COMPETENT *A. TUMEFACIENS*

The lab prepared *A. tumefaciens* electro-competent cells were transformed through electroporation. An aliquot of 50 uL cells were thawed on ice and mixed with 15 ng of purified plasmid DNA. The mixture was left to stand for 2 min on ice and then transferred to a pre-chilled ice-cold electroporation cuvette (Bio-Rad, 1 mm gap) and electroporated with one single pulse at 2.5 kV with a MicroPulser (Bio-Rad). Immediately after that, 1 mL of liquid LB medium was added to the cuvette and mixed with a pipette before transferring to a sterile 15 mL test tube. The culture was incubated at 28°C at 220 rpm for three hours before spreading on LB agar plates with appropriate antibiotics. The plate was incubated in the dark at 28°C for 48 hours to allow colonies to grow.

2.7.6 SUBCULTURE AND GROWTH OF CASSAVA

Wild-type and transgenic cassava plants were grown from *in vitro* plants. Buds with 8-10 mm of stem were cut from *in vitro* cassava plants and subcultured in sterile pots with CBM (4.4 g/L Murashige-Skoog medium with vitamin, 20 g/L sucrose, 2 µM CuSO4, and 3 g/L Gelrite (Duchefa), pH = 5.9). *In vitro* plants were grown in a controlled environment suite (24°C, relative humidity 50%, 16 hours daylight, Weiss-Gallenkamp) and subcultured every 6 months. Cassava lines to be transferred to soil were subcultured into CBM pots with 2.5 g/L Gelrite and grown *in vitro* for 7-10 days for the root and stem to develop. The subcultured plantlets were transferred to soil (three parts of Levington's M2 compost and one part of perlite). The cassava plants were then kept and grown in the University glasshouse with supplemented lighting at 30°C during the day and 17°C during the night.

2.7.7 CASSAVA TRANSFORMATION

Cassava transformation was carried out on friable embryogenic calli (FEC) with several modifications following the methods from Bull *et al.* (2009). Modifications to the protocol was the use of the plasmid pRNAi-GG instead of pCambia 1305.1. Due to this, paromomycin and geneticin were used as the plant selectable markers instead of hygromycin. The concentration of paromomycin in the GD medium used during the recovery and maturation of transformed FEC was 27.5 uM for 3 weeks. After that, 25 uM of geneticin was used for the regeneration of transgenic plants. The concentration of carbenicillin was kept the same at 250 uM for both the maturation and regeneration

steps. Preliminary screening using the rooting experiment was not done. Transgenic plants were identified through PCR genotyping with primers designed specific to the RNAi construct.

2.8 NUCLEIC ACID MANIPULATION

2.8.1 CASSAVA LEAF DNA EXTRACTION

The DNA extraction protocol without the use of chloroform from Huang *et al.* (2013) was used. 100 mg of cassava leaf sample was ground up to powder with liquid nitrogen (LN_2) in a 1.5 mL Eppendorf tube. 1 mL of lysing buffer (2% (w/v) CTAB, 200 nM Tris, 2 M NaCl, 25 mM EDTA, 1% (w/v) LSS, 20 mM Borax, pH = 8.0) that was preheated at 65°C was added to the tissue powder and mixed thoroughly by vortexing. The homogenate was incubated for 30 min at 65°C and vortexed every 10 min. The homogenate was then centrifuged at 13,400 rpm for 10 min in a MiniSpin microfuge (Eppendorf) at room temperature (RT). 500 uL of supernatant was transferred to a clean 1.5 mL Eppendorf tube and mixed with 500 uL of 70% isopropanol. The mixture was stored at -20°C for at least 30 min and then centrifuged at 13,400 rpm for 10 min at RT to pellet nucleic acids. The supernatant was discarded, and the pellet was washed with 1 mL of 75% EtOH (4°C). The supernatant was discarded, and the pellet was washed another time with 100% EtOH. The pellet was air-dried at room temperature for 10 min and suspended in 50 uL of MilliQ water. The DNA concentration and quality were checked with a NanoVue spectrophotometer (GE).

2.8.2 SOUTHERN BLOT

2.8.2.1 PREPARATION OF DIG-LABELLED PROBE

Specific primers (Table 2.1) for the probe were designed from the T-DNA region of the plasmid pRNAi-GG. The probe was prepared in a 0.2 mL tube and the reaction was carried out in a Thermo Cycler PTC-200 (MJ Research). The composition of the probe and the reaction conditions are listed in Tables 2.3 and 2.4, respectively. The PCR product was run on a 1.2% (w/v) agarose gel electrophoresis for 60 min at 60 V and purified using the protocol provided by QIAquick Gel Extraction Kit (Qiagen). The gel-purified probe was eluted in 50 μL of sterile MilliQ water and stored at -20°C until use.

Table 2.3 **Composition of reaction for the probe**

Component	Volume (uL)
10x standard Taq reaction buffer	5
10x DIG dNTPs	5
Taq DNA polymerase (5 U/uL)	1
10 µM probe-forward primer	0.5
10 µM probe-reverse primer	0.5
15ng/ul Plasmid DNA (pRNAi-GG)	1
Nuclease-free water	33
Total volume	50

Table 2.4 **Probe PCR reaction conditions**

Step	Temperature (°C)	Time	
Initial denaturation	94	2 min	
Denaturation	94	30 sec	**30 Cycles**
Annealing	57	30 sec	
Elongation	72	1 min	
Final elongation	72	5 min	
Final hold	12	Forever	

2.8.2.2 REAGENTS USED FOR SOUTHERN BLOTTING

Reagents used in southern blotting are 20X SSC, 20% SDS, depurinating solution, denaturing solution, neutralisation solution, washing solution 1 (W1), washing solution 2 (W2), washing solution 3 (W3), washing buffer (WB), B1 solution, B2 solution and B3 solution (McCabe *et al.*, 1997). The compositions of the reagents are available in Appendix I.

2.8.2.3 SOUTHERN BLOT

Day 1: Digestion reaction

10 ug of cassava genomic DNA was digested with the restriction enzyme *Hind*III (Promega) in a 0.2 mL tube with a total volume of 150 uL at 37°C overnight in a Thermo Cycler PTC-200 (MJ Research). The master mix composition for the digestion reaction is listed in Table 2.5. pRNAi-GG plasmid DNA was used as a positive control.

Table 2.5 **Master mix composition for the digestion using *Hind*III**

Component	Volume (uL)
Cassava genomic DNA (500 ng/uL)	20
Promega Buffer E	15
10x BSA	5
*Hind*III (10 U/uL)	10
MilliQ water	10
Total volume	150

Day 2: Ethanol Precipitation

After the overnight digest, the reaction was heat-inactivated at 80°C for 20 min. The digested sample (150 uL) was transferred to a 1.5 mL tube (Eppendorf) and 15 uL of NaOAC (pH 4.8) and 350 uL of 95% (v/v) ethanol were added to the mixture and incubated at -20°C for 2 hours. The mixture was then centrifuged at 13,000 rpm, RT for 15 min and the pellet was washed with 1 mL of 70% (v/v) ethanol. The pellet was air-dried at RT for 10 min and then resuspended in 15 uL of MilliQ water. The resuspended pellet was incubated at 65°C for 10 min to allow the pellet to dissolve and was vortexed and spun down.

Agarose gel electrophoresis

A DIG-labelled ladder was prepared (2 uL of DNA Molecular Weight Marker VII, DIG-labelled (Sigma-Aldrich), 5 uL of 5x loading dye and 13 uL of MilliQ water). 5 uL of 5x loading dye was added to 15 uL of digested DNA to make a total volume of 20 uL to be loaded onto the gel. The samples and DIG-labelled ladder were run overnight on a 0.8% (w/v) agarose gel (no EtBr) at 1 V/cm in Tris-Borate-EDTA (TBE) as the running buffer.

Day 3: DNA visualisation under UV light

The gel was stained with 0.3 mg/ml Ethidium Bromide for 15 min and then de-stained in water for 20 min. The gel was observed under UV light for DNA visualisation and documentation.

Capillary transfer of DNA to membrane

The gel was cut at the bottom right corner to mark its orientation. The gel was incubated in 300 mL of depurination buffer for 30 min on a shaker (The Belly Dancer, Stovall Life Science). It was then washed twice with sterile MilliQ water and incubated in 300 mL of denaturation buffer for 30 min on the shaker. The gel was washed twice with sterile MilliQ water and was incubated in 300 mL of neutralisation buffer for 15 min twice. The gel was set up for overnight capillary transfer of DNA to a nylon membrane (GE Healthcare, Life Sciences) (Southern, 2006).

Day 4: Hybridisation with DIG-labelled probe

The membrane was cut at the bottom right to mark its orientation (same corner as the gel). The membrane was placed onto fresh Whatman paper and dried at RT for 5 min and then placed under 254 nm UV light for 2 min to undergo UV-crosslinking. The membrane was placed in a hybridisation tube and 10 mL of DIG easy hybridisation solution (Rohce) was added. The membrane was pre-hybridized at 42°C for 4.5 hours (minimum of 1 hour). The DIG easy hybridisation solution was discarded and 10 mL of fresh DIG easy hybridisation solution containing 50 µL of pre-heated (95°C, 5 min) DIG-labelled probe was added to the tube and incubated at 42°C overnight.

Day 5: Washing membrane

The DIG easy hybridisation solution containing the probe was poured into a 15 mL tube and stored at -20°C. All washing and incubation steps were done on a shaker. The membrane was placed into a clean container and washed 3 times (3x) with 300 mL W1 solution for 15 min each, followed by 300 mL of pre-warmed (68°C) W2 for 15 min and then 300 mL of pre-warmed (72°C) W3 solution for 30 min. The container was changed with each wash. The membrane was then incubated in 300 mL of WB solution for 3 min and then in 160 mL of B2 solution for a minimum of 1 hour. The B2 solution was discarded and fresh B2 solution with 1.5 µL of antibodies (Rohce) was added to incubate the membrane for 30 min. The membrane was then washed 5x with WB solution (1x for 5 min, 2x for 10 min, 1x for 1 hr and 1x for 30 min). After washing, the membrane was incubated in 93 mL of B3 solution for 5 min, and then overlaid with 7 mL of B3 solution containing 40 µL of CDP star (Rohce) for 5 min. The membrane was dried on Whatman paper for 5 min at RT. Once the membrane was completely dry, it was sealed with Saran plastic wrap (DowBrands, USA) and put in a film cassette (Kodak BioMax light film, Sigma Aldrich) with a film and was incubated at 37°C overnight.

Day 6: Developing film

The film was developed in a dark room using OPTIMAX X-Ray Film Processor Modell 1170-1-0000 (PROTEC GmbH & Co KG, Germany).

2.8.3 CASSAVA RNA EXTRACTION

The Spectrum™ Plant Total RNA Kit (Sigma-Aldrich) was used to extract cassava storage root RNA. 50 mg of root sample was ground up with LN_2 in a pestle and mortar that were pre-baked at 200°C to denature RNase. 1000 uL of the lysis solution and 2-mercaptoethanol mixture were used per sample that was then incubated at room temperature for 3 minutes. To bind the RNA to the column, Protocol A was used and due

to the low levels of RNA, the binding solution was increased to 750 uL. To remove residual genomic DNA from the RNA samples, on-column DNAse digestion was performed using the On-Column DNAse I Digest Set (Sigma-Aldrich) as per protocol provided by the kit. The concentration and quality of the RNA samples were verified with a NanoVue spectrophotometer (GE) and checked on gel for degradation. If the RNA concentration was too low (<15 ng/uL), the purified RNA samples was vacuum-concentrated using a Savant SpeedVac vacuum concentrator (Thermo Fisher). The RNA samples were then stored at -80°C until further use.

2.8.4 cDNA PREPARATION

The High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems) with oligo dT (18) primer (New England Biolabs) were used to synthesise cDNA from RNA samples. 150 ng of total RNA was used in each reaction and the protocol provided was followed with an additional 1 uL of 50 uM oligo dT (18) primer in each reaction. The reaction was incubated in a Thermo Cycler PTC-200 (MJ Research).

2.8.5 *E. COLI* PLASMID EXTRACTION

A single *E. coli* colony was inoculated in 5 mL of sterile LB medium with appropriate antibiotic. The culture was incubated in a shaking incubator overnight at 37°C, 200 rpm to allow the cells to multiply. The overnight bacterial culture was centrifuged at 2000 rpm to pellet the cells in a Universal 32R centrifuge (Hettich). The supernatant was discarded, and the plasmid extracted from the bacterial cell pellet with the QIAprep Spin Miniprep Kit (Qiagen) and eluted with 35 uL to 50 uL of sterile MilliQ water. The quality and concentration of the plasmid DNA were verified with a NanoVue spectrophotometer (GE).

2.8.6 POLYMERASE CHAIN REACTION (PCR)

Taq DNA Polymerase and Q5® Hot Start High-Fidelity DNA Polymerase were both used, and they were obtained from New England Biolabs (NEB). All reactions were carried out in a Thermo Cycler PTC-200 (MJ Research). Taq DNA polymerase was used for genotyping and screening to determine the presence or absence of a target DNA sequence (cDNA or genomic DNA). The PCR preparation and reaction set up are listed in Table 2.6 and 2.7. When amplifying a DNA fragment for vector construction and sequencing, Q5® Hot Start High-Fidelity DNA Polymerase was used due to its proof-

reading ability to minimise errors during amplification. The PCR preparation and reaction set up are listed in Table 2.8 and 2.9.

Table 2.6 **Taq DNA polymerase Reaction Mixture**

Component	Volume (25 µl reaction)	Final concentration
10x standard Taq reaction buffer	2.5 µl	1x
10 mM dNTPs	2 µl	200 µM each
Taq DNA polymerase	0.125 µl	0.025 U/µl
10 µM forward primer	1 µl	0.4 µM (0.05-1 µM)
10 µM reverse primer	1 µl	0.4 µM (0.05-1 µM)
Template DNA	1 µl	< 1000 ng
Nuclease-free water	to 25 µl	

Table 2.7 **Taq DNA Polymerase Set-up**

Step	Temperature (°C)	Time	
Initial denaturation	94	2 min	
Denaturation	94	30 sec	**30 Cycles**
Annealing	50-65	30 sec	
Elongation	72	1 min/kb	
Final elongation	72	5 min	
Final hold	14	Forever	

Table 2.8 **Q5 High Fidelity DNA Polymerase Reaction Mixture**

Component	25 µL Reaction	50 µL Reaction	Final concentration
5x Q5 Reaction Buffer	5 µl	10 µl	1x
10 mM dNTPs	2 µl	4 µl	200 µM each
Q5 High-Fidelity DNA polymerase	0.25 µl	0.5 µl	0.02 U/µL
10 µM forward primer	1.25 µl	2.5 µl	0.5 µM (0.05-1 µM)
10 µM reverse primer	1.25 µl	2.5 µl	0.5 µM (0.05-1 µM)
Template DNA	1.25 µl	2.5 µl	< 1000 ng
Nuclease-free water	to 25 µl	to 50 µl	

Table 2.9 **Reaction set up using Q5 High-Fidelity DNA polymerase**

Step	Temperature (°C)	Time	
Initial denaturation	94	2 min	
Denaturation	94	30 sec	**30 Cycles**
Annealing	50-65	1 min	
Elongation	72	30-40 sec/kb	
Final elongation	72	5 min	
Final hold	14	Forever	

2.9 PLASMID MANIPULATION

2.9.1 DNA RESTRICTION ENDONUCLEASE DIGESTION

Restriction enzymes from Promega were used unless stated otherwise. The protocol for the restriction digest provided by the manufacturer was used and up to 1 ug of DNA was used in each reaction unless stated otherwise.

2.9.2 DNA ELECTROPHORESIS DNA GEL PURIFICATION

Agarose electrophoresis was used to analyse nucleic acid samples. The gel was prepared by dissolving agarose (Bioline) in 70 mL or 120 mL (depending on the gel rack used) of 1x TAE buffer (40 mM Tris-Acetate and 1 mM EDTA) in the microwave oven. The solution was cooled and mixed with ethidium bromide solution to a final concentration of 0.25 mg/mL. The solution was poured into a sealed gel rack with appropriate combs and left for 20 min to solidify. The gel was placed in an electrophoresis tank with 1x TAE buffer and nucleic acid samples mixed with 1x loading dye (8% sucrose and 0.05% bromophenol blue) were loaded into the wells. The voltage used and running time depended on the size of nucleic acid fragments. The gel was visualised with a GDS 7500 UV transilluminator (UVP) and Grab-IT 2.0 imaging software (Synoptics Ltd).

2.9.3 DNA GEL PURIFICATION

A 0.8% gel was used and the gel was run at 70 V for 60 min for DNA gel purification. The DNA band was excised from the gel with a sterile razor and kept in a sterile microfuge tube. The QIAquick Gel Extraction Kit (Qiagen) was used to extract the DNA from the gel following the protocol from the manufacturer. The DNA was eluted in 50 µL of MilliQ water.

2.9.4 RNAi CONSTRUCT

The plasmid pRNAi-GG (Yan *et al.*, 2012) was used to create the RNAi constructs following the protocol provided. Further information in chapter 4.

2.9.5 DNA SEQUENCING

Plasmids were purified from *E. coli* (refer to 2.8.5). 15 uL of plasmid DNA samples with a concentration of 70 – 100 ng/uL and 15 uL of appropriate primers (10 pmol/µl) were

sent to Eurofins for sequencing. The results were analysed with the Geneious R8 software.

2.10 PPD ASSESSMENT

2.10.1 CASSAVA HARVEST

Cassava plants were grown in the glasshouse for 9 months and then harvested for the storage roots. 3 biological replicates were selected for each line with 1-4 individual plants per replicate. The selected cassava plants were measured for their height (cm) and weight of roots per plant (g). After the height of the plant had been measured, the fibrous roots were removed, and the storage roots were washed in the glasshouse to remove soil before transferring to the lab. The stem and leaves of the plants were kept in -80ºC for further analyses. The storage roots were then weighed (g). The storage roots were then immersed in 10% (v/v) bleach solution for 10 minutes for sterilisation to minimise bacterial and fungal contamination during the PPD induction process.

2.10.2 PPD INDUCTION IN CASSAVA ROOTS

Cassava PPD was induced using the method developed by Buschmann *et al.* (2000b), where the roots were sliced into discs with a thickness of 10-15 mm. The sliced root discs were placed in a Petri dish (at least 3 slices per petri dish) lined with a single dry filter paper (Whatman). This was done in the laminar hood and all surfaces and equipment were sterilised with 70% ethanol to prevent fungal and bacterial contamination. The plates with the root samples were then incubated in a growth chamber with the lid on in the dark at 25 ± 2°C but no control over relative humidity. PPD evaluation was carried out on 3 biological replicates with 3 technical replicates each. Samples were collected at 4 timepoints postharvest: day 0 (fresh), day 2, day 4 and day 6. To determine the PPD level, the roots were photographed and the discolouration score measured with a MatLab imaging program based on the grayscale of the image (Vanderschuren *et al.*, 2014). The root discs had the dry areas sliced off and photographed (Pentax K20D) with the following settings: Exposure time: 1/180 sec, ISO-speed: 400, Aperture: f/5.6, Focal length: 50, Metering mode: Centre Weighted Average, Flash mode: OFF, White balance: manual.

2.10.3 PPD MEASUREMENT

The MatLab program (Vanderschuren *et al.*, 2014) measures PPD level by converting the colour images of the user selected area (parenchyma tissue) to grey scale and analysing the frequency distribution of the selected area to generate a histogram with two sets of values as quantile 2.5% and 97.5%. The program then calculates the PPD score based on the formula:

PPD score = (Quantile 97.5% - Quantile 2.5%) / Quantile 97.5%

2.11 BIOCHEMICAL ANALYSIS

2.11.1 SAMPLE PREPARATION

500 mg of cassava sample was ground to fine powder with LN_2 with a pestle and mortar and the powder was transferred into a 5 mL screw-cap glass vial (Fisher). 2.0 mL of 99.9% (v/v) ethanol (HPLC grade) containing 4-methylumbelliferone (4-MU), the first internal control, was added to the ground samples and incubated for 16 hours overnight on a shaker (IKA-VIBRAX-VXR) at room temperature at 10 rpm. 1.2 mL of the extract (liquid phase) was filtered through a 13 mm, 0.22 um PVDF syringe filter (Kinesis) into a 2 mL screw-cap microfuge tube (Thermo Fisher). The filtrate was then vacuum-concentrated in a Savant SpeedVac vacuum concentrator (Thermo Fisher) for approximately 3 to 4 hours until the solvent had completely evaporated. 200 uL of 10% (v/v) methanol (HPLC grade) containing 1 ug/mL of scoparone, the second internal control, was added to the dried filtrate and incubated overnight for 16 hours at 4ºC to dissolve the extract. The extract was vortexed vigorously for at least 10 seconds before incubation. After the incubation, it was centrifuged at 13,400 rpm for 5 min to spin down floating particulates and 150 ul of the extract was transferred into a 300 uL screw-cap vial (Waters) which was sent for HPLC-MS in the Department of Pharmacy and Pharmacology, University of Bath.

2.11.2 STANDARD PREPARATION

Standard solutions for the standard curve were prepared in 10% (v/v) methanol (HPLC-grade) at six concentrations: 2000 ug/mL, 1000 ug/mL, 500 ug/mL, 250 ug/mL, 125 ug/mL, 62.5 ug/mL and 31.25 ug/mL in a total volume of 600 uL per vial. The working concentrations were adjusted depending on the number of standards to be included so

that the final concentration would be as listed above. If the compound has a low solubility in the solvent (methanol), DMSO was added to allow the compound to dissolve.

2.12 QUANTITATIVE GENE EXPRESSION ANALYSIS

2.12.1 PRIMER DESIGN AND STANDARD CURVE

Primers were designed specifically for qPCR analysis (Table 2.1). The parameters for designing the primers were: melting temperatures of 55-62ºC with less than 5ºC difference between the primer pair, GC content of 50-60%, amplicon size between 70-150bp, primer spans an exon-exon junction and no four same nucleotides consecutively. The primers for F6H gene family were obtained from Liu *et al.* (2017). The primers were all tested on cDNA with RT-PCR to make sure they work before proceeding for qPCR.

Relative standard curve was performed to test the efficiency of each primer pair before using them for gene expression analysis. Firstly, a cDNA sample was diluted to five different concentrations (1x, 0.2x, 0.04x, 0.008x, 0.0016x), in order to create a standard curve. The qPCR reaction was set up using the StepOne Software V2.3 (Applied Biosystems, UK). The program set-up for the relative curve were 95°C for 30 sec, 94°C for 20 sec, 60°C for 20 sec and 72°C for 20 sec, with 40 cycles in total, 95°C for 15 sec and 40°C for 2 min. The set-up for the reaction was 1x PCR master mix containing 10 uL of fast SYBR green master mix, 0.4 uL forward primer (200 nM), 0.4 uL reverse primer (200 nM) and 8.2 uL of MilliQ water was prepared in a 96 well reaction plate (Applied Biosystems, UK). All preparations were done on ice and the plate was centrifuged at 2,000 rpm for 2 min at RT to remove bubbles. The results were then analysed on the same software and primer pairs with the highest efficiency were used for comparative CT experiment for gene expression analysis.

2.12.2 COMPARATIVE C_T EXPERIMENT

The reaction master-mix was prepared in the reaction plate with the same components and procedure at above (2.12.1). The comparative C_T experiment was used for quantitative gene expression analysis, so the cDNA template was used directly without any dilutions. The program settings for this reaction are: 95°C for 30 sec, 94°C for 20 sec, 60°C for 20 sec and 72°C for 20 sec, with 40 cycles in total, then, 95°C for 15 sec and 40°C for 2 min.

The housekeeping gene used as a reference for this experiment was ubiquitin 10 (ubq10); the primer sequences are listed in Table 2.1. The control used was wild-type cassava storage root sample at day 0.

The relative quantity (RQ) of the target genes are quantified with the $\Delta\Delta C_T$ method which has the formula:

ΔC_TE = TE (Target Experimental) – HE (Housekeeping Experimental)

ΔC_TC = TC (Target Control) – HC (Housekeeping Control)

$\Delta\Delta C_T = \Delta C_T E - \Delta C_T C$

Relative Quantity (RQ) = 2^-$\Delta\Delta C_T$

3 ARE PHENYLPROPANOID METABOLISM GENES INVOLVED IN SCOPOLETIN BIOSYNTHESIS FOUND IN CASSAVA?

3.1 INTRODUCTION

3.1.1 BACKGROUND

The accumulation of scopoletin increases by 150-200 times in cassava storage roots after harvest, strongly indicating that scopoletin is involved in PPD development (Wheatley and Schwabe, 1985). Peroxidase has a high affinity to scopoletin, oxidising it to a blue-black insoluble metabolite, which correlates with the blue-black discolouration, known as vascular streaking, observed in the roots (Reilly *et al.*, 2007). Phenylalanine ammonia-lyase (PAL), the entry enzyme of the phenylpropanoid metabolism, which is also responsible for the biosynthesis of scopoletin, increased during the PPD response (Góez-Vásquez *et al.*, 2004), suggesting *de novo* synthesis of scopoletin. Further evidence supporting the oxidation of scopoletin playing a role in PPD is that the exclusion of oxygen by waxing the roots after harvest effectively delays PPD development (Reilly *et al.*, 2003). Scopoletin is most likely a quencher to ROS compounds such as H_2O_2, which accumulate during the oxidative burst which is triggered by wounding (Reilly *et al.*, 2003).

3.1.2 SCOPOLETIN BIOSYNTHESIS IN CASSAVA AND ARABIDOPSIS

The phenylpropanoid metabolism is responsible for the biosynthesis of many secondary metabolites, including scopoletin, which is a hydroxycoumarin. Within the phenylpropanoid metabolism, there are three pathways that lead to the biosynthesis of scopoletin in cassava (figure 3.1); to simplify the explanation, coenzyme A is omitted from the metabolites. These are via (1) hydroxycinnamate, (2) hydroxycaffeate and (3) ferulate intermediates of which the dominant pathway is via the (3) ferulate intermediate where 90% of total scopoletin in cassava is produced (Bayoumi *et al.*, 2008). All three pathways start with the conversion of *L*-phenylalanine to (*E*)-cinnamate by PAL and then cinnamate 4-hydroxylase (C4H) catalyses the cinnamate to yield (*E*)-*p*-coumarate. After this, pathway 1 diverges as (*E*)-*p*-coumarate is 6'-hydroxylated to form (*E*)-2',4'-dihydroxycinnamate, which is then E-Z isomerised to (*Z*)-2',4'-dihydroxycinnamate. It is

then lactonise to form umbelliferone and then 3'-hydroxylated to from esculetin. Esculetin can then be 3'-methylated to form scopoletin through an *O*-methyltransferase enzyme. In pathway 2 and 3, *p*-coumarate is first 3'-hydroxylated to (*E*)-caffeate. Pathway 2 then branches from (*E*)-caffeate which is 6'-hydroxylated into (*E*)-6'-hydroxycaffeate, then isomerised to (*Z*)-6'-hydroxycaffeate and lactonised into esculetin. Like pathway 1, esculetin is then converted into scopoletin through an *O*-methyltransferase enzyme. In pathway 3, (*E*)-caffeate is 3'-methylated to ferulate and then 6'-hydroxylated to yield (*E*)-6'-hydroxyferulate. It is then E-Z isomerised to (*Z*)-6'-hydroxyferulate which is lactonised to produce scopoletin. Scopoletin can then be glucosylated to its glucoside, scopolin, through scopoletin-glucosyltransferase (scopoletin-GT); and deglucosylated from scopolin back into scopoletin through scopolin-β-glucosidase (scopolin-βG) (Ahn *et al.*, 2010; Gachon *et al.*, 2004).

Like cassava, the model plant, *A. thaliana*, which has been well studied and characterised, also synthesises scopoletin and scopolin in its roots. This makes it a valuable tool in studying the genes involved in scopoletin biosynthesis as it can act as a reference to understand said pathways in cassava (Kai *et al.*, 2006). Although other alternative pathways involved in scopoletin biosynthesis have not been thoroughly explored in *A. thaliana*, it is likely that the major routes will be shared between cassava and *A. thaliana* (Bayoumi *et al.*, 2008; Kai *et al.*, 2006). This is due to the majority of pathways in the phenylpropanoid metabolism being conserved in higher plants because the acquisition of phenylpropanoid metabolism occurred early on the evolution of terrestrial plants, aiding the adaptation to new environments and opportunities of life out of water (Tohge *et al.*, 2013). As the majority of scopoletin is biosynthesised through the major pathway, using *A. thaliana* as a reference, the final enzyme involved in the biosynthesis of scopoletin, F6'H1 (feruloyl CoA 6'-hydroxylase 1, gene ID At3g13610, NCBI reference sequence: NC_003074.8), which is responsible for the 6'-hydroxylation of feruloyl CoA, is of great interest. When F6'H1 is knocked-out in *A. thaliana* by a T-DNA insertion, the mutant does not accumulate scopoletin (Kai *et al.*, 2008). These data confirm that the dominant biosynthetic pathway of scopoletin in cassava is the same as the major pathway (3) in cassava. Using *A. thaliana* as a reference, homologous F6'H genes (Me10291, Me33240, Me10292, Me10381, Me27567, Me10376 and Me30526) were identified in cassava and, with the intent to reduce scopoletin biosynthesis and thus delaying/preventing PPD, RNAi knock-down cassava lines of homologous F6'H genes were generated (Liu *et al.*, 2017). This managed to significantly reduce scopoletin accumulation and root deterioration in the RNAi knock-down cassava lines but not completely; therefore, PPD still occurred, albeit delayed. The reason for this may be due

to there being seven members in the F6'H gene family in cassava so the RNAi construct may not be able to target all the cassava F6'H genes although it should have been able to target most of the genes due to the high similarities between them.

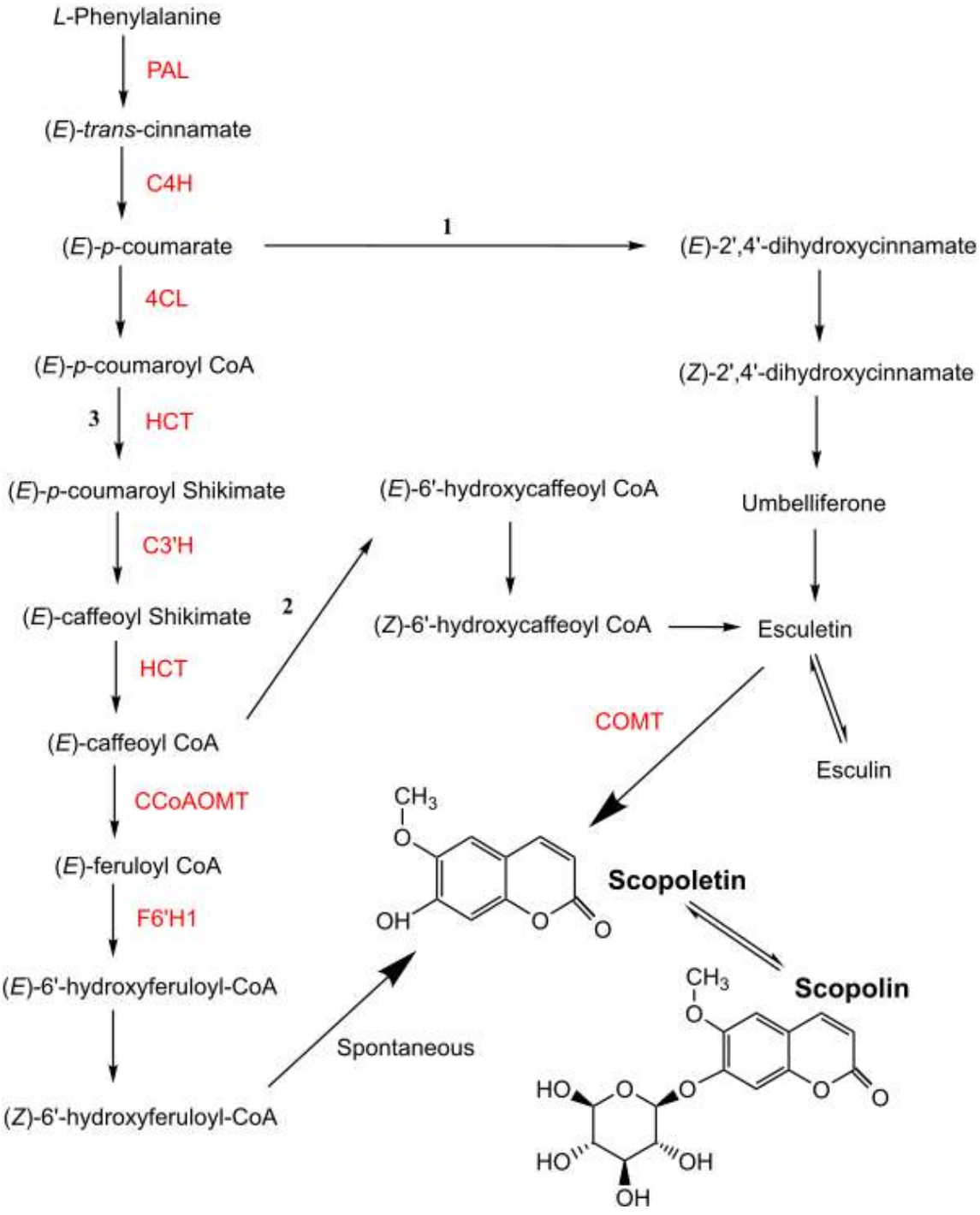


Figure 3.1 **The biosynthesis of scopoletin through three alternative pathways in the phenylpropanoid metabolism.** The enzymes of interest are highlighted in red.

Another reason for the accumulation of scopoletin in the transgenic lines may be due to the production of scopoletin through the alternative pathways. Although the alternative pathways normally account for less than 10% of the scopoletin produced during PPD (Bayoumi *et al.*, 2008), there might be a feedback mechanism that causes an increase in scopoletin production due to the inhibition of the major pathway. The positive results from the F6'H RNAi knock-down cassava mutants confirm that the suggested biosynthetic pathways of scopoletin is correct, and that *A. thaliana* is a good reference to identify other homologous genes involved in the biosynthesis of scopoletin. Understanding the other genes in the phenylpropanoid metabolism involved in the biosynthesis of scopoletin may provide answers as to why the cassava knock-down lines still accumulate scopoletin. It may also identify other potential candidate genes to be manipulated with the aim of fully inhibiting scopoletin biosynthesis.

3.1.3 PHENYLPROPANOID METABOLISM GENES INVOLVED IN SCOPOLETIN BIOSYNTHESIS IN ARABIDOPSIS

Phenylpropanoids are a diverse group of compounds that are derived from the carbon skeleton of phenylalanine and are involved in plant structural support, defence and survival (Vogt, 2010). Although not every secondary metabolite's function is understood, the importance of phenylpropanoids for plant survival is undeniable. Especially the production of monolignols (lignin monomers) that serve as building blocks of lignin which provides structural support, vascular integrity and defence in plants (Boerjan *et al.*, 2003). Lignin is formed through the dehydrogenative polymerisation of three monolignols, *p*-coumaroyl alcohol, coniferyl alcohol and sinapyl alcohol, which respectively give rise to *p*-hydroxyphenyl (H), guiacyl (G) and syringyl (S) units of the lignin polymer (Figure 1.10). Other secondary metabolites synthesised include flavonoids which are important plant pigments needed for flower colouration to attract pollinators as well as coumarins such as scopoletin and esculetin which play a role in plant defence (Lepiniec *et al.*, 2006; D'Auria and Gershenzon, 2005). Figure 3.1 shows the biosynthesis of scopoletin through the three alternative pathways in cassava and includes the enzymes involved through reference from *A. thaliana* which will be further discussed in this section.

3.1.3.1 PHENYLALANINE AMMONIA-LYASE (PAL)

The entry point of the phenylpropanoid pathway is the amino acid *L*-phenylalanine (*L*-Phe) which is the end product from the shikimate pathway. PAL catalyses the initial reaction in which *trans*-cinnamate ((*E*)-cinnamate) and ammonia are produced through

the deamination of phenylalanine. There are four PAL isoforms (PAL1 (At2g37040), PAL2 (At3g53260), PAL3 (At5g04230) and PAL4 (At3g10340)) in *A. thaliana* (Raes et al., 2003). Expression studies performed in *A. thaliana* have shown that PAL1 and PAL2 are the main PAL enzymes as there are specific promoter elements related to PAL1 and PAL2 but not PAL3 and PAL4 (Raes *et al.*, 2003). To further support the observation that PAL1 and PAL2 are the main isoforms in *A. thaliana*, studies by Rohde *et al.* (2004) have shown that *pal1* and *pal2* mutants have no noticeable phenotype as PAL2 is upregulated in the *pal1* mutant and *vice versa*. The *pal1pal2* double mutant has reduced lignin content with a higher syringyl (S) to guaiacyl (G) ratio, accumulates phenylalanine and is also deficient in anthocyanin and tannin biosynthesis (Rohde *et al.*, 2004). PAL4 is up-regulated in both *pal1* and *pal2* mutants as well as the double mutant but it only partly compensates the loss of the main two PAL genes. The *pal1 pal2 pal3 pal4* mutant has stunted growth with reduced lignin content but still has residual PAL activity which could mean that there are other unidentified PAL-like genes (Huang *et al.*, 2010).

3.1.3.2 CINNAMATE 4-HYDROXYLASE (C4H)

C4H is a member of the cytochrome P450-dependent monooxygenase family (P450s) that catalyses the hydroxylation of (*E*)-cinnamate into (*E*)-*p*-coumarate (also known as 4-coumarate). C4H is the first of three P450s in the phenylpropanoid metabolism that is involved in lignin biosynthesis and is only encoded by a single gene, CYP73A5 (At2g30490), in *A. thaliana* (Werck-Reichhart et al., 2002). C4H plays an important role in the biosynthesis of lignin as there is a high expression level of C4H in cells and roots that are undergoing lignification (Bell-Lelong *et al.*, 1997). Besides that, C4H expression is also upregulated by light, wounding and pathogen infections in plants (Chapple, 1998). A viable mutant for C4H has only recently been identified in *A. thaliana* (Schilmiller *et al.*, 2009; Ruegger and Chapple, 2001) as the complete loss of C4H is lethal to the plant due to its importance in vascular development (Bell-Lelong *et al.*, 1997). There are three *reduced epidermal fluorescence 3* (*ref3*) mutants with a missense mutation that leads to amino acid substitutions in the highly conserved structural motifs of the C4H protein that result in reduced enzyme activity. The mutants (*ref3-1*, *ref3-2* and *ref3-3*) all have lower concentrations of sinapoylmalate compared to wild-type plants. *ref3* mutants have lower lignin deposition compared to wild-type plants as well as reduced levels of condensed tannins in the seeds (Schilmiller *et al.*, 2009).

3.1.3.3 4-COUMARATE COA LIGASE (4CL)

p-coumarate is then catalysed by 4-coumarate:CoA ligase (4CL) to form *p*-coumaroyl CoA in an ATP dependent formation of the thioester bond between coenzyme A (CoA) and its carboxyl group. There are four isoforms of 4CL, 4CL1 (At1g51680), 4CL2

(At3g21240), 4CL3 (At1g65060) and 4CL4 (At3g21230), in *A. thaliana* (Costa *et al.*, 2005; Hamberger and Hahlbrock, 2004; Ehlting *et al.*, 1999; Lee *et al.*, 1995). The expression of 4CL1 correlates with pathogen attack, stem bolting and lignin deposition in seedlings (Lee *et al.*, 1995). 4CL2 has a similar expression pattern in stems to 4CL1 as they are 83% identical (Hamberger and Hahlbrock, 2004). 4CL3 is only 61% identical to 4CL1 and is primarily expressed in siliques and unlike 4CL1 and 4CL2, 4CL3 is induced by UV radiation and not pathogenic infiltration (Ehlting *et al.*, 1999). Expression profiles and data from 4CL homologues in other plant species suggest that 4CL1 and 4CL2 are likely to be involved in lignin biosynthesis and 4CL3 to be involved in flavonoid biosynthesis (Ehlting *et al.*, 1999). 4CL4 is the last isozyme to be identified and it has a completely different metabolic function to the other three isozymes as it shows favoured activity towards sinapate and ferulate instead of *p*-coumarate (Hamberger and Hahlbrock, 2004).

3.1.3.4 HYDROXYCINNAMOYL-COA SHIKIMATE:QUINATE HYDROXYCINNAMOYL-TRANSFERASE (HCT)

Hydroxycinnamoyl-CoA shikimate:quinate hydroxycinnamoyl-transferase (HCT) is involved in two reactions in the phenylpropanoid metabolism. Firstly, it catalyses the transfer of the *p*-coumaroyl group in *p*-coumaroyl CoA to shikimate to produce *p*-coumaroyl shikimate (Hoffmann *et al.*, 2003). *p*-coumaroyl shikimate 3' hydroxylase (C3'H) then catalyses the hydroxylation of *p*-coumaroyl shikimate to produce caffeoyl shikimate (Franke *et al.*, 2002a). HCT then proceeds to catalyse the transfer of the caffeoyl moiety back to the CoA to form caffeoyl CoA (Hoffmann *et al.*, 2003). HCT can use quinate instead of shikimate but not very efficiently. At5g48930 is the is only gene encoding HCT in *A. thaliana* and *A. thaliana* mutants where HCT accumulation is inhibited by RNA-mediated gene silencing show a dwarf phenotype with accumulation of flavonoids and have reduced S lignin in comparison to wild-type plants (Hoffmann *et al.*, 2004).

3.1.3.5 *P*-COUMAROYL SHIKIMATE 3' HYDROXYLASE (C3'H)

C3'H is the second out of three P450s involved in lignin biosynthesis in the phenylpropanoid metabolism and it catalyses the hydroxylation of *p*-coumaroyl shikimate to produce caffeoyl shikimate (Franke *et al.*, 2002a). It is also one of the three members belonging in the CYP98 family (CYP98A3) in *A. thaliana* (Franke *et al.*, 2002a; Schoch *et al.*, 2001). C3'H is able to act on both the shikimate ester and the quinate ester of *p*-coumarate but the rate of conversion for the shikimate ester is four times higher than the latter (Schoch *et al.*, 2001). The CYP98A3 mutant, *ref8*, has a single mutation which leads to a G444D substitution that greatly reduces C3'H activity (Franke *et al.*, 2002a).

Like the other *ref* mutants, *ref8* plants accumulates lower concentrations of sinapoylmalate in comparison to wild-type plants and deposit H lignin instead of G and S lignin. It also has a dwarfed phenotype and over-accumulates flavonoids (Franke *et al.*, 2002b).

3.1.3.6 CAFFEOYL COA 3-*O*-METHYLTRANSFERASE (CCoAOMT)

Caffeoyl CoA 3-*O*-methyltransferase (CCoAOMT) catalyses the transfer of the methyl group of caffeoyl CoA to yield feruloyl CoA. There is only one functionally confirmed CCoAOMT gene in *A. thaliana* (CCoAOMT1) but there are seven other CCoAOMT-like genes. T-DNA knock-out mutants have collapsed xylem vessels as well as a lower stem lignin concentration than wild-type plants (Do *et al.*, 2007). Kai *et al.* (2008) also showed that the lack of CCoAOMT activity significantly affected the biosynthesis of coumarins; the levels of scopoletin and scopolin compared to wild-type roots were approximately 30% and 15% lower respectively.

3.1.3.7 FERULOYL COA 6'-HYDROXYLASE (F6'H)

Feruloyl CoA 6'-hydroxylase 1 (F6'H1) is an *A. thaliana* gene which regulates the hydroxylation of feruloyl CoA to 6'-hydroxyferuloyl CoA and is required for scopoletin biosynthesis (Kai *et al.*, 2008). F6'H1 belongs to the 2-oxoglutarate dependent dioxygenase (2OGD) enzyme family where 2OGD integrates both atoms of an O_2 molecule to a 2-oxoglutarate molecule and its substrate. For scopoletin biosynthesis, F6'H1 incorporates a hydroxyl group to feruloyl-CoA, turning it to 6'-hydroxyferuloyl CoA. When this gene was knocked-out through a T-DNA insertion in *A. thaliana*, scopoletin failed to accumulate in the roots, confirming that the dominant scopoletin biosynthetic pathway in cassava is the same in *A. thaliana* (Kai *et al.*, 2008).

3.1.3.8 CAFFEIC ACID *O*-METHYLTRANSFERASE (COMT)

In both the alternative pathways, esculetin is methoxylated to produce scopoletin, which means this reaction is catalysed by an *O*-methyltransferase enzyme. A study by Kim *et al.* (2006) showed that *E. coli* expressing an *O*-methyltransferase gene from poplar (*Populus deltoids* Marsh), *POMT-9*, was able to produce scopoletin, isoscopoletin and scoparone from esculetin. The homologue of POMT-9 in *A. thaliana* is caffeic acid *O*-methyltransferase 1 (COMT1). COMT methoxylates at the C5 position of the phenolic ring and is responsible for the biosynthesis of sinapyl alcohol, the precursor of the syringyl (S) units. An *A. thaliana* mutant with the *COMT1* gene knock-out shows lignins lacking in S units and contains 5-OH-G units derived from 5'-hydroxyconiferyl alcohol although it did not show reduction in lignin content (Goujon *et al.*, 2003).

3.2 AIM OF RESEARCH

This study intends to further the understanding into the function and gene expression patterns of the phenylpropanoid enzymes to gain an insight into the biosynthesis of scopoletin. Since the major pathway of scopoletin biosynthesis is shared between cassava and *A. thaliana* and phenylpropanoid enzymes are highly conserved (Tohge *et al.*, 2013), homologous genes in cassava can be identified using functionally confirmed *A. thaliana* genes of said enzymes (PAL, C4H, 4CL, HCT, C3'H, CCoAOMT and COMT). This will be done through BLASTing *A. thaliana* reference genes in the cassava genome and identifying active sites, binding sites and signature patterns in the cassava candidate genes to confirm functionality. Gene specific primers will be designed for the cassava candidate genes and tested on cassava storage root and leaf cDNA to understand the expression profiles. Ultimately, selected cassava candidate genes may be knocked-down using RNA-mediated silencing to produce transgenic cassava with the aim to inhibit scopoletin accumulation.

3.3 RESULTS

3.3.1 CASSAVA CANDIDATE GENES

The *A. thaliana* reference sequences were retrieved from NCBI and peptide sequences were used to BLAST for cassava homologous genes in the Phytozome database (Goodstein *et al.*, 2012) with the default algorithmic parameters. The candidate genes were evaluated under two parameters, the Expect (E) value and score. The *E*-value describes the number of hits to be expected by chance when searching a database of a certain size. Essentially, the *E*-value describes the background noise and the lower the *E*-value (the closer it is to 0), the more 'significant' the match is. The score indicates the similarity between the query sequence and the hit sequence and the *E*-value is expected to decrease exponentially as the score of the match increases (Goodstein *et al.*, 2012). The scores and *E*-values of the cassava candidate genes and the *A. thaliana* reference genes are shown in Appendix II. Active sites, binding sites and signature pattern obtained from PROSITE are also observed to make sure the candidate genes are true homologs of the reference genes. Although these domain signatures are generally accurate, there are false positive hits as well as false negative sequences, hence these signature patterns can only be used alongside the phylogenetic analyses as a selection criterion and do not confirm functionality of the cassava candidate genes.

3.3.1.1 PHENYLALANINE AMMONIA-LYASE (PAL)

The four PAL sequence - PAL1 (At2g37040), PAL2 (At3g53260), PAL3 (At5g04230) and PAL4 (At3g10340) from *A. thaliana* were used as references to identify homologous genes in the cassava genome. Six genes from cassava (Cassava4.1_002628, Cassava4.1_034377, Cassava4.1_002709, Cassava4.1_002591, Cassava4.1_028434 and Cassava4.1_003117) were good candidates as the E values were 0 and the scores were above 80% for all except Cassava4.1_003117 which has an identity score of 59.3% to 64.5% depending on which PAL sequence was used. These genes were then compared with PAL genes from dicotyledons, monocotyledons and bryophytes in a Neighbour-Joining consensus phylogenetic tree with a bootstrap replicate of 1000 (Figure 3.2). The phylogenetic analysis includes dicotyledons (cassava (*Manihot esculenta*), *Arabidopsis thaliana*, *Citrus sinensis*, *Malus domestica, Solanum tuberosum*, *Boechera stricta*, *Brassica rapa* and *Populus tricocarpa*), monocotyledons (*Sorghum bicolor, Brachypodium distachyon*, *Setaria italica* and *Oryza sativa*) and bryophyte species (*Physcomitrella patens*). Cassava4.1_002628, Cassava4.1_034377, Cassava4.1_002709, Cassava4.1_002591 and Cassava4.1_028434 are most closely related to the *A. thaliana* PAL genes within the dicotyledon cluster. Cassava4.1_3117 is more distantly related and is most closely related to a bryophyte (*P. patens*) which suggests that this gene is perhaps older than the others (Figure 3.2). Cassava4.1_003117 does not have an intron whereas the other candidate genes all possess one intron, but the coding sequence of all the candidate genes are all relatively the same length. Despite not having an intron, the coding sequence of Cassava4.1_003117 is aligned to the other cassava candidate genes.

To confirm that the retrieved cassava genes are good candidates that are likely to share the same functions as the *A. thaliana* reference genes, the active site, binding site and PAL signature pattern have been investigated. The active site of the *A. thaliana* reference PAL1 is tyrosine (Y) at position 117, the binding site is arginine (R) at position 363 and the PAL signature pattern is [GS]-[STG]-[LIVM]-[STG]-[SAC]-S-G-[DH]-L-x-[PN]-L-[SA]-x(2,3)-[SAGVTL] (Langer *et al.*, 1994; Schuster and Retey, 1994; Taylor and McInnes, 1994). The symbol ‘x’ is used when any amino acid is acceptable and amino acids indicated between a square bracket [ ] means any of those amino acids can be accepted, for example, [STG] means either serine, threonine and glycine can be accepted. These conserved sites are all present in the six cassava candidate genes (Figure 3.3), indicating that they should share the same function and activity as the reference genes.

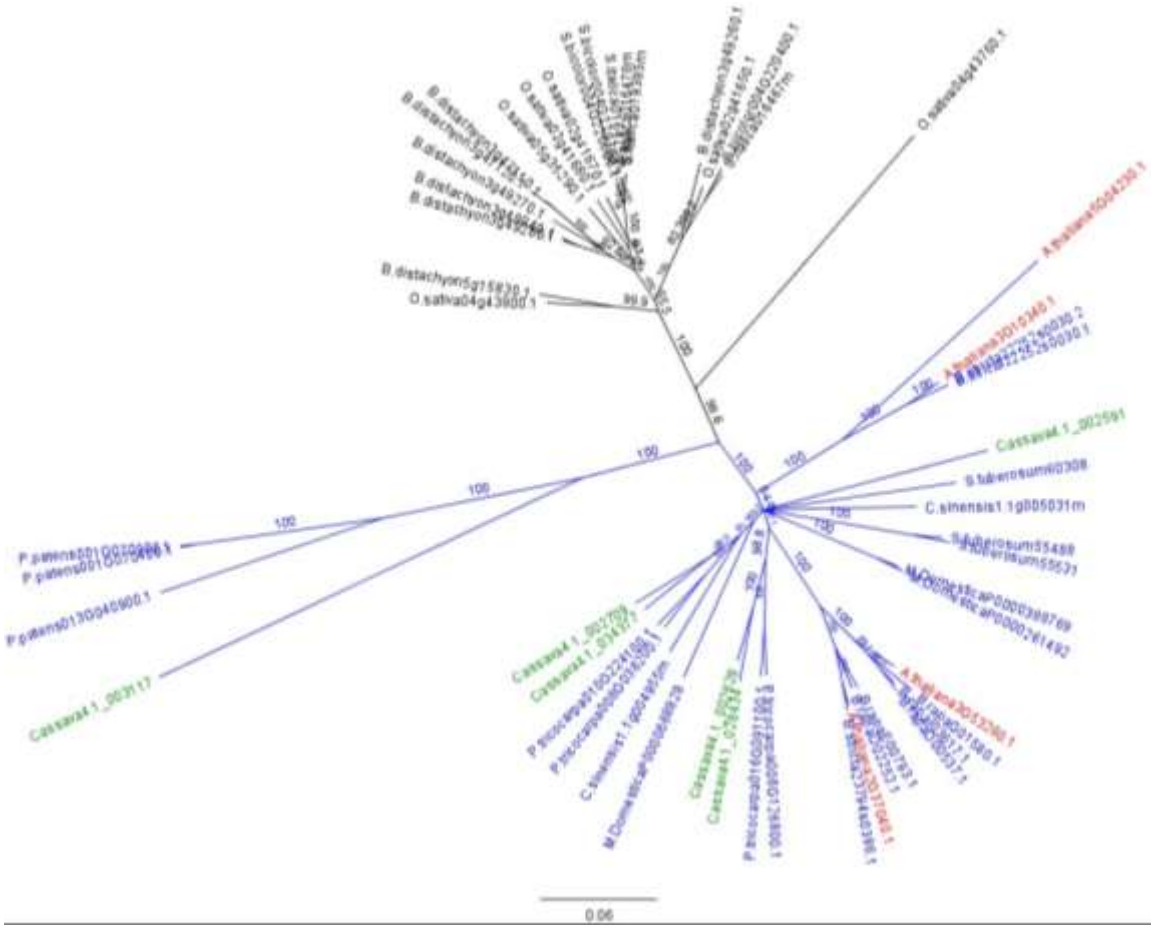


Figure 3.2 **Neighbour joining consensus phylogenetic analysis of phenylalanine ammonia-lyase (PAL) peptide sequences with 1000 bootstrap replicates.** *A. thaliana* reference sequences (PAL1 (At2g37040), PAL2 (At3g53260), PAL3 (At5g04230) and PAL4 (At3g10340)) are highlighted in red and the cassava candidate genes (Cassava4.1_002628, Cassava4.1_034377, Cassava4.1_002709, Cassava4.1_002591, Cassava4.1_028434 and Cassava4.1_003117) are in green. The blue highlight shows the node from which the cassava candidate genes were selected. The plant species included in this analysis are the dicots cassava (*Manihot esculenta*), *Arabidopsis thaliana*, *Citrus sinensis*, *Malus domestica, Solanum tuberosum*, *Boechera stricta*, *Brassica rapa* and *Populus tricocarpa,* monocots *Sorghum bicolor, Brachypodium distachyon*, *Setaria italica* and *Oryza sativa and* bryophyte species *Physcomitrella patens.* The scale bar represents 0.06 substitutions per amino acid.

A

```
Cassava4.1_003117      TYGVTTGFGATSHRRTSKTGDLQTELIRFLNAGVI--GKEN------LPTSYSKAAMLVR
Cassava4.1_002591      SYGVTTGFGATSHRRTKQGAALQQELIRFLNAGIF--GKGA-ESGHTLPRTATRAAMLVR
PAL1-At2g37040         SYGVTTGFGATSHRRTKNGVALQKELIRFLNAGIF--GSTK-ETSHTLPHSATRAAMLVR
PAL2-At3g53260         SYGVTTGFGATSHRRTKNGTALQTELIRFLNAGIF--GNTK-ETCHTLPQSATRAAMLVR
Cassava4.1_028434      SYGVTTGFGATSHRRTKEGGALQRELIRFLNAGIF--GKGT-ESCHTLSHSATRAAMLVR
Cassava4.1_002628      SYGVTTGFGATSHRRTKQGGALQRELIRFLNAGIF--GNKT-ESCHTLSHSATRAAMLVR
Cassava4.1_034377      SYGVTTGFGATSHRRTKQGGALQRELIRFLNAGIF--GNGQ-ESCHTLSHSATRAAMLVR
Cassava4.1_002709      TYGITTGFGATSHRRTKQGGALQRELIRFLNAGIF--GNGA-ESSHTLPHSGTRAAMLVR
PAL3-At5g04230         TYGITTGFGSSSRRRTDQGAALQKELIRYLNAGIFATGNEDDDRSNTLPRPATRAAMLIR
PAL4-At3g10340         SYGVTTGFGATSHRRTKQGGALQNELIRFLNAGIF--GPGAGDTSHTLPKPTTRAAMLVR
                       :**:*****::*.***.:   ** ****:****::  *        *. . :.****:*
```

B

```
Cassava4.1_003117      TNTLMQGYSGIRWEILESIAKLMNENLIPKLPLRGTITASGDLVPLSYIAGLLTGRHNSK
Cassava4.1_002591      INTLLQGYSGIRFEILEALTKLINHNITPCLPLRGSITASGDLVPFSYIAGLLTGRPNST
PAL1-At2g37040         INTLLQGFSGIRFEILEAITSFLNNNITPSLPLRGTITASGDLVPLSYIAGLLTGRPNSK
PAL2-At3g53260         VNTLLQGYSGIRFEILEAITSLLNHNISPSLPLRGTITASGDLVPLSYIAGLLTGRPNSK
Cassava4.1_028434      INTLLQGYSGIRFEILEAITKFLNHNITPCLPLRGTITASGDLVPLSYIAGLLTGRPNSK
Cassava4.1_002628      INTLLQGYSGIRFEILDAITKLLNHNITPCLPLRGTITASGDLVPLSYIAGLLTGRPNSK
Cassava4.1_034377      INTLLQGYSGIRFEIMEAITKFINNNVTPRLPLRGTITASGDLVPLSYIAGLLTGRPNSK
Cassava4.1_002709      INTLLQGYSGIRFEILEAISKFINNNITPCLPLRGTITASGDLIPLAYIAGLLTGRPNSK
PAL3-At5g04230         VNTLLQGYSGIRFEILEAITTLLNCKITPLLPLRGTITASGDLVPLSYIAGFLIGRPNSR
PAL4-At3g10340         VNTLLQGYSGIRFEILEAITKLLNHEITPCLPLRGTITASGDLVPLSYIAGLLTGRPNSK
                        ***:**:****:**::::::.::* :: * *****:*******:*::****:* ** **
```

C

```
Cassava4.1_003117      TKPKQDRYAIRTSPQWLGPQIEVIRAATHSIEREINSVNDNPLIDVARDLALHGGNFQGT
Cassava4.1_002591      QKPKQDRYALRTSPQWLGPQIEVIRFSTKSIEREINSVNDNPLIDVSRNKALHGGNFQGT
PAL1-At2g37040         QKPKQDRYALRTSPQWLGPQIEVIRYATKSIEREINSVNDNPLIDVSRNKAIHGGNFQGT
PAL2-At3g53260         QKPKQDRYALRTSPQWLGPQIEVIRQATKSIEREINSVNDNPLIDVSRNKAIHGGNFQGT
Cassava4.1_028434      QKPKHDRYALRTSPQWLGPQIEVIRFSTKAIEREINSVNDNPLIDVSRNKALHGGNFQGT
Cassava4.1_002628      QKPKQDRYALRTSPQWLGPQIEVIRFSTKSIEREINSVNDNPLIDVSRNKALHGGNFQGT
Cassava4.1_034377      QKPKQDRYALRTSPQWLGPQIEVIRTATKMIEREINSVNDNPLIDVSRNIALHGGNFQGT
Cassava4.1_002709      QKPRQDRYALRTSPQWLGPQIEVIRTATKMIEREINSVNDNPLIDVSRNKALHCGNFQGT
PAL3-At5g04230         QKPKQDRYALRTSPQWLGPQIEVIRAATKMIEREINSVNDNPLIDVSRNKAIHGGNFQGT
PAL4-At3g10340         QKPKQDRYALRTSPQWLGPQIEVIRAATKMIEREINSVNDNPLIDVSRNKALHGGNFQGT
                        **.:****:*************** :*: ***************:*: *:* ******
```

Figure 3.3 **Protein map of PAL and MUSCLE alignment of the cassava candidate genes (Cassava4.1_002628, Cassava4.1_034377, Cassava4.1_002709, Cassava4.1_002591, Cassava4.1_028434 and Cassava4.1_003117) and the *A. thaliana* PAL genes (PAL1 (At2g37040), PAL2 (At3g53260), PAL3 (At5g04230) and PAL4 (At3g10340)). A)** Tyrosine (Y) is the active site for PAL enzymes. **B)** The PAL signature pattern. **C)** Arginine (R) is the binding site of PAL enzymes.

3.3.1.2 CINNAMATE 4-HYDROXYLASE (C4H)

The *A. thaliana* gene, At2g30490, was used as a reference to identify homologous genes in the cassava genome. Two genes from cassava (Cassava4.1_005978 and Cassava4.1_005452) were selected as candidates because the *E*-values were both 0 and the scores were 827.0 and 580.1 respectively. These genes were then compared with C4H genes from dicotyledons (cassava (*Manihot esculenta*), *Arabidopsis thaliana*, *Boechera stricta* and *Linum usitatissimum*), monocotyledons (*Sorghum bicolor*, *Brachypodium distachyon*, *Oryza sativa* and *Zea mays*) and bryophyte (*Physcomitrella patens*) species in a Neighbour-Joining consensus phylogenetic tree with a bootstrap replicate of 1000 (Figure 3.4). The phylogenetic analysis shows that bryophyte, monocotyledons and dicotyledons are in separate clusters and Cassava4.1_005978 is more closely related to the reference At2g30490 which is why it was selected as the homologous gene and Cassava4.1_005452 was disregarded.

To confirm that the retrieved cassava genes are good candidates that will share the same functions as the *A. thaliana* C4H reference, the C4H signature pattern [FW]-[SGNH]-x-[GD]-{F}-TG-[RKHPT]-{P}-C-[LIVMFAP]-[GAD] where cysteine (C) is the heme iron ligand is also present in the peptide sequence of Cassava4.1_005978, indicating the shared function between them (Figure 3.5). The symbol ‘x’ is used when any amino acid is acceptable and amino acids indicated between a square bracket [ ] means any of those amino acids can be accepted, for example, [STG] means either serine, threonine and glycine can be accepted. The curly brackets {x} indicate any amino acids can be accepted except x, for example, {G} means apart from glycine, all other amino acids are acceptable.

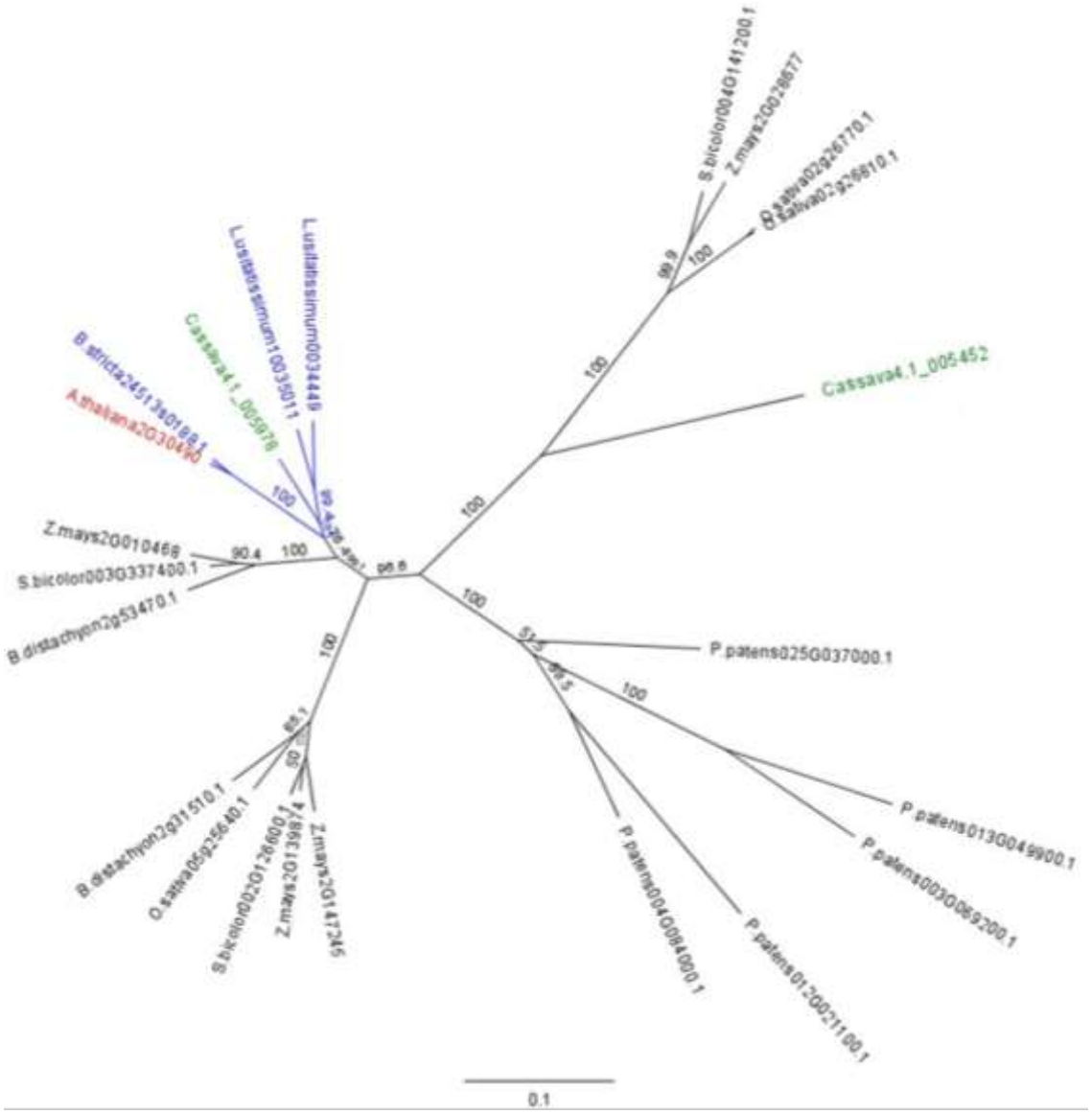


Figure 3.4 **Neighbour joining consensus phylogenetic analysis of cinnamate 4-hydroxylase peptide sequences with 1000 bootstrap replicates.** *A. thaliana* reference sequence (At2g3049) is highlighted in red and the cassava candidate genes (Cassava4.1_005978 and Cassava4.1_005452) are in green. The blue highlight shows the node from which the cassava candidate genes were selected from. The plant species analysed are the dicots, (cassava (*Manihot esculenta*), *Arabidopsis thaliana*, *Boechera stricta* and *Linum usitatissimum*), monocots (*Sorghum bicolor*, *Brachypodium distachyon*, *Oryza sativa* and *Zea mays*) and bryophyte (*Physcomitrella patens*). Cassava4.1_005978 is the best candidate with *A. thaliana* At2g30490 as the reference gene. The scale bar represents 0.1 substitutions per amino acid.

Figure 3.5 **Protein map of C4H and MUSCLE alignment of the cassava candidate gene's peptide sequence, Cassava4.1_005978, and the *A. thaliana* reference C4H, At2g30490.** The amino acids highlighted (FGVGRRSCPG) are the C4H signature pattern.

3.3.1.3 4-COUMARATE COA LIGASE (4CL)

The four 4CL isozymes are encoded by four genes in *A. thaliana*, 4CL1 (At1g51680), 4CL2 (At3g21240), 4CL3 (At1g65060) and 4CL4 (At3g21230). These were used as references to identify homologous genes in the cassava genome. Four genes from cassava (Cassava4.1_005006, Cassava4.1_005014, Cassava4.1_004658 and Cassava4.1_004136) were identified as having low enough E values with both high identity and score (Appendix II). These genes were then compared with 4CL genes from dicotyledons (cassava (*Manihot esculenta*), *Arabidopsis thaliana*, *Prunus persica*, *Citrus sinensis*, *Aquilegia coerulea* and *Linum usitatissimum*), monocotyledons (*Sorghum bicolor*, *Brachypodium distachyon* and *Oryza sativa*) and a bryophyte species (*Physcomitrella patens*) in a Neighbour-Joining consensus phylogenetic tree with a bootstrap replicate of 1000 (Figure 3.6). The phylogenetic tree shows that Cassava4.1_005006 and Cassava4.1_005014 were more closely related to 4CL1 (At1g51680), 4CL2 (At3g21240) and 4CL4 (At3g21230), whereas Cassava4.1_004658 and Cassava4.1_004136 were more closely related to 4CL3 (At1g65060).

To confirm that the retrieved cassava genes are good candidates that will share the same functions as the *A. thaliana* reference genes, the nucleotide binding sites and 4CL signature pattern were present in the candidate genes (Figure 3.7). The 4CL signature, which is an AMP-binding domain signature, is [LIVMFY]-{E}-{VES}-[STG]-[STAG]-G-[ST]-[STEI]-[SG]-x-[PASLIVM]-[KR]. Within the signature pattern, SSGTTGLPK is a nucleotide binding site where a mutagenesis study showed that enzyme activity is greatly reduced when lysine (K) is replaced by serine (S) (Stuible *et al.*, 2000). The other highlighted sequences are binding sites which have shown reduction in enzyme activity

and substrate specificity when mutated; the activity of 4CL is completely inhibited when lysine (K) is replaced by asparagine (N) (Figure 3.7D) (Stuible *et al.*, 2000).

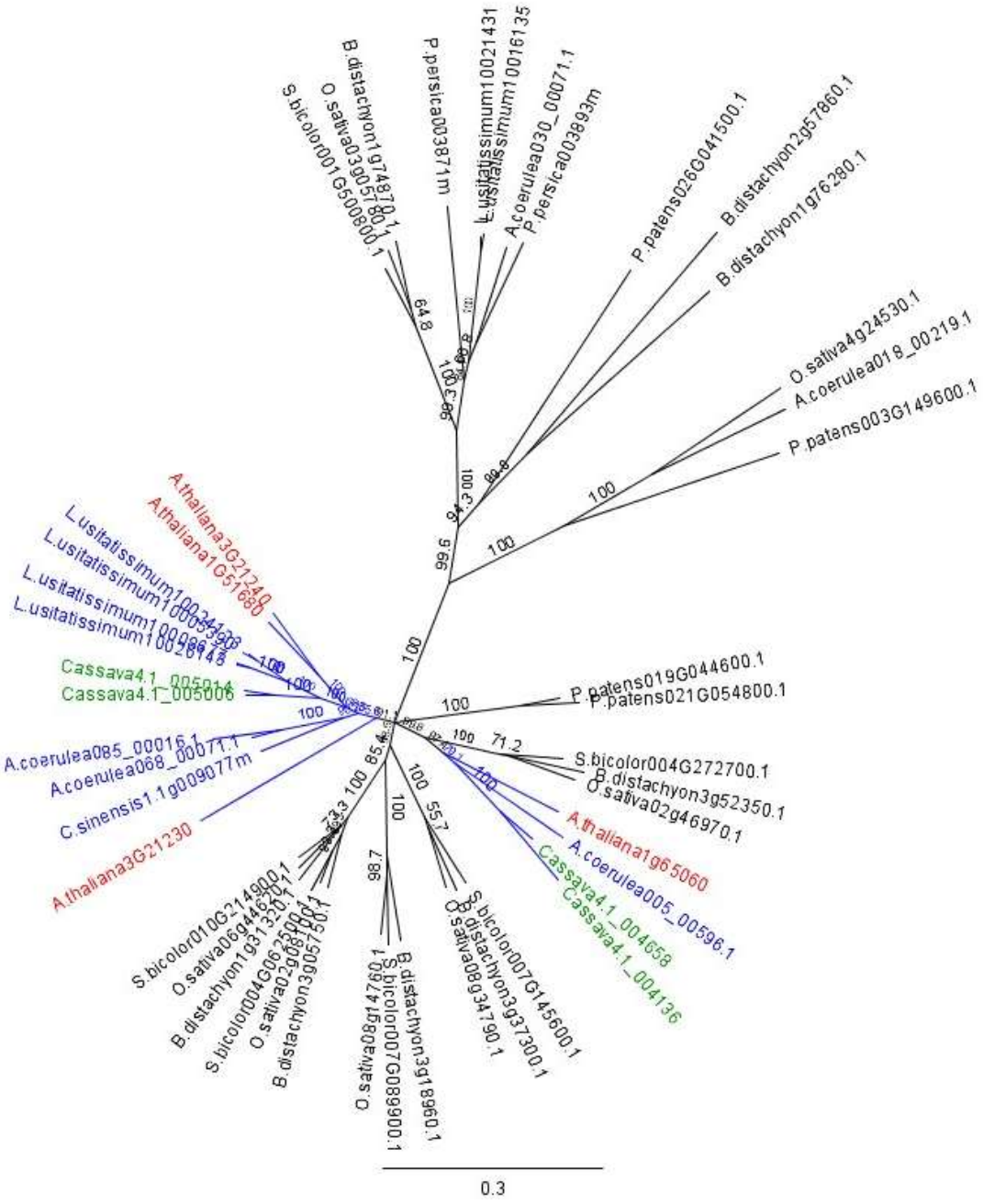


Figure 3.6 **Neighbour joining consensus phylogenetic analysis of 4 coumarate CoA ligase (4CL) peptide sequences with 1000 bootstrap replicates.** *A. thaliana* reference sequences 4CL1 (At1g51680), 4CL2 (At3g21240), 4CL3 (At1g65060) and 4CL4 (At3g21230) are highlighted in red and the cassava candidate genes' peptide sequence (Cassava4.1_005006, Cassava4.1_005014, Cassava4.1_004658 and Cassava4.1_004136) are in green. The blue highlight shows the node from which the cassava candidate genes were selected from. The plant species analysed are the dicots cassava (*Manihot esculenta*), *Arabidopsis thaliana*, *Prunus persica*, *Citrus sinensis*, *Aquilegia coerulea* and *Linum usitatissimum*, monocots *Sorghum bicolor*, *Brachypodium distachyon* and *Oryza sativa*, and bryophyte *Physcomitrella patens*. The scale bar represents 0.3 substitutions per amino acid.

A

```
4CL3-At1g65060          TTDEPT--------PENCLPFSTLITDDET-NPFQETVDIGGDDAAALPFSSGTTGLPKG
Cassava4.1_004658       TIDDP---------PENCLHFTVLSEAKES---EIPDVTIHPDDPVALPFSSGTTGLPKG
Cassava4.1_004136       TIDDP---------PENCLPLTVLTEANES---EVPDVTINPDDPVALPFSSGTTGLPKG
4CL4-At3g21230          CLDDDGDNGVVSSSDDGCVSFTELTQADET---ELLKPKISPEDTVAMPYSSGTTGLPKG
Cassava4.1_005006       CIDSA---------PDGCLHFSELSEADER---DMPDVDISPEDVVALPYSSGTTGLPKG
Cassava4.1_005014       CIDSA---------PDGCLHFSELSEADER---DMPDVNISSEDVVALPYSSGTTGLPKG
4CL1-At1g51680          CIDDNESVPI----PEGCLRFTELTQSTTEASEVIDSVEISPDDVVALPYSSGTTGLPKG
4CL2-At3g21240          TTDSDAI-------PENCLRFSELTQSEEPRVDSIPE-KISPEDVVALPFSSGTTGLPKG
                          *.          :.*: :: *              *  :* .*:*:**********
```

B

```
4CL3-At1g65060          QDSLRRRLPQAILGQGYGMTEAGPVLSMSLGFAKEPIPTKSGSCGTVVRNAELKVVHLET
Cassava4.1_004658       EDALRSRVPQAILGQGYGMTEAGPVLSMCLGFAKQPFPTKSGSCGTVVRNAELKVIDPET
Cassava4.1_004136       EDAFRSRVPQATLGQGYGMTEAAAAITLCLGFAKQPFPTKLGSCGTVLRNAELKVIDPET
4CL4-At3g21230          EDAVRLKFPNAIFGQGYGMTESGTV-AKSLAFAKNPFKTKSGACGTVIRNAEMKVVDTET
Cassava4.1_005006       EDTVRAKFPQATLGQGYGMTEAGPVLAMCLAFAKEPFNIKPGACGTVVRNAEMKIVDPET
Cassava4.1_005014       EDTVRAKFPQATLGQGYGMTEAGPVLAMCLAFAKEPFSIKPGACGTVVRNAEMKIVDPET
4CL1-At1g51680          EDAVNAKFPNAKLGQGYGMTEAGPVLAMSLGFAKEPFPVKSGACGTVVRNAEMKIVDPDT
4CL2-At3g21240          EDAISAKFPNAKLGQGYGMTEAGPVLAMSLGFAKEPFPVKSGACGTVVRNAEMKILDPDT
                        :*:.  ..*:* :********:... : .*.***:*:  * *:****:****:*::  :*
```

C

```
4CL3-At1g65060          RLSLGYNQPGEICIRGQQIMKEYLNDPEATSATIDEEGWLHTGDIGYVDEDDEIFIVDRL
Cassava4.1_004658       GCSLGYNQPGEICIRGQQIMKGYLNDLEATANTIDVEGWLHTGDIGYVDDDEEVFIVDRV
Cassava4.1_004136       GCSLGYNQPGEICIRGQQIMKAYLDDLETTLNTIDVEGWLHTGDIGYVDDDEEVFIVDRI
4CL4-At3g21230          GISLPRNKSGEICVRGHQLMKGYLNDPEATARTIDKDGWLHTGDIGFVDDDDEIFIVDRL
Cassava4.1_005006       GDSLPRNQPGEICIRGDQIMKGYLNDPEATARTIDKQGWLHTGDIGYIDDDDELFIVDRL
Cassava4.1_005014       GESLPRNQPGEICIRGDQIMKGYLNDPEATARTIDKQGWLHTGDIGYIDDDDELFIVDRL
4CL1-At1g51680          GDSLSRNQPGEICIRGHQIMKGYLNNPAATAETIDKDGWLHTGDIGLIDDDDELFIVDRL
4CL2-At3g21240          GDSLPRNKPGEICIRGNQIMKGYLNDPLATASTIDKDGWLHTGDVGFIDDDDELFIVDRL
                          **  *:.****:** *:** **::  :*  *** :*******:* :*:*:*:*****:
```

D

```
4CL3-At1g65060              KEYVAKQVVFYKRLHKVFFVASIPKSPSGKILRKDLKAKLC------
Cassava4.1_004658           KDYIAKQVVFYKKLHKVYFVHAIPKSPSGKILRKDIKAKLASTSSVS
Cassava4.1_004136           KEYIEKQVVFYKKLSKVYFVHAIPKSPSGKILRKDLKARLAAASLLP
4CL4-At3g21230              KSYVNKQVVHYKRIKMVFFIEVIPKAVSGKILRKDLRAKLETMCSK-
Cassava4.1_005006           KQYISKQVVFYKRIGRVFFMEAIPKAPSGKILRKNLREKLATDFQK-
Cassava4.1_005014           KQYIYKQVVFYKRIGKVFFVESIPKAPSGKILRKILRENLTSDFQK-
4CL1-At1g51680              KQFVSKQVVFYKRINKVFFTESIPKAPSGKILRKDLRAKLANGL---
4CL2-At3g21240              KQFVSKQVVFYKRINKVFFTDSIPKAPSGKILRKDLRARLANGLMN-
                            *.:: ****.**.:  *:*   ***: ******* :. .*
```

Figure 3.7 **Protein map of 4CL and MUSCLE alignment of the cassava candidate peptide sequences, Cassava4.1_005006, Cassava4.1_005014, Cassava4.1_004658 and Cassava4.1_004136, and the *A. thaliana* references, 4CL1 (At1g51680), 4CL2 (At3g21240), 4CL3 (At1g65060) and 4CL4 (At3g21230). A)** The 4CL signature pattern and nucleotide binding region (SSGTTGLPK). (**B)** GYGMTE is a nucleotide binding region. **C)** Aspartic acid (D) and arginine (R) are nucleotide binding sites. **D)** Lysine (K) is a nucleotide binding site which results in activity loss if mutated (Stuible *et al.*, 2000).

3.3.1.4 HYDROXYCINNAMOYL-COA SHIKIMATE:QUINATE HYDROXYCINNAMOYL-TRANSFERASE (HCT)

The peptide sequence of the *A. thaliana* HCT reference gene, At5g48930, was used as a reference to identify homologous genes in the cassava genome. Two cassava sequences were identified (Cassava4.1_008045 and Cassava4.1_008063) as good candidates as the *E*-values were 0 and they have a score of 755.7 and 750.7 respectively (Appendix II). These peptide sequences were then compared with HCT sequences from dicotyledons (cassava (*Manihot esculenta*), *Arabidopsis thaliana*, *Populus tricocarpa*, *Cucumis sativus* and *Brassica rapa*), monocotyledons (*Sorghum bicolor*, *Brachypodium distachyon*, *Oryza sativa*, *Setaria italica* and *Panicum virgatum*) and bryophyte (*Physcomitrella patens*) species in a Neighbour-Joining consensus phylogenetic tree with a bootstrap replicate of 1000 (Figure 3.8). The phylogenetic analysis shows that Cassava4.1_008045 and 008063 are closely related to the *A. thaliana* HCT gene reference.

To confirm that the retrieved cassava genes are good candidates that will likely share the same functions as the *A. thaliana* HCT reference gene, the motif HxxxD which is part of the active site (histidine (H)) has been identified in the candidate genes (Figure 3.9A). The x indicates any amino acid is acceptable. The aspartic acid (D) (Figure 3.9B) is also an active site (proton acceptor) (Levsh *et al.*, 2016).

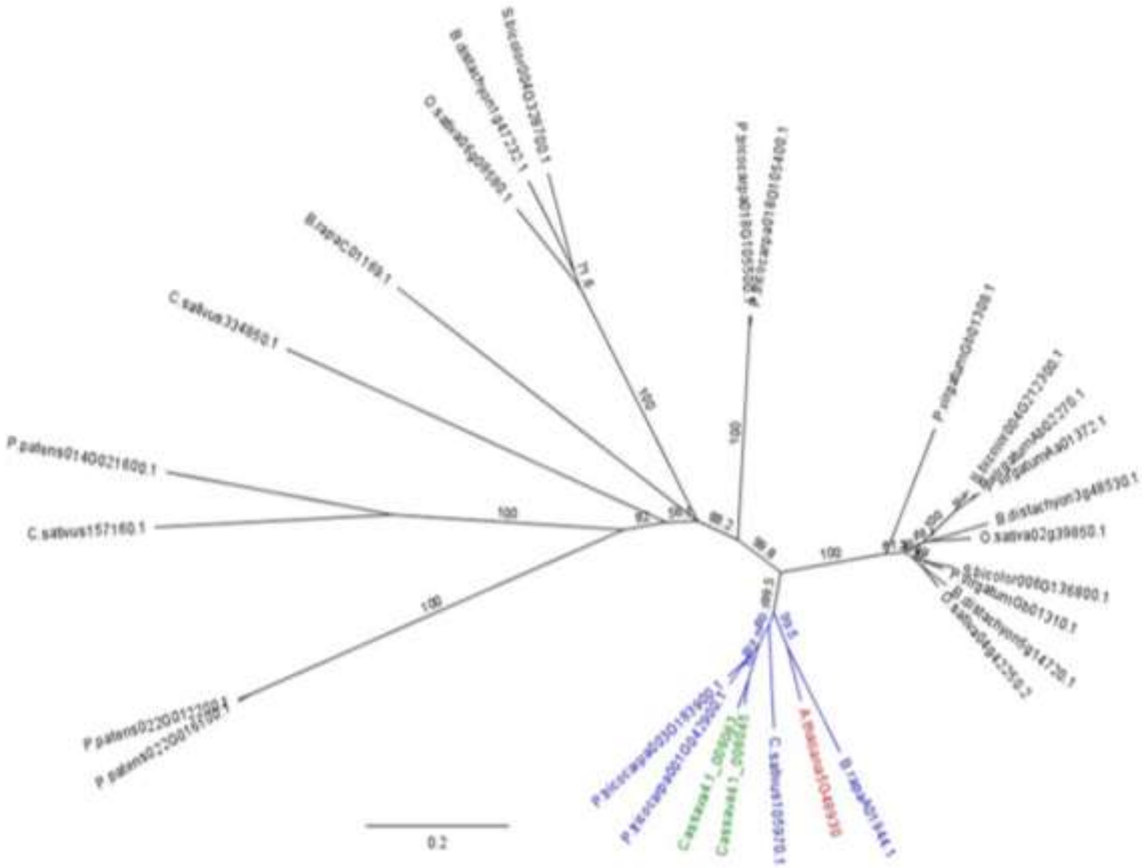


Figure 3.8 **Neighbour joining consensus phylogenetic analysis of hydroxycinnamoyl-CoA shikimate:quinate hydroxycinnamoyl-transferase (HCT) peptide sequences with 1000 bootstrap replicates.** The *A. thaliana* reference sequence (At5g48930) is highlighted in red and the cassava candidate genes' peptide sequences (Cassava4.1_008045 and Cassava4.1_008063) are in green. The blue highlight shows the node from which the cassava candidate genes was selected from. The plants species analysed are the dicots, cassava (*Manihot esculenta*), *Arabidopsis thaliana*, *Populus tricocarpa*, *Cucumis sativus* and *Brassica rapa*, the monocots, *Sorghum bicolor*, *Brachypodium distachyon*, *Oryza sativa*, *Setaria italica* and *Panicum virgatum* and bryophyte species, *Physcomitrella patens*. The scale bar represents 0.2 substitutions per amino acid.

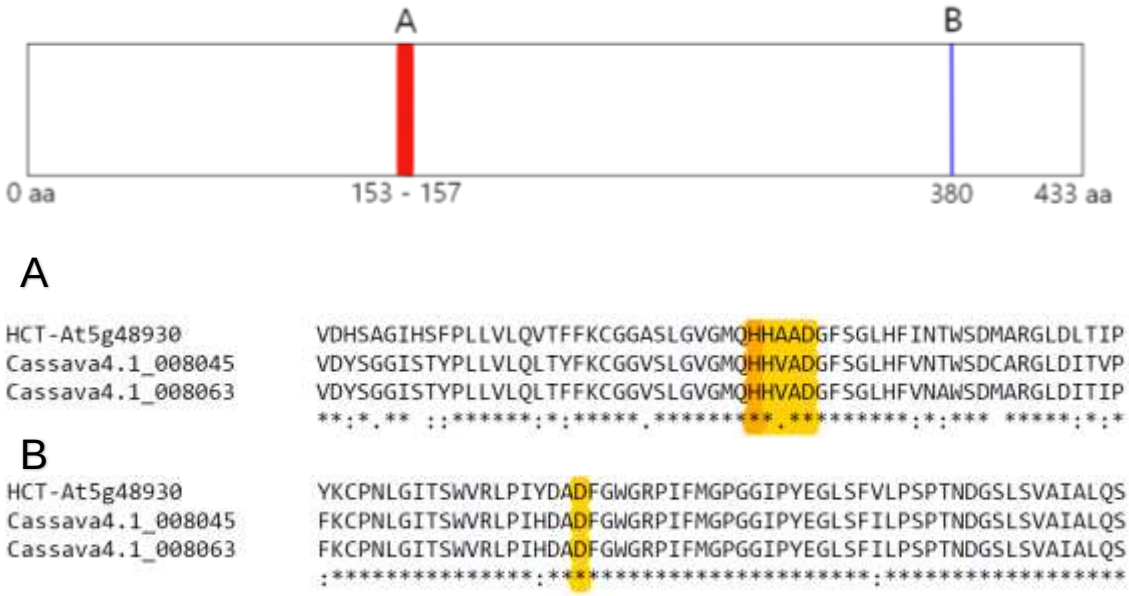


Figure 3.9 **Protein map of HCT and MUSCLE alignment of the cassava candidate peptide sequences, Cassava4.1_008045 and Cassava4.1_008063, and the *A. thaliana* reference, At5g48930. A)** HXXXD a motif of the HCT enzyme with histidine (H) being the active site. **B)** An active site (Levsh *et al.*, 2016).

3.3.1.5 P-COUMAROYL SHIKIMATE 3' HYDROXYLASE (C3'H)

C3'H is encoded by one gene in *A. thaliana* (At2g40890), which has been used as a reference to identify homologous genes in the cassava genome. Only one candidate gene from cassava (Cassava4.1_005910) was identified as having a low enough E value (0.0) and high score (853.2). The cassava candidate peptide sequence was then compared with C3'H sequences from dicotyledons (cassava (*Manihot esculenta*), *Arabidopsis thaliana*, *Populus tricocarpa*, *Solanum tuberosum* and *Brassica rapa*), monocotyledons (*Sorghum bicolor*, *Brachypodium distachyon*, *Oryza sativa* and *Setaria italica*) and a bryophyte (*Physcomitrella patens*) in a Neighbour-Joining consensus phylogenetic tree with a bootstrap replicate of 1000 (Figure 3.10).

To confirm that the retrieved cassava genes are good candidates that will likely share the same functions as the *A. thaliana* C3'H reference gene, the C3'H signature pattern which is the cytochrome P450 cysteine heme-iron ligand signature, [FW]-[SGNH]-x-[GD]-{F}-[RKHPT]-{P}-C-[LIVMFAP]-[GAD], where the cysteine (C) in the signature is the heme iron ligand has been investigated in cassava. The signature is conserved in the cassava candidate gene Cassava4.1_005910 (Figure 3.11). The glycine (G) amino acid which causes loss of enzyme activity in *A. thaliana* when mutated to aspartic acid (D)

(Franke *et al.*, 2002b) is also conserved between the reference sequence and the cassava candidate sequence.

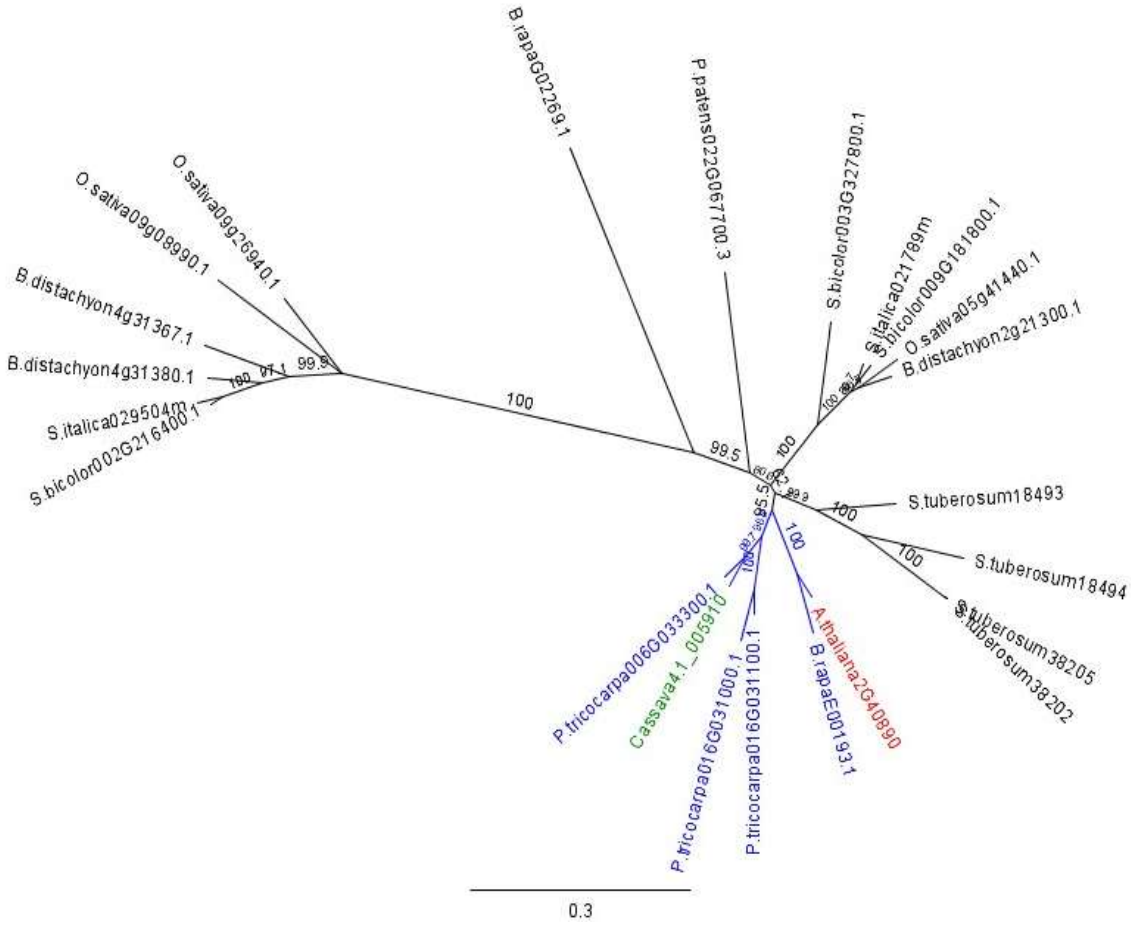


Figure 3.10 **Neighbour joining consensus phylogenetic analysis of *p*-coumaroyl shikimate 3' hydroxylase (C3'H) peptide sequences with 1000 bootstrap replicates.** The *A. thaliana* reference sequence (At2g40890) is highlighted in red and the cassava candidate gene's peptide sequence (Cassava4.1_005910) is in green. The blue highlight shows the node from which the cassava candidate gene was selected from. The peptide sequences from other plants species analysed are the dicots, cassava (*Manihot esculenta*), *Arabidopsis thaliana*, *Populus tricocarpa*, *Solanum tuberosum* and *Brassica rapa*, the monocots, *Sorghum bicolor*, *Brachypodium distachyon*, *Oryza sativa* and *Setaria italica* and a bryophyte, *Physcomitrella patens*. The scale bar represents 0.3 substitutions per amino acid.

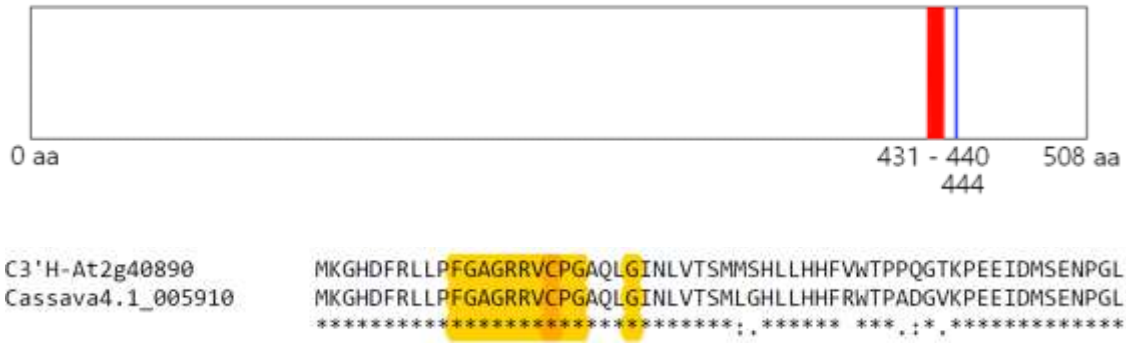


Figure 3.11 **Protein map of C3'H and MUSCLE alignment of the cassava candidate peptide sequence, Cassava4.1_005910, and the *A. thaliana* reference sequence, At2g40890.** FGAGRRVCPG where cysteine (C) is the heme iron ligand is the signature pattern of cytochrome P450 cysteine heme-iron ligand which includes the C3'H enzyme. The highlighted glycine (G) three sequences from the highlighted signature pattern is shown to cause loss of C3'H activity in *A. thaliana* when mutated to aspartic acid (D) (Franke *et al.*, 2002b).

3.3.1.6 CAFFEOYL COA 3-*O*-METHYLTRANSFERASE (CCoAOMT)

CCoAOMT is encoded by one gene (At4g34050) in *A. thaliana* which has been used as a reference to identify homologous genes in the cassava genome. Two genes from cassava (Cassava4.1_011832 and Cassava4.1_014783) were selected as candidates as the E values were low enough, 7.40E-162 and 4.60E-159 with a score of 453.4 and 446.4 respectively. The peptide sequences of the cassava candidate genes were then compared with CCoAOMT peptide sequences from dicotyledons (cassava (*Manihot esculenta*), *Arabidopsis thaliana*, *Populus tricocarpa*, *Solanum tuberosum*, and *Boechera stricta*), monocotyledons (*Sorghum bicolor*, *Brachypodium distachyon*, *Oryza sativa* and *Setaria italica*) and a bryophyte (*Physcomitrella patens*) in a Neighbour-Joining consensus phylogenetic tree with a bootstrap replicate of 1000 (Figure 3.12).

To confirm that the retrieved cassava genes are good candidates and will likely share the same functions as the *A. thaliana* CCoAOMT reference gene, both the domain and binding sites of the CCoAOMT enzyme are present in the cassava candidate peptide sequence (Figure 3.13).

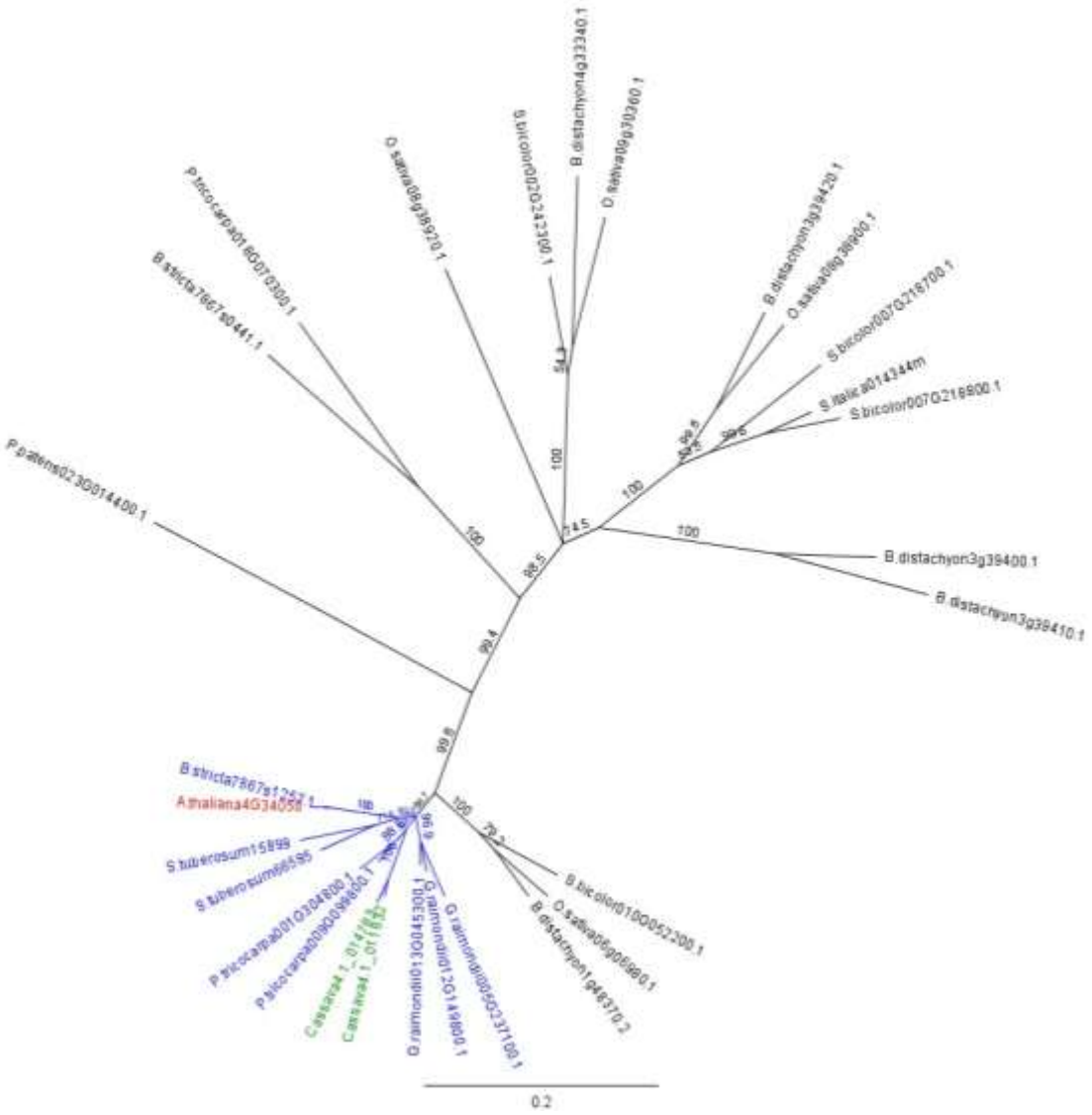


Figure 3.12 **Neighbour joining consensus phylogenetic analysis of caffeoyl CoA 3-*O*-methyltransferase (CCoAOMT) peptide sequences with 1000 bootstrap replicates.** The *A. thaliana* reference sequence (At4g34050) is highlighted in red and the cassava candidates' sequences (Cassava4.1_011832 and Cassava4.1_014783) are in green. The blue highlight shows the node from which the cassava candidate genes were selected from. The plants species analysed are the dicots, cassava (*Manihot esculenta*), *Arabidopsis thaliana*, *Populus tricocarpa*, *Solanum tuberosum*, and *Boechera stricta*, the monocots, *Sorghum bicolor*, *Brachypodium distachyon*, *Oryza sativa* and *Setaria italica* and a bryophyte, *Physcomitrella patens*. The scale bar represents 0.2 substitutions per amino acid.

Figure 3.13 **Protein map of CCoAOMT and MUSCLE alignment of the cassava candidate peptide sequence, Cassava4.1_011832 and Cassava4.1_014783, and the *A. thaliana* reference sequence, At4g34050.** The highlighted peptide sequences are all binding sites of the CCoAOMT enzyme.

3.3.1.7 CAFFEIC ACID *O*-METHYLTRANSFERASE (COMT)

The POMT9 peptide sequence from *P. deltoids* which was found to produce scopoletin from esculetin in *E. coli* was used to identify the *A. thaliana* gene, At5g54160, also known as COMT1. This was then used as a reference to identify homologous genes in the cassava genome. Two genes from cassava, Cassava4.1_010187 and Cassava4.1_010203, were selected as candidates because the *E*-values were both 0 and the scores were 612.5 and 606.3 respectively. These candidate genes' peptide sequences were then compared with other COMT sequences from dicotyledons (cassava (*Manihot esculenta*), *Arabidopsis thaliana*, *Populus deltoids*, *Solanum tuberosum*, *Brassica rapa* and *Malus domestica*), monocotyledons (*Oryza sativa*, *Ananas comosus* and *Zea mays*) and bryophyte (*Physcomitrella patens*) species in a Neighbour-Joining consensus phylogenetic tree with a bootstrap replicate of 1000 (Figure 3.14).

To confirm that the retrieved cassava genes are good candidates and will likely share the same functions as the *A. thaliana* COMT1 reference gene, both the binding and active sites of the COMT1 enzyme are present in the cassava candidates' peptide sequences (Figure 3.15). The two glutamate amino acids highlighted in dark orange are active sites that have been confirmed by Byeon *et al.* (2014).

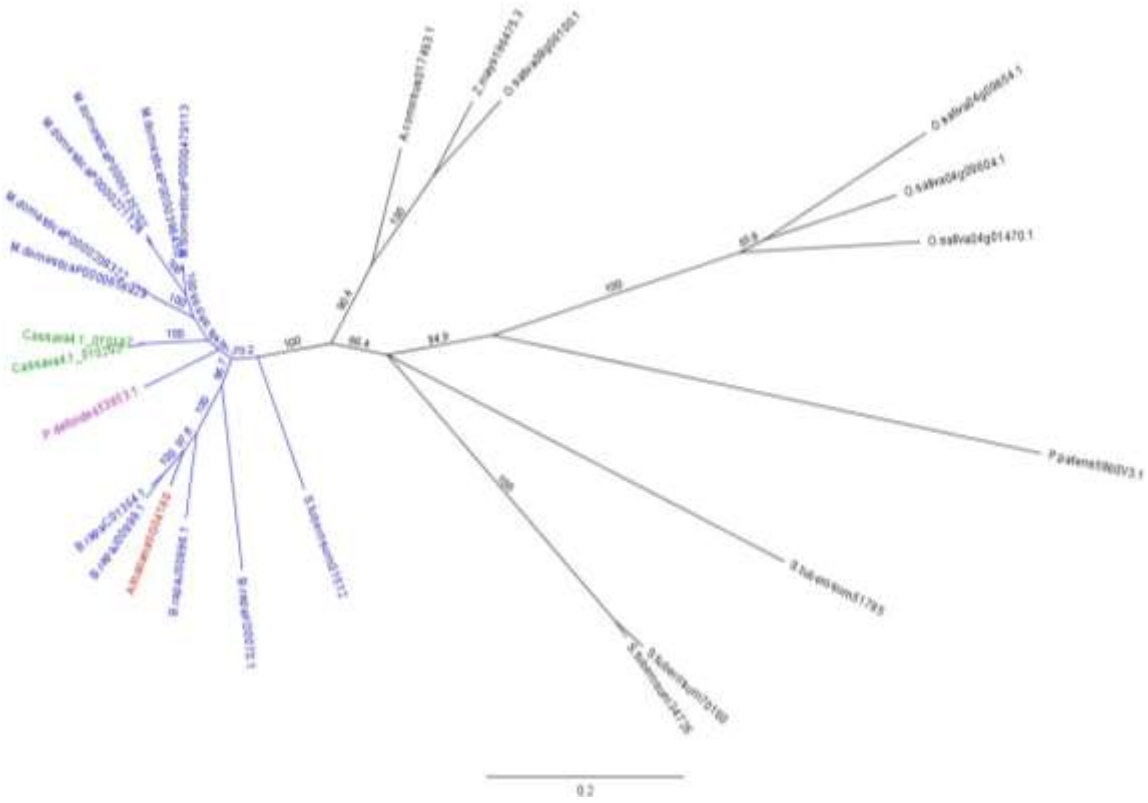


Figure 3.14 **Neighbour joining consensus phylogenetic analysis of caffeic acid *O*-methyltransferase (COMT) peptide sequences with 1000 bootstrap replicates.** The *A. thaliana* reference sequence (At5g54160) is highlighted in red, the *P. deltoids* sequence is highlighted in purple and the cassava candidate genes' peptide sequences (Cassava4.1_010187 and Cassava4.1_010203) are in green. The blue highlight shows the node from which the cassava candidate sequences were selected from. The plants species analysed are the dicots, cassava (*Manihot esculenta*), *Arabidopsis thaliana*, *Populus deltoids*, *Solanum tuberosum*, *Brassica rapa* and *Malus domestica*), monocots, *Oryza sativa*, *Ananas comosus* and *Zea mays*, and bryophyte, *Physcomitrella patens*. The scale bar represents 0.2 substitutions per amino acid.

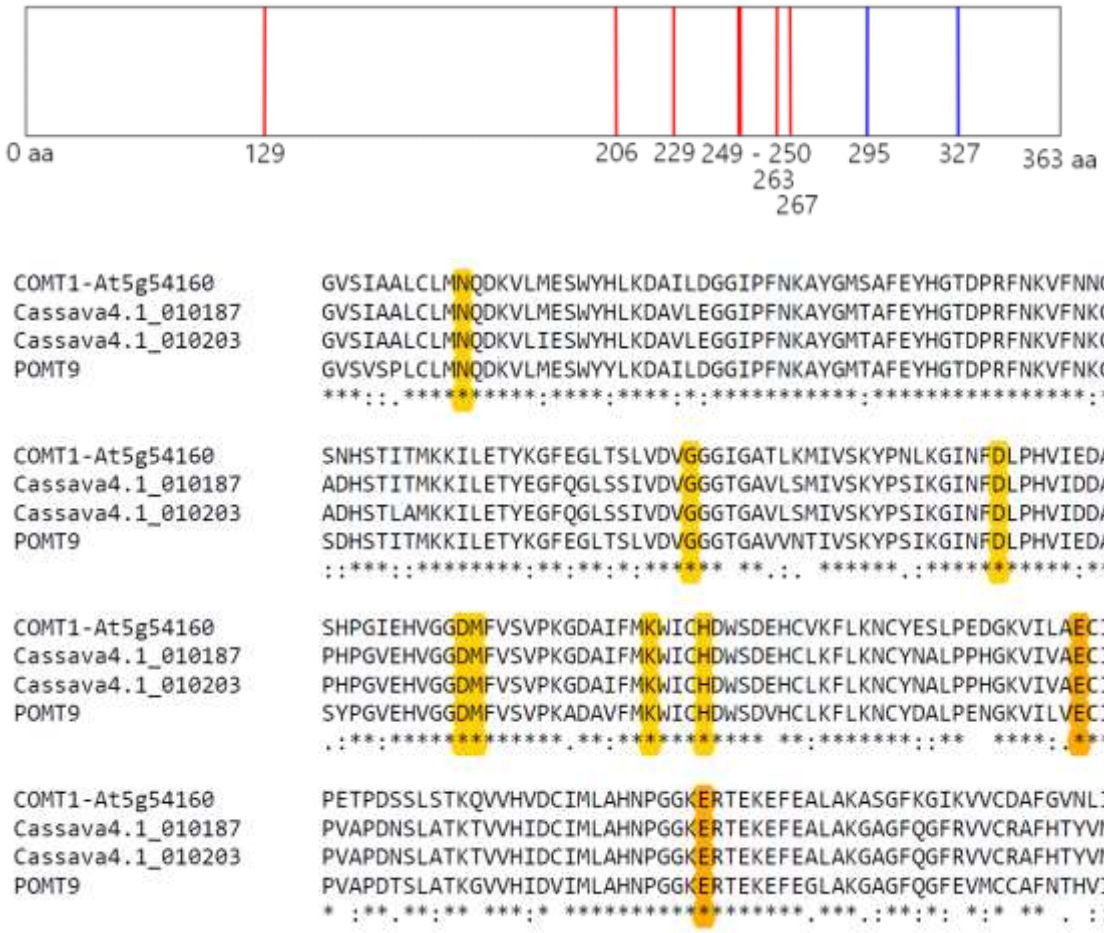


Figure 3.15 **Protein map of COMT and MUSCLE alignment of the cassava COMT candidate peptide sequence, Cassava4.1_010187 and Cassava4.1_010203, and the *A. thaliana* reference sequence, At5g54160 and POMT9 from *Populus deltoids*.** The light orange highlighted amino acids are binding sites of the COMT enzyme. The darker orange highlights are active sites of COMT (Byeon *et al.*, 2014).

3.3.2 CASSAVA CANDIDATE GENE EXPRESSION PROFILES

Gene expression of cassava candidate genes was analysed to gain an insight into the role the genes played during PPD in different tissues (storage root and leaf). cDNA was synthesised from storage roots over the course of PPD development (0 hour (fresh), 24 hours, 48 hours and 72 hours) and leaves. Gene specific primers were designed for all candidate genes except for CCoAOMT candidate genes (Cassava4.1_011832 and Cassava4.1_014783) and COMT candidate genes (Cassava4.1_010187 and Cassava4.1_010203) as these genes were too similar to one another to allow primers to target the specific individual gene (Table 2.1). The primers were designed using the Geneious v8.0.5 software and they span an intron so that the product size will be different for genomic DNA and cDNA to be certain that the amplicon for the gene expression analysis will definitely be from cDNA and not genomic DNA. Other criteria for the primers were that the melting temperature (Tm) must be above 52ºC and below 62ºC with a GC content of 40% to 62%. Hairpins and dimer structures were also examined to ensure that the temperatures in which these structures will form were considerably lower than the Tm. The properties were checked through OligoAnalyzer 3.1 from Integrated DNA Technologies (https://www.idtDNA.com/calc/analyzer). All primers were first tested on cassava genomic DNA and optimised through gradient RT-PCR and were confirmed to be working before being performed on cDNA. The RT-PCR results are qualitative and not quantitative and only show whether the gene is expressed or not, so expression levels are not comparable (Table 3.1). Cassava candidate genes were renamed to facilitate discussion and referencing (Table 3.2).

Table 3.1 **RT-PCR results performed on cassava storage roots over the PPD timescale (0 hour, 24 hours, 48 hours and 72 hours) and leaf with gene specific primers excluding CCoAOMT and COMT which used primers that targeted both genes in the family.** A tick (✓) represents the gene is expressed and a cross (×) represents no gene expression.

Gene Family	Gene	Leaf	Root 0 h	Root 24 h	Root 48 h	Root 72 h
PAL	Cassava4.1_034377	✓	×	✓	✓	✓
	Cassava4.1_002628	✓	✓	✓	✓	✓
	Cassava4.1_002709	✓	✓	✓	✓	✓
	Cassava4.1_028434	✓	×	✓	✓	✓
	Cassava4.1_002591	✓	✓	✓	✓	✓
	Cassava4.1_003117	✓	×	✓	✓	✓
C4H	Cassava4.1_005978	✓	✓	✓	✓	✓
4CL	Cassava4.1_005006	✓	×	✓	✓	✓
	Cassava4.1_005014	✓	✓	✓	✓	✓
	Cassava4.1_004658	✓	×	✓	✓	✓
	Cassava4.1_004136	×	×	×	×	×
HCT	Cassava4.1_008045	✓	×	✓	✓	✓
	Cassava4.1_008063	✓	✓	✓	✓	✓
C3'H	Cassava4.1_005910	✓	✓	✓	✓	✓
CCoAOMT	Cassava4.1_011832	✓	✓	✓	✓	✓
	Cassava4.1_014783					
COMT	Cassava4.1_010187	✓	✓	✓	✓	✓
	Cassava4.1_010203					

Table 3.2 **Names given to cassava candidate genes encoding relevant enzymes.**

Enzyme	Cassava Candidate Gene	New Name
PAL	Manes.08G008400.1 (cassava4.1_002628m.g)	MePAL1
	Manes.10G047500.1 (cassava4.1_034377m.g)	MePAL2
	Manes.09G063700.1 (cassava4.1_028434m.g)	MePAL3
	Manes.07G098700.1 (cassava4.1_002709m.g)	MePAL4
	Manes.04G018000.1 (cassava4.1_002591m.g)	MePAL5
	Manes.16G098200.1 (cassava4.1_003117m.g)	MePAL6
C4H	Manes.18G126900.1 (cassava4.1_005978m.g)	MeC4H1
4CL	Manes.11G071800.1 (cassava4.1_005006m.g)	Me4CL1
	Manes.04G095300.1 (cassava4.1_005014m.g)	Me4CL2
	Manes.09G127000.1 (cassava4.1_004658m.g)	Me4CL3
	Manes.08G066200.1 (cassava4.1_004136m.g)	Me4CL4
HCT	Manes.11G067800.1 (cassava4.1_008045m.g)	MeHCT1
	Manes.04G101700.1 (cassava4.1_008063m.g)	MeHCT2
C3'H	Manes.08G063400.1 (cassava4.1_005910m.g)	MeC3'H1
CCoAOMT	Manes.10G078800.1 (cassava4.1_011832m.g)	MeCCoAOMT1
	Manes.07G075700.1 (cassava4.1_014783m.g)	MeCCoAOMT2
COMT	Manes.01G043700.1 (cassava4.1_010187m.g)	MeCOMT1
	Manes.01G043600.1 (cassava4.1_010203m.g)	MeCOMT2

All PAL candidate genes are expressed in the leaf and are not exclusive to the root which is expected as it is the first enzyme in the phenylpropanoid metabolism. Three genes, MePAL1, MePAL4 and MePAL5 are constitutively expressed in the storage root over the PPD timescale and are not induced by wounding. MePAL2, MePAL3 and MePAL6 are likely to be induced by wounding in the storage root as they are not expressed in fresh (0h) storage roots and expression is only seen from 24 hours post-harvest.

There is only one candidate gene for C4H which is MeC4H1. Probably due to the importance of the C4H enzyme, this gene is constitutively expressed in all tissue types over the PPD time course and in leaf as well (Figure 3.16).

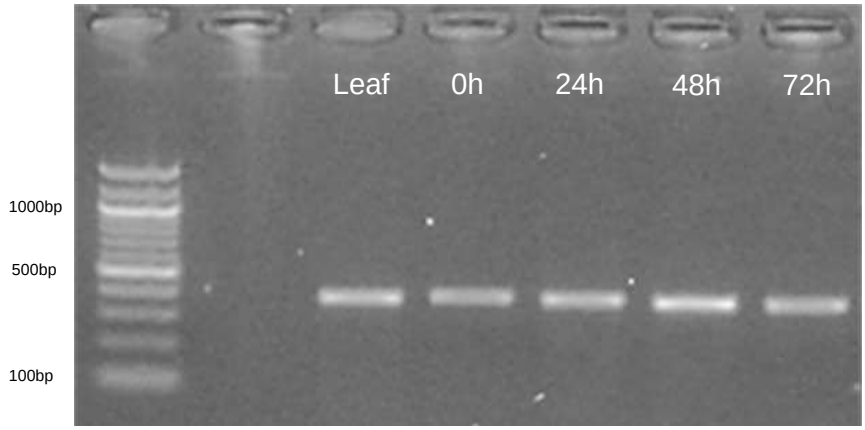


Figure 3.16 **Expression of the C4H candidate gene, Cassava4.1_005978 (MeC4H1).** This gene is expressed constitutively in the storage root from fresh (0h) to 72h post-harvest as well as the leaf. The amplicon size is 397bp.

The 4CL gene, Me4CL4, has no expression in all tissues but is amplified in genomic DNA, which suggests that this could be due to gene redundancy after a genome duplication event in cassava's history. Two 4CL candidate genes, Me4CL3 (Figure 3.17) and Me4CL1, are induced by wounding in the storage root as there is no expression in fresh cassava root. The other 4CL gene, Me4CL2, are constitutively expressed in the leaf and storage root.

Figure 3.17 **Expression of the 4CL candidate gene, Cassava4.1_004658 (Me4CL3).** This gene is expressed in leaf and 24h, 48h and 72h post-harvest cassava storage root and not in fresh root.

One of the HCT candidate gene, MeHCT2, is constitutively expressed in the leaf and storage root and the other (MeHCT1) is only expressed in the leaf and 24h, 48h and 72h post-harvest storage root.

The only candidate gene for C3'H is MeC3'H1 and it is expressed in the leaf and constitutively in the storage root.

Due to the similarity of the two candidate genes of CCoAOMT (MeCCoAOMT1 and MeCCoAOMT2), gene specific primers cannot be designed which is why the primers are designed to amplify both genes. The CCoAOMT genes are expressed in the leaf and constitutively expressed in the storage root.

Like the CCoAOMT candidate genes, the candidate genes for COMT are also highly identical (MeCOMT1 and MeCOMT2), hence primers are designed to amplify both genes. The RT-PCR results show that the COMT genes are constitutively expressed in the storage root over the development of PPD and in the leaf.

3.4 DISCUSSION

Enzymes leading to the biosynthesis of scopoletin were identified from the phenylpropanoid metabolism and only the best candidates with the highest identity and score were taken forward as cassava candidate genes (Table 3.2). Certain enzymes have a larger gene family; for example, PAL and 4CL have a number of isozymes whereas C4H, HCT, C3'H, CCoAOMT and COMT are only encoded by a single gene in *A. thaliana*. Cassava is an allotetraploid (Nassar, 2002), which explains why cassava has a larger number of isoforms for all enzymes except C4H and C3'H compared to *A. thaliana*. Genome duplication events can lead to duplicated genes becoming a pseudogene through loss-of-function mutations which is likely to be the case for Me4CL4.

The ideal characteristics of a candidate gene for RNAi knock-down will be root specific, expressed throughout the PPD timescale and come from a small gene family. The likely candidates will be found further along the biosynthesis pathway of scopoletin as knocking down an enzyme found early in the pathway will have more adverse effects on the plant's fitness. In cassava, there are six PAL isoforms which would mean that to knock-down PAL activity, the RNAi construct will have to target all the isoforms and this may be hard or even impossible to achieve. PAL is not an ideal option as it is the first enzyme of the phenylpropanoid metabolism and is extremely important to the plant. RT-PCR results show that all candidate genes for PAL are expressed in the leaf and only three genes

(MePAL2, MePAL3 and MePAL6) are induced by wounding. Huang *et al* (2010) also described residual PAL activity in the *A. thaliana pal1 pal2 pal3 pal4* quadruple mutant which suggests there are other PAL-like genes in the genome.

The second enzyme in the phenylpropanoid metabolism is C4H which is only encoded by a single gene in both *A. thaliana* and cassava genomes. *A. thaliana* mutants for C4H have adverse phenotypes including stunted growth and reduced lignin deposition. This gene is so important that a complete knock-out is lethal to the plant due to its importance in vascular development (Bell-Lelong *et al.*, 1997). The expression of C4H in cassava from the RT-PCR results affirms its importance as it is strongly and constitutively expressed in both leaf tissue and storage root over the PPD development timescale. Due to the importance of this enzyme, knocking down its gene will likely affect the plant's survival significantly.

4CL is the third enzyme in the pathway and it has five isoforms (one may be a pseudogene) and is very important to plant survival. It is responsible for the biosynthesis of *p*-coumaroyl CoA which is the starting point of many secondary metabolites including lignin (Lee *et al.*, 1995). Apart from Me4CL4 which is not expressed at all and is likely to be a pseudogene, all the other genes are expressed in the leaf and storage root, with Me4CL1 and Me4CL3 being induced by wounding. Due to the high number of isoforms, knocking-down them all with the RNAi construct will be a challenge and the importance of this enzyme would greatly reduce the plant's integrity and growth.

HCT has two isoforms in which one gene (MeHCT2) is constitutively expressed in all tissues although 0h storage root and 72h fibrous root appear to have lower expression levels compared to the other tissues and the other gene (MeHCT1) is exclusive to the root and is induced by wounding. RNAi silencing was performed in *A. thaliana* and the mutant was shown to be dwarfed (Hoffmann *et al.*, 2004). C3'H is only encoded by a single gene in the cassava genome and is expressed in both leaf tissue and storage root tissue over the PPD timescale. A missense mutation in the *A. thaliana ref8* mutant has stunted growth, accumulates flavonoids and shows a complete loss of apical dominance (Bonawitz and Chapple, 2013). Although these two enzymes have a small gene family which is ideal for RNAi silencing, the studies performed in *A. thaliana* mutants show that knocking out these genes negatively affects the plant's growth. This would not be useful in cassava as it may adversely affect storage root development and the plant's fitness.

The last two enzymes responsible for scopoletin biosynthesis, apart from F6'H1 which has been studied (Liu *et al.*, 2017), are CCoAOMT (MeCCoAOMT1 and MeCCoAOMT2) and COMT (MeCOMT1 and MeCOMT2). These two enzymes are good potential

candidates for RNAi silencing due to having a small gene family with a high similarity between the genes. RT-PCR results show constitutive expression in the storage roots over the PPD development timescale and the leaf for both enzymes. The functions of these genes have also been confirmed through complementation of CCoAOMT and COMT T-DNA knockout lines for *A. thaliana* (Alhalasseh, 2017). Unlike other enzymes in the phenylpropanoid metabolism where the *A. thaliana* knockout resulted in growth defects and even fatality, the CCoAOMT and COMT knockout mutants did not show any adverse effects in its growth and phenotype. However, studies showed that knocking out CCoAOMT led to collapsed xylem vessels, lower stem lignin content and reduction in coumarin biosynthesis, in particular, scopoletin (Kai *et al.*, 2008; Do *et al.*, 2007). The COMT T-DNA knockout mutant COMT1 was missing S lignin as lignin is made up of three subunits: H, G and S lignin (Boerjan *et al.*, 2003), and reduced sinapoylmalate levels in leaves, stem and seedlings but showed no morphological differences (Goujon *et al.*, 2003). These two enzymes are the best candidates among the other phenylpropanoid enzymes as they have small gene families and the *A. thaliana* knockout mutants do not show a reduction in fitness despite some changes in the plants' phenotype.

More information on these genes and strategies involving RNAi silencing will be discussed in chapter 4.

4 GENERATING CASSAVA KNOCK-DOWN LINES THROUGH RNAI SILENCING

4.1 INTRODUCTION

Scopoletin plays a key role during the development of PPD in cassava storage roots and by inhibiting the accumulation of scopoletin, PPD can be delayed (Liu *et al.*, 2017). Studies have shown that phenylpropanoid metabolism is highly conserved in plants, including cassava and *A. thaliana* (Tohge *et al.*, 2013). Cassava has three possible pathways leading to scopoletin biosynthesis, with one dominant pathway through which the majority of scopoletin is biosynthesised through (Bayoumi *et al*., 2008). Apart from the dominant pathway, the other alternative pathways involved in scopoletin biosynthesis have not been thoroughly explored in *A. thaliana*, but it is likely that the major routes will be shared between cassava and *A. thaliana* (Kai *et al*., 2006). This makes *A. thaliana* a good reference to identify homologous genes of interest in cassava.

The onset of PPD in cassava storage roots was delayed when scopoletin biosynthesis was partially inhibited with RNAi which led to the reduced accumulation of scopoletin in the storage roots. This was shown by Liu *et al.* (2017) where RNAi knock-down cassava lines targeting the family of F6'H genes (Me10291, Me33240, Me10292, Me10381, Me27567, Me10376 and Me30526) in cassava were generated. Due to the considerable number of genes in the family, the RNAi construct may be unable to target all the cassava F6'H genes, although it should have been able to target most of the genes due to the high similarities between them. The reason this gene family was selected to be inhibited was because when *F6'H1* was knocked-out in *A. thaliana* by a T-DNA insertion, the mutant did not accumulate scopoletin (Kai *et al.*, 2008). Unfortunately, scopoletin biosynthesis is more complicated in cassava as there are three pathways and larger gene families, hence a complete inhibition of scopoletin accumulation has not been achieved. Phenylpropanoid metabolism is a complex network of pathways and understanding the other genes involved in scopoletin biosynthesis may provide answers as to why the cassava F6'H knock-down lines still accumulated scopoletin. Homologous genes in the phenylpropanoid metabolism involved in scopoletin biosynthesis were identified in the previous chapter with the aim to select for candidates to be knocked-down with an RNAi construct. Caffeoyl CoA *O*-methyltransferase (MeCCoAOMT1 and MeCCoAOMT2) and caffeic acid *O*-methyltransferase (MeCOMT1 and MeCOMT2) were

selected as ideal candidates to be targeted for RNAi silencing with the aim of further understanding the biosynthesis of scopoletin in cassava storage roots.

To confirm the functional identity of CCoAOMT and COMT cassava candidate genes, *A. thaliana* T-DNA mutants with the reference genes knocked-out were complemented with said genes. Due to the high similarity between MeCCoAOMT1 and MeCCoAOMT2, and between MeCOMT1 and MeCOMT2, only MeCCoAOMT1 and MeCOMT1 were selected to complement the *A. thaliana* T-DNA knock-out mutants (Alhalasseh, 2017). Scopoletin and scopolin concentrations were reduced by 60% in the root extracts of the *A. thaliana* CCoAOMT mutant plant and through complementation with MeCCoAOMT1, the accumulation of both scopoletin and scopolin were restored. In the *A. thaliana* COMT mutant, scopoletin and scopolin concentrations were reduced by 20% and the accumulation of both were also restored following the complementation with MeCOMT1 (Alhalasseh, 2017). This proves that the cassava candidate genes selected share the same functions as the reference genes from *A. thaliana*. Individual members of the gene family for CCoAOMT and COMT were not tested for functionality due to the difficulty in designing gene specific primers because of the high similarity between them. This is ideal as both genes are aimed to be knocked-down by the RNAi construct and the higher the similarity between them, the greater the chances of both genes being targeted.

To understand the biosynthesis of scopoletin in cassava better, in addition to the F6'H cassava RNAi lines (pRNAi-F6'H) generated by Liu *et al.* (2017), two RNAi constructs were generated, an RNAi construct targeting both CCoAOMT and F6'H (pRNAi-CCoAOMT/F6'H) and another RNAi construct targeting all CCoAOMT, F6'H and COMT (pRNAi-CCoAOMT/F6'H/COMT) (Figure 4.1). These transgenic cassava lines should be able to give answers as to why scopoletin still accumulates in the pRNAi-F6'H lines, and provide further understanding into the pathways responsible for scopoletin biosynthesis through comparison of scopoletin accumulation in each of the three RNAi mutants during PPD. For example, if the double construct (2x), pRNAi-CCoAOMT/F6'H, lines accumulate less scopoletin during PPD than pRNAi-F6'H (1x) and pRNAi-CCoAOMT/F6'H/COMT (3x), this would suggest that the pRNAi-F6'H construct was not able to inhibit the whole F6'H gene family and one or more gene family members are still being expressed thus leading to scopoletin biosynthesis, and the other two pathways do not substantially contribute to scopoletin accumulation. On the other hand, if the triple construct (3x), pRNAi-CCoAOMT/F6'H/COMT, lines accumulate less scopoletin during PPD compared to pRNAi-F6'H (1x) and pRNAi-CCoAOMT/F6'H (2x), this would mean that scopoletin is being biosynthesised through the other non-dominant pathways.

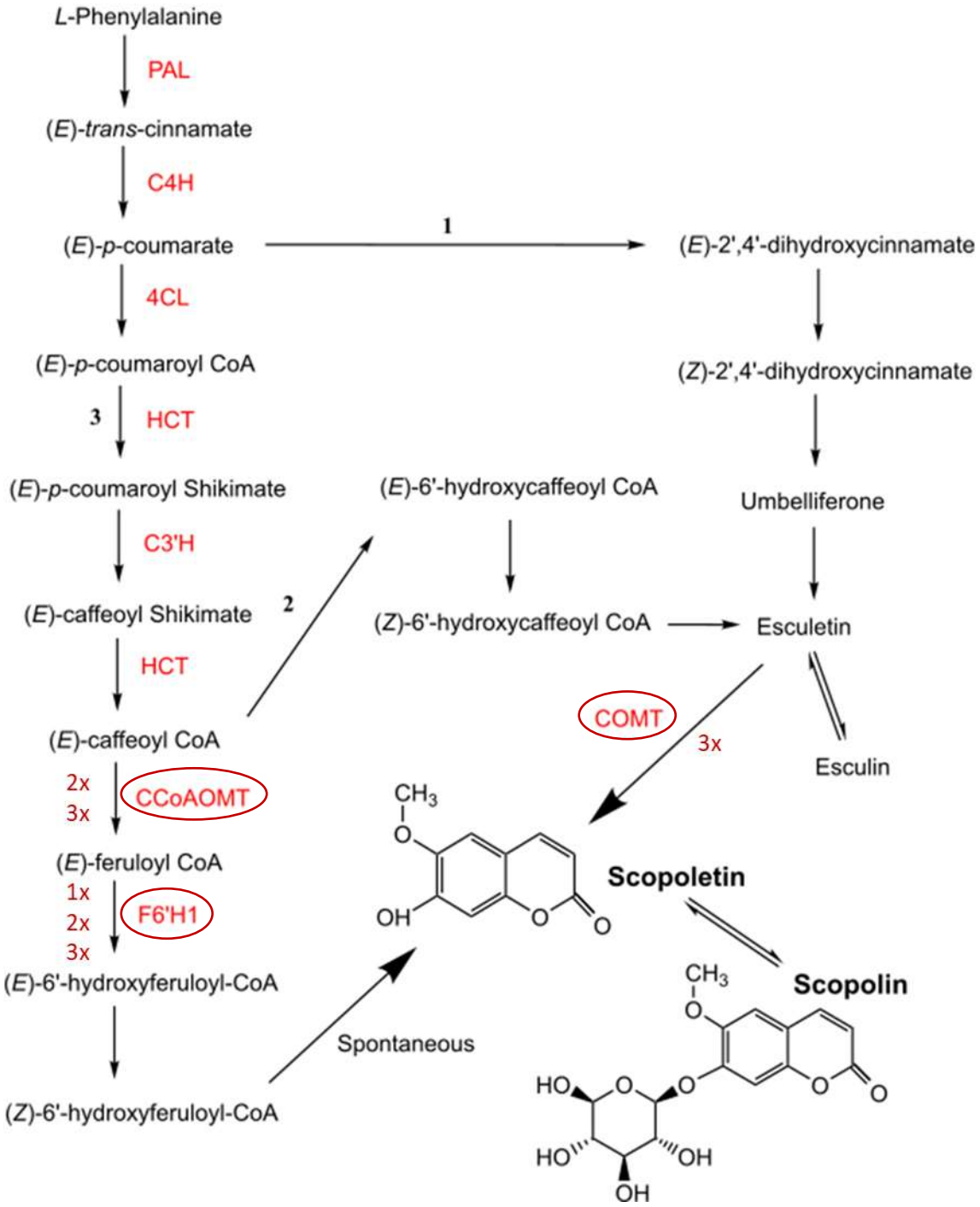


Figure 4.1 **Biosynthesis of scopoletin through three pathways in the phenylpropanoid metabolism.** The enzymes targeted to be knocked-down with RNAi (CCoAOMT, F6'H and COMT) have been circled in red. The lines in which these genes have been knocked-down are also listed: 1x (pRNAi-F6'H), 2x (pRNAi-CCoAOMT/F6'H) and 3x (pRNAi-CCoAOMT/F6'h/COMT).

4.2 AIMS OF STUDY AND RESEARCH STRATEGIES

The aim of this chapter is to generate transgenic cassava lines expressing pRNAi-CCoAOMT/F6'H and pRNAi-CCoAOMT/F6'H/COMT RNAi constructs. The following research strategies were implicated to achieve the aim of this study:

1. Design primers targeting regions of high identity for the CCoAOMT and COMT gene families.
2. Prepare the selected CCoAOMT, F6'H and COMT genes for RNAi using the pRNAi-GG protocol developed by Yan *et al.* (2012), which will be using the Golden Gate cloning technology.
3. Transform wild-type (TMS60444) cassava friable embryogenic callus (FEC) and identify individual transgenic lines.
4. Multiply and grow transgenic lines to obtain storage roots.

4.3 RESULTS

4.3.1 CREATION OF RNAi CONSTRUCTS

Interference hairpin RNAi (ihpRNA) constructs were made using the Golden Gate technology and the pRNAi-GG vector developed by Yan *et al.* (2012) (Figure 4.2). Firstly, primers were designed for the candidate genes of CCoAOMT and COMT, targeting regions with high identity (Figure 4.3). The primers designed will target the CCoAOMT region with 90% identity and the COMT region with 98% identity. The gene specific primers designed by Liu *et al.* (2017) to make the F6'H RNAi construct were reused.

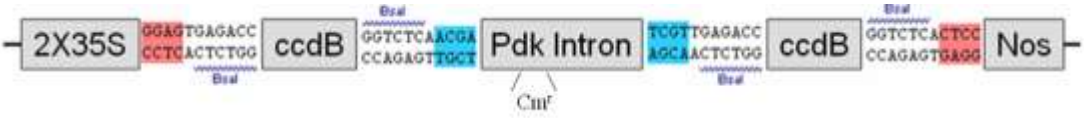


Figure 4.2 **The cassette of pRNAi-GG.** The two 35S CaMV promoter, two copies of the ccdB genes, the Pdk intron with the chloramphenicol-resistant gene (Cm^r) and the four BsaI sites with specific adaptors are cloned between the HindIII and SacI of the T-DNA vector pBI121. The nucleotides highlighted in different colours (blue and red) show the specific adaptors and the adaptors with the same colour have the same sequences but the opposite orientation (Yan et al., 2012).

A

```
Identities = 378/418 (90%),
Positives = 378/418 (90%), Gaps = 2/418 (0%)

cassava4.1_011832m extraction   524 CTCTTGCTCTTCCTGATGATGGAAAGATCTTGGCCATGGACATAAACAGAGAAAATTATG 58
                                    CTCTTGCTCTTCCTGATGATGGAAAGATCTTGGC ATGGACATAAACAGAGA AA TATG
cassava4.1_014783m extraction   290 CTCTTGCTCTTCCTGATGATGGAAAGATCTTGGCGATGGACATAAACAGAGAGAACTATG 34

cassava4.1_011832m extraction   584 AGCTGGGTCTGCCTGTTATCCAGAAAGCTGGTGTTGCACACAAGATTGATTTCAAAGAAG 64
                                    AGCTGGGTCTGCCTGTTATCCAGAAAGC GGTGTTGC CACAAG TTGATTTCA AGAAG
cassava4.1_014783m extraction   350 AGCTGGGTCTGCCTGTTATCCAGAAAGCAGGTGTTGCTCACAAGGTTGATTTCAGAGAAG 40

cassava4.1_011832m extraction   644 GCCCTGCTTTGCCTGTGCTTGATCAAATGATTGCAGAAGGCAAGTATCATGGAACTTTTG 70
                                    GCCCTGCTTTGCCTGT CTTGA CAAATGATTGC GAAGG AAGTA CATGGAACTTTTG
cassava4.1_014783m extraction   410 GCCCTGCTTTGCCTGTTCTTGACCAAATGATTGCCGAAGGGAAGTACCATGGAACTTTTG 46

cassava4.1_011832m extraction   704 ACTTCATATTTGTCGACGCTGATAAGGACAATTATCTGAACTACCACAAGAGATTGATTG 76
                                    A TTCAT TTTGT GACGCTGATAA GACAATTATCTGAACTACCACAAGAGATTGATTG
cassava4.1_014783m extraction   470 ATTTCATCTTTGTGGACGCTGATAAAGACAATTATCTGAACTACCACAAGAGATTGATTG 52

cassava4.1_011832m extraction   764 AATTAGTGAAAGTTGGAGGAGTGATCGGGTACGACAACACACTGTGGAACGGGTCTGTGG 82
                                    A TT GT AAAGTTGG GGA T ATCGG TACGACAACAC CT TGGAA GG TCTGTGG
cassava4.1_014783m extraction   530 AGTTGGTCAAAGTTGGTGGACTAATCGGCTACGACAACACCCTATGGAATGGTTCTGTGG 58

cassava4.1_011832m extraction   824 TGGCCCCACCTGATGCACCCCTAAGGAAGTATGTCAAGTACTACAGGGACTTTGTGATGG 88
                                    TGGC  CACCTGATGCACCC TAAGGAAGTATGT AAGTACTACAGGGACTT GTGATGG
cassava4.1_014783m extraction   590 TGGCAGCACCTGATGCACCCTTAAGGAAGTATGTTAAGTACTACAGGGACTTCGTGATGG 64

cassava4.1_011832m extraction   884 AGCTCAATAAGGCCCTACCAGCTGATCCAAGGATTGAGATTT-GCATGCTACCGGTTG 940
                                    AGCT AA AAGGC CTACCAGC GA CC AGGATTGAGATTT GCATGCT CCGGTTG
cassava4.1_014783m extraction   650 AGCTTAACAAGGCTCTACCAGCGGACCCCAGGATTGAGATTT-GCATGCTTCCGGTTG 706
```

B

```
Identities = 325/330 (98%),
Positives = 325/330 (98%), Gaps = 0/330 (0%)

cassava4.1_010203m extraction   123 GCTCGACCTGCTTGAAATCATCGGTAAAGCTGGTCCTGGTGCGTTCCTGTCGCCGACCGA 18
                                    GCTCGACCTGCTTGAAATCATCGGTAAAGCTGGTCCTGGTGCGTTC TGTCGCCGACCGA
cassava4.1_010187m extraction   123 GCTCGACCTGCTTGAAATCATCGGTAAAGCTGGTCCTGGTGCGTTCTTGTCGCCGACCGA 18

cassava4.1_010203m extraction   183 CATCGCTTCTCAGTTGCCCACCAGCAATCCAGATGCTCCAGTGATGTTGGACCGTATCTT 24
                                    CATCGCTTCTCAGTTGCCCACCAGCAATCCAGATGCTCCAGTGATGTTGGACCGTATCTT
cassava4.1_010187m extraction   183 CATCGCTTCTCAGTTGCCCACCAGCAATCCAGATGCTCCAGTGATGTTGGACCGTATCTT 24

cassava4.1_010203m extraction   243 GCGACTCTTAGCCAGCTACTCCATACTTAAATTTTCTCCGAGAACCCTTCCTGATGGCAA 30
                                    GCGACTCTTAGCCAGCTACTCCATACTTAAATTTTCTC GAGAACCCTTCCTGATGGCAA
cassava4.1_010187m extraction   243 GCGACTCTTAGCCAGCTACTCCATACTTAAATTTTCTCTGAGAACCCTTCCTGATGGCAA 30

cassava4.1_010203m extraction   303 AGTTGAGAGGCTGTATGGTCTTGCTCCTGTATGCAAATTCTTGACGAAAAATGAAGATGG 36
                                    AGTTGAGAGGCTGTATGGTCTTGCTCCTGTATGCAAATTCTTGACGAA AATGAAGATGG
cassava4.1_010187m extraction   303 AGTTGAGAGGCTGTATGGTCTTGCTCCTGTATGCAAATTCTTGACGAAGAATGAAGATGG 36

cassava4.1_010203m extraction   363 CGTCTCTATTGCCGCTCTTTGTCTGATGAACCAGGACAAAGTCCTCATAGAAAGCTGGTA 42
                                    CGTCTCTATTGCCGC CTTTGTCTGATGAACCAGGACAAAGTCCTCAT GAAAGCTGGTA
cassava4.1_010187m extraction   363 CGTCTCTATTGCCGCCCTTTGTCTGATGAACCAGGACAAAGTCCTCATGGAAAGCTGGTA 42

cassava4.1_010203m extraction   423 TCACTTGAAAGATGCAGTTCTTGAAGGAGG 452
                                    TCACTTGAAAGATGCAGTTCTTGAAGGAGG
cassava4.1_010187m extraction   423 TCACTTGAAAGATGCAGTTCTTGAAGGAGG 452
```

Figure 4.3 **Nucleotide Muscle alignment of RNAi target regions. A)** CCoAOMT gene family alignment of region targeted by the ihpRNA construct. Length of sequence: 418bp; Identical sites: 378; Identity: 90%. **B)** COMT gene family alignment of region targeted by the ihpRNA construct. Length of sequence: 330bp; Identical sites: 325; Identity: 98%.

The universal structure for the primer design is:

Forward primer (FP): 5′- ACCA (protective bases) - GGTCTC (BsaI) – AGGAG (adaptor for pRNAi-GG) - gene specific sequence -3′

Reverse primer (RP): 5′- ACCA (protective bases) - GGTCTC (BsaI) – ATCGT (adaptor for pRNAi-GG) - gene specific sequence -3′

GGAG (FP) and TCGT (RP) are adaptors specific to the pRNAi-GG vector (Figure 4.4).

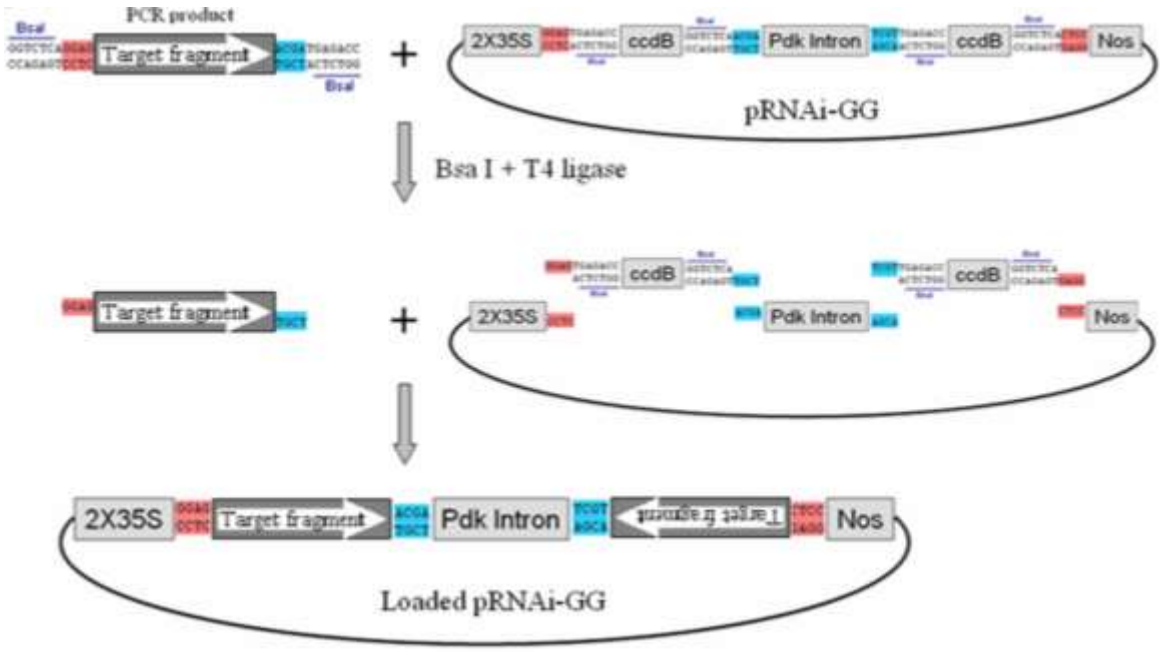


Figure 4.4 **ihpRNA construction with the pRNAi-GG vector using Golden Gate technology.** One-step construction of the ihpRNA where the gene of interest is PCR amplified with gene-specific primers with the BsaI site and complementary adaptors to the pRNAi-GG vector. The purified PCR product, the pRNAi-GG vector, BsaI restriction enzyme and T4 ligase are mixed in one tube for a one-step restriction ligation reaction (Yan *et al.*, 2012).

To achieve simultaneous silencing, primers must be designed so that the adaptor sequences will complement one another and ligate (Figure 4.5). The primers designed for the two constructs, pRNAi-CCoAOMT/F6'H and pRNAi-CCoAOMT/F6'H/COMT are in the Table 4.1.

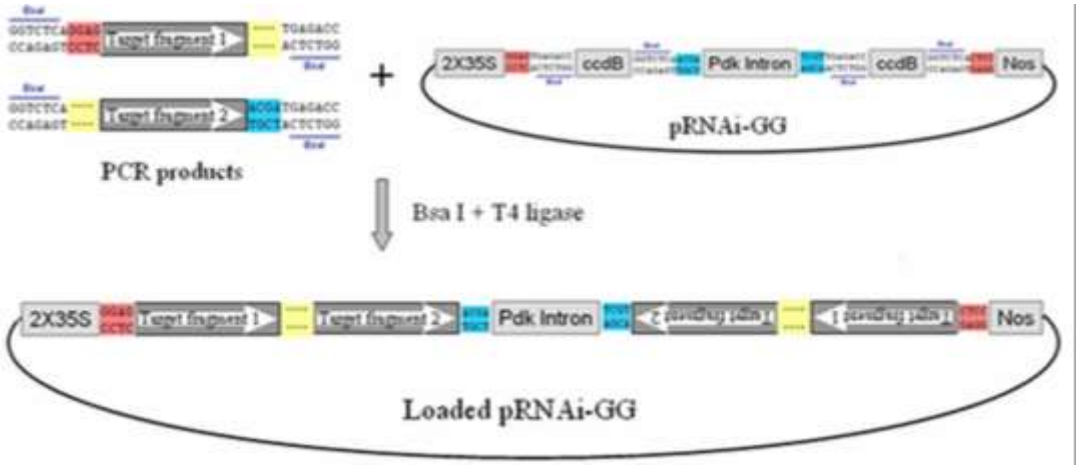


Figure 4.5 **ihpRNA construction with the pRNAi-GG vector through Golden Gate technology for simultaneous silencing of two genes.** Two PCR products of genes of interest were cloned into the pRNAi-GG vector simultaneously through a single restriction-ligation reaction following the same principles as Figure 4.2 (Yan *et al.*, 2012).

Table 4.1 **Primers used in the construction of pRNAi-CCoAOMT/F6'H and pRNAi-CCoAOMT/F6'H/COMT.**

Primer name	Primer sequence (5' – 3')
CCoAOMT_pRNAi-CCoAOMT/F6'H_Fwd	ACCAGGTCTCAGGAGCTCTTGCTCTTCCTGATGATGG
CCoAOMT_pRNAi-CCoAOMT/F6'H_Rev	ACCAGGTCTCAGCTGAACCGGAAGCATGCAAATCTC
F6'H_pRNAi-CCoAOMT/F6'H_Fwd	ACCAGGTCTCACAGCCCAACACTTGCAGAATCAGCC
F6'H_pRNAi-CCoAOMT/F6'H_Rev	ACCAGGTCTCATCGTATTAGCCTCGTCGTCGGAGA
CCoAOMT_pRNAi-CCoAOMT/F6'H/COMT_Fwd	ACCAGGTCTCAGGAGCTCTTGCTCTTCCTGATGATGG
CCoAOMT_pRNAi-CCoAOMT/F6'H/COMT_Rev	ACCAGGTCTCAGCTGAACCGGAAGCATGCAAATCTC
F6'H_pRNAi-CCoAOMT/F6'H/COMT_Fwd	ACCAGGTCTCACAGCCCAACACTTGCAGAATCAGCC
F6'H_pRNAi-CCoAOMT/F6'H/COMT_Rev	ACCAGGTCTCAACGGATTAGCCTCGTCGTCGGAGA
COMT_pRNAi-CCoAOMT/F6'H/COMT_Fwd	ACCAGGTCTCACCGTGCTCGACCTGCTTGAAATC
COMT_pRNAi-CCoAOMT/F6'H/COMT_Rev	ACCAGGTCTCATCGTCCTCCTTCAAGAACTGCATC

A two-step reaction was used due to the size of the amplicons and number of products to undergo restriction-ligation. For the first step, the amplified PCR products were checked on gel electrophoresis and visualised under a UV transilluminator. The PCR products were then gel purified and assembled with BsaI and T4 ligase. The assembled products were then amplified with CCoAOMT_pRNAi-CCoAOMT/F6'H forward primer (FP) and F6'H_pRNAi-CCoAOMT/F6'H reverse primer (RP) for the double knock-down construct (pRNAi-CCoAOMT/F6'H) and CCoAOMT_pRNAi-CCoAOMT/F6'H/COMT FP and COMT_pRNAi-CCoAOMT/F6'H/COMT RP for the triple knock-down construct (pRNAi-CCoAOMT/F6'H/COMT) and checked on gel electrophoresis.

The assembled PCR products were gel purified and then used in the second restriction-ligation reaction with the pRNAi-GG vector. The assembled pRNAi-GG vectors were chemically transformed into MAX Efficiency™ DH5α™ Competent Cells (>1x10^9) (Thermo Fisher Scientific). Colony PCR was performed to identify the correct clones with pRNAi-GG specific primers (Table 4.2; Figure 4.6). The primers (P24 and P25) also identify the orientation of the intron, but intron orientation should not affect the silencing efficiency (Yan *et al.*, 2012). The sequences were confirmed by PCR (Figure 4.7) and DNA sequencing with P22 and P23.

Table 4.2 **Primer sequences used to identify recombinants, intron orientation and sequencing for the pRNAi-GG vector (Yan *et al.*, 2012)**.

Primer	Sequence (5' – 3')	Description
P21	ACCATTTACGAACGATAGCC	Recombinants identification
P22	GTAAAACGACGGCCAGTG	Recombinants identification and sequencing
P23	CGAATCTCAAGCAATCAAGC	Recombinants sequencing
P24	CATTTTAGCTTCCTTAGCTCC	Recombinants and intron orientation identification
P25	CATTTGGATTGATTACAGTTGG	Recombinants and intron orientation identification

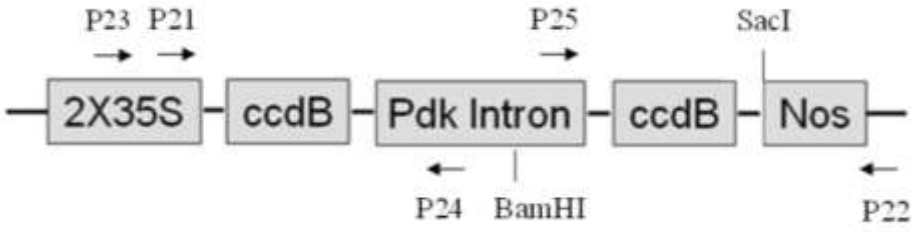


Figure 4.6 **The pRNAi-GG primers position and restriction sites.** P21, P22, P24 and P25 are used in recombinants identification. P22 and P23 are used in recombinants sequencing. P24 and P25 can also be used to identify intron orientation. BamHI and SacI can also be used in intron orientation identification.

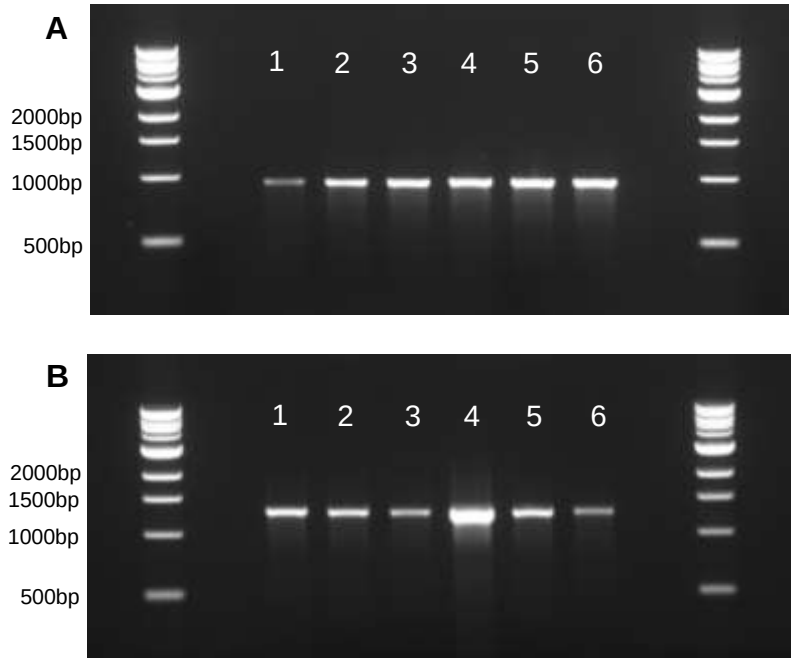


Figure 4.7 **Gel electrophoresis results of recombinant pRNAi-GG with primers P21 and P25. A)** Antisense transformants of pRNAi-CCoAOMT/F6'H identified in colony PCR (Colonies 1-6). **B)** Antisense transformants of pRNAi-CCoAOMT/F6'H/COMT identified in colony PCR (Colonies 1-6).

4.3.2 CASSAVA TRANSFORMATION AND GROWTH

Cassava friable embryogenic calli (FEC) were provided by Dr Herve Vanderschuren from ETH-Zürich and were multiplied and transformed with Agrobacterium at the University of Bath (Bull *et al.*, 2009). Cassava plantlets were successfully regenerated from the post-transformation FEC after eight months of intensive care. Not all regenerated cassava plantlets will be transgenic which is why the plantlets must be screened through PCR to confirm that the insert is present with primers P21 and P25 (Figure 4.8). Southern blot was then performed on the transgenic cassava plantlets to determine independent lines through identifying insert location and insert copy number (Figure 4.9, Table 4.3). Three pRNAi-CCoAOMT/F6'H (2x) lines and five pRNAi-CCoAOMT/F6'H/COMT (3x) lines were then propagated in CBM agar. They were then transferred to soil (M2 medium + perlite, 3:1) 8-14 days after propagation depending on the rate of root growth. This was done in the growth room under controlled conditions (26°C, 16h light). One month after the plantlets have been transferred into soil, they were then transferred to the glasshouse (25-28°C, 30-50% relative humidity and 16h light) and transplanted into 1 L sized pots with the same soil composition. 35 plants were grown for each transgenic line with a total of 10 lines including wild-type (TMS60444), 35S-L (1x) line (F6'H RNAi transgenic line

obtained from Liu *et al.* (2017)), pRNAi-CCoAOMT/F6'H lines (3) and pRNAi-CCoAOMT/F6'H/COMT (5).

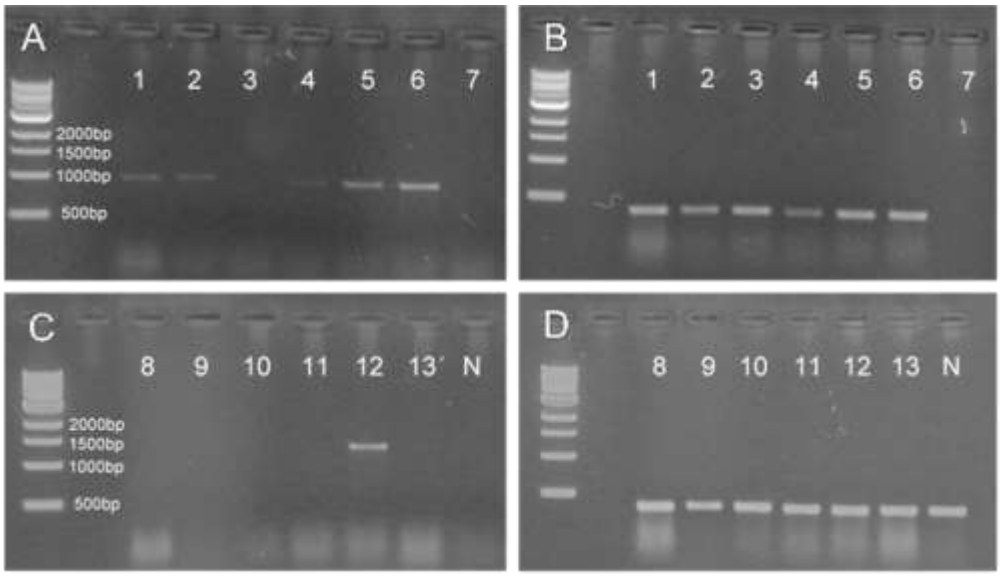
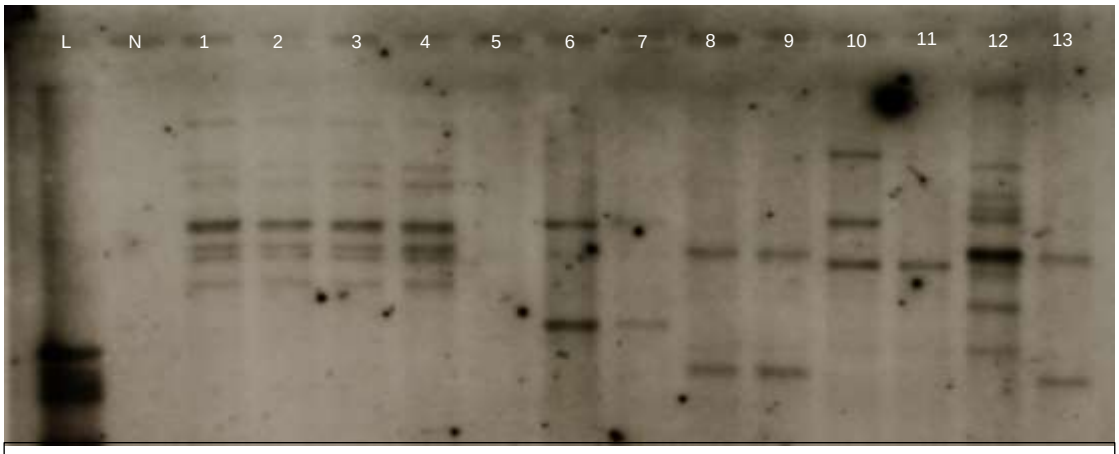


Figure 4.8 **Confirmation of the RNAi insert by PCR using DNA from plantlets regenerated from transformed FEC with primers P21 and P25. A)** pRNAi-CCoAOMT/F6'H transgenic lines (1, 2, 4, 5, 6). **B)** Positive control (PP2A) for lines 1 – 6; PCR failed for line 7. **C)** pRNAi-CCoAOMT/F6'H/COMT transgenic line (12); N is negative control with wt (TMS60444). **D)** Positive control (PP2A) for lines 8 - 13; N is negative control with wt (TMS60444).



L: DNA Molecular Weight Marker III-DIG labelled Roche.
N: pRNAi-GG (Negative control)
pRNAi-CCoAOMT/F6'H: 1-7
pRNAi-CCoAOMT/F6'H/COMT: 8-13

Figure 4.9 **Southern blot to identify independent lines and transgene copy number.** Southern blots results of *Hind*III-digested cassava genomic DNA of pRNAi-CCoAOMT/F6'H and pRNAi-CCoAOMT/F6'H/COMT lines using insert specific probe (Table 2.1). Lanes 1-4 have multiple inserts and are the same line (clones). Lane 5 failed. Lane 6 and 7 are have double inserts. Lanes 8,9 and 13 are the same line with double inserts. Lane 10 is a triple insert line. Lane 11 is a single insert line. Lane 12 is a multiple insert line.

Table 4.3 **Cassava transgenic lines and the number of transgene insertion.** 2x represents the double knock-down lines (pRNAi-CCoAOMT/F6'H) and 3x represents the triple knock-down lines (pRNAi-CCoAOMT/F6'H/COMT).

Cassava Transgenic Line	Transgene Copy Number
2xA	7
2xB	2
2xE	3
3xA	2
3xB	3
3xD	9
3xE	1
3xG	3

4.4 DISCUSSION

Cassava is a difficult plant to transform due to its long lifecycle, high levels of genetic heterozygosity, variable flowering patterns and low seed set and germination (Jennings and Iglesias, 2002). Due to this, cassava transformation is performed with friable embryogenic calli (FEC) via *Agrobacterium*. This is widely used and shown to be the most efficient method to generate transgenic cassava. However, this method remains a labour-intensive and time-consuming procedure which took over a year to successfully regenerate plantlets from FEC to a sufficiently mature enough stage to be transferred to the glasshouse. Besides that, low regeneration of plantlets from somatic embryos (Baba *et al.*, 2008) and highly variable numbers of transgenic events (Koehorst-van Putten *et al.*, 2012) are the two main obstacles for cassava FEC transformation. Simultaneous gene-silencing through RNAi was the approach taken due to the efficiency and length taken to regenerate transgenic cassava plantlets after transformation.

In order to compare the double and triple knock-down cassava lines with the F6'H knock-down cassava plants generated by Liu *et al.* (2017), the same primers were used in this study for the F6'H gene family. Since CCoAOMT and COMT have only two members in their gene families and are highly identical to one another, the RNAi primers designed will target both members thus preventing leakage of gene expressions. Simultaneously knocking down genes through a single RNAi construct through pRNAi-GG (Yan *et al.*, 2012) has not been done in cassava before. This approach was chosen because separate transformation events to generate the double and triple knock-down mutants would be inefficient and unreliable due to the low transformation success rate and

regeneration of cassava plantlets from FEC. Due to the size of the inserts of approximately 300bp to 500bp in length per insert (two inserts for pRNAi-CCoAOMT/F6'H and three for pRNAi-CCoAOMT/F6'H/COMT), the RNAi construct had to be made in two steps instead of one and had to be transformed into competent cells with at least greater than $1x10^9$ efficiency due to the large plasmid size (approximately 14,000bp).

Since a different vector to what was used in the cassava transformation protocol developed by Bull *et al.* (2009), the protocol had to be modified. Hygromycin was used in the original protocol as hygromycin can achieve complete growth inhibition at low concentrations thus resulting in selection of a higher number of transformed callus lines. pRNAi-GG has *nptII* rather than *hptII,* which is why geneticin selection was used as cassava has high levels of inherent resistance to kanamycin that is usually used in *nptII* selection (Schöpke *et al.*, 1996). The protocol uses the GUS assay to track transformation success but unfortunately, pRNAi-GG vector does not have the GUS gene, so this could not be done. Another issue faced was that the rooting test did not work well with geneticin selection as growth inhibition was not strong enough and resulted in non-transformed plants successfully developing roots. In response to this, all regenerated plantlets were analysed through PCR to determine and confirm integration of the genes of interest. Without the preliminary screening through the rooting test, more time was taken to extract DNA and analyse each plant through PCR. Although it took longer to generate transgenic cassavas due to the modifications to the protocol, three lines were successfully generated for pRNAi-CCoAOMT/F6'H (2x) and five lines were generated for pRNAi-CCoAOMT/F6'H/COMT (3x).

5 DO THE TRANSGENIC CASSAVA PLANTS DIFFER IN TERMS OF GROWTH AND THE PPD RESPONSE?

5.1 INTRODUCTION

In the previous chapters, three enzymes, CCoAOMT, F6'H and COMT, were identified as targets for silencing through RNAi. RNAi constructs were made with the intention to knock-down all members of the gene families encoding for said enzymes. CCoAOMT, F6'H and COMT are part of the phenylpropanoid metabolism, which is responsible for important secondary metabolites such as monolignols, flavonoids, coumarins, and many more. The hydroxycoumarin scopoletin, which is a product of the phenylpropanoid metabolism, has been shown to play a vital role in cassava PPD development, showing significant accumulation during PPD (Wheatley and Schwabe, 1985). Although this dramatic accumulation of scopoletin is not solely responsible for PPD, its oxidation by H_2O_2 and peroxidase to produce the blue-black pigment definitely plays an important role in the discolouration of cassava storage root (Beeching *et al.*, 2000). Inhibiting the accumulation of scopoletin could delay PPD, which is why the enzymes (CCoAOMT, F6'H and COMT) responsible for the biosynthesis for scopoletin are targeted for silencing.

As described in the previous chapter, pRNAi-CCoAOMT/F6'H (2x) and pRNAi-CCoAOMT/F6'H/COMT (3x) transgenic lines were generated to understand the significance of the alternative pathways in scopoletin biosynthesis. These lines will be compared to the single knock-down line for F6'H generated by Liu *et al.* (2017). The model plant *A. thaliana* is a good reference to use in the study of scopoletin biosynthesis in cassava (Kai *et al.*, 2006); although the two alternative pathways involved in scopoletin biosynthesis in cassava (Figure 5.1) have not been thoroughly explored in *A. thaliana*, it is likely that the major routes will be shared between them (Bayoumi *et al.*, 2008; Kai *et al.*, 2006) as the majority of pathways in the phenylpropanoid metabolism are conserved in higher plants (Tohge *et al.*, 2013).

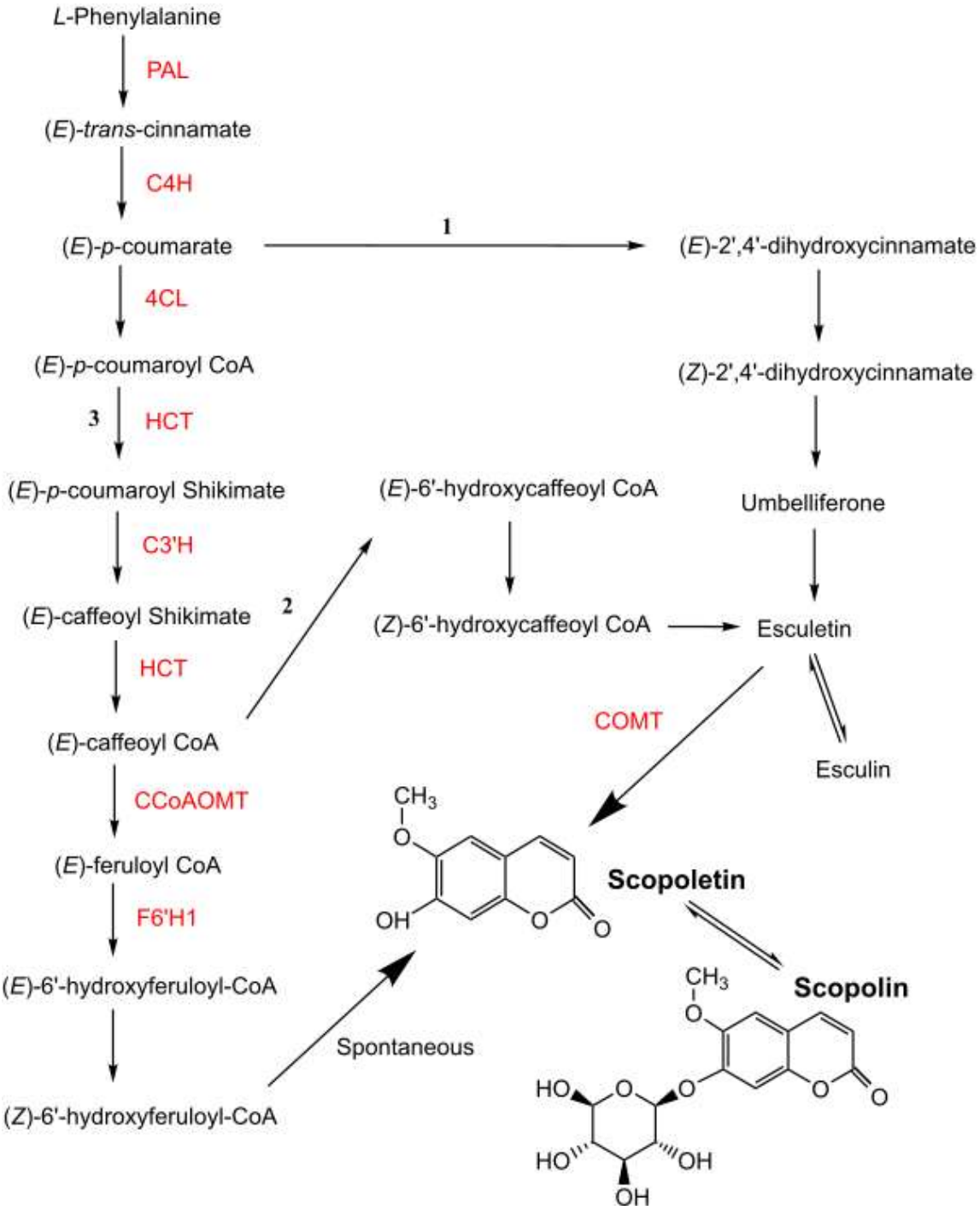


Figure 5.1 **The biosynthesis of scopoletin through three alternative pathways in the phenylpropanoid metabolism.**

When F6'H1 is knocked-out with a T-DNA insertion in *A. thaliana*, the mutant does not accumulate scopoletin (Kai *et al.*, 2008). Unfortunately, this was not the case in cassava as the RNAi knock-down cassava lines did not completely inhibit scopoletin biosynthesis therefore PPD still occurred, albeit delayed (Liu *et al.*, 2017). This is hypothesised to either be due to incomplete knock-down of F6'H as there are seven members in the gene family, or scopoletin is being synthesised through the alternative pathway where esculetin is methoxylated by COMT to produce scopoletin. To answer these questions, CCoAOMT and COMT were selected as targets for RNAi knock-down. T-DNA mutants for CCoAOMT (CCoAOMT1), COMT (COMT1) and double mutants for CCoAOMT and COMT (CCoAOMT1 COMT1) were generated in *A. thaliana*. The CCoAOMT1 mutants have collapsed xylem vessels and a lower stem lignin concentration, in particular, the reduction of guaiacyl (G) monomers, compared to wild-type plants (Do *et al.*, 2007). It also has approximately 30% and 15% lower levels of scopoletin and scopolin in the roots, respectively (Kai *et al.*, 2008). The COMT1 mutant lacked syringyl (S) unit lignin and had 5-OH-G units derived from 5'-hydroxyconiferyl alcohol but, unlike CCoAOMT1, it did not show any reduction in lignin content (Goujon *et al.*, 2003). In *A. thaliana*, CCoAOMT1 is involved in the 3-*O*-methylation of lignin subunits but is mainly responsible for the biosynthesis of G monomers with the cooperation of COMT1 to methylate the precursors of coniferyl alcohol. COMT1 is also involved in the biosynthesis of the S monomer by acting as a 5-*O*-methyltransfrase.

Despite having changes to the lignin composition, the *A. thaliana* CCoAOMT1 and COMT1 mutants do not have significant phenotypic differences compared to the wild-type plants under optimal growth conditions (Do *et al.*, 2007). Unfortunately, this was not the case for the CCoAOMT1 COMT1 double mutant as the F2 progeny seedlings suffered developmental arrest 4 days post germination and turned yellow (Do *et al.*, 2007). The double mutants had significantly enriched H non-methoxylated units in the lignin composition and the biosynthesis of G and S subunits were severely affected, suggesting that CCoAOMT and COMT act cooperatively to methylate the C3 position. The double-mutant plantlets also lack isorhamnetin, which is the methylated form of quercetin, but this is unlikely to be the cause of the developmental defects as the *tt4* mutant which has a mutated chalcone synthase gene has normal development despite the absence of flavonoids (Shirley *et al.*, 1995). It is still unknown as to why the double mutant suffers from developmental arrest, but CCoAOMT and COMT are likely to have an important role in plant development. Fortunately, unlike *A. thaliana*, the cassava 3x mutant lines did not appear to suffer from developmental arrest and were successfully transplanted into soil as mentioned in the previous chapter. The cassava triple RNAi

mutants are not complete knockouts like those in *A. thaliana,* which means there will still be residual gene activity that may explain why the mutants survived. The effects of the gene knock-down on cassava storage root development are still unknown but there are multiple lines of 3x transgenic cassava with different insert locations, which probably result in different expression levels of the RNAi construct. This should provide a range of transgenic lines with different phenotypes, so the most appropriate or useful lines can be selected for analysis. Besides developing a further understanding on the biosynthesis of scopoletin in cassava PPD regarding the importance of the alternative pathways, this study may also provide an insight to the roles of CCoAOMT and COMT in the development of cassava.

5.2 RESULTS

5.2.1 DO THE TRANSGENIC CASSAVA PLANTS HAVE MORPHOLOGICAL DIFFERENCES?

5.2.1.1 HEIGHT

35 plants were grown for each transgenic line with a total of 10 lines including WT line (TMS60444), 35S-L line (F6'H RNAi transgenic line obtained from Liu *et al.* (2017) that was renamed 1xA for convenience), pRNAi-CCoAOMT/F6'H lines (2xA, 2xB and 2xE) and pRNAi-CCoAOMT/F6'H/COMT (3xA, 3xB, 3xD, 3xE, 3xG). Plant height was measured at 5, 7 and 9 months before harvesting the plants to analyse PPD in the storage root. One-way ANOVA was performed on the transgenic plants (1xA, 2xA, 2xB, 2xE, 3xA, 3xB, 3xD, 3xE and 3xG) in relation to the wild-type plants at 5, 7 and 9 months (Figure 5.2). The triple knock-down lines (3xA, 3xB, 3xD, 3xE, 3xG) showed significant difference at all ages and had a reduced growth rate compared to the wild-type control. There is especially high variation within the 3xA, 3xE and 3xG lines where a proportion of the plants have lost apical dominance and are severely stunted in growth (Figure 5.3). Some plants managed to recover from this problem and grew normally, whereas others remained dwarfed until 9 months, which was the time of harvest. This indicates that down-regulation of the three gene families, CCoAOMT, F6'H and COMT, can have adverse effects on plant development.

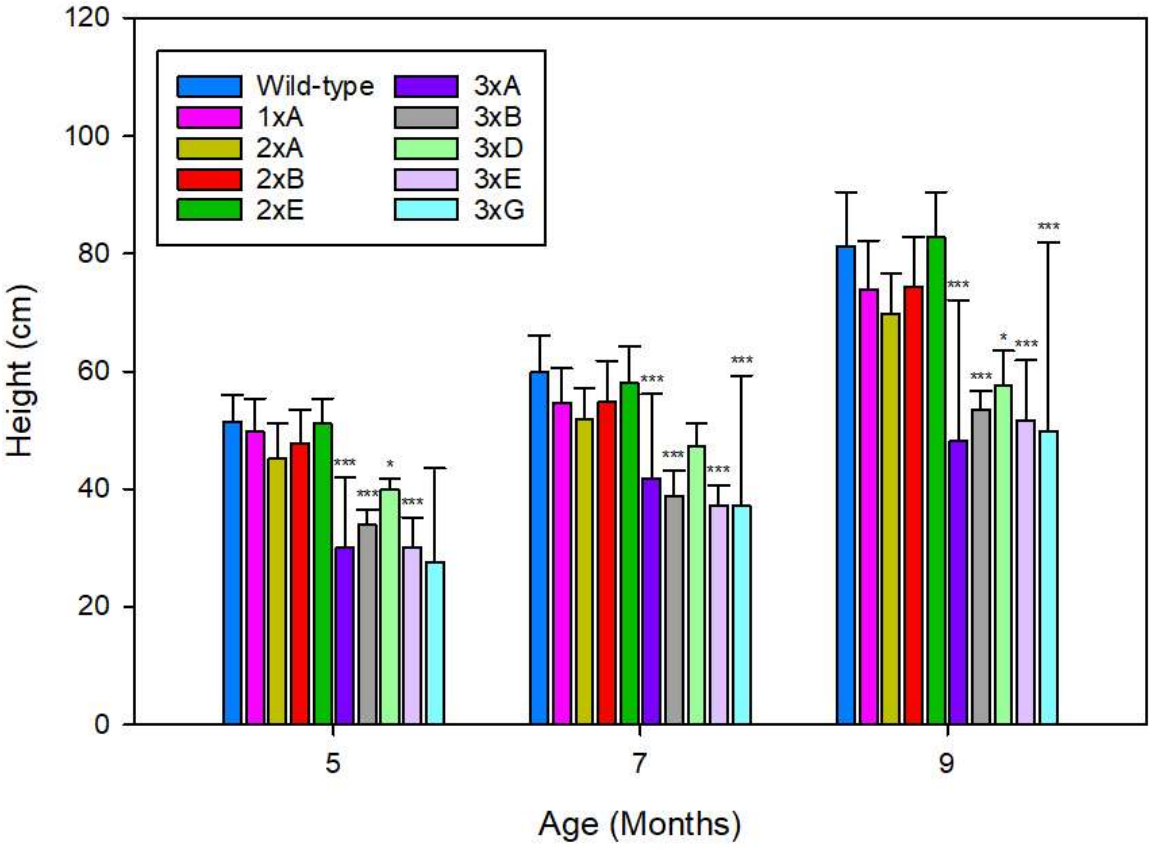


Figure 5.2 **Height of cassava plants at 5, 7 and 9 months.** Data were collected from 10 plants for each line and the height shown is the average of 10 plants ± standard deviation (SD). Statistical analysis (One-way ANOVA) in relation to the wild-type: $p<0.05$ = *, $p<0.01$ = **, $p<0.001$ = ***.

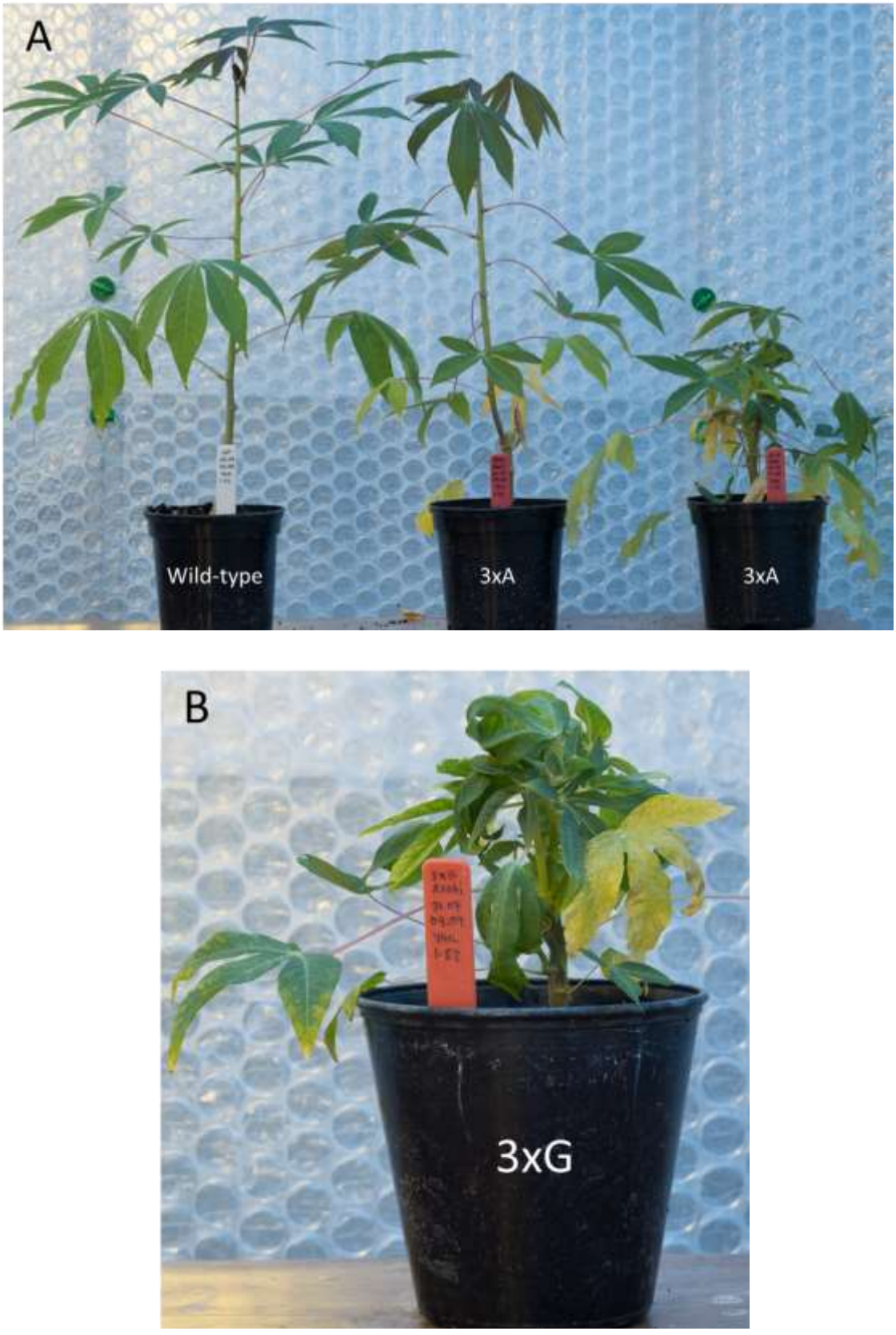


Figure 5.3 **Comparison of wild-type and 3x (pRNAi-CCoAOMT/F6'H/COMT) plants at 9 months old. A)** 9 months old wild-type, 3xA plant with relatively normal growth and 3xA plant with dwarfing and loss of apical dominance. **B)** Close-up of a 3xG plant with dwarfing and loss of apical dominance.

5.2.1.2 LEAF SHAPE

There were morphlogical differences observed in the leaf shape of the transgenic plants in comparison to the wild-type control plants. The leaves of 2x and 3x transgenic plants have a claw-like appearance (Figure 5.4) when they first emerge but may return to normal when older. The 1xA plants have a different leaf morphology where they have broader lobes and a darker green colour (Figure 5.5) which is consistent with the findings from Liu *et al.* (2017).

Figure 5.4 **Wild-type cassava and 3xD (pRNAi-CCoAOMT/F6'H/COMT D) transgenic cassava at 9 months old to show difference in leaf morphology.** The 3xD plant has claw-like leaves and appear smaller compared to the wild-type.

Figure 5.5 **Comparison of wild-type cassava and 1xA (pRNAi-F6'H A) transgenic cassava leaf morphology. A)** Side-by-side comparison between the wild-type plant and 1xA plant at 9 months old. **B)** Close-up pictures taken at 9 months comparing the leaf morphology between wild-type and 1xA transgenic cassava plants (Liu, 2016).

5.2.1.3 ROOT WEIGHT

The cassava plants were harvested at 9 months old and the roots were weighed before proceeding for PPD assessment. Fibrous roots were removed, and the storage roots were washed to remove soil and then weighed. All the transgenic lines showed significant reduction in weight where $p<0.001$ in comparison to the wild-type roots when tested with ANOVA (Figure 5.6). 3xA, 3xE and 3xG had the lowest weight of all the transgenic lines and there were plants from these lines where storage roots failed to develop. This could be related to the developmental issues these transgenic lines faced. Besides 1xA and 3xD, the other transgenic plants all had storage roots that were less than half the weight of the wild-type (116.1 ± 31).

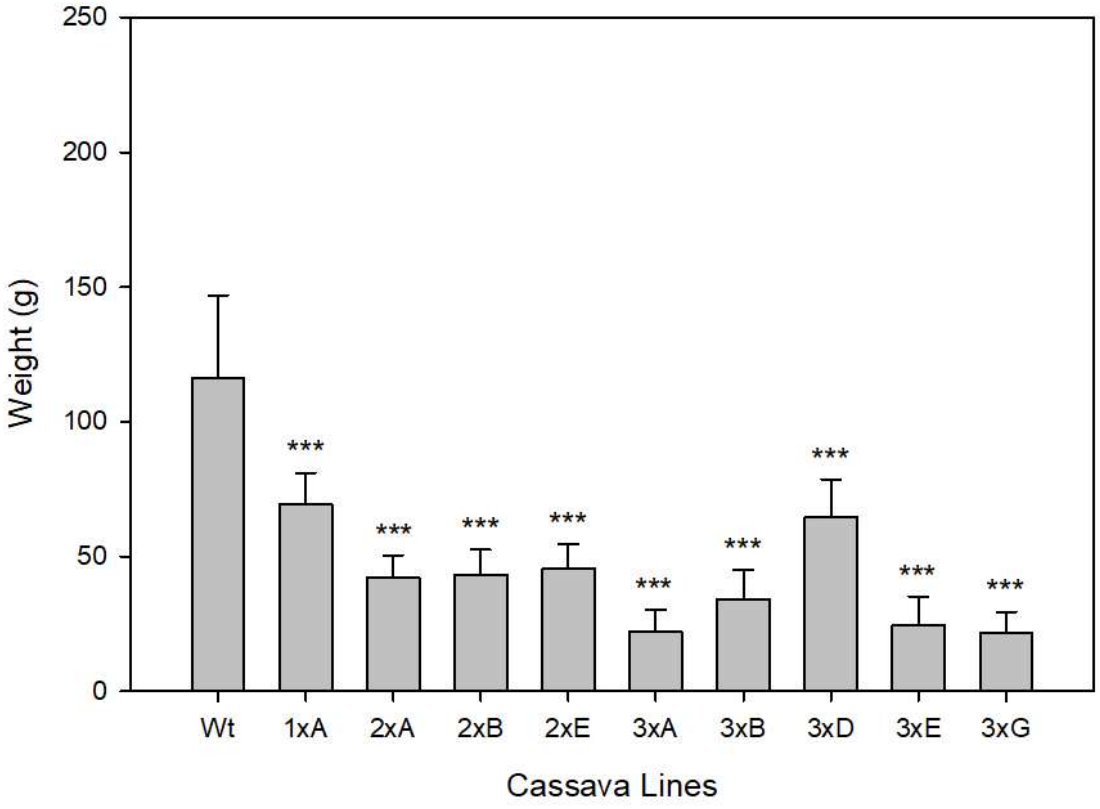
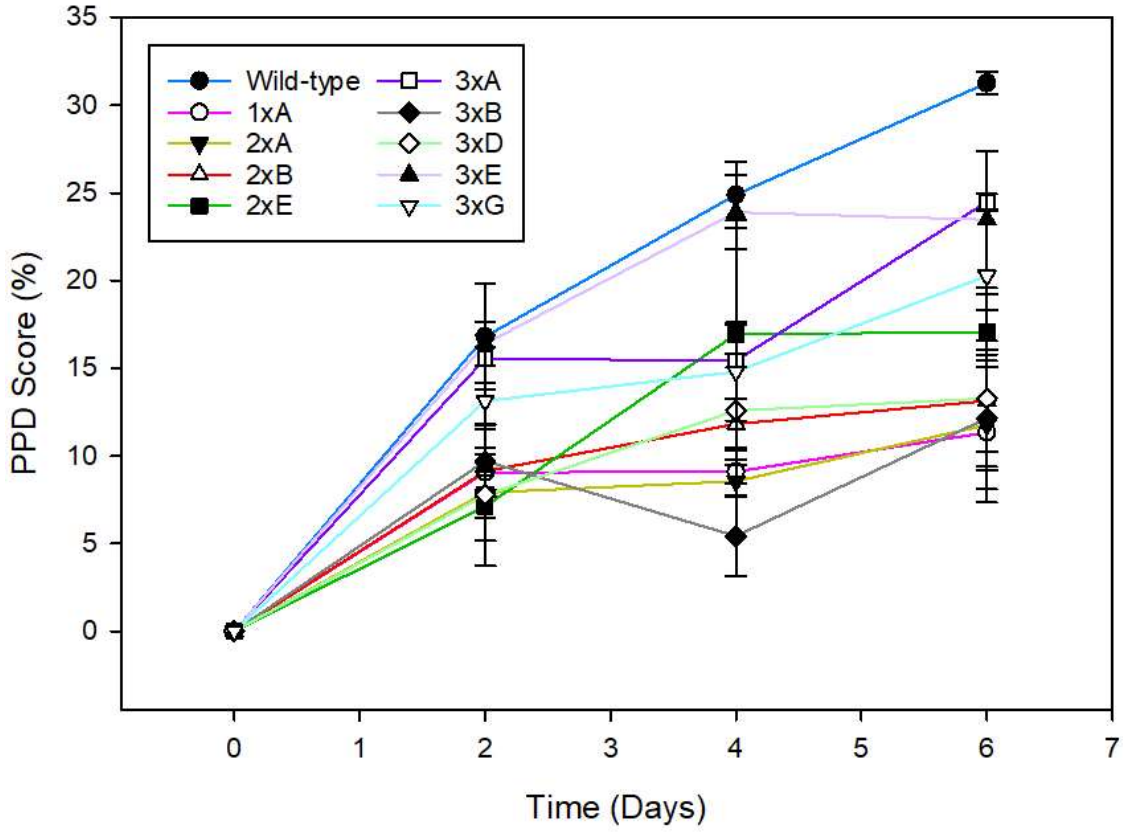


Figure 5.6 **Weight of cassava lines storage root at 9 months.** Storage roots were harvested from 9 plants per cassava line and fibrous roots were removed and then weighed. The weight shown is the average of 9 plants ± SD (standard deviation of 9 samples). Statistical analysis (One-way ANOVA) was performed on the transgenic lines in relation to the WT control: $p<0.05$ = *, $p<0.01$ = **, $p<0.001$ = ***.

5.2.2 IS PPD AFFECTED BY THE DOWNREGULATION OF F6'H, CCOAOMT AND COMT?

Storage roots of the 10 lines (WT, 1xA, 2xA, 2xB, 2xE, 3xA, 3xB, 3xD, 3xE, 3xG) of similar size were harvested, washed, weighed and then cut up into discs for PPD development. PPD discolouration was then evaluated as described in Chapter 2. The PPD score (%) was evaluated and normalised to the score at day 0 (Figure 5.7). Besides the triple knock-down lines 3xA and 3xE, the other RNAi transgenic lines show significant reduction in PPD symptoms over the 6 days (Figure 5.8). The lines that showed the largest reduction in terms of discolouration were 1xA, 2xA, 2xB, 3xB and 3xD, as their PPD scores were approximately half of the wild-type's and the ANOVA test showed P values of less than 0.001 (***). 3xA and 3xE were the lines that experienced developmental issues and some plants remained dwarfed and developed very small storage roots, if any at all. The rate of PPD varied with the size of storage root which was why roots of the similar size were used in the experiment. Unfortunately, due to the lack of storage roots for lines 3xA and 3xE, the smaller roots had to be used for the PPD assessment. This may be the reason why the PPD score is higher than the other transgenic lines. The results between 3xA and 3xE with the other transgenic lines are hard to compare because of the inconsistency in root sizes. Although 3xG showed a significant difference to the wild-type control with a P value of less than 0.05, there was high variation within the 3xG population because like 3xA and 3xE, 3xG also had developmental issues where some plants lost apical dominance and remained stunted in growth. Unlike 3xA and 3xE, 3xG had some plants that produced storage roots that are of reasonable size which may be why difference in PPD was witnessed. Lines 3xA and 3xB had a decrease in PPD score from day 2 to day 4 although this should not be possible. This may be due to the high variation observed within the population and the lack of similar sized roots. Ideally, the stunted and non-stunted populations within the 3xA and 3xD lines would be assessed separately, but unfortunately, there were not enough samples to make these comparisons.



Lines	Day 2	Day 4	Day 6
1xA	*	***	***
2xA	**	***	***
2xB	*	**	***
2xE	**	-	***
3xA	-	*	-
3xB	*	***	***
3xD	**	**	***
3xE	-	-	-
3xG	-	*	*

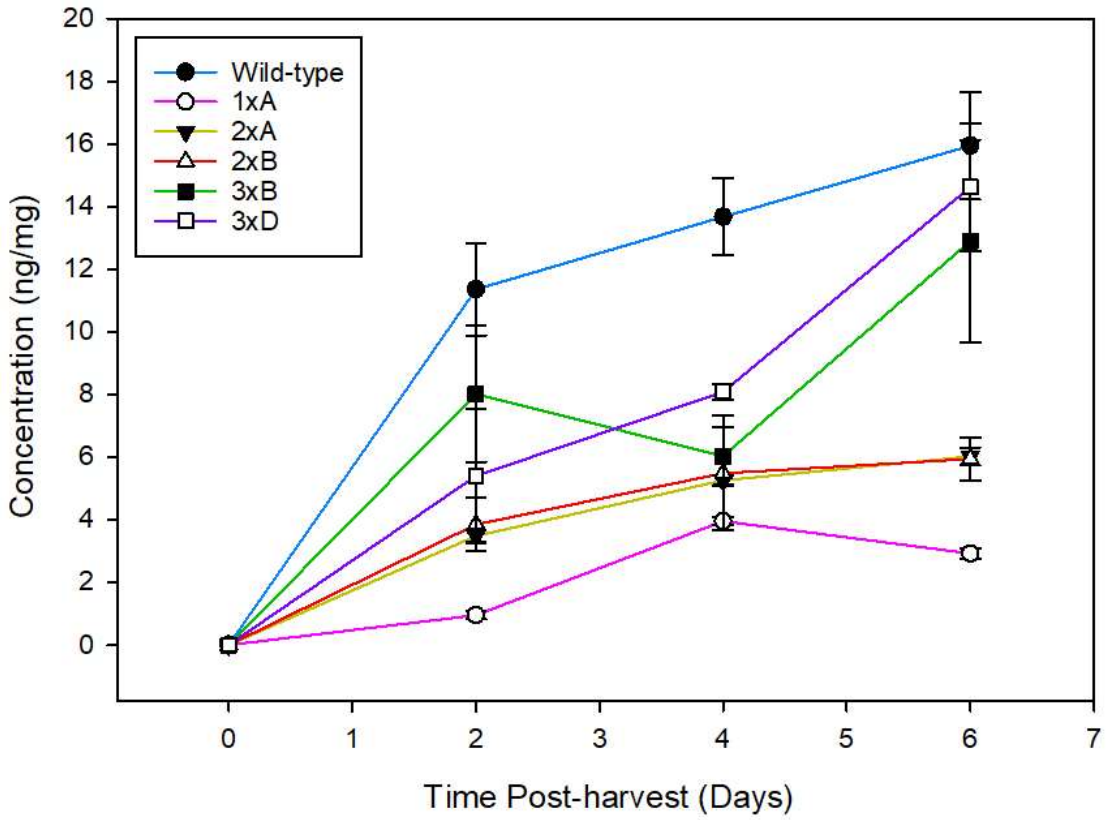
Figure 5.7 **The discolouration of the wild-type cassava compared with nine RNAi transgenic lines during the PPD time course.** The table presents the statistical analysis (one-way ANOVA) in relation to the wild-type where * is $p<0.05$, ** is $p<0.01$ and *** is $p<0.001$ in relation to the wild-type.

Figure 5.8 **PPD development in the wild-type, 2xA and 3xB cassava storage roots.** The pictures were taken at day 0 (fresh), day 2, day 4 and day 6 post-harvest.

5.2.3 WAS THE ACCUMULATION OF SCOPOLETIN REDUCED IN TRANSGENIC CASSAVA LINES?

Previous research has identified scopoletin as the major contributor to the discolouration observed in the storage roots during PPD. As a result of the PPD assessment conducted in 5.2.2, cassava lines (1xA, 2xA, 2xB, 3xB, 3xD) that had the lowest PPD score (%) and the wild-type control were selected for biochemical analysis through HPLC-MS to determine whether changes in scopoletin and scopolin accumulation were seen. Figure 5.9 shows the difference in scopoletin accumulation in the six cassava lines mentioned. The one-way ANOVA analysis showed that except for 3xD at day 6, all the transgenic lines have significantly reduced scopoletin accumulation compared to the wild-type. The line with the least scopoletin accumulation throughout the PPD process is 1xA, followed by the double construct (pRNAi-CCoAOMT/F6'H) lines, 2xA and 2xB. The P-value for these three lines throughout day 0, 2 and 6 are less than 0.000 which meant the results are greatly significant. The triple knock-down (pRNAi-CCoAOMT/F6'H/COMT) lines 3xB and 3xD had a less drastic reduction in scopoletin accumulation. In the wild-type, scopoletin concentration increased greatly from day 0 to day 2 and then resumed a gradual increase until day 6. For 1xA, 2xA and 2xB lines, there was no spike in scopoletin accumulation, only a gradual increase which peaked at day 4 for 1xA and day 6 for 2xA and 2xB. On the other hand, 3xB had a significant increase from day 0 to day 2 which

then decreased in day 4 and the spiked up again on day 6. 3xD had significant scopoletin reduction until day 4 and then there was an increase in day 6 that was not significantly less than the wild-type. Although the glucoside scopolin was included in this analysis, there was no change in scopolin levels throughout the PPD timescale for all lines including the wild-type where it remained at 0 ng/mg.



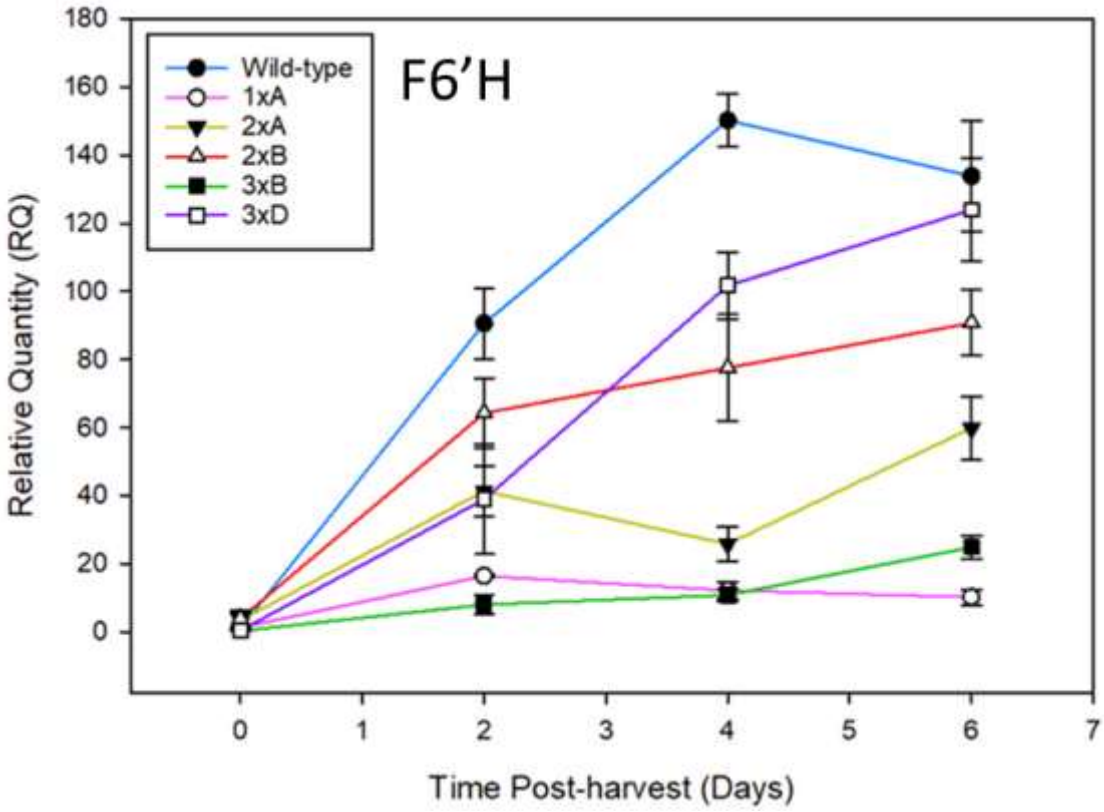
Lines	Day 2	Day 4	Day 6
1xA	***	***	***
2xA	***	***	***
2xB	***	***	***
3xB	***	***	**
3xD	***	***	-

Figure 5.9 **Scopoletin concentration (ng/mg) of the 6 lines over the PPD time course.** The table presents the statistical analysis (one-way ANOVA) in relation to the wild-type where * is $p<0.05$, ** is $p<0.01$ and *** is $p<0.001$ in relation to the wild-type.

5.2.4 DID THE RNAi CONSTRUCTS SUCCESSFULLY KNOCK-DOWN GENES OF INTEREST?

To investigate whether the two RNAi construct (pRNAi-CCoAOMT/F6'H and pRNAi-CCoAOMT/F6'H/COMT) successfully knocked-down gene expression for the respective genes of interest, quantitative gene expression of CCoAOMT, F6'H and COMT were obtained through qRT-PCR. The relative expression was calculated using the Livak method, also known as $\Delta\Delta C_T$, and was presented as relative quantity ($2^{-\Delta\Delta Ct}$). The relative quantity (RQ) method normalises the C_T value of the target (CCoAOMT, F6'H and COMT) against the housekeeping gene, ubiquitin 10 (Moreno *et al.*, 2011).

The F6'H gene family in cassava comprises seven members and the RNAi constructs aim to target all of them and downregulate their expression. Due to the high similarity between the F6'H gene members, it was impossible to design primers to differentiate between the individual genes and a pair of primers capable of targeting all of the genes was used (Liu *et al.*, 2017). Figure 5.10 shows the relative quantity (RQ) of the F6'H family in the six cassava lines (wild-type, 1xA, 2xA, 2xB, 3xB and 3xD) and the results from the one-way ANOVA analysis. At day 0, the RQ for F6'H was around 0 for the wild-type and transgenic lines. The RQ of wild-type cassava then increased to 90.6 ± 10.4 on day 2 and peaked on day 4 at 150.2 ± 7.9. There was no difference at day 0 but all the transgenic lines showed significant reduction in RQ from day 2 onwards except 3xD at day 6 where the RQ was similar to the wild-type. 1xA and 3xB lines had the greatest reduction in expression compared to the wild-type with the RQ at day 4 being 12.2 ± 2.4 and 10.8 ± 1.8 respectively. 2xA and 2xB also showed significant difference with a P-value of at least less than 0.05. 3xD had good reduction in expression on day 2 but caught up to the wild-type at day 6 ($p>0.05$).



Lines	Day 0	Day 2	Day 4	Day 6
1xA	-	***	***	***
2xA	-	***	***	***
2xB	-	*	***	**
3xB	-	***	***	***
3xD	-	***	***	-

Figure 5.10 **Relative quantification ($2^{-\Delta\Delta Ct}$) of F6'H transcripts in transgenic and wild-type storage roots over the PPD timescale using qRT-PCR.** RQ is calculated based on C_T mean values obtained from three biological replicates, normalised to the reference gene (Ubq10) and the control (WT day 0). The table presents the statistical analysis (one-way ANOVA) where * is $p<0.05$, ** is $p<0.01$ and *** is $p<0.001$ in relation to the wild-type.

CCoAOMT has two genes which are highly identical hence, like F6'H, it is not possible to design primers to target specific genes. A primer pair designed to target both genes was used (Table 2.1). There was no significant difference between all the transgenic lines and the wild-type (Figure 5.11). There were no significant changes in the gene expression level of CCoAOMT in any lines, including the wild-type throughout the PPD period. In comparison to F6'H which saw a huge spike in RQ during PPD where it peaked at day 4 at 150.2 ± 7.9 for the wild-type roots, the highest RQ for CCoAOMT was line 2xB at day 6 with an RQ of 8.2 ± 4.5, which was significantly lower than the RQ observed for all F6'H samples. 2x and 3x transgenic lines will have these genes knocked-down but

due to the lack of expression throughout the PPD timescale, it is hard to determine to what extent the genes had been downregulated.

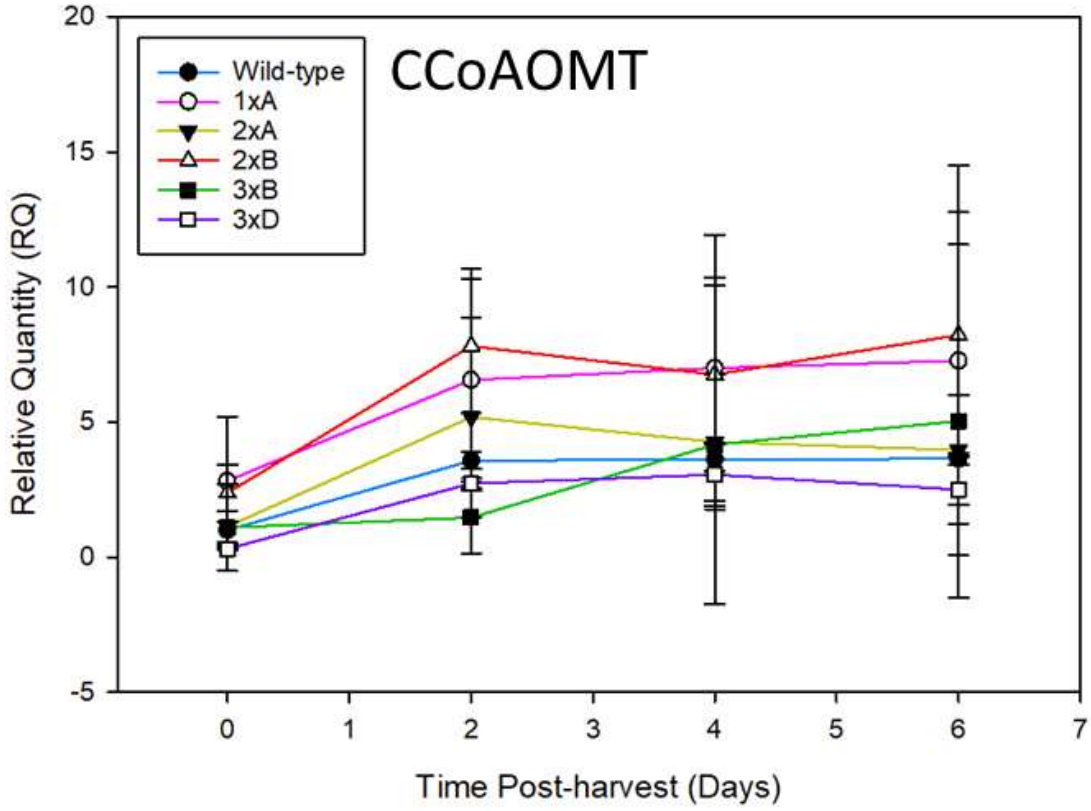


Figure 5.11 **Relative quantification ($2^{-\Delta\Delta Ct}$) of CCoAOMT transcripts in transgenic and wild-type storage roots over the PPD timescale using qRT-PCR.** RQ is calculated based on C_T mean values obtained from three biological replicates, normalised to the reference gene (Ubq10) and the control (WT day 0). No significant difference was observed in any lines in relation to the wild-type.

COMT is also encoded by two highly identical genes which is why, like F6'H and CCoAOMT, a primer pair has been designed to target both members (Table 2.1). Like CCoAOMT, there was no significant difference in the RQ values between the transgenic roots and wild-type roots over the PPD time course (Figure 5.12). A significant change in RQ values for COMT was not seen throughout the PPD process and the maximum RQ value was below 10, similar to the results of CCoAOMT.

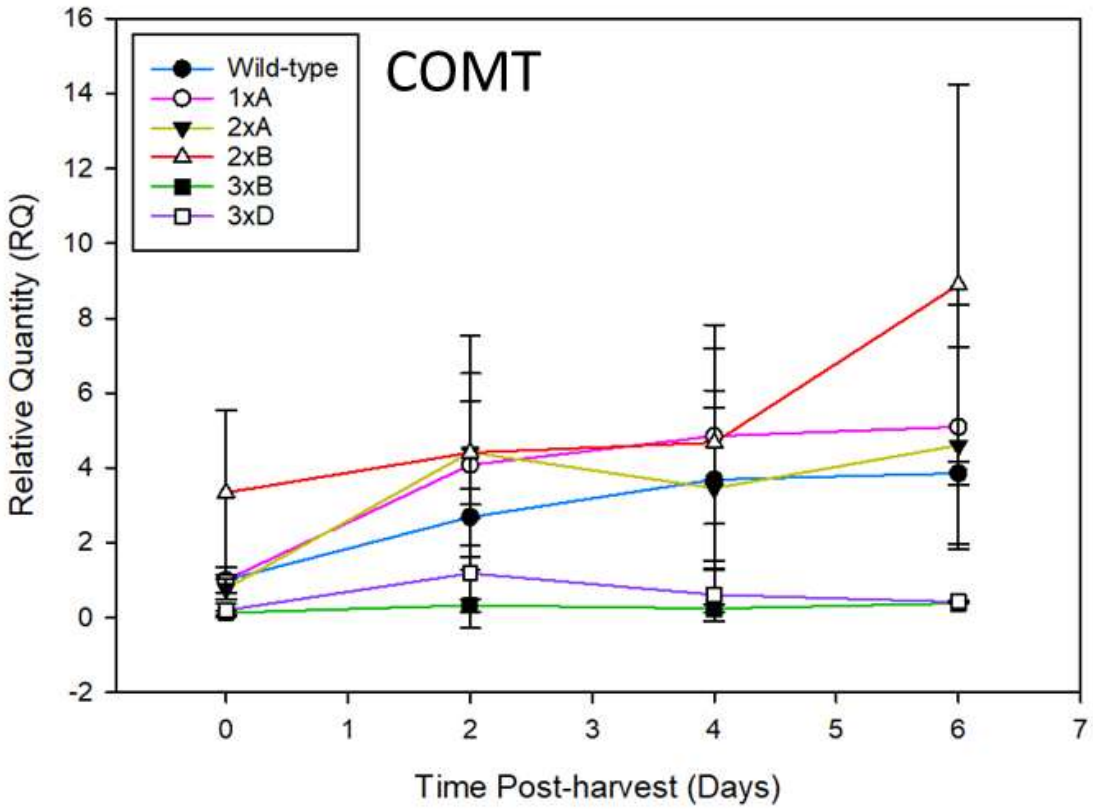


Figure 5.12 **Relative quantification ($2^{-\Delta\Delta Ct}$) of COMT transcripts in transgenic and wild-type storage roots over the PPD timescale using qRT-PCR.** RQ is calculated based on C_T mean values obtained from three biological replicates, normalised to the reference gene (Ubq10) and the control (WT day 0). No significant difference was observed in any lines in relation to the wild-type.

5.3 DISCUSSION

With the aim to reduce scopoletin accumulation by inhibiting genes responsible for the biosynthesis of scopoletin, the genes of interest (CCoAOMT, F6'H and COMT) were targeted by RNAi constructs (pRNAi-F6'H (1x), pRNAi-CCoAOMT/F6'H (2x) and pRNAi-CCoAOMT/F6'H/COMT (3x)). These RNAi transgenic lines (1xA, 2xA, 2xB, 2xE, 3xA, 3xB, 3xD, 3xE and 3xG) were generated and data were collected on the phenotype of the transgenic plants, storage root PPD progression, scopoletin concentration throughout PPD and gene expression analysis of said genes.

The results show that all three RNAi constructs have significant results on delaying PPD development, with some lines having better results than others. Of all the lines generated, only six lines (1xA, 2xA, 2xB, 3xB and 3xD) were selected for further analyses as they showed the best improvement in PPD symptoms. Two lines (3xA and 3xE) from the triple knock-down plants did not show a significant delay in PPD in comparison to the wild-type control. 3xA, 3xE and 3xG had the lowest weight and height compared to all

the other lines and suffered from developmental problems where a proportion of plants lost apical dominance and were stunted in growth. The plants that lost apical dominance had a bush-like appearance due to the development of the axillary buds (Figure 5.3). Some plants managed to recover and continued growing but others remained dwarfed. This resulted in a high variation within the population of 3xA, 3xE and 3xG as some plants were relatively normal but others were less than half of the height of the normal plants (Figure 5.3). This was unexpected as the plants within a line were clones that were propagated from an individual plant. Besides that, these lines had different insert locations and copy numbers and other 3x lines with the same construct (pRNAi-CCoAOMT/F6'H/COMT) such as 3xB and 3xD did not have this phenotype. This could be due to somaclonal variation as these plants were regenerated from friable embryogenic calli (FEC) and there have been studies that showed phenotypic effects on first generation plants (Larkin and Scowcroft, 1981). Although this is a possibility which could be solved by growing subsequent generations of plants through conventional stem cuttings, it is unlikely to be the case as only plants from the 3x RNAi construct where CCoAOMT, F6'H and COMT are knocked-down simultaneously that have this phenotype. This suggests that the phenotype observed is due to the knock-down of said genes, specifically CCoAOMT and COMT.

CCoAOMT and COMT are involved in the biosynthesis of the monolignols coniferyl alcohol (guaiacyl (G lignin)) and sinapyl alcohol (syringyl (S lignin)). When CCoAOMT was knocked-out in *A. thaliana* using T-DNA insertion, the mutant plant had a collapsed xylem and lower stem lignin content but did not face any severe drawbacks to its fitness (Kai *et al.*, 2008; Do *et al.*, 2007). The COMT T-DNA mutant had no morphological differences but had reduced sinapoylmalate levels and was lacking in S subunits (Goujon *et al.*, 2003). The study by Do *et al.* (2007) showed that when these two genes (CCoAOMT1 and COMT1) were knocked-out simultaneously in the *A. thaliana* double-mutant, the mutant was arrested in the plantlet stage (day 4 post-germination). CCoAOMT and COMT are involved in the methylation steps necessary for the biosynthesis of monolignols (G and S units) and the *ccomt1 comt1* double mutant is mainly composed of *p*-hydroxylphenyl (H) units and severely lacking in G and S units. It is suggested that CCoAOMT1 and COMT1 act together to methylate the monolignols' phenolic ring (C3 position) in *A. thaliana*. Besides that, they are also involved in the biosynthesis of isorhamnetin and sinapoyl malate. It is still uncertain as to what causes the developmental issues experienced by the *ccomt1 comt1* double mutant, but the early growth arrest of the double mutant suggests that these enzymes are essential for plant development.

Bonawitz and Chapple (2013) studied enzymes in the phenylpropanoid pathway leading to lignin biosynthesis with the aim to understand causes of dwarfing and developmental abnormalities observed in lignin biosynthetic mutants. Lignin is essential for normal plant development and *A. thaliana* mutants lacking in lignin fibres are unable to stand erect but do not have a dwarf phenotype (Mitsuda *et al.*, 2007; Zhong *et al.*, 1997). This suggests that loss of lignin in fibres is not the cause of dwarfism. There are suggestions that dwarfism may be due to the collapse of water-transporting cells caused by the negative pressure generated by transpiration in xylems with reduced lignin content (Jones *et al.*, 2001; Piquemal *et al.*, 1998). However, it is unlikely that C3'H, C4H and CCoAOMT/COMT *A. thaliana* mutants that arrest at seedlings would be due to vascular collapse as the negative pressure generated through transpiration would unlikely be enough to cause it (Bonawitz and Chapple, 2013). Other experimental observations showed that *A. thaliana* mutants blocked at different steps of the phenylpropanoid metabolism did not have the same morphological phenotype. If lignin deficiency was the sole cause of dwarfism, mutants would expectedly fall within a phenotypic spectrum of mild to severe dwarfism depending on the level of lignin deficiency (Bonawitz and Chapple, 2013). However, this is not the case as different mutants show varied morphological features such as differences in leaf colour and shape, seedling morphology, apical dominance and fertility. For example, HCT-deficient and C3'H-deficient *A. thaliana* plants that are blocked at sequential steps of the phenylpropanoid pathway display a similar degree of dwarfing but are very morphologically distinguishable from one another (Besseau *et al.*, 2007; Hoffmann *et al.*, 2004; Franke *et al.*, 2002a; Franke *et al.*, 2002b). Like the cassava transgenic plants, C3'H-deficient *A. thaliana* plants show a near complete loss of apical dominance whereas the HCT-deficient plants have strong apical dominance. Besides that, C3'H-deficient plants are sterile unlike HCT-deficient plants which have fertile flowers, though relatively few in numbers (Hoffmann *et al.*, 2004; Franke *et al.*, 2002a). Although these phenotypes may be due to differences in lignin composition, it remains highly possible that these phenotypes come from the loss or disruption of other phenylpropanoid secondary metabolites (Bonawitz and Chapple, 2013).

Phenylpropanoid metabolism is responsible for many secondary metabolites in addition to lignin such as flavonoids, suberin, coumarins which include scopoletin, hydroxycinnamates and many more (Vogt, 2010; Lepiniec *et al.*, 2006; D'Auria and Gershenzon, 2005). These secondary metabolites are essential to different aspects of plant fitness such as reproduction, defence and stress responses but they do not appear to be the cause of dwarfing (Bonawitz and Chapple, 2013). An alternative explanation is

that blocking the phenylpropanoid pathway usually leads to the accumulation of pathway intermediates, products or derivatives. For example, C4H-deficient plants hyperaccumulate cinnamoylated compounds; C3'H-deficient plants hyperaccumulate *p*-coumaroyl esters, anthocyanins and flavanols (Franke *et al.*, 2002a); and 5-hydroxyguaiacyl substituted compounds in COMT-deficient plants (Vanholme *et al.*, 2010). Some of these intermediates are cytotoxic and may inhibit growth at high concentrations, or the hyperaccumulation of these compounds may interfere with signalling pathways involved in plant development (Bonawitz and Chapple, 2013). Flavonols, for example, have been shown to inhibit polar auxin transport (Brown *et al.*, 2001) but the study by Li *et al.* (2010) show that they are not responsible for the dwarfing phenotype. Nevertheless, it is a possibility that the hyperaccumulation of some other pathway product is responsible for the growth inhibition, either through auxin inhibition or some other mechanism. These metabolites may have different biological activity or toxicity effects which is why the hyperaccumulation of these products may cause variation in their morphological phenotypes (Bonawitz and Chapple, 2013). The disruption of auxin signalling is a valid hypothesis as the phenylpropanoid mutants show phenotypes (short petiole, changes in leaf morphology and loss of apical dominance) that correlate with cell elongation inhibition, which is an auxin-regulated process (Cleland, 1987).

Relating back to the cassava triple knock-down mutants with loss of apical dominance and dwarf phenotype, the studies mentioned provide a few alternative explanations to a rather complicated problem. The cassava mutants have the genes knocked-down with RNAi and not knocked-out like in the *A. thaliana* examples which may be why the transgenic plants are able to grow past the plantlet stage unlike the *ccomt1 comt1* T-DNA knock-out mutants (Do *et al.*, 2007). The difference between a T-DNA insertion mutant and RNAi transgenic is that in the former, the gene is knocked-out and no mRNA is produced, while in the latter the mRNA is produced but is titrated out by the production of the RNA from the RNAi construct, which leads to its destruction. Therefore, if the RNAi construct is inserted into a transcriptionally active region of the genome, there could be a 100% reduction of the mRNA, however if it is inserted into a transcriptionally inactive part of the genome, the reduction of the mRNA may be incomplete.

Six lines of the triple RNAi transgenic cassava with copy numbers ranging from one to nine were generated but despite having the same transgene, these lines had different phenotypes. Within the population of 3xA, 3xE and 3xG, despite all plants in the individual line being genetically identical, a proportion of plants had the dwarfing and loss of apical dominance phenotype whereas other plants continued to grow normally. Firstly,

each line will have some variation, whether it can be observed with the eye or not, as the insert location will affect the expression of the transgene. The genomic DNA surrounding the transgene may have local regulatory elements such as an enhancer that could interact with the regulatory elements in the transformation construct that controls transcription thus changing the expression profile of the transgene (Kohli *et al.*, 2010). Secondly, the transgene copy number affects the stability of the transgene. It is natural to assume that with increasing transgene copies, the transgene product will increase accordingly. However, studies have shown that the higher the copy number, the more likely it is for the transgene to be epigenetically silenced through DNA methylation at the locus. For example, the experiment conducted by Meyer *et al.* (1992) where the maize *A1* gene was introduced into petunia plants that produced white flowers. The expression of the transgene would produce pelargonidin, giving red pigment to the white flowers. The experiment showed that red flowers mostly appeared on plants with only a single transgene copy, whereas plants with multiple copy numbers of the transgene had variegated or white flowers. On the other hand, some studies in cereals show that multiple transgene copies do not necessarily lead to epigenetic silencing and may even enhance expression levels in correlation to its copy number. (Gahakwa *et al.*, 2000; Stoger *et al.*, 1998). An explanation for these contrasting results is that epigenetic silencing may be triggered by extremely high levels of expression of the transgene, which may correlate with the transgene copy number (Schubert *et al.*, 2004; Vaucheret *et al.*, 1998; Lindbo *et al.*, 1993).

Coming back to the transgenic cassava plants, the lines which had the dwarfing and loss of apical dominance phenotype all had less than three transgene copies. This could be because the transgene(s) may have integrated in a local environment which is favourable for transgene expression which resulted in a high transgene expression level. As the product of the transgene is the RNA construct which will lead to silencing of the genes CCoAOMT, F6'H and COMT, the higher the expression level of the transgene, the greater the silencing effect on said genes. As mentioned previously, CCoAOMT and COMT are required for healthy plant development (Do *et al.*, 2007) therefore the silencing of these genes will have adverse effects on the plant's development (dwarfing and loss of apical dominance). A plausible explanation for why a proportion of plants within the line did not have this phenotype or have managed to recover from this initial growth defect may be due to epigenetic silencing. The high expression levels of the transgene in these plants may have triggered silencing which is why the plant can recover from the phenotype caused by CCoAOMT/COMT-deficiency. This can also be used to explain why lines 3xB and 3xD did not show phenotypes associated with

CCoAOMT/COMT-deficiency. 3xB had three transgene copies like 3xG but perhaps due to an unfavourable insert location, the gene expression levels could be lower than 3xG, or the transgene could have been silenced before the CCoAOMT/COMT-deficiency could affect the plants' fitness. 3xD had nine transgene copies and was very likely to have been silenced which meant the transgene product which would cause RNAi silencing of said genes would not be produced in levels as high as the other lines therefore phenotype associated with CCoAOMT/COMT-deficiency was not observed.

For the double RNAi knock-down lines 2xA, 2xB and 2xE, there were no loss in fitness but there were differences in leaf morphology which could be due to the deficiency of CCoAOMT/F6'H or somaclonal variation. These results are consistent with the study by Do *et al.* (2007) where *A. thaliana* CCoAOMT mutant, *ccomt1*, had no morphological difference in comparison to the wild-type. Like the 3x plants, there will be differences in expression levels of the transgene due to the same reasons therefore the gene expressions of the CCoAOMT and F6'H will differ in each line.

Previous studies have shown that scopoletin accumulation has a positive correlation with root discolouration which is an accurate representative of deterioration (Fathoni, 2017; Liu, 2016). Most of the transgenic lines had significant improvements in deterioration compared to the wild-type. The results obtained from the 3xA, 3xE and 3xG plants with developmental problems are not reliable due to the poor fitness of the plants causing reduced storage root development. Liu (2016) showed that storage root sizes affected PPD development, with smaller roots showing a higher degree of discolouration than larger roots of the same line. The six lines with the lowest PPD score (least amount of discolouration) were 1xA, 2xA, 2xB, 3xB and 3xD. These lines were selected for biochemical analysis to determine scopoletin concentration and gene expression analyses to find out how successful the transgene was at silencing said genes of interest. The biochemical analysis data showed that these six lines, which had significant reduction in PPD score compared to the wild-type, had significant reduction in scopoletin accumulation, except for 3xD at day 6. This could be because the knock-down was incomplete or insufficient to inhibit scopoletin biosynthesis in the later stage of PPD where scopoletin accumulation was approximately fifteen times the concentration of fresh root (day 0) samples.

Finally, gene expression levels of CCoAOMT, F6'H and COMT for the six lines were determined through qRT-PCR using relative quantity (RQ). The results showed that only F6'H was upregulated during PPD which made sense as studies have shown that F6'H is the main enzyme responsible for scopoletin biosynthesis, with the *A. thaliana* F6'H1

mutant failing to accumulate scopoletin (Kai *et al.*, 2008). When F6'H was knocked-down in cassava (1xA), scopoletin accumulation was significantly reduced thus delaying PPD symptoms (Liu *et al.*, 2017). The results in this chapter confirms this and provides evidence that COMT does not play a major role in scopoletin biosynthesis in PPD development as the gene expression of COMT remained at a very low expression level in both the wild-type and the transgenic lines. CCoAOMT was targeted for RNA silencing because it was hard to target all members of the F6'H family and it would be easier to target CCoAOMT instead. This would potentially produce a bottleneck effect where feruloyl CoA production would be reduced thus in turn reducing scopoletin biosynthesis. Although the 2x and 3x transgenic lines did manage to reduce F6'H expression therefore showing significant reduction in scopoletin accumulation and PPD discolouration, the line showing the greatest delay in PPD and scopoletin accumulation was the 1xA single copy number line generated by Liu *et al.* (2017). Some possible reasons for why 2x and 3x lines were not able to knock-down gene expression of F6'H to the extent that 1xA managed to may be because to achieve the same level of silencing as F6'H in 1xA, the plant may suffer from fitness problems due to deficiency in CCoAOMT for 2x lines and CCoAOMT/COMT in 3x lines. 2x and 3x lines with high silencing activity may have suffered from developmental issues and failed to develop storage roots or may not have made it past the plantlet stage due to early growth arrest. In order to get a viable phenotype, some degree of silencing of the multiple RNAi construct may be necessary, therefore all targeted genes would be affected in parallel. Due to CCoAOMT and COMT having a more important role than F6'H in plant development, the degree of silencing of the transgene to permit viability may be higher than that which would be obtained were F6'H be targeted alone, hence the 1xA line had a greater reduction in F6'H gene expression than the 2x and 3x transgenic lines.

Overall, this study has demonstrated that the F6'H gene family is actively involved in the biosynthesis of scopoletin as shown by Liu *et al.* (2017) and that the alternative pathway with COMT does not contribute towards the accumulation of scopoletin throughout PPD development. It also shows the importance of CCoAOMT and COMT in cassava development and why the manipulation of phenylpropanoid enzymes are risky and potentially detrimental to plant fitness. This investigation was conducted in a controlled environment (glasshouse) and in 1 L pots therefore the results obtained may differ from if the plants were grown in the field, which would include various environmental stresses such as pests, temperature, water availability and others. In conclusion, field trials would be necessary to determine how these plants will behave in real-life scenarios and give clearer insights into the PPD process.

6 GENERAL DISCUSSION

This study aims to further the understanding of scopoletin biosynthesis in cassava during the PPD process. Building upon results from previous studies, this investigation was designed to extend that work to provide a deeper insight into PPD development with the ultimate goal to extend the shelf-life of cassava for the benefit of resource-poor farmers. The strategy used was centred on reducing the accumulation of scopoletin by inhibiting scopoletin biosynthesis through silencing genes involved using RNAi. There are three alternative pathways to scopoletin biosynthesis within the general phenylpropanoid metabolism. The dominant pathway involves the 6'-hydroxylation of feruloyl CoA by F6'H to synthesise scopoletin. The alternative pathways both involve the *O*-methylation of esculetin into scopoletin by COMT.

Scopoletin accumulation was successfully reduced when F6'H genes were knocked-down with RNAi (1x) which resulted in delayed PPD development (Liu *et al.*, 2017). However, scopoletin still accumulated in the F6'H mutants, and this was hypothesised to be due to an incomplete knock-down of F6'H genes or, alternatively, scopoletin was being synthesised through the alternative pathways (Liu, 2016). The attempt to answer this was to generate transgenic cassava lines to simultaneously silence CCoAOMT, F6'H with a double RNAi construct (2x) and CCoAOMT, F6'H and COMT with a triple RNAi construct (3x). Using *A. thaliana* as a reference, the homologous genes for CCoAOMT, F6'H and COMT were identified in cassava. Wild-type, 1x, 2x and 3x cassava were generated and storage roots were harvested. 1x, 2x and 3x transgenic cassava all had significantly reduced scopoletin accumulation and PPD discolouration throughout the PPD response. The 3x transgenic plants also had some interesting phenotypes that could be due to the silencing of said genes. Only F6'H showed significant increase in gene expression during PPD; CCoAOMT and COMT had no significant changes in gene expression levels throughout the PPD response in both wild-type and transgenic cassava plants, suggesting that these genes did not contribute towards scopoletin biosynthesis during PPD. This would confirm that scopoletin is biosynthesised mainly from the dominant pathway and the alternative pathway does not contribute to the accumulation of scopoletin observed during the PPD response.

6.1 THE ROLE OF SCOPOLETIN IN PLANTS

Scopoletin plays a major role in the development of PPD in cassava storage roots and by reducing scopoletin accumulation, the discolouration that is observed during PPD has been successfully delayed (Fathoni, 2017; Liu *et al.*, 2017). The blue-black discolouration is the most visible aspect of PPD and is likely to be due to the oxidation of scopoletin and other phenolics by peroxidase and H_2O_2 (Wheatley and Schwabe, 1985; Edwards *et al.*, 1997). Due to the reactive nature of scopoletin, it is stored as an inactive glucoside known as scopolin. Scopoletin is actively accumulated through *de novo* synthesis and the deglycosylation of its inactive glucoside, scopolin. Although scopoletin contributes towards the short shelf-life of cassava, it is beneficial to the plant because of its probable antioxidant function and antifungal and antimicrobial properties (Rodríguez *et al.*, 2000). Reducing scopoletin biosynthesis may have adverse effect on the fitness of the plant and make it more susceptible to fungal and bacterial attacks. Scopoletin acts as a phytoalexin in plants including cassava, tobacco, *A. thaliana* and sunflower (*Helianthus annuus*). The biosynthesis of scopoletin can be triggered by entomo-chemical or mechanical stimulation (Blagbrough *et al.*, 2010; Kai *et al.*, 2006). Many coumarin-derived compounds including scopoletin have antifungal properties, especially against *Fusarium* species (Ojala *et al.*, 2000; Giesemann *et al.*, 1986). For example, scopoletin can inhibit the germination of a range of fungi (Olson and Roseland, 1991); in tobacco, scopoletin levels increase in the upper stems and leaves upon fungal infection which is not observed in the absence of stress (Reuveni and Cohen, 1978). The study by Gutiérrez-Mellado *et al.* (1996) shows that scopoletin biosynthesis is triggered by stress as only sunflowers growing in the wild accumulated scopoletin in comparison to plants grown in a controlled environment suite. The mechanism of the antifungal activity is rather complicated, it involves the interaction and inhibition of DNA and enzymatic activities which involves the substitution of functional groups on the aromatic rings (Sardari *et al.*, 2000).

Besides that, another study has shown that the scopoletin biosynthesis pathway and scopoletin is crucial for iron (Fe) mobilisation. Fe is a mineral nutrient that is essential for healthy plant growth; plants have evolved mechanisms to increase the solubility of Fe so that Fe can be extracted from Fe-deficient soil and soil with an alkaline pH (Römheld and Marschner, 1984). In *A. thaliana*, the scopoletin pathway is reprogrammed to produce and secrete coumarins, in particular fraxetin, with Fe-mobilising properties upon Fe-deficiency. This is especially important for plants in alkaline soil as free ion activity is very low and can lead to Fe chlorosis. Scopoletin plays an important role as it is hydroxylated at the C8 position by scopoletin 8-hydroxylase (S8H) to produce fraxetin. S8H is induced

by low Fe availability and proceeds to convert scopoletin into fraxetin, which is secreted into the soil solution to aid in Fe uptake (Tsai *et al.*, 2018). The study also showed that when the *A. thaliana* mutant *F6'H1* was grown in Fe-deficient soil, it suffered from severely reduced growth with significant difference of $P<0.0001$ from the wild-type.

Due to the cassava plants being grown in a controlled environment, they have not been exposed to the environmental stresses field-grown cassava would experience. The controlled environment will have its own stresses, especially since cassava thrives in a high humidity environment and there have been some issues with the maintenance of the humidity level; this may lead to different observations as a result. Scopoletin is involved in plant defence, especially in antifungal and antimicrobial activities, and in Fe mobilisation. The transgenic cassava plants with lower scopoletin accumulation may be more susceptible to fungal attacks and may not do as well in the field, especially in alkaline soil with low Fe levels, in comparison to wild-type plants. Although reducing scopoletin biosynthesis may extend the shelf-life of cassava, if this comes at the cost of yield, it won't be a plausible solution. To address these potential problems, field trials will be necessary.

6.2 THE ROLE OF CCOAOMT, F6'H AND COMT

CCoAOMT, F6'H and COMT are enzymes within the general phenylpropanoid metabolism but hold a special interest in this study because they are involved in the biosynthesis of scopoletin. The phenylpropanoid metabolism produces many secondary metabolites and is highly conserved in the majority of higher plant species because it arose early in the evolution of terrestrial plants, aiding the adaptation of plants to the new environment out of water. One of the more obvious key benefits of the phenylpropanoid metabolism is the production of monolignols, also known as hydroxycinnamyl alcohols, which are the building blocks of lignin. This gave plants the ability to grow upright and transport water through the vascular system. Other secondary metabolites that are produced through the phenylpropanoid metabolism are phenolic compounds such as flavonoids which are divided into six subclasses: flavones, flavanones, flavonols, flavanols, isoflavones and anthocyanidins. The most significant function of phenolics is UV photo-protection where flavonoids counteracts UV-B induced oxidative damage by reducing ROS production. (Kusano *et al.*, 2011; Landry *et al.*, 1995). Other roles of phenolic compounds include biotic stress response as phytoalexins against pathogens and herbivores, abiotic stress such as high carbon and nitrogen deficient growth conditions, flower pigmentation for the attraction of insect pollinators and many others

(Samanta *et al.*, 2011). As mentioned above, the coumarin scopoletin plays an important role in Fe mobilisation in which it is hydroxylated by S8H to form fraxetin. When F6'H1 was knocked-out in *A. thaliana*, the mutant suffered from Fe-deficiency and had significantly reduced growth. This may well be the case as well for the transgenic cassava plants with significantly lower scopoletin accumulation compared to the wild-type.

Lignin is made through the dehydrogenative polymerisation of three monolignols known as *p*-coumaroyl alcohol, coniferyl alcohol and sinapyl alcohol, which respectively give rise to *p*-hydroxyphenyl (H), guiacyl (G) and syringyl (S) units of the lignin polymer. CCoAOMT and COMT are responsible for the biosynthesis of the G and S units of the lignin polymer. These two enzymes are involved in the methylation of the lignin precursors; CCoAOMT is responsible for the 3-*O*-methylation of lignin precursors and cooperates with COMT to yield G monomers. In addition, COMT acts as a 5-*O*-methytransferase to synthesise S units (Do *et al.*, 2007). When these genes (CCoAOMT1 and COMT1) were knocked-out by T-DNA insertions in *A. thaliana*, the double mutant *ccomt1 comt1* had reduced G and S lignin and was composed mainly of H non-methoxylated units. The *ccomt1 comt1* plantlet also suffered from growth arrest at 4 days old. This was hypothesised to be due to either the lack of lignin in the xylem thus leading to vascular collapse caused by the negative pressure generated by transpiration; or the hyperaccumulation of pathway intermediates, products or derivatives which may be cytotoxic or interfere with signalling pathways involved in plant development (Bonawitz and Chapple, 2013).

In cassava, CCoAOMT has two genes, MeCCoAOMT1 and MeCCoAOMT2, F6'H has seven genes, MeF6'H1 – MeF6'H7, and COMT has two genes, MeCOMT1 and MeCOMT2. Due to the high similarities between gene members within the gene family, they were ideal to be silenced by siRNA as this can down-regulate multiple genes that share similar sequences (Jackson *et al.*, 2006). As discussed in chapter 5, the transgenic cassava plants had significant reduction in scopoletin accumulation and delayed PPD symptoms. And similar to the *A. thaliana* double mutant *ccomt1 comt1*, the triple knock-down (3x) cassava plants where CCoAOMT, F6'H and COMT had been knocked-down through RNAi suffered developmental issues, such as loss of apical dominance and dwarfism. Although the double and triple knock-down transgenic plants may not be the solution for PPD due to developmental issues, they have provided an insight into the biosynthesis of scopoletin. This study confirms that scopoletin is biosynthesised predominantly through the major pathway through F6'H and the alternative pathways via COMT hardly contribute towards the accumulation of scopoletin observed during PPD.

In addition, the study by Fathoni (2017) showed that despite knocking down the genes for scopoletin-glucosyltransferase, which are responsible for the glucosylation of scopoletin to scopolin, and up-regulating scopolin beta-glucosidase, which is responsible for the deglucosylation of scopolin back into scopoletin, scopoletin still accumulated during PPD. This suggested that scopoletin is mainly synthesised *de novo* during PPD and not through the deglucosylation of scopolin. Reducing the expression of genes involved in scopoletin biosynthesis had a significantly negative effect on the yield (mass) of the storage root. Even the single F6'H knock-down line, which showed no developmental issues except the morphologically different leaves, had significantly ($p<0.001$) reduced mass compared to the wild-type.

This study was only conducted in a controlled glasshouse environment and cassava plants were grown in 1 L pots for 9 months to allow the storage root to mature sufficiently for experimental purposes. The transgenic plants may behave differently under field conditions where they will be able to grow to full maturity and be exposed to biotic and abiotic stresses. This will especially be important for the single F6'H knock-down as no visible negative phenotype was observed, but this might be different in the field especially if it is grown on alkaline soil with a low availability of Fe. Therefore, no accurate conclusion can be drawn from this study alone without a field study.

6.3 APPROACH TO SOLVING PPD

This study has focused on understanding PPD through the inhibition of scopoletin accumulation, to develop strategies to control PPD. Unfortunately, scopoletin is biosynthesised through the phenylpropanoid metabolism which is involved in many aspects of plant fitness. Through the results obtained, though scopoletin has been reduced and PPD successfully delayed, this came at the cost of the plant's health and the yield (height and weight) of the cassava crop, which is significantly reduced in comparison to the wild-type. This suggests that perhaps attempting to prevent PPD through the inhibition of scopoletin may not be feasible.

A large-scale proteomic study conducted by Vanderschuren *et al.* (2014) on cassava PPD found that the majority of over-represented proteins' molecular function included antioxidant and redox activities, suggesting that oxidative stress is an important process that is involved in, and possibly, induces PPD and drives its progression. The study identified the glutathione/ascorbate cycles as having key roles, especially the enzyme ascorbate peroxidase 3 (APX3) that had upregulated protein levels 6 hours post-harvest.

Unfortunately, APX3 is constrained by the availability of ascorbate, which may be why the accumulation of H_2O_2 is observed during PPD development (Xu *et al.*, 2013; Reilly *et al.*, 2003). Enzymes that use glutathione to detoxify H_2O_2 such as glutathione peroxidases (GPX) have also been detected during PPD but their activity did not significantly change throughout the PPD process. Transgenic cassava with an overexpression of the *A. thaliana* cytosolic GPX in the storage roots were generated and showed a significant delay in PPD symptoms from 6 hours until 48 hours after harvest. The increase in GPX activity seemed to limit lipid peroxidation and decrease the accumulation of H_2O_2 during the PPD progression (Vanderschuren *et al.*, 2014).

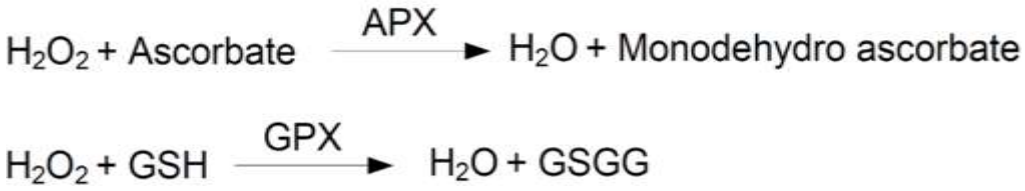


Figure 6.1 **The mode of action of enzymatic ROS scavenging** (Apel and Hirt, 2004).

In addition to scopoletin biosynthesis and reactive oxygen species detoxification, another pathway of interest associated with PPD development is the biosynthesis of ethene. Ethene is involved in wound response signalling and triggers defensive compounds. Ethene is biosynthesised using L-Methionine (L-Met) as a substrate by two enzymes, S-adenosyl-L-methionine (SAM) synthetase and ACC synthase, and the final step is catalysed by ACC oxidase (Wang *et al.*, 2002). Both SAM synthetase and ACC oxidase showed significant upregulation during PPD, which was consistent with the increase in ethene biosynthesis, but ACC synthase which converts SAM into ACC was not detected in the proteomics study (Vanderschuren *et al.*, 2014). The activation of the Met pathway also correlates with the increase in ethene biosynthesis but this may be due to protein synthesis which occurs during PPD (Beeching *et al.*, 1998). As ethene biosynthesis is consistent with the development of PPD, perhaps reducing the production of ethene would delay PPD; but this may interfere with responses to biotic and abiotic stresses and affect plant fitness.

6.4 FUTURE DIRECTIONS FOR CASSAVA PPD

This study has determined that the alternative pathway where esculetin is methoxylated into scopoletin by COMT does not contribute to the *de novo* synthesis of scopoletin during PPD; it has been suggested that the majority of scopoletin accumulated during the PPD progression is through *de novo* synthesis and not through the deglucosylation of scopolin (Fathoni, 2017). This study has also confirmed that F6'H genes are the main contributors to scopoletin biosynthesis, and that the down-regulation of these genes have successfully reduced scopoletin accumulation and delayed PPD progression. The use of RNAi to completely knock-out F6'H activity was unsuccessful due to the large gene family size; perhaps simultaneous targeting of the F6'H genes using RNAi or other gene silencing techniques such as CRISPR-Cas9 would be a better choice to completely inhibit F6'H activity. Field trials will also have to be implemented to get reliable results as environmental factors are important in terms of cassava development and may affect how the roots respond to PPD especially since scopoletin is involved in plant defence and Fe uptake.

Scopoletin undoubtedly plays an important role in the development of PPD in cassava, especially the aspect of discolouration. However, the biosynthesis of scopoletin is part of the phenylpropanoid metabolism and through this study in cassava and other studies in *A. thaliana*, the manipulation of the genes in the phenylpropanoid pathway will likely negatively impact the plant's fitness. An aspect that may be worth investigating is the hydroxylation of scopoletin into fraxetin by scopoletin 8-hydroxylase (S8H) as the study showed that scopoletin and scopolin concentration in the roots of *A. thaliana* with an upregulation of S8H was lower than the wild-type. This suggests that S8H activity seems to mediate the biosynthesis of fraxetin/fraxin pair at the expense of scopoletin/scopolin (Tsai *et al.*, 2018). This may be interesting to investigate if cassava possesses the same mechanism in which fraxetin is used to mobilise Fe in Fe-deficient alkaline soil.

In conclusion, this study has not only added to the understanding of scopoletin biosynthesis in cassava PPD development, but also demonstrates the importance of the phenylpropanoid metabolism in plant development.

REFERENCES

Ahn, Y.O., Shimizu, B., Sakata, K., Gantulga, D., Zhou, Z., Bevan, D.R. & Esen, A., 2010. Scopolin-hydrolyzing β-glucosidases in roots of *Arabidopsis*. *Plant and Cell Physiology,* 51, pp. 132-43.

Alhalasseh, L., 2017. *Phytochemistry of hydroxycoumarins from Manihot esculenta Euphorbiaceae (cassava).* PhD, University of Bath.

Alicai, T., Omongo, C.A., Maruthi, M.N., Hillocks, R.J., Baguma, Y., Kawuki, R., Bua, A., Otim-Nape, G.W. & Colvin, J., 2007. Re-emergence of cassava brown streak disease in Uganda. *Plant Disease,* 91, pp. 24-9.

Allem, A.C., 1994. The origin of *Manihot esculenta* Crantz (Euphorbiaceae). *Genetic Resources and Crop Evolution,* 41, pp. 133-50.

Allem, A.C., 2002. The origins and taxonomy of cassava. In: R.J. Hillocks, J. Thresh & A. Bellotti, eds. *Cassava: Biology, Production and Utilization.* Wallingford: CABI, pp. 1-16.

Alves, A.A.C., 2002. Cassava botany and physiology. In: R.J. Hillocks, J. Thresh & A. Bellotti, eds. *Cassava: Biology, Production and Utilization.* Wallingford: CABI, pp. 67-89.

Apel, K. & Hirt, H., 2004. Reactive oxygen species: Metabolism, oxidative stress, and signal transduction. *Annual Review of Plant Biology,* 55, pp. 373-99.

Apostol, I., Heinstein, P.F. & Low, P.S., 1989. Rapid stimulation of an oxidative burst during elicitation of cultured plant cells. *Plant Physiology,* 90, pp. 109-16.

Baba, A.I., Nogueira, F.C.S., Pinheiro, C.B., Brasil, J.N., Jereissati, E.S., Jucá, T.L., Soares, A.A., Santos, M.F., Domont, G.B. & Campos, F.A.P., 2008. Proteome analysis of secondary somatic embryogenesis in cassava (*Manihot esculenta*). *Plant Science,* 175, pp. 717-23.

Baltas, M., Lapeyre, C., Bedos-Belval, F., Maturano, M., Saint-Aguet, P., Roussel, L., Duran, H. & Grima-Pettenati, J., 2005. Kinetic and inhibition studies of cinnamoyl-CoA reductase 1 from *Arabidopsis thaliana*. *Plant Physiology and Biochemistry,* 43, pp. 746-53.

Bayoumi, S.A.L., Rowan, M.G., Beeching, J.R. & Blagbrough, I.S., 2008. Investigation of biosynthetic pathways to hydroxycoumarins during post-harvest physiological deterioration in cassava roots by using stable isotope labelling. *ChemBioChem,* 9, pp. 3013-22.

Bayoumi, S.A.L., Rowan, M.G., Beeching, J.R. & Blagbrough, I.S., 2010. Constituents and secondary metabolite natural products in fresh and deteriorated cassava roots. *Phytochemistry,* 71, pp. 598-604.

Beeching, J.R., Buschmann, H., Gómez-Vásquez, R., Han, H., Iglesias, I., Li, H., Reilly, K. & Rodriguez, M.X., 1999. An abiotic stress response in cassava: post-harvest physiological deterioration. . In: M.F. Smallwood, C.M. Calvert & D.J. Bowles, eds. *Plant Responses to Environmental Stress.* Oxford: BIOS.

Beeching, J.R., Dodge, A., Moore, K., Phillips, H. & Wenham, J.E., 1994. Physiological deterioration in cassava: possibilities for control. *Tropical Science,* 34, pp. 335-43.

Beeching, J.R., Han, Y., Gómez-Vásquez, R., Day, R.C. & Cooper, R.M., 1998. Wound and defense responses in cassava as related to post-harvest physiological deterioration. *Recent Advances in Phytochemistry,* 32, pp. 231-48.

Beeching, J.R., Reilly, K., Gómez-Vásquez, R., Li, H., Han, Y., Rodriguez, M.X., Buschmann, H., Taylor, N., Fauquet, C. & Tohme, J., 2000. Post-harvest physiological deterioration of cassava. In: M. Nakatani & K. Komaki, eds. *12th Symposium of the International Society for Tropical Root Crops: Potential of root crops for food and industrial resources.* Tsukuba: Cultio.

Bell-Lelong, D.A., Cusumano, J.C., Meyer, K. & Chapple, C., 1997. Cinnamate-4-hydroxylase expression in *Arabidopsis* (Regulation in response to development and the environment). *Plant Physiology,* 113, pp. 729-38.

Bellotti, A.C., Smith, L. & Lapointe, S.L., 1999. Recent advances in cassava pest management. *Annual Review of Entomology,* 44, pp. 343-70.

Besseau, S., Hoffmann, L., Geoffroy, P., Lapierre, C., Pollet, B. & Legrand, M., 2007. Flavonoid accumulation in *Arabidopsis* repressed in lignin synthesis affects auxin transport and plant growth. *The Plant Cell,* 19, pp. 148-62.

Best, R. & Henry, G., 1992. Cassava: towards the year 2000. In: W.M. Roca & T.A. M., eds. *Report of the First Meeting of the International Network for Cassava Genetic Resources.* Cali, Columbia: Centro Internacional de Agricultura Tropical, pp. 3-11.

Blagbrough, I.S., Bayoumi, S.A., Rowan, M.G. & Beeching, J.R., 2010. Cassava: An appraisal of its phytochemistry and its biotechnological prospects. *Phytochemistry,* 71, pp. 1940-51.

Boerjan, W., Ralph, J. & Baucher, M., 2003. Lignin biosynthesis. *Annual Review of Plant Biology,* 54, pp. 519-46.

Bonawitz, N.D. & Chapple, C., 2013. Can genetic engineering of lignin deposition be accomplished without an unacceptable yield penalty? *Current Opinion in Biotechnology,* 24, pp. 336-43.

Borin, K. & Frankow-Lindberg, B.E., 2005. Effects of legumes-cassava intercropping on cassava forage and biomass production. *Journal of Sustainable Agriculture,* 27, pp. 139-51.

Brown, D.E., Rashotte, A.M., Murphy, A.S., Normanly, J., Tague, B.W., Peer, W.A., Taiz, L. & Muday, G.K., 2001. Flavonoids act as negative regulators of auxin transport in vivo in *Arabidopsis*. *Plant Physiology,* 126, pp. 524-35.

Bull, S.E., Ndunguru, J., Gruissem, W., Beeching, J.R. & Vanderschuren, H., 2011. Cassava: constraints to production and the transfer of biotechnology to African laboratories. *Plant Cell Reports,* 30, pp. 779-87.

Bull, S.E., Owiti, J.A., Niklaus, M., Beeching, J.R., Gruissem, W. & Vanderschuren, H., 2009. Agrobacterium-mediated transformation of friable embryogenic calli and regeneration of transgenic cassava. *Nature Protocols,* 4, pp. 1845-54.

Buschmann, H., Reilly, K., Rodriguez, M.X., Tohme, J. & Beeching, J.R., 2000a. Hydrogen peroxide and flavan-3-ols in storage roots of cassava (*Manihot esculenta* Crantz) during postharvest deterioration. *Journal of Agricultural and Food Chemistry,* 48, pp. 5522-9.

Buschmann, H., Rodriguez, M.X., Tohme, J. & Beeching, J.R., 2000b. Accumulation of hydroxycoumarins during post-harvest deterioration of tuberous roots of cassava (*Manihot esculenta* Crantz). *Annals of Botany,* 86, pp. 1153-60.

Byeon, Y., Lee, H.Y., Lee, K. & Back, K., 2014. Caffeic acid *O*-methyltransferase is involved in the synthesis of melatonin by methylating N-acetylserotonin in *Arabidopsis*. *Journal of Pineal Research,* 57, pp. 219-27.

Carabali, A., Bellotti, A.C., Montoya-Lerma, J. & Cuellar, M.E., 2005. Adaptation of *Bemisia tabaci* biotype B (Gennadius) to cassava, *Manihot esculenta* (Crantz). *Crop Protection,* 24, pp. 643-9.

Cardoso, A.P., Mirione, E., Ernesto, M., Massaza, F., Cliff, J., Rezaul Haque, M. & Bradbury, J.H., 2005. Processing of cassava roots to remove cyanogens. *Journal of Food Composition and Analysis,* 18, pp. 451-60.

Ceballos, H., Iglesias, C.A., Pérez, J.C. & Dixon, A.G., 2004. Cassava breeding: opportunities and challenges. *Plant Molecular Biology,* 56, pp. 503-16.

Chandrasekara, A. & Josheph Kumar, T., 2016. Roots and tuber crops as functional foods: A review on phytochemical constituents and their potential health benefits. *International Journal of Food Science,* 2016, pp. 1-15.

Chapple, C., 1998. Molecular-genetic analysis of plant cytochrome P450-dependent monooxygenases. *Annual Review of Plant Physiology and Plant Molecular Biology,* 49, pp. 311-43.

Chapple, C.C., Vogt, T., Ellis, B.E. & Somerville, C.R., 1992. An Arabidopsis mutant defective in the general phenylpropanoid pathway. *The Plant Cell,* 4, pp. 1413-24.

Chong, J., Baltz, R., Schmitt, C., Beffa, R., Fritig, B. & Saindrenan, P., 2002. Downregulation of a pathogen-responsive tobacco UDP-Glc:phenylpropanoid glucosyltransferase reduces scopoletin glucoside accumulation, enhances oxidative stress, and weakens virus resistance. *The Plant Cell,* 14, p. 1093.

Churngchow, N. & Rattarasarn, M., 2001. Biosynthesis of scopoletin in *Hevea brasiliensis* leaves inoculated with *Phytophthora palmivora*. *Journal of Plant Physiology,* 158, pp. 875-82.

Cleland, R.E., 1987. Auxin and cell elongation. In: D.P. J, ed. *Plant hormones and their role in plant growth and development.* Dordrecht: Springer, pp. 132-48.

Cliff, J., Muquingue, H., Nhassico, D., Nzwalo, H. & Bradbury, J.H., 2011. Konzo and continuing cyanide intoxication from cassava in Mozambique. *Food and Chemical Toxicology,* 49, pp. 631-5.

Cock, J.H., 1982. Cassava: a basic energy source in the tropics. *Science,* 218, pp. 755-62.

Cock, J.H., 1985. *Cassava: new potential for a neglected crop.* London: Westfield Press.

Cooke, R.D., Blake, G.G. & Battershill, J.M., 1978. Purification of cassava linamarase. *Phytochemistry,* 17, pp. 381-3.

Costa, M.A., Bedgar, D.L., Moinuddin, S.G.A., Kim, K.W., Cardenas, C.L., Cochrane, F.C., Shockey, J.M., Helms, G.L., Amakura, Y., Takahashi, H., Milhollan, J.K., Davin, L.B., Browse, J. & Lewis, N.G., 2005. Characterization in vitro and in vivo of the putative multigene 4-coumarate:CoA ligase network in *Arabidopsis*: syringyl lignin and sinapate/sinapyl alcohol derivative formation. *Phytochemistry,* 66, pp. 2072-91.

Cottle, W. & Kolattukudy, P.E., 1982. Abscisic acid stimulation of suberization: Induction of enzymes and deposition of polymeric components and associated waxes in tissue cultures of potato tuber. *Plant Physiology,* 70, pp. 775-80.

Couée, I., Sulmon, C., Gouesbet, G. & El Amrani, A., 2006. Involvement of soluble sugars in reactive oxygen species balance and responses to oxidative stress in plants. *Journal of Experimental Botany,* 57, pp. 449-59.

D'Auria, J.C. & Gershenzon, J., 2005. The secondary metabolism of *Arabidopsis thaliana*: growing like a weed. *Current Opinion in Plant Biology,* 8, pp. 308-16.

Derikvand, M.M., Sierra, J.B., Ruel, K., Pollet, B., Do, C.-T., Thévenin, J., Buffard, D., Jouanin, L. & Lapierre, C., 2008. Redirection of the phenylpropanoid pathway to feruloyl malate in *Arabidopsis* mutants deficient for cinnamoyl-CoA reductase 1. *Planta,* 227, pp. 943-56.

Dixon, R.A., Achnine, L., Kota, P., Liu, C.-J., Reddy, M.S.S. & Wang, L., 2002. The phenylpropanoid pathway and plant defence—a genomics perspective. *Molecular Plant Pathology,* 3, pp. 371-90.

Do, C.T., Pollet, B., Thévenin, J., Sibout, R., Denoue, D., Barrière, Y., Lapierre, C. & Jouanin, L., 2007. Both caffeoyl coenzyme A 3-*O*-methyltransferase 1 and caffeic acid *O*-methyltransferase 1 are involved in redundant functions for lignin, flavonoids and sinapoyl malate biosynthesis in *Arabidopsis*. *Planta,* 226, pp. 1117-29.

Du, L., Bokanga, M., Møller, B.L. & Halkier, B.A., 1995. The biosynthesis of cyanogenic glucosides in roots of cassava. *Phytochemistry,* 39, pp. 323-6.

Ecker, J.R. & Davis, R.W., 1987. Plant defense genes are regulated by ethylene. *Proceedings of the National Academy of Sciences,* 84, pp. 5202-6.

Edwards, R., Stones, S.M., Gutiérrez-Mellado, M.C. & Jorrin, J., 1997. Characterization and inducibility of a scopoletin-degrading enzyme from sunflower. *Phytochemistry,* 45, pp. 1109-14.

Eggum, B.O., 1970. The protein quality of cassava leaves. *British Journal of Nutrition,* 24, pp. 761-8.

Ehlting, J., Büttner, D., Wang, Q., Douglas, C.J., Somssich, I.E. & Kombrink, E., 1999. Three 4-coumarate:coenzyme A ligases in *Arabidopsis thaliana* represent two evolutionarily divergent classes in angiosperms. *The Plant Journal,* 19, pp. 9-20.

Ejechi, B.O. & Souzey, J.A., 1999. Inhibition of biodeterioration of yam tuber *Dioscorea rotundata* Poir in storage with phenolic extract of *Acalypha hispida* Burm. f. leaves. *Journal of Stored Products Research,* 35, pp. 127-34.

El-Sharkawy, M.A., 2003. Cassava biology and physiology. *Plant Molecular Biology,* 53, pp. 621-41.

El-Sharkawy, M.A. & Cock, J.H., 1990. Photosynthesis of cassava (*Manihot esculenta*). *Experimental Agriculture,* 26, pp. 325-40.

Ephraim, N., Yona, B., Evans, A., Sharon, A. & Titus, A., 2015. Effect of cassava brown streak disease (CBSD) on cassava (*Manihot esculenta* Crantz) root storage components, starch quantities and starch quality properties. *International Journal of Plant Physiology and Biochemistry,* 7, pp. 12-22.

Ernesto, M., Cardoso, A.P., Nicala, D., Mirione, E., Massaza, F., Cliff, J., Haque, M.R. & Bradbury, J.H., 2002. Persistent konzo and cyanogen toxicity from cassava in northern Mozambique. *Acta Tropica,* 82, pp. 357-62.

FAO, 2017. Food Outlook: Biannual Report on Global Food Markets. Available: http://www.fao.org/3/a-I8080e.pdf [Accessed 2018/09/11].

Fargette, D., Konaté, G., Fauquet, C., Muller, E., Peterschmitt, M. & Thresh, J.M., 2006. Molecular ecology and emergence of tropical plant viruses. *Annual Review of Phytopathology,* 44, pp. 235-60.

Farmer, E.E. & Ryan, C.A., 1992. Octadecanoid precursors of jasmonic acid activate the synthesis of wound-inducible proteinase inhibitors. *The Plant Cell Online,* 4, pp. 129-34.

Fathoni, A., 2017. *The role of interconversion of scopoletin and scopolin in cassava postharvest physiological deterioration (PPD).* PhD, University of Bath.

Ferrer, J.L., Austin, M.B., Stewart, C. & Noel, J.P., 2008. Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. *Plant Physiology and Biochemistry,* 46, pp. 356-70.

Franke, R., Hemm, M.R., Denault, J.W., Ruegger, M.O., Humphreys, J.M. & Chapple, C., 2002a. Changes in secondary metabolism and deposition of an unusual lignin in the *ref8* mutant of *Arabidopsis*. *The Plant Journal,* 30, pp. 47-59.

Franke, R., Humphreys, J.M., Hemm, M.R., Denault, J.W., Ruegger, M.O., Cusumano, J.C. & Chapple, C., 2002b. The *Arabidopsis REF8* gene encodes the 3-hydroxylase of phenylpropanoid metabolism. *The Plant Journal,* 30, pp. 33-45.

Gachon, C., Baltz, R. & Saindrenan, P., 2004. Over-expression of a scopoletin glucosyltransferase in *Nicotiana tabacum* leads to precocious lesion formation during the hypersensitive response to tobacco mosaic virus but does not affect virus resistance. *Plant Molecular Biology,* 54, pp. 137-46.

Gahakwa, D., Maqbool, S.B., Fu, X., Sudhakar, D., Christou, P. & Kohli, A., 2000. Transgenic rice as a system to study the stability of transgene expression: multiple heterologous transgenes show similar behaviour in diverse genetic backgrounds. *Theoretical and Applied Genetics,* 101, pp. 388-99.

Garcia, D., Sanier, C., Macheix, J.J. & D'Auzac, J., 1995. Accumulation of scopoletin in *Hevea brasiliensis* infected by *Microcyclus ulei* (P. Henn.) V. ARX and evaluation of its fungitoxicity for three leaf pathogens of rubber tree. *Physiological and Molecular Plant Pathology,* 47, pp. 213-23.

Giesemann, A., Biehl, B. & Lieberei, R., 1986. Identification of scopoletin as a phytoalexin of the rubber tree *Hevea brasiliensis*. *Journal of Phytopathology,* 117, pp. 373-6.

Góez-Vásquez, R., Day, R., Buschmann, H., Randles, S., Beeching, J.R. & Cooper, R.M., 2004. Phenylpropanoids, phenylalanine ammonia lyase and peroxidases in elicitor-challenged cassava (*Manihot esculenta*) suspension cells and leaves. *Annals of Botany,* 94, pp. 87-97.

Gomez, M.A., Lin, Z.D., Moll, T., Chauhan, R.D., Renninger, K., Beyene, G., Taylor, N.J., Carrington, J.C., Staskawicz, B.J. & Bart, R., 2018. Simultaneous CRISPR/Cas9-mediated editing of cassava *eIF4E* isoforms *nCBP-1* and *nCBP-2* reduces cassava brown streak disease symptom severity and incidence. *Plant Biotechnology Journal*, pp. 1-33.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N. & Rokhsar, D.S., 2012. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research,* 40, pp. 1178-86.

Goujon, T., Sibout, R., Pollet, B., Maba, B., Nussaume, L., Bechtold, N., Lu, F., Ralph, J., Mila, I., Barrière, Y., Lapierre, C. & Jouanin, L., 2003. A new *Arabidopsis thaliana* mutant deficient in the expression of *O*-methyltransferase impacts lignins and sinapoyl esters. *Plant Molecular Biology,* 51, pp. 973-89.

Gutiérrez-Mellado, M.C., Edwards, R., Tena, M., Cabello, F., Serghini, K. & Jorrín, J., 1996. The production of coumarin phytoalexins in different plant organs of sunflower (*Helianthus annuus* L.). *Journal of Plant Physiology,* 149, pp. 261-6.

Hamberger, B. & Hahlbrock, K., 2004. The 4-coumarate:CoA ligase gene family in *Arabidopsis thaliana* comprises one rare, sinapate-activating and three commonly occurring isoenzymes. *Proceedings of the National Academy of Sciences,* 101, pp. 2209-14.

Hillocks, R.J. & Jennings, D.L., 2003. Cassava brown streak disease: A review of present knowledge and research needs. *International Journal of Pest Management,* 49, pp. 225-34.

Hirose, S., Data, E. & Maturan, E., 1984. Relation of respiration and ethylene production to postharvest deterioration in cassava roots from pruned and unpruned plants. In: I. Uritani & E.D. Reyes, eds. *Tropical Root Crops: Postharvest Physiology and Processing.* Tokyo: Japan Scientific Societies Press.

Hoekema, A., Hirsch, P.R., Hooykaas, P.J.J. & Schilperoort, R.A., 1983. A binary plant vector strategy based on separation of vir- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. *Nature,* 303, pp. 179-80.

Hoffmann, L., Besseau, S., Geoffroy, P., Ritzenthaler, C., Meyer, D., Lapierre, C., Pollet, B. & Legrand, M., 2004. Silencing of hydroxycinnamoyl-coenzyme A shikimate/quinate hydroxycinnamoyltransferase affects phenylpropanoid biosynthesis. *The Plant Cell,* 16, pp. 1446-65.

Hoffmann, L., Maury, S., Martz, F., Geoffroy, P. & Legrand, M., 2003. Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism. *Journal of Biological Chemistry,* 278, pp. 95-103.

Holmes, G.J. & Stange, R.R., 2002. Influence of wound type and storage duration on susceptibility of sweet potatoes to Rhizopus soft rot. *Plant Disease,* 86, pp. 345-8.

Hu, W., Kong, H., Guo, Y., Zhang, Y., Ding, Z., Tie, W., Yan, Y., Huang, Q., Peng, M., Shi, H. & Guo, A., 2016. Comparative physiological and transcriptomic analyses reveal the actions of melatonin in the delay of postharvest physiological deterioration of cassava. *Frontiers in Plant Science,* 7, p. 736.

Huang, J., Gu, M., Lai, Z., Fan, B., Shi, K., Zhou, Y.H., Yu, J.Q. & Chen, Z., 2010. Functional analysis of the *Arabidopsis PAL* gene family in plant growth, development, and response to environmental stress. *Plant Physiology,* 153, pp. 1526-38.

Huang, Q.X., Wang, X.C., Kong, H., Guo, Y.L. & Guo, A.P., 2013. An efficient DNA isolation method for tropical plants. *African Journal of Biotechnology,* 12, pp. 2727-32.

Hunt, L., Wholey, D. & Cock, J., Growth physiology of cassava (*Manihot esculenta* Crantz). In: *Field Crop Abstracts*, 1977. pp. 77-91.

Ikediobi, C.O., Chelvarajan, R.L. & Ukoha, A.I., 1989. Biochemical aspects of wound healing in yams (*Dioscorea spp*). *Journal of the Science of Food and Agriculture,* 48, pp. 131-9.

Imaseki, H., Uhttani, I. & Stahmann, M.A., 1968. Production of ethylene by injured sweet potato root tissue. *Plant and Cell Physiology,* 9, pp. 757-68.

Isendahl, C., 2011. The domestication and early spread of manioc (*Manihot esculenta* Crantz): A brief synthesis. *Latin American Antiquity,* 22, pp. 452-68.

Iyer, S., Mattinson, D.S. & Fellman, J.K., 2010. Study of the early events leading to cassava root postharvest deterioration. *Tropical Plant Biology,* 3, pp. 151-65.

Jackson, A.L., Burchard, J., Schelter, J., Chau, B.N., Cleary, M., Lim, L. & Linsley, P.S., 2006. Widespread siRNA “off-target” transcript silencing mediated by seed region sequence complementarity. *RNA,* 12, pp. 1179-87.

Jansson, C., Westerbergh, A., Zhang, J., Hu, X. & Sun, C., 2009. Cassava, a potential biofuel crop in (the) People’s Republic of China. *Applied Energy,* 86, pp. S95-S9.

Jennings, D.L. & Iglesias, C., 2002. Breeding for crop improvement. In: R.J. Hillocks, J. Thresh & A. Bellotti, eds. *Cassava: Biology, Production and Utilization.* New York: CABI Publishing, pp. 149-66.

Jones, L., Ennos, A.R. & Turner, S.R., 2001. Cloning and characterization of irregular xylem4 (irx4): a severely lignin-deficient mutant of *Arabidopsis*. *The Plant Journal,* 26, pp. 205-16.

Kai, K., Mizutani, M., Kawamura, N., Yamamoto, R., Tamai, M., Yamaguchi, H., Sakata, K. & Shimizu, B., 2008. Scopoletin is biosynthesized via ortho-hydroxylation of feruloyl CoA by a 2-oxoglutarate-dependent dioxygenase in *Arabidopsis thaliana*. *The Plant Journal,* 55, pp. 989-99.

Kai, K., Shimizu, B., Mizutani, M., Watanabe, K. & Sakata, K., 2006. Accumulation of coumarins in *Arabidopsis thaliana*. *Phytochemistry,* 67, pp. 379-86.

Kato, C. & Uritani, I., 1976. Changes in carbohydrate content of sweet potato in response to cutting and infection by black rot fungus. *Japanese Journal of Phytopathology,* 42, pp. 181-6.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P. & Drummond, A., 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics,* 28, pp. 1647-9.

Kim, B.G., Lee, Y., Hur, H.G., Lim, Y. & Ahn, J.H., 2006. Production of three *O*-methhylated esculetins with *Escherichia coli* expressing *O*-methyltransferase from poplar. *Bioscience, Biotechnology, and Biochemistry,* 70, pp. 1269-72.

Kim, E.K., Kwon, K.B., Shin, B.C., Seo, E.A., Lee, Y.R., Kim, J.S., Park, J.W., Park, B.H. & Ryu, D.G., 2005. Scopoletin induces apoptosis in human promyeloleukemic cells, accompanied by activations of nuclear factor κB and caspase-3. *Life Sciences,* 77, pp. 824-36.

Kim, H., Van Bo, N., Long, H., Trong Hien, N., Ceballos, H. & Howeler, R., 2017. Current situation of cassava in Vietnam. *A New Future for Cassava in Asia: Its Use as Food, Feed and fuel to Benefit the Poor.* Vientiane, Lao PDR, pp. 100-12.

Kim, S.J., Kim, K.W., Cho, M.H., Franceschi, V.R., Davin, L.B. & Lewis, N.G., 2007. Expression of cinnamyl alcohol dehydrogenases and their putative homologues during *Arabidopsis thaliana* growth and development: Lessons for database annotations? *Phytochemistry,* 68, pp. 1957-74.

Koehorst-van Putten, H.J.J., Sudarmonowati, E., Herman, M., Pereira-Bertram, I.J., Wolters, A.M.A., Meima, H., de Vetten, N., Raemakers, C.J.J.M. & Visser, R.G.F., 2012. Field testing and exploitation of genetically modified cassava with low-amylose or amylose-free starch in Indonesia. *Transgenic Research,* 21, pp. 39-50.

Kohli, A., Miro, B. & Twyman, R.M., 2010. Transgene Integration, Expression and Stability in Plants: Strategies for Improvements. In: C. Kole, C.H. Michler, A.G. Abbott & T.C. Hall, eds. *Transgenic Crop Plants: Principles and Development.* Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 201-37.

Kusano, M., Tohge, T., Fukushima, A., Kobayashi, M., Hayashi, N., Otsuki, H., Kondou, Y., Goto, H., Kawashima, M. & Matsuda, F., 2011. Metabolomics reveals comprehensive reprogramming involving two independent metabolic responses of *Arabidopsis* to UV-B light. *The Plant Journal,* 67, pp. 354-69.

Landry, L.G., Chapple, C.C. & Last, R.L., 1995. *Arabidopsis* mutants lacking phenolic sunscreens exhibit enhanced ultraviolet-B injury and oxidative damage. *Plant Physiology,* 109, pp. 1159-66.

Langer, M., Reck, G., Reed, J. & Retey, J., 1994. Identification of serine-143 as the most likely precursor of dehydroalanine in the active site of histidine ammonia-lyase. A study of the overexpressed enzyme by site-directed mutagenesis. *Biochemistry,* 33, pp. 6462-7.

Larkin, P.J. & Scowcroft, W.R., 1981. Somaclonal variation—a novel source of variability from cell cultures for plant improvement. *Theoretical and Applied Genetics,* 60, pp. 197-214.

Lauvergeat, V., Lacomme, C., Lacombe, E., Lasserre, E., Roby, D. & Grima-Pettenati, J., 2001. Two cinnamoyl-CoA reductase (CCR) genes from *Arabidopsis thaliana* are differentially expressed during development and in response to infection with pathogenic bacteria. *Phytochemistry,* 57, pp. 1187-95.

Lebot, V., 2009. *Tropical root and tuber crops: cassava, sweet potato, yams and aroids.* Cabi.

Lee, D., Ellard, M., Wanner, L.A., Davis, K.R. & Douglas, C.J., 1995. The *Arabidopsis thaliana* 4-coumarate:CoA ligase (4CL) gene: stress and developmentally regulated expression and nucleotide sequence of its cDNA. *Plant Molecular Biology,* 28, pp. 871-84.

Legg, J.P. & Fauquet, C.M., 2004. Cassava mosaic geminiviruses in Africa. *Plant Molecular Biology,* 56, pp. 585-99.

Legg, J.P., Owor, B., Sseruwagi, P. & Ndunguru, J., 2006a. Cassava mosaic virus disease in East and Central Africa: Epidemiology and management of a regional pandemic. *Advances in Virus Research,* 67, pp. 355-418.

Legg, J.P., Owor, B., Sseruwagi, P. & Ndunguru, J., 2006b. Cassava mosaic virus disease in east and central Africa: Epidemiology and management of a regional pandemic. *Advances in Virus Research.* Academic Press, pp. 355-418.

Lepiniec, L., Debeaujon, I., Routaboul, J.-M., Baudry, A., Pourcel, L., Nesi, N. & Caboche, M., 2006. Genetics and biochemistry of seed flavonoids. *Annual Review of Plant Biology,* 57, pp. 405-30.

Levsh, O., Chiang, Y.C., Tung, C.F., Noel, J.P., Wang, Y. & Weng, J.K., 2016. Dynamic conformational states dictate selectivity toward the native substrate in a substrate-permissive acyltransferase. *Biochemistry,* 55, pp. 6314-26.

Li, X., Bonawitz, N.D., Weng, J.K. & Chapple, C., 2010. The growth reduction associated with repressed lignin biosynthesis in *Arabidopsis thaliana* is independent of flavonoids. *The Plant Cell*, pp. 1-13.

Lindbo, J.A., Silva-Rosales, L., Proebsting, W.M. & Dougherty, W.G., 1993. Induction of a highly specific antiviral state in transgenic plants: implications for regulation of gene expression and virus resistance. *The Plant Cell,* 5, pp. 1749-59.

Liu, S., 2016. *The role of scopoletin in cassava post-harvest physiological deterioration.* PhD, University of Bath.

Liu, S., Zainuddin, I.M., Vanderschuren, H., Doughty, J. & Beeching, J.R., 2017. RNAi inhibition of feruloyl CoA 6′-hydroxylase reduces scopoletin biosynthesis and post-harvest physiological deterioration in cassava (*Manihot esculenta* Crantz) storage roots. *Plant Molecular Biology,* 94, pp. 185-95.

Liu, X., Zhang, L., Fu, X., Chen, K. & Qian, B., 2001. Effect of scopoletin on PC3 cell proliferation and apoptosis. *Acta Pharmacologica Sinica,* 22, pp. 929-33.

Lokko, Y., Anderson, J.V., Rudd, S., Raji, A., Horvath, D., Mikel, M.A., Kim, R., Liu, L., Hernandez, A., Dixon, A.G.O. & Ingelbrecht, I.L., 2007. Characterization of an 18,166 EST dataset for cassava (*Manihot esculenta* Crantz) enriched for drought-responsive genes. *Plant Cell Reports,* 26, pp. 1605-18.

Low, P. & Merida, J., 1996. The oxidative burst in plant defense: function and signal transduction. *Physiologia Plantarum,* 96, pp. 533-42.

Lozano, J.C., Toro, J., Castro, A. & Bellotti, A., 1977. *Production of cassava planting material.* Cali, Colombia: CIAT.

Lulai, E.C. & Corsini, D.L., 1998. Differential deposition of suberin phenolic and aliphatic domains and their roles in resistance to infection during potato tuber (*Solanum tuberosum* L.) wound-healing. *Physiological and Molecular Plant Pathology,* 53, pp. 209-22.

Lulai, E.C., Suttle, J.C. & Pederson, S.M., 2008. Regulatory involvement of abscisic acid in potato tuber wound-healing. *Journal of Experimental Botany,* 59, pp. 1175-86.

Ma, Q., Zhang, T., Zhang, P. & Wang, Z.Y., 2016. Melatonin attenuates postharvest physiological deterioration of cassava storage roots. *Journal of Pineal Research,* 60, pp. 424-34.

Mann, C., 1997. Reseeding the green revolution. *Science,* 277, pp. 1038-43.

Manu-Aduening, J., Lamboll, R., Dankyi, A. & Gibson, R., 2005. Cassava diversity in Ghanaian farming systems. *Euphytica,* 144, pp. 331-40.

Maruthi, M.N., Hillocks, R.J., Mtunda, K., Raya, M.D., Muhanna, M., Kiozia, H., Rekha, A.R., Colvin, J. & Thresh, J.M., 2005. Transmission of Cassava brown streak virus by *Bemisia tabaci* (Gennadius). *Journal of Phytopathology,* 153, pp. 307-12.

May, M.J., Vernoux, T., Leaver, C., Montagu, M.V. & Inzé, D., 1998. Glutathione homeostasis in plants: implications for environmental sensing and plant development. *Journal of Experimental Botany,* 49, pp. 649-67.

McCabe, M.S., Power, J.B., de Laat, A.M.M. & Davey, M.R., 1997. Detection of single-copy genes in DNA from transgenic plants by nonradioactive southern blot analysis. *Molecular Biotechnology,* 7, pp. 79-84.

Meyer, K., Cusumano, J.C., Somerville, C. & Chapple, C.C., 1996. Ferulate-5-hydroxylase from *Arabidopsis thaliana* defines a new family of cytochrome P450-dependent monooxygenases. *Proceedings of the National Academy of Sciences,* 93, pp. 6869-74.

Meyer, P., Linn, F., Heidmann, I., Meyer, H., Niedenhof, I. & Saedler, H., 1992. Endogenous and environmental factors influence 35S promoter methylation of a maize A1 gene construct in transgenic petunia and its colour phenotype. *Molecular and General Genetics MGG,* 231, pp. 345-52.

Mitsuda, N., Iwase, A., Yamamoto, H., Yoshida, M., Seki, M., Shinozaki, K. & Ohme-Takagi, M., 2007. NAC transcription factors, NST1 and NST3, are key regulators of the formation of secondary walls in woody tissues of *Arabidopsis*. *The Plant Cell,* 19, pp. 270-80.

Mittler, R., Vanderauwera, S., Gollery, M. & Van Breusegem, F., 2004. Reactive oxygen gene network of plants. *Trends in plant science,* 9, pp. 490-8.

Moreno, I., Gruissem, W. & Vanderschuren, H., 2011. Reference genes for reliable potyvirus quantitation in cassava and analysis of cassava brown streak virus load in host varieties. *Journal of Virological Methods,* 177, pp. 49-54.

Nassar, N. & Ortiz, R., 2010. Breeding cassava to feed the poor. *Scientific American,* 302, pp. 78-84.

Negrel, J., Javelle, F. & Paynot, M., 1993. Wound-induced tyramine hydroxycinnamoyl transferase in potato (*Solanum tuberosum*) tuber discs. *Journal of Plant Physiology,* 142, pp. 518-24.

Negrotto, D., Jolley, M., Beer, S., Wenck, A.R. & Hansen, G., 2000. The use of phosphomannose-isomerase as a selectable marker to recover transgenic maize plants (*Zea mays* L.) via Agrobacterium transformation. *Plant Cell Reports,* 19, pp. 798-803.

Nguyen, T.L.T., Gheewala, S.H. & Garivait, S., 2007. Full chain energy analysis of fuel ethanol from cassava in Thailand. *Environmental Science & Technology,* 41, pp. 4135-42.

Nishikawa, F., Kato, M., Hyodo, H., Ikoma, Y., Sugiura, M. & Yano, M., 2005. Effect of sucrose on ascorbate level and expression of genes involved in the ascorbate biosynthesis and recycling pathway in harvested broccoli florets. *Journal of Experimental Botany,* 56, pp. 65-72.

Noon, R.A. & Booth, R.H., 1977. Nature of post-harvest deterioration of cassava roots. *Transactions of the British Mycological Society,* 69, pp. 287-90.

Ojala, T., Remes, S., Haansuu, P., Vuorela, H., Hiltunen, R., Haahtela, K. & Vuorela, P., 2000. Antimicrobial activity of some coumarin containing herbal plants growing in Finland. *Journal of Ethnopharmacology,* 73, pp. 299-305.

Olsen, K. & Schaal, B., 2001. Microsatellite variation in cassava (*Manihot esculenta*, Euphorbiaceae) and its wild relatives: further evidence for a southern Amazonian origin of domestication. *American Journal Of Botany,* 88, pp. 131-42.

Olsen, K.M. & Schaal, B.A., 1999. Evidence on the origin of cassava: Phylogeography of *Manihot esculenta*. *Proceedings of the National Academy of Sciences,* 96, pp. 5586-91.

Olson, M.M. & Roseland, C.R., 1991. Induction of the coumarins scopoletin and ayapin in sunflower by insect–feeding stress and effects of coumarins on the feeding of sunflower beetle (Coleoptera: Chrysomelidae). *Environmental Entomology,* 20, pp. 1166-72.

Onwueme, I.C., 1978. *The tropical tuber crops: yams, cassava, sweet potato, and cocoyams.* Chichester: John Wiley and Sons.

Owiti, J., Grossmann, J., Gehrig, P., Dessimoz, C., Laloi, C., Hansen, M.B., Gruissem, W. & Vanderschuren, H., 2011. iTRAQ-based analysis of changes in the cassava root proteome reveals pathways associated with post-harvest physiological deterioration. *The Plant Journal,* 67, pp. 145-56.

Passam, H. & Noon, R., 1977. Deterioration of yams and cassava during storage. *Annals of Applied Biology,* 85, pp. 436-40.

Patil, B.L. & Fauquet, C.M., 2009. Cassava mosaic geminiviruses: actual knowledge and perspectives. *Molecular Plant Pathology,* 10, pp. 685-701.

Piquemal, J., Lapierre, C., Myton, K., O'connell, A., Schuch, W., Grima-pettenati, J. & Boudet, A.M., 1998. Down-regulation of cinnamoyl-CoA reductase induces significant changes of lignin profiles in transgenic tobacco plants. *The Plant Journal,* 13, pp. 71-83.

Pujol, B., Guillaume, G., Laurent, G., Marina, P.K., Marianne, E., Hossaert-McKey, M. & McKey, D., 2002. Germination ecology of cassava (*Manihot esculenta* Crantz, Euphorbiaceae) in traditional agroecosystems: Seed and seedling biology of a vegetatively propagated domesticated plant. *Economic Botany,* 56, pp. 366-79.

Qin, Y., Djabou, A.S.M., An, F., Li, K., Li, Z., Yang, L., Wang, X. & Chen, S., 2017. Proteomic analysis of injured storage roots in cassava (*Manihot esculenta* Crantz) under postharvest physiological deterioration. *PLOS ONE,* 12, p. e0174238.

Raes, J., Rohde, A., Christensen, J.H., Van de Peer, Y. & Boerjan, W., 2003. Genome-wide characterization of the lignification toolbox in Arabidopsis. *Plant Physiology,* 133, pp. 1051-71.

Ravi, V., Aked, J. & Balagopalan, C., 1996. Review on tropical root and tuber crops I. Storage methods and quality changes. *Critical Reviews in Food Science and Nutrition,* 36, pp. 661-709.

Ravindran, G. & Wanasundera, J., 1992. Chemical changes in yam tubers (*Dioscorea alata* and *D. esculenta*) during storage. *Tropical Science,* 33, pp. 57-62.

Ray, R.C. & Ravi, V., 2005. Post harvest spoilage of sweet potato in tropics and control measures. *Critical Reviews in Food Science and Nutrition,* 45, pp. 623-44.

Reilly, K., Bernal, D., Cortés, D.F., Gómez-Vásquez, R., Tohme, J. & Beeching, J.R., 2007. Towards identifying the full set of genes expressed during cassava post-harvest physiological deterioration. *Plant Molecular Biology,* 64, pp. 187-203.

Reilly, K., Gómez-Vásquez, R., Buschmann, H., Tohme, J. & Beeching, J.R., 2003. Oxidative stress responses during cassava post-harvest physiological deterioration. *Plant Molecular Biology,* 53, pp. 669-85.

Reuveni, M. & Cohen, Y., 1978. Growth retardation and changes in phenolic compounds, with special reference to scopoletin, in mildewed and ethylene-treated tobacco plants. *Physiological Plant Pathology,* 12, pp. 179-89.

Revers, F. & García, J.A., 2015. Molecular biology of Potyviruses. *Advances in Virus Research,* 92, pp. 101-99.

Rhodes, J. & Wooltorton, L., 1978. The biosynthesis of phenolic compounds in wounded plant storage tissues. In: G. Kahl, ed. *Biochemistry of Wounded Plant Tissues.* Berlin: Walter de Gruyter, pp. 243-86.

Rickard, J.E., 1985. Physiological deterioration of cassava roots. *Journal of the Science of Food and Agriculture,* 36, pp. 167-76.

Rodríguez, M.X., Buschmann, H., Iglesias, C. & Beeching, J., Production of antimicrobial compounds in cassava (*Manihot esculenta* Crantz) roots during post-harvest physiological deterioration. In: L. Carvalho, A. Thro & A. Vilarinho, eds. *Cassava Biotechnology: IVth International Scientific Meeting of the Cassava Biotechnology Network*, 2000, Embrapa, Brazil. pp. 320-8.

Rogers, D.J. & Appan, S.G., 1973. *Flora Neotropica monograph No 13. Manihot Manihotoides (Euphorbiaceae).* New York: Hafner Press.

Rohde, A., Morreel, K., Ralph, J., Goeminne, G., Hostyn, V., De Rycke, R., Kushnir, S., Van Doorsselaere, J., Joseleau, J.P., Vuylsteke, M., Van Driessche, G., Van Beeumen, J., Messens, E. & Boerjan, W., 2004. Molecular phenotyping of the *pal1pal2* mutants of *Arabidopsis thaliana* reveals far-reaching consequences on phenylpropanoid, amino acid, and carbohydrate metabolism. *The Plant Cell,* 16, pp. 2749-71.

Römheld, V. & Marschner, H., 1984. Plant-induced pH changes in the rhizosphere of “Fe-efficient” and “Fe-inefficient” soybean and corn cultivars. *Journal of Plant Nutrition,* 7, pp. 623-30.

Rudi, N., Norton, G.W., Alwang, J. & Asumugha, G., 2010. Economic impact analysis of marker-assisted breeding for resistance to pests and post-harvest deterioration in cassava. *African Journal of Agriculture Resource Economics,* 4, pp. 110-22.

Ruegger, M. & Chapple, C., 2001. Mutations that reduce sinapoylmalate accumulation in *Arabidopsis thaliana* define loci with diverse roles in phenylpropanoid metabolism. *Genetics,* 159, pp. 1741-9.

Sakai, T. & Nakagawa, Y., 1988. Diterpenic stress metabolites from cassava roots. *Phytochemistry,* 27, pp. 3769-79.

Sakai, T., Nakagawa, Y., Uritani, I. & Data, E.S., 1986. Occurrence of various kinds of metabolites in physiologically and microbially damaged cassava (*Manihot esculenta* Crantz) roots. *Agricultural and Biological Chemistry,* 50, pp. 2905-7.

Salcedo, A. & Siritunga, D., 2011. Insights into the physiological, biochemical and molecular basis of postharvest deterioration in cassava (*Manihot esculenta*) roots. *American Journal of Experimental Agriculture,* 1, pp. 414-31.

Samanta, A., Das, G. & Das, S.K., 2011. Roles of flavonoids in plants. *International Journal of Pharmaceutical Sciences and Technology,* 100, pp. 12-35.

Sánchez, T., Chávez, A.L., Ceballos, H., Rodriguez-Amaya, D.B., Nestel, P. & Ishitani, M., 2006. Reduction or delay of post-harvest physiological deterioration in cassava roots with higher carotenoid content. *Journal of the Science of Food and Agriculture,* 86, pp. 634-9.

Sánchez, T., Dufour, D., Moreno, J.L., Pizarro, M., Aragón, I.J., Domínguez, M. & Ceballos, H., 2013. Changes in extended shelf life of cassava roots during storage in ambient conditions. *Postharvest Biology and Technology,* 86, pp. 520-8.

Santana, M.a.A., Vásquez, V., Matehus, J. & Aldao, R.R., 2002. Linamarase expression in cassava cultivars with roots of low- and high-cyanide content. *Plant Physiology,* 129, pp. 1686-94.

Saravanan, R., Ravi, V., Stephen, R., Thajudhin, S. & George, J., 2014. Post-harvest physiological deterioration of cassava (*Manihot esculenta*) – A review. *Indian Journal of Agricultural Sciences,* 86, pp. 1383-90.

Sardari, S., Nishibe, S. & Daneshtalab, M., 2000. Coumarins, the bioactive structures with antifungal property. *Studies in Natural Products Chemistry,* 23, pp. 335-93.

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. & Zhang, P., 2011. The Biocassava Plus program: Biofortification of cassava for Sub-Saharan Africa. *Annual Review of Plant Biology,* 62, pp. 251-72.

Schadel, W.E. & Walter Jr, W.M., 1981. Localization of phenols and polyphenol oxidase in 'Jewel' sweet potatoes (*Ipomoea batatas* 'Jewel'). *Canadian Journal of Botany,* 59, pp. 1961-7.

Schilmiller, A.L., Stout, J., Weng, J.K., Humphreys, J., Ruegger, M.O. & Chapple, C., 2009. Mutations in the cinnamate 4-hydroxylase gene impact metabolism, growth and development in *Arabidopsis*. *The Plant Journal,* 60, pp. 771-82.

Schoch, G., Goepfert, S., Morant, M., Hehn, A., Meyer, D., Ullmann, P. & Werck-Reichhart, D., 2001. CYP98A3 from *Arabidopsis thaliana* is a 3′-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. *Journal of Biological Chemistry,* 276, pp. 36566-74.

Schöpke, C., Taylor, N., Cárcamo, R., Konan, N.D.K., Marmey, P., Henshaw, G.G., Beachy, R.N. & Fauquet, C., 1996. Regeneration of transgenic cassava plants (*Manihot esculenta* Crantz) from microbombarded embryogenic suspension cultures. *Nature Biotechnology,* 14, pp. 731-5.

Schubert, D., Lechtenberg, B., Forsbach, A., Gils, M., Bahadur, S. & Schmidt, R., 2004. Silencing in *Arabidopsis* T-DNA transformants: the predominant role of a gene-specific RNA sensing mechanism versus position effects. *The Plant Cell,* 16, pp. 2561-72.

Schuster, B. & Retey, J., 1994. Serine-202 is the putative precursor of the active site dehydroalanine of phenylalanine ammonia lyase. Site-directed mutagenesis studies on the enzyme from parsley (*Petroselinum crispum* L.). *FEBS Lett,* 349, pp. 252-4.

Shirley, B.W., Kubasek, W.L., Storz, G., Bruggemann, E., Koornneef, M., Ausubel, F.M. & Goodman, H.M., 1995. Analysis of *Arabidopsis* mutants deficient in flavonoid biosynthesis. *The Plant Journal,* 8, pp. 659-71.

Sibout, R., Eudes, A., Pollet, B., Goujon, T., Mila, I., Granier, F., Séguin, A., Lapierre, C. & Jouanin, L., 2003. Expression pattern of two paralogs encoding cinnamyl alcohol dehydrogenases in *Arabidopsis*. Isolation and characterization of the corresponding mutants. *Plant Physiology,* 132, pp. 848-60.

Siritunga, D. & Sayre, R.T., 2003. Generation of cyanogen-free transgenic cassava. *Planta,* 217, pp. 367-73.

Southern, E., 2006. Southern blotting. *Nature Protocols,* 1, pp. 518-25.

Stahmann, M.A., Clare, B.G. & Woodbury, W., 1966. Increased disease resistance and enzyme activity induced by ethylene and ethylene production of black rot infected sweet potato tissue. *Plant Physiology,* 41, p. 1505.

Stephenson, K., Amthor, R., Mallowa, S., Nungo, R., Maziya-Dixon, B., Gichuki, S., Mbanaso, A. & Manary, M., 2010. Consuming cassava as a staple food places children 2-5 years old at risk for inadequate protein intake, an observational study in Kenya and Nigeria. *Nutrition Journal,* 9, pp. 1-6.

Stoger, E., Williams, S., Keen, D. & Christou, P., 1998. Molecular characteristics of transgenic wheat and the effect on transgene expression. *Transgenic Research,* 7, pp. 463-71.

Stuible, H., Buttner, D., Ehlting, J., Hahlbrock, K. & Kombrink, E., 2000. Mutational analysis of 4-coumarate:CoA ligase identifies functionally important amino acids and verifies its close relationship to other adenylate-forming enzymes. *FEBS Lett,* 467, pp. 117-22.

Stupak, M., Vanderschuren, H., Gruissem, W. & Zhang, P., 2006. Biotechnological approaches to cassava protein improvement. *Trends in Food Science & Technology,* 17, pp. 634-41.

Tanaka, Y., Data, E.S., Hirose, S., Taniguchi, T. & Uritani, I., 1983. Biochemical changes in secondary metabolites in wounded and deteriorated cassava roots. *Agricultural and Biological Chemistry,* 47, pp. 693-700.

Taylor, R.G. & McInnes, R.R., 1994. Site-directed mutagenesis of conserved serines in rat histidase. Identification of serine 254 as an essential active site residue. *Journal of Biological Chemistry,* 269, pp. 27473-7.

Tohge, T., Watanabe, M., Hoefgen, R. & Fernie, A.R., 2013. The evolution of phenylpropanoid metabolism in the green lineage. *Critical Reviews in Biochemistry and Molecular Biology,* 48, pp. 123-52.

Tsai, H.H., Rodríguez-Celma, J., Lan, P., Wu, Y.C., Vélez-Bermúdez, I.C. & Schmidt, W., 2018. Scopoletin 8-hydroxylase-mediated fraxetin production is crucial for iron mobilization. *Plant Physiology,* 177, pp. 1-37.

Tumuhimbise, R., Melis, R. & Shanahan, P., 2015. Genetic variation in cassava for postharvest physiological deterioration. *Archives of Agronomy and Soil Science,* 61, pp. 1333-42.

Uarrota, V.G., Nunes, E.d.C., Peruch, L.A.M., Neubert, E.d.O., Coelho, B., Moresco, R., Domínguez, M.G., Sánchez, T., Meléndez, J.L.L., Dufour, D., Ceballos, H., Becerra Lopez-Lavalle, L.A., Hershey, C., Rocha, M. & Maraschin, M., 2015. Toward better understanding of postharvest deterioration: biochemical changes in stored cassava (*Manihot esculenta* Crantz) roots. *Food Science & Nutrition,* 4, pp. 409-22.

Valle, T., López, J.L., Hernández, J.M. & Corchete, P., 1997. Antifungal activity of scopoletin and its differential accumulation in *Ulmus pumila* and *Ulmus campestris* cell suspension cultures infected with *Ophiostoma ulmi* spores. *Plant Science,* 125, pp. 97-101.

van Oirschot, Q.E.A., O'Brien, G.M., Dufour, D., El-Sharkawy, M.A. & Mesa, E., 2000. The effect of pre-harvest pruning of cassava upon root deterioration and quality characteristics. *Journal of the Science of Food and Agriculture,* 80, pp. 1866-73.

Vanderschuren, H., Nyaboga, E., Poon, J.S., Baerenfaller, K., Grossmann, J., Hirsch-Hoffmann, M., Kirchgessner, N., Nanni, P. & Gruissem, W., 2014. Large-scale proteomics of the cassava storage root and identification of a target gene to reduce postharvest deterioration. *The Plant Cell,* 26, pp. 1913-24.

Vanholme, R., Ralph, J., Akiyama, T., Lu, F., Pazo, J.R., Kim, H., Christensen, J.H., Van Reusel, B., Storme, V. & De Rycke, R., 2010. Engineering traditional monolignols out of lignin by concomitant up-regulation of F5H1 and down-regulation of COMT in *Arabidopsis*. *The Plant Journal,* 64, pp. 885-97.

Vaucheret, H., Béclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J.B., Mourrain, P., Palauqui, J.C. & Vernhettes, S., 1998. Transgene-induced gene silencing in plants. *The Plant Journal,* 16, pp. 651-9.

Vlaar, P., van Beek, P. & Visser, R., 2007. Genetic modification and its impact on industry structure and performance: post-harvest deterioration of cassava in Thailand. *Journal on Chain and Network Science,* 7, pp. 133-42.

Vogt, T., 2010. Phenylpropanoid biosynthesis. *Molecular Plant,* 3, pp. 2-20.

Wang, K.L., Li, H. & Ecker, J.R., 2002. Ethylene biosynthesis and signaling networks. *The Plant Cell,* 14, pp. S131-S51.

Webster, G.L., 1994. Synopsis of the genera and suprageneric taxa of Euphorbiaceae. *Annals of the Missouri Botanical Garden,* 81, pp. 33-144.

Werck-Reichhart, D., Bak, S. & Paquette, S., 2002. Cytochromes P450. *The Arabidopsis Book* [Online]. Available: https://doi.org/10.1199/tab.0028 [Accessed 2018/08/25].

Wheatley, C., Orrego, A., Sanchez, T. & Granados, E., 1993. Quality evaluation of the cassava core collection at CIAT. In: R. WM & T. AM, eds. *First international scientific meeting of the Cassava Biotechnology Network*, 25-28 August 1992, Cartegena, Colombia. Cali, Colombia: Centro Internacional de Agricultura Tropical, pp. 255-64.

Wheatley, C.C. & Schwabe, W.W., 1985. Scopoletin involvement in post-harvest physiological deterioration of cassava root (*Manihot esculenta* Crantz). *Journal of Experimental Botany,* 36, pp. 783-91.

Wydra, K., Fanou, A. & Rudolph, K., 2001. Effect of cassava bacterial blight on cassava growth parameters and root yield in different ecozones and influence of the environment on symptom development. In: *7th Triennial Symposium of International Society of Tropical Root Crops, Africa Branch (ISTRC-AB)*, 2001, Cotonou, Benin. ISTRC-AB and Government of Benin, pp. 562-9.

Xu, J., Duan, X., Yang, J., Beeching, J.R. & Zhang, P., 2013. Enhanced reactive oxygen species scavenging by overproduction of superoxide dismutase and catalase delays postharvest physiological deterioration of cassava storage roots. *Plant Physiology,* 161, pp. 1517-28.

Yan, P., Shen, W., Gao, X., Li, X., Zhou, P. & Duan, J., 2012. High-throughput construction of intron-containing hairpin RNA vectors for RNAi in plants. *PLoS One,* 7, pp. 1-8.

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S. & Madden, T.L., 2012. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics,* 13, p. 134.

Young, V.R. & Pellett, P.L., 1994. Plant proteins in relation to human protein and amino acid nutrition. *The American Journal of Clinical Nutrition,* 59, pp. 1203S-12S.

Zhang, C., Han, W., Jing, X., Pu, G. & Wang, C., 2003a. Life cycle economic analysis of fuel ethanol derived from cassava in southwest China. *Renewable and Sustainable Energy Reviews,* 7, pp. 353-66.

Zhang, P., Bohl-Zenger, S., Puonti-Kaerlas, J., Potrykus, I. & Gruissem, W., 2003b. Two cassava promoters related to vascular expression and storage root formation. *Planta,* 218, pp. 192-203.

Zhang, P., Vanderschuren, H., Fütterer, J. & Gruissem, W., 2005. Resistance to cassava mosaic disease in transgenic cassava expressing antisense RNAs targeting virus replication genes. *Plant Biotechnology Journal,* 3, pp. 385-97.

Zhong, R., Taylor, J.J. & Ye, Z.-H., 1997. Disruption of interfascicular fiber differentiation in an *Arabidopsis* mutant. *The Plant Cell,* 9, pp. 2159-70.

Zidenga, T., Leyva-Guerrero, E., Moon, H., Siritunga, D. & Sayre, R., 2012. Extending cassava root shelf life via reduction of reactive oxygen species production. *Plant Physiology,* 159, pp. 1396-407.

Ziska, L.H., Runion, G.B., Tomecek, M., Prior, S.A., Torbet, H.A. & Sicher, R., 2009. An evaluation of cassava, sweet potato and field corn as potential carbohydrate sources for bioethanol production in Alabama and Maryland. *Biomass and Bioenergy,* 33, pp. 1503-8.

APPENDICES

Appendix I. Southern blot reagents (McCabe *et al.*, 1997).

20x SSC	**1 L**
300 nM Sodium citrate dehydrate	88.23 g
3 M NaCl	175.32 g
pH 7.6, autoclave and store at RT	
20% SDS	**1 L**
Sodium dodecyl sulphate	200 g
Store at RT	
Depurination solution	**1 L**
250 nM HCl (37%)	25 mL
MilliQ water	975 mL
Store at RT	
Denaturation solution	**1 L**
500 nM NaOH	20 g
1.5 M NACl	87.66 g
Neutralisation solution	**1 L**
1mM EDTA	0.37 g
500 mM Tris	60.57 g
1.5 M NaCl	87.66 g
ph 7.2, autoclave and store at RT	
W1 solution	**1 L**
0.1% SDS	5 mL
2x SSC	100 mL
MilliQ water	895 mL
Prepare fresh, store at RT	
W2 solution	**1 L**
0.1% SDS	5 mL
0.2x SSC	10 mL
MilliQ water	985 mL
Prepare fresh, store at RT	
W3 solution	**1 L**
0.1% SDS	5 mL
0.1x SSC	5 mL
MilliQ water	990 mL
Prepare fresh, store at RT	

WB solution	**1 L**
0.3% Tween20	3 mL
B1 solution	997 mL
Prepare fresh, store at RT	

B1 solution	**1 L**
150 mM NaCl	8.77 g
100 mM Maleic acid	11.6 g
pH 7.5, autoclave and store at RT	

B2 solution	**250 mL**
Blocking powder	2.5 g
B1 solution	250 mL
Prepare fresh	

B3 solution	**100 mL**
50 mM $MgCl_2$	2.5 mL
100 mM Tris-HCl (pH 9.5)	5 mL
100 mM NaCl	5.85 g
MilliQ water	92.5 mL
Prepare fresh	

Appendix II. Scores and *E*-values of cassava candidate genes using *A. thaliana* as a reference.

Cassava Candidate Genes	*Arabidopsis thaliana* Reference Genes											
PAL	**PAL1 (At2g37040)**			**PAL2 (At3g53260)**			**PAL3 (At5g04230)**			**PAL4 (At3g10340)**		
	Score	**Identity (%)**	**E-value**	**Score**	**Identity (%)**	**E-value**	**Score**	**Identity (%)**	**E-value**	**Score**	**Identity (%)**	**E-value**
Manes.08G008400.1 (cassava4.1_002628m.g)	1236.1	85.2	0.0	1259.6	86.0	0.0	1042.0	73.8	0.0	1152.1	80.3	0.0
Manes.10G047500.1 (cassava4.1_034377m.g)	1232.6	84.5	0.0	1243.0	84.3	0.0	1070.5	75.0	0.0	1180.6	81.7	0.0
Manes.09G063700.1 (cassava4.1_028434m.g)	1222.2	83.5	0.0	1242.3	84.2	0.0	1031.9	72.1	0.0	1141.0	79.1	0.0
Manes.07G098700.1 (cassava4.1_002709m.g)	1206.8	83.1	0.0	1220.7	82.9	0.0	1059.7	73.8	0.0	1177.5	81.2	0.0
Manes.04G018000.1 (cassava4.1_002591m.g)	1165.2	79.5	0.0	1179.9	80.2	0.0	1033.1	72.8	0.0	1122.8	79.4	0.0
Manes.16G098200.1 (cassava4.1_003117m.g)	870.2	64.5	0.0	886.3	63.9	0.0	818.5	59.3	0.0	874.4	64.3	0.0
C4H	**C4H (At2g30490)**											
	Score			**Identity (%)**			**E-value**					
Manes.18G126900.1 (cassava4.1_005978m.g)	827.0			85.9			0.0					
4CL	**4CL1 (At1g51680)**			**4CL2 (At3g21240)**			**4CL3 (At1g65060)**			**4CL4 (At3g21230)**		
	Score	**Identity (%)**	**E-value**	**Score**	**Identity (%)**	**E-value**	**Score**	**Identity (%)**	**E-value**	**Score**	**Identity (%)**	**E-value**
Manes.11G071800.1 (cassava4.1_005006m.g)	807.7	75.9	0.0	819.7	75.6	0.0	530.4	64.9	0.0	666.0	63.6	0.0
Manes.04G095300.1 (cassava4.1_005014m.g)	790.4	74.3	0.0	798.5	72.2	0.0	525.8	63.6	0.0	648.3	61.8	0.0
Manes.09G127000.1 (cassava4.1_004658m.g)	674.9	64.0	0.0	716.8	63.5	0.0	627.9	82.9	0.0	568.2	55.6	0.0
Manes.08G066200.1 (cassava4.1_004136m.g)	630.9	60.5	0.0	676.4	59.7	0.0	586.6	66.0	0.0	540.8	54.4	0.0
Manes.14G151400.1 (cassava4.1_027178m.g)	558.1	68.6	0.0	576.6	68.7	0.0	465.7	64.2	4.60E-160	459.5	57.0	1.50E-156
HCT	**HCT (At5g48930)**											
	Score			**Identity (%)**			**E-value**					
Manes.11G067800.1 (cassava4.1_008045m.g)	755.7			82.4			0.0					
Manes.04G101700.1 (cassava4.1_008063m.g)	750.7			81.3			0.0					
C3'H	**C3'H (At2g40890)**											
	Score			**Identity (%)**			**E-value**					
Manes.08G063400.1 (cassava4.1_005910m.g)	853.2			82.1			0.0					
CCoAOMT	**CCoAOMT (At4g34050)**											
	Score			**Identity (%)**			**E-value**					
Manes.10G078800.1 (cassava4.1_011832m.g)	453.4			86.7			7.40E-162					
Manes.07G075700.1 (cassava4.1_014783m.g)	446.4			85.5			4.60E-159					
COMT	**COMT (At5g54160)**											
	Score			**Identity (%)**			**E-value**					
Manes.01G043700.1 (cassava4.1_010187m.g)	612.5			79.1			0.0					
Manes.01G043600.1 (cassava4.1_010203m.g)	606.3			78.0			0.0					