Phytochemistry of hydroxycoumarins from *Manihot esculenta* Euphorbiaceae
(Cassava)

Lidia Kamal Jameel Alhalaseh

A thesis submitted for the degree of Doctor of Philosophy

University of Bath
Department of Pharmacy and Pharmacology

August 2017

**COPYRIGHT**

Attention is drawn to the fact that copyright of this thesis rests with its author. A copy of this thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and they must not copy it or use material from it except as permitted by law or with the consent of the author.

This thesis may be made available for consultation within the University Library and may be photocopied or lent to other libraries for the purposes of consultation. Signed on behalf of the Faculty of Science.

Signed: .................................. Date: ........................
Abstract

This is an interdisciplinary research project on cassava (*Manihot esculenta* Crantz, Euphorbiaceae) ultimately working towards producing cassava roots which are long-lasting, free of post-harvest physiological deterioration (PPD). It aims to contribute to ensuring food security. In cassava, scopoletin and its β-glycoside scopolin are considered phytoanticipants, not phytoalexins, due to their increasing accumulation during the PPD process compared to their barely detectable levels in fresh roots.

Starting with a focussed literature review on the potential of cassava, contrasted with its limitations on harvesting due to PPD, and biosynthesis along the phenylpropanoid pathway of key hydroxycoumarins, e.g. scopoletin and esculetin, the associated gaps in our current knowledge have been set out. Whether scopoletin is biosynthesized *de novo* from L-phenylalanine in response to stress, or whether stress prompts its release from the corresponding glycoside is unknown. Therefore, assessing hydroxycoumarin biosynthesis and quantifying their accumulation patterns have been undertaken in wild-type and transgenic plants in order to elucidate the divergence in scopoletin biosynthetic pathways. The identification of key genes on each pathway leading to scopoletin in cassava, and then exploring their functional identities using the model plant *A. thaliana* and genetically engineered *E. coli*, where the genes were isolated, cloned, and expressed, were also undertaken. Transgenic *A. thaliana* lines with no activity of the key enzymes on the proposed pathway, namely F6′H1, CCoAOMT, and EOMT, were developed. Competition feeding experiments using stable isotopically labelled potential biosynthetic intermediates showed the incorporation of labelled ferulate into scopoletin in transgenic A.t-F6′H1 and M.e-F6′H. This confirmed the activity of other hydroxylase enzymes rather than F6′H1 in the *ortho*-hydroxylation steps. The hydroxycoumarins of interest were isolated, characterized, and quantified in the wild type and mutant lines using chromatographic and spectroscopic techniques, mainly NMR, HR-MS, and LC-MS.

Taken together, a significant contribution to knowledge about hydroxycoumarin biosynthesis has been made.
This Thesis is dedicated to my late father
Acknowledgements

I express my deepest gratitude to my main supervisor Dr Ian S. Blagbrough for giving me the chance to explore freely during the 3-labworking years, and for his tireless efforts in proofreading my dissertation during writing-up. Also, for his guidance, persistent assistance, and support especially in family-related issues.

I also thank my co-supervisors Dr John R. Beeching and Dr Michael G. Rowan without their valuable constructive suggestions, constant support and guidance over the course of my PhD the completion of this project would not have been achievable.

Dr Shaun B. Reeksting has my gratefulness for his assistance with the LC-MS spectroscopic experiments, and for being patient and welcoming despite my fussy comments and endless repetitive analysis. I thank Dr Tim J. Woodman for his assistance with the NMR spectroscopic experiments. Thank you also to Don Perry, Andrew James, and all the technical staff in the Department of Pharmacy and Pharmacology and the Department of Biology and Biochemistry, University of Bath, for helping me to overcome many technical issues.

I gratefully thank Mu’tah University, Jordan for granting me this PhD scholarship.

Very special thanks to my colleagues in 5W3.20 and 4S1.52 for the friendly working environment and for all the useful scientific discussions we had which contributed considerably to this project. A special thanks to Dr Soad A. Bayoumi and Dr Shi Liu whose previous work got me inspired in cassava.

Finally, my deepest thanks to my parents, parents in law, sisters and brother who were always encouraging and supportive. My greatest thanks go to my husband Dr Mohannad J. Al-Zregat for showing faith in me, for his selfless love, and for his moral and financial support. Words will never be enough to express how grateful I am to my daughters Diana and Katia for having the magical tools to draw a smile on my face even in the most difficult and stressful moments I have ever experienced.
<table>
<thead>
<tr>
<th>Abbreviations</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>µg</td>
<td>microgram</td>
</tr>
<tr>
<td>4CL</td>
<td>4-coumarate CoA ligase</td>
</tr>
<tr>
<td>4-MU</td>
<td>4-methylumbelliferone</td>
</tr>
<tr>
<td>ACMV</td>
<td>African Cassava Mosaic Virus</td>
</tr>
<tr>
<td>aq.</td>
<td>aqueous</td>
</tr>
<tr>
<td>att</td>
<td>attachment site</td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>C3´H</td>
<td>p-coumaroyl shikimate/quinate 3´-hydroxylase</td>
</tr>
<tr>
<td>C4´H</td>
<td>cinnamate-4´-hydroxylase</td>
</tr>
<tr>
<td>CAMBIA</td>
<td>Centre for the Application of Molecular Biology to International Agriculture</td>
</tr>
<tr>
<td>CBM</td>
<td>cassava basic medium</td>
</tr>
<tr>
<td>CCoAOMT</td>
<td>caffeoyl CoA 3´-O-methyltransferase</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>COMT</td>
<td>caffeic acid methyltransferase</td>
</tr>
<tr>
<td>CTAB</td>
<td>cetyl trimethyl ammonium bromide</td>
</tr>
<tr>
<td>Da</td>
<td>Dalton</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethylsulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylene diamine tetra acetic acid</td>
</tr>
<tr>
<td>ESI</td>
<td>electrospray ionization</td>
</tr>
<tr>
<td>F6´H</td>
<td>feruloyl CoA 6´-hydroxylase</td>
</tr>
<tr>
<td>Fw</td>
<td>forward</td>
</tr>
<tr>
<td>GMO</td>
<td>genetically modified organisms</td>
</tr>
<tr>
<td>GST</td>
<td>glutathione-S-transferase</td>
</tr>
<tr>
<td>GT</td>
<td>glucosyltransferase</td>
</tr>
<tr>
<td>GUS</td>
<td>β-glucuronidase gene</td>
</tr>
<tr>
<td>h</td>
<td>hour</td>
</tr>
<tr>
<td>HCT</td>
<td>shikimate: quinate hydroxycinnamoyl transferase</td>
</tr>
<tr>
<td>HPLC</td>
<td>high performance liquid chromatography</td>
</tr>
<tr>
<td>HRMS</td>
<td>high resolution mass spectrometry</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>HSQC</td>
<td>heteronuclear single quantum coherence</td>
</tr>
<tr>
<td>Hz</td>
<td>Hertz</td>
</tr>
<tr>
<td>IAA</td>
<td>iso-amyl alcohol</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani medium</td>
</tr>
<tr>
<td>LB1</td>
<td>left border primer</td>
</tr>
<tr>
<td>LC-MS</td>
<td>liquid chromatography-mass spectrometry</td>
</tr>
<tr>
<td>LLD</td>
<td>lower limit of detection</td>
</tr>
<tr>
<td>LLQ</td>
<td>lower limit of quantification</td>
</tr>
<tr>
<td>LP</td>
<td>left primer</td>
</tr>
<tr>
<td>M</td>
<td>molar</td>
</tr>
<tr>
<td>m/z</td>
<td>mass over charge</td>
</tr>
<tr>
<td>MCS</td>
<td>multiple cloning sites</td>
</tr>
<tr>
<td>MES</td>
<td>2-((N-morpholino)ethane-sulfonic acid</td>
</tr>
<tr>
<td>mg</td>
<td>milligram</td>
</tr>
<tr>
<td>MHz</td>
<td>mega-Hertz</td>
</tr>
<tr>
<td>ml</td>
<td>millilitre</td>
</tr>
<tr>
<td>MS</td>
<td>mass spectrometry</td>
</tr>
<tr>
<td>NASC</td>
<td>Nottingham Arabidopsis Stock Centre</td>
</tr>
<tr>
<td>NCBI</td>
<td>the National Centre for Biotechnology Information</td>
</tr>
<tr>
<td>ng</td>
<td>nanogram</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
</tr>
<tr>
<td>ºC</td>
<td>degree Celsius</td>
</tr>
<tr>
<td>OD</td>
<td>optical density</td>
</tr>
<tr>
<td>OMT</td>
<td>O-methyl transferase</td>
</tr>
<tr>
<td>PAL</td>
<td>phenylalanine ammonia lyase</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PPD</td>
<td>post-harvest physiological deterioration</td>
</tr>
<tr>
<td>ppm</td>
<td>parts per million</td>
</tr>
<tr>
<td>PVP</td>
<td>poly vinyl pyrrolidone</td>
</tr>
<tr>
<td>RCF</td>
<td>relative centrifugal force</td>
</tr>
<tr>
<td>RE</td>
<td>restriction enzyme</td>
</tr>
<tr>
<td>Rev</td>
<td>reverse</td>
</tr>
<tr>
<td>Rf</td>
<td>retention factor</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>RH</td>
<td>relative humidity</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RNAi</td>
<td>RNA interference</td>
</tr>
<tr>
<td>RNase</td>
<td>ribonuclease</td>
</tr>
<tr>
<td>ROS</td>
<td>reactive oxygen species</td>
</tr>
<tr>
<td>RP</td>
<td>right primer</td>
</tr>
<tr>
<td>RT</td>
<td>reverse transcriptase</td>
</tr>
<tr>
<td>SAM</td>
<td>S-adenosyl-L-methionine</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SOC</td>
<td>super optimal broth with catabolite repression</td>
</tr>
<tr>
<td>SOD</td>
<td>superoxide dismutase</td>
</tr>
<tr>
<td>SPE</td>
<td>solid phase extraction</td>
</tr>
<tr>
<td>TAE</td>
<td>tris-acetate EDTA</td>
</tr>
<tr>
<td>TIGER</td>
<td>The Institute for Genome Research</td>
</tr>
<tr>
<td>TLC</td>
<td>thin layer chromatography</td>
</tr>
<tr>
<td>Tris</td>
<td>tris(hydroxymethyl)methylamine</td>
</tr>
<tr>
<td>UDP-glucose</td>
<td>uridine diphosphate-glucose</td>
</tr>
<tr>
<td>UTR</td>
<td>untranslated region</td>
</tr>
<tr>
<td>v/v</td>
<td>volume/volume</td>
</tr>
<tr>
<td>w/v</td>
<td>weight/volume</td>
</tr>
<tr>
<td>w/w</td>
<td>weight/weight</td>
</tr>
<tr>
<td>X-Gal</td>
<td>5-bromo-4-chloro-3-indoyl-β-D-galactoside</td>
</tr>
</tbody>
</table>
## Contents

Abstract 2  
Dedication 3  
Acknowledgements 4  
Abbreviations 5  

Chapter 1. General introduction and literature review 10  
  1.1. Aims and objectives 10  
  1.2. Cassava 12  
  1.3. Phenylpropanoids 21  

Chapter 2. The roles of hydroxycoumarins in *M. esculenta* and *A. thaliana* under normal and abiotic stress conditions 32  
  2.1. Introduction 32  
  2.2. Materials 34  
  2.3. General methods 36  
  2.4. Results and discussion 41  

Chapter 3. Complementation of *Arabidopsis thaliana* mutants through expressing cassava *O*-methyltransferase genes 58  
  3.1. Introduction 58  
  3.2. Materials 61  
  3.3. General methods 67  
  3.4. Results and discussion 84  

Chapter 4. Tracing scopoletin biosynthesis in wild type and transgenic *M. esculenta* and *A. thaliana* 100  
  4.1. Introduction 100  
  4.2. Materials 103  
  4.3. General methods 103  
  4.4. Results and discussion 108
Chapter 5. Functional expression of cassava O-methyltransferase enzymes in transgenic *Escherichia coli*  

5.1. Introduction  

5.2. Materials  

5.3. General methods  

5.4. Results and discussion  

Chapter 6. General conclusions  

References  

Appendices
Chapter 1. General introduction and literature review

1.1. Aims and objectives

The ultimate goal of this interdisciplinary research project on cassava (*Manihot esculenta* Crantz, Euphorbiaceae) is to produce cassava roots which are long-lasting, free of PPD signs for an extended period. This work is a contribution towards ensuring food security and it will help to decrease economic wastage of a valuable foodstuff.

The strategic aims of this project are to examine the phytochemical aspects of the biosynthetic pathways leading to scopoletin accumulation in cassava roots undergoing PPD. In order to achieve these aims, the identification of key genes on each pathway leading to scopoletin biosynthesis in cassava, and then also exploring their functional identities using the model plant *A. thaliana* and the genetically engineered *E. coli* will be undertaken.

The molecular mechanisms behind PPD are poorly understood, and the divergent but connected pathways to hydroxycoumarins are not fully elucidated in cassava. While the accumulation of scopoletin and its β-glycoside scopolin are confirmed during PPD, their biosynthesis still needs further investigation. It is currently unknown and merits clarification whether scopoletin is biosynthesized *de novo* from L-phenylalanine in response to stress, or whether stress prompts its release from the corresponding glycoside. Thus, assessing hydroxycoumarin biosynthesis and quantifying the accumulation patterns, particularly in cassava under PPD and generally in plants under abiotic stress, are necessary to confirm the roles of hydroxycoumarins in stress responses (chapter 2). To elucidate the different pathways leading to scopoletin biosynthesis, transgenic *A. thaliana* lines with no activity of the key enzymes on the proposed pathway, namely F6’H1, CCoAOMT, and EOMT, will be developed. The hydroxycoumarins of interest will be isolated, characterized, and quantified in the wild type and the mutant lines using modern chromatographic and spectroscopic techniques (mainly NMR, HR-MS, and LC-MS).

Even though the cassava genome has been investigated and published (Wang *et al.*, 2014), most of the genes involved in the alternative hypothetical pathways are not yet identified. The identities of selected enzymes in scopoletin biosynthesis will be characterized using bioinformatics tools. Their functions and
relative importance will be explored by investigating their ability to complement the mutation in *A. thaliana* plants through restoring the deficiency in scopoletin level (chapter 3).

The biosynthetic flux from the early precursor L-phenylalanine, through hydroxycinnamate to the production of scopoletin and its β-glycoside, scopolin, will be investigated in wild type and transgenic plants using different commercially available and synthesized (in house) stable isotopically labelled intermediates. This will allow the elucidation of the divergence in scopoletin biosynthetic pathways. The relative importance of each pathway as represented by key intermediates and key enzymes along each pathway will be assessed (chapter 4).

Another objective for this research project is to assess the biochemical profiles in a transgenic cassava plant recently produced by Dr Liu in the Beeching laboratories, University of Bath, by quantifying the hydroxycoumarin levels and exploring the origin of the biosynthesized scopoletin (chapter 4). This later objective will also be achieved by designing and carrying out competition feeding experiments with stable isotopically labelled potential intermediates.

In related microbiological experiments, the genes of interest will be isolated, cloned, and expressed in transgenic *E. coli*. These genetically modified organisms (GMO) will be used to assess the activities of enzymes, translated from selected genes, towards several substrates on the phenylpropanoid pathways in order to investigate their specificities (chapter 5).

Taken together, these five objectives will enable a significant contribution to be made to knowledge about hydroxycoumarin biosynthesis. In particular, genetically engineered *E. coli*, gene knock outs in the model plant *A. thaliana*, transgenic cassava lacking F6’H enzyme, and of course, wild type cassava (*M. esculenta* Crantz) will be the chosen organisms and species for this laboratory-based research project.

A focused literature review on the potential of cassava, contrasted with the limitations on harvesting cassava, has been used to set the scene. The biosynthesis, along the phenylpropanoid pathway, of key hydroxycoumarins, e.g. scopoletin and esculetin, and the associated gaps in our current knowledge have been set out. Thus, having established the clear potential for achieving these important aims and objectives, the roles of the hydroxycoumarins in selected plants under normal and abiotic stress conditions will be investigated.
1.2. Cassava

Cassava (Manihot esculenta Crantz Euphorbiaceae) is an important starch rich crop which plays a major role in food security in many African developing countries. It is also an income-generating crop for resource-limited small holder farmers, and a renewable low-input biofuel source in many developed countries including Brazil, Thailand, and China (Ziska et al., 2009; El-Sharkawy 2012; Zidenga et al., 2012). Worldwide, cassava is the sixth most important crop with respect to starch production after wheat Triticum aestivum, rice Oryza sativa, maize Zea mays, potato Solanum tuberosum, and barley Hordeum vulgare (Lebot, 2008).

Cassava (Figure 1.1) is a perennial shrub (up to 1-5 m tall). It is grown mainly for its edible starchy roots (30-60% dry matter), and could be used to help in ensuring food security for the increasing global population (>120% is expected in the next 30-35 years in sub-Saharan Africa) especially after complicated climate changes and diminished agricultural land issues (De Souza et al., 2017).

Fig. 1.1. Cassava shrubs and its edible roots.
http://sattvamji.blogspot.co.uk/2014/10/78-cassava-or-maniok-root-one-more.html
(accessed on 01.08.2017).

Cassava is gaining an increased attention due to its ability to grow and efficiently produce starch (more than 163 x 10^6 ton of cassava starch annually) under suboptimal conditions of poor infertile soil and low rainfall (Teerawanichpan et al., 2008). Other remarkable characteristics such as high paste clarity, recurrent freezing-thawing stability, and lack of odour, in addition to the low input need,
make it an important industrial precursor for different goods including paper, textiles, pharmaceuticals and bio-products (Teerawanichpan et al., 2008).

**Constraints and limitations**

Despite its importance, commercialization of cassava production in Africa is challenged by different constraints. These include pathogens such as African Cassava Mosaic Virus (ACMV), cassava bacterial blight *Xanthomonas axonopodis*, cassava mealy bug, and cassava brown streak disease (Hillocks and Jennings, 2003; Sayre et al., 2011; Bull et al., 2011). Vegetative propagation by stem cuttings, not by seeds, is an artificially selected propagation method and is favoured by farmers because cassava possesses knobby thickened stems so 5-10 cuttings can be obtained from a single plant, and this facilitates cassava planting. The free exchanging of stems between farmers for planting cassava, facilitates the spread of pathogens, and even the insects hosting harmful viruses. This contributes in magnifying those devastating plant diseases (Bull et al., 2011; Moyib et al., 2015; Chavarriaga-Aguirre et al., 2016).

High content of cyanogenic glycosides, such as linamarin and lotaustralin, (up to 1.5 mg/kg) is also among the major nutritive drawbacks of cassava. Consumption of raw or inefficiently processed cassava roots is associated with an irreversible paralytic disease called konzo, affecting mostly children and women (Onabolu et al., 2000; Ngudi et al., 2002). Konzo afflicts African populations where a monotonous cassava diet is taken (Ngudi et al., 2002). This is because cassava is a protein-poor crop. It lacks the essential sulfur-containing amino acids to detoxify the cyanide released from the cyanogenic glycosides. However, proper processing and preparation of the cassava roots significantly reduces or even eliminates this toxicity (Onabolu et al., 2001; Mkumbira et al., 2003). Cossettes is one of the most popular cassava products in Nigeria, prepared by domestics by softening the bitter roots (with their higher cyanogenic glycoside contents) by soaking them in water for ~3 days, when they are peeled and then sundried (2-5 days), in order to obtain a cyanide-free cassava product with extended storage time (Jorgensen et al., 2011; Nyirendaa et al., 2011; Mlingi et al., 2011).

Generally in plants containing cyanogenic glycosides, including cassava, bitter almonds, and rubber, the cyanogenic glycosides, in particular linamarin and lotaustralin, are biosynthesised from the amino acids valine and isoleucine. These
precursors are oxidized to form the oxime-intermediates valox and ilox respectively via CYP79 enzymes. The oximes are metabolized to produce the main cyanohydrin intermediates, acetone cyanohydrin and 2-hydroxy-2-methylbutyronitrile from the precursors valine and isoleucine respectively. The cyanohydrin is glycosylated before the final storage in the plant vacuoles by the activity of a glucosyltransferase enzyme. After insect chewing or upon wounding by harvest, linamarase enzyme mediates the hydrolysis of the stored cyanogens linamarin and lotaustralin to produce glucose and the corresponding intermediates, acetone-cyanohydrin and 2-hydroxy-2-methylbutyronitrile respectively. Then the intermediates are further metabolised to produce hydrogen cyanide (HCN) and acetone or butanone (Figure 1.2). The last step could occur spontaneously in an acidic medium (pH <5) or at high temperature (>35 ºC), or it could be enzymatically mediated via hydroxynitrile lyase (HNL) in leaves. Hydrogen cyanide will not be released in the intact cells, as the glycosides linamarin and lotaustralin are stored in the vacuoles and the deglucosidase enzymes, e.g. linamarase, are present in the cell wall (Hughes et al., 1994).

Fig. 1.2. Cyanogenic glycosides biosynthesis and metabolism in cassava.

Linamarin and lotaustralin are biosynthesized from the amino acids L-valine and L-isoleucine respectively. Hydrolysis of the stable cyanogenic glycosides will release the cyanohydrin which collapses to the corresponding ketone with the release of HCN.

Detoxification of the hydrogen cyanide can be mediated by the enzyme rhodanese to produce the soluble thiocyanate (SCN⁻) after reaction with the sulfur containing-essential amino acids such as L-cysteine or L-methionine, or via
synthase enzyme to produce the corresponding proteins from the amino acids, e.g. with cysteine to produce cyanoalanine, then asparagine then protein (Figure 1.3).

In cases of protein insufficiency, where the sulfur-containing amino acids are low, cyanide will be converted into cyanate (OCN⁻) which cause the neurodegenerative disease (Withthon and Naumann, 1987; Tor-Agbidye et al., 1999; Jorgensen et al., 2011; Zidenga et al., 2017).

Fig. 1.3. Detoxification of hydrogen cyanide in the presence of sulfur-containing amino acids. HCN could be assimilated into protein synthesis via the cyanoalanine synthase enzyme, or alternatively detoxified into the soluble form thiocyanate via rhodanese enzyme.

Post-harvest physiological deterioration (PPD)

A major unresolved limitation in cassava production and commercialization is post-harvest physiological deterioration (PPD). PPD is a biological result of the mechanical damage during harvesting the roots. It progresses from the proximal site of the wound to the distal end, encouraging the symptoms of vascular streaking and blue-black discolouration to appear only 2-3 days after harvesting. The discolouration arises from the occlusions in the vascular parenchyma perhaps by the oxidation of some phenolic secondary metabolites, in particular hydroxycoumarins (Beeching et al., 1993; Han et al., 2001; Zidenga et al., 2012; Xu et al., 2013; Urrolota et al., 2016). Figure 1.4 shows a trans-section of fresh cassava root with clear parenchyma and the blackish discolouration in the parenchyma of the deteriorated root.
Fig. 1.4. A cross-section of a cassava root. A shows the main anatomical parts of the storage root. B shows the vascular streaking of the cassava tuber under PPD (after Yao et al., 2014).

These visual symptoms of PPD shorten the shelf-life of the harvested storage roots, and render them unpalatable and unmarketable. This therefore limits their economic and industrial potential (Sanchez et al., 2006; Bayoumi et al., 2008a). While different environmental factors affect the deterioration process, such as temperature, humidity and oxygen, locals tried several attempts to delay the symptoms of PPD through manipulating these conditions, e.g. by freezing the fresh harvested roots or dipping them into hot water, or by excluding the oxygen by paraffin-waxing, or storing the roots in polyethylene bags immediately after harvest. Harvesting on market demand also helps in reducing the loss, although delaying harvest time will harden the storage roots. Rapid processing of the edible roots, e.g. into cassava flour, increases the storage life and facilitate trans-nation transport (Reilly et al., 2003; Moyib et al., 2015). Delaying PPD symptoms is achievable, although not preferable, by cutting the higher stems and leaves of the growing cassava plants 10-30 days before harvesting the roots, in other words pruning. Pruning is highly efficient in delaying the visual symptoms of PPD up to 20 days after harvest. However, pruning has the significant drawback causing a decrease in the root-weight due to a decrease in the desirable starch content, which is presumably metabolized into sugars rendering the roots with a reduced organoleptic characteristic (van Oirschot et al., 2000; Reilly et al., 2003; Iyer et al., 2010).

Conventional breeding to reduce PPD has been challenged by limited flowering, high heterozygosity, slow propagation, and positive correlation between
PPD susceptibility and the desirable dry matter content (Jennings and Iglesias, 2002; Ceballos et al., 2004; Morante et al., 2010). However, breeding has been used successfully to improve ACMV resistance (Legg and Fauquet, 2004), but has not applied to improve cassava PPD response.

Cassava roots deteriorate in two separate processes. An initial physiological deterioration (PPD), followed by a secondary microbiological process (Buschmann et al., 2000b). PPD considered to be an abiotic stress response, hence the observed morphological changes in cassava are comparable to the wound responses in other plants, e.g. potato Solanum tuberosum (Lulai et al., 2008). Even though these responses could have an efficient healing ability to the wounded but still attached cassava roots, but it is impaired, too late, or inadequate in the harvested roots. This is because the roots have lost their efficient wound repair mechanisms during evolution, and/or because the roots has no biological function after harvest (Cortes et al., 2002; Reilly et al., 2003).

Reactive oxygen species (ROS) overproduction is reported to be the earliest response to root wounding by harvesting (Buschmann et al., 2000a; Reilly et al., 2003). Generally, in plants, ROS is a by-product of the natural plant aerobic respiration and several scavengers are normally produced to counteract ROS production and prevent its toxicity. ROS activity is modulated by catalase, peroxidase, and superoxidase dismutase (SOD) enzymes (Buschmann et al., 2000a, Reilly et al., 2003, Iyer et al., 2010). An oxidative burst occurs when the equilibrium between ROS production and scavenging capability is disturbed (Apostol et al., 1989; Apel and Hirt, 2004; Zidenga et al., 2012).

PPD is an active endogenous oxidative process involving changes at the molecular level, including altering gene expression and protein synthesis. Changes in the activities of the oxidative enzymes resulted in the accumulation of phenol and polyphenol antioxidants in cassava roots. These include flavonoids and anthocyanin derivatives such as catechine, leucoanthocyanidins, delphinidin, and cyanidin, which could polymerize to form tannins (Akinrele, 1964; Reilly et al., 2007; Zidenga 2012; Sanchez et al., 2013). Cassava stems are also found to accumulate phenolic antioxidants during PPD, including isovanillin, syringaldehyde, coniferaldehyde, ficusol, ethamivan, and $p$-coumaric acid (Yi et al., 2011) (Figure 1.5).
Fig. 1.5. Phenols and polyphenols isolated from cassava roots and stems during PPD. Flavonois and anthocyanin derivatives, phenols and phenylpropanoid antioxidants isolated from cassava roots and stems during PPD.

Fig. 1.6. The hydroxycoumarins involved in PPD. Scopoletin, esculetin and their respective β-glycosides, scopolin and esculin.
PPD is still an unresolved issue, although it is known that the undesirable blue-black discolouration of the harvested roots linked to the oxidation of phenolics in particular, hydroxycoumarins. Hydroxycoumarins are available in abundance in the deteriorated cassava roots compared to their levels in the fresh cassava roots (Tanaka et al., 1983; Buschmann et al., 2000b; Bayoumi et al., 2008a). Scopoletin, esculetin, and their corresponding β-glycosides scopolin and esculin (Figure 1.6) are the hydroxycoumarins evidently involved in the PPD process because their concentrations in the deteriorated cassava roots dramatically increased compared to those in the fresh roots (Buschmann et al., 2000a; Bayoumi et al., 2008a).

Fig. 1.7. Possible phenoxide radicals produced by one electron oxidation of scopoletin. Scopoletin radicals could neutralize the toxic ROS intermediates into different oxidized products (after Leite et al., 2015).

Scopoletin, along with other hydroxycoumarins in cassava, displays free-radical scavenging properties toward the ROS intermediates including the
superoxide ion (\(\cdot\)O\(_2\)), the hydroxyl radical (HO\(^*\)), and hydrogen peroxide (H\(_2\)O\(_2\))

which could be produced as a consequence of excessive reduction of O\(_2\) to H\(_2\)O as a

result of the electron transport chain leakage (Schutzendubel and Polle, 2002; Xu et al., 2013).

None of the oxidation products of scopoletin has yet been isolated from
deteriorated cassava roots, but they are considered to be responsible for the blue-
blackish discolouration. This is supported by in vitro experimental results of
scopoletin oxidation mediated by horseradish peroxidase enzyme, where a blue-
coloured product was produced (Miller et al., 1975). The possible free-radical
scavenging process and the oxidation products of scopoletin are shown in Figure
1.7.

**Biotechnological attempts to produce better cassava**

Improvement of cassava yield and concomitant retail quality are critical to
meet the big demand for starch-rich food in the African countries. The failure of
the conventional methods to alleviate PPD syndromes and to extend cassava shelf
life, its poor nutritional values, and the microbiological challenges, encourage
research for potential biotechnological solutions in order to overcome these
challenges and to improve cassava.

The BioCassava Plus (BC +) programme was established in 2005 aiming to
improve the nutritional values of cassava to obtain genetically modified biofortified
cassava rich in the essential food elements, e.g. zinc, iron, protein, and vitamin A,
and to reduce the cyanogen content, increase pathogen resistance, and to increase
the shelf-life of the roots after harvesting (Sayre et al., 2011).

The ROS accumulation after harvest is naturally attenuated by plants by
overproduction of antioxidants such as flavonoids and coumarins. Also cassava
cultivars with naturally higher carotene levels, which has antioxidant properties,
show better resistance to PPD (Sanchez et al., 2006). Enzymatic detoxification
mechanisms are included by the action of superoxide dismutase (SOD) which
converts the superoxide (\(\cdot\)O\(_2\)) into hydrogen peroxide H\(_2\)O\(_2\), and the catalase
enzymes which detoxify the resultant H\(_2\)O\(_2\) into water (Apel and Hert, 2004;
Saravanan et al., 2016). A delay in PPD symptoms up to 10 days was achieved by
overproduction of both the Cu/Zn superoxide dismutase and catalase enzymes in
transgenic cassava by enhancing the detoxification of the ROS molecules, and
therefore tolerance to oxidative stress (Xu et al., 2013).
Delaying the appearance of PPD symptoms up to 21 days was achieved by reducing the accumulation of ROS in cassava roots after harvest by overexpressing the *A. thaliana* oxidase gene, cyanide-insensitive alternative oxidase (AOX1) in transgenic cassava. However, the delay in PPD symptoms was accompanied by a decrease in biomass-yield of the field-grown transgenic plants which is mostly due to the heterozygosity of the expressed gene (Zidenga *et al.*, 2012).

In cassava, the oxidative burst occurs soon after harvesting (~15 min), and the ROS burst could be triggered by the release of the cyanide after linamarin hydrolysis (Zidenga *et al.*, 2017). Reduced cyanogenic accumulation delayed the PPD symptoms in the produced transgenic cassava by Sayre and co-workers up to 21 days, but the transgenic cassava was not able to grow without ammonia to compensate for the role of linamarin in transporting reduced nitrogen from leaves to roots (Siritunga and Sayre, 2004; Zidenga *et al.*, 2017).

### 1.3. Phenylpropanoids

The plant phenylpropanoids originate from the shikimate pathway, considered as the backbone of the plant secondary metabolites involved in diverse plant functions including cell wall strengthening as in lignin, plant defence such as coumarins, and flavonoids (Vogt, 2010). Starting from phosphoenolpyruvate (PEP) and erythrose-4-phosphate, the plant produces shikimic acid, the direct precursor of the aromatic amino acids including L-phenylalanine. Phenylalanine ammonia lyase (PAL) catalyses the deamination of the amino acid to produce cinnamates, then *p*-coumarates, the direct precursors of several plant secondary metabolites (Figure 1.8).
Scopoletin biosynthesis in cassava

The increase in hydroxycoumarin accumulation, namely scopoletin, esculetin and their respective β-glycosides scopolin and esculin, is concomitant with the increase of the expression of specific genes along the phenylpropanoid pathways leading to scopoletin biosynthesis, e.g. phenylalanine ammonia lyase (PAL), the enzyme which catalyses the amino acid deamination into the cinnamates, the early precursor the hydroxycoumarins, is upregulated in the cassava roots after harvest (Wheatley and Schwabe, 1985; Reilly et al., 2003). In the last decade, several attempts have been carried out to understand and investigate the hydroxycoumarins, in particular scopoletin, biosynthesis along the phenylpropanoid pathway.

In cassava roots, Bayoumi et al. defined the following pathway to be major in scopoletin biosynthesis, on the basis of stable isotope labelling experiments,
starting from deamination of the amino acid L-phenylalanine to produce E-cinnamate, then para-hydroxylation to produce p-coumarate, meta-hydroxylation to E-caffeate, then 3’-O-methylation to produce E-ferulate, the ferulate is further ortho-hydroxylated to produce E-6’-hydroxyferulate, followed by E-Z-isomerization, then lactonization steps to produce scopoletin (Figure 1.9) (Bayoumi et al., 2008a; Bayoumi et al., 2008b).

**Fig. 1.9. Scopoletin biosynthetic pathway in cassava roots.** Scopoletin is biosynthesised via the phenylpropanoid pathway after hydroxylation, O-methylation, isomerization, and lactonization steps

**Scopoletin biosynthesis in A. thaliana**

In the model plant *Arabidopsis thaliana*, scopoletin biosynthesis was investigated at the molecular level and different intermediates were found to mediate the downfield carbon-flux from phenylalanine into scopoletin. The major enzymes involved were also investigated and defined (Kai et al., 2006; Bourgaud et al., 2006; Kai et al., 2008; Matsumoto et al., 2012) (Figure 1.10). Deamination of
phenylalanine into cinnamates is mediated by phenylalanine ammonia lyase (PAL), the first hydroxylation of cinnamate into \( p \)-coumarate is catalysed by cinnamates 4´-hydroxylase (C4´H) enzyme a CYP 450 monooxygenase enzyme involved also in both lignin and flavonoid biosynthesis, downregulation the C4´H activity in \( A. \) thaliana resulted in severe phenotyping including male sterility and dwarfism (Schilmiller et al., 2009). 4-Coumarate: CoA ligase (4CL) catalyse the formation of the CoA thioester intermediate \( p \)-coumaroyl CoA, which represents the early precursor not only for hydroxycoumarins, but also for a variety of secondary metabolites including flavonoids, isoflavonoids, stilbenes, aurones, lignin, and coumarins (Figure 1.8) (Vogt, 2010).

In the hydroxycoumarin biosynthetic pathway, the CoA group on \( p \)-coumaroyl CoA is transferred into \( p \)-shikimate to produce \( p \)-coumaroyl shikimate, and to a lesser extent \( p \)-coumaroyl quinate by the action of hydroxycinnamoyl CoA shikimate: quinate hydroxycinnamoyl transferase (HCT) (Fraser and Chapple, 2011). The second hydroxylation is mediated by \( p \)-coumaroyl shikimate 3´-hydroxylase (C3´H) to produce caffeoyl shikimate, and to less extent caffeoyl quinate. The affinity toward producing the shikimate ester is four times higher than the corresponding quinate ester (Schoch et al., 2001).

The shikimate/quinate intermediates were found to be essential in scopoletin biosynthesis after \textit{in planta} experimental results where knocking-out C3´H in \( A. \) thaliana significantly reduced scopoletin accumulation in the roots compared to its level in the wild type (~97%), but also dwarf phenotype plants were produced as the C3´H involves also in lignin formation (Kai et al., 2006). The methyl transferase reaction is catalysed by caffeoyl CoA 3´-O-methyltransferase (CCoAOMT) enzyme to add a methyl group on the 3-hydroxyl- of the caffeoyl CoA to produce feruloyl CoA. The third and last hydroxylation step in scopoletin biosynthetic pathway is mediated by feruloyl CoA 6´-hydroxylase (F6´H) to produce \( E \)-6´-hydroxyferuloyl CoA which isomerizes into the \( Z \)-isomer, and lactonizes to produce scopoletin. F6´H activity is essential in scopoletin biosynthesis as eliminating its activity significantly reduced (~97%) the accumulation of scopoletin in \( A. \) thaliana mutant (Kai et al., 2008).

Umbelliferone is a simple \( p \)-coumaroyl CoA derivative biosynthesised after hydroxylation by the \( p \)-coumaroyl CoA 2´ hydroxylase activity (C2´H), \( E \)-Z-
isomerization, and then lactonization steps (Figure 1.10) (Matsumoto et al., 2012; Vialart et al., 2012).

**Fig. 1.10. Hydroxycoumarins biosynthesis pathways in A. thaliana.**

Enzymes: phenylalanine lyase (PAL), cinnamates 4'-hydroxylase (C4’H), 4-coumarate: CoA ligase (4CL), hydroxycinnamoyl CoA shikimate: quinate hydroxycinnamoyl transferase (HCT), p-coumaroyl CoA 2’ hydroxylase (C2’H), p-coumaroyl shikimate 3’ hydroxylase (C3’H), caffeoyl CoA 3’-O-methyltransferase (CCoAOMT), and feruloyl CoA 6’-hydroxylase (F6’H) (after Kai et al., 2006; Bourgaud et al., 2006).

Umbelliferone serves as a pivotal precursor in the synthesis of different coumarin derivatives including furanocoumarins and pyranocoumarins. Furanocoumarins are potent phytoalexins and classified according to the furan ring attachment into linear (ring attached into C6 and C7) and angular (ring attached into C7 and C8) while pyranocoumarins have no defined functions in the plant, but
they possibly have phytoalexin activity as they bear a structural resemblance to furanocoumarins. The linear furanocoumarins: psoralen, xanthotoxol, and bergaptol, and the angular furanocoumarins angelicin, sphenandin and pimpinellin are known to be potent phytoalexins (Figure 1.11) (Bourgaud et al., 2006; Lin et al., 2013).

Fig. 1.11. Furanocoumarins. Psoralen, xanthotoxol, bergaptol, sphondin, and pimpinellin are potent furanocoumarin phytoalexins biosynthesised from umbelliferone (after Bourgaud et al., 2006).

Artificial biosynthetic pathways for hydroxycoumarins

The exact entire hydroxycoumarin biosynthetic pathways are still under debate as there is comparatively little information known, and some is ambiguous or even conflicting. An artificial pathway for scopoletin biosynthesis was recently assembled in genetically engineered E. coli to explore the hydroxycoumarins, in particular scopoletin, biosynthesis and the key enzymes involved. In bacteria, tyrosine, rather than phenylalanine, is the entrance precursor for both umbelliferone and scopoletin biosynthesis. Genetically engineered E. coli cells with the 4CL, and
C2′H isolated from *A. thaliana* were able to produce umbelliferone (4 mg/l) from the host’s tyrosine and CoA molecules. Tyrosine which is naturally biosynthesized in the bacterial cells, is deaminated by the native tyrosine ammonia lyase (TAL) enzyme to produce *p*-coumaric acid. By the exogenous 4CL, C2′H activities, the intermediates *p*-coumaroyl CoA and 2′-hydroxycoumaroyl CoA were produced respectively. The latter is isomerised and then lactonized to produce umbelliferone (Lin *et al*., 2013) (Figure 1.12).

**Fig. 1.12. An artificial biosynthetic pathway of umbelliferone in genetically engineered *E. coli*.** *p*-Cumaric acid is produced naturally in the bacterial cells after deamination of tyrosine by tyrosine ammonia lyase (TAL), adding the native CoA was mediated by the exogenous 4CL to produce *p*-coumaroyl CoA, hydroxylation on C2′ is mediated by the exogenous C2′H, then E-Z-isomerization, and lactonization steps to produce umbelliferone (after Lin *et al*., 2013). Bacterial native enzymes and molecules are in blue, plant exogenous plant enzymes are in red.

In an advanced artificial bioconversion of the phenylpropanoid acids into scopolentin, *A. thaliana* genes proposed to be involved in scopolentin biosynthesis were functionally expressed in genetically modified *E. coli* cells. The production of scopolentin (27 mg/l) was highly dependent on the introduced plant enzymes in two connected pathways, the interaction between the bacterial endogenous enzymes and metabolites allowed the use of the bacterial tyrosine as an early precursor for scopolentin biosynthesis (Lin *et al*., 2013).

Tyrosine could be either deaminated by TAL into *p*-coumaric acid or hydroxylated on the C-3′ to produce L-DOPA by the bacterial 4HPA3H, deamination of L-DOPA into caffeic acid is also mediated by TAL, and alternatively, caffeic acid could be produced as a hydroxylated product of *p*-coumaric acid by 4HPA3H. The intermediates *p*-coumaric acid, L-DOPA, and
caffeic acid are naturally occurring in *E. coli*. By the action of the exogenous plant enzymes, both *p*-coumaric acid and caffeic acid intermediates could be transferred into their CoA ester forms by the 4CL activity using the host CoA molecules. Caffeoyl CoA could also be produced indirectly from the *p*-coumaroyl CoA after producing the shikimate/quinate intermediate by HCT, then adding a hydroxyl-group onto C3' by C3'H activity. The produced caffeoyl CoA ester is O-methylated by CCoAOMT into feruloyl CoA, ortho-hydroxylated into 6'-hydroxyferulyol CoA, undergoes E-Z-isomerization, and finally lactonization to produce scopoletin (Figure 1.13).

![Artificial biosynthetic pathway of scopoletin in genetically engineered *E. coli*](image)

**Fig. 1.13. An artificial biosynthetic pathway of scopoletin in genetically engineered *E. coli*.** *p*-Coumaric acid, L-DOPA, and caffeic acid are natural metabolites from tyrosine biosynthesized by TAL and 4HPA3H. Scopoletin was synthesized by introducing the plant enzymes: 4CL, HCT, C3'H, CCoAOMT, and F6'H. Bacterial native enzymes and molecules are in blue, plant exogenous plant enzymes are in red (after Lin et al., 2013).

In a separate study, the hydroxycoumarins umbelliferone, esculetin, and scopoletin were artificially synthesized in genetically engineered *E. coli* with 4CL and F6'H enzymes isolated from *A. thaliana*, from the added precursors *p*-
cumaric acid, caffeic acid, and ferulic acid respectively (Yang et al., 2015). The study confirms the importance of producing the thioester intermediates and the non-selective activity of the A. thaliana F6′H towards p-coumaroyl CoA, caffeoyl CoA, in addition to its known activity toward feruloyl CoA (Kai et al., 2008) (Figure 1.14).

![Artificial biosynthetic pathways of simple hydroxycoumarins in genetically engineered E. coli](image)

**Fig. 1.14. Artificial biosynthetic pathways of simple hydroxycoumarins in genetically engineered E. coli.** Umbelliferone, esculetin, and scopoletin are synthesized from their corresponding phenylpropanoid acids p-coumaric acid, caffeic acid, and ferulic acid respectively. The plant enzymes are in red and the endogenous molecules are in blue.

The interconnected pathways and the non-selectivity of the involved enzymes open the possibilities of the occurrence of other branched pathways in the hydroxycoumarins, in particular scopoletin, biosynthesis perhaps, by different arrangement of the hydroxylation and methylation steps.
Hydroxycoumarins

The name coumarins is derived after *coumarou*, the French term for the Tonka bean *Coumarouma odorata* from where coumarin was first isolated in 1820 (Borges *et al*., 2005). Hydroxycoumarins, the plant-derived bioactive products, have attracted research interest to explore their diverse and promising biological activities in the plant and human systems. Coumarin derivatives are also used in industry as aroma enhancers in tobacco and some alcoholic beverages (Matos *et al*., 2015).

In plants, hydroxycoumarins are natural phytoalexins and protective compounds against several abiotic and biotic threats. Their application in treating plant pathogens is promising, e.g. their antibacterial activity towards the *Ralstonia solanacearum*, has recently been established. The complex pathogenicity of *R. solanacearum* towards several essential crops including potato, tomato, and eggplant has been reduced by simple coumarins umbelliferone, esculetin, xanthotol, and daphnetin (Figure 1.15) possibly by cell-membrane destruction and biofilm formation inhibition mechanisms (Ribera and Zuniga, 2012; Yang *et al*., 2016). Scopoletin also shows an antifungal activity towards the tobacco pathotype *Alternariaalternate* (Sun *et al*., 2014).

Their low cytotoxicity on human cells, high availability, and structural simplicity are behind their known and potential diverse pharmacological and therapeutic activities of coumarins (Kawase *et al*., 2003; Patil *et al*., 2013). Warfarin, the vitamin K antagonist, is the most notable coumarin derivative used in current therapeutics for its anticoagulant activity (Daly, 2013). Promising antipsychotic and anti-neurodegenerative disease therapies from some natural plant-derived coumarins follow their *in vitro* and *in vivo* activity in inhibiting monoamine oxidase (MAO) enzyme. Psoralen, bergapten, and xanthotoxin isolated from the roots of *Peucedanum japonicum*, and lacinartin isolated from the stems of *Zanthoxylum schinifolium*, are examples of naturally occurring coumarins with MAO inhibitor activity (Figure 1.15) (Huong *et al*., 1999; Lee *et al*., 2000; Jo *et al*., 2002; Patil *et al*., 2013). Umbelliferone, isolated from *Potentilla evestita*, showed *in vivo* anti-nociceptive and anti-inflammatory activities (Rauf *et al*., 2014). Scopoletin showed *in vivo*, a potential therapeutic activity to the gastro-oesophageal inflammation probably through enhancing mucosal sections (Mahattanadul *et al*., 2011). Naturally occurring scopoletin from *Gelsemium*
*sempervirens* showed a potential cytotoxic activity against human melanoma cells A375 (Khuda-Bukhsh *et al.*, 2010). Moreover, scopoletin exhibits antifungal and hepatoprotective activities (Carpinella *et al.*, 2005; Noh *et al.*, 2011). These studies exemplify the importance of hydroxycoumarins and in particular of scopoletin, esculetin and their structurally-related compounds.

**Fig. 1.15. Coumarins with biological activities.**
Chapter 2. The roles of hydroxycoumarins in *M. esculenta* and *A. thaliana* under normal and abiotic stress conditions

### 2.1. Introduction

Plants are sessile organisms, therefore, they have evolved mechanisms to adapt to changing environmental challenges and avoid unfavourable conditions of stress which could be biotic such as pathogens or herbivores, or abiotic such as soil salinity, drought or lack of nutrients. Plant responses involve changes at the cellular and molecular levels and accumulation of secondary metabolites (Rocco *et al.*, 2013; Sham *et al.*, 2015). Different polyphenols, including coumarins, were biosynthesized *de novo* or released from the storage vacuoles after plants were triggered by stress. Coumarins could be phytoalexins, synthesized and accumulated in plants and then released upon exposure to unfavourable conditions (Taguchi *et al.*, 2000; Gomez-Vasquez *et al.*, 2004; Gnonlonfin *et al.*, 2011) or phytoanticipants, biosynthesized *de novo* after being triggered with different types of stresses (Ribera and Zuniga, 2012). In this chapter, the involvement of the hydroxycoumarins (scopoletin, esculetin and their respective β-glycosides) in stress responses were investigated in cassava and *A. thaliana* under normal and abiotic stress conditions.

Cassava roots undergo different changes at the cellular and molecular levels mediated by several regulatory genes in response to wounding and harvesting (Reilly *et al.*, 2003). These wound responses confirm that the post-harvest physiological deterioration (PPD) symptoms observed in cassava roots are responses to abiotic stress (Han *et al.*, 2001). Plants, in general, respond to wounding by producing different signalling molecules, e.g. ethylene, hydrogen peroxide, polyphenols, salicylic acid, as part of their wound repair and defence mechanism (Buschmann *et al.*, 2000b; Han *et al.*, 2001). While cassava roots are not propagules, the incomplete wound response after harvesting gives rise to the symptoms of PPD.

Different plants show different responses to unfavourable changes in their environments and interacting with abiotic stresses may affect their biotic resistance in different ways, e.g. *Solanum lycopersicum* (tomato) and *Hordeum vulgare* (barley) show increased resistance to pathogenic bacteria after acquiring increase tolerance to drought, salinity, and osmolality abiotic stresses (Wiese *et al.*, 2004;
Achuo et al., 2006). This contrasts the A. thaliana behaviour which showed higher susceptibility to pathogens after drought stress (Mohr and Cahill, 2003).

Being a model plant, A. thaliana responses to different abiotic stresses have been extensively studied (Munns, 2002; Dubois et al., 2013; Trontin et al., 2014). A. thaliana is a glycophyte thus it has limited ability to accommodate and survive salinity stress (Guo et al., 2014). Salinity limits the growth and productivity as a result of inhibition of metabolic processes and vital enzymes, salt-responsive genes were recently reported (Shavrukov, 2013). Salinity perception starts from the roots and then the signal is transmitted to the aerial parts of the plant (Yadav et al., 2011). In the roots, several energy-related proteins were found to be upregulated during the plant salinity responses: glycolysis proteins including glyceraldehyde-3-phosphate dehydrogenase, electron transporting chain (ATPase subunit 1), and proton transporting chain (ATPase subunit B1) (Zhao et al., 2013).

Temperature variability including plant exposure to low and high extremes affects plant cells viability, but plants can acquire adaptation and tolerance if pre-exposed to gradient continuous change in temperature (Clarke et al., 2004, Larkindale et al., 2005). Short-term exposure generally induces rapid metabolic responses (Rocco et al., 2013), alteration of membrane composition, biosynthesis of antioxidants (Li et al., 2011), and increase in ion permeability (Goulas et al., 2006, Gao et al., 2009).

A rapid but transient accumulation in reactive oxygen species (ROS) is a common stress consequence. This results in oxidative stress (Schutzendubel and Polle 2002; Horvath et al., 2007). The over production of ROS could disturb the plant redox system, cause cellular dysfunctions, and eventually lead to plant death (Moller, 2007). In order to adapt to stress and its consequences, plants tend to enhance their anti-oxidative mechanisms by changing gene expression and producing protective secondary molecules. Hydroxycoumarin biosynthesis is reported to be induced by stress, for example: scopoletin and ayapin were accumulated to higher levels in sunflowers after fungal infection and tissue injuries (Roseland, 1991).

The aims of these experiments are to isolate, quantify, and to explore the role of the hydroxycoumarins, scopoletin, esculetin and their respective β-glycosides scopolin and esculin in cassava roots and A. thaliana during stress
conditions and to investigate whether the scopoletin was biosynthesised *de novo* in response to stress conditions or released by hydrolysis of its glycoside scopolin.

### 2.2. Materials

**Plant material**

Wild type (WT) cassava TMS 60444 plants were propagated and grown in the University of Bath glass house under the following conditions: temperature 25-30 °C, relative humidity 80% and 16 h light. Other commercial cassava roots were purchased locally in Bath (Morrisons). Wild type *Arabidopsis thaliana* ecotype Columbia-0 (Col-0) seeds were obtained from the cassava research group in the Department of Biology and Biochemistry at the University of Bath and grown in the University of Bath growth room in the following conditions: temperature 21-22 °C, relative humidity 50-60% and 16 h light.

**Chemicals**

All solvents were purchased from Thermo Fisher Scientific, UK. All chemicals and reagents were purchased from Sigma-Aldrich Chemical Co. Ltd. UK, unless otherwise stated.

**Instrumentation**

**High Performance Liquid Chromatography (HPLC)**

Separation, characterization, and isolation of coumarins from commercial cassava were carried out using a High Performance Liquid Chromatography (HPLC) instrument with a Jasco PU-980 pump and monitored at $\lambda = 360$ nm using a Jasco UV-975 detector. Pre-packed C18 reverse phase analytical (4.6 mm x 150 mm) and semi-preparative (10 mm x 250 mm) Gemini columns from Phenomenex, USA were used. Samples (loop size 40 and 100 µl) were injected manually. Isocratic elution used acidified 16% aqueous acetonitrile (0.05% formic acid) at flow rates of 1 ml/min (analytical column) and 3 ml/min (semi-preparative column). Water was pre-filtered through Milli-Q plus PF using a Whatman cellulose nitrate filter with a membrane diameter of 47 mm and pore size of 0.45 µm. The mobile system was degassed (30 min) using a Decon ultra-sonicator, UK.
The chromatographic traces were recorded on a Goerz Metrawatt Servogor 120 recorder. Peaks observed at 7.5, 9.0, 14.0, and 27.5 min were collected and characterised.

**Mass Spectrometry (MS)**

High Resolution Electrospray Ionisation Mass Spectrometry (HR-ESI MS) was carried out on a Bruker Daltonics (micrOTOF), Germany in the Department of Pharmacy and Pharmacology, University of Bath.

**Nuclear Magnetic Resonance (NMR) Spectroscopy**

NMR spectra were obtained using Bruker Avance III spectrometer, Germany, the spectrometer operating at 500.13 MHz for $^1$H and 125.77 MHz for $^{13}$C. The data obtained were processed using MestReNova NMR, MestreLab Research software. The sample was dissolved in deuterated solvents (DMSO-d$_6$) purchased from Goss Scientific, UK. Chemical shifts (δ) are given in parts per millions (ppm) and referenced to the residual solvent peaks. Coupling constants are reported in Hertz (Hz), multiplicities are abbreviated as: singlet (s), doublet (d), doublet of doublet (dd), and multiplet (m). The HPLC-purified peak at 9.0 min was identified as scopolin (Figure 2.1). $^1$H NMR δ: 3.18 (1H, br s, H-2’), 3.30 (1H, br s, H-4’), 3.44 (1H, br s, H-3’), 3.45 (1H, br s, H-5’), 3.46 (1H, m, H-6’b), 3.70 (1H, br s, H-6’a), 3.82 (3H, s, 6-OMe), 5.08 (1H, d, $J=10.0$, H-1’), 6.33 (1H, d, $J=9.5$, H-3), 7.16 (1H, s, H-8), 7.30 (1H, s, H-5), 7.96 (1H, d, $J=9.5$, H-4). $^{13}$C-NMR δ: 56.16 (6-OMe), 60.78 (C-6’), 69.73 (C-4’), 73.19 (C-2’), 76.89 (C-5’), 77.26 (C-3’), 99.72 (C-1’), 103.14 (C-8), 109.81 (C-5), 112.39 (C-4a), 113.45 (C-3), 144.36 (C-4), 146.12 (C-6), 149.07 (C-8a), 150.04 (C-7), 160.40 (C-2)

![Fig. 2.1. Numbering of scopolin.](image)
Incubator

A MaxQ4000 bench-top incubator (Thermo Fisher Scientific, UK) was used to expose the A. thaliana plants to a constant heating system at 40 ºC for 3 days.

Freeze dryer

Concentration of the plant extracts for analytical purposes was carried out using a SuperModulyo freeze dryer with an ice-condenser section and an Edwards’ vacuum pump attached to a heat controlled Speedvac chamber to hold 2 ml plastic vials (Stratech Scientific, London).

UHPLC/ESI-MS

Separation and quantitative analysis of coumarins were carried out using an UltiMate High Performance Liquid Chromatograph (UHPLC) from Thermo Fisher Scientific coupled to electrospray ionisation Time-of-Flight (ESI-TOF) HR-MS from Bruker (Germany). Pre-packed C18-reversed phase column Acquity UPLC BEH (1.7 µm, 2.1 x 50 mm) from Waters (USA) was used. Samples (10 µl) were injected using an RS auto-sampler, gradient elution with degassed methanol/water system was carried out at a flow rate of 0.3 ml/min. The data obtained were processed and analysed using Data Analysis 4.3 software from Bruker (Germany).

2.3. General methods

Growing plants

*M. esculenta* Crantz plants

Small fresh cassava buds (6-10 mm) were cut from *in vitro* cassava plants and allowed to grow for 2-4 weeks in pre-prepared sterile Cassava Basic Medium (CBM) pots containing a nutrient-agar medium (per litre: Murashige-Skoog MS (4.4 g), sucrose (20 g), CuSO₄ (2 x 10⁻⁶ M, 0.32 g) and Gelrite (3 g) pH adjusted to 5.9 using NaOH 0.1 M) at 24 ºC, relative humidity 50% and 16 h light. Once ready, the healthy grown roots and stems were washed out from the agar and transferred into soil (3:1 Levington’s M2: perlite), moved to the University tropical glass house and allowed to grow for 6-8 months (30 ºC, 16 h light and 17 ºC, 8 h dark).
A. thaliana plants

A. thaliana seeds were handled, germinated and plants were allowed to grow according to the instructions of the Arabidopsis Biological Resource Centre ABRC. Seeds (40-60) were surface sterilized before being planted in soil or sown over agar plates (initial washing for 5 min with 70% ethanol (800 µl), sterilizing for 5 min with freshly prepared bleach solution (800 µl of 50% bleach, 0.1% tween 20) and rinsing 3 times with Milli-Q water (1 ml). Finally, the treated seeds were stratified in phyto-agar solution (1 ml of 0.15%) at 4 ºC for 2-3 days to enhance their germination https://abrc.osu.edu (accessed on 01.08.17).

For root harvesting purposes, stratified A. thaliana seeds were germinated and aseptically grown for 3-4 weeks in Petri dishes contain nutrient phyto-agar medium (0.8%, pH 5.9) prepared by adding Murashige and Skoog salt with composing vitamins (4.4 g), 2-(N-morpholino)ethane-sulfonic acid (MES) (0.5 g), sucrose (10 g) and Milli-Q water up to 1 l.

For other experiments (collecting seeds, nucleic acids isolation, and A. tumefaciens–mediated transformation), stratified A. thaliana seeds were germinated and allowed to grow for 7-8 weeks in a pre-treated nutrient A. thaliana soil with an aqueous insecticide solution (0.2%) of Intercept 70 WG® (75% w/w imidacloprid) purchased from Scotts, UK.

Plants were grown in soil trays and Petri dishes for the required period in the University of Bath growth room in the following conditions: temperature 22 ºC, relative humidity 50-60% and 16 h light.

Hydroxycoumarin isolation

Based on Bayoumi et al. 2010 protocol, commercial cassava roots (1.5 kg) were washed under running tap water, dried, peeled (1.2 kg) and cut into cubes (~1 cm³), and allowed to deteriorate for 4 days in a controlled environment (20 ºC, 80% humidity). The deteriorated roots were then crushed using a porcelain mortar and pestle and double extracted with ethanol (2 mg/ml). The ethanolic extracts were combined, filtered, and concentrated under reduced pressure (35-40 ºC). The dried extract was fractionated between water and chloroform, and then subjected to further purification analysis methods using solid phase extraction (SPE), thin layer chromatography (TLC) and HPLC.
The chloroformic fraction: anhydrous sodium sulfate was added in a sufficient quantity to the combined chloroformic layers, shaken vigorously, filtered and concentrated under reduced pressure. The chloroformic extract (76 mg) was further purified by SPE as follows: Varian C18 SPE columns were washed (5 ml methanol) and conditioned (5 ml of 50% aq. methanol), before the prepared extract (1 mg/ml) was loaded and eluted with 50% aq. methanol. The eluent was concentrated under reduced pressure (35-40 ºC) and subjected to TLC analysis. The aqueous fraction: the combined aqueous layers were concentrated under reduced pressure (8.2 g) and a methanolic solution of 1 mg/ml was prepared, filtered and subjected to further purification and analysis by TLC and HPLC.

Thin layer chromatography (TLC)

Primary identification of the extracted coumarins was carried out by spotting diluted samples of the extracts on the analytical TLC plates purchased from Merck (Germany) and comparing the retention time with the standard reference coumarins. The analytical TLC plates used were pre-coated aluminium backed with silica gel 60 F254 of 0.20 mm thickness. The mobile phase used was chloroform: ethyl acetate: methanol (2:3:1 v/v/v). The developed spots were visualised under UV light using short- and long-wave radiation \(\lambda = 254\) and 365 nm respectively. Each spot was methanol-extracted directly from the TLC plate using the CAMAG TLC-MS interface (Switzerland), after precise positioning of the spot using the integrated laser head. A flow-rate of 0.1-0.2 ml/min was used using an HPLC pump, the extract was collected and subjected to HR-MS analysis to determine the accurate molecular weight.

Quantitative determination of coumarin levels in wild-type plants using LC/MS

Coumarins from *M. esculenta* roots during PPD

Wild type cassava roots were harvested, washed, peeled and cut into cubes (~1 cm³). The cubes were distributed into 8 groups of 17 g each. One group, labelled as fresh roots, were immediately crushed and extracted with ethanol. The other groups were placed in Petri dishes and stored in a controlled environment (20 ºC, 80% humidity) to induce PPD. Then, on a daily basis, one group was crushed and extracted with ethanol (60 ml). The 8 extracts (fresh and day 1-day 7) were
separately concentrated under reduced pressure, then 100 mg of each extract was re-dissolved in water (500 µl) before being filtered through a Fisherbrand™ non-sterile PTFE syringe filter with a membrane diameter of 25 mm, and pore size of 0.45 µm), and then subjected to LC/MS analysis to quantify the coumarins of interest. Each experiment was performed with at least three biological replicates and each sample was analysed with three technical replicates.

**Coumarins from *A. thaliana***

Sterilized *A. thaliana* seeds were allowed to grow (4-5 weeks) in nutrient-agar square plates vertically to ease root harvesting in the University of Bath growth room before the aerial parts and roots were harvested, thoroughly washed under running tap water, and then dried (55 ºC, 30 min). Dried roots (10 mg) and dried leaves (100 mg) were separately crushed and extracted with methanol (2 ml) which afforded effective extraction of the coumarins of interest. 4-Methylumbelliferone (100 ng/ml) was added to the extraction solvent as an internal standard to improve data reliability between the three biological replicates. The mixture was shaken (20 ºC, 16 h) before being filtered through a Fisherbrand™ non-sterile PTFE syringe filter with a membrane diameter of 13 mm, and pore size of 0.2 µm, and lyophilized. The extract was dissolved in methanol (200 µl) with scoparone (100 ng/ml) as a second internal standard to improve data reliability between the three technical replicates. The samples were subjected to LC/MS analysis to quantify the four selected hydroxycoumarins: scopolin, scopolin, esculetin, and esculin. Each experiment was performed with at least three biological replicates and each sample was analysed with three technical replicates.

**Induction of abiotic stress in wild type *A. thaliana***

**Salinity and osmolality stress**

*A. thaliana* plants were allowed to grow vertically in agar plates under their normal growing conditions as stated above to allow harvesting the roots. After 3 weeks, each plate was flooded with filter-sterilized NaCl aqueous solution (4 ml, 150 mM, pH 7) or filter sterilized mannitol aqueous solution (4 ml, 150 mM, pH 7). Solutions were filtered through a sterile Sartorius™ Minisart™ NY syringe filter with a diameter of 25 mm and a pore size of 0.2 µm. The plants were allowed
to grow horizontally for 3 more days to allow better utilisation of the solution. The concentration of the saline solution was chosen in accordance with the literature (Horvath et al., 2007) and likewise the concentration of the mannitol solution (Claeys et al., 2014; Trontin et al., 2014. After that, the leaves and roots were separated, harvested, rinsed briefly under tap water, and dried (55 °C for 30 min). The dried plant materials were homogenised (manually or by using a clean plastic rod in an Eppendorf tube) and subjected to LC/MS analysis.

**Heat shock**

*A. thaliana* plants were allowed to grow vertically in agar plates under their normal growing conditions as stated above. After 3 weeks, plates were placed in a MaxQ4000 bench top incubator (40 °C) for three days. The temperature was chosen according to a previous report (Zhao et al., 1998). After that, the leaves and roots were separated, harvested, and dried (55 °C for 30 min). The dried plant materials were separately homogenised and subjected to LC/MS analysis.

**UPLC/ESI-MS**

Methanolic and aqueous samples were separated using a reverse phase (RP) C18 column (UPLC BEH C18, 1.7 µm, 2.1 x 50 mm) from Waters, USA. Gradient elution used acidified methanol (0.1% formic acid) and acidified water (0.1% formic acid) in a total run-time of 10 min with the following gradient: aqueous methanol solution (1% until 2 min), linear gradient (40% until 5 min), followed by increasing gradient (99% until 8 min), followed by a washing-out stage (99%, 2 min) with methanol for re-equilibration to the initial conditions. The flow rate was 0.3 ml/min and the injection volume was 10 µl. Detection was performed by HR-TOF MS for the coumarins of interest (scopoletin, esculetin and their respective β-glycosides scopolin and esculin) obtained on a Bruker Daltonic micrOTOF spectrometer using ESI in positive-ion mode. Data obtained were processed using Data Analysis software 4.3 from Bruker (Germany). The detected mass peaks of the compounds of interest [M+H]+ and [M+Na]+ were typically matched with their corresponding theoretical values within a 5 ppm error range unless otherwise stated. The theoretical [M+H]+ mass/charge (m/z) ratios are: scopoletin (C_{10}H_{9}O_{4}) 193.0495, scopolin (C_{16}H_{10}O_{5}) 355.1024, esculetin (C_{9}H_{7}O_{4}) 179.0339, and esculin (C_{15}H_{17}O_{9}) 341.0867.
2.4. Results and discussion

The coumarins of interest were isolated and characterized from deteriorated cassava roots. The crude ethanolic extracts of the fresh and deteriorated cassava roots were analysed by TLC. Blue fluorescent spots were visualized under UV light with the same Rf values obtained for the reference standards: scopoletin, esculetin and esculin. The molecular weights of the three hydroxycoumarins in addition to an additional TLC spot identified as scopolin, were unambiguously confirmed by HR-MS after extracting the TLC spots using the TLC-MS interface. The protonated molecular ion for scopoletin [M+H]+ \( \text{C}_{10}\text{H}_{9}\text{O}_{4} \) requires 193.0495, found 193.0491. The deprotonated molecular ion for esculetin \( \text{C}_{9}\text{H}_{5}\text{O}_{4} \) [M-H]− requires 177.0188, found 177.0213. The deprotonated molecular ion for esculin \( \text{C}_{15}\text{H}_{15}\text{O}_{9} \) [M-H]− requires 339.0716, found 339.0724. The sodium adduct molecular ion for scopolin \( \text{C}_{16}\text{H}_{18}\text{O}_{9}\text{Na} \) [M+Na]+ requires 377.0849, found 377.0853 (Figure 2.2).

![Image](image.jpg)

**Fig. 2.2.** HR-MS data of the four hydroxycoumarins identified by TLC-MS extraction.

The accumulation of the hydroxycoumarins in cassava roots undergoing PPD process was expected and is a confirmation of their role in the PPD process as they are known to have potential phytoanticipant or phytoalexin activity (Wheatley and Schwabe, 1985; Buschmann *et al.*, 2000a). These TLC and HR-MS results confirmed the accumulation of scopoletin, scopolin, esculetin, and esculin in the deteriorated cassava roots compared to the fresh roots (t=0). Chemical structures of the four hydroxycoumarins are shown in Figure 2.3.

41
Fig. 2.3. Hydroxycoumarins isolated from deteriorated cassava roots.

The chloroformic fraction was further concentrated by SPE to afford scopoletin identical by TLC to the authentic sample. The other hydroxycoumarins could be too polar to be obtained by extraction into chloroform or are not detectable because they were present in much lower concentration than scopoletin (Wheatley and Schwabe, 1985; Buschmann et al., 2000a; Bayoumi et al., 2008a).

Different peaks were obtained from the aqueous fraction analysed by HPLC (Figure 2.4). Peaks observed at 7.5, 14.0, and 27.5 min showed identical retention times with the reference standards of esculin, esculetin, and scopoletin respectively (Figure 2.3). These, in addition to an extra peak (9.0 min), were collected, concentrated, and their identities were confirmed by HR-MS. The peak at 27.5 min was confirmed to be scopoletin not iso-scopoletin as shown by the retention time difference between the two authentic standards (Figure 2.5). The HPLC peak (9 min) was collected from ~100 injections (100 µl loop), lyophilized to yield 3.5 mg and identified as scopolin using HR-MS within 1 ppm mass error range $C_{16}H_{18}O_9Na [M+Na]^+$ requires 377.0849, found 377.0853.
Fig. 2.4. HPLC trace of ethanolic extract of deteriorated cassava roots. The compounds were separated by semi-preparative C18 column using 3 ml/min flow rate. Esculin, esculetin and scopoletin were identified by comparing their retention times with the standards. Traces were recorded with a chart speed of 1 cm/min.

All peak identities were confirmed using HR-MS.

Fig. 2.5. HPLC trace of scopoletin and isoscopoletin co-injection. The compounds were separated by semi-preparative C18 column using 3 ml/min flow rate, and were identified by comparing their retention time with the standards. Traces were recorded with a chart speed of 1 cm/min.
The compound identity was confirmed using $^{13}$C (Figure 2.6) and $^1$H NMR spectroscopy (Figure 2.7). Four peaks were detected in the aromatic region in the $^1$H NMR spectrum, two doublets with a coupling constant of 9.5 Hz and two singlets. H4 was shifted downfield because of the deshielding effect of the benzene ring, H3 was shifted upfield because of the carbonyl group electron withdrawing effect. Methoxy group appeared at 3.82. The anomeric proton chemical shift was more downfield (5.08) and the observed coupling constant ($J = 10.0$) confirmed the β-configuration (Drake and Brown, 1977). The sugar protons signals were assigned according to the two-dimensional HSQC-NMR and in reference to the NMR spectra of pure glucose NMR and the published data (Drake and Brown, 1977). An enhanced number of scans (1048) was run to obtain the $^{13}$C signals, 16 carbon signals were obtained in addition to the residual solvent signals (C$_2$D$_6$SO at $\delta = 39.5$, CD$_3$OD at $\delta = 49.1$) (Figure 2.7). The two-dimensional HSQC-NMR was run to match the proton-carbon signals (Figure 2.8). Chemical shifts were in agreement with previous reports (Kuroyanagi et al., 1986; Fliniaux et al., 1997).

Since scopolin was isolated using an acidified mobile phase and then used as an analytical reference over a period of 3 years, the β-glycosidic bond became partially hydrolysed and scopolin therefore decomposed to its aglycone, scopoletin. This issue was corrected for mathematically when the compound was required as a quantitative reference, e.g. for the construction of calibration curves, the proportion of scopolin remaining was calculated in order to apply as appropriate when used as a reference in a calibration curve.

Fig. 2.6. $^{13}$C-NMR spectrum for scopolin (C$_2$D$_6$SO).
Fig. 2.7. $^1$H NMR spectrum for scopolin. A shows the full $^1$H NMR spectrum (C$_2$D$_6$SO), B shows the expanded aromatic region, and C shows the expanded sugar region (ns=128).
Fig. 2.8. HSQC spectrum of scopolin. A Shows the expanded aromatic region, B shows the expanded sugar region.
Quantification of coumarin levels in fresh cassava roots and during PPD

The concentration of scopoletin was found to be minimal in fresh roots and there was an increasing pattern after harvesting, reaching its maximum on the second day (~778.0 ± 4.4 ng/mg fresh weight). A gradual decline in scopoletin accumulation was found over the following three days which could be due mainly to the exhausted tissues with their lower ability to replenish the loss in scopoletin by de novo synthesis. In general, the increase in scopoletin accumulation within the early few days after harvest is linked to the initial physiological discolouration of the vascular tissues while the second increase in scopoletin level measured on the days 6 and 7 could be related to other biological and microbial tissue breakdown and rotting (Wheatley and Schwabe, 1985) (Figure 2.9). Buschmann et al. (2000b) also reported that some cassava cultivars exhibited this bimodal increase in scopoletin accumulation at the second and after the 5th days, these cultivar varieties include the low-PPD susceptible cultivar MBRA 337, and the high PPD-susceptible cultivar MNGA 2. Liu et al. have reported a similar trend in scopoletin accumulation pattern, but with higher overall concentration, e.g. the concentration measured in days 3-5 was ~12 ng/mg ± 6.0 (Liu et al., 2017). These variations in concentration of scopoletin in fresh root may arise from using different extraction methods. Other cassava varieties also showed increasing levels of scopoletin in roots after harvesting, reaching a peak on day 2 (Buschmann et al., 2000b)

![Fig. 2.9. Scopoletin level in wild type (TMS 60444) cassava roots.](image-url)
The accumulation of scopoletin (n=3) in cassava fresh roots and during PPD.
These results confirmed the correlation between scopoletin accumulation in cassava roots after harvesting and the symptoms of PPD, mainly root discolouration (Figure 2.10). The blue-blackish discolouration could be linked, in addition to other factors, to the oxidation of scopoletin (Tanaka et al., 1983; Blagbrough et al., 2010).

![Fig. 2.10. Fresh and deteriorated cassava roots.](image)

Images show cross-sections of fresh (A) root, and black-discolouration after: day 2 (B), day 3 (C), day 4 (D), and day 5 (E) cassava root harvesting.

An increasing accumulation of scopolin was measured in cassava roots during PPD compared to the undetectable levels (i.e. below the LLD) in the fresh roots. The peak accumulation was measured at day 2 ~1.38 ± 0.02 μg/g fresh root ($P$ values between days 1, 2 and between days 2, 3 were < 0.001) and then, a gradual decline was observed over the next five days (Figure 2.11). Scopolin accumulated to a higher extent than scopoletin (~1.9x) early in PPD, e.g. the average concentration of scopoletin in days 1-3 ~632.5 ± 104 ng/g, while the average concentration of scopolin in the same period ~1212.9 ± 162.6 ng/g.
The scopolin accumulation pattern was compatible with the pattern from the aglycone scopolin. No scopolin was detected directly after harvesting the roots. From day 1, a gradual increase in accumulation was measured until it reached the maximum on day 2 (1.38 ± 0.02 µg/g). Later in the PPD process, a decline in the level was measured starting from day 3 (on day 3 ~996 ± 0.07 ng/g, and days 5-7 ~250 ± 1.5 ng/g). These results support the hypothesis that both scopolin and scopolin were de novo biosynthesized in response to stress early in PPD. The possibility of prompt scopolin deglycosylation to release the antioxidant scopolin to encounter the massive increase in ROS and H$_2$O$_2$ released early after harvesting the roots is not excluded (Reilly et al., 2003). The second increase in scopolin level on the 6th and 7th days despite the exhausted tissues may originate from scopolin deglycosylation. The ratio scopolin: scopolin obtained from the used cassava cultivar TMS 60444 agrees with the ratio obtained previously from other cassava varieties (Buschmann et al., 2000a).

Other hydroxycoumarins, namely esculetin and its glycoside esculin, were quantified and their accumulations during the PPD process were minimal compared to scopolin and scopolin (Figure 2.12). For example, within the PPD peak period (days 2 and 3), the measured scopolin concentration was 12x the measured esculin concentration, while the level of esculetin was below the limit of quantification. This less significant accumulation after harvest is potentially indication of the lower impact esculetin and esculin might play in cassava roots discolouration. Interestingly, an increasing concentration of esculin was measured late in PPD,
peaking at the 6th and 7th days ~1 µg/g (P value between days 6 and 7 = 0.77), perhaps this is indicative of to the need for more free radical scavengers to counter the increase in peroxidase activity during days 3-7 (Buschmann et al., 2000b).

Fig. 2.12. Esculetin and esculin levels in wild type cassava. The accumulation of scopoletin (in the extract of fresh 20g roots) in fresh roots and during PPD.

The overall pattern of the four studied hydroxycoumarins was in agreement with the literature for the early stage in PPD (the first 5 days), while the increase in scopoletin and esculin levels later in PPD matched the recent findings in the same cassava cultivars (Liu et al., 2017) and using other cassava varieties (Buschmann et al., 2000a). The overall conclusion is that there is an active involvement of scopoletin and its glycoside scopolin, in the process of PPD in wild type cassava. The biosynthesis of scopoletin was therefore thoroughly investigated in detail (chapters 3 and 4).

Quantification of hydroxycoumarins in A. thaliana

The aerial parts and the roots were harvested from 4-5 weeks old wild type A. thaliana growing in the University of Bath growth room. The levels of different hydroxycoumarins from each plant tissues were analysed separately.

In many of the assays, reproducible absolute concentrations could not be determined and in these cases the ratio between the accumulations of the hydroxycoumarins of interest between different plant parts or between different plant lines (chapters 3, 4) rather than the absolute concentration per tissue weight were recorded. This is because a variation was found in the absolute concentration per dried tissue weight obtained while optimizing the analytical methods and even
after repeating the same experiments after different time intervals but using the same methods. These challenges are due to different biological and technical limitations. These limitations include: the tiny *A. thaliana* root size 100-150 µm in diameter (Sena et al., 2011), plus the ease of contamination, created the need to prepare multiple agar plates to obtain the required root weight, in a limited space available. Also, the time limitation, where it took 4-5 weeks until harvesting the plant tissues and then 1 more week in preparing the samples. Due to the big demands on the available LC/MS, the plant materials mostly stored at -20 °C until analysed (sometimes, for up to 1 month). The intensity of the coumarin peaks tended to decrease after storage. Figure 2.13 shows the change in peak intensities after injecting the authentic standards freshly and after storing for 1 week at -4 °C. Accordingly, the ratios rather than the absolute amounts were used to compare the differences in coumarins accumulation in *A. thaliana*.

**Fig. 2.13. Difference in coumarin peaks intensities during storage.**

Figures show the changes in esculetin, scopoletin, scopolin and esculin peak areas relative to the internal standard (4-MU) after 7 days storage at -4 °C. Blue lines for the fresh standards injections, orange lines for the stored standards injections.

However, in wild type *A. thaliana* roots the measured concentrations of scopoletin, scopolin, and esculin were found to be: ~13.8 ± 1.35 ng/mg, ~3.6 ± 0.07 µg/mg and ~4.5 ± 0.8 ng/mg respectively. Other hydroxycoumarins, esculetin
and umbelliferone were also detected. The level of esculetin was below the lower limit of quantification (LOQ), where the signal to noise ratio less than 10. Umbelliferone concentration was inconsistent and variable. Hardly any of the hydroxycoumarins was detected in the aerial parts of the A. thaliana plants except for scopolin which was found at ~2.0 ng/mg, possibly because the hydroxycoumarins were biosynthesized in leaves and stored in the roots until needed (Ahn et al., 2010) (Figure 2.14). Other workers have reported widely varying levels of hydroxycoumarins in A. thaliana roots: scopoletin level in fresh roots was reported to be ~2.8 ng/mg and ~3.4 ng/mg by Kai et al. and Siwinska et al. respectively, while Liu, et al. reported much higher concentration (~150 ng/mg fresh weight). Scopolin was reported to be 424 ng/mg by Kai et al. (Kai et al., 2006; Siwinska et al., 2014; Liu et al., 2017). The variation is mainly attributed to the use of different extraction and quantification methods and different cassava cultivars.

In the wild type A. thaliana roots, the scopolin accumulated to a much higher extent than scopoletin. Presumably due to the activity of scopoletin glucosyltransferase to store scopolin in a more transportable form within the plant parts, decrease its toxicity, and facilitate its prompt release from its glycosidic form upon exposure to stress by the action of scopolin glycosidase (Bowles et al., 2005; Bourgaud et al., 2006; Chong et al., 1999). Scopolin undergoes deglycosylation just before or right after being exported across the plasma membrane, the translocation or movement of scopolin occurs against the concentration gradient, suggesting that scopolin uptake involves active transport into the cell (Taguchi et al., 2000; Fraser and Chapple, 2011). This assumption could be applied to esculetin and esculin (Schmid et al., 2014).

The accumulation pattern of the hydroxycoumarins of interest varies between cassava and A. thaliana roots. Cassava tends to synthesize the coumarins in response to wounding, while scopoletin and scopolin were detected in A. thaliana roots at harvest. Nevertheless, A. thaliana was used as a model plant to investigate the responses to salinity and osmolality abiotic stressors which mimic the effect of drought on plants, and the overheating effect, and therefore, their influence on the hydroxycoumarin accumulation profiles.
**Fig. 2.14. Hydroxycoumarin levels in wild type A. thaliana.** Histograms show the concentrations (n=3) of scopoletin (A), scopolin (B), and esculin (D) in root and aerial tissues. C shows the accumulation of both scopoletin and scopolin in the A. thaliana roots.

**Salinity effect**

*A. thaliana* plants (4-5 weeks) were flooded with sterile NaCl solution (150 mM, 4 ml) for 3 days before the roots were harvested and the hydroxycoumarins were quantified (Figure 2.15). Serial concentrations (50-200 mM) were also used to optimize the method. The average concentration of the scopoletin (n=3) was not significantly different in the stressed roots compared to the unstressed roots ($P = 0.12$), while the average concentration of the scopolin (n=3) showed a small but significant decrease (0.9x) in the stressed roots ($P < 0.001$). The experiment was repeated twice with three biological replicates each. The unexpected experimental results may indicates the complex response to salinity, which is not only by a direct increase in the antioxidants. Salinity is a multifactorial stress (Guo *et al*., 2014); it mimics the drought effect by causing water depletion of the medium as a result of reduction in osmotic potential of the medium, in addition to the cellular Na$^+$ toxicity that have an inhibitory effect on the plant metabolism (Guerrier *et al*., 1996; Verslues *et al* 2006; Claeys *et al*., 2014). Other major processes in the plant will be affected such as photosynthesis, mineral uptake and growth (Yadav, 2011).
**Fig. 2.15. Salinity effect in hydroxycoumarin levels in wild type A. thaliana roots.** Graphs show scopoletin (A) and scopolin (B) levels in the normal and salinity-stressed A. thaliana root tissues.

It has been reported that many polyols accumulated to considerable levels in response to osmotic adjustment and osmotic protection in plants (Bohnert *et al.*, 1995). However, the unexpected changes in both scopoletin and scopolin levels obtained in the osmotically stressed A. thaliana plants perhaps referred to the effect of salt-shock upon the plants where the mature plants were treated with toxic concentrations of NaCl over a short period (3 days), rather than allowing the plants to adjust itself to adapt the unfavourable conditions by increasing the antioxidants levels. Salt-shock mostly caused the plants to suffer plasmolysis, which will lead to different expression profile of the stress inducing genes (Shavrukov, 2013).

**Mannitol effect**

*A. thaliana* plants (4-5 weeks) were flooded with sterile mannitol solution (150 mM, 4 ml) for 3 days before the roots were harvested and the hydroxycoumarins of interest were quantified. Serial concentrations (50-200 mM) were also used to optimize the method. The measured average concentrations of the scopoletin and scopolin (n=3) were significantly ($P < 0.001$) lower in the untreated roots (~1/3x) (Figure 2.16). Scopoletin concentration in the unstressed wild type root extract was ~11.6 ± 1.9 ng/mg and ~ 4.3 ± 0.55 ng/mg in the mannitol stressed roots. Scopolin concentration in the unstressed wild type root extract was ~3.6 ± 0.07 µg/mg and ~0.97 ± 0.5 µg/mg in the stressed root extract.

Mannitol is an abiotic and osmotic stressor which acts by causing a reduction in the potential water available in the medium (Trontin *et al.*, 2014) and it is also considered as a biotic stressor regarding to its production by many fungi.
(Claeys et al., 2014). These dual stress effects were expected to induce the production of coumarins, but the experimental data did not match the expectations. The reduction in both scopoletin and scopolin levels could be explained by the growth arrest triggering effect of mannitol on A. thaliana (Trontin et al., 2014).

The expression of the stress-induced genes is not always predictable in plant in response to different stressors, e.g. ethylene-responsive transcription factor 5 ERF5 is highly induced by mannitol but not NaCl although, they both shared the osmotic effect (Munns, 2002; Dubois et al., 2013). The plants were shocked with a high concentration of mannitol solution over a short period rather than responding to environmental changes gradually and this may explain the unpredicted behaviour.

Heat shock

A. thaliana plants growing in agar plates (4 weeks old), were placed for three days in an incubator at 40 °C, data showed a dramatic decrease in scopoletin and scopolin levels in response to the increase in temperature compared to the plant growing at normal temperature (22 °C). The measured average (n=3) concentration of scopoletin in the untreated roots was ~13.87 ± 1.35 ng/mg and in the treated roots ~4.5 ng/mg, while the average concentration of scopolin was ~3.6 ± 0.07 µg/mg in the unstressed roots and ~2.17 ± 0.24 µg/mg in the heat-shocked root extract. The decrease in scopoletin (3x) and scopolin (1.6x) levels were unexpected because of the reported increase in stress enzyme expression levels in
response to heat (Panchuk et al., 2002) which presumably lead to increase in the antioxidant levels, but it is not necessarily that the information getting from transcriptomic data will always match the cellular protein complement (Fernie and Stitt, 2012). On the third day under heat shock, the plants started to turn chlorotic and they died on the fourth, these observation were concomitant with previous reports (Zhao et al., 1998), this could turn the plant tissue exhausted and unable to replenish the deficiency in scopolin. Scopolin released the stored scopolin by the action of scopolin glycosidase enzyme and the later got oxidized to attenuate the increase in ROS (Ahn et al., 2010; Buschmann et al., 2000a) (Figure 2.17).

**Fig. 2.17. Heat shock effect on hydroxycoumarin levels in wild type A. thaliana roots.** Graphs show A scopoletin and B scopolin levels in the normal and heat-shocked (40 ºC for three days) A. thaliana root tissues.

The overall conclusions from these experiments is that the hydroxycoumarins scopoletin and its β-glycoside scopolin are actively involved in the appearance of PPD symptoms a few days after cassava root harvesting. Scopoletin showed a bimodal increasing pattern, the first was 2-3 days after harvest due to the direct respond to wounding, and the second was later after harvest (5-7 days) which could result from a biological stress response. Different cassava varieties vary in their PPD response and in the amount of hydroxycoumarins accumulated after harvest, but in general, they share the same pattern in increasing levels of both scopoletin and scopolin compared to their non-detectable levels immediately after harvesting.

Modelling the cassava responses to abiotic stress in A. thaliana showed unexpected results. Decreased levels of the hydroxycoumarins were measured in
the mannitol-stressed and heat-shocked A. thaliana roots compared to the untreated roots, while no significant changes were measured with NaCl-stressed roots. There are differences in how different plants respond to stress and how the same plant responds to different stressors. For example, while some plants tend to alter their physiology to survive the stress, in other plants the same stress can prove fatal especially in extreme cases. The overproduction of the defensive secondary metabolites in cassava is perhaps related to the protective mechanisms against the expected (by the plant) secondary microbial infection after wounding by harvesting. Mimicking the abiotic stress in A. thaliana by increasing the osmolality and the temperature resulted in changes at the level of plant biochemistry where the hydroxycoumarin levels unexpectedly decreased, rather than requiring these molecules for defensive roles.

In order to explore the biosynthetic pathways leading to scopoletin and scopolin accumulation in cassava, and to investigate the key enzymes involved, experiments were designed and performed (chapter 3) to quantify the hydroxycoumarins of interest in transgenic A. thaliana plants which lack the OMT genes essential for scopoletin biosynthesis. In these mutant plants, the hydroxycoumarin levels will be quantified and compared to their levels in the wild-type plant. Homologous cassava OMT genes will then be overexpressed in the mutant plants in order to explore their functions.
Chapter 3. Complementation of *Arabidopsis thaliana* mutants through expressing cassava *O*-methyltransferase genes

3.1. Introduction

PPD is an active endogenous oxidative process which starts immediately after wounding the roots by harvesting, and is observed as a blue-black discolouration of the parenchyma (Buschmann *et al*., 2000a). In cassava, as in other plants, the disruption of the physical barrier will be followed by enhanced oxygen levels, increased risk of microbial invasion, and alteration in the metabolome including gene expression levels, e.g. catalase CAT1 and catalase CAT2 are upregulated in cassava roots during PPD (Reilly *et al*., 2007). Among the common wound-responses in cassava is the enhanced production of defence compounds, mainly hydroxycoumarins (Beeching *et al*., 1993; Sanchez *et al*., 2006). Scopoletin and its β-glycoside scopolin, along with other hydroxycoumarins such as esculetin and its β-glycoside esculin, are known to be involved in the PPD process due to the increase in their concentrations in the days immediately after harvesting the roots (chapter 2). The oxidation and perhaps the polymerization of these small molecules may well be responsible for the discolouration of the roots (Bayoumi *et al*., 2008a; Blagbrough *et al*., 2010; Moyib *et al*., 2015).

Scopoletin is derived from L-Phe in the phenylpropanoid pathway which begins with deamination by PAL into *E*-cinnamate. Several regiospecific hydroxylation and *O*-methylation reactions follow to produce scopoletin which is stored in the plant vacuoles in its water soluble form, scopolin, after glycosylation mediated by scopoletin glucosyltransferase (GT) (Bourgaud *et al*., 2006; Ahn *et al*., 2010). Different hydroxylation and *O*-methylation reaction sequences will produce different possible pathways along which scopoletin could be biosynthesised (Figure 3.1). These possible pathways occur via three different intermediates: via 4´-hydroxy-3´-methoxycinnamate (ferulate) (pathway 1), or also via 3´,4´-dihydroxy-cinnamate (caffeate), but then 6´-hydroxylation to *E*-6´-hydroxycaffeate (pathway 2), or via 2´,4´-dihydroxy-cinnamate, isomerisation and then lactonization leads to 7-hydroxycoumarin (umbelliferone) (pathway 3). In *A. thaliana*, the major pathway to scopolin was defined after several *in planta* studies as pathway 1 (Figure 3.1). The genes potentially involved were isolated and characterised (Kai *et al*., 2006; Kai *et al*., 2008). Whereas in cassava, using the current update of the
cassava genome published in the Phytozome search tools (Wang et al., 2014), the genes involved in scopoletin biosynthesis were identified (after translation) by their amino acid sequence similarities with the reference *A. thaliana* proteins.

![Diagram of scopoletin biosynthesis]

**Fig. 3.1. Three possible metabolic pathways in scopoletin biosynthesis.**

Scopoletin could be synthesised from *E*-p-coumarate after 3′-hydroxylation by C3′H, 3′-*O*-methylation by CCoAOMT, and 6′-hydroxylation by F6′H1, then isomerization, and lactonization (pathway 1); or after 3′- then 6′-hydroxylation, isomerization, and lactonization to esculetin (pathway 2); or after 2′-hydroxylation, isomerization, and lactonization to umbelliferone, then 6′-hydroxylation to esculetin (pathway 3). Regioselective *O*-methylation of the 6′-hydroxy of esculetin will produce scopoletin (pathways 2 and 3).
None of the hydroxylase and methyltransferase enzymes on pathways 2 and 3 has been reported in cassava.

**Plant O-methyl transferase genes**

About 40 sequences for plant O-methyltransferase (OMT) genes, found in the sequence databases, encode functional proteins that have been published (Chiron et al., 2000). OMTs catalyse the methyl group transfer from one-carbon donor amino acid S-adenosyl-L-methionine (SAM) onto an oxygen of an acceptor molecule to yield the corresponding methyl ether derivatives. These O-methylated compounds have an essential role in regulating plant growth, development, and defence (Lam et al., 2007). Most of these OMT genes are involved in secondary metabolite biosynthesis (Pichersky and Gang, 2000).

Plant OMTs are classified according to their affinity to methylate different chemical groups within plants: phenylpropanoids (Busam et al., 1971; Bugos et al., 1991), flavonoids (Maxwell et al., 1993; Gauthier et al., 1996), benzyl isoquinoline alkaloids (Takeshita et al., 1995; Frick and Kutchan, 1999), and myo-inositol (Vernon and Bonhert, 1992).

Kim et al. (2005) made a more detailed classification according to specific substrate groups: caffeoyl CoA in lignin formation, caffeic and 5′-hydroxyferulic acids, flavonoids synthesised under stress conditions, and alkaloids. In another classification, according to their molecular weight, OMTs are categorized into two major groups: lower molecular weight (23,000-27,000 Da) enzymes whose catalytic activities depend on Mg$^{2+}$ cations and higher molecular weight (38,000-43,000 Da) which are not Mg$^{2+}$ dependent (Ibdah et al., 2003).

In this chapter, these three possible scopoletin biosynthetic pathways (Figure 3.1) were explored in genetically modified *A. thaliana* by analysing and quantifying scopoletin along with other hydroxycoumarins. Mutant plants missing the CCoAOMT enzyme (referred to as A.t-CCoAOMT), missing the EOMT enzyme (referred to as A.t-EOMT), and a double mutant *A. thaliana* plant missing both enzymes will be made. In parallel, cassava OMTs will be identified not only based on amino acid sequence similarities with the reference proteins using bioinformatics tools, but also after elucidation of function through their ability to complement the decrease in scopoletin biosynthesis in mutant *A. thaliana* plants.
3.2. Materials

Plants

Wild type (WT) cassava TMS 60444 plants and wild type A. thaliana ecotype Columbia-0 (Col-0) were propagated and grown in the University of Bath glass house and growth rooms as stated in chapter 2. A. thaliana seeds of the T-DNA insertion lines Salk-055103 and Salk 135290 were purchased from The Nottingham Arabidopsis Stock Centre (NASC), University of Nottingham, and grown in the University of Bath growth room under the following conditions: temperature 21-22 °C, relative humidity 50-60%, and 16 h light.

Bacterial strains

High efficiency NEB 10-Beta chemically competent Escherichia coli cells with a transformation efficiency of 1-3 x 10^9 cfu/µg pUC19 DNA, harbouring a streptomycin resistant gene and sensitive to each of the following antibiotics: ampicillin, chloramphenicol, kanamycin, nitrofurantoin, spectinomycin and tetracyclin, were used for the transformation and cloning of the desired insert genes in different destination vectors depending on the experiment’s purpose. The cells were purchased from New England Biolabs (NEB). Wild type Agrobacterium tumefaciens GV3101 cells harbouring rifampicin and gentamicin resistant genes were supplied from the cassava research group at the University of Bath and used to prepare high efficiency electrocompetent A. tumefaciens cells.

Plasmid vectors

TOPO® TA cloning vector

The pCR2.1®-TOPO® TA vector of 3.9 kbp size directed by the T7 promoter was used to clone the 3’-deoxyadenosine (A) over-hanged PCR purified product. The plasmid vector was purchased from Fisher. Selection of the successful cloning products depends on both conferring resistance to ampicillin and kanamycin to the insert, and to the blue/white colour selection due to the presence of the LacZ-α-peptide that allows β-galactosidase activity. TOPO® TA has a universal M13 reverse priming site to aid sequencing the insert. The featured map is shown (Figure 3.2).
Fig. 3.2. Map of the pCR2.1®-TOPO® TA vector. This map shows kanamycin and ampicillin resistant genes, LacZ-α selectable marker, M13 reverse priming site, multiple cloning sites and 5’-thymidine residues. [Link](https://www.thermofisher.com/uk/en/home/life-science/cloning/topo/topo-ta-cloning/topo-ta-for-subcloning.html) (accessed on 01.08.17).

Gateway® Technology vectors

The Gateway® Technology enables quick and highly efficient (> 99%) directional cloning of the desired insert into a variety of protein expression systems adapted to this technology (Freuler et al., 2008).

a. Entry clone/Gateways® pDONR™ vector

The Gateway® adapted vector of 4.291 kbp size facilitates the ease of shuffling the gene of interest into a variety of expression vectors using this technology by recombining the desired insert into an attL-flanked entry clone containing the gene of interest as a result of a combination of attB-flanked PCR product with an attP-flanked entry clone (pDONR™). Selection of the successful recombinant depends on the conferred resistance to kanamycin and zeocin offered by the plasmid to the insert. pDONR™ vector has universal M13 priming sites to ease sequencing the insert. The plasmid vector was purchased from Thermo Fisher Scientific. The featured map is shown (Figure 3.3).
Fig. 3.3. Map of the Gateway® pDONR™ vector. This map shows kanamycin and zeocin resistant genes, M13 priming sites and the counter selectable ccdB genes flanked by the attP sequences. 

https://tools.thermofisher.com/content/sfs/manuals/gateway_pdonr_vectors.pdf (accessed on 01.08.17).

b. Over-expression clone/pCAMBIA-1305.1

The binary vector of 11.846 kbp size from the Centre for Application of Molecular Biology to International Agriculture (pCAMBIA-1305.1, Gen Bank: AF354045) was used to mediate cassava DNA transformation to the A. thaliana http://www.cambia.org. pCAMBIA directed by the 35S promotor was used to overexpress the selected cassava genes. The conferred resistance to the antibiotic hygromycin enabled the selection of transformed plants. The conferred resistance to the antibiotic kanamycin enabled the selection of the successful bacterial transformation. The vector was purchased from Invitrogen® and modified by the cassava research group/Beeching’s laboratory, University of Bath (Page, 2009). The featured map is shown (Figure 3.4).
The plasmid vector pCAMBIA 1305.1 was modified to meet Gateway\textsuperscript{®} cloning requirements. With the aid of restriction and ligase enzymes, Gus Plus reporter gene was replaced with Gateway\textsuperscript{®} cassette reading frame so \textit{att}R1 and \textit{att}R2 sequences were added to enable LR recombination reaction (refer to Gateway\textsuperscript{®} cloning section 3.3) (Page, 2009).

**Bioinformatics software**

The National Centre for Biotechnology Information (NCBI) [http://www.ncbi.nlm.nih.gov/](http://www.ncbi.nlm.nih.gov/) was used for collecting data about the selected genes. Phytozome search tools [http://www.phytozome.net/search.php](http://www.phytozome.net/search.php) were used for finding the homologous genes in \textit{Manihot esculenta} (Goodstein et al., 2012). Generation and analysis of the phylogenetic trees, alignment of nucleotide and amino acid sequences, and designing primers for different purposes were performed through Geneious\textsuperscript{®} V 5.3 “bioinformatics software.” Analyses of the characteristics such as melting temperature, GC percentage, hairpin structure of the designed primers, and finding the complementary primers were performed via the online software [http://eu.idtdna.com(calc/analyser](http://eu.idtdna.com/calc/analyser) (accessed on 01.08.17) (Ye et al., 2012).
Oligonucleotide primers

Oligonucleotide primers of an average of 20 nucleotides in length, annealing temperature of 40-60 °C, and with a GC content of not less than 40% and not more than 60%, were designed according to the experiment’s aim. Primers were used in different reactions to detect the presence of the gene, to check its expression, to isolate the full-length cDNA, or to confirm the gene’s identity. All the primers used were designed using Geneious® 5.3 software and optimized using Oligo Analyser 3.1 Integrated DNA Technologies, INC (US) [https://www.idtdna.com/calc/analyzer](https://www.idtdna.com/calc/analyzer), ordered from, and synthesized by Fisher.

Polymerase enzymes

Both Taq DNA polymerase and Q5® High-Fidelity DNA polymerase enzymes used in PCR reactions were purchased from NEB.

S.O.C. medium

Super-optimal broth medium with the addition of glucose, S.O.C. (super-optimal broth with catabolite repression) was used in the recovery step of the E. coli high-efficiency competent cells transformation to maximize the transformation efficiency. S.O.C. medium was prepared by adding (per litre): tryptone (20 g), yeast extract (5 g), NaCl (0.01 M, 0.584 g), KCl (0.0025 M, 0.186 g), MgCl₂ (0.01 M, 0.952 g), MgSO₄·4H₂O (0.01 M, 1.92 g) and glucose (0.02 M, 3.6 g). The pH was adjusted to 7 using NaOH (1 M), and the medium was sterilized by autoclaving (121 °C, 20 min), cooled and then stored at 4 °C until required.

Luria-Bertani (LB) medium and plates

The Luria-Bertani (LB) medium used to grow the plasmid culture was prepared by adding (per litre): tryptone (10 g), yeast extract (5 g) and NaCl (10 g). Low-salt LB medium was prepared (only 5 g of NaCl) for zeocin selective plates. The pH was adjusted to 7 using NaOH (1 M), and the medium was sterilized by autoclaving (121 °C, 20 min), cooled and then stored at 20 °C until required. The LB solid medium was prepared by adding 15 g/l of bacteriological agar to the LB medium before autoclaving. After autoclaving, the medium was cooled to 40 °C and kanamycin (50 μg/ml) was added, poured and allowed to harden in 10-cm circular pre-sterilized plates, then stored at 4 °C until required.
Chemicals

All solvents were purchased from Fisher. All chemicals and reagents were purchased from Sigma-Aldrich, unless otherwise stated.

PCR

All the polymerase chain reactions (PCR) carried out in this research project, either to synthesize a new complementary strand of the offered DNA or for heat-controlled incubation purposes, were set up using a 96-well thermocycler (MJ Research PTC-200 DNA Engine) equipped with thermally controlled blocks that facilitate repeated cycling in an accurate pattern.

Electrophoresis and trans-illuminator

Electrophoresis of the nucleotides loaded into agarose gel wells (0.5-1 cm) was run under 70-80 V for 40-60 min and the fluorescence bands were visualized and recorded by a GDS 7500 UV trans illuminator (UVP) system.

Centrifuges

Eppendorf MiniSpin plus centrifuge from Beckman Coulter, with a maximum relative centrifugal force (rcf) of 14,000, was used for nucleic acids sedimentation where small volume samples (up to 2 ml) and high rotational speed (up to 13,400 rpm) were recommended. Universal 32R centrifuge (Hettich) with a maximum rcf of 9,418, was used for bacterial cells sedimentation where large volume samples (up to 85 ml) and low rotational speed (up to 5,000 rpm) were recommended.

Incubators

A Heraeus microbiological incubator was used for growing bacteria in Petri dishes at 37 °C. A MaxQ4000 bench-top incubator (Fisher) was used to maintain the growth of bacterial cultures that required continuous vertical shaking.

MicroPulser™ Electroporator

High voltage (up to 3,000 V) from a MicroPulser™ electroporator/Bio-Rad was used in the transformation of the modified A. tumefaciens electrocompetent cells.
3.3. General methods

Nucleic acid manipulation

Cassava genomic DNA isolation

The protocol for genomic cassava DNA leaf extraction was carried out based-on and modified after Huang et al., 2013. Fresh young leaves (~100 mg) of cassava plants were mixed with poly vinyl pyrrolidone (PVP, 2%) before being ground under liquid nitrogen. Preheated lysis buffer at 60 °C (1 ml of: Cetyl Trimethyl Ammonium Bromide (CTAB) (2%), tris-HCl (0.1 M, pH 8), PVP (2%), Ethylene Diamine Tetra Acetic acid (EDTA) (0.025 M), NaCl (0.002 M) and β-mercaptoethanol (2%)) was added to the homogenously mixed powder. The mixture was heated (65 °C, 30 min) with frequent shaking (every 3-5 min) before centrifuging (13,400 rpm, 10 min, 4 °C). The supernatant (600 µl) was transferred to a new Eppendorf tube and an equal volume of chloroform: iso-amyl alcohol (IAA) (24:1 v/v) was added to the supernatant and mixed thoroughly before centrifuging (5 min, 13,400 rpm, 4 °C). The supernatant (500 µl) was mixed by vortex with an equal volume of isopropanol, the mixture was cooled (-20 °C, 15 min), and then centrifuged (13,400 rpm, 5 min, 4 °C). Finally, the supernatant was discarded and the DNA-pellet was washed twice (1 ml of aqueous 75% ethanol then 1 ml of 100% ethanol), centrifuged (13,400 rpm, 15 min), the ethanol was discarded, and the pellet was dried (20 °C, 30 min). The DNA pellet was reconstituted in Milli-Q water (50 µl) and stored at -20 °C until required.

A. thaliana genomic DNA isolation

Genomic A. thaliana DNA leaf extraction was carried out based-on and modified after Liu et al., 1995. Fresh young leaves of 3-4 week old A. thaliana plants (~20 mg) were mixed with A. thaliana extraction buffer (200 µl of D-sorbitol (0.14 M), EDTA (0.022 M, pH 8), tris-HCl (0.22 M, pH 8), NaCl (0.8 M), CTAB (0.8%), and N-lauroyl sarcosine (0.1%)) before grinding using a plastic rod attached to a pillar drill machine. To maximize the extraction efficiency, the homogenised leaves were heated (65 °C, 5 min) before brief vortex mixing (30-60 s) with chloroform (100 µl), centrifuged (13,400 rpm, 5 min). The clear supernatant (170 µl) was mixed with isopropanol (70%, 150 µl), stored (20 °C, 15
min), then centrifuged (13,400 rpm, 20 min). Finally, the clear pellet was washed gently with ethanol (500 µl, 70% in aq.), centrifuged (13,400 rpm, 15 min), and then the supernatant was discarded and the pellet was dried (20 °C, 30 min). The DNA pellet was reconstituted in Milli-Q water (50 µl) and stored at -20 °C until required.

Cassava total RNA isolation

Cassava total RNA isolation was carried out based-on and modified after Chang et al., 1993. For leaf RNA isolation: young leaves (~100 mg) were picked and kept under liquid nitrogen. For root RNA: cassava root (6 months) was harvested from the glass house, cut into small cubes (~1 cm³) and distributed into 4 groups. One group was labelled as fresh (0 h), and the other groups were allowed to deteriorate for 24 h, 48 h, and 5 days in Petri dishes over the surface of filter paper under a controlled environment (24 °C, 50% humidity). Cassava tissues (~100 mg) from each group were ground using a chilled pestle and mortar under liquid nitrogen to minimize the degradation by RNase activity before vigorously mixing with a preheated extracting buffer (600 µl, 60 °C, CTAB (2%), PVP (2%), tris-HCl (0.1 M, pH 8), EDTA (0.025 M), NaCl (2 M) with β-mercaptoethanol (to make a final concentration of 2% added immediately before use). The homogenised tissues were stored (0 °C, 15 min) with frequent shaking (for 30 s every 5 min) before mixing with chloroform: IAA (500 µl, 24:1 v/v) solution. The mixture was then centrifuged (13,400 rpm, 10 min). The clear supernatant was mixed with an equal volume of chloroform:IAA (24:1 v/v) and then centrifuged (13,400 rpm, 10 min). The clear supernatant was mixed with an equal volume of isopropanol, LiCl (~150 µl of 10 M to make a final concentration of 2 M) was added. The mixed solution was cooled (4 °C, 16 h) and then centrifuged (13,400 rpm, 30 min). After discarding the supernatant, the pellet was treated with nuclease-free water (1 ml) and LiCl (250 µl, 10 M), cooled (0 °C, 3 h), and then centrifuged (13,400 rpm, 10 min). After discarding the supernatant, the pellet was treated with nuclease-free water (250 µl), sodium acetate (25 µl, 3 M) and absolute ethanol (1 ml), stored (~20 °C, 3 h) and then centrifuged (13,400 rpm, 10 min). Finally, the supernatant was discarded and the pellet re-suspended in nuclease-free water (50 µl) and stored at -80 °C until required.
**A. thaliana** total RNA isolation

Total RNA was isolated from fresh young *A. thaliana* leaves (~30 mg) using the instructions of a Spin or Vacuum (SV) Total RNA isolation system purchased from Promega. Leaves were ground under liquid nitrogen using a pre-chilled mortar and pestle and immediately transferred to an Eppendorf tube (1.5 ml) and mixed with the RNA lysis buffer (175 µl containing guanidine isothiocyanate (4 M), tris (0.01 M, pH 7.5), and β-mercaptoethanol (0.97%). RNA dilution buffer (350 µl) was added to the mixture, mixed and then centrifuged (13,400 rpm, 10 min). The clear lysate was mixed with ethanol (95%, 200 µl) and transferred to the spin-basket provided with the kit, centrifuged briefly (13,400 rpm, 1 min) and the eluate was discarded. RNA wash solution (600 µl containing potassium acetate (0.6 M), tris (10x10-3 M, pH 7.5), ethanol (60% in aqueous solution) was added to the spin-basket and centrifuged briefly (13,400 rpm, 1 min). Freshly prepared DNase incubation mix (50 µl) composed of yellow core buffer (40 µl containing tris (0.0225 M, pH 7.5), NaCl (1.125 M), yellow dye (0.0025%), MnCl₂ (5 µl, 0.09 M) and DNase (5 µl), was added to the basket’s membrane and stored (20 °C, 15 min). DNase stop solution (200 µl containing guanidine isothiocyanate (2 M), tris (4x10-3 M, pH 7.5, ethanol (57%) was added and centrifuged (13,400 rpm, 1 min). RNA wash solution (250 µl) was added and the mixture then centrifuged (13,400 rpm, 2 min). The spin-basket was transferred to an elution tube provided with the kit before nuclease-free water (100 µl) was added to elute the isolated RNA which was stored at -70 °C until required. [https://www.promega.com/-/media/files/resources/protocols/technical-manuals/0/sv-total-rna-isolation-system-protocol.pdf](https://www.promega.com/-/media/files/resources/protocols/technical-manuals/0/sv-total-rna-isolation-system-protocol.pdf) (accessed on 01.08.17).

**Total RNA purification**

Turbo™ DNase enzyme (Ambion, life technology) was used for the purpose of purifying the isolated RNA from residual genomic DNA and therefore inhibiting the genomic DNA templates from interfering with the vitro transcription reaction results. Following the manufacturer’s protocol, isolated RNA (35 µl) was mixed with 5x Turbo DNase buffer (5 µl), nuclease-free water (4 µl) and Turbo DNase (1 µl, 2 U/µl). The mixture was heated (37 °C, 30 min), the enzyme was deactivated with chloroform (50 µl) and then centrifuged (13,400 rpm, 3 min). The aqueous layer was mixed with sodium acetate (5 µl, 3 M) and ethanol (160 µl, 75% aqueous...
solution), cooled (-80 °C, 4 h), and then centrifuged (13,400 rpm, 20 min). After discarding the supernatant, the pellet was re-constituted in nuclease-free water (50 µl) and the purified RNA was stored at -80 °C until required.

**Complementary DNA (cDNA) synthesis**

Converting the purified RNA samples into their corresponding double-stranded (ds) DNA was performed using a High-Capacity cDNA Reverse Transcription (RT) kit (Applied Biosystems) and following the manufacturer’s protocol. RT reaction was performed by adding an equal volume of a purified RNA (10 µl contains 100-150 ng) to a freshly prepared RT master mix (RT buffer (2 µl), dNTP (0.8 µl, 0.1 M), random primers (2 µl), MultiScribe™ Reverse Transcriptase (1 µl), and nuclease-free water (4.2 µl)). The reaction was performed in a thermal cycler (25 °C for 10 min, 37 °C for 120 min, and 85 °C for 5 s) and then stored at -20 °C until required.

**Nucleic acid assessment**

The quality of the nucleic acids was assessed by measurement of the absorbance at 230 nm, 260 nm, and 280 nm using NanoDrop™ micro-volume spectrophotometer. A260/A280 of ~1.8 for DNA, and ~2.0 for RNA, indicates a sufficiently low level of protein contamination. Whilst A260/A230 indicates that the sample is free from residual solvent from the extraction process or residual guanidine used during the isolation and purification process. [www.nanodrop.com](http://www.nanodrop.com).

**Polymerase Chain Reaction (PCR)**

Each PCR was set up and performed in a Thermo Cycler PTC-200 (MJ Research) according to the manufacturer’s protocol. All oligonucleotide primer sequences and their working conditions are listed in Table 3.1.

**PCR for general genomic screening for the genes of interest**

For general genomic screening, the reaction mix contained *Taq* DNA polymerase enzyme (0.125 µl, 5 U/µl), template DNA (1 µl contains 10-100 ng) forward primer (1 µl, 10 x 10^{-6} M), reverse primer (1 µl, 10 x 10^{-6} M), and a freshly prepared master mix to give a final volume of 25 µl. The master mix contained: loading dye (3 µl with sucrose (40% w/v) and bromophenol blue (0.25%)), dNTPs
mix (2 µl, 0.01 M), 10x strength standard Taq polymerase buffer pH 8.3 (2.5 µl containing tris-HCl (0.01 M), MgCl₂ (0.0015 M) and KCl (0.05 M)), and Milli-Q water (14.375 µl). The accurate volume was measured using Eppendorf® micro-pipettes. The PCR programme was as follows: initial denaturation (94 °C, 2 min) followed by 30 cycles of denaturation (94 °C, 30 s), annealing (45-68 °C, 30 s), and elongation (72 °C, 1 min/kbp), then final extension (72 °C, 5 min). When the screening required bacterial culture as the template, denaturation required a longer time (5 min) to ease the releasing of the plasmid DNA from the cells.

PCR for specific gene amplification and gene expression study (RT-PCR)

For the purpose of cloning and amplifying the gene of interest, or to check the expression of the selected gene, the reaction mix contained Q5® High-Fidelity DNA polymerase enzyme (0.25 µl), the template cDNA (1.25 µl containing 100 ng), forward primer (1.25 µl, 10 µM), reverse primer (1.25 µl, 10 µM), and a freshly prepared master mix (dNTPs (2 µl, 0.01 M), 5x strength Q5 buffer (5 µl), and Milli-Q water (14.00 µl) to give a total reaction volume of 25 µl. The PCR reaction was programmed as follows: initial denaturation (94 °C, 2 min) followed by 30 cycles of denaturation (94 °C, 30 s), annealing (45-68 °C, 30 s), and elongation (72 °C, 40 s/kbp), then final extension (72 °C, 5 min).

Agarose gel electrophoresis

Estimation of the size of the nucleic acid in a PCR product by reference to molecular weight markers (0.1 and 1 kbp DNA ladders) was performed using agarose gel electrophoresis, where both the sample and the marker were run on the agarose gel simultaneously. The gel was prepared by adding agarose powder into TAE buffer pH 8.3 (tris-acetate (0.04 M), acetic acid (0.02 M) and EDTA (0.001 M)) to make a final concentration of 1.5% for best separation of nucleic acids (0.3-3.0 kbp). The agarose was dissolved by several microwave heat pulses (800W, 3 pulses of 30 s) then ethidium bromide was added (0.25 mg/ml) before the solution was cooled and solidified into special welled-gel trays (20 °C, 25 min). Electrophoresis was run with the gel-tray dipped in TAE buffer under 7-8 V/cm for 40-60 min as required. The stained nucleic acids moved according to their molecular weight and the bands were visualized under UV light. The images were documented using imaging software (Grab-IT 2.0) purchased from Synoptics.
Table 3.1.Designed PCR primer pairs and their working conditions. *att* sequences shown in italics.

<table>
<thead>
<tr>
<th>name</th>
<th>nucleotide sequences (5’-3’)</th>
<th>amplicon size (bp)</th>
<th>template</th>
<th>temp. (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M.e-CCoAOMT1 specific</td>
<td>&gt;Fs GTCAGTTGTCTGACTACACAA&lt;br&gt;Rev GAAACAGGATGAAGTGTCTGGGT</td>
<td>252</td>
<td>cassava DNA/cDNA</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>&gt;Fs GGAATGGTTCTGTGGTTG&lt;br&gt;Rev GTTCAATACACACACGTGGAT</td>
<td>351</td>
<td>cassava DNA/cDNA</td>
<td>56</td>
</tr>
<tr>
<td>M.e-EOMT specific</td>
<td>&gt;Fs TCTCTTCTTCCGCTTCTTCTTGAA&lt;br&gt;Rev TACAGCTCTCAACTTTGCA</td>
<td>370</td>
<td>cassava DNA/cDNA</td>
<td>54</td>
</tr>
<tr>
<td>M.e-CCoAOMT2 full</td>
<td>&gt;Fs CTGATCGCCCATCTACTCTCC&lt;br&gt;Rev GTTCAATACACACACGTGGAT</td>
<td>996</td>
<td>cassava cDNA</td>
<td>56</td>
</tr>
<tr>
<td>M.e-EOMT full</td>
<td>&gt;Fs TCTCTTCTTCCGCTTCTTCTTGAA&lt;br&gt;Rev CAATCGTGTGCGAGTTCCTCAT</td>
<td>1245</td>
<td>cassava cDNA</td>
<td>54</td>
</tr>
<tr>
<td>M.e-CCoAOMT2-attB</td>
<td>&gt;attB1 AAAAAAGCAGGGCTTCTGATCGCCCATCTACTCTCTC&lt;br&gt;attB2 AGAAAAGCTGGTGTTGCTCAATACACCGTGGAT</td>
<td>783+att seq</td>
<td>M.e-CCoAOMT2</td>
<td>56</td>
</tr>
<tr>
<td>M.e-EOMT-attB</td>
<td>&gt;attB1 AAAAAAGCAGGGCTTCTTCTTCCGCTTCTTCTTGAA&lt;br&gt;attB2 AGAAAAGCTGGTGTTGCTCAATACACCGTGGAT</td>
<td>1098+att seq</td>
<td>M.e-EOMT</td>
<td>54</td>
</tr>
<tr>
<td>attB adapter</td>
<td>&gt;attB1 GGGGACGATTTGTACAAAAAGCAAGGGT&lt;br&gt;attB2 GGGGACGATTTGTACAAAAAGCAAGGGT</td>
<td>gene+adapter</td>
<td>gene+12attB</td>
<td>n/a</td>
</tr>
<tr>
<td>M13</td>
<td>&gt;Fs GTAAAACGACGGCCAGT&lt;br&gt;Rev AACAGCTATGACCATG</td>
<td>n/a</td>
<td>M13 sequence in vector</td>
<td>n/a</td>
</tr>
<tr>
<td>pCam35S</td>
<td>&gt;Fs TGCAGTTAGGAGATACGAC&lt;br&gt;Rev TAAATCGTCAAGACCGGCA</td>
<td>n/a</td>
<td>over-expression vector</td>
<td>n/a</td>
</tr>
<tr>
<td>CCoAOMT-T-DNA</td>
<td>&gt;LP1 TAAAGGGCAACGAAACGATCC&lt;br&gt;RP1 GATGTTTTGTGCTGACTTCC</td>
<td>1211</td>
<td>CCoAOMT-T-DNA</td>
<td>54</td>
</tr>
<tr>
<td>EOMT-T-DNA</td>
<td>&gt;LP2 TGGAAACTCTCCTGGTGTTG&lt;br&gt;RP2 AATTCTTCTGATGGGATTCC</td>
<td>1000-1200</td>
<td>EOMT T-TDNA</td>
<td>54</td>
</tr>
<tr>
<td>Primer Set</td>
<td>Forward Primer</td>
<td>Reverse Primer</td>
<td>Length</td>
<td>Label</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------</td>
<td>----------------</td>
<td>--------</td>
<td>---------------------</td>
</tr>
<tr>
<td>RP1+LB1</td>
<td>&gt;RP1 GATGTTTTTGCTGTCACCTCCC</td>
<td>&gt;LB1 GCGTGGACCGCTTTGCTGCAACT</td>
<td>1141</td>
<td>CCoAOMT-TDNA</td>
</tr>
<tr>
<td></td>
<td>&gt;RP2 AATTCTTGATGGGATTCCC</td>
<td>&gt;LB1 GCGTGGACCGCTTTGCTGCAACT</td>
<td>504-804</td>
<td>EOMT T-TDNA</td>
</tr>
<tr>
<td>A.t-CCoAOMT</td>
<td>&gt;Fw CACAGAGAGAGAAAGAGAGAGA</td>
<td>&gt;Rev ACTGTGGACAAACTCAGAGTA</td>
<td>837</td>
<td>A.t-CCoAOMT</td>
</tr>
<tr>
<td>A.t-EOMT</td>
<td>&gt;Fw GCAGAGACACAATTAACCTCCGG</td>
<td>&gt;Rev AGAGCAGCAATGGAAACACC</td>
<td>362</td>
<td>A.t-EOMT</td>
</tr>
<tr>
<td>A.t-F6'H1</td>
<td>&gt;Fw CAAATCCAGCTGAAGTAACCG</td>
<td>&gt;Rev GAAGAATTTGAGTGCAGCAG</td>
<td>305</td>
<td>A.t-F6'H1</td>
</tr>
</tbody>
</table>
**A. thaliana** plant preparation

**Homozygous** *A. thaliana* lines verification

*A. thaliana* mutant seeds (T-DNA insertion Salk-lines) were purchased from NASC and screened to confirm the homozygosity of the selected mutant. According to the specifications and instructions from the Salk Institute, genomic DNA was isolated from young fresh *A. thaliana* leaves (2-3 weeks old) and two PCR reactions were run with two different pairs of primers to distinguish between wild type, heterozygous and homozygous lines. A left and right primer pair (LP and RP) was used to look for a wild-type band. A right and left border (LB1) primer pair (RP and LB1) was used to check the presence of T-DNA insertion. According to the Salk Institute definition, the homozygous line must show one single insertion band (RP and LB1 reaction), while those that show double bands (RP and LB1, RP and LP) are verified as heterozygous lines. Single band (RP and LP reaction) means no insertion was detected (Figure 3.5). [http://signal.salk.edu/cgi-bin/tdnaexpress](http://signal.salk.edu/cgi-bin/tdnaexpress) (accessed on 01.08.17). To confirm homozygosity, RT-PCR was performed using synthesized cDNA as a template, absence of expression-band clarifies non-functional gene or pseudogene that has no protein-coding ability (Vanin, 1985). Those lines revealed to be homozygous were used for further experiments.

---

**Fig. 3.5. Simplified scheme to show T-DNA insert and primers positions in the *A. thaliana* genomic DNA.** WT was confirmed after a single band of LP + RP reaction. A homozygous mutant was confirmed after a single band of RP+LB1 reaction. A heterozygous was confirmed after two bands for LP+RP, RP+LB1 reactions. After [http://signal.salk.edu/Help/iSectToolHelp.html](http://signal.salk.edu/Help/iSectToolHelp.html) (accessed on 01.08.17).
Producing double homozygous T-DNA insertion lines

A double mutant *A. thaliana* plant was created by reciprocal cross-pollination between two homozygous parent lines. A young flower was emasculated 48 h before being pollinated by gently opening the petals and the sepals of an immature flower bud (2-3 mm) from the mother plant and the anthers were carefully cut away, using a tiny forceps, to avoid any self-pollination. The stigma was allowed to ripen (48 h) before cross pollination with exogenous pollen by gently tapping the mature pollens from the other parent onto the mature stigma in which the papillae had extended and were therefore more receptive. The emasculation and pollination processes were undertaken using a 40X magnifier lens of a dissecting microscope (Figure 3.6). New pods were developed 2-3 days later, and the F1 generation seeds were ready to be collected after 14 days. The F1 plants were allowed to grow and self-pollinate to produce F2 generation plants, from which homozygous double-mutant plants were selected. Screening for the homozygous double-mutant plants among the F2 generation was carried out by running three PCR reactions for each isolated DNA using the same method described above.

Fig. 3.6. Cross breeding between two homozygous *A. thaliana* mutants. A shows the stigma of one parent plant after emasculation. B shows the mature pollen over the anther of the second parent. Images were taken under 40X magnifier lens.

Cassava (OMT) candidate genes preparation

Cassava genes detection, expression, and isolation

Homologous cassava genes coding for possible CCoAOMT activity were selected based on their sequence similarity with the reference *A. thaliana* gene published in NCBI [http://www.ncbi.nlm.nih.gov/](http://www.ncbi.nlm.nih.gov/). Homologous cassava and *A.
*Arabidopsis thaliana* genes coding for EOMT were selected based on their sequence similarity with the reference *Populus deltoids* gene which has been reported to exhibit methylation selectivity with esculetin (Kim et al., 2006). Primers were designed using Geneious® 5.3 to detect (screening PCR) and check the expression (RT-PCR) of the selected genes (Table 3.1). The PCR product was visualized under UV light and an image was saved using the UVP system. The genes that showed full expression in roots and leaves were used for further experiments. The full-length cDNA of the expressed cassava gene was amplified by a proof-reading polymerase enzyme (HF-Q5) from the synthesized cDNA in a PCR reaction using primers designed from the untranslated regions (5'-UTR and 3'-UTR). The PCR product was visualized under UV light and the image was saved using the UVP system.

**Cassava genes amplification**

The isolated cDNA for each OMT gene from cassava was amplified and cloned into a plasmid-construct for overexpression in *A. thaliana* mutants using TOPO® TA and GATEWAY® cloning technology as described above.

**Producing fully attB flanked PCR product for over-expression constructs**

Producing a fully attB flanked cDNA was performed in two PCR reactions. Firstly, by adding a 12-attB flank to the full-length amplified cDNA using HF-Q5 PCR to minimize the amplification errors. PCR was programmed as follows: initial denaturation (94 °C, 10 min), followed by 10 cycles of denaturation (94 °C, 30 s), annealing (40-60 °C, 1 min), then final extension (72 °C, 1 min). Secondly, by adding 25 attB adapter nucleotides to produce a fully attB flanked full-length cDNA. The PCR reaction was programmed as follows: initial denaturation (94 °C, 2 min) followed by 5 cycles of denaturation (94 °C, 30 s), annealing (45 °C, 30 s) then elongation (72 °C, 2 min), after that 20 cycles of denaturation (94 °C, 30 s), annealing (55 °C, 30 s), elongation (72 °C, 1.2 min), and finally a further extension step (72 °C, 5 min). The PCR product was visualized under UV light and the image was saved using the UVP system. All primer sequences are listed in Table 3.1.

**cDNA gel purification**

The fully attB flanked cDNA PCR product was electrophoresed in adjusted conditions for better separation of the cDNA fragments (agarose gel 0.8%, 70 V for
Following the manufacturer’s instructions, QIAquick® Gel extraction kit (QIAGEN) was used. The DNA band was excised under UV light using a sterile blade, and transferred immediately into a clean labelled Eppendorf tube containing QG buffer (300 μl/100 mg of guanidine thiocyanate (Gu-SCN) (5.5 M), tris-HCl (0.02 M, pH 6.6). The mixture was then stored (50 °C, 10 min) with frequent shaking (for 10 s every 2 min) to facilitate gel solubilizing. Then isopropanol (100 μl/100 mg) was added and mixed gently, by pipetting, before transferring into a QIAGEN spin column, and centrifuged (13,400 rpm, 1 min). QG buffer (500 μl) was then added and centrifuged (13,400 rpm, 1 min). Finally, PE washing buffer (750 μl of tris-HCl (0.1 M, pH 7.5, aqueous ethanol (80%)) was added to the spin column, centrifuged (13,400 rpm, 1 min), and the column was transferred to a sterile Eppendorf tube before the extracted cDNA was eluted with Milli-Q water (40 μl) and stored at -20 °C until required.

Adenylation and TOPO® TA cloning reaction

Cloning the fully attB-flanked cDNA into pCR™2.1-TOPO® plasmid vector provided a highly efficient ligation between the 3’-deoxyadenosine (A) end in the desired insert and the single 3’-deoxythymidine (T) in the linearized TOPO vector (Figure 3.7). As the HF-Q5 polymerase enzyme used in the PCR amplification reaction did not possess the overhang 3’-deoxyadenosine (A), adenylation of the desired DNA insert is a requirement. An A-tailing reaction was therefore performed by adding Taq DNA polymerase (0.2 μl) to the blunt-ended PCR product (~240 ng), Taq standard buffer (5 μl), dNTPs (0.01 M, 1 μl), and then Milli-Q water up to 50 μl. Adenylation was allowed to occur in a thermal cycler (72 °C, 20 min) and the reaction product was used immediately for subsequent TA cloning. The freshly-adenylated DNA was ligated into the pCR™2.1-TOPO® plasmid by forming hydrogen bonds with the overhanging 3’-deoxythymidine (T) residues in the vector. According to the manufacturer’s specification, the ligation reaction used fresh PCR product (3 μl) and TOPO® vector (1 μl) with the addition of a salt solution (1 μl of NaCl (1.2 M) and MgCl₂ (0.06 M)) to enhance the reaction efficacy and Milli-Q water (1 μl). After gentle mixing and short storing (20 °C, 5 min), the ligated product was used to transform highly competent E. coli cells.
Fig. 3.7. A-T hydrogen bonds. Hydrogen bonds were formed between one nucleic acid and the other nucleic acid in Watson-Crick pairing.

Transformation reaction

High efficiency chemically competent 10-Beta E. coli cells (50 µl) were gently mixed with the TA cloning reaction product (2 µl), stored in ice (30 min), before being heat-shocked (42 °C, 30 s). The transformed cells were allowed to rest in ice (2 min) before a nutrient S.O.C. medium (950 µl) was added. The cells were then allowed to grow (37 °C, 60 min) with continuous vertical shaking (200 rpm). Then aliquots (100 µl) of the bacterial culture were spread over pre-prepared selective LB agar plates containing kanamycin (50 µg/ml), and X-gal (5-bromo-4-chloro-3-indoyl-β-D-galacto-pyranoside, 40 µl) was spread over to allow blue/white colour selection. Colonies were allowed to develop (37 °C, 16 h). A white colony, in which β-galactosidase enzyme activity was rescued by the α-complementation of the vector, was picked and allowed to grow (37 °C, 16 h, 200 rpm) in a selective LB medium (3 ml). The grown culture (1 µl) was used as a template for colony PCR reaction using primers designed from the untranslated regions (UTR) to screen for the successful transformation. The DNA was purified from the plasmid (pDNA) and its identity was further confirmed by gene sequencing. The efficiency of the cells is their ability to take up and express the extracellular DNA. This was evaluated according to the number of successfully transformed colonies per the amount of DNA used.

Plasmid-DNA purification

Following the manufacturer’s specification for QIAprep® Miniprep (QIAGEN), the bacterial culture (3 ml) was centrifuged (10,000 rpm, 10 min, 4 °C) and the pellet was suspended in buffer P1 (250 µl, tris-HCl (0.05 M, pH 8), EDTA (0.01 M), RNase (100 µg/ml)) and mixed thoroughly before the lysing P2 buffer (250 µl,
NaOH (0.2 M), SDS (1%), and N3 neutralizing buffer (350 µl, guanidine-HCl (4.2 M), potassium acetate (0.9 M), pH 4.8), were added consecutively. The mixture was centrifuged (13,400, 10 min), and the clear supernatant (170 µl) was transferred into a new QIAprep spin column before being centrifuged (13,400 rpm, 1 min). Then the column was washed with PE buffer (750 µl), centrifuged (13,400 rpm, 1 min), and the outflow was discarded. The purified pDNA was eluted with Milli-Q water (50 µl) and stored at -20 °C until required.

Purified pDNA sequencing

The purified pDNA (15 µl of ~50-100 ng) was sequenced in both directions by the Eurofin Company using the universal M13 primers (15 µl, 0.01 M), https://www.eurofinsgenomics.eu (accessed on 01.08.2017). The sequences obtained were aligned with the reference gene sequence using Geneious® software. M13 forward and reverse primer sequences are listed in Table 3.1.

Gateway® cloning

The purified plasmid DNA was further cloned into the universal Gateway® technology to ensure an efficient shuttle into multiple vectors regulated with promotors compatible with A. tumefaciens. Two successive site-specific recombination reactions were performed following the manufacturer’s specification (Invitrogen).

a. Building a Gateway® entry clone

The first directional recombination reaction was performed to produce a Gateway® entry clone. The entry clone is the product of ligating the attB flanked cDNA with the pDNOR™ plasmid (Invitrogen®) which is doubly overhanged with the attP sequence. The ligation reaction is mediated by a Gateway® BP Clonase™ II enzyme mix (Invitrogen® proprietary formulation). The attL flanked cDNA (referred to as pENTER™) is the product in which the desired cDNA replaced the bacterial lethal ccdB sequence (counter-selectable marker), and allowed the survival of correctly transformed E. coli cells on zeocin/LB solid media. A simplified scheme for the BP recombination reaction is shown in Figure 3.8.
Fig. 3.8. **pENTER™ clone construction.** The fully attR flanked cDNA replaces the lethal ccdB sequence to produce a pENTER™ plasmid. The reaction is mediated by Gateway® BP Clonase™ II enzyme mix.

Following the manufacturer’s specifications, the fully attB flanked DNA fragments (15-150 ng) was mixed with pDNOR® vector (1 µl), and TE buffer (pH 8.0, up to 8 µl). Then freshly thawed BP Clonase™ enzyme mix (2 µl) was added, and the reaction mixture was stored (25 ºC, 16 h). The reaction was terminated by the addition of proteinase K solution (1 µl, 37 ºC, 10 min). The reaction product (2 µl) was transformed into competent E. coli cells (50 µl). The transformed E. coli cells were allowed to grow and produce colonies (37 ºC, 16 h) on selective solid agar medium containing the antibiotic zeocin (50 µg/ml). Single colonies were cultured, the plasmid DNA was purified, and colony PCR was performed to screen for the correct transformation. Gene sequencing, performed by Eurofin, was used for confirmation.

**Producing a Gateway® overexpression clone**

The second directional recombination reaction was performed to produce a Gateway® overexpression clone. This clone is the product of ligating the attL flanked cDNA with the modified pCAMBIA 1305.1 plasmid vector which is doubly overhanded with attR sequence. The ligation reaction is mediated by LR Clonase™ enzyme mix (Invitrogen® proprietary formulation). The attB flanked cDNA (referred to as overexpression clone) is the product in which the desired cDNA replaced the bacterial lethal ccdB sequence, and allowed the survival of the transformed E. coli cells on kanamycin/LB solid media. A simplified scheme for LR recombination reaction is shown in Figure 3.9.
**Fig. 3.9. Overexpression clone construction.** The *attL* flanked cDNA replaces the lethal *ccdB* sequence in pCAMBIA 1305.1 to produce an overexpression clone. The reaction is mediated by LR Clonase™ II enzyme mix.

Following the manufacturer’s specification, the doubly *attL*-flanked DNA fragment (50-150 ng) was mixed with the modified pCAMBIA 1305.1 vector (1 µl), and TE buffer (pH 8.0, up to 8 µl). Then, the freshly thawed LR Clonase™ II enzyme mix (2 µl) was added, and the reaction mixture was stored (25 ºC, 16 h). The reaction was terminated by the addition of proteinase K solution (1 µl, 37 ºC, 10 min). The reaction product (2 µl) was transformed into competent *E. coli* cells (50 µl). The transformed *E. coli* cells were allowed to grow and produce colonies (37 ºC, 16 h) on selective solid agar medium (50 µg/ml kanamycin). Single colonies were cultured, and the plasmid DNA was purified. Colony PCR was performed to screen for the correct transformation. Gene sequencing, performed by Eurofin, was used for confirmation. *A. thaliana* β-glucuronidase gene (pENTER-gus) purchased from NEB was used as a positive control on a separate selective plate for each cloning reaction. This served as an accurate measure of the life of the clonase enzyme as well as of experimental error.

**A. tumefaciens** electrocompetent cells

*Agrobacterium* is widely used in mediating different plant transformations (Stanton, 2003). The *Agrobacterium tumefaciens* (GV3101) wild-type cells were modified to mediate transforming the prepared cassava candidate genes into T-DNA insert *Arabidopsis thaliana* plants.

*Agrobacterium* cells were rendered electrocompetent following Hofgen and Willmitzer’s 1988 protocol. Wild-type cells were allowed to grow in a selective 2YT medium (5 ml) with vertical shaking (28 ºC, 200 rpm, 16 h). The selective 2YT medium contains: tryptone pectone (1.6%), yeast extract (1.0%) and NaCl (0.5%), and the pH was adjusted to 7 using aqueous NaOH solution (0.1 M). The grown
culture was scaled-up using fresh 2 YT medium to make a final volume of 100 ml, and allowed to grow until OD$_{600}$ = 0.9-1.0, under the same stated conditions. The cells were harvested by centrifuging (2500 rpm, 20 min, 4 °C) and washed twice with pre-chilled sterile Milli-Q water (20 ml). The washed cells was re-suspended in an aqueous glycerol solution (10 ml of 10%). The cells were harvested, re-suspended in an aqueous glycerol solution (400 μl of 10%), divided into small aliquots (40 μl), and stored at -80 °C until required (Hofgen and Willmitzer, 1988).

**Electro-transformation of cassava genes into *A. tumefaciens* competent cells**

The purified cassava cDNA (15 ng) was added to the freshly thawed competent *A. tumefaciens* cells (40 μl) and electrically shocked in a MicroPulser™ electroporator (2.5 kV, 5 s). Selective LB medium (1 ml) which contains: gentamycin (25 μg/ml), rifampicin (30 μg/ml) and kanamycin (50 μg/ml), was added immediately to the cells. Then, the cell culture was allowed to grow with shaking (28 °C, 200 rpm, 3 h). The cells (100 μl) were allowed to grow and develop colonies (28 °C, 48 h) on selective solid agar medium containing the same concentrations of antibiotics. Single colonies were used to inoculate selective LB medium (5 ml), and then allowed to grow (28 °C, 200 rpm, 16 h). The grown culture was scaled-up (100 ml), and allowed to grow until OD$_{600}$ = 0.9-1.0 (28 °C, 200 rpm, 16 h). The transformed cells were harvested and suspended in freshly pre-prepared aqueous buffer (sucrose 5% and MgCl$_2$ (0.01 M)). The density of the cells was optimized (OD$_{600}$ = 0.7). The surfactant Silwet L-77 (0.05% v/v) was added to increase plant transformation efficacy by facilitating the transgenic bacterial infection of plant cells (Mano et al., 2014).

**A. thaliana transformation**

The transformation of the selected cassava into T-DNA inserted *A. thaliana* plants was mediated by *A. tumefaciens* cells following Clough and Bent (1998) floral dip protocol. The *A. thaliana* immature flower clusters (4-5 week old plants) were twice dipped (3 s) with gentle agitation in the buffer contains the transgenic *A. tumefaciens* cells (20 s in between) after clipping out the already formed siliques. The treated plants were kept in the dark and damp conditions by covering the whole plant with black plastic bags (16-24 h). Then, the plants were allowed to grow and to develop seeds for 3 more weeks (22 °C). The harvested seeds (~3000, 10 mg)
were surface sterilised as stated before, and then sown on selective solid agar plates (hygromycin 30 µg/ml) to allow screening for the positive transformants. The plates were stratified and germinated in the dark (4 °C for 24 h and then 22 °C for 4 days) before being allowed to grow (2 weeks, 22 °C). The plants which developed longer hypocotyl (7-8 mm), which is the stem part that comes directly above the roots, were considered to be transgenic (Zhang et al., 2006). The genomic DNA was isolated from the transgenic plants, PCR reaction was performed to confirm cassava gene transformation. Successful transgenic plants were allowed to grow in soil and develop seeds for further experiments.

Quantification of hydroxycoumarins in wild type and transgenic plants using LC/MS

Under the same conditions stated in chapter 2, wild type and transgenic A. thaliana plants were allowed to grow vertically in agar plates in aseptic conditions for 4-5 weeks. The roots were harvested, dried (55 °C, 30 min), homogenized under liquid nitrogen using a pestle and mortar or manually using a clean plastic rod. The dried root tissues (10 mg) were extracted with methanol (2 ml) containing 4-methylumbelliferone (100 ng/ml) as an internal standard. The mixture was shaken (20 °C, 16 h), and then syringe-filtered (diameter 13 mm, pore size 0.2 µm). Samples were lyophilized, resuspended in methanol (200 µl) containing scoparone (100 ng/ml) as a second internal standard, and then syringe-filtered (diameter 13 mm, pore size 0.2 µm). The two internal standards were added to improve the data reliability. The hydroxycoumarins of interest: scopoletin, esculetin, and their corresponding β-glycosides scopolin and esculin, were analysed by LC-MS using a reverse phase C18 column, in a gradient elution system of acidified methanol (0.1% formic acid) and acidified water (0.1% formic acid) in a total run time of 10 min as stated in chapter 2. Detection was performed by HR-TOF MS for scopoletin and its β-glycoside scopolin using ESI in positive-ion mode. The detected mass peaks of the compounds of interest [M+H]⁺ and [M+Na]⁺ were typically matched with their corresponding theoretical values within a 5 ppm mass error range. The theoretical m/z ratios are: scopoletin C₁₀H₉O₄ 193.0495, C₁₀H₉O₄Na 215.0315, scopolin C₁₆H₁₉O₉ 355.1024, C₁₆H₁₈O₉Na 377.0843.
3.4. Results and discussion

Characterization of CCoAOMT and EOMT genes which could be involved in the last steps of scopoletin biosynthesis

Nucleotide sequences encoding the genes of interest were retrieved from the NCBI database. Multiple sequence alignments for their corresponding amino acids were performed to generate neighbour-joining phylogenetic trees using Geneious® V 5.3 software. These phylogenetic trees were designed to find homologous cassava genes related to the reference genes. Genes laying within the same cluster show similarity in their evolution history (Gori et al., 2016). However, for sequence similarity with the reference gene, other parameters were used according to NCBI definitions, the expect value (E-value) which describes the random background noise, and the score value which indicates the resemblance between the gene and the reference. Based on these two parameters, the genes with higher score and lower E-values were chosen. Although there is no clear range of acceptable values, the score values of more than 200 and E-values of not more than 1 x 10^-50 were the default parameters used to identify the similar genes.

For CCoAOMT, two cassava genes: cassava 4.1_011832 named M.e-CCoAOMT1 (score 453.4 and E-value 7.7 x10^-162) and cassava 4.1_014783 named M.e-CCoAOMT2 (score 4464 and E-value 4.7 x 10^-159) show the highest similarities to the A. thaliana reference gene (AT4G34050.2) named A.t-CCoAOMT. The other cassava genes, found in other clusters, show weaker relation to the reference, which has been chosen after direct NCBI-key word enquiry for the common name. The phylogenetic tree, constructed using Geneious®, for the relation between the reference A.t-CCoAOMT and cassava OMT genes is shown in Figure 3.10.

For EOMT, two cassava genes: cassava 4.1_010187 named M.e-EOMT1 (score 421.8 and E-value 6.1 x 10^-146) and cassava 4.1_010203 named M.e-EOMT2 (score 417.5 and E-value 2.5 x 10^-144), and the A. thaliana gene (ATG54160.1) named A.t-EOMT showed the highest similarity to the poplar (Populus deltoids) reference gene (Potri.014G106600). The poplar gene was used as a reference to assign the cassava and A. thaliana homologous EOMT genes, on the basis of the published in vivo experimental results, after recombinant expression in E. coli, showing its activity in transferring a methyl group to esculetin to form scopoletin, iso-scopoletin, and scoparone (Kim et al., 2006).
Fig. 3.10. Phylogenetic tree (neighbour-joining) illustrating amino acid sequence similarity of CCoAOMT cassava genes to the *A. thaliana* reference. The reference A.t-CCoAOMT (AT4G34050) is in red. Cassava4.1_011832 was chosen and renamed to M.e-CCoAOMT1 and cassava4.1_014783 was chosen and renamed to M.e-CCoAOMT2. The bootstrap percentage between the reference gene and both cassava4.1_011832 and cassava4.1_014783 genes is 100% after 10,000 replicates.

The cassava 4.1_031311 lays in the same distance as the chosen genes with the reference gene, but it was not selected because it has lower score and higher E-values (score 330 and E-value 7.8 x 10^{-110}) than the selected cassava genes. The phylogenetic tree, constructed using Geneious®, showing the relation between the reference poplar OMT, cassava OMTs, and *A. thaliana* OMTs genes is shown in Figure 3.11.
Fig. 3.11. Phylogenetic tree illustrating amino acid sequence similarity between EOMT cassava and *A. thaliana* genes in reference to poplar OMT gene. The reference Potri.014G106600.1 is in red. Cassava4.1_010187 was chosen and renamed to M.e-EOMT1; cassava4.1_010203 was chosen and renamed to M.e-EOMT2; and *A. thaliana* AT5G54160.1 was chosen and renamed to A.t-EOMT. The bootstrap percentage between the reference gene and the chosen genes is 73.5% after 10,000 replicates.

Although M.e-EOMT2 (cassava 4.1_010203) showed a high similarity to the reference gene, it was not used for downstream experiments due to the difficulty in designing specific gene primers to amplify the gene. This might be due to the absence of the 5′-UTR region and the incomplete sequences of the gene retrieved from the Phytozome 10 database (2014). The genomic sequence for M.e-EOMT2 is shown in Appendix 1.
The phylogenetic trees (Figures 3.10 and 3.11) were constructed early in the research project (2014), and the genes used for downstream experiments were chosen based on it. As the completion of the cassava genome is still a work in progress, new cassava genes have been discovered and published in Phytozone research tool during this project (to 2017). Accordingly, the newest phylogenetic trees was constructed. The update in the genome has no effect on the relation between the reference A.t-CCoAOMT (AT4G34050.2) and the chosen M.e-CCoAOMT genes (cassava4.1_011832 and cassava4.1_014783), therefore the updated phylogenetic tree is shown in Figure 3.10.

Fig. 3.12. New phylogenetic tree illustrating amino acid sequence similarity between EOMT cassava and A. thaliana genes in reference to poplar OMT gene. The reference Potri.014G106600.1 is in red. The chosen genes were shown in green. The bootstrap percentage between the reference gene and the chosen genes is 73.5% after 10,000 replicates.
For the EOMT phylogenetic tree, the update in the cassava genome added new cassava genes laying in the same distance (bootstrap 73.5%) to the reference poplar genes (Potri.014G106600.1) with the chosen candidates. The updated phylogenetic tree (2017) for the relation between the reference poplar OMT with both the cassava and A. thaliana OMTs is shown in Figure 3.12. Anyhow, the chosen genes cassava4.1_010187, and cassava4.1_010203 are still the best candidates, as they are closely related, laying in the same clusters with the A. thaliana genes (bootstrap 100%) AT1G33030.1 and AT5G54160.1 (for the last 2) respectively. The resemblance with the A. thaliana genes has a significant importance in the downstream complementation experiments. Therefore, two cassava CCoAOMT genes M.e-CCoAOMT1 (cassava 4.1_011832) and M.e-CCoAOMT2 (cassava 4.1_014783), and one cassava EOMT gene M.e-EOMT (cassava 4.1_010187) were selected, isolated and used for further experiments.

Detection, isolation and expression study of the selected cassava genes

To confirm the presence of the selected genes in the cassava genome, PCR screening reactions were performed using the isolated cassava genomic DNA as a template. The primer pairs M.e-CCoAOMT1 specific, M.e-CCoAOMT2 specific, and M.e-EOMT specific were used in the screening reactions for the genes M.e-CCoAOMT1, M.e-CCoAOMT2, and M.e-EOMT respectively. All the primer sequences and their working conditions are listed in Table 3.1. Gel electrophoresis was run for each PCR product and a band with the correct size measured using a 100 bp DNA ladder as a molecular weight marker was visualised which revealed the presence of the three selected cassava genes in the cassava genome. Furthermore, the detected genes were subjected to RT-PCR with the synthesized cassava cDNA as a template, using the same detecting primer pairs, and under the same RT-PCR reaction conditions as stated above (section 3.3), to elucidate their expression in leaves and root tissues during various time-points of PPD (fresh, 48 h, and 5 days).

Both M.e-CCoAOMT2 and M.e-EOMT were fully expressed in leaves and roots. M.e-CCoAOMT1 is not expressed neither in leaves nor in roots, indicating non-functionality which means its inability to code protein despite the high similarity in sequences with the functional M.e-CCoAOMT2 gene. This might be due to accumulation of multiple mutations that renders it to a pseudo non-translated gene (Vanin, 1985). DNA bands amplified from the cDNA with the correct expected size
revealed the expression of functional genes in leaves (Figure 3.13) and roots during the PPD process (Figure 3.14).

![Expression profile in cassava leaves](image)

**Fig. 3.13. Expression profile in cassava leaves.** Gel image for the expression of selected cassava genes using 100 bp DNA ladder as molecular marker. Each reaction was performed twice with two different synthesized cDNA samples. M.e-CCoAOMT1 (lanes 1 and 6) no bands reveal no expression, M.e-EOMT (lanes 2 and 5), and M.e-CCoAOMT2 (lanes 3 and 4).

![Expression profile in fresh cassava roots and during PPD](image)

**Fig. 3.14. Expression profile in fresh cassava roots and during PPD.** M.e-CCoAOMT1 (lane 1 = fresh, lane 2 = 48 h, and lane 3 = 5 days). M.e-EOMT (lane 4 = fresh, lane 5 = 48 h, and lane 6 = 5 days). M.e-CCoAOMT2 (lane 7 = fresh, lane 8 = 48 h, and lane 9 = 5 days). Positive control (lane + genomic DNA).

Although the strength of gene expression is not entirely dependent on the band’s strength (Vanderschuren et al., 2014), M.e-CCoAOMT expression was significantly upregulated after root harvesting and during PPD signalling. The functionality of both genes during PPD suggests they may have roles in PPD. The DNA bands were purified from the agarose gel using QIAquick® Gel extraction kit (QIAGEN), amplified, and cloned into multiple vectors to be consistent with the
plant transformation carrier, *Agrobacterium tumefaciens*, to ensure the success of the gene transformation into the corresponding *A. thaliana* mutants.

Isolating full-length cDNA

Obtaining the complete coding regions of the selected genes was performed using HF-Q5 polymerase to minimize the chance of amplification errors. PCR was performed using primers flanking the functional genes (UTR regions) and the synthesized cassava cDNA was used as a template under the same reaction conditions as stated above (section 3.3). For M.e-CCoAOMT1 amplification, PCR was performed using primer pair M.e-CCoAOMT (full) and for M.e-EOMT amplification, PCR was performed using primer pair M.e-EOMT (full). The size of the isolated cDNA was measured using a 1 kbp DNA ladder as a molecular weight marker, for M.e-CCoAOMT2 a full-length cDNA of 996 bp, and for M.e-EOMT a full-length cDNA of 1245 bp were isolated. The gel image is shown in Figure 3.15.

![Gel Image](image_url)

Fig. 3.15. Full-length cDNA. M.e-CCoAOMT2 with an expected size of 996 bp (lane 1), M.e-EOMT with an expected size of 1245 bp (lanes 2 and 4 with different cDNA) and positive control (lane 3).

Cassava candidate genes were cloned into multiple Gateway® vectors

Fully attB-flanked DNA was produced in two stages. Firstly, adding 12-attB sequences. For M.e-CCoAOMT2, PCR reaction was performed using M.e-CCoAOMT2-attB primer pair, and for M.e-EOMT, PCR reaction was performed using M.e-EOMT-attB primer pair. Secondly, by adding 25 attB adapter nucleotides to produce a fully attB flanked full-length cDNA, using attB adapter primer pair under the same PCR reaction conditions as stated above (section 3.3). All primer pairs are listed in Table 3.1. The reason for adding the full adapter in two stages is
to optimize the primers physical properties by not exceeding the 12-22 bp optimal oligonucleotide sequence length. The attB-wrapped cDNA was then amplified into pCR2.1®-TOPO® TA vector after a simple adenylation reaction, getting the benefit of Taq-polymerase’s ability to add terminal adenine (A) nucleotides to the 3’-ends. This allows the topoisomerase I to build a hydrogen bond with the thymidine (T)-over-hanging vector to enable TA cloning. The ligation product was then transformed into competent 10-Beta E. coli cells which were allowed to grow and produce colonies on a selective solid media (kanamycin 50 µg/ml). The antibiotic selectivity coupled with the blue/white colour screening facilitate the identification of a successfully cloned colony. A white colony was picked and allowed to grow in kanamycin-LB media, colony PCR was performed with the attB adapter primers to check the presence of the gene and, for double confirmation, the isolated plasmid DNA was sequenced using the universal M13 primers by Eurofin.

Preliminary cloning into TOPO plasmid non-expressing vector, eases gene manipulation and affords overproduction of the DNA fragments for sequencing confirmation, protein purification, and later cloning in expression vectors (Hartley, 2006). The purified TOPO pDNA was cloned into pDNOR™ vector in a BP Clonase™ mediated recombination reaction between the attB-flanked insert and attP-flanked pDNOR™ vector to build pENTER™ construct. The pENTER™ serves as stage one Gateway® reaction product. The recombination product was then transformed, plated, and allowed to grow and produce colonies on selective solid media (50 µg/ml zeocin). Cultured pENTER™, which survived the antibiotic selectivity, was PCR screened using attB primers and the identity of the insert genes was accurately confirmed by sequencing in both directions by Eurofin using the universal M13 primers (Table 3.1). Purified plasmid DNA was used as a template to create an overexpression clone.

The recombination product was then transformed, plated, and allowed to grow and produce colonies on selective solid media (kanamycin 50 µg/ml). Cultured overexpression construct, which survived the antibiotic selectivity, was PCR screened using attB primers and the identity of the insert genes was accurately confirmed by sequencing in both directions using pCam 35S primer pair by Eurofin (Table 3.1). The pCAMBIA construct carrying M.e-CCoAOMT2 gene was used to transform A. thaliana CCoAOMT T-DNA insert single mutant. The pCAMBIA
construct carrying M.e-EOMT1 gene was used to transform *A. thaliana* EOMT T-DNA insert single mutant.

**Homozygous *A. thaliana* mutants were transformed with cassava genes**

Cassava candidate genes, which are under the control of mosaic 35S promotor, were transformed into *A. tumefaciens* by applying electric pulses across the electrocompetent Agrobacterium cells (1-2 mm) rendering their cell membranes more permeable to host the exogenous DNA (Sugar and Neumann, 1984). The transformed Agrobacterium cells were cultured and plated on selective agar media containing the antibiotics gentamycin (25 µg/ml), rifampicin (30 µg/ml), and kanamycin (50 µg/ml). The colonies which survived were considered to be transgenic due to the conferred resistance to kanamycin offered by the carrying vector pCAMBIA. Wild type Agrobacterium cells were used as control. For double confirmation, cultured Agrobacterium was PCR screened using attB adapters (Table 3.1) and a gel band with the correct size was obtained.

Agrobacterium cells carrying the correct insert were used to transform homozygous *A. thaliana* mutant using the in planta floral dip transformation method modified after Clough and Bent (1998). This provides an easy, less time-consuming genetic modulation of *A. thaliana* plants and helps to avoid complicated in vitro tissue culture methods (Wiktorak-Smagur et al., 2009). *A. thaliana* T-DNA insert lines (Salk lines) are Salk-055103 renamed to CCoAOMT-T-DNA, Salk 135290 renamed to EOMT-T-DNA which were chosen in reference to the *A. thaliana* genes A.t-CCoAOMT and A.t-EOMT respectively from the Salk Institution website [http://signal.salk.edu/cgi-bin/tdnaexpress](http://signal.salk.edu/cgi-bin/tdnaexpress) (accessed on 01.08.17). The transformed *A. thaliana* plants were allowed to grow in normal conditions and the seeds were harvested and screened by plating them on selective agar plates (30 µg/ml hygromycin). The transformed seeds maintained the hygromycin selectivity offered by pCAMBIA and succeeded in producing longer hypocotyl (7-8 mm) (Figure 3.16). However, the low, but detectable successful transformants (12-15/3000 seeds) is due mainly to the poor efficiency of the Agrobacterium-mediated transformation method (Feldmann and Zoecon, 1987; Valvekens et al., 1988).
Before being transformed with cassava candidate genes, Salk lines homozygosity were verified using PCR. The line is considered homozygous when the isolated DNA showed a PCR product (500-800 bp) for LB1 and RP reaction which indicates the presence of a T-DNA insertion sequence, but no product for LP and RP reaction which means that the whole gene is getting too large to be amplified under the same reaction conditions. To ensure that the function of the knocked-out gene is inactivated, no gel bands were visualised after RT-PCR reactions (Figure 3.17) using A.t-CCoAOMT primer pair for CCoAOMT-T-DNA line and using A.t-EOMT primer pair for EOMT-T-DNA line (Table 3.1).

Single homozygous mutant plants were crossed (A.t-CCoAOMT X A.t-EOMT) and the product seeds were allowed to grow in soil to produce F1 heterozygous plants. The F1 plants were allowed to grow in soil and self-pollinated to produce F2 plants. Double homozygous lines were selected. Function inactivation of both genes was confirmed via RT-PCR.
The selection of homozygous double mutant plants. DNA produced PCR product for RP1+LB1 and RP2+LB1 indicates the presence of T-DNA sequence in both CCoAOMT and EOMT genes respectively. No band was detected for RP1+LP1 or RP2+LP2 (not shown). Plants were confirmed to be homozygous double mutants.

Quantification of coumarin levels in wild-type and transgenic *A. thaliana* plants

Hydroxycoumarin levels in the methanolic extracts of the wild-type and transgenic *A. thaliana* roots (10 mg) were separately analysed using LC-MS. Their accumulations in the roots were quantified by measuring the LC peak areas of the coumarins of interest relative to the LC peak of the added internal standard 4-methylumbelliferone (4-MU) to avoid underestimation of the coumarin levels due to operator or technical errors.

A significant reduction ($P < 0.05$) of $\sim 60\%$ in scopoletin level in the root extract of A.t-CCoAOMT mutant plant was measured compared to its level in the wild type, while only a slight reduction of $\sim 20\%$ was measured in the A.t-EOMT mutant roots. No significant differences ($P = 0.33$) between scopoletin levels in the A.t-EOMT and the double mutant (CCoAOMT x EOMT) root extracts were measured (Figure 3.18). The drop in scopoletin concentration after knocking out the A.t-CCoAOMT gene in *A. thaliana* is compatible with previous findings where a $\sim 30\%$ decrease in scopoletin was reported in the same mutant line (Kai *et al*., 2006). The results obtained clearly confirmed that scopoletin biosynthesis via esculetin O-methylation makes a lower contribution to the overall accumulation of scopoletin in *A. thaliana* roots. Even though it was expected that double mutation will further
decrease the level of scopoletin, the obtained results showed an unexplained accumulation pattern which could result from activation of other O-methyltransferase genes to compensate for the drop in scopoletin level, or other scopoletin biosynthesis routes could be involved.

**Fig. 3.18.** Scopoletin levels in wild-type and transgenic *A. thaliana* plants. Histograms show the relative accumulation of scopoletin in the root extracts to the internal standard 4-methylumbelliferone.

**Fig. 3.19.** Scopolin levels in wild-type and transgenic *A. thaliana* plants. Histograms show the relative accumulation of scopolin in the root extracts to the internal standard 4-methylumbelliferone.
Scopolin showed the same reduced biosynthesis (~40%) as scopoletin in the A.t-CCoAOMT root extract compared to the wild type (Figure 3.19). This significant reduction ($P < 0.05$) in the measured concentration was expected as scopolin is biosynthesized by adding a glucose unit to scopoletin by the action of scopoletin glucosyltransferase (Ahn et al., 2010). Also, in the root extract of A.t-EOMT mutant plants, reduction in the scopolin level was comparable with the reduction in the scopoletin level in this line. The reduced level of scopolin in the double mutant line (by ~40%) compared to the wild type, could partially explain the high level of scopoletin unexpectedly obtained. This reduction could result from glycolysis of the stored scopolin to compensate for the reduction of scopoletin after knocking out the genes for both OMT enzymes. The reduction of scopolin in all mutant lines compared to its level in the wild-type roots was significant ($P < 0.05$), but there was no significant difference between the measured scopolin concentration in the root extracts of the single-mutant A.t-CCoAOMT and the double-mutant lines ($P = 0.23$).

Inhibition of esculetin $O$-methylation in both A.t-EOMT and the double mutant plants resulted in accumulation of esculetin in their roots to a higher level than in the wild type root extract (Figure 3.20). A significant increase ($P < 0.05$) in esculetin level was measured in both the single-mutant A.t-EOMT and the double-mutant line (1.7x and 1.4x respectively) despite the low concentrations of esculetin in all the plant lines examined.

Umbelliferone levels were analysed in the root extracts of different mutant lines, but its accumulation pattern showed a big variation and unreproducible results, despite the reported slight increase in its accumulation in the A.t-CCoAOMT roots (Kai et al., 2006).

These analysis experiments were performed with different chromatographies while optimizing the method. They were repeated several times in the same optimised method. Despite the variation in the absolute hydroxycoumarin concentrations obtained between the different trials, the pattern in coumarin accumulations between the different lines, with reference to the wild type, was not affected.
Fig. 3.20. Esculetin levels in wild-type and transgenic *A. thaliana* plants.
Histograms show the relative accumulation of esculetin in the root extracts to the internal standard 4-methylumbelliferone.

Confirmation of cassava OMT candidate genes functional identity

Cassava OMT candidate genes M.e-CCoAOMT2 and M.e-EOMT were separately overexpressed in the *A. thaliana* mutant lines A.t-CCoAOMT and A.t-EOMT respectively. The transformation reactions were mediated by the transgenic *A. tumefaciens* cells.

Scopoletin and scopolin levels varied between the different transformed *A. thaliana* A.t-CCoAOMT mutant lines (Figure 3.21).

Fig. 3.21. Levels of hydroxycoumarins after overexpressing cassava M.e-CCoAOMT2 gene in A.t-CCoAOMT mutant plants. Histograms show scopoletin (A) and scopolin (B) levels in the root extracts of the transformed lines compared to the mutant and the wild-type lines.
The accumulation of both scopoletin and scopolin were restored in the root extracts of the transformed line 1, but no significant variations in these hydroxycoumarin levels were measured in the other transformed lines. The ability of M.e-CCoAOMT2 gene to complement the mutation in the A.t-CCoAOMT plants, therefore restoring both scopoletin and scopolin levels, confirmed its function in scopoletin biosynthetic pathway, not only depending on sequence similarity with the reference A. thaliana CCoAOMT gene.

After transforming A.t-EOMT mutant plants with M.e-EOMT cassava gene, the measured levels of scopoletin in the root extracts of the three transformed lines were significantly higher \((P\) values between each transformed line and A.t-EOMT < 0.05) than its level in the mutant line and also than in the wild type. Scopolin levels were also restored in the transformed lines: line 2 and line 3 (Figure 3.22). The ability of M.e-EOMT gene to complement the mutation in the A.t-EOMT plants and therefore restoring both scopoletin and scopolin levels confirmed its functional identity.

![Fig. 3.22. Levels of hydroxycoumarins after overexpressing cassava M.e-EOMT gene in A.t-EOMT mutant plants.](image)

Histograms show scopoletin (A) and scopolin (B) levels in the root extracts of the transformed lines compared to the mutant and the wild-type lines.

The variations in the different transformed lines are not unexpected because each transformed line could have a different expression level. Moreover, it is unpredictable whether a single or multiple copies of the candidate gene were inserted in the A. thaliana genome. The effect of multiple copies could increase, decrease, or even inhibit the expression of the inserted gene (Jorgensen, 1990; Hobbs et al., 1993).
The overall conclusions of these experiments agreed with previous findings (Kai et al., 2006; Kai et al., 2008) that scopoletin is mainly biosynthesized through the ferulate pathway (pathway 1, Figure 3.1) in wild type A. thaliana. This was confirmed after measuring a significant reduction (by ~60%) in its concentration in the root extracts of mutant plants which lack the CCoAOMT activity, a key enzyme in pathway 1. Scopoletin was also biosynthesized (to a lesser extent to ~20%) in the wild type plants through pathways 2 and 3 (Figure 3.1) after esculetin O-methylation reaction. This was mediated by A.t-EOMT enzyme which was identified via bioinformatics tools; its function was confirmed by these in planta experiments. The ability of the cassava OMTs to restore scopoletin and scopolin levels in these mutant A. thaliana lines verified their OMT functions.

To further explore the biosynthetic origin of the scopoletin in each A. thaliana mutant line, in addition to the cassava wild-type and transgenic roots, feeding experiments with different stable isotopically labelled intermediates were designed, and the carbon flux in each different pathway was followed (chapter 4).
Chapter 4. Tracing scopoletin biosynthesis in wild type and transgenic *M. esculenta* and *A. thaliana*

4.1. Introduction

The activity of PAL, a key entry enzyme on the phenylpropanoid pathway, is upregulated soon after harvesting the cassava roots (Tanaka *et al.*., 1983). Therefore, scopoletin and its β-glycoside, scopolin accumulate in cassava roots undergoing PPD as part of the plant responses to wounding (Beeching *et al.*, 1993; Bayoumi *et al.*, 2010). Their accumulation, and possibly, oxidation and polymerization are responsible for the undesirable blue/black discolouration (Uarrota *et al.*, 2016). Cassava roots exhibit PPD symptoms soon (2-3 days) after harvesting, rendering them unpalatable and unmarketable (Blagbrough *et al.*, 2010). While PPD is entirely due to an endogenous oxidative processes, i.e. not microbiological in nature, several procedures are carried out by producers to delay the oxidation and therefore to prolong the shelf life of the roots. Atmospheric oxygen is excluded through storing the roots in polyethylene bags or by coating them with wax immediately after harvesting (Reilly *et al.*, 2003).

**Fig. 4.1. The major biosynthetic pathway of scopoletin and scopolin.** Scopoletin and its β-glycoside scopolin are biosynthesized mainly from L-phenylalanine via ferulate.
Scopoletin is derived from L-phenylalanine after several hydroxylation and O-methylation reactions followed by isomerization and lactonization steps. Indeed, there is an agreement in the literature that scopoletin is biosynthesized on a pathway via ferulate as a major intermediate (Kai et al., 2008; Matsumoto et al., 2012). The biosynthesis pathway is via: deamination of L-phenylalanine by PAL to produce E-cinnamate, hydroxylation by C4’H to produce p-coumarate, hydroxylation by C3’H to produce E-caffeate, O-methylation by CCoAOMT to produce E-ferulate, hydroxylation by F6’H to produce E-6’-hydroxyferulate, followed by E-Z-isomerization to produce Z-6’-hydroxyferulate which could occur spontaneously (Lin et al., 2012) or via an isomerase enzyme (Bayoumi et al., 2008a). Finally, the Z-6’-hydroxyferulate lactonizes to give scopoletin. Scopoletin is glycosylated in the cytosol into its β-glycoside scopolin by scopoletin glucosyltransferase (GT), and is then stored in the vacuoles (Bourgaud et al., 2006; Gnonlonfin et al., 2011) (Figure 4.1).

Downregulation of the F6’H1 gene in A. thaliana (Kai et al., 2008) and of a homologous gene in M. esculenta known as F6’H (Liu et al., 2017), led to a significant reduction in scopoletin accumulation in the roots, and of the black/blue discolouration in cassava roots after harvesting. These in planta observation, confirmed that F6’H is a key enzyme in scopoletin biosynthesis in both A. thaliana and M. esculenta.

**F6’H**

Two feruloyl CoA 6’-hydroxylase genes were expressed in A. thaliana root tissues: AT3g13610 referred to as A.t-F6’H1, and AT4g37410 referred to as A.t-F6’H2. Downregulation of A.t-F6’H1 resulted in severe reduction (~97%) in scopoletin and scopolin accumulations compared to the wild type roots. No effect was noticed after downregulation of A.t-F6’H2 in A. thaliana (Kai et al., 2008). Homologous sequences for the A.t-F6’H1 gene were identified in sweet potato Ipomoea batatas. Their involvement in scopoletin biosynthesis was confirmed after in vivo experiments by expressing the I. batatas F6’H gene in E. coli (Matsumoto et al., 2012).

Depending only on 59% amino acid sequence similarity with A.t-F6’H1, the hydroxylase gene isolated from rue, the herb of grace Ruta graveolens, showed, after being functionally expressed in E. coli, an equal affinity to produce scopoletin and
umbelliferone from feruloyl CoA and \( p \)-coumaroyl CoA respectively (Vilarat et al., 2012). Another study based on sequence similarities with A.t-F6'H1 in tobacco *Nicotina tabacum*, revealed that silencing tobacco F6'H1 gene resulted in a dramatic decrease in scopoletin level and therefore, weaker blue fluorescence under UV was noticed in response to pathogens. This shows that tobacco F6'H1 is strongly involved in scopoletin biosynthesis (Sun et al., 2014).

Recently, Liu et al. created transgenic cassava plants with no or at least downregulated activity of F6'H1 enzyme using RNA interference (RNAi) technology. According to its amino acid sequence similarity with A.t-F6'H1, a small family of 7 hydroxylase genes was identified in cassava. Three of these genes were able to complement an F6'H1 T-DNA insertion mutant in *A. thaliana* plants, thereby restoring scopoletin accumulation. Knocking-out or downregulating these genes showed a significant reduction in scopoletin accumulation in cassava roots undergoing PPD, compared to the wild type, and therefore a visible reduction of the undesirable blue-black discolouration (Liu et al., 2017). As part of ongoing research on cassava in the Beeching research group, University of Bath, Dr Liu generously offered the best transgenic lines of cassava that exhibited the most reduction in scopoletin level for these further studies. This transgenic cassava plant is referred to as M.e-F6'H mutant.

The aims of these experiments are to investigate the source of the scopoletin biosynthesized in both A.t-F6'H1 and M.e-F6'H mutants, along with another *A. thaliana* mutant (A.t-EOMT) that has been created and discussed before in chapter 3, by tracing the biosynthesized products of wild type and transgenic plants grown in media rich in stable isotopically labelled intermediates.
4.2. Materials

Plant materials

Transgenic cassava (M.e-F6’H) created and generously provided by Dr Liu in the Beeching research group, University of Bath, were propagated and grown in the University of Bath glass house under the following conditions: 25-30 °C, relative humidity 80%, and 16 h light. The transgenic cassava is a modification of the wild type TMS 60444 after being transformed with an RNAi construct designed to silence the feruloyl CoA 6′-hydroxylase (F6’H) enzyme in cassava roots (Liu et al., 2017).

The seeds of A. thaliana T-DNA insertion line Salk-129938, in which the A.t-F6’H1 was knocked-out, were purchased from the Nottingham Arabidopsis Stock Centre (NASC), University of Nottingham, and grown in the University of Bath growth room under the following conditions: 21-22 °C, relative humidity 50-60%, and 16 h light.

Chemicals and instruments

All solvents were purchased from Fisher, UK. All chemicals were purchased from Sigma-Aldrich, UK unless otherwise stated. E-Cinnamic-2,3,2′,3′,4′,5′,6′-d7 acid (98 atom % D) and 4-hydroxy-3-methoxy-d3-cinnamic acid (3-OCD3) (99.9 atom % D) were purchased from CDN-Isotopes, Canada via their distributor QMX, UK.

4.3. General methods

Homozygous A.t-F6’H line verification

A. thaliana (Salk-129938) mutant plants were allowed to grow for 2-3 weeks. The genomic DNA was isolated and used as a template for the PCR homozygosity verification reactions. Total RNA was isolated and the corresponding cDNA was synthesised. DNA and RNA isolation, cDNA synthesis, and homozygous lines identification were performed as stated before in chapter 3 (section 3.3). Primer pairs used for the detection of the T-DNA insert were designed through Salk Institute genomic analysis website, http://signal.salk.edu/tdnaprimers.2.html (accessed on 01.08.2017). Primer pair used for the RT-PCR reaction was designed and quantified using Oligo Analyser 3.1 Integrated DNA Technologies, INC (US)
A DNA-gel band using the primer pair (RP + LB1) verified the T-DNA insertion, and the absence of a DNA-gel band using the primer pair (F6’H-T-DNA) confirmed mutant homozygosity. RT-PCR was performed using the synthesized cDNA as a template using the primer pair (A.t-F6’H1) to confirm that the gene is not functional. Primer sequences are listed in Table 4.1. Those lines revealed to be homozygous, were allowed to grow (5-6 more weeks) and to develop seeds. The seeds were harvested and used for further experiments.

Table 4.1. Designed PCR primers sequences

<table>
<thead>
<tr>
<th>primer</th>
<th>sequence (5´-3´)</th>
<th>amplicon size (bp)</th>
<th>annealing temp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>F6’H-T-DNA</td>
<td>&gt;RP3 AGAAGATGGTGAGGAGGC &gt;LP3 GGTCGGGATTCTAATCTCAGC</td>
<td>1244</td>
<td>53 °C</td>
</tr>
<tr>
<td>RP3+LB1</td>
<td>&gt;RP3 AGAAGATGGTGAGGAGGC &gt;LB1 CGTGGACCGCTTGCTGCAACT</td>
<td>573-873</td>
<td>53 °C</td>
</tr>
<tr>
<td>A.t-F6’H1</td>
<td>&gt;Fw AAATCCAGCTGAAGTAACCAG &gt;Rev GAAGAACTTTTGAGTCGCAG</td>
<td>304</td>
<td>54 °C</td>
</tr>
</tbody>
</table>

Quantitative determination of coumarin levels in wild type and transgenic plants using LC/MS

Coumarins from *M. esculenta* roots

Wild type and M.e-F6’H cassava roots (6-8 months) were harvested, washed, peeled, and cut (~1 cm³). PPD was induced in both plant types for 5 days in a controlled environment (20 °C, 80% humidity). Then the cubes were crushed and extracted with ethanol (60 ml, 16 h). Each extract (wild type and transgenic) was separately concentrated under reduced pressure, then 100 mg of each extract was re-dissolved in water (0.5 ml) before being filtered (0.45 μm pore size), and then subjected to LC/MS analysis to quantify the hydroxycoumarins of interest.

Coumarins from *A. thaliana*

Roots were harvested from wild type and A.t-F6’H1 *A. thaliana* plants (4-5 weeks) grown in sterile agar plates. The roots were dried (10 mg, 55 °C, 30 min), crushed and extracted with methanol (2 ml). 4-Methylumbelliferone (100 ng/ml)
was added to the extraction solvent as an internal standard to improve data reliability between the three biological replicates. The mixture was shaken (20 °C, 16 h), filtered (0.2 µm pore size), and lyophilized. The dried powder was dissolved in methanol (200 µl) with scoparone (100 ng/ml) as a second internal standard to improve data reliability between the three technical replicates, and subjected to LC/MS analysis to quantify the hydroxycoumarins of interest.

NMR characterization

1H NMR spectra of ferulic (Figure 4.2) and ferulic-OCD3 acids were obtained using a Bruker Avance III spectrometer operating at 500.13 MHz. Samples were dissolved in D2O (0.6 ml). Chemical shifts (ppm) for ferulic acid were: 3.86 (3H, s, O-Me), 6.28 (1H, d, J = 20.0, H-2), 6.78 (1H, d, J = 10.0, H-5’), 7.03 (1H, d, J = 10.0, H-6’), 7.15 (1H, s, H-2’), 7.56 (1H, d, J = 20.0, H-3). For ferulic-OCD3 acid, identical proton chemical shifts were obtained with the absence of the methoxy signal at δ = 3.86 ppm (Figure 4.2).

Fig. 4.2. Ferulic acid.

1H NMR and 2H NMR (deuterium-observed) spectra of cinnamic-d7 acid were obtained (500.13 MHz and 76.75 MHz respectively) using CDCl3 and CHCl3 solvents respectively. Chemical shifts obtained from the non-deuterated cinnamic acid (Figure 4.3) 1H NMR spectrum were used as reference: 6.45 (1H, d, J = 16.0, H-2), 7.39-7.41 (3H, m, H-3’, H-4’, H-5’), 7.55 (2H, dd, J = 2.5 and 6.5, H-2’, H-6’), 7.79 (1H, d, J = 16.0, H-3). For cinnamic-d7 acid: no aromatic (styryl) proton signals were observed in 1H NMR and enhanced deuterium signals were obtained by 2H NMR spectra at the corresponding chemical shifts.

Fig. 4.3. Cinnamic acid.
Incorporation in esculetin and ferulic acid

Based on, and following the precedent literature, ¹⁸O-enrichment of esculetin and ferulic acid separately, was performed as follows: esculetin (20 mg, 0.112 mmol) or ferulic acid (20 mg, 0.103 mmol) was mixed with H₂¹⁸O (500 µl, 99 atom % ¹⁸O), and acidified acetonitrile (1 ml containing 10 µl conc. HCl) (Namt and Valentine, 1993; Niles et al., 2009). The reaction mixtures were stored (70 °C) for 8 and 10 days for esculetin and ferulic acid respectively, and monitored (every 2 days) by HR-MS. The ¹⁸O-esculetin was dissolved in DMSO (0.5 ml), diluted with water (2.5 ml), and then the pH was adjusted to 7 using aqueous NaOH (0.1 M). Likewise, the ¹⁸O-ferulic acid was dissolved in water (3 ml), and the pH was adjusted to 7 using aqueous NaOH (0.1 M). For A. thaliana feeding purposes, the solutions were filter-sterilized (twice) through sterile Sartorious™ Minisart™ NY syringe filters (diameter of 25 mm and a pore size of 0.2 µm). The theoretical m/z ratios are:

- esculetin C₁₀H₁₀O₄ 177.0339, ¹⁸O-esculetin C₁₀H₁₀O₃¹⁸O 181.0381,
- ferulic acid C₁₀H₉O₄ 193.0506, ¹⁸O-ferulic acid C₁₀H₉O₃¹⁸O 195.0549, C₁₀H₉O₂¹⁸O₂ 197.0590.

Feeding cassava roots with isotopically labelled compounds

Following and modified after Bayoumi et al., 2008a protocol, cassava roots were harvested, peeled, cut into small cubes (~1 cm³), and divided into two groups (17-20 g). One group was manually sprayed with an aqueous solution (4 ml) which contained an isotopically labelled compound (0.5 mg/ml of cinnamic-d₇ acid or ferulic-OCD₃ acid or ¹⁸O esculetin). The aqueous solutions of cinnamic-d₇ acid and ¹⁸O-esculetin were prepared by dissolving 2 mg in 0.5 ml DMSO, then diluting with water up to 4 ml. The pH was adjusted to 7 using aqueous NaOH (0.1 M). The other group was not sprayed and was therefore used as a negative control. The cassava cubes were maintained for 5 days (20 °C, 80% humidity) on filter papers in Petri dishes in a vacuum desiccator to allow PPD to occur. During the process, the pressure was reduced using a water pump (3 times for 10 min on the first day, then 2 times for 10 min daily) to aid the exogenous materials in penetrating the root tissues. On the ⁵th day, the root cubes were crushed using a pestle and mortar, extracted with ethanol (2x 60 ml, 16 h) with continuous stirring using a magnetic stirrer. The combined extracts were filtered and concentrated under reduced pressure at 40 °C. The crude extract (100 mg) was dissolved in water (0.5 ml) for LC/MS analysis.
Feeding *A. thaliana* plants with isotopically labelled compounds

*A. thaliana* plants were allowed to grow vertically in agar plates in the University of Bath growth room for 3 weeks. Under aseptic conditions, the plant water was removed and each plate was flooded with a sterilized aqueous solution containing the labelled compounds (4 ml, 0.5 mg/ml). The volume and concentration of the feeding solution were chosen after several attempts to optimize the method. The plants were placed horizontally in the growth room to allow better utilization of the exogenous compounds, and the plants were allowed to grow for 10 more days. Untreated wild-type plants were used as a control. Plant tissues were harvested, rinsed briefly under running tap water, dried (55 °C, 30 min), gently crushed (manually or by using a clean plastic rod in an Eppendorf tube), and then stored at -20 °C. Root tissues (20-30 mg) and leaf tissues (100 mg) were each extracted twice with methanol (2x 1 ml for roots and 2x 2 ml for leaves) with continues shaking (15 rpm, 16 h). The combined extracts were filtered through Fisherbrand™ non-sterilized syringe filter (pore size 0.45, diameter 25 mm), lyophilized, and the extracts were dissolved in either methanol (200 µl) or water (200 µl) for LC/MS analysis.

UPLC/ESI-MS

Using the UPLC/ESI-MS methods, as stated before in section 2.3, methanolic and aqueous samples were separated by LC using a reverse phase (RP) C18 column, and the coumarins of interest were detected by HR-TOF MS using ESI in positive ion mode. Data obtained were processed using Data Analysis software 4.3. The detected mass peaks of the compounds of interest [M+H]^+ and [M+Na]^+ and their isotopes were typically matched with their corresponding theoretical values within a 10 ppm error range unless otherwise stated. The mass error was calculated as: Δ mass (ppm) = (found mass - required mass)/required mass x 10⁶. The required *m/z* values are: scopoletin C₁₀H₉O₄ 193.0495, scopoletin-d₁ C₁₀H₈O₄D₁ 194.0558, scopoletin-d₂ C₁₀H₇O₄D₂ 195.0621, ¹⁸O-scopoletin C₁₀H₇O₄¹⁸O 195.0538, scopoletin-d₃ C₁₀H₆O₄D₃ 196.0684, scopoletin-d₄ C₁₀H₅O₄D₄ 197.0746, scopolin C₁₆H₁₉O₉ 355.1024, scopolin-d₁ C₁₆H₁₈O₉D₁ 356.1086, ¹⁸O-scopolin C₁₆H₁₉O₈¹⁸O 357.1066, scopolin-d₂ C₁₆H₁₇O₉D₂ 357.1149, scopolin-d₃ C₁₆H₁₆O₉D₃ 358.1212, scopolin-d₄ C₁₆H₁₅O₉D₄ 359.1275, esculin C₁₅H₁₇O₉ 341.0872, ¹⁸O-esculin C₁₅H₁₇O₈¹⁸O 343.0914, esculetin C₉H₇O₄ 179.0339, and ¹⁸O-esculetin C₁₀H₁₀O₃¹⁸O
181.0381. The mass spectrometric data (m/z ratios) were obtained by adding the isotopic weights minus an electron (0.000548597), the final calculated mass was rounded to 4 decimal places. \textsuperscript{1}H requires 1.007825, \textsuperscript{2}H requires 2.014102, \textsuperscript{12}C requires 12.000000, \textsuperscript{16}O requires 15.994915, and \textsuperscript{18}O requires 17.999160. The LC peaks of scopoletin (not unexpectedly overlapping with its isotopes) had retention times (Rt) = 5.9 min, scopolin (and its isotopomers) Rt = 5.3 min, esculin (and its isotopomers) Rt = 4.9 min.

4.4. Results and discussion

\textbf{A.t-F6´H1}

Homozygous A.t-F6´H1 plants were verified after Taq-polymerase PCR reactions, genomic DNA was isolated from 2-3 weeks old plants and used as a template. No gel band was detected after a PCR reaction using the primer pair (T-DNA-F6´H), and a clear band was detected after a PCR reaction using the primer pair (RP3+LB1). This confirmed the presence of the T-DNA insertion that renders the gene size too big to be amplified. No expression band was detected after a RT-PCR reaction using the primer pair (A.t-F6´H1), and this confirm that the gene is non-functional.

The coumarins of interest were extracted with methanol (2 ml) and quantified from the A.t-F6 H1 mutant roots. Their accumulations were compared with those isolated from the wild type \textit{A. thaliana} roots. The selective peak-areas were normalized to the internal standard 4-methylumbelliferone peak area to avoid any underestimation of the hydroxycoumarin levels due to technical or operator errors. 4-Methylumbelliferone and scoparone (Figure 4.4) share the nucleus structure of the coumarins of interest, but it was not detected in the plant crude extracts. In the chromatography system used, the Rt values were: scopolin 5.3 min, scopoletin 5.9 min, scoparone 6.15 min, and 4-methylumbelliferone 6.2 min.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{molecules.png}
\caption{The internal standards 4-methylumbelliferone and scoparone}
\end{figure}
Scopoletin and scopolin accumulation levels were dramatically decreased in the *A. thaliana* roots after knocking-out the F6’H1 gene (Figure 4.5). This confirms that *o*-hydroxylation of feruloyl CoA is a key step in scopoletin, and therefore scopolin, biosynthesis in *A. thaliana*. On the other hand, F6’H1 gene was reported to have, after in vivo experiments in *E. coli*, an activity towards *p*-coumaroyl CoA to produce umbelliferone, even though it was weaker than its activity towards feruloyl CoA, a possible direct precursor of scopoletin (Kai *et al.*, 2008). Also, the F6’H1 homologue, in *Ruta graveolens*, showed an equal affinity towards both *p*-coumaroyl CoA and feruloyl CoA (Vialart *et al.*, 2012). Accordingly, the dramatic decrease in scopoletin and therefore scopolin accumulation after knocking-out the F6’H1 gene could result from the inhibition of all the three pathways (Figure 3.1) by knocking-out the feruloyl CoA 6’-hydroxylase (F6’H1) in *A. thaliana*.

Fig. 4.5. Accumulation of scopoletin and scopolin in *A. thaliana* roots.

Histograms show the dramatic decrease in both scopoletin (A) and scopolin (B) levels in A.t-F6’H1 mutant compared to the wild type.

**M.e-F6’H**

Transgenic cassava plants with a downregulated or inhibited activity of F6’H1 were allowed to grow for 6-8 months before the roots were harvested. Then PPD was induced and the hydroxycoumarins were extracted with ethanol, isolated, and quantified. An overall reduction in scopoletin level was detected in the transgenic cassava roots compared to the wild type *M. esculenta* (Figure 4.6). Early in PPD, scopoletin was found to be significantly reduced ~1/2x (*P* values between wild type and M.e-F6’H on each measured point were < 0.05 except for day 4, *P* = 0.51). In addition to the overall decrease in scopoletin accumulation, there was a clear delay in the appearance of the PPD symptoms, which is concomitant with the
shifting of the peak accumulation from day 2 in the wild type roots to days 4-5 in the transgenic roots. PPD symptoms were previously quantified in M.e-F6´H by measuring the discoloration on cassava roots, where the highest discoloration scores were reported on days 4 and 5 in the transgenic plant. Despite the differences in the detailed day to day measured concentrations, i.e. Liu et al. found a ~70% reduction in scopoletin accumulation in days 1 and 2, and ~ 55% reduction in days 3-5 after harvest in the M.e-F6´H compared to the wild type, Liu and co-workers recently reported a very similar trend in scopoletin accumulation patterns in both wild type and transgenic cassava roots (Liu et al., 2017). These variations mainly reflect the use of different extraction and sample preparation techniques.

**Fig. 4.6. Accumulation of scopoletin in cassava roots.** Lines show a decrease in scopoletin accumulation in M.e-F6´H mutant compared to the wild type. Scopolin showed the same reduced accumulation pattern as scopoletin in the initial stage of PPD; a significant reduction in its measured concentration (~1/2x) in the transgenic compared to the wild type roots was measured (P values between wild type and M.e-F6´H on each measured point were < 0.05) (Figure 4.7). The reduced level was an expected consequence of reducing scopoletin biosynthesis, as scopoletin is a biosynthetic precursor to scopolin. Scopolin is biosynthesized by the addition of glucose unit to scopoletin in a reaction mediated by uridine diphosphate-glucose (UDP-Glc) phenylpropanoids glucosyltransferase enzymes, and then stored in the plant vacuoles. Scopoletin glucosyltransferase has been investigated in other
plants, e.g. tobacco (Chong et al., 2002) and in A. thaliana (Ahn, et al., 2010), but not yet identified in cassava.

The small size (3 cm length) and weight (5 g) of the harvested transgenic roots in some pots (Figure 4.8) and the time limit (up to 8 months until harvesting the roots), impede repeating the analysis with a larger number of samples, but scopolin concentrations by day 6 are comparable between wild type and transgenic cassava (Figure 4.7).

![Graph showing scopolin level over days after roots harvest]

**Fig. 4.7. Accumulation of scopolin in cassava roots.** Lines show a decrease in scopolin level in the initial stage of PPD (days 1-3) in the transgenic cassava roots (red line) compared to the wild type roots (blue line). There is an unexplained increase late in PPD (days 3-6) in scopolin level in M.e-F6´H mutant compared to the wild type.

![Image of M.e-F6´H cassava root]

**Fig. 4.8. M.e-F6´H cassava root.** Image shows the harvested transgenic cassava root (8 months, 3 cm length, 5 g).
A decrease in esculin levels in response to inhibited or downregulated F6ʼH activity in cassava roots was also discovered. The average (n=3) measured concentration of esculin was found to be ~314 ± 17.8, 586 ± 64.3, and 1039 ± 44.6 ng/g fresh wild type roots and ~199 ± 89, 186 ± 14.3, and 396 ± 84 ng/g fresh transgenic roots for days 4, 5, and 6 after harvest respectively. Esculin accumulation was significantly reduced late in PPD (after day 3) (P values between wild type and M.e-F6ʼH on each measured point were < 0.001). In the transgenic cassava line, a steady low concentration of esculin was detected (P value between days 4 and 5 = 0.38) until day 6 after harvesting the roots where its accumulation increased (P value between days 5 and 6 < 0.001) (Figure 4.9).

![Graph showing esculin level](image)

**Fig. 4.9. Accumulation of esculin in cassava roots.** Lines show a decrease in esculin level in M.e-F6ʼH mutant compared to the wild type.

Esculetin was detected at very low levels ~30-70 ng/g in both wild type and transgenic fresh roots.

The overall decrease in scopoletin and its β-glycoside scopolin, after downregulating the F6ʼH enzyme in cassava, confirms the crucial role of the O-hydroxylation step of feruloyl CoA mediated by the enzyme F6ʼH1 in scopoletin biosynthesis and therefore the appearance of PPD. The decreased level of esculin strengthens the hypothesis that F6ʼH controls the three pathways by its activity towards coumaroyl CoA in addition to feruloyl CoA. 2ʼ,4ʼ-dihydroxycinnamic acid is the expected hydroxylated product of coumaroyl CoA, therefore the biosynthesis
of umbelliferone and possibly 6´-hydroxyferulate, the intermediates on pathways 2 and 3, will be affected by F6´H downregulation.

\[ ^{18}\text{O-Enrichment of esculetin and ferulic acid} \]

Enrichment of esculetin with an \(^{18}\text{O}\)-atom was performed to introduce a labelled oxygen into the lactone carbonyl by acid catalysed exchange reaction with \(\text{H}_2^{18}\text{O}\) (Figure 4.10).

\[ \text{esculetin} \xrightarrow{70 \degree \text{C}} \text{esculetin}^{^{18}\text{O}} \]

\[ \text{esculetin} \xrightarrow{H^+} \text{esculetin}^{^{18}\text{O}} \]

**Fig. 4.10. Incorporation of an \(^{18}\text{O}\)-labelled oxygen into esculetin via an acid-catalysed exchange reaction.** Blue oxygen represents \(^{16}\text{O}\) and red oxygen represents \(^{18}\text{O}\).

The reaction was monitored by LC/MS over a period of 8 days. The percentage incorporations of the average area under LC/MS peak (n=3) for \(\text{C}_9\text{H}_7\text{O}_3^{^{18}\text{O}}\): \(\text{C}_9\text{H}_7\text{O}_4\) were: \(~38\%\) on day 2, \(~71\%\) on day 4 and \(~100\%\) on day 8 (Figure 4.11).
**Fig. 4.11. Monitoring esculetin $^{18}$O-enrichment reaction using LC-MS.** Images show the incorporation of the labelled oxygen into esculetin on days 2, 4, and 8. Yellow-filled peaks represent C$_9$H$_7$O$_3^{18}$O.

Enrichment of ferulic acid with two $^{18}$O atoms was performed to introduce two $^{18}$O-labelled oxygen atoms into the propionic-C1 carboxylic acid by double acid-catalysed exchange reactions with H$_2^{18}$O (Figure 4.12). The first exchange reaction was with the carbonyl oxygen proceeding to the second with the acid oxygen because of the excess of H$_2^{18}$O atoms in the acidified reaction medium.

---

**Fig. 4.12. Incorporation of two $^{18}$O-labelled oxygen into ferulic acid via double acid-catalysed exchange reactions.** Blue oxygen represents $^{16}$O and red oxygen represents $^{18}$O.
The reaction was monitored by LC/MS over a period of 10 days. The ratios of the average area under LC/MS peak (n=3) for the \( C_{10}H_{9}O_4 \), \( C_{10}H_{9}O_3^{18}O \), \( C_{10}H_{9}O_2^{18}O_2 \) to the total ferulic acid in the reaction were: ~18%: 49%: 33% on day 6, ~12%: 43%: 45% on day 8, and ~12%: 32%: 56% on day 10 (Figures 4.13 and 4.14).

**Fig. 4.13.** Monitoring ferulic acid double \(^{18}\text{O}\)-enrichment reaction using LC-MS. Image shows the incorporation of the labelled oxygens into ferulic acid on days 6, 8, and 10.

**Fig. 4.14.** HR-MS spectra showing the relative peak intensities for ferulic acid and its isotopomers. HR-MS data for ferulic acid (193.0506), ferulic-\(^{18}\text{O}\) acid (195.0549) and ferulic-\(^{18}\text{O}_2\) acid (197.0591) on days 6, 8, and 10. HR-MS (negative ion mode) for [M-H\(^-\)] were found within error ranges of 1 ppm.

The \(^{18}\text{O}\)-rich esculetin and ferulic acid were then used to explore their incorporation in scopoletin biosynthetic pathways in both wild type and transgenic *A. thaliana* and cassava plants.
Characterizations of the authentic cinnamic-d$_7$ and ferulic-OCD$_3$ acids

The identity of the commercial *trans*-cinnamic-d$_7$ acid was checked and confirmed using HR-MS, the theoretical *m/z* ratio for [M-H], $C_9D_7O_2$ requires 154.0885, found 154.0862. The deuterium atom positions (3-phenyl-d$_5$-propenoic acid-2,3-d$_2$) were confirmed using $^1$H NMR and deuterium-observe NMR spectroscopy (Figure 4.15).

Fig. 4.15. Cinnamic-d$_7$ acid NMR characterization. A shows the $^1$H NMR spectrum (500 MHz) for $C_9H_8O_2$ in CDCl$_3$. B shows the $^1$H NMR spectrum (500 MHz) for $C_9HD_7O_2$ in CDCl$_3$. C shows deuterium-observed NMR spectrum (76.75 MHz) for $C_9HD_7O_2$ in CHCl$_3$ and a trace (100 µl) of CDCl$_3$ for reference.
Non-deuterated cinnamic acid \(^1\)H NMR spectrum was used as a reference to assign the proton signals chemical shifts. Proton signals were in the noise level in the \(^1\)H NMR spectrum of the deuterated cinnamic acid \(\text{C}_9\text{D}_8\text{O}_2\). This is because the proton signals are limited to the residual protons in the sample, the commercial cinnamic-d7 acid is sold as 98 atom % D. To confirm that the deuterium atoms replaced their corresponding protons, deuterium-observed spectroscopy was carried out in non-deuterated \(\text{CHCl}_3\). A small volume of \(\text{CDCl}_3\) (100 \(\mu\)l) was added to enable chemical shift referencing. Peaks with identical chemical shifts with the reference cinnamic acid were observed because they have the same magnetic environments. The peaks were slightly broad and this is a result of coupling and poor deuterium resolution (Schory et al., 2004). NMR data confirmed the structure of \(E\)-3-phenyl-d5-propenoic acid-2,3-d2.

![Diagram of ferulic and ferulic-OCD3 acids](image)

**Fig. 4.16.** \(^1\)H NMR spectra for ferulic and ferulic-OCD3 acids (500 MHz). A shows the ferulic-OCD3 acid spectrum with no signal for O-\(\text{CH}_3\) at \(\delta = 3.86\) ppm, B shows the reference ferulic acid spectrum.

Feeding the plants with labelled ferulic acid enables the investigation of the incorporation of ferulic acid in scopoletin biosynthesis in both wild type and transgenic plants. It is important to confirm the position of the added three deuterium atoms on the ferulic acid molecule to predict the labelled-scopoletin product and interpret the produced data. The commercial \(E\)-ferulic-OCD3 acid identity was checked and confirmed including by HR-MS, the theoretical \(m/z\) ratio for [M-H] \(\text{C}_{10}\text{H}_{6}\text{D}_3\text{O}_4\) requires 196.0689, found 196.0680. The deuterium atom
positions were unambiguously confirmed to be on the methoxy group rather than on the phenyl ring using $^1$H NMR spectroscopy (Figure 4.16). The structure of the labelled ferulic acid was confirmed to be $E$-4'-hydroxy-3'-methoxy-d$_3$-cinnamic acid by the absence of the methoxy signal at $\delta = 3.86$ ppm.

Following the flux in wild type and transgenic cassava during PPD

Cassava roots were harvested, peeled, and cut into cubes (~1 cm$^3$). Different solutions (pH was adjusted to 7 using 0.1 M aqueous NaOH) containing labelled intermediates were fed separately to the cubes using a simple hand-pumped aerosoliser. PPD was induced (5 days, 20 ºC, 80% humidity), cubes were crushed, and the coumarins of interest were extracted with ethanol. The scopoletin and scopolin along with their isotopomers were analysed and quantified using LC-ESI-MS. Authentic scopoletin and the previously isolated scopolin (chapter 2) were used as reference standards.

The incorporation of cinnamic-d$_7$ acid in scopoletin biosynthesis in both wild type and transgenic cassava roots

After feeding the wild type and the transgenic cassava cubes with aqueous solution contains cinnamic-d$_7$ acid, scopoletin-d$_3$ and scopolin-d$_3$ peaks were detected as the major isotope peaks as well as peaks assigned to the naturally occurring scopoletin and scopolin in both wild type and transgenic cassava root extracts. No MS peaks for scopoletin-d$_2$ and scopolin-d$_4$ were detected. Scopoletin and scopolin LC peaks not unexpectedly overlapping their isotopomers are shown in Figure 4.17. HR-MS m/z ratios were: scopoletin C$_{10}$H$_9$O$_4$ requires 193.0495, found 193.0506; $^{13}$C-scopoletin C$_{9}$CH$_6$O$_4$ requires 194.0529, found 194.0539; scopolin-d$_3$ C$_{10}$H$_6$D$_3$O$_4$ requires 196.0684, found 196.0693; C$_{10}$H$_8$O$_4$Na requires 215.0315, found 215.0325; scopolin C$_{16}$H$_{19}$O$_9$ requires 355.1024, found 355.1035; scopolin-d$_{3}$ C$_{16}$H$_{16}$D$_3$O$_9$ requires 358.1212, found 358.1217; C$_{16}$H$_{18}$O$_9$Na requires 377.0843, found 377.0851 (Figure 4.18).

The percentage incorporations of the isotopically labelled precursor, cinnamic-d$_7$ acid, into scopoletin to the naturally occurring scopoletin were similar in both the wild type and the transgenic root extracts: [scopoletin+I]$^+$ ~9.8%, and scopoletin-d$_3$ ~13.0%. For scopolin in the wild type roots: [scopolin+I]$^+$ ~18.0%, and scopolin-d$_3$ ~25.0% in wild and 52.0% in the transgenic roots. In the control
(wild type with no feeding), [scopoletin+1]$^+$ and [scopolin+1]$^+$ were detected in a percentage incorporation of ~10%, no other isotopomers were detected. The feeding experiment with labelled cinnamic-d$_7$ acid was repeated three times with the wild type and twice with the transgenic roots, and then, the ratios were obtained from three injections of the extracts. Ratios naturally varied depending on the already biosynthesized and accumulated coumarins in the plant tissues before harvesting.

**Fig. 4.17.** Selective LC-MS chromatograms for scopoletin and scopolin after feeding with cinnamic-d$_7$ acid. A shows scopoletin and scopolin overlapped with their isotopomers. B shows an enlarged chromatogram of scopolin (C$_{16}$H$_{19}$O$_6$), scopolin-d$_1$, and scopolin-d$_3$.

**Fig. 4.18.** HR-MS of scopoletin and scopolin products after feeding wild type cassava roots with cinnamic-d$_7$ acid. A shows a spectrum for scopoletin and its isotopomers, B shows a spectrum for scopolin and its isotopomers.
The percentage of [scopoletin+1] and [scopolin+1] ions, recovered from each extract, to the unlabelled scopoletin and scopolin respectively, were similar to the $^{13}$C abundance incorporation (1.1% for each carbon). Theoretically, $^{13}$C carbon incorporation in scopoletin C$_{10}$H$_8$O$_4$ could require 11%. Experimentally it was found to be $\sim$9.8%. Scopolin C$_{16}$H$_{19}$O$_9$ could require 17.6% and experimentally it was found to be $\sim$18%. In addition, the resolution of the LC-MS instrument is limited to 0.005 Da mass error, the mass error defined as $\Delta$ mass (mDa) = (found mass – required mass). Therefore, it was not practical to distinguish unambiguously between scopoletin-d$_1$, scopolin-d$_1$ and their corresponding naturally occurring $^{13}$C-isotopes because their molecular masses overlapped within the 0.005 Da mass error accuracy (Figure 4.19). Scopoletin-d$_1$ requires 194.0558 (± 0.005 Da: 194.0508 - 194.0608), $^{13}$C-scopoletin requires 194.0529 (± 0.005 Da: 194.0479 - 194.0579), scopolin-d$_1$ requires 356.1086 (± 0.005 Da: 356.1036 - 356.1136), and $^{13}$C-scopolin requires 356.1057 (± 0.005 Da: 356.1007 - 356.1107). It is unlikely that the cinnamic-d$_7$ acid will lose 6 deuterium ions to produce scopoletin-d$_1$ and scopolin-d$_1$, therefore the (M+1) ions were assigned as the natural $^{13}$C isotopes.

![Accurate molecular mass within 0.005 Da mass error](image)

**Fig. 4.19. Accurate molecular mass within 0.005 Da mass error.** A scopoletin-d$_1$ and $^{13}$C-scopoletin. B scopolin-d$_1$ and $^{13}$C-scopolin.

It was expected that scopoletin-d$_3$ and scopolin-d$_3$ would be detected after feeding both the wild type and the transgenic roots with cinnamic-d$_7$ acid. This is because scopoletin and therefore, scopolin are derived from cinnamic acid as an early precursor in the phenylpropanoid pathway (Figure 4.20). Three aromatic deuteriums will be lost during the 3 hydroxylation steps regardless of which of the three pathways was followed and it could be two or even possibly three pathways.
The fourth deuterium is lost from position C-2 during the E-Z-isomerization step (Bayoumi et al., 2008b). However, the complete mechanism of this isomerization process is not yet delineated.

Fig. 4.20. Pathways for scopoletin biosynthesized from the exogenous labelled cinnamic acid. The scheme shows the incorporation of the labelled cinnamic-d₁₇ acid in the three possible pathways to produce scopoletin-d₃ after three different hydroxylation, E-Z-isomerization, and lactonization steps. In the transgenic plant, α-hydroxylation of ferulate-d₅ (red arrow) is inhibited.
The higher incorporation of the labelled deuterium in scopolin extracted from the transgenic roots compared to the wild type extracts (scopolin-d$_3$ ~52% in the transgenic and ~25% in the wild type cassava roots), is likely to reflect a smaller pool of the unlabelled scopolin in the transgenic roots.

Feeding experiments with a direct precursor or precursors in the latest steps to scopoletin in each pathway were therefore designed and performed to gain insight into the actual incorporation along each pathway in scopoletin biosynthesis, and to investigate the source of the residual scopoletin biosynthesized in the mutant plants after knocking-out key enzymes representing each pathway. Wild type and M.e-F6´H cassava plants were fed with ferulic-OCD$_3$ acid and $^{18}$O-esculetin separately.

The incorporation of ferulic-OCD$_3$ acid in scopoletin biosynthesis in both wild type and transgenic cassava roots

Cassava cubes under PPD were fed with deuterium-rich ferulic acid (ferulic-OCD$_3$), which is a late precursor in scopoletin biosynthesis through the ferulate pathway (Figure 4.1). It was expected to detect a higher incorporation of the labelled material in scopoletin biosynthesis after feeding with ferulic-OCD$_3$ acid in the wild-type plants than the percentage incorporated after feeding with labelled cinnamic-d$_7$ acid because of the diversity of hydroxycinnamate utilization along various biosynthetic pathways, e.g. into flavonoids and lignin biosynthetic pathways (Vogt, 2010). On the other hand, no labelled scopoletin was expected to be detected in the M.e-F6´H mutant due to the lack of the O-hydroxylating enzyme (Figure 4.21). However, the obtained results did not completely match these expectations.

![Chemical reactions](image_url)

Fig. 4.21. The expected scopoletin-d$_3$ product after feeding with E-ferulate-OCD$_3$. Scopoletin-OCD$_3$ is expected as a result of o-hydroxylation by F6´H enzyme, E-Z-isomerization, and then lactonization steps. No labelled scopoletin is expected to be synthesized in the absence of the F6´H enzyme in the transgenic cassava roots.
In the wild type extracts, scopoletin-d₃ and scopolin-d₃ peaks were detected by LC-MS (Figure 4.22). Scopoletin and scopolin LC peaks were overlapped with their isotopes, HR-MS m/z values were: scopoletin C₁₀H₉O₄ requires 193.0495, found 193.0497; C₁₀H₆D₃O₄ requires 196.0684, found 196.0685; scopoletin-d₃ C₁₀H₆D₃O₄ requires 197.0718, found 197.0725; scopolin C₁₆H₁₉O₉ requires 355.1024, found 355.1030; scopolin-d₃ C₁₆H₁₆D₃O₉ requires 358.1212, found 358.1217; C₁₆H₁₅O₉Na requires 377.0849, found 377.0849; and C₁₆H₁₅D₃O₉Na requires 380.1031, found 380.1037. The percentage incorporation of scopoletin-d₃: scopoletin was ~43%, and the percentage incorporation of scopolin-d₃: scopolin was ~63%.

Fig. 4.22. HR-MS of scopoletin and scopolin products after feeding wild type cassava roots with ferulic-OCD₃. A shows a spectrum for scopoletin and its isotopomers, B shows a spectrum for scopolin and its isotopomers.

In the M.e-F6´H deteriorated root extracts, feeding the fresh transgenic cassava cubes with ferulic-OCD₃ acid afforded a high deuterium incorporation in both scopoletin (scopoletin-d₃: scopoletin ~4:3) and scopolin (scopolin-d₃: scopolin ~2:1) despite the assumed inactivity of the F6´H enzyme. Scopoletin and scopolin LC peaks are shown in Figure 4.23. HR-MS m/z values were: scopoletin C₁₀H₉O₄
requires 193.0495, found 193.0498; scopoletin-d$_3$ C$_{10}$H$_6$D$_3$O$_4$ requires 196.0684, found 196.0686; C$_{10}$H$_8$O$_4$Na requires 215.0315, found 215.0316; C$_{10}$H$_5$D$_3$O$_4$Na requires 215.0503, found 218.0505; scopolin [M+H]$^+$ C$_{16}$H$_{16}$O$_9$ requires 355.1024, found 355.1024; C$_{16}$H$_{16}$D$_3$O$_9$ requires 358.1212, found 358.1222; C$_{16}$H$_{18}$O$_9$Na requires 377.0849, found 377.0853; and C$_{16}$H$_{15}$D$_3$O$_9$Na requires 380.1031, found 380.1040 (Figure 4.24).

![Fig. 4.23. Selective LC-MS chromatograms for scopoletin and scopolin after feeding transgenic roots with ferulic-OCD$_3$ acid.](image)

A shows scopoletin, scopolin overlapped with their isotopomers. B shows an enlarged chromatogram of scopolin overlapping scopolin-d$_3$.

This unexpected incorporation of the labelled ferulic-OCD$_3$ acid into scopoletin and therefore, into scopolin biosynthesized in the transgenic plants is explained either by the failure of the RNAi technology to knock-out the F6´H activity completely (although, it significantly downregulated it), or/and by the involvement of other hydroxylase enzymes in feruloyl CoA hydroxylation.
Fig. 4.24. HR-MS spectra of scopoletin and scopolin products after feeding transgenic cassava roots with ferulic-OCD$_3$. A shows a spectrum for scopoletin and its isotopomers, B shows a spectrum for scopolin and its isotopomers.

To get a better insight into the origin of the labelled hydroxycoumarins produced, labelled ferulic acid was fed to the completely knocked-out F6'H1 enzyme in the model plant A. thaliana, and the residual biosynthesized scopoletin (~3%) was investigated (next section). The lack or diminished activity of F6'H1 enzyme, could result in a decrease in the naturally biosynthesized substrate, ferulate as a part of a negative feedback inhibition process. Accordingly, the feeding with the OCD$_3$-labelled exogenous ferulate resulted in the observed high incorporation of the labelled deuteriums into scopoletin and scopolin (~1.3-fold and ~2-fold respectively) in the transgenic cassava root experiments.

The incorporation of $^{18}$O-esculetin in scopoletin biosynthesis in both wild type and transgenic cassava roots

For a more comprehensive exploration of the origin of the scopoletin biosynthesized in the mutant roots, a further feeding experiment was carried out using $^{18}$O-labelled esculetin. The percentage incorporation of the labelled scopoletin
and scopolin related to their natural occurrence in both the wild type and transgenic cassava roots were as follows: in the wild type roots, ~2.5% of $^{18}$O-scopoletin, and ~22% of $^{18}$O-scopolin, in the transgenic in the wild type roots ~1% $^{18}$O-scopoletin and ~18.5% $^{18}$O-scopolin. HR-MS $m/z$ values in the wild type extract were: scopoletin $\text{C}_{10}\text{H}_{9}\text{O}_{4}$ requires 193.0495, found 193.0489; $^{13}$C-scopoletin $\text{C}_9^{13}\text{CH}_9\text{O}_4$.requires 194.0529, found 194.0521; $^{18}$O-scopoletin $\text{C}_{10}\text{H}_{9}\text{O}_3^{18}\text{O}$ requires 195.0538, found 195.0537. In the transgenic extract: scopoletin $\text{C}_{10}\text{H}_{9}\text{O}_4$ requires 193.0495, found 193.0486; and $^{18}$O-scopoletin $\text{C}_{10}\text{H}_{9}\text{O}_3^{18}\text{O}$ requires 195.0538, found 195.0634 (50 ppm) (Figure 4.25).

Fig. 4.25. HR-MS spectra of scopoletin and its isotope $^{18}$O-scopoletin after feeding with $^{18}$O-esculetin. A shows a spectrum for the wild type extract, B shows a spectrum of the transgenic extract. The naturally occurring $^{13}$C-scopoletin $\text{C}_9^{13}\text{CH}_9\text{O}_4$ was detected in both spectra.
For scopolin in the wild type extract: scopolin $C_{16}H_{19}O_9$ requires 355.1024, found 355.1019; $^{13}$C-scopolin $C_{15}$ $^{13}$CH$_{19}$O$_9$ requires 356.1057, found 356.1059; $C_{16}H_{19}O_8^{18}$O requires 357.1066, found 357.1072. In the transgenic extract: scopolin $C_{16}H_{19}O_9$ requires 355.1024, found 355.1015; $^{13}$C-scopolin $C_{15}$ $^{13}$CH$_{19}$O$_9$ requires 356.1057, found 356.1053; $C_{16}H_{19}O_8^{18}$O requires 357.1066, found 357.1097 (Figure 4.26). Not unexpectedly, neither $^{18}$O-scopoletin nor $^{18}$O-scopolin was detected in the control (wild type with no feeding), but the expected naturally occurring $C_{10}H_9O_4$ and $C_{16}H_{19}O_9$ were detected (in the negative control) showing that the experiments were working.

![HR-MS spectra of scopolin and its isotope $^{18}$O-scopolin after feeding with $^{18}$O-esculetin.](image)

**A** shows a spectrum of the wild type extract, **B** shows a spectrum of the transgenic extract. The naturally occurring $^{13}$C-scopolin $C_{15}$ $^{13}$CH$_{19}$O$_9$ was detected in both spectra.

The obtained results showed little or no scopoletin obtained through direct methylation of esculetin in both the wild type and the transgenic roots. The LC-MS error limits up to 0.005 Da fails to separate the $^{18}$O-scopoletin $C_{10}H_9O_3^{18}$O and the naturally occurring 2 x $^{13}$C-scopoletin $C_8$ $^{13}$C$_2$H$_9$O$_4$ within the 0.005 Da mass error.
range; C$_{10}$H$_9$O$_3$$^{18}$O requires 195.0538 ($\pm$ 0.005 Da: 195.0488-195.0588), and 2 x$^{13}$C-scopoletin C$_8$$^{13}$C$_2$H$_9$O$_4$ requires 195.0548 ($\pm$ 0.005 Da: 195.0498-195.0598) (Figure 4.27). The theoretical natural occurrence of 2 x$^{13}$C-scopoletin is 1.2% (1.1 x 1.1%), whilst the scopoletin peaks found for [M +2]$^+$ are ~1%.

![Graph showing molecular weights of scopolin and scopoletin isomers](image)

Fig. 4.27. The overlap between $^{18}$O-scopoletin C$_{10}$H$_9$O$_3$$^{18}$O and 2 x$^{13}$C-scopoletin C$_8$$^{13}$C$_2$H$_9$O$_4$ accurate molecular weights within 0.005 Da mass error.

The detection of labelled scopolin in both wild type and transgenic roots, clearly confirms the involvement of other alternative pathways in scopolin biosynthesis. The accumulation and unambiguous detection of $^{18}$O-scopolin (~20%) in cassava root could be explained by other pathways such as those shown in Figure 4.28. Glycosylation of esculetin to produce iso-esculin (7-$\beta$-glucoside of esculetin), then 6-$\beta$-methylation to yield scopolin, which could de-glycosylated to release the aglycone scopoletin if needed. Another possibility of double $\beta$-glycosylation of esculetin to yield 6,7-esculetin-di-$\beta$-glycoside, and then losing one glucose unit from the 6 position to produce scopolin. These possibilities could partly explain the accumulation of scopolin in the transgenic roots despite the decrease in scopoletin levels (Figure 4.7).

$\beta$-Glycosylation of esculetin into its $\beta$-glycoside, esculin or isoesculin or even into esculin-diglycoside is mediated by glucosyltransferase (GT) enzymes, to produce a more soluble and less toxic form which is then stored in the plant vacuoles (Bowles et al., 2005). The hydrolysis of the glycosidic form is mediated by glucoside hydrolase (GH) enzymes to release the aglycone form. According to the plant carbohydrate-active enzymes database PlantCAZyme, published in 2014, (http://cys.bios.niu.edu/plantcazyme/, accessed on 01.08.2017), there are more than
153 glucosyltransferase genes and more than 58 glycoside hydrolase genes in cassava. Anyhow, none of the scopoletin and esculetin glucosyltransferase enzymes, scopolin, esculin, isoesculin, and 6,7-esculetin-di-O-glycoside, glycoside hydrolase enzymes are fully investigated yet in cassava and the identities of the enzymes were based on their sequence similarities with reference genes in other plants. (Ekstrom et al., 2004). The investigation of GT and GH enzymes is beyond the aims of this project.

**Fig. 4.28. Proposed metabolic pathway in the biosynthesis of scopoletin.**

Esculetin is biosynthesized either from umbelliferone or from 6′-hydroxycaffeate, esculetin could glycosylate on the 7-hydroxy to produce isoesculin, followed by 6-\(O\)-methylation to produce scopolin which is deglycosylated to produce scopoletin. Also esculin could be glycosylated into 7-\(O\)-glucoesculin, deglycosylated into isoesculin, then \(O\)-methylated to scopolin.
However, the accumulation of both scopoletin and scopolin through interconnected pathways merits further investigations to define the main pathway and to explore the plant responses after silencing key enzymes on each pathway. To explore this issue, transgenic *A. thaliana* plants, in which the enzymes of interest were fully knocked-out, were fed separately with stable isotopically labelled intermediates: cinnamic-d7 acid, ferulic-OCD3 acid, 18O-ferulic acid, and 18O-esculetin to follow the flux and to determine the origin of the scopoletin biosynthesized in each mutant plant. The transgenic lines used were A.t-F6´H1, with no feruloyl 6´-hydroxylase activity, and A.t-EOMT, with no esculetin O-methyltransferase activity. Wild type *A. thaliana* plant was used as a control.

The incorporation of cinnamic-d7 acid in scopoletin and scopolin biosynthesis in wild type and transgenic *A. thaliana* plants

Scopoletin-d3 and scopolin-d3 were detected as the major isotope peaks in the leaf and root extracts, as well as peaks assigned to the naturally occurring scopoletin and scopolin in the wild type and the A.t-EOMT *A. thaliana* plants. As the results obtained were similar to those from cassava, scopoletin was concluded to be biosynthesized through the same pathway/s in both *A. thaliana* and cassava. The loss of the fourth deuterium favours the hypothesis of the enzymatic isomerization of the hydroxycinnamates (pathway-dependent) over the spontaneous occurrence in *A. thaliana* (Bayoumi et al., 2008b).

The percentage incorporation of the labelled to the naturally occurring scopoletin (scopoletin-d3: scopoletin) in the leaf extracts were found to be higher than those in the root extracts in both the wild type and A.t-EOMT plants. For example in the wild type leaf extract, scopoletin-d3: scopoletin was ~70%, and in the root extract the ratio found to be ~30% (Figure 4.29). Neither scopoletin nor scopolin was detected in the in the A.t-F6´H1 extracts after feeding with cinnamic-d7 acid, this perhaps refers to the negative feedback inhibition of the downstream reactions after the increase in the cinnamates level, a major precursor.

Ratios varied between different plant line extracts, and between biological replicates within the same line, but scopoletin-d3 and scopolin-d3 were the major isotopomers found in all extracts. The higher ratio detected in the leaf extracts was expected because scopoletin is biosynthesized and then transported to be stored in
the plant vacuoles in its β-glycoside form (Rataboul et al., 1985; Ahn et al., 2010), therefore, their minimal accumulation in the leaves accounted for the high ratios.

Fig. 4.29. HR-MS spectra of scopoletin and its isotopomers after feeding the wild type A. thaliana with cinnamic-d7 acid. A shows the leaf extract and B shows the root extract spectra. The naturally occurring 13C-scopoletin (C913CH2O4) was detected in the root extract as indicated.

After feeding the wild type A. thaliana with cinnamic-d7 acid, HR-MS m/z ratios in the leaf extract (Figure 4.29 A) were: scopoletin C10H9O4 requires 193.0495, found 193.0493; scopoletin-d3 C10H6D3O4 requires 196.0684, found 196.0677. Scopoletin-d1, scopoletin-d2, and scopoletin-d4 were not detected in any of the extracts within the allowed 10 ppm mass errors. Scopoletin-d1 C10H8D1O4 requires 194.0558, found 194.1173 (317 ppm); scopoletin-d2 C10H7D2O4 requires 195.0621, found 195.1220 (307 ppm); scopoletin-d4 C10H5D4O4 requires 197.0746, found 197.1282 (272 ppm). Moreover, their LC peaks did not overlap the naturally occurring scopoletin peak (Figure 4.30). From the wild type root extract (Figure 4.29 B), HR-MS m/z ratios were: scopoletin C10H9O4 requires 193.0495, found
193.0478; $^{13}$C-scopoletin $C_9^{13}$CH$_2$O$_4$ requires 194.0529, found 194.0510; scopoletin-d$_3$ $C_{10}$H$_6$D$_3$O$_4$ requires 196.0684, found 196.067.

Fig. 4.30. Selective LC-MS chromatogram for scopoletin and its isotopomers in the root extract after feeding wild type *A. thaliana* with cinnamic-d$_7$ acid. The naturally occurring scopoletin, $^{13}$C-scopoletin, and the feeding product scopoletin-d$_3$ peaks overlapped.

The amount of scopolin (ng) was in the noise level in the leaf extracts of the transgenic lines, although low levels of the naturally occurring scopolin were detected in the wild type leaf extract. Scopolin-d$_3$ was detected as the major peak in the root extracts of both the wild type and A.t-EOMT plants, beside the naturally occurring scopolin.

From the wild type root extract, HR-MS $m/z$ ratios were: scopolin $C_{16}$H$_{19}$O$_9$ requires 355.1024, found 355.1034; scopolin-d$_3$ $C_{16}$H$_{16}$D$_3$O$_9$ requires 358.1212, found 358.1209. In the A.t-EOMT root extract, $m/z$ ratios were: scopolin $C_{16}$H$_{19}$O$_9$ requires 355.1024, found 355.1044; scopolin-d$_3$ $C_{16}$H$_{16}$D$_3$O$_9$ requires 358.1212, found 358.1208 (Figure 4.31). None of the other peaks was considered as a scopolin isotopomer because of the high mass error. Variable ratios were obtained between the different lines and even between the biological replicates, but it was higher in the leaf extracts in all samples.

The biosynthesis of scopoletin-d$_3$ and scopolin-d$_3$, after feeding the growing *A. thaliana* plants with cinnamic-d$_7$ acid, indicated the success of the designed feeding method with respect to the ability of growing *A. thaliana* to take up the exogenous intermediates. In the control wild-type plant (no feeding), no scopoletin
and scopolin isotopomers were detected except the natural abundance $^{13}$C-scopoletin and $^{13}$C-scopolin.

![HR-MS spectra of scopolin and scopolin-d$_3$ after feeding cinnamic-d$_7$ acid](image)

**Fig. 4.31.** HR-MS spectra of scopolin and scopolin-d$_3$ after feeding cinnamic-d$_7$ acid. A: the root extract of the wild type. B: the root extract of A.t-EOMT.

In a series of designed experiments, ferulic-OCD$_3$ acid, $^{18}$O-ferulic acid, and $^{18}$O-esculetin were then fed separately to the wild type and transgenic *A. thaliana* plants in order to explore the origin of the residual scopoletin biosynthesized in the absence of key enzymes on the phenylpropanoid pathways.

The incorporation of ferulic-OCD$_3$ acid in scopoletin and scopolin biosynthesis in both wild type and A.t-F6’H1 lines

After feeding the wild type *A. thaliana* with ferulic-OCD$_3$ acid, both scopoletin-d$_3$ and scopolin-d$_3$ were detected as the major isotope peaks in the leaf (Figure 4.32) and root (Figure 4.33) wild type extracts, as well as peaks assigned to the naturally occurring scopoletin and scopolin. The high incorporation of the
labelled ferulic-OCD₃ acid in scopoletin biosynthesis in the wild-type plant confirms its role as a late precursor to scopoletin in the main biosynthetic pathway, compared to the lower incorporation of the early precursor cinnamate acid (Figure 4.1).

From the wild type leaf extract after feeding ferulic-OCD₃ acid, HR-MS m/z ratios were: scopoletin C₁₀H₉O₄ requires 193.0495, found 193.0489; scopoletin-d₃ C₁₀H₉D₃O₄ requires 196.0684, found 196.0671; scopolin C₁₆H₁₉O₉ requires 355.1024, found 355.1033; scopolin-d₃ C₁₆H₁₆D₃O₉ requires 358.1212, found 358.1213 (Figure 4.32).

Fig. 4.32. HR-MS spectra from leaf extract of the wild type A. thaliana after feeding ferulic-OCD₃ acid. A shows scopoletin and scopoletin-d₃, B shows scopolin and scopolin-d₃.

From the wild type root extract (Figure 4.33), HR-MS m/z ratios were: scopoletin C₁₀H₉O₄ requires 193.0495, found 193.0490; ¹³C-scopoletin C₉¹³CH₉O₄
requires 194.0529, found 194.0523; scopoletin-d$_3$ C$_{10}$H$_6$D$_3$O$_4$ requires 196.0684, found 196.0678; scopolin C$_{16}$H$_{19}$O$_9$ requires 355.1024, found 355.1027; $^{13}$C-scopolin C$_{15}$CH$_{16}$O$_9$ requires 356.1057, found 356.1057; scopolin-d$_3$ C$_{16}$H$_{19}$D$_3$O$_9$ requires 358.1212, found 358.1213.

Fig. 4.33. HR-MS spectra from the wild type root extract after feeding ferulic-OCD$_3$ acid. A shows scopoletin and its isotopomers. B shows scopolin and its isotopomers.

It was expected to detect both scopoletin-d$_3$ and scopolin-d$_3$ in the wild type extracts after feeding ferulic-OCD$_3$ acid. The exogenous ferulate-OCD$_3$ could be converted to its CoA ester by a CoA ligase enzyme, then feruloyl-OCD$_3$ CoA or/and ferulate-OCD$_3$ will be 6'-hydroxylated to form 6'-hydroxyferulate-OCD$_3$ or/and 6'-hydroxyferuloyl-OCD$_3$ CoA, by F6'H1, the product will be isomerized, and then ring closed to the lactone to produce scopoletin-d$_3$ which will be further glycosylated into scopolin-d$_3$ by scopoletin GT enzyme (Figure 4.34).
In A.t-F6'H1 plant, where the activity of the F6'H1 enzyme is knocked out, no labelled scopoletin was expected to be detected after feeding ferulic-OCD$_3$. This is because the *ortho*-hydroxylation step is essential in scopoletin biosynthesis. However, the obtained experimental results did not match these expectations.

![Chemical diagram]

**Fig. 4.34. The incorporation of ferulic-OCD$_3$ acid into scopoletin-d$_3$ and scopolin-d$_3$ in wild type *A. thaliana.*** The labelled ferulate or/and its CoA ester will be *ortho*-hydroxylated to form 6'-hydroxyferulate-OCD$_3$, or/and 6'-hydroxyferuloyl-OCD$_3$ CoA, isomerized, and then lactonized to form scopoletin-d$_3$ then scopolin-d$_3$.

After feeding the transgenic A.t-F6'H1 line with ferulic-OCD$_3$ acid, the lower limit of detecting scopoletin and scopolin peaks was in the noise level. No scopoletin and scopolin isotopomers were detected in the A.t-F6'H1 leaf extract. In the root extract of A.t-F6'H1, a low level of scopolin was detected. Scopoletin-d$_3$ was the major isotope peak detected in the root extract of A.t-F6'H1 plant, in addition to the naturally occurring scopoletin. Moreover, a high percentage incorporation of the labelled intermediate was observed, scopoletin-d$_3$: scopoletin was ~140% (Figure 4.35). From the A.t-F6'H1 root extract, HR-MS *m/z* ratios were: scopoletin C$_{10}$H$_6$O$_4$ requires 193.0495, found 193.0488; scopoletin-d$_3$ C$_{10}$H$_6$D$_3$O$_4$ requires 196.0684, found 196.0676.

The detection of labelled scopoletin in the absence of F6'H1 activity after feeding a late biosynthetic precursor, can only be explained by the involvement of other enzymes with similar hydroxylase activity. In *A. thaliana*, F6'H2, with 77%
amino acid sequence similarity to F6´H1, has a slight, but consistent, expression in
the roots, even though, knocking it out did not reduce scopoletin concentration in the
roots. The abiotic stressor 2,4-dichlorophenoxyacetic acid also failed to induce its
expression (Kai et al., 2008), but it could be involved in ferulate hydroxylation, and
therefore in scopoletin biosynthesis.

![HR-MS spectra of scopoletin and scopoletin-d3 from the A.t-F6´H1 root extract after feeding ferulic-OCD3 acid.](image)

Fig. 4.35. HR-MS spectra of scopoletin and scopoletin-d3 from the A.t-F6´H1 root extract after feeding ferulic-OCD3 acid.

The high scopoletin-d3: scopoletin ratio (1.4: 1) reflects the low pool of the
naturally occurring scopoletin in the transgenic roots (~3% of those in the wild type).
These findings confirm the involvement of another hydroxylase enzyme in the
biosynthesis of the residual scopoletin in A.t-F6´H1 mutant plants. In order to
investigate if other routes are also involved, e.g. by O-methylation of esculetin, a
feeding experiment with labelled esculetin was designed and performed in this
mutant plant line.

The incorporation of $^{18}$O-ferulic acid in scopoletin and scopolin biosynthesis in A.t-
EOMT plant

As expected, $^{18}$O-scopoletin and $^{18}$O-scopolin were detected in the root
extract of A.t-EOMT plants after feeding with $^{18}$O-ferulic acid, beside the naturally
occurring scopoletin and scopolin and their $^{13}$C-isotopomers. HR-MS m/z ratios
were: scopoletin \(C_{10}H_9O_4\) requires 193.0495, found 193.0493; \(^{13}\text{C}\)-scopoletin \(C_9^{13}\text{CH}_9O_4\) requires 194.0529, found 194.0524; \(^{18}\text{O}\)-scopoletin \(C_{10}H_9O_3^{18}\text{O}\) requires 195.0538, found 195.0538; scopolin \(C_{16}H_{19}O_9\) requires 355.1024, found 355.1024; \(^{13}\text{C}\)-scopolin \(C_{15}^{13}\text{CH}_9O_9\) requires 356.1057, found 356.1054; \(^{18}\text{O}\)-scopolin \(C_{16}H_{19}O_8^{18}\text{O}\) requires 357.1066, found 357.1070 (Figure 4.36).

\[\text{Fig. 4.36. HR-MS spectra of scopoletin, scopolin from the A.}\text{-EOMT root extract after feeding }^{18}\text{O-ferulic acid.}\]

A shows scopoletin and its isotopomers. B shows scopolin and its isotopomers.

The incorporation of \(^{18}\text{O}\)-esculetin acid in scopoletin, scopolin, and esculin biosynthesis in the wild type and the transgenic lines

From the previous findings (chapter 3), biosynthesis of scopoletin after esculetin \(O\)-methylation has a low impact on scopoletin accumulation, ~20% in A. \textit{thaliana} roots. For a more comprehensive view of the possible role of esculetin, feeding labelled esculetin experiments were designed and performed in the A.
*Arabidopsis thaliana* wild type and transgenic lines. The incorporation of $^{18}$O-labelled esculetin in scopoletin and therefore in scopolin biosynthesis was investigated.

After feeding the wild type *A. thaliana* with $^{18}$O-labelled esculetin, $^{18}$O-scopoletin, $^{18}$O-scopolin, and $^{18}$O-esculin were detected as the major isotope peaks beside the naturally occurring scopoletin, scopolin, esculin, and their corresponding $^{13}$C-isotopomers.

In the wild type leaf extract (Figure 4.37), HR-MS $m/z$ ratios were: scopoletin $C_{10}H_{9}O_{4}$ requires 193.0495, found 193.0487; $^{13}$C-scopoletin $C_9^{13}CH_9O_4$ requires, 194.0529 found 194.0520; $^{18}$O-scopoletin $C_{10}H_9O_4^{18}$O requires 195.0538, found 195.0532; scopolin $C_{16}H_{19}O_9$ requires 355.1024, found 355.1006; $^{13}$C-scopolin $C_{15}^{13}CH_{19}O_9$ requires 356.1057, found 356.1037; $^{18}$O-scopolin $C_{16}H_{19}O_8^{18}$O requires 357.1066, found 357.1044; $C_{15}^{13}CH_{19}O_8^{18}$O requires 358.1085; esculin $C_{15}H_{17}O_9$ requires 341.0873, found 341.0842; $^{18}$O-esculin $C_{15}H_{17}O_8^{18}$O requires 343.0915, found 343.0886; $C_{15}^{13}CH_{17}O_8^{18}$O requires 344.0949, found 344.0931.

In the wild type root extract (Figure 4.38), $^{18}$O-scopoletin, $^{18}$O-scopolin, and $^{18}$O-esculin were detected as the major isotope peaks beside the naturally occurring scopoletin, scopolin, esculin, and their corresponding $^{13}$C-isotopomers. HR-MS $m/z$ ratios were: scopoletin $C_{10}H_{9}O_{4}$ requires 193.0495, found 193.0490; $^{13}$C-scopoletin $C_9^{13}CH_9O_4$ requires, 194.0529 found 194.0521; $^{18}$O-scopoletin $C_{10}H_9O_4^{18}$O requires 195.0538, found 195.0531; scopolin $C_{16}H_{19}O_9$ requires 355.1024, found 355.0997; $^{13}$C-scopolin $C_{15}^{13}CH_{19}O_9$ requires 356.1057, found 356.1027; $^{18}$O-scopolin $C_{16}H_{19}O_8^{18}$O requires 357.1066, found 357.1037; $C_{15}^{13}CH_{19}O_8^{18}$O requires 358.1100, found 358.1072; esculin $C_{15}H_{17}O_9$ requires 341.0873, found 341.0840; $^{13}$C-esculin $C_{14}^{13}CH_{17}O_9$ requires 342.0906, found 342.0871; $^{18}$O-esculin $C_{15}H_{17}O_8^{18}$O requires 343.0915, found 343.0883; $C_{14}^{13}CH_{17}O_8^{18}$O requires 344.0949, found 344.0913.
Fig. 4.37. HR-MS spectra of scopoletin, scopolin, and esculin from the wild type leaf extract after feeding $^{18}$O-esculetin. A shows scopoletin and its isotopomers. B shows scopolin and its isotopomers. C shows esculin and its isotopomers.
Fig. 4.38. HR-MS spectra of scopoletin, scopolin, and esculin from the wild type root extract after feeding $^{18}$O-esculetin. A shows scopoletin and its isotopomers. B shows scopolin and its isotopomers. C shows esculin and its isotopomers.

Both $^{18}$O-scopoletin and $^{18}$O-esculin were biosynthesized after adding a methyl group or a glucose unit on the 6-hydroxyl group of the exogenous esculentin by EOMT and GT respectively. The detection of $^{18}$O-scopolin could be explained
either by the 7-O-glycosylation of $^{18}$O-scopoletin, or by other routes such as those explained in Figure 4.28. The β-glycoside of esculetin was considered as esculin according to previous studies on cassava and A. thaliana (Buschmann et al., 2000a; Bayoumi et al., 2008a). The detected and isolated amount (ng) was insufficient for NMR spectroscopy to confirm whether it is esculin or isoesculin (cicohorin, chicorin, chikorin, 7-O-glucoesculetin), and the unavailability of the standard isoesculin did not allow identification by HPLC separation. Whether the labelled esculetin was incorporated into $^{18}$O-esculin or $^{18}$O-isoesculin, the product could be further metabolised into $^{18}$O-scopolin: by direct O-methylation of $^{18}$O-isoesculin on the 6-hydroxyl group by OMT, or by glycosylation of esculin into 7-O-glucoesculin by GT, followed by deglycosylation of the diglycoside to produce isoesculin by GH, then adding a methyl group on the 6-hydroxyl group of the latter product to produce $^{18}$O-scopolin (Figure 4.28). An isomerisation reaction mediated by transglucosidase (TG) to yield the lower energy, therefore more stable isomer 6-O-glucoesculetin (esculin) from 7-O-glucoesculetin (isoesculin), has been previously reported in chickory Cichorium intybus flowers (Sato and Hasegawa, 1972; Ueno et al., 1985). Both esculin and isoesculin were previously isolated and characterised in elegant biosynthetic studies in Solanum pinnatisectum (Harborne, 1960). These results confirm the involvement of different pathways in scopoletin biosynthesis in wild type A. thaliana.

From earlier results, another hydroxylase enzyme was concluded to be involved in the residual scopoletin biosynthesized in the absence of F6’H1 activity in the mutant plant A.t-F6’H1. In order to investigate whether other routes are involved, e.g. O-methylation of esculetin, the mutant plant was fed with $^{18}$O-esculetin and the formation of labelled scopoletin, scopolin, and esculin was explored.

In the leaf extract of A.t-F6’H1 plant, high incorporations of the labelled intermediate into scopoletin, scopolin, and esculin were detected, after feeding with $^{18}$O-labelled esculetin. $^{18}$O-Scopoletin, $^{18}$O-scopolin, and $^{18}$O-esculin were the major isotope peaks detected in addition to the peaks assigned for the naturally occurring corresponding hydroxycoumarins. The high ratios of $^{18}$O-scopoletin: scopoletin and $^{18}$O-scopolin: scopolin indicate the small endogenous pool of the hydroxycoumarins of interest occurring naturally in the leaves (Figure 4.39).
Fig. 4.39. HR-MS spectra of scopoletin, scopolin, and esculin from the A.t-F6'H1 leaf extract after feeding $^{18}$O-esculetin. A shows scopoletin and its isotopomers. B shows scopolin and its isotopomers. C shows esculin and its isotopomers.

In the leaf extract of A.t-F6'H1 plant after feeding with $^{18}$O-esculetin, HR-MS $m/z$ ratios were: scopoletin C$_{10}$H$_{9}$O$_{4}$ requires 193.0495, found 193.0479; $^{18}$O-scopoletin C$_{10}$H$_{9}$O$_{4}^{18}$O requires 195.0538, found 195.0523; C$_{9}$H$_{7}$O$_{3}^{18}$O requires 196.0572, the peak found 196.0595 (11.7 ppm); scopolin C$_{16}$H$_{19}$O$_{9}$ requires
355.1024, found 355.1019; $^{13}$C-scopolo C$_{15}^{13}$CH$_{19}$O$_9$ requires 356.1057, found 356.1069; $^{18}$O-scopolo C$_{16}$H$_{19}$O$_8^{18}$O requires 357.1066, found 357.1047; esculin C$_{15}$H$_{17}$O$_8$ requires 341.0873, found 341.0844; $^{18}$O-esculin C$_{15}$H$_{17}$O$_8^{18}$O requires 343.0915, found 343.0890; C$_{14}^{13}$CH$_{17}$O$_8^{18}$O requires 344.0949, found 344.0951. None of the other detected peaks indicate other isotopomers because their LC peaks did not overlap with the naturally occurring hydroxycoumarin peaks, in addition to the high mass error. HR-MS $m/z$ ratios were: C$_9^{13}$CH$_9$O$_4$ requires 194.0529, the peak found 194.1160 (325 ppm); $^{13}$C-esculin C$_{16}^{13}$CH$_{17}$O$_9$ requires 342.0906, the peak found 342.1017 (32 ppm), C$_{13}^{13}$C$_2$H$_{17}$O$_8^{18}$O requires 345.0982, found 345.1410 (124 ppm).

The same major isotope peaks were detected in the root extract of A.t-F6´H1 plants after $^{18}$O-esculetin feeding (Figure 4.40). HR-MS $m/z$ ratios were: scopolo C$_{10}$H$_9$O$_4$ requires 193.0495, found 193.0479; $^{13}$C-scopolo C$_{15}^{13}$CH$_9$O$_4$ requires 194.0529, found 194.0510; $^{18}$O-scopolo C$_{16}$H$_{19}$O$_9^{18}$O requires 195.0538, found 195.0522; C$_9^{13}$CH$_9$O$_3^{18}$O requires 196.0572, found 196.0556; scopolo C$_{16}$H$_{19}$O$_9$ requires 355.1024, found 355.0997; $^{13}$C-scopolo C$_{15}^{13}$CH$_{19}$O$_9$ requires 356.1057, found 356.1029; $^{18}$O-scopolo C$_{16}$H$_{19}$O$_8^{18}$O requires 357.1066, found 357.1045; C$_{15}^{13}$CH$_{17}$O$_8^{18}$O requires 358.1100, found 358.1090; C$_{14}^{13}$C$_2$H$_{19}$O$_8^{18}$O requires 359.1133, found 359.1095; esculin C$_{15}$H$_{17}$O$_9$ requires 341.0873, found 341.0842; $^{13}$C-esculin C$_{14}^{13}$CH$_{17}$O$_9$ requires 342.0906, found 342.0873; $^{18}$O-esculin C$_{13}$H$_{17}$O$_8^{18}$O requires 343.0915, found 343.0885; C$_{14}^{13}$CH$_{17}$O$_8^{18}$O requires 344.0949, found 344.0917. None of the other peaks in the mass spectra indicate other isotopomers because their LC peaks did not overlap with the naturally occurring hydroxycoumarin peaks, in addition to the high mass error observed.

In order to investigate the other possible routes in scopolo biosynthesis through esculin O-methylation and/or glycosylation, the transgenic A. thaliana plants A.t-EOMT where the methylation step of esculetin to scopolo is inhibited, was fed with labelled esculetin and the flux was followed. As expected, no labelled scopolo was detected in the methanolic extracts of both leaf and root parts of A.t-EOMT plant. No labelled scopolo was detected in the leaf extract, but a ratio of 1:10 ($^{18}$O-scopolo: scopolo) was measured in the root extract. In both the plant parts extracts, a high incorporation of the labelled esculetin into its β-glycoside esculin was found; the ratio $^{18}$O-esculin: esculin was 1.4:1.
Fig. 4.40. HR-MS spectra of scopoletin, scopolin, and esculin from the A.t-F6’H1 root extract after feeding $^{18}$O-esculetin. A shows scopoletin and its isotopomers. B shows scopolin and its isotopomers. C shows esculin and its isotopomers.

In the A.t-EOMT leaf extract after feeding $^{18}$O-esculetin, HR-MS $m/z$ ratios were: esculin $C_{15}H_{17}O_9$ requires 341.0873, found 341.0869; $^{13}$C-esculin $C_{14}^{13}CH_{17}O_9$ requires 342.0906, found 342.0901; $^{18}$O-esculin $C_{15}H_{17}O_8^{18}$O requires 343.0915, found 343.0912; $C_{14}^{13}CH_{17}O_8^{18}$O requires 344.0949, found 344.0942;
C_{13}^{13}C_{2}H_{17}O_{8}{^{18}}O requires 345.0982, found 345.0957 (Figure 4.41). The LC chromatogram of the naturally occurring esculin and its isotopomers is shown in Figure 4.42.

![Figure 4.41. HR-MS spectra of esculin and its isotopomers from the A.t-EOMT leaf extract after feeding $^{18}$O-esculetin.](image1)

![Figure 4.42. Selective LC-MS chromatogram for esculin and its isotopomers in the A.t-EOMT leaf extract after feeding $^{18}$O-esculetin. The split peak indicated by the red star could be isoesculin.](image2)

In the At-EOMT root extract, HR-MS m/z ratios were: scopolin C_{16}H_{19}O_{9} requires 355.1024, found 355.1003; $^{13}$C-scopolin C_{15}^{13}CH_{19}O_{9} requires 356.1057, found 356.1033; $^{18}$O-scopolin C_{16}H_{19}O_{8}{^{18}}O requires 357.1066, found 357.1044; C_{15}^{13}CH_{19}O_{8}{^{18}}O requires 358.1100, found 358.1094; esculin C_{15}H_{17}O_{9} requires
341.0873, found 341.0872; $^{13}$C-esculin $C_{14}^{13}CH_{17}O_9$ requires 342.0906, found 342.0903; $^{18}$O-esculin $C_{15}H_{17}O_8^{18}O$ requires 343.0915, found 343.0915; $C_{14}^{13}CH_{17}O_8^{18}O$ requires 344.0949, found 344.0946; $C_{13}^{13}C_2H_7O_8^{18}O$ requires 345.0982, found 345.0960 (Figure 4.43). The LC chromatogram of the naturally occurring esculin and its isotopomers is shown in Figure 4.44.

Fig. 4.43. HR-MS spectra of scopolin and esculin from the A.t-EOMT root extract after feeding $^{18}$O-esculetin. A shows scopolin and its isotopomers. B shows esculin and its isotopomers.
Fig. 4.44. Selective LC-MS chromatogram for esculin and its isotopomers in the A. t-EOMT root extract after feeding $^{18}$O-esculetin. The split peak indicated by the red star could be isoesculin.

Using the same chromatography (Figure 4.45), authentic esculin did not show the split in the LC peak as the isolated esculin. The split peak could be evidence for the occurrence of the natural structural isomer isoesculin in the plant extract (Sato and Hasegawa, 1972). This supports the involvement of the proposed scopolin biosynthetic pathway shown in Figure 4.28.

Fig. 4.45. LC chromatogram of authentic esculin. The chromatogram shows a non-split peak of the authentic esculin (1 µg/ml).

The obtained results confirmed the specificity of the selected gene towards esculetin $O$-methylation, and indicated $O$-glycosylation of scopoletin may not be the only biosynthetic pathway to scopolin.
The overall conclusions of the results obtained from the feeding experiments in transgenic and wild type cassava and *A. thaliana* plants confirm that scopoletin is biosynthesized in a similar pattern in both cassava and *A. thaliana*. Three deuterium atoms are lost in three oxidative hydroxylation steps. The loss of the 4th deuterium from the exogenous intermediate cinnamic-d7 acid could be during the E-Z-isomerization step, suggesting that it is enzyme catalysed and not a spontaneously occurring photochemical reaction (Bayoumi *et al*., 2008b). Inhibiting or reducing the activity of the cassava F6′H enzyme should significantly reduce scopoletin accumulation, and therefore reduce the physiological symptoms of PPD. Scopoletin is biosynthesised via interconnected pathways, even though the pathway through ferulate hydroxylation has the biggest share, but other alternative pathways do participate including: via *E*-caffeate, and/or *E*-2′,4′-dihydroxycinnamate intermediates as well as after esculetin O-glycosylation.

From the findings in chapter 3, the functions of cassava CCoAOMT and cassava EOMT were confirmed after their ability to complement the mutation in the model plant *A. thaliana*. This therefore restored the ability to biosynthesize scopoletin. To confirm further the functions of these two cassava OMTs, and the cassava EOMT regiospecificity as well, *in vivo* experiments were designed and performed (chapter 5) in order to explore the O-methylation affinity of each enzyme towards different intermediates on the phenylpropanoid pathways after separately expressing these OMTs in chemically competent *E. coli* cells BL21.
Chapter 5. Functional expression of cassava O-methyltransferase enzymes in transgenic *Escherichia coli*

5.1. Introduction

In chapters 3 and 4 the activities of CCoAOMT and EOMT enzymes in both *A. thaliana* and cassava were explored through studying their influence on the profiles of the coumarins of interest by using different genetic tools and tracing the flux through each pathway using compounds labelled with stable isotopes. In this chapter, the activities of cassava CCoAOMT and EOMT towards different potential substrates along the phenylpropanoid pathways were explored through the expression of the selected genes in *Escherichia coli*. The use of such a transgenic bacterial expression system both complements and adds dimensions to the *in planta* research studies.

**OMTs in cassava and *A. thaliana***

The cassava genome sequence was recently generated and published in Phytozome research tool [https://phytozome.jgi.doe.gov/pz/portal.html](https://phytozome.jgi.doe.gov/pz/portal.html) (accessed on 01.08.2017) (Wang *et al.*, 2014). The published cassava genome sequences allow the prediction of cassava OMT function in reference to the *A. thaliana* OMT-genes depending on the sequence homology between them. However, sequence homology does not necessarily result in functional similarity due to the fact that only one or a few different amino acids between the similar proteins could change the substrate preferences or the overall protein activity (Gauthier *et al.*, 1998; Frick and Kutchan, 1999). Determination of the exact function of the selected gene could be possible by exploring gene-to-metabolite correlation through metabolite accumulation analysis and RT-PCR studies of the levels of gene expression (Tohge *et al.*, 2007).

The *A. thaliana* ecotype Columbia (Col-0) genome sequence was first completed and published in The Institute for Genome Research (TIGER) database in 2000 and most of the genes have been extensively studied (Lin *et al.*, 2000; Lin *et al.*, 2003; Bevan and Walsh, 2005). The boost in genomic technologies and the relatively low-cost sequencing facilities allow frequent and continuous updating in gene structures and functions of *A. thaliana* (Lin *et al.*, 2003).
CCoAOMT

The enzyme CCoAOMT catalyses the transfer of a methyl group from SAM onto caffeoyl CoA ester to produce feruloyl CoA ester. In the *A. thaliana* genome, seven related CCoAOMT-like genes encoding CCoAOMT enzymes were found (Raes *et al*., 2003; Sibout *et al*., 2005), but only one of them was confirmed as CCoAOMT (Do *et al*., 2007). Several separate *in planta* and in *E. coli* studies have been reported to elucidate its function either by analysing the effect of CCoAOMT downregulation in plants or by studying the gene-metabolite-correlation in transgenic *E. coli* (Kai *et al*., 2006; Li *et al*., 2013; Lin *et al*., 2013). Downregulation of A.t-CCoAOMT decreases the stem lignin content in *A. thaliana* plants compared to the wild type (Do *et al*., 2007). The same results were obtained in maize straw after downregulation of CCoAOMT (Li *et al*., 2013). Kai and co-workers, found that downregulation of A.t-CCoAOMT gene showed a decrease scopoletin accumulation in *A. thaliana* roots (~30%) compared to the wild type (WT) (Kai *et al*., 2008). In chapter 3, a ~50% decrease in scopoletin and scopolin concentrations was found in A.t-CCoAOMT mutant line compared to the wild type. These results contribute to confirming the role of the methylated product feruloyl CoA in scopoletin biosynthesis. This leads to the conclusion that CCoAOMT is involved in hydroxycoumarin biosynthesis in addition to its reported participation in lignin formation. This biosynthetic machinery for cassava CCoAOMT genes has not reported from studies in transgenic *E. coli*.

EOMT

There is still a significant lack of agreement in the scientific literature about the function of A.t-EOMT (ATG54160.1). The complete sequence of A.t-EOMT is published in Phytozome research tool and its function is annotated as OMT. https://phytozome.jgi.doe.gov/pz/portal.html#!gene?search=1&detail=1&method=2296&searchText=transcriptid:19667404 (accessed on 01.08.17).

In 1985, Higuchi reported a catalytic activity of A.t-EOMT towards 5´-hydroxyferulic acid and caffeic acid (Higuchi, 1985). Zhang *et al*. also annotated this gene as caffeic acid *O*-methyltransferase (CAOMT) due to its high sequence similarity to *P. tremuloides* (Zhang *et al*., 1997). Humphreys *et al*. used the same notation based on the simple identification of the ferulic acid methylated product by TLC (Humphreys *et al*., 1999). Later on, this was challenged after analysis of the
gene recombinant-protein metabolites from *E. coli* studies. It was re-identified as 5′-hydroxyconiferaldehyde OMT (HCAld-5-OMT) with regards to its role in sinapyl alcohol (SAlc) and syringyl lignin formation (Osakabe *et al*., 1999; Lin *et al*., 1999). Simplified pathways for the biosynthesis of sinapate and sinapyl aldehyde mediated by CAOMT are shown in Figure 5.1. However, the recombinant gene in *E. coli* cannot *O*-methylate caffeic acid (Muzac *et al*., 2000). Therefore, the previous annotation is no longer accepted. After downregulation of A.t-EOMT in *A. thaliana*, the lignin was found to have more 5-hydroxy-guaiacyl units and a decrease in syringyl units (Goujon *et al*., 2003). These findings confirmed that A.t-EOMT is involved in lignin biosynthesis (Figure 5.1).

**Fig. 5.1. The involvement of A.t-OMT in lignin component formation.** OMT substrates include: *E*-caffeate, a precursor to ferulate, 5′-hydroxyferulate, a precursor to sinapate, 5′-hydroxyconiferyl aldehyde, a precursor to sinapyl aldehyde, and 5′-hydroxyconiferyl alcohol to produce sinapyl alcohol. Abbreviations: caffeic acid *O*-methyltransferase (CAOMT), hydroxyconiferaldehyde-5-OMT (CAld5OMT), *Arabidopsis thaliana O*-methyltransferase1 AtOMT1, ferulate hydroxylase (FH), and coniferyl aldehyde 5-hydroxylase (CAlld-5H). Steps mediated by A.t-EOMT are shown in red.
After recombinant-protein studies in *E. coli*, A.t-EOMT was also reported to have specific 3′-O-methylation activity towards quercetin and myricetin among the flavonoids, and no activity was reported towards luteolin which lacks the flavonoid 3′-hydroxy-group (Muzac *et al*., 2000) (Figure 5.2).

![Diagram of flavonoid substrates for A.t-EOMT](image)

**Fig. 5.2. Potential flavonoid substrates for A.t-EOMT.**

In chapter 3, it was shown that downregulation of A.t-EOMT in an *A. thaliana* mutant resulted in a slight decrease in scopoletin accumulation compared to that found in the wild-type plant. In chapter 4, traces of 18O-scopoletin and 18O-scopolin were detected after feeding the wild type *A. thaliana* with stable isotopically labelled 18O-esculetin. From these findings, it was concluded that A.t-EOMT is involved in hydroxycoumarin biosynthesis in *A. thaliana*.

While the identification of the genomic functionality is not entirely dependent on gene sequence homology, the cassava OMT genes, M.e-CCoAOMT2 and M.e-EOMT will be studied in order to explore their function through expressing the separate recombinant genes in genetically engineered *E. coli*. These two recombinant OMT enzymes will be fed with different potential substrates from the phenylpropanoid pathways. These experiments were undertaken using intermediate compounds that were substituted with at least one phenolic group in order to render them potentially good substrates for OMT activity. As enzymes in general, including plant enzymes, are both regio- and stereospecific (Lin *et al*., 2013), the possibility of different isomeric enzymatic products will also be investigated.
5.2. Materials

Cassava OMT genes

Full length cDNA was purified from the overexpression plasmid construct (pCAMBIA 1305.1) synthesized earlier (chapter 3).

Bacterial strains

Chemically competent *E. coli* cells BL21 with a transformation efficiency of 1-5 x 10^7 cfu/µg PUC 19 DNA were used. The cells are sensitive to each of the following antibiotics: ampicillin, carbenicillin, chloramphenicol, kanamycin, streptomycin, and tetracycline. The BL21 competent cells have no Lon protease activity responsible for foreign protein degradation and no OmpT protease activity responsible for extracellular protein degradation (Rosano and Ceccarelli, 2014). Thus, they are used for transformation and protein expression purposes. The cells were purchased from New England Biolabs (NEB). High efficiency NEB 10-Beta *E. coli* cells were used for vector transformation.

Plasmid vector: Glutathione S-Transferase (GST) gene fusion system: PGEX-4T-1

The PGEX-4T-1 overexpression vector of 4.9 kbp size, driven by the tac promoter, was used to induce and express the cDNA insert (up to 2 kbp) as a fusion product with glutathione S-transferase gene (GST). The expanded multiple cloning sites (MCS) allowed the unidirectional cloning of the insert. A double restriction reaction ensured the correct orientation of the insert. A series of site-specific proteases were designed to allow the purification and cleavage of the fusion system. Detection of the expressed protein was allowed by direct labelling of the tagged thrombin protein *in vitro*. The PGEX-4T-1 vector provides the three translational reading frames starting at the *Eco* RI restriction site. Selection of the successful construct depends upon conferring resistance to ampicillin to the insert. The plasmid vector was purchased from GE Healthcare, UK; the vector map is shown in Figure 5.3.
**Fig. 5.3. Map of the PGEX-4-T-1 vector.** Map shows the ampicillin resistant gene, Ptac promoter, thrombin tagged protein, and GST fusion protein. The MCS starting from the reading frame of EcoRI until the stop codon. The map is adapted from PGEX vectors handbook/GE Healthcare.

**DNA restriction digestion and ligation enzymes**

Directional cloning of the selected cDNA was optimized by using two restriction enzymes (RE) to maintain the right orientation of the vector. Endonuclease enzymes were chosen carefully to linearize the vector and not to cut within the insert.

**BamH I**

An endonuclease restriction enzyme derived from *Bacillus amyloliquefaciens* was used to cleave the vector DNA sequence at its recognition site 5'GGATCC3'. The BamH I enzyme produced a sticky end at the edge of the linearized vector.

**Sal I**

An endonuclease restriction enzyme derived from *Streptomyces albus* was used to cleave the DNA sequence of the vector at its recognition site 5'GTCGAC3'. The Sal I enzyme produced a sticky end at the edge of the linearized vector.
T4 DNA ligase enzyme

A ligation enzyme derived from *E. coli* was used to catalyse the generation of a phosphodiester bond between the 5'-phosphate group and the 3'-hydroxyl group. The enzyme ligase was used to join the sticky ends of the vector restriction fragments with the recognition sites of the DNA of interest. All the three enzymes were purchased from New England Biolabs (NEB).

Primers

Short oligonucleotide sequences were designed to add appropriate restriction sites on both the forward and the reverse primers. The restriction sequences should be absent in the sequence of the gene of interest. Six extra guanine nucleotides were added to allow the endonuclease activity of the restriction enzymes. Primers were synthesized and ordered from Fisher. All primer sequences used in this experiment are listed (Table 5.1).

Chemicals and instruments

All solvents were purchased from Fisher. All chemicals and reagents were purchased from Sigma-Aldrich unless otherwise stated. Instruments used for these experiments are described in chapter 3 unless otherwise stated.

Spectrophotometer

Spectronic Helios Gamma UV-Vis spectrophotometer (Fisher) capable of measuring wavelengths between 190-1100 nm in fixed 2 nm steps, was used to measure the optical density (OD) of bacterial cultures at \( \lambda = 600 \) nm.

Plate reader

SkanIt microplate reader from Fisher was used to measure the absorbance of the bacterial cultures growing in 96-well plate over 24 h. Absorbance was recorded every 30 min. and the results were analysed using SkanIT 3.1 software (Fisher).
5.3. General methods

Cloning M.e-OMT genes into PGEX-4-T-1 vector

The full-length cDNAs for both M.e-CCoAOMT2 and M.e-EOMT were amplified by HF-Q5 polymerase enzyme in a thermocycler reaction (40 cycles, gradient TM 52-60 °C, elongation time 1 min). The primer pairs RE-CCoAOMT and RE-EOMT were used to add the specific restriction sites to the M.e-CCoAOMT2 and M.e-EOMT cDNAs respectively. The pCAMBIA 1305.1 constructs were used as templates as described in chapter 3. The primer pairs and their sequences are listed in Table 5.1.

Table 5.1. Designed PCR primer pairs. Restriction sequences are showed in italics

<table>
<thead>
<tr>
<th>primer</th>
<th>sequence (5ʹ-3ʹ)</th>
<th>amplicon size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RE-CCoAOMT</td>
<td>&gt;FwGGGGGGGAGATCCATGGCATCCAACCTGAACAG&lt;br&gt;                             &gt;RevTCACCTTGATCCGACGCGACGTCCGACGGGGGGG</td>
<td>783+RE</td>
</tr>
<tr>
<td>RE-EOMT</td>
<td>&gt;FwGGGGGGGAGATCCATGGGTTCAACCGCC&lt;br&gt;                             &gt;RevTCAGGCCACTTTCCAGAATTCGTCGACGGGGGGG</td>
<td>1098+RE</td>
</tr>
<tr>
<td>PGEX</td>
<td>&gt;FwGGGCTGGCAAGCCACGTGGTG&lt;br&gt;                                                         &gt;RevCCGGGAGCTGCATGTGTCAGAGG</td>
<td>n/a</td>
</tr>
</tbody>
</table>

The PCR amplified product was electrophoresed (0.8% agarose, 70 V, 60 min) and the correct gel-band was excised and purified using QIAquick® Gel extraction kit. The gel-purified DNA was used for performing the ligation reactions. PGEX-4-T-1 plasmid vector (1 µl) was transformed into 10-Beta E. coli competent cells. The cells were plated, cultured, and the plasmid DNA was purified. The purified plasmid DNA was sequenced by Eurofin using the primer pair PGEX (Table 5.1) to confirm that no mutations had been introduced by the PCR or other reactions. The correct cultures (200 µl) were glycerol-stocked in several aliquots by adding an equal amount of 50% aq. glycerol solution, and then stored at -70 °C until required.

The purified PGEX vector plasmid DNA was linearized by double digestion reaction using both BamHI and SalI restriction enzymes. NEBuffer 3.1 was used because it showed the least star activity for both enzymes, therefore, it preserves the specificity of the restriction enzyme activity. The restriction enzyme star activity is
the alteration of the enzyme specificity in non-optimal reaction conditions (Nasri and Thomas, 1986).

According to the manufacturer’s specifications, a double digestion reaction (50 µl) was performed by mixing the purified plasmid DNA (1 µl, 100 ng) with NEBuffer 3.1 (5 µl), *BamH* I (10 U, 1 µl), *SalI* I (10 U, 1 µl), and Milli-Q water (42 µl). The reaction was undertaken at 37 ºC for 15 min before being terminated by heat inactivation (65 ºC, 20 min) to stop *SalI* I digestion activity followed by DNA purification using QIAquick PCR purification kit (QIAGEN) to stop the *BamH* I digestion activity.

According to QIAGEN’s specifications, the vector double-digestion product (100 µl) was mixed with PCR purification buffer (PB) (500 µl, guanidine hydrochloride (5 M), isopropanol (30%)), and mixed thoroughly before being placed into a QIAquick spin column and centrifuged (13,400 rpm, 60 s) to bind the DNA to the column membrane. The membrane was washed with the PE washing buffer (750 µl, tris-HCl (0.1 M, pH 7.5), aqueous ethanol (80%)) and centrifuged twice (13,400 rpm for 60 s each) to ensure that no residual buffer was left. The washed membrane carrying the desired DNA was eluted with Milli-Q water (30 µl). The purified vector was used to build a recombinant construct containing the cDNA of interest.

The ligation reaction between the gel-purified DNA and the linearized PGEX vector was performed according to NEB’s specifications. A total reaction of 20 µl consisted of: T4 DNA buffer (2 µl, tris-HCl (0.005 M), MgCl₂ (0.01 M), dithiothreitol (0.01 M), ATP (1 mM, pH 7.5), linearized vector (7 µl of 70 ng), DNA insert (10 µl, 100-200 ng) and freshly thawed T-4 ligase enzyme (1 µl, 400 cohesive-end U). The reaction mix was stored (16 ºC, 16 h) before being transformed (2 µl) into LB21 competent *E. coli* cells. The transformed cells were allowed to grow (200 rpm, 37 ºC, 1 h) in selective LB medium containing the antibiotic ampicillin (100 µg/ml). The grown cells (100 µl) were allowed to grow (37 ºC, 16 h) and to produce colonies on solid agar medium (ampicillin 100 µg/ml). Single colonies were used to inoculate selective LB medium (5 ml) containing the same concentration of the antibiotic, and the inoculated medium were allowed to grow with shaking (200 rpm, 37 ºC, 16 h). Colony PCR reaction, to check the correct insertion, was performed using the grown medium (1 µl) as a template and the primer pairs: M.e-CCoAOMT2 (full) and M.e-EOMT (full) for M.e-CCoAOMT2 and M.e-EOMT genes.
respectively. Primers, sequences, and their working conditions are listed in Table 3.1.

**Bacterial growth curve**

Measurement of the transgenic bacterial growth rate and determination of the different life phases in terms of OD was performed by plotting a bacterial growth curve of the cells viability, by measuring its OD$_{600}$ versus time (in min). A suspension of transgenic *E. coli* cells carrying the gene of interest, sub-cloned into PGEX vector, was prepared by inoculating selective LB medium (30 ml, 100 µg/ml ampicillin) with a single colony from fresh agar plate. The inoculated medium was incubated with vertical shaking (200 rpm, 16 h) and then the cells were diluted to a final OD$_{600}$= 0.2. The diluted cell-culture was allowed to grow in a 96-well microplate (100 µl/well) and the absorbance was measured and recorded every 30 min for 24 h using a SKAnit plate reader.

**Enzymatic activity assay**

The transgenic *E. coli* cells were allowed to grow in a selective LB medium (3 ml, 50 µg/ml ampicillin) under normal growth conditions (37 °C, 200 rpm, 16 h). The seed culture (1 ml) was used to inoculate fresh LB medium (20 ml, 50 µg/ml ampicillin) and the culture was allowed to grow until OD$_{600}$=0.5-0.6 (37 °C, 200 rpm, 3 h). Then isopropyl β-D-1-thiogalacto-pyranoside IPTG (8 µl of 100 mM/ml) was added to induce the expression activity. The induced transformant was allowed to grow for 4 h at 25 ºC before the cells were harvested by centrifuging (5,000 rpm, 10 min, 4 ºC). The harvested cells were re-suspended in fresh LB medium (50 ml, 50 µg/ml ampicillin). Different substrates were added separately in a final concentration of 150 µM, to examine the ability of M.e-CCoAOMT2 and M.e-EOMT to O-methylate them, and the culture was then incubated with continuous shaking (25 ºC, 200 rpm, 16 h). Then the cells were harvested and the supernatant was collected and extracted with ethyl acetate (2 x 1:1 v/v). The combined organic layers were evaporated under reduced pressure using a Buchi Rotavapor R-114 (30-40 ºC). Finally, the residue was dissolved in methanol (1 ml) for analysis by LC/MS quantification analysis. Substrates used to feed the transgenic culture were: caffeic acid, *p*-coumaric acid, esculetin, and umbelliferone. Non-transgenic *E. coli* and PGEX plasmid cultures were used as negative controls.
UPLC/ESI-MS

Ethyl acetate extracts of the transgenic bacterial cultures were separated by liquid chromatography (LC) using a reverse phase RP (C18) column under the same conditions stated in chapter 2. Gradient elution used acidified methanol: water for the separation of scopoletin, isoscopoletin, scoparone, esculetin, umbelliferone, and 4-methylumbelliferone. Gradient elution used acetonitrile: water for the separation of p-cinnamic, 4’-methoxycinnamic, caffeic, and ferulic acids. Detection was performed by High Resolution Time-of-Flight spectroscopy (HR-TOF) using Electrospray Ionisation (ESI) in positive ion-mode for the hydroxycoumarins and in negative ion-mode for the acids.

The theoretical mass/charge (m/z) ratios are: scopoletin C_{10}H_{9}O_{4} 193.0495, isoscopoletin C_{10}H_{9}O_{4} 193.0495, scoparone C_{11}H_{11}O_{4} 207.0652, esculetin C_{9}H_{7}O_{4} 179.0339, p-hydroxycinnamic acid C_{9}H_{6}O_{3} 163.4000, 4’-methoxycinnamic acid C_{10}H_{11}O_{3} 177.0557, umbelliferone C_{9}H_{9}O_{3} 163.0390, 7-methoxyumbelliferone C_{10}H_{8}O_{3} 177.0546, caffeic acid C_{9}H_{7}O_{4} 179.0350, and ferulic acid C_{10}H_{9}O_{4} 193.0506.

5.4. Results and discussion

Amplification of Cassava OMTs genes

The full-length cDNA for both M.e-CCoAOMT2 and M.e-EOMT were amplified by HF-Q5 polymerase enzyme using the primer pairs RE-CCoAOMT2 and RE-EOMT respectively (Table 5.1). Sites for the restriction enzymes had been added to flank the cDNA. The PCR products were electrophoresed (70V, 60 min) and the corresponding gel-bands with the correct size were gel-purified. Non-specific PCR products of multiple bands were produced, this is due to the strong tendency of the extra guanine nucleotides added to the primers to form hydrogen-bonds within the cDNA-vector templates. Primer dimers were also expected to be produced as by-products due to the relatively long (27-33 bp) sequence and high GC ratios (57.6-70.4%) of the used primer-pairs. Dimers were produced after the hybridization of the forward and reverse primer molecules which were mostly elongated by the polymerase activity (Brownie et al., 1997).
The correct sized DNA-bands were purified and ligated separately with the linearized PGEX vector before being transformed into BL21 *E. coli* cells. A representative gel image for the PCR products is shown in Figure 5.4.

**Fig. 5.4. Amplification of the cassava OMT-cDNA.** Gel image showing the PCR products after gradient PCR reactions (TM 52.6-59.0 °C), to amplify the cDNA of interest plus the added restriction-nucleotides from their corresponding pCAMBIA-templates. Band-size was measured using a 100 bp DNA ladder as molecular marker. Each lane shows the PCR product at a definite annealing temperature in °C. The correct M.e-CCoAOMT bands are underlined in blue, the correct M.e-EOMT bands are underlined in yellow, and the primer dimers bands are surrounded in red, non-specific by-products are indicated with green arrows.

**Cassava-OMT genes cloned into PGEX vector**

All the *E. coli* cultures that survived the ampicillin selectivity in the agar medium were successfully transformed with the genes of interest. This was confirmed by colony PCR reactions using the primer pairs CCoAOMT2 (full) and M.e-EOMT (full) listed in Table 3.1. The different phases of the bacterial growth were determined for each transgenic culture depending on the generated growth
curve. The selected genes were fused into a glutathione (GSH) system to enhance the solubility and stability (Yang et al., 2015).

No differences in the growth rate between PGEX-CCoAOMT and PGEX-EOMT were noticed. An OD<sub>600</sub>=0.4-0.5 was chosen as the time point to induce the inserted gene because the bacterial cells will then be the exponential phase and undergoing active multiplication (Figure 5.5). Cells were fed with different substrates early in the exponential phase, and allowed to grow until reaching the stationary phase. Elongation of the bacterial life-cycle was optimized by decreasing the incubation temperature from 37 ºC to 25 ºC.

![Graph of Transgenic E. coli growth curve](image)

**Fig. 5.5. Transgenic E. coli growth curve.** The graph shows the transgenic E. coli growth curves for both PGEX-M.e-CCoAOMT (red line) and PGEX-M.e-EOMT (blue line) cultures. Lag phase is where the bacteria are adapting to the media (not shown), log (exponential) phase is where the bacteria are undergoing active duplication, stationary phase is where the growth and death rates are equal, the death phase starts after ~24 h. Both cultures showed similar life-cycle activities.

Reactions products were detected and determined

Bacterial metabolites were analysed by LC/MS to identify and quantify the expected O-methylated products in each feeding reaction. Potential substrates were chosen from the intermediates along the different pathways (Figure 5.6). Due to the difficulty in synthesizing coumaroyl CoA, caffeoyl CoA, and feruloyl CoA esters, the methylation ability of CCoAOMT and EOMT enzymes towards their
corresponding carboxylic acids have been studied. Furthermore, *E. coli* lacks the 4-coumarate: CoA ligase (4CL) enzyme activity which is required to produce the CoA esters even though the host cells are able to provide the CoA molecules. SAM which is required for OMT enzyme activity was provided by the host (Lin *et al.*, 2013).

**Fig. 5.6. Potential substrates for the OMT enzymes on the phenylpropanoid pathways.** The corresponding acids of *E*-p-coumarate and *E*-caffeate, umbelliferone, and esculetin were fed separately to CCoAOMT and EOMT enzymes. Substrates are shown in red, enzymes in blue.
Controls

Not unexpectedly, all the three negative controls showed flat lines for the O-methylated products. None of the hydroxycoumarins or their O-methylated products was among the bacterial metabolites. The wild type BL21 cells and BL21 cells carrying an empty PGEX plasmid also showed no methylation activity towards any of the tested substrates, esculetin, caffeic acid, p-coumaric acid, and umbelliferone (Figure 5.7).

Fig. 5.7. The negative controls chromatograms. LC chromatograms of the two negative control extracts: wild type BL21 cells (orange line) and an empty PGEX vector (brown line) after feeding with esculetin. The peak of the authentic standards scopoletin is shown (purple line). Scopoletin was not detected in the negative control extracts.

Esculetin (6,7-dihydroxycoumarin)

The growing bacterial cultures, containing M.e-CCoAOMT2 and M.e-OMT separately, were fed with esculetin, allowed to grow until OD_{600}=0.8, then the LB medium was extracted with ethyl acetate, and analysed by LC-MS. Both engineered E. coli extracts showed a regioselective enzymatic activity to catalyse the addition of a methyl group onto the 6-hydroxy group of the exogenous esculetin added to the cultures. None of the 7-O-methylated (isoscopoletin) or 6,7-di-O-methylated (scoparone) was detected (Figure 5.8).
Fig. 5.8. Esculetin O-methylation. Three possible O-methylated products of esculetin: scopoletin, isoscopoletin, and scoparone. Scopoletin was the only product detected after the O-methylation reactions catalysed by M.e-CCoAOMT and M.e-EOMT separately.

Different concentrations of the authentic standards scopoletin, isoscopoletin, and scoparone were prepared, and used as reference measurements of the concentration of the biologically synthesized coumarins in the transgenic E. coli cells after feeding esculetin. The lower limit of detection (LLD) of the hydroxycoumarins: scopoletin, isoscopoletin, and scoparone, where the signal: noise ratio is 3:1, was 7.5 ng/ml. The lower limit of quantification (LLQ) of the three coumarins, where the signal: noise is 10:1 was 25 ng/ml in the chromatography used (Figure 5.9).

Fig. 5.9. The calibration curves of scopoletin and scoparone. Different concentrations of the authentic standards scopoletin (A) and scoparone (B) were used to construct the calibration curves.

Neither isoscopoletin nor scoparone were detected in the extracted cultures of both genetically modified cell types. Scopoletin LC peak was detected and quantified in both bacterial culture extracts (Figure 5.10). For PGEX-CCoAOMT,
scopoletin concentration (n=3) was 14.94 ± 4.42 µg/l. For PGEX-EOMT, the measured scopoletin concentration (n=3) was 8.18 ± 0.98 µg/l (Figure 5.10).

Therefore, according to the results obtained, both cassava enzymes M.e-CCoAOMT2 and M.e-EOMT were confirmed to have a regioselective activity towards esculetin to produce 6-O-methylesculetin (scopoletin), confirming that both enzymes are involved in scopoletin biosynthesis.

Fig. 5.10. Enzymatic activity of PGEX-EOMT and PGEX-CCoAOMT recombinant proteins towards esculetin. Scopoletin is biosynthesised in both PGEX-EOMT (A) and PGEX-CCoAOMT (C) cultures. B shows an expanded chromatogram for PGEX-EOMT reaction extract with a flat line for scoparone.

Several publications have reported esculetin O-methylation catalysed by plant and human OMTs. In 1981, Collendavelloo et al., reported the activity of the three O-di-phenyl-O-methyltransferases purified from infected tobacco leaves towards esculetin and other O-di-phenolic substrates including 5'-hydroxyferulic and caffeic acids (Collendavelloo et al., 1981). In 2006, Kim et al., reported an affinity of poplar OMT isolated from Populus deltoids towards esculetin to produce scopoletin, isoscopoletin, and scoparone (Kim et al., 2006). In 2014, attempts to find an inhibitor to prevent levodopa degradation, the E. coli recombinant protein of the isolated peripheral human catechol OMT from total brain RNA showed a specificity to produce scopoletin from esculetin (Kurkela et al., 2004). In 2016, from other recombinant DNA studies in E. coli, the activity of a CCoAOMT-like isoform (PaOMT2) purified from the liver wort plant Plagiochasma appendiculatum, towards esculetin to produce scopoletin and isoscopoletin, and towards caffeoyl CoA to produce feruloyl CoA was reported (Xu et al., 2016).
Caffeic acid (3´,4´-dihydroxycinnamic acid)

The ethyl acetate-extracted transgenic bacterial cultures of PGEX-CCoAOMT and PGEX-EOMT, after feeding caffeic acid, were analysed by LC-MS. Enzymatic assay based on each of the recombinant genes separately, confirms the activity of cassava CCoAOMT towards caffeic acid to produce the corresponding 3´-O-methyl caffeic acid (ferulic acid) (Figure 5.11). The methyl group donor SAM was afforded by the host cells. The biosynthesis of the 3´-O-methylated caffeic acid in a reaction mediated by M.e-CCoAOMT2 meets the expectations as the enzyme activity was predicted due to its amino acid similarity to the reference A.t-CCoAOMT in A. thaliana which is known to exhibit such activity with caffeic acid (Kai et al., 2008).

![Regiospecific O-methylation of caffeic acid](image)

**Fig. 5.11. Regiospecific O-methylation of caffeic acid.** Ferulic acid was biosynthesized by 3´-O-methylation of caffeic acid in a reaction catalysed by M.e-CCoAOMT enzyme. SAM is afforded by the *E. coli* host cells.

The concentration of the biosynthesized ferulic acid was measured according to the calibration curve of different concentrations of the authentic standard ferulic acid (Figure 5.12). The LLD for ferulic acid is 4 ng/ml and the LLQ is 13 ng/ml in the chromatography system used. The measured concentration (n=3) of the biosynthesized ferulic acid in PGEX-CCoAOMT system was 336.9 ± 50.5 µg/l. Cassava EOMT enzyme showed no activity towards O-methylation of caffeic acid (Figure 5.13). Other possible methylated products: 3´-hydroxy-4´-methoxy- and 3´,4´-dimethoxy-cinnamic acids were not studied because the main pathway through ferulic acid is well identified in both *A. thaliana* (Kai et al., 2008) and cassava (Bayoumi et al., 2008b).
Fig. 5.12. The calibration curves of caffeic and ferulic acids. Different concentrations of the authentic standards caffeic (A) and ferulic (B) acids were used to construct the calibration curves.

Fig. 5.13. Enzymatic activity of PGEX-EOMT and PGEX-CCoAOMT recombinant proteins towards caffeic acid. A shows the LC peaks of the authentic standards. B shows a flat line of ferulic acid in the PGEX-EOMT reaction overlaid on authentic ferulic acid. C shows the added caffeic acid and the biosynthesized ferulic acid in the PGEX-CCoAOMT recombination reaction mediated by M.e-CoAOMT2.

*p-Coumaric acid (4’-hydroxycinnamic acid)*

The ethyl acetate-extracted transgenic bacterial cultures of PGEX-CCoAOMT and PGEX-EOMT after feeding 4’-hydroxycinnamic acid to both cultures separately, were analysed by LC-MS. The expected O-methylated product,
if any of the cassava OMT enzymes has activity towards the substrate, is 4′-methoxycinnamic acid as shown in Figure 5.14.

![Diagram of 4'-hydroxycinnamic acid and 4'-methoxycinnamic acid](image)

**Fig. 5.14. Possible O-methylation product of 4′-hydroxycinnamic acid.**
4′-Methoxycinnamic acid was not detected after separately feeding both PGEX-CCoAOMT and PGEX-EOMT cultures with 4′-hydroxycinnamic acid.

Different concentrations of the authentic standards *p*-coumaric acid (4′-hydroxycinnamic acid) and 4′-methoxycinnamic acid were used to draw calibration curves for quantification purposes. The LLD for 4′-hydroxycinnamic acid and for 4′-methoxycinnamic acid is 7.8 ng/ml and the LLQ is 26 ng/ml in the chromatography system used (Figure 5.15).

![Calibration curves of 4′-hydroxycinnamic and 4′-methoxycinnamic acids](image)

**Fig. 5.15. The calibration curves of 4′-hydroxycinnamic and 4′-methoxycinnamic acids.** Different concentrations of the authentic standards 4′-hydroxycinnamic (A) and 4′-methoxycinnamic (B) acids were used to construct the calibration curves.

Based on the enzymatic assay results, none of the cassava OMT has an *O*-methylation activity towards 4′-hydroxycinnamic acid. A flat line was obtained from the LC chromatograms of the PGEX-CCoAOMT and PGEX-EOMT separately (Figure 5.16).
**Fig. 5.16. Enzymatic activity of PGEX-CCoAOMT recombinant protein towards 4′-hydroxycinnamic acid.** A shows the LC peaks of the authentic standards. B shows a flat line of 4′-methoxycinnamic acid in the PGEX-CCoAOMT reaction overlaid on authentic standard.

**Umbelliferone (7-hydroxycoumarin)**

The ethyl acetate-extracted transgenic bacterial cultures of PGEX-CCoAOMT and PGEX-EOMT after feeding umbelliferone to both cultures separately, were analysed by LC-MS. The expected O-methylated product, if any of the cassava OMT enzymes has activity towards the substrate, is 7-methoxycoumarin as shown in Figure 5.17.

**Fig. 5.17. Possible O-methylation product of 7-hydroxycoumarin (umbelliferone).** 7-Methoxycoumarin was not detected after separately feeding both PGEX-CCoAOMT and PGEX-EOMT cultures with 7-hydroxycoumarin.

Different concentrations of the authentic standards 7-hydroxycoumarin and 7-methoxycoumarin were used to construct calibration curves for quantification purposes. The LLD of umbelliferone and 7-methoxycoumarin is 4 ng/ml and the LLQ is 13 ng/ml in the chromatography system used (Figure 5.18).
Fig. 5.18. The calibration curves of 7-hydroxycoumarin (umbelliferone) and 7-methoxycoumarin. Different concentrations of the authentic standards 4’-hydroxycinnamic (A) and 4’-methoxycinnamic (B) acids were used to construct the calibration curves.

Based on the enzymatic assay results, none of the cassava OMT has O-methylation activity towards 7-hydroxycoumarin. Therefore, flat lines were obtained in the LC chromatograms from PGEX-CCoAOMT and PGEX-EOMT separately when they were fed with 7-hydroxycoumarin (Figure 5.19).

Fig. 5.19. Enzymatic activity of PGEX-EOMT recombinant protein towards 7-hydroxycoumarin. A shows the LC peaks of the authentic standards. B shows the added 7-hydroxycoumarin (umbelliferone) and a flat line of 7-methoxycoumarin in the PGEX-EOMT reaction, overlaid on the authentic standards.

The overall conclusions of these microbiological experiments are that it has been confirmed that the amino acid sequences similarity is not sufficient to predict the enzymatic function. Regarding M.e-CCoAOMT, the experimental results confirmed its O-methylation activities towards caffeic acid and esculetin as predicted.
from its sequence similarities to A.t-CCoAOMT (Kai et al., 2006). For the M.e-EOMT, the results showed regiospecificity of this cassava enzyme to biosynthesize scopoletin after feeding the transgenic E. coli with esculetin, despite the sequences similarities with the poplar OMT which is able to catalyse the biosynthesis of isoscopoletin and scoparone along with scopoletin (Kim et al., 2006).

The inactivity of any of the M.e-CCoAOMT and M.e-EOMT cassava enzymes towards both p-coumaric acid and 7-hydroxycoumarin (umbelliferone) showed that none of these OMTs catalyses the biosynthesis of isoscopoletin and scoparone from umbelliferone (Figure 5.20).

**Fig. 5.20. Hypothetical pathway for the biosynthesis of isoscopoletin and scoparone.** Red dashed line indicates the O-methylation step which is not mediated by either M.e-CCoAOMT or M.e-EOMT cassava enzymes.
Chapter 6. General conclusions

Cassava is an important starch-rich crop in the tropics due to its ability to grow and survive in the adverse environmental conditions of poor soil and drought. One of the major constraints to large-scale commercial cassava production is the development of blue-black discoloration soon after harvesting the roots (24-72 h). This unfavourable discoloration is attributed to an internal oxidative-burst leading to the accumulation and oxidation of secondary metabolites known as post-harvest physiological deterioration (PPD). It is established that hydroxycoumarins, in particular scopoletin and its glycoside scopolin, are among the secondary metabolites which undergo dramatic accumulations in the roots after harvest. To a lesser extent, the closely related esculetin and esculin hydroxycoumarins are also involved.

In general, plant responses to stress involve production of defensive antioxidant compounds to counteract the stark increase in ROS (Cortez et al., 2002). PPD is a genetically active process involving changes in gene expression and protein synthesis (Reilly et al., 2007). Despite the recent discovery and annotation of the cassava genome (Wang et al., 2014), the entire pathway leading to scopoletin biosynthesis in cassava is not yet fully elucidated. The observed physiological changes after harvesting the roots are a result of wound responses, and therefore PPD is an abiotic stress response.

The four hydroxycoumarins known to be involved in the visible signs of PPD, scopoletin, esculetin and their corresponding β-glycosides scopolin, and esculin were detected, isolated, and quantified from fresh and from deteriorated cassava roots. Scopoletin showed an increasing pattern in its accumulation in the deteriorated roots compared to its barely detectable levels in fresh roots. Scopolin, was also not detected directly after harvest, but an increasing concentration was measured over 7 days after harvest. Esculetin and esculin were only detected in very low concentrations later on (days 4-7) in PPD. The undetectable levels of these hydroxycoumarins in fresh cassava roots (t=0) defines them as phytoalexins as they are biosynthesized in cassava roots after harvest. It has been established that scopoletin, in particular, is biosynthesized de novo from hydroxycinnamate (p-coumarate) in the phenylpropanoid pathway rather than released from its glycoside. The role of hydroxycoumarins as defence compounds has been confirmed.
Although the model plant *A. thaliana* exhibited different patterns in scopoletin biosynthesis, where both scopoletin and scopolin were found in detectable levels in the fresh roots, not leaves, nevertheless it shares the same phenylpropanoid pathways with cassava. Thus, it was used for modelling the abiotic stress responses and for exploring the key enzymes on the pathways. Unexpectedly, the measured concentrations of both scopoletin and scopolin in *A. thaliana* roots after inducing the plant with drought-mimic effectors, were either comparable to the non-stressed plants as in the salinity stressor, or were much reduced as after exposure to heat shock. Different plant responses could vary. In some cases, the plants adapt their physiology to the unfavoured environment while, for others, extreme conditions proved fatal.

To get more insight into the scopoletin biosynthetic pathways, and to understand the intertwined convergent pathways that cassava could employ to maintain appropriate levels of scopoletin, its biosynthetic pathways were modelled in *A. thaliana*. Each proposed pathway was interrupted at the molecular level by knocking-out a specific key-enzyme in the latter part of the alternative scopoletin biosynthetic pathways, thereby only altering scopoletin accumulation, whilst ensuring that the synthesis of other vital secondary metabolites was not affected. In *A. thaliana*, knocking-out CCoAOMT, which is a key enzyme on and therefore representative of the 4′-hydroxy-3′-methoxycinnamate (ferulate) pathway, significantly reduced scopoletin, and therefore scopolin, accumulation in the roots (~50%). Knocking-out EOMT, which is representative of both 2′,4′-dihydroxycinnamate and 3′,4′-dihydroxycinnamate (caffeate) pathways, showed only a slight effect on scopoletin profile (~20%). Accordingly, the biosynthetic pathway through ferulate is defined as the major pathway, at least in the model plant *A. thaliana*. Transgenic *A. thaliana* roots with double mutations in both CCoAOMT and EOMT did not show any further reduction in scopoletin levels. While CCoAOMT has been previously reported to be involved in scopoletin biosynthesis (Kai *et al.*, 2006), this is the first time the activity of the A.t-EOMT gene towards esculetin has been reported following such *in planta* experiments.

These findings in *A. thaliana* were used to explore experimentally the functional identity of cassava genes involved in scopoletin biosynthesis. Cassava candidate genes were identified according to their (translated) amino acid sequence similarities with the reference *A. thaliana* genes. These putative cassava genes were
isolated, cloned, and their expression in both cassava leaves and roots was studied. Those genes expressed in cassava tissues were sub-cloned using Gateway® technology, and then used to complement the deficiency in scopoletin biosynthesis in the T-DNA inserted A. thaliana mutant plants. Increasing levels of both scopoletin and scopolin were measured after complementing the mutant plants with cassava candidate genes. Thus, both M.e-CCoAOMT and M.e-EOMT genes were confirmed to be involved in scopoletin biosynthesis in cassava. Key genes involved in the main hydroxycoumarin biosynthetic pathways were identified and studied.

Despite the progress with establishing the cassava genome, M.e-EOMT is the first gene confirmed to be involved in the alternative scopoletin biosynthetic pathways in cassava through its ability to mediate esculetin O-methylation into scopoletin. The functional identities of M.e-EOMT, in addition to M.e-CCoAOMT, were further confirmed after in vivo expression of both cassava genes separately in E. coli. M.e-CCoAOMT was found to have O-methylating activity towards producing ferulic acid from caffeic acid. Both M.e-EOMT and M.e-CCoAOMT were found to have a regioselective activity towards adding an O-methyl group to esculetin to produce scopoletin. This non-specificity could explain the slight decrease in scopoletin levels after inhibiting the activity of EOMT alone in A. thaliana. These results confirmed that scopoletin is mainly biosynthesised by 6′-hydroxylation of ferulate, E-Z-isomerization, and then lactonization steps, in both cassava (Bayoumi et al., 2008a) and A. thaliana (Kai et al., 2008) wild type plants.

Hydroxycinnamate is a key entry molecule in scopoletin biosynthesis, channelling carbon flux early from L-phenylalanine to scopoletin through different proposed pathways. In order to investigate the origin of the residual scopoletin biosynthesised in the mutant lines, competition feeding experiments were designed and performed with different stable isotopically labelled intermediates. Feeding with cinnamic-d7 acid served as a marker of the plant’s ability to uptake and utilize the exogenous material. All the tested lines afforded scopoletin-d3 and scopolin-d3 after feeding with cinnamic-d7 acid. Three deuteriums were lost in the 3 hydroxylation reactions mediated by the oxidative enzymes C3′H, C4′H, and F6′H1 respectively. The fourth deuterium from cinnamic-d7 acid was lost during the E-Z-isomerization step by an as yet unknown mechanism. Feeding cubes of wild type cassava roots with ferulic-OCD3 acid afforded scopoletin-d3 and scopolin-d3 where the three deuteriums were conserved on the O-methyl group. Even though it was not expected
to detect any labelled scopoletin after feeding the transgenic M.e-F6'H cassava roots with ferulic-d₃ acid, a high incorporation of the labelled intermediate was measured. This high percentage resulted from the low-level of the already biosynthesized scopoletin and either the failure of the RNAi technology used to knock-out the F6'H activity completely or the possibility of other hydroxylase enzymes adding a hydroxyl group into the ortho-position, carbon C6'. This puzzle was solved experimentally by modelling the process in the A.t-F6'H1 mutant plant. The measured levels of scopoletin and scopolin were minimal (~3% of the levels measured in the wild type); RT-PCR confirmed that the complete knock-out of the gene was achieved, but the detection of scopoletin-d₃ (from the 3%) in A.t-F6'H1 mutant after feeding with ferulic-OCD₃ acid strongly supports the hypothesis of alternative hydroxylase enzyme activity. Moreover, in transgenic cassava, a significant but not complete reduction in scopoletin level was achieved (~50%). Thus, both the incomplete knock-out and the presence of other hydroxylase options are possible. Feeding labelled ferulic acid (either ferulic-OCD₃ or ¹⁸O-ferulic) to the wild type and A.t-EOMT plants afforded labelled scopoletin and scopolin. Competition feeding with ¹⁸O-esculetin in cassava afforded a low percentage of ¹⁸O-scopoletin incorporation (~1%), indeed a level close to the natural abundance of the 2x ¹³C-scopoletin (~1%), but a higher labelled ¹⁸O-scopolin (~20%) level was detected in both the wild type and the transgenic cassava lines. Therefore, the conclusion from these results is that a fourth, though not favoured, biosynthetic pathway may exist via the glycosylation of esculetin into isoesculin (cichorin, chicorin), then 6-O-methylation to produce scopolin which could de-glycosylate into scopoletin. This could involve di-O-glycosylation of esculetin on O-C6 and O-C7, regiospecific O-C6 deglycosylation, followed by O-methylation on O-C6 to yield scopolin as another possible pathway. Feeding labelled esculetin to different A. thaliana lines produced labelled scopolin in different incorporation ratios. The low incorporation of the labelled ¹⁸O-esculetin in A.t-EOMT knock-out plants, which is almost near the natural abundance ratio of 2x ¹³C-scopoletin, confirms the functional identity of the EOMT gene.

The overall findings of this research project contribute towards a better understanding of the interconnectivity of the different scopoletin biosynthetic pathways. These results may ultimately contribute towards the production of transgenic cassava with less or delayed signs of PPD, a contribution to food security.
In conclusion, with a focus on food security, cassava is grown for its starch-rich roots, a staple food for millions of people across Africa. Cassava is the sixth most important crop for starch production after wheat, rice, maize, potato, and barley. Moreover, that it can be harvested year round, grows in poor soil, and its drought-tolerance make it a promising reservoir in famine. Fortifying cassava with essential nutrients, such as vitamin A, is ongoing research in biotechnology.

Despite its importance, growing cassava has several limitations, some of which are biological, e.g. ACMV and bacterial blight pathogens. These have been resolved by (artificial) selection, by farmers, who chose the best pathogen-resistant cultivar, and also by using the conventional cross-breeding techniques.

High-cyanide content is a major nutritive drawback of cassava, besides its low content of proteins, vitamins, and minerals. Consumption of raw or inefficiently processed cassava is associated with the neurodegenerative disease Konzo. Vegetative propagation of the less bitter (lower cyanide content) and proper processing of the edible parts contribute to decreasing the cyanide toxicity.

A major, but still unresolved limitation in cassava production is PPD causing a significant loss ~20% of the produced cassava roots. Conventional breeding fails to resolve the problem of PPD due to the complex heterozygous nature of cassava, e.g. the desirable dry weight content is positively correlated to the PPD trait. A few trials in producing transgenic cassava with extended shelf-life have been recently reported, but unfortunately none of them succeeded in produced normal phenotype healthy plants in field trials with no or delayed PPD (Zidenga et al., 2012; Xu et al., 2013). As PPD is an endogenous oxidative process, resulting in a burst of reactive oxygen species (ROS), the accumulation of many phenolic and polyphenolic antioxidants maintains homeostasis, along with upregulation of many stress-responsive enzymes such as catalase and dismutase.

These studies have served to confirm that the hydroxycoumarin scopoletin is positively correlated to the development of PPD symptoms due to the increase in its concentration in the deteriorated roots compared to the essentially non-detectable level in the fresh roots. A better understanding of scopoletin biosynthesis in cassava has been achieved. Creating and modelling transgenic cassava in A. thaliana was performed in order to enable further research in producing transgenic cassava which targets the later steps in scopoletin biosynthesis without affecting the normal phenotype or other vital biosynthetic processes in this important starch-rich plant.
References (in the style of Phytochemistry, 2017)


Frick, S., Kutchan, T.M., 1999. Molecular cloning and functional expression of O-
methyltransferases common to isoquinoline alkaloid and phenylpropanoid pathways. 
Plant J. 17, 329-339.
Frick, S., Ounaroon, A., Kutchan, T.M., 2001. Combinatorial biochemistry in plants: 
Gauthier, A., Gulick, P.J., Ibrahim, R.K., 1996. cDNA cloning and characterization of a 
3'/5'-O-methyltransferase for partially methylated flavonols from Chrysosplenium 
americanum. Plant Mol. Biol. 32, 1163-1169.
which encode O-methyltransferases for the methylation of both flavonoid and 
phenylpropanoid compounds. Arch. Biochem. Biophys. 351, 243-249.
in scopoletin concentration in cassava chips from four varieties during storage. J. Sci. 
Food Agric. 91, 2344-2347.
Gomez-Vasquez, 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. Ann. 
Bot. 94, 87-97.
Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., 
Goujon, T., Sibout, R., Pollet, B., Maba, B., Nussaume, L., Bechtold, N., Lu, F., Ralph, 
mutant deficient in the expression of O-methyltransferase impacts lignins and 
sinapoyl esters. Plant Mol. Biol. 51, 973-989.
Goulas, E., Schubert, M., Kieselbach, T., Kleczkowski, L.A., Gardestrom, P., Schoeder, 
W., Hurry, V., 2006. The chloroplast lumen and stromal proteomes of Arabidopsis 
thaliana show differential sensitivity to short- and long-term exposure to low 
Guerrier, G., 1996. Fluxes of Na\(^+\), K\(^+\) and Cl\(^-\) and osmotic adjustment in *Lycopersicon pimpinellifolium* and *L. esculentum* during short- and long-term exposures to NaCl. Physiol. Plant 97, 583-591.


Li, X., Chen, W., Zhao, Y., Xiang, Y., Jiang, H., Zhu, S., Cheng, B., 2013. Downregulation of caffeoyl-CoA O-methyltransferase (CCoAOMT) by RNA


coenzyme A involved in formation of umbelliferone and scopoletin in sweet potato,

eexpression of alfalfa isoliquiritigenin 2′-O-methyltransferase, an enzyme specifically
involved in the biosynthesis of an inducer of *Rhizobium meliloti* nodulation genes.
Plant J. 4, 971-981.

Miller, R.W., Sirois, J.C., Morita, H., 1975. The reaction of coumarins with horseradish

Mkumbira, J., Chiwona-Karltn, L., Lagercrantz, U., Mahungu, N.M., Saka, J., Mhone,
cassava into ‘bitter’ and ‘cool’ in Malawi: from farmers’ perception to
caracterisation by molecular markers. Euphytica 132, 7-22.

Konzo in southern Tanzania: Rehabilitation and prevention using the wetting

Mohr, P.G., Cahill, D.M., 2003. Abscisic acid influences the susceptibility of
*Arabidopsis thaliana* to Pseudomonas syringae pv. tomato and *peronospora parasitica.*


Morante, N., Sanchez, T., Ceballos, H., 2010. Tolerance to post harvest physiological
deterioration in cassava roots. Crop Sci. 50, 1333-1338.

Moyib, K., Mkumbira, J., Odunola, O.A., Dixon, A.G., Akoroda, M.O., Kulakow, P.,
2015. Genetic variation of postharvest deterioration susceptibility in a cassava
germplasm. Crop Sci. 55, 2701-2711.

25, 239-250.

expression of an *Arabidopsis* cDNA clone encoding a flavonol 3′-O-
375, 385-388.

Nakatsubo, T., Kitamura, Y., Sakakibara, N., Mizutani, M., Hattori, T., Sakurai, N.,


Patil, P.O., Bari, S.B., Firke, S.D., Deshmukh, P.K., Donda, S.T., 2013. A comprehensive review on synthesis and designing aspects of coumarin derivatives as


www.sigmaaldrich.com/content/dam/sigma-aldrich/docs/Sigma/General_Information/applications_dnmr.pdf (accessed on 09.08.17).


methyltransferase from cultured cells of *Coptis japonica*. Plant Cell Physiol. 36, 29-36.
proteomics of the cassava storage root and identification of a target gene to reduce postharvest deterioration. Plant Cell 26, 1913-1924.


Appendices

Appendix 1.

Nucleotide sequences of cassava 4.1_010203 gene coding for CCoAOMT. Exons are highlighted in blue and the UTR regions in pink. N indicates non-sequenced nucleotides.
Appendix 2.

Presentations arising from these studies.

The Academy of Pharmaceutical Sciences Conference, Nottingham, UK, 7\textsuperscript{th}-9\textsuperscript{th} September 2015. Phytochemistry of hydroxycoumarins in cassava (\textit{Manihot esculenta}) during post-harvest physiological deterioration (Poster presentation).

Trends in natural products research: A young scientists meeting of the Phytochemical Society of Europe PSE and IUNG-PIB Conference, Pulawy, Poland, 30\textsuperscript{th} May-2\textsuperscript{nd} June 2016. Using transgenic plants to investigate the pathways of hydroxycoumarins biosynthesis (Oral and Poster presentations). Winner of the TEVA Award for the best Oral presentation.

The Academy of Pharmaceutical Sciences Conference, Glasgow, UK, 5\textsuperscript{th}-7\textsuperscript{th} September 2016. Hydroxycoumarin biosynthesis along different phenylpropanoid pathways (Oral and Poster presentations).

The Academy of Pharmaceutical Sciences Conference, London, UK, 5\textsuperscript{th}-7\textsuperscript{th} September 2017. Tracing scopoletin biosynthesis in wild type and transgenic \textit{Arabidopsis thaliana} (Poster presentation).

The Teva Award

This is to certify that

Lidia K. Alhalaseh

has been awarded the Prize from Teva Czech Industries

for Best Oral Presentation

www.phytochemicalsociety.org
Phytochemistry of hydroxycoumarins in cassava (Manihot esculenta) during post-harvest physiological deterioration

Lidia K. AlHalaseh, Ian S. Blagbrough, John R. Beeching
Department of Pharmacy and Pharmacology, and *Department of Biology and Biochemistry.

Abstract - Cassava is the most important crop in the tropics but liable to a rapid post-harvest deterioration due to an internal oxidative burst which leads to the accumulation of hydroxycoumarins biosynthesized by different proposed pathways. These metabolites were detected and isolated from deteriorated roots using chromatographic methods and the genetic material for the enzymes responsible for their biosynthesis were identified and isolated, their expressions have been checked and cloned and ready to be inserted in a homozygous T-DNA mutants A. Thaliana to evaluate the different biosynthetic pathways.

INTRODUCTION

Manihot esculenta is the most important root crop in the tropics. It has high carbohydrate content and adaptability to diverse and challenging environments. However, the post-harvest physiological deterioration (PPD) of the root is a major constraint to commercial production. PPD is mainly due to the accumulation of the hydroxycoumarins scopoletin, esculetin and their glucosides within 24-72 h of harvesting the roots. Despite the importance of these secondary metabolites but major details in their biosynthesis remain unresolved.

MATERIALS AND METHODS

Cassava roots from different wild cultivars were collected from local shops in Bath/UK, and cassava culture TMS 60444 was grown in the University of Bath green house. T-DNA lines were obtained from Salk * Company. Pruners were obtained from Fisher institute UK. Taq polymerase enzyme and Q5 high fidelity polymerase enzyme were purchased from New England Biolabs. All Reagents and solvents were purchased from Sigma-Aldrich.

HPLC data were collected from HPLC instrumentation was a Jasco PU-980 pump with the post-column output monitored at λ = 360 nm by a Jasco UV-975 detector. The columns were pre-packed reverse-phase HPLC columns purchased from Phenomenex Inc... Phenomenex Gemini (C18 5 μm packing) semi-preparative (10 mm x 250 mm) and analytical (4.6 mm x 150 mm) columns. Detection was by UV absorbance at λ = 360 nm. Spectroscopic measurements were recorded on a Bruker Daltonics "microTOF" electrospray ionisation mass spectrometer (ESI-MS).


CONCLUSIONS

PPD is mainly due to the massive accumulation and oxidation of the secondary metabolite hydroxycoumarins scopoletin, esculetin and their glucosides observed within 24-72 h of harvesting the roots. Different pathways contribute in scopoletin biosynthesis including methylation of esculetin and methylation of caffeoyl CoA due to the positive expression of both cassava 4.1_014783 and cassava 4.1_010187 (encoding for CCaOMT and esculetin OMT respectively), but in unknown proportions.

ACKNOWLEDGMENTS

We thank Mr. A. Alkal for his support during the study.

REFERENCES

Using transgenic plants to investigate the pathways of hydroxycoumarin biosynthesis

Lidia K. Alhalaseh a, Ian S. Blagbrough a, John R. Beeching b, Michael G. Rowan a

a Department of Pharmacy and Pharmacology, University of Bath, Bath BA2 7AY, U.K.
b Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, U.K.

The hydroxycoumarins scopoletin, esculetin and their glycosides accumulate up to two hundred-fold in harvested cassava roots. Their oxidation with 2-5 days leads to the blue-blackish vascular streaking, known as post-harvest physiological deterioration (PPD), which renders this starch food-stuff unpalatable and unmarketable. Cassava is the world’s sixth most important crop and has a major role to play in terms of food security in often drought-prone sub-Saharan African countries [1]. We aim to gain insights into the control the coumarin biosynthesis through alternative biosynthetic pathways within general phenylpropanoid metabolism by producing single and double mutant plants (e.g. in Arabidopsis thaliana as a model plant) in which genes for key enzymes in these pathways are knocked out.

Salk lines with T-DNA insertion of the two O-methyl transferase enzymes (caffeoyl CoA O-methyl transferase CCoAOMT encoding by the at4g34050 gene and esculetin O-methyl transferase enzyme E-OMT [2] encoding by the at5g54160 arabidopsis gene) were allowed to grow under controlled conditions. The two homozygous plants were crossed so as to produce a double mutant plant. The accumulation of these hydroxycoumarins was determined in the single, double mutants and wild-type plants. Additionally, homologous cassava genes in (cassava4.1_014783 and cassava4.1_010187 coding for CCoAOMT and E-OMT correspondingly) were isolated, amplified, and overexpressed in mutant Arabidopsis plants in order to confirm their functional identity through their ability to complement the mutations.

Interfering in the main phenylpropanoid pathway by knocking out CCoAOMT decreased the level of scopoletin and its glycoside scopolin by 50%. Inhibiting the second pathway (via caffeic acid) and the third pathway (via umbelliferone) had little effect on scopoletin and scopolin concentrations. The concentrations of other hydroxycoumarins were altered by these mutations (Figure 1). These data imply that the major pathway of scopoletin biosynthesis in A. thaliana goes via ferulic acid.

We thank Mu'tah University for financial support (PhD studentship to LKA).

Hydroxycoumarin biosynthesis along different phenylpropanoid pathways

Lidia K. Allhasala1, Michael G. Rowan1, Ian S. Blagbrough1, and John R. Beeching2
1Department of Pharmacy and Pharmacology, University of Bath, Bath BA2 7AY, U.K.
2Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, U.K.

Abstract – Single mutations of caffeoyl CoA O-methyltransferase (CCoAOMT) and esculetin O-methyltransferase enzyme (E-OMT) in A. thaliana showed (CCoAOMT) 50% decreased levels of scopoletin and scopoletin, whereas E-OMT mutation had little effect.

INTRODUCTION

Cassava is the world’s 6th most important crop and plays a major role in terms of food security in often drought-prone sub-Saharan Africa [1], but it suffers from internal oxidation processes 2-3 days after harvesting with the accumulation of secondary metabolites and blue-blackish vascular streaking known as post-harvest physiological deterioration (PDD).

The hydroxycoumarins, esculetin, isoscopel, and their glycosides, play a major role in PPD as they accumulate up to 200-fold in harvested cassava roots. We aim to gain insights into the control of hydroxycoumarin biosynthesis through alternative biosynthetic pathways within general phenylpropanoid metabolism by producing single and double mutant plants (e.g. in Arabidopsis thaliana as a model plant) in which genes for key enzymes in these hydroxycoumarin biosynthesis pathways are knocked out.

MATERIALS AND METHODS

Cassava culture TMS 60444 was grown in the University of Bath (UoB) glass house under controlled conditions. Arabidopsis seeds with T-DNA insertion were obtained from SALK8 institutions and grown in the UoB growth room under controlled conditions. Molecular biology data for CCoAOMT and E-OMT genes and finding the homologous genes in cassava are currently under analysis via bioinformatics software GENIUS®. The National Centre for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov). HPLC data were collected using a Jasco PU-980 pump with post-column output monitored at λ = 360 nm by a Jasco UV-975 detector. The columns were pre-packed RP-HPLC columns (Phenomenex Inc.: Phenomenex Gemini C18 5 μ semi-preparative (10 mm x 250 mm) and analytical (4.6 mm x 150 mm) columns. Spectrometric measurements were recorded on a Bruker Daltonics “micrOTOF” electrospray ionization mass spectrometer (ESI-MS).

RESULTS AND DISCUSSION

Scopoletin can be biosynthesized by three different phenylpropanoid pathways in unknown proportions, starting from the deamination of L-phenylalanine to cinnamic acid by the action of phenylalanine ammonia lyase enzyme via three different major intermediates: 7-hydroxycoumarin (umbelliferone), 3,4-dihydroxycoumarin (caffeate), or 4-hydroxy-3-methoxycoumarin (ferulate) in arabidopsis [2]. We knocked out the esculetin O-methyltransferase enzyme (E-OMT) to inhibit the methylation of esculetin to scopoletin in both the first and second pathways, and the caffeoyl CoA O-methyltransferase (CCoAOMT) to inhibit the synthesis of ferulic acid, the direct precursor for scopoletin, in the third pathway to assess the difference in hydroxycoumarin concentration compared to the wild type. Remarkable decreases in both scopoletin and scopoletin concentrations (both ~50%) in the CCoAOMT knocked-out mutant plant compared with the E-OMT mutant plant (Figure 1) leads to the conclusion that the major pathway of scopoletin biosynthesis in A. thaliana is via ferulic acid.

Figure 1. Coumarin levels in arabidopsis root extracts

COnCLUSIONS

Interfering in the main phenylpropanoid pathway (via ferulic acid), by knocking out CCoAOMT in A. thaliana decreased the level of scopoletin and its glycoside scopoletin by 50%. Knocking out E-OMT, an enzyme common to pathways both via caffeic acid and via umbeliferone, had little effect on scopoletin and scopoletin concentrations. The concentrations of other hydroxycoumarins (Figure 1) were altered by both of these single mutations showing their significance to the phenylpropanoid pathways.

ACKNOWLEDGEMENTS

We thank Mu’tah University for financial support (PhD studentship to LKA).

REFERENCES

Tracing scopoletin biosynthesis in wild type and transgenic Arabidopsis thaliana

Lidia K. Alhalaseh1, Michael G. Rowan1, Ian S. Blagbrough1, and John R. Beeching2
1Department of Pharmacy and Pharmacology, University of Bath, Bath BA2 7AY, UK
2Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, UK

A R T I C L E   I N F O

*Corresponding author.
Tel.: +44 1225 386795
Fax: +44 1225 384114
E-mail: prsisb@bath.ac.uk

KEYWORDS: isotopes, phenylpropanoid, scopoletin

S U M M A R Y

There are potentially three phenylpropanoid pathways leading to scopoletin biosynthesis. To reveal the contribution of each pathway, wild type and T-DNA insertion mutant Arabidopsis plants were fed with different isotopically labelled intermediates before the methanolic extracts of the roots were subjected to LC/MS analysis. An Arabidopsis F6’H1 mutant in which the ferulate pathway was inhibited, an Arabidopsis E-OMT mutant in which both the caffeate and 2’,4’-dihydroxycinnamate pathways were inhibited, and wild type plants as control, were fed separately with deuterated cinnamic and ferulic acids and 18O-esculetin. All the three labelled intermediates were successfully incorporated in a pattern that shows that the three pathways are operating in Arabidopsis, and that the alternative pathways could compensate, in part, for deficiencies in the other pathways in the biosynthesis of scopoletin.

INTRODUCTION

Hydroxycoumarins are considered as phytoalexins because their biosynthesis is stimulated under plant stress situations (Chong, 2002). Scopoletin and its β-glucoside scopolin, which are known to have a key role in plant defence mechanisms are biosynthesized and accumulated in A. thaliana roots by one dominant (Kai, 2008) and two possibly hypothetical pathways within phenylpropanoid metabolism. These pathways starting from phenylalanine are different in the main intermediate: 2’,4’-dihydroxycinnamate, 3’,4’-dihydroxycinnamte (caffeate) and 4’-hydroxy-3’-methoxycinnamate (ferulate) (Blagbrough, 2010). The aim of this research is to understand the contribution of each pathway in scopoletin and scopolin synthesis in wild type and in mutants in which genes for key enzymes in different pathways were inactivated, through tracing the biosynthesized products of plants growing in media rich in isotopically labelled intermediates.

MATERIALS AND METHODS
Wild type *A. thaliana* ecotype Columbia-0 (Col-0) and T-DNA insertion lines were purchased from The Nottingham Arabidopsis Stock Centre (NASC). Wild type and homozygote mutant plants were fed with different isotopically labelled compounds: *trans*-cinnamic-\(^d_7\) acid (CDN isotopes), ferulic-\(^d_3\) acid (Sigma), \(^{18}\)O-esculetin (synthesized). The separation and quantitative analysis were carried out using UPLC/HRMS (Bruker).

**RESULTS AND DISCUSSION**

The incorporation of exogenous isotopically labelled intermediates into the wild type and mutant plants indicates the presence of active enzymes in the corresponding pathway required to produce labelled scopoletin and scopolin.

The already high accumulation of coumarins in the wild type explains the result of the low ratio (heavy:light) of scopoletin-\(^d_3\) and \(^{18}\)O-scopoletin produced and reveals the contribution of the three pathways in its biosynthesis. The production of scopoletin-\(^d_3\) after ferulic acid feeding in F6’H1 mutants indicates the activation of other hydroxylase enzymes while the production of \(^{18}\)O-scopoletin after esculetin feeding (Fig. 1) clarifies the role of other pathways.

![HRMS of scopoletin product after feeding F’6H mutant plant with \(^{18}\)O-esculetin.](image)

**CONCLUSIONS**

The dramatic decrease in scopoletin accumulation found after knocking out the Feruloyl CoA 6’-Hydroxylase (F6’H1) defines the pathway from cinnamic acid via ferulate leading to scopoletin as the major one (Kai, 2008). However, the high incorporation of \(^{18}\)O-esculetin in scopoletin biosynthesis in wild type and mutants unambiguously demonstrate that all the three proposed pathways are active in Arabidopsis.

**ACKNOWLEDGEMENTS**

We thank Mu’taq University, Jordan, for the financial support of this research (scholarship to LKA).

**REFERENCES**


Functional expression of cassava O-methyltransferase enzymes in *Escherichia coli*

Lidia K. Alhalaseh¹, Michael G. Rowan¹, Ian S. Blagbrough⁴¹, and John R. Beeching²

¹Department of Pharmacy and Pharmacology, University of Bath, Bath BA2 7AY, UK
²Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, UK

**ARTICLE INFO**

*Corresponding author.
Tel.: +44 1225 386795
Fax: +44 1225 386114
E-mail: prsisb@bath.ac.uk

KEYWORDS: cassava, biotransformation, O-methyl-transferase

**SUMMARY**

The affinity of O-methyl transferase enzymes for different potential substrates on phenylpropanoid pathways was explored through the expression of two selected cassava genes, cassava4.1_014783 coding for CCoAOMT and cassava4.1_010187 coding for Me-OMT, in transgenic *Escherichia coli*. LC/MS analysis of reaction products revealed the regiospecific methylation of esculetin, its conversion into scopoletin, in both genetically engineered cultures. Only CCoAOMT had the ability to methylate caffeic acid into the corresponding O-methylated product ferulic acid. Neither of the two selected enzymes showed activity towards umbelliferone or 4'-hydroxy-cinnamic acid.

**INTRODUCTION**

About 40 sequences for plant O-methyl transferase (OMT) genes, found in published sequence databases, encode functional proteins (Chiron, 2000). Some of these enzymes are involved in the synthesis of different secondary metabolites, in particular hydroxycoumarins that have a crucial role in plant defence and growth regulation (Ribera, 2012). Hydroxycoumarins are biosynthesised through different pathways within general phenylpropanoid metabolism (Matsumoto, 2012). Many of the intermediates substituted with at least one hydroxyl group render them potentially good substrates for enzymatic activity. *In vivo* assays for activity of OMTs isolated from *Manihot esculenta* roots were performed to check the affinity and specificity of cassava enzymes towards different potential substrates in the phenylpropanoid pathways.

**MATERIALS AND METHODS**

Total RNA was isolated from cassava roots and cDNA was synthesised using High-Capacity cDNA Reverse Transcription kit (Applied Biosystems). O-Methyl transferase enzymes (cassava4.1_014783 coding for CCoAOMT and cassava4.1_010187 coding for Me-OMT) Full length cDNAs for O-methyl transferase enzymes (cassava4.1_014783 coding for CCoAOMT and cassava4.1_010187 coding for Me-OMT) were amplified by PCR and cloned into a GST-tagged PGEX-5X-1 plasmid vector (Healthcare) and expressed in *E. coli* LB21 (NEB). The transformant was allowed to grow in selective LB media (ampicillin 50 µg/ml) and a range of
potential substrates (150 mM) were added. Ethyl acetate extracts of the LB media were analysed by UPLC/HRMS (Bruker) in order to detect and quantify the methylation products.

RESULTS AND DISCUSSION

LC/MS profiles of esculetin assay reactions from both transgenic cultures revealed the selective production of scopoletin, but not isoscopoletin or scoparone (Fig. 1.). The caffeic acid assay reaction in E. coli expressing the CCoAOMT gene, but not the Me-OMT, showed the production of the O-methylated corresponding compound, ferulic acid. No reaction products were detected after umbelliferone and p-hydroxycinnamic acid feeding. Four control reactions were performed: bacterial culture with no construct, PGEX-5X-1 vector culture with no insert, CCoAOMT recombinant culture with no feeding, and Me-OMT culture with no feeding. All the controls gave negative results.

Fig. 1. LC/MS profile of the esculetin reaction shows the production of scopoletin (6-O-methylesculetin) in the PGEX-CCoAOMT culture.

CONCLUSIONS

E. coli expressing cassava CCoAOMT and Me-OMT can regiospecifically transfer a methyl group from S-adenosyl-L-methionine (AdoMet) onto esculetin for 6-O-methylesculetin (scopoletin), not 7-O-methylesculetin (isoscopoletin) nor 6,7-di-O-methylesculetin (scoparone). Only CCoAOMT has the enzymatic affinity to methylate 3',4',5'-dihydroxycinnamic acid (caffeic acid) converting it into 4'-hydroxy-3'-methoxycinnamic acid (ferulic acid) and none of the cassava O-methyl transferase enzymes has catalytic activity towards umbelliferone and 4'-hydroxycinnamic acid.

ACKNOWLEDGEMENTS

We thank Mu’tah University, Jordan, for the financial support of this research (scholarship to LKA).

REFERENCES

