Doctoral Thesis

Molecular and biochemical understanding of post-harvest physiological deterioration in cassava roots and approaches for its modulation

Author(s):
Adhiambo Owiti, Judith

Publication Date:
2009

Permanent Link:
https://doi.org/10.3929/ethz-a-006023037

Rights / License:
In Copyright - Non-Commercial Use Permitted
Molecular and Biochemical Understanding of Post-harvest Physiological Deterioration in Cassava Roots and Approaches for its Modulation

A dissertation submitted to the ETH Zurich for the degree of Doctor of Science

Presented by Judith Adhiambo Owiti MSc in Biochemistry Jomo Kenyatta University of Agriculture and Technology (JKUAT) Nairobi, Kenya born April 8, 1976 Citizen of Kenya

Accepted on the recommendation of:

Prof. Dr. Wilhelm Gruissem, examiner
Dr. John Beeching, co-examiner
Dr. Peng Zhang, co-examiner

2009
Table of Contents

Table of Content .................................................................................................................. i
Acknowledgments .................................................................................................................. iv
Abstract ................................................................................................................................. vii
Résumé ................................................................................................................................... ix
List of Tables .......................................................................................................................... xi
Abbreviations ......................................................................................................................... xiv

Chapter 1 General Introduction ............................................................................................. 1
1.1 The Cassava Plant: origin, distribution and special characteristics ................................. 1
1.2 Ecology, propagation, production and economics ............................................................... 2
1.3 Importance of Cassava ...................................................................................................... 3
1.4 Constraints to Cassava Research and Production ............................................................... 4
  1.4.1 Biotic constraints .......................................................................................................... 5
  1.4.2 Abiotic constraints ......................................................................................................... 5
1.5 Cassava and PPD problem ................................................................................................. 6
  1.5.1 Economic Impact of PPD to cassava production ............................................................ 7
  1.5.2 Molecular and biochemical understanding of PPD ..................................................... 9
1.6 Conventional breeding and transgenic approaches to delay PPD ...................................... 19
1.7 Rationale .......................................................................................................................... 20
  1.7.1 Aim of the thesis .......................................................................................................... 21

Chapter 2 General Materials and Methods .......................................................................... 23
2.1 Plant samples and PPD set up ......................................................................................... 23
  2.1.1 Plant Material ............................................................................................................... 23
  2.1.2 PPD experiment set up ................................................................................................. 23
2.2 Molecular Biology and Cassava Transformation Methods ............................................... 24
  2.2.1 Cassava total RNA and protein extraction ................................................................... 24
  2.2.2 cDNA synthesis ........................................................................................................... 25
  2.2.3 PCR ............................................................................................................................... 26
  2.2.4 Agarose Gel Eleetrophoresis ....................................................................................... 27
  2.2.5 Purification of amplification fragments ....................................................................... 28
  2.2.6 TOPO cloning ............................................................................................................. 28
  2.2.7 Construction of transformation vectors ........................................................................ 28
2.2.8 Transformation into E. coli ................................................................. 29
2.2.9 Plasmid preparation ................................................................. 30
2.2.10 DNA and RNA yield determination ........................................ 30
2.2.11 Sequence analysis of the DNA ................................................... 30
2.2.12 Agrobacterium tumefaciens and Cassava transformation ........... 30

2.3 Proteomics Methods ........................................................................ 31
2.3.1 Sample clean up ........................................................................ 31
2.3.2 Protein quantification ................................................................. 31
2.3.3 Determination of protein concentration by Bradford Assay .......... 32
2.3.4 Protein–cyanine dye labelling ...................................................... 32
2.3.5 Rehydration of strips and Running the IEF ................................... 33
2.3.6 Casting and running of 2DE ........................................................ 35
2.3.7 Fixation, staining and destaining of the gels .................................. 36
2.3.8 Gel imaging scanning and analysis ............................................. 36
2.3.9 Spot Picking and In-gel protein digestion ...................................... 37
2.3.10 Peptide desalting and MALDI plate spotting ............................ 37
2.3.11 MALDI analysis of protein spots and database searches .......... 38
2.3.12 Reduction, denaturation and blocking of cysteine residues ........ 38
2.3.13 Protein Digestion ..................................................................... 38
2.3.14 Labeling the protein digests with iTRAQ Reagents ................... 38
2.3.15 Strong cation exchange separation of peptides (HPLC) ............ 39
2.3.16 Nano-LC separation and MALDI target spotting of tryptic peptides 40
2.3.17 MALDI-TOF/TOF mass spectrometry ....................................... 40
2.3.18 Peptide and protein identification by database searching .......... 41
2.3.19 Statistical analysis of the iTRAQ data ....................................... 41

2.4 Transcript analysis ........................................................................ 42
2.4.1 Bioinformatics ........................................................................... 42
2.4.2 Primers for real time PCR ........................................................ 42
2.4.3 Real time PCR (qPCR) ............................................................... 43

2.5 Biochemical assays ........................................................................ 44
2.5.1 ROS enzyme assays ................................................................. 44
2.5.2 Cell wall enzyme assays ........................................................... 44

Chapter 3 Results .............................................................................. 46
3.1 Proteomic Results ........................................................................ 46
3.1.1 Analysis of protein changes during cassava PPD using 2D gels ................................................. 46
3.1.2 Proteins regulated 12hs after PPD, analysis by normal 2D gels ...................................................... 48
3.1.3 Analysis of protein changes during cassava PPD using DIGE gels ............................................. 51
3.1.4 Proteins found to be regulated 24h and 48h after PPD in the DIGE experiment ............................... 52
3.1.5 Classification of regulated proteins into pathways by Mapman ..................................................... 56
3.1.6 Analysis of protein changes during PPD by iTRAQ ................................................................. 58
3.1.7 Determination of regulation cut-off for the iTRAQ identified proteins ......................................... 61
3.1.8 Regulated proteins from iTRAQ experiment .................................................................................. 63
3.2 Transcript analysis .............................................................................................................................. 73
3.2.1 Comparison of regulated cassava gene products with their Arabidopsis orthologs ..................... 73
3.2.2 Quantitative PCR (qPCT) analysis of CSD and APX genes ......................................................... 76
3.2.3 Quantitative PCR (qPCR) analysis of genes involved in cell wall metabolism ............................ 79
3.3 Enzyme assays .................................................................................................................................. 82
3.3.1 Changes in activity of APX during PPD ......................................................................................... 82
3.3.2 Changes in activity of SOD during PPD ......................................................................................... 83
3.3.3 Change in activity of Beta 1, 3-glucanase during PPD ................................................................. 84
3.3.4 Change in activity of pectin methyl esterase ............................................................................. 85
3.4 Cassava transformation ...................................................................................................................... 88

Chapter 4 General Discussion ............................................................................................................. 94

Conclusion .......................................................................................................................................... 110

References .......................................................................................................................................... 112
Acknowledgments

I extend my heartfelt gratitude to Prof. Wilhelm Gruissem for the wonderful opportunity to be part of his plant biotechnology research group, and for enabling me to gain access to the state-of-the-art research facilities both at ETH Zurich and at the Functional Genomics Center Zurich. I am also thankful for the advice and mentorship he offered during my research study. It was a great and worthy opportunity to meet and work under the close supervision of Dr. John Beeching. My special thanks to you. Although based in the UK, and thus had to contend with frequent trips to Switzerland for meetings, presentations and discussions, all these proved fruitful and contributed immensely to helping me develop a clear focus for the project. Thanks too to Dr. Peng Zhang for the guidance especially during the initial inception of the project and reviewing the thesis manuscript.

I have to specifically single out the contribution of my immediate supervisor, and current cassava group leader Dr. Herve Vanderschuren, especially in providing new direction and novel research ideas for the project. Your enthusiasm and interest in the PPD project stimulated motivation that enabled me to explore areas that I had only imagined possible.

The inclusion of Dr. Christophe Laloi on board towards the end of the research study was quite a boost to the project. Your enormous expertise in Transcript analysis enabled us to view the project in a different context.

I thank all the proteomics experts at the FGCZ especially Peter Gehrig, Simoen Barrkov, Mike Scott, and Bernd Roschitzki for assistance with the experiments and operating the proteomics instruments. I am grateful especially to Jonas Grossmann also from the FGCZ for help with analysis of all the proteomics data, and for the great company during those early lunches at FGCZ. Indeed your willingness to share knowledge and openness is a virtue to be admired.

I extend my great appreciation for the help, advice, encouragement and support from Simon Bull. The ‘wonderful’ time spent doing tissue culture at the bench as well as the discussions revolving around PPD, cassava tissue culture and other stuff will not quickly fade.

Many thanks to all the members of the Gruissem group both past and present for the extremely conducive and warm working atmosphere. The academic as well as social atmosphere was instrumental in helping me reach this far. My special thanks Dr. Christophe Sautter, Dr. Lars H, Dr. Johannes Füterer for the scientific advice and discussions. To all the former and present
Post-docs as well as PhD students, notably: Carmen, Martin, Sonja, Evelyne, Lorenzo, Yec’han, Andrea, Yvonne, Huan, Ruben, Masha, Alessandro, Carolina, Christina, Diana, Noemi, Daniel, Gilles, Isabel, Sylvain, Ruben, Christina, Lily, Sasha, Evans, Charles, Kwan-Te and Huahong for the assistance, support and discussions. The list is definitely longer than this, but I do appreciate the support from all the members of the lab. All you guys contributed to my being wiser and comfortable in Zurich in one way or another.

I also extend a special thanks to Mike Page from University of Bath for assistance with the enzyme assays, Cecile for the cell wall enzyme assays, Christophe Dessimoz for help with the bioinformatics, Kim and Doris for ensuring the availability of consumables, to Michael for the helping hand with cassava transformation, and Maria for the transcript analysis. I do also acknowledge Sabine Klarer and Katalyn Konya for the care and maintenance of cassava plants in the greenhouse in Eschikon, Matthias and Sven for their timeless assistance with computer related issues, Katharina Holzinger for the overwhelming assistance with administrative issues right from my first day in the lab through to the end of my studies at ETH.

I am also grateful to the Rockefeller foundation, for providing me with the Scholarship to pursue my PhD studies in Switzerland, I particularly single out the contribution of Prof. Deborah Delmer in facilitating the scholarship award.

To the Kenyan community in Zurich, I owe my social sanity to your constant invitations for ‘nyama chomas’, dinners and the lively discussions we had from time to time. Although far from Kenya, you brought Kenya right back to my doorstep, and the true spirit of ‘harambee’ was rekindled in Zurich. For this I am especially grateful to James and Margaret Wafula, from the time I arrived in Zurich, right to the end you became my second family in Switzerland. Many thanks to James Wainaina and family, Maina and his family and to Barasa. To my friend Connie, and Wairimu, many thanks for the many opportunities we had to share experiences and encourage one another.

I would like to thank the almighty God for the strength, guidance and good health throughout my studies. I am very grateful for the support, constant encouragement and love from my family especially my parents, brother Ron, sister Caro, my parents in law, sisters-in-law Keryne
and Roselyne and their families, Gordon and Becky as well as my entire in law family for the encouragement throughout the years of my study in Switzerland.

And last but definitely not least; this work would not have been complete without one person who has been the pillar to my emotional stability. To the one person I am greatly indebted to and who has had to sacrifice the most throughout my study period in Switzerland, the most important person in my life, I want to thank my dearest and beloved husband Edwin, with whose perpetual support I have come this far. Indeed, it has been a long-winded journey, but Edwin, your love, patience, tolerance, understanding and commitment has see me walk this far.
Abstract

An important constraint that limits the full realization of cassava’s potential in developing countries is the short storage life of harvested roots. Cassava roots undergo rapid deterioration 24 - 48 hours after harvest, a phenomenon referred to as post harvest physiological deterioration (PPD) which renders the roots unpalatable and reduces the marketability of the roots. PPD is a polygenic trait attributed to a cascade of signaling events which are triggered by oxidative damage at the wound site. To gain a broader understanding of the molecular and biochemical events associated with PPD, we have undertaken a proteomics profiling experiment to analyze changes in cassava root proteins during PPD.

In this experiment, the isobaric tags for relative and absolute quantification (iTRAQ) analysis coupled with mass spectrometry was adopted to investigate the changes in protein expression patterns in cassava roots in response to PPD. Proteins samples were extracted from cassava roots during early PPD (0, 6, 12 and 24h) and compared. Additionally protein samples from the late PPD time points (0, 48 72 and 96h) were also compared. Proteins were extracted from both cytosolic and membrane fractions of the early and late PPD time points and the entire experiment repeated. The analysis generated 8 datasets, and a total of 8, 960 proteins were identified from cassava. Due to a high degree of redundancy in the protein identifications, the data was filtered by generating Arabidopsis thaliana orthologs of the proteins identified. A total of 1110 unique proteins were obtained, and 711 proteins had no known orthologs in Arabidopsis. A total of 60 proteins were found to be significantly regulated in both replicates in the same direction and in corresponding sample identities. Of these 33 proteins were shown to be up-regulated and 27 proteins were down-regulated relative to the 0h control.

Real time-PCR (qPCR) analysis and enzyme assays were used to confirm if transcriptional activity and enzyme activity of a selection of the proteins identified by iTRAQ correlated with the changes depicted at the protein level. Transcript analysis revealed the up-regulation of pectin methyl esterase during PPD, which was in agreement with the change in abundance of this protein observed during PPD, and also the increase in the enzyme activity. The up-regulation of Cu-Zn superoxide dismutase (CSD1.2), and beta 1,3-glucanase protein during early PPD (6-12h) revealed by iTRAQ was also confirmed via qPCR, however the total enzyme activity of superoxide dismutase and beta 1,3-glucanase increased throughout during PPD.
The proteomics experiment led to the identification of proteins that are regulated in cassava during PPD. These included mainly enzymes involved in ROS detoxification, cell wall metabolism, wound and stress response, and secondary metabolite biosynthesis. In the study documented in this thesis, the combination of transcriptomics, proteomics, and enzymatic assays has led to the generation of comprehensive data on transcriptional and post-translational modulation of key enzymes involved in PPD.
Résumé

Une des contraintes les plus importantes pour la production de manioc dans les pays en développement demeure la courte durée de conservation après la récolte. Les racines de manioc subissent une rapide décomposition physiologique (en 24 à 48h) après la récolte, un phénomène identifié sous le nom de dégradation physiologique post-récolte. Ce phénomène réduit l’appétence ainsi que la valeur marchande des racines de manioc. La dégradation physiologique post-récolte est un caractère multigénique attribué à une cascade d’évènements qui sont initiés par un stress oxydatif au niveau du site de lésion. Pour mieux comprendre les évènements moléculaires et biochimiques associés avec la dégradation physiologique post-récolte, nous avons développé une étude protéomique afin d’analyser les changements du protéome au cours de la dégradation physiologique post-récolte dans les racines de manioc.

Dans l’étude présente, la technologie iTRAQ (isobaric tags for relative and absolute quantification) couplée à la spectrométrie de masse a été utilisée pour caractériser les changements des profils d’expression des protéines dans les racines de manioc au cours de la dégradation physiologique post-récolte. Les échantillons protéiques ont été prélevés sur des racines de manioc au début de la dégradation post-récolte (0, 6, 12 et 24h) et à des stades avancées de la dégradation post-récolte (0, 48, 72 et 96h). L’ensemble des échantillons ont été comparés sur base de leurs profils d’expression protéique. Les fractions cytosolique et non soluble ont également été extraites séparément et analysées. L’ensemble de l’expérience a été répétée. L’analyse a généré 8 bases de données et un total de 8960 protéines identifiées. A cause d’un niveau de redondance élevé dans les protéines identifiées, les bases de données ont été filtrées en recherchant les orthologues d’*Arabidopsis thaliana* pour l’ensemble des protéines identifiées. Au total, 1110 protéines uniques (orthologues d’*Arabidopsis*) ont été obtenues après l’application du filtre ainsi que 711 protéines n’ayant pas d’orthologues dans le génome d’*Arabidopsis*. Soixante protéines ont montré le même type de régulation (positive ou négative, temps post-récolte) dans les deux réplicas biologiques. Parmi ces 60 protéines, 33 ont montré une augmentation significative de leur niveau d’expression tandis que 27 ont montré une baisse significative de leur niveau d’expression. Des analyses par PCR semi-quantitative en temps réel et par réactions enzymatiques ont également été utilisées afin de confirmer les
changements au niveau des transcripts et des activités enzymatiques pour certaines protéines sélectionnées afin de caractériser une corrélation potentielle avec les résultats obtenus par iTRAQ. L’analyse des transcripts a révélé l’accumulation de transcripts codant pour la méthylestérase de pectine au cours de la dégradation physiologique post-récolte. Cette observation est corrélée avec l’accumulation de cette enzyme au cours de la dégradation physiologique post-récolte. Les augmentations de la Cu-Zn superoxide dismutase et de la glucanase β-1,3 identifiées par l’iTRAQ au début de la dégradation physiologique post-récolte (6-12h) ont également été confirmées par PCR semi-quantitative en temps réel. Les activités enzymatiques mesurées pour ces enzymes ont révélé une augmentation tout au long de la dégradation physiologique post-récolte.

Les expériences protéomiques ont conduit à l’identification de protéines qui soit s’accumulent ou sont dégradées dans les racines de manioc au cours de la dégradation physiologique post-récolte. Il s’agit essentiellement d’enzymes impliquées dans la détoxification des espèces réactives d’oxygène, dans le métabolisme de la paroi, dans la réponse aux lésions et au stress, ainsi que la biosynthèse des métabolites secondaires. Dans cette thèse, la combinaison d’approches protéomique, transcriptomique et enzymatique a généré un ensemble de données qui permettent de caractériser les modulations transcriptionnelles et post-transcriptionnelles d’enzymes-clés impliquées dans la dégradation physiologique post-récolte des racines de manioc.
List of Tables

Table 1: Primer combination sequences used to amplify target DNA. ............................................. 27
Table 2: Cassava DIGE labeling design ............................................................................................. 33
Table 3: The IEF parameters used for focusing for the 24 cm strip pH 3-10 ......................................... 34
Table 4: Primers used for real time PCR ............................................................................................. 43
Table 5: Representative proteins unique to 0h PPD ......................................................................... 49
Table 6: Representative proteins down regulated 12hs after PPD ..................................................... 50
Table 7: A selection of proteins up-regulated in both 24h and 48h PPD compared to 0h from DIGE experiment .................................................................................................................. 53
Table 8: A selection of proteins down-regulated in both 24h and 48h PPD compared to 0h from DIGE experiment .................................................................................................................. 54
Table 9: Percentage numbers of proteins that show a 2 fold change in the soluble and membrane protein fractions .................................................................................................................. 63
Table 10: Number of soluble proteins found to be regulated .............................................................. 64
Table 11: Number of membrane proteins found to be regulated ....................................................... 64
Table 12: A selection of proteins found to be up-regulated in both iTRAQ experiments ..................... 66
Table 13: A selection of proteins found to be down-regulated in both experiments ......................... 67
List of Figures

Figure 1: Cassava root after harvest ................................................................. 6
Figure 2: Cassava root 48hr PPD ................................................................. 6
Figure 3: Reactive oxygen species scavenging system in eukaryotic organisms .............. 14
Figure 4: Cassava roots slices used for PPD experiment .................................... 23
Figure 5: Destination vector (modified from pCAMBIA 1301)................................ 29
Figure 6: Expression clone generated from the destination vector. ......................... 29
Figure 7: General workflow of the iTRAQ experiment. ..................................... 39
Figure 8: 2D analysis of cassava root proteins during PPD .................................. 47
Figure 9: Magnified views of protein spots shown to be unique to 0h (A) and down regulated in 12h (B) ................................................................. 48
Figure 10: Representative images from DIGE analysis of cassava proteins ............... 52
Figure 11: Magnified images from 2-D DIGE analysis showing spot in 0h (control), down regulated 24h and 48h after PPD ................................................................. 56
Figure 12: Overview of cassava proteins differentially expressed during PPD ............ 57
Figure 13: Overlap between protein from experiment 1 and experiment 2 ............... 60
Figure 14: Overlap between proteins from soluble and membrane fractions ............... 60
Figure 15: Overlap between protein from early and late PPD time points in both experiments 61
Figure 16: GO categories of the up-regulated proteins from the iTRAQ experiment ........ 70
Figure 17: GO categories of the down-regulated proteins from the iTRAQ experiment .... 71
Figure 18: Histograms showing the abundance ratio of the regulated proteins selected for downstream analysis ................................................................. 72
Figure 19: APX gene family in cassava ............................................................ 74
Figure 20: Cu-Zn SOD (CSD) gene family in cassava ........................................ 75
Figure 21: Changes in transcript levels of the APX isoenzymes during PPD ............... 77
Figure 22: Changes in transcript levels of the CSD 1.1 and CSD 1.2 isoenzymes during PPD ................................................................. 78
Figure 23: Changes in transcript levels of the CSD2 and CSD3 isoenzymes during PPD .... 79
Figure 24: Changes in transcript level of beta-1, 3-glucanase during PPD......................... 80
Figure 25: Changes in transcript level of pectin methyl esterase during cassava PPD. ............. 81
Figure 26: Changes in total APX activity during PPD ............................................................ 82
Figure 27: Changes in total SOD activity during PPD .......................................................... 83
Figure 28: Beta 1,3-glucanase activity in cassava roots during PPD........................................ 85
Figure 29: Gel diffusion assay revealing PME activity during PPD. ...................................... 86
Figure 30: Changes in PME activity in cassava during PPD.................................................. 87
Figure 31: PCR fragments amplified from Arabidopsis and cassava cDNA. ............................. 91
Figure 32: Putative cassava transgenic plants ...................................................................... 92
Figure 33: Transgenic cassava shoots culture on CBM media................................................. 93
Abbreviations

2-DE: Two-dimensional gel electrophoresis
AA : Ascorbate/ ascorbic acid
ACN: Acteonitrile
BLAST: Basic local alignment search tool
BSA: Bovine serum albumin
CBM: Cassava basal media
CHAPs: 3[[3-Cholamidopropyl]dimethylammonio]-propanesulfonic acid
DEPC : Diethylenepyrocarbonate
DIGE: Two-dimensional difference gel electrophoresis
dNTP: deoxyribonucleotide Triphosphates
EDTA: Ethylene diamine tetraacetic acid
EST: Expressed sequence tag
FEC: Friable embryogenic calli
GPS: Global Peoteome server
GUS : β-glucuronidase
HEPES: N-2-Hydroxyethylpiperazone-n-2-Ethanesulfonic Acid
HRPG: hydroxyproline-rich cell-wall glycoprotein
HSP : Heat shock protein
IAA: Iodo acetic acid
ICAT: Isotope coded affinity tags
IEF: Isoelectric focusing
IPG: Immobilized pH gradients
IPTG: Isopropyl β-D-1-thiogalactopyranoside
iTRAQ™: Isobaric tagging technology for relative and absolute quantitation
KCI: Potassium cholride
LB: Luria broth
LC: Liquid chromatography
LC-MS/MS: Liquid chromatography tandem mass spectrometry
LiCl: Lithium chloride
MALDI: Matrix-assisted laser desorption/ionization
MOWSE: Molecular Weight Search
MS/MS: tandem mass spectrometry
MS: Mass spectrometry
NaCl: Sodium chloride
PAGE: Polyacrylamide gel electrophoresis
PCR: Polymerase chain reaction
PPD: Post harvest physiological deterioration
PR: Pathogenesis-related
PVP: Polyvinyl pyrrolidone
RNA: Ribonucleic acid
SDS: Sodium dodecyl sulphate
TAE: Tris-acetate-EDTA
TBP: Tributyl phosphine
TFA: Trifluoro acetic acid
TOF: Time-of-flight
X-gal: bromo-chloro-indolyl-galactopyranoside
Chapter 1 General Introduction

1.1 The Cassava Plant: origin, distribution and special characteristics

Cassava (*Manihot esculenta* Crantz) is a perennial vegetatively propagated shrub that belongs to the family *Euphorbiaceae* and subfamily *Crotonoideae*. This family also includes milkweed and castor bean however, cassava is the only member of the *Euphorbiaceae* family that is cultivated as a food crop (Fauquet 1990). Cassava originated in South America and was transported to Africa by the early European traders during the 16th and 17th centuries. Cassava cultivation has spread to the tropical and subtropical regions where it is grown from sea level up to altitudes of 1800 m above sea level within the equatorial belt between 30 ° north and 30 ° south of the equator (Cock 1985). Cassava serves as the most important root crop in the tropics and sub-tropics especially sub-Saharan Africa. It is a major carbohydrate source after rice, sugarcane and maize for over 500 million people (Cock 1985). The roots contain 30-40 % dry matter and of which 85 % is starch and only 1-2 % protein. Due to the high starch content, the crop has taken on more importance as a source of starch for industry and food processing, and as animal feed in South East Asia and South America (Ceballos, 2002).

Cassava is a rustic crop able to thrive in marginal agro-climatic conditions where few other crops could survive (Cock 1982). This ability to produce under sub-optimal conditions is related to various physiological traits. For instance cassava lacks the critical growth periods such as anthesis observed in grain crops. The lack of critical growth stages in cassava is linked to the characteristic of simultaneous growth habit in which growth and development of the economically useful plant parts (roots) and the structures required to supply them with energy occur simultaneously (Cock 1984). This contrasts sharply with the cereal grains, which have a sequential growth habit; initially they build leaf area, roots and support structures and then later develop and fill reproductive organs. The sequential growth habit in the cereal crops implies that in the presence of biotic or abiotic stresses, major crop failure or even total crop loss could occur.
Cassava can survive on impoverished soils and some varieties have been shown to be drought tolerant. The crop has a flexible planting and harvesting time which explains its critical role for food security especially in regions prone to drought. It provides more dietary calories per hectare than any other staple crop and can be easily incorporated in various cropping systems (Onwueme 1978; Fregene 2000). The growth in small holder plots is advantageous to the farmers since they can grow a lot of cultivars with special agronomic traits such as superior taste, early maturing, pest/disease resistant cultivars, drought tolerant cultivars and better processing characteristics (Salick 1997). In contrast to the capital intensive and input-demanding green revolution cereal crops such as wheat, rice and maize, the special traits of cassava make it popular among small holder farmers within the tropics where there is limited access to agricultural inputs (Fregene 2000).

1.2 Ecology, propagation, production and economics

Cassava is vegetative propagated using mature stem cuttings which are planted horizontally, vertically or inclined on ridged or flat soils, (Keating 1988). It is usually intercropped with other staples such as maize, sorghum or grain legumes, although single cropping production also occurs to a lesser extent. Cassava suffers from many pests and diseases which also affect the quantity and quality of planting material. Pests and diseases may be transmitted from one vegetative cycle to the next; therefore the selection of healthy, disease-free and pest-free material is critical.

Since cassava is a lowland tropical plant, it requires a warm moist climate with average temperature of 24-30 °C for optimum growth and production, and for maximum leaf photosynthesis (El-Sharkawy 1992a). Seeds in cassava are used mainly in breeding programs. Although the use of seeds in commercial cassava production is a promising option to obviate some of the production constraints associated with vegetative propagation, (Iglesias 1994) plants arising from sexual seeds are normally weaker than those obtained from cuttings since they are usually homozygous for recessive and prejudicial genes (Nassar 2007). Plants originating from cuttings have much more vigor due to the heterozygotic nature of cuttings. The ideal soils for cassava are light sandy loam with medium fertility. Although resilient to drought and capable of growth in low nutrient soils,
cassava cannot tolerate high concentrations of salts with a pH above 8, excess soil moisture, or temperatures below 10 °C (Onwueme 1978; Lozano 1980; IITA 2001; Nassar 2005). Storage roots are generally harvested 7-24 months after planting depending on the variety and growing conditions. The mature roots can be left in the ground for a long period making it a useful crop against famine. The world production of cassava has increased to 212 million tonnes (FAO 2007). Yields in cassava vary depending on the variety and soil type (El-Sharkawy 2004), for example in Uganda the yield varies from 10-30 tons/ha while in Kenya yield averages of 10.6 tons/ha have been reported when environmental factors are considered (Fermont 2009). To achieve maximum yields, cassava requires high solar radiation, high mean day temperature, sufficient supply of all required nutrients, good rainfall distribution during crop establishment and possibly a dry period before harvesting (El-Sharkawy 2004).

1.3 Importance of Cassava

World consumption of cassava for food, either fresh or processed, is concentrated in the developing countries (FAO 1999), but its utilization varies greatly in different parts of the world (FAO 1999). In Africa about 70 % of the tuberous roots of cassava are used for human consumption either directly as food after cooking or in processed forms. The remaining 30% is used for animal feed and other industrial applications such as starch, glucose and alcohol production. A typical diet based on cassava starch provides less than 30 % of the minimum daily requirement for protein and only 10-20 % of the required amounts of iron, zinc, vitamin A and vitamin E. In some parts of Africa where diets are typically based on starchy staples, the leaves of cassava are harvested and used as a vegetable or a constituent of a sauce taken along with the main staple meals (Lancaster 1983). The cassava leaves thus serve as a supplement to the dietary requirements for proteins and vitamins as well as other minerals (Hahn 1988; FAO 1993). The leaves are especially useful in the dry season when other green vegetables are in short supply (FAO 1993).

In Latin America and the Caribbean, on the average, 42 % of cassava production is used for human consumption as fresh either boiled or dried, roasted flour known as farinha de mandioca, and as a pre-cooked flour called farinha da mesa (FAO 1999). About a third of
cassava produced in the Americas is used for animal feed, and to a lesser extent for starch-based and cassava-based snacks (Westby 2002). In Asia, over 40 % of cassava produced is for human consumption, with much of the remaining exported as chips and pellets to the European Union for use as animal feed (Westby 2002). India and Indonesia are the two leading cassava-consuming countries in Asia. In India, baked roots are turned into small chips; flour and sago while in Indonesia, where 57 % of cassava production is used as food, the roots are eaten boiled or steamed and processed into dried chips called *gaplek*. In Thailand, a greater percentage of the production is exported as cassava chips; nonetheless, some are processed into starch and consumed as cakes, noodles and pastry (FAO 1999). The industrial applications of cassava is expected to increase markedly especially in Asia where rapid economic growth is stirring up demand for starch and ethanol (FAO 2007). With the persistent increase in crude oil prices, cassava is soon becoming a commercially attractive and viable feedstock for energy production. In some parts of Asia notably China, Philippines, and Thailand, heavy investments in cassava as feedstock for bio-fuel production have been initiated (FAO 2007).

1.4 Constraints to Cassava Research and Production

Cassava has long been considered a primitive crop of resource-poor farmers, and despite the growing demand and its production potential, it has continuously received very little attention from researchers, funding organizations and governments alike thus making it an “orphan crop”. Until three decades ago the knowledge base on cassava was scarce compared to that of most cereal crops (FAO 2000) and it is only recently perhaps with the realization of the crop’s great potential to combat food insecurity and its significant role in the starch industry that scientific attention has shifted towards improving cassava both in terms of production and technology and advancing understanding of the crop physiology. Despite cassava being a hardy crop, its increased production is hampered by a myriad of biotic (diseases, insects, mites and weeds) and abiotic (soil, climate, post harvest deterioration and agronomic factors) constraints. The fact that cassava is a long season crop able to grow in diverse agro-ecological zones contributes largely to its exposure to these constraints (Dixon 1992; Mahungu 1994).
1.4.1 Biotic constraints

In Africa and parts of Asia cassava production is hindered by pathogens and pests which reduce yields and availability of planting material leading to loss of food and income for the farming communities (Dixon 2003). Diseases such as Cassava Brown Streak Disease (CBSD) and Cassava Mosaic Disease (CMD) transmitted by whiteflies (Bemisia tabaci) and spread by cuttings, Bacterial blight caused by (Xanthomonas axonopodis pv manitotis) and anthracnose (Colletotrichum gleosporoides) are the most widespread in cassava. The white flies cause direct damage to the plant by feeding on the phloem of the leaves, inducing chlorosis and leaf fall. They can also induce honeydew on which sooty mould can develop (Bellotti 1999). Several pests feed on cassava including the cassava mealybug (Phenococcus manihoti), cassava green mite (Mononychellus tanajoa) and nemtaodes particularly (Meloidogyne spp).

1.4.2 Abiotic constraints

Environmental factors that impact on cassava production include soil erosion and low soil fertility, infestation of the fields with weeds and poor climatic conditions. Low winter temperature is for instance a climatic constraint in the sub-tropics, mainly affecting southern China and North Vietnam; however cassava’s relatively poor adaptation to cool temperatures is compensated by its tolerance to poor soils in these areas (FAO 2000). Being a low-value crop grown in fragile or otherwise marginal ecosystems with low or unpredictable rainfall, cassava is pushed more and more into the least favourable areas, and as such, the use of improved varieties and cultural practices does not impact positively on yields (FAO 2000). The growing of cassava continuously for many years without adequate fertilization in soils that tend to be of low fertility and susceptible to erosion also contributes to a further decline in soil fertility for example in Thailand (FAO 2000). Cassava is quite susceptible to competition from weeds, especially at the early growth stage. Failure of timely weeding can cause a total loss of harvest. In most parts of Thailand weed control is traditionally done by animal drawn implements or by hand. Labor for weeding accounts for about 40% of total labor used, thus representing a major part of production cost. Additionally, since cassava is vegetatively propagated, it is usually planted within one to two
months after the roots of the previous crop have been harvested. Farmers therefore produce their own planting material and the shortage of clean planting material could further impact on the yield output.

1.5 Cassava and PPD problem

Upon harvesting, the starchy roots of cassava suffer a rapid deterioration that renders them unpalatable within 24-72 hours depending on the cultivar and environmental conditions (Fig 1 and 2). This deterioration process can be classified into two phases. The first phase known as primary deterioration initiates from the central vascular bundles of the root. Its first visible signs are a blue-black discoloration, or vascular streaking, beginning at the broken or cut surfaces. Subsequently, the deterioration spreads to the adjacent storage parenchyma and the stored starch undergoes structural changes (Plumbley 1991). This initial deterioration process is a physiological process that does not involve microorganisms (Averre 1967; Noon 1977). The second phase, also known as secondary deterioration, is due to infection with microorganisms leading to fermentation and softening of the root tissue (Plumbley 1991; Wehnam 1995). It is the former, known as post-harvest physiological deterioration (PPD), that is the constraint to the development of cassava necessitating the prompt consumption or processing of the roots after harvesting as the roots become unpalatable and unmarketable (Beeching 1994).

Figure 1: Cassava root after harvest

Figure 2: Cassava root 48hr PPD

PPD is initiated by injury and wounding during harvesting and the histological and biochemical changes observed in cassava during PPD parallels the wound responses in other
plant systems (Bennett 1994). Plant wounding induces the production of signaling components that initiate the wound response: this includes the production of defensive compounds and enzymes, the preparation of the plant for the potential extension of wounding, and culminates in wound repair which is followed by the inhibition of signals. These aspects of wound response are present in cassava root except that the wound repair and down modulation of the signaling cascade are inadequate or non-existent (Booth 1976; Rickard 1981).

In addition to the visual symptoms of PPD, there are increases in respiration (Hirose 1986), changes in lipid composition (Lalaguna 1989), secondary metabolite accumulation (Rickard 1981; Uritani 1983; Wheatley 1985), and the synthesis of the phytohormone ethylene (Hirose 1984). There is also an increase in the activity of a range of enzymes including phenylalanine ammonia-lyase (PAL), acid invertase, catalase, dehydrogenases, peroxidases and polyphenol oxidase (Booth 1976; Rickard 1981). PPD is further compounded by the environmental conditions and the genetic background of cassava variety (Booth 1976; Wehnam 1995).

1.5.1 Economic Impact of PPD to cassava production

PPD necessitates the prompt consumption or processing of the cassava roots after harvesting. Because of their perishability, most roots are usually consumed or marketed close to the centers of production. In the rural set up, where cassava is grown and consumed or processed immediately after harvest, PPD is not a significant problem. However, with increasing urbanization and agribusiness attention focusing on cassava as a source of raw material, PPD has become a critical constraint to the use and development of cassava in terms of production, consumption and processing as well as extensive commercialization of cassava.

Additionally, since cassava is grown mostly by smallholder farmers who are often located in areas with poor transport infrastructure, time and distance between production fields and processing industries or market further exacerbates the problem. The bulkiness and perishability of the roots results in increased marketing and transport costs, making the product relatively expensive in urban areas, a constraint that could be overcome through
the development of processing close to the production centers. Estimated yield losses that stem from cassava PPD range from 5 - 25% of harvested roots (Janssen 1985) depending on genotype and environmental conditions. Deteriorated roots fetch low market returns as producers are forced to sell the products at discounted prices: consumers on the other hand turn to alternative carbohydrate sources (Janssen 1985; Vlaar 2001). The visible coloration of the root tissue associated with PPD development is used as an indication of the culinary value and taste, making it difficult to sell the roots. Most of the deteriorated roots end up as animal feed with a reduction in price greater than 50%. PPD also leads to reductions of starch quality and quantity in the storage roots (Uritani 1984), thus indirectly leading to reduced cassava root utilization as a raw material for the starch industry. These problems have immense negative implications on the rural and national economies where cassava is produced as an economic crop.

Several methods have been developed to extend cassava shelf life, including processing roots into more stable traditional or industrial products. Since PPD occurs due to an endogenous oxidative process that initiates within the roots, exclusion of oxygen by storing and transporting the roots in plastic bags, or coating individual roots with paraffin wax have been attempted to delay the deterioration for three - four weeks. Traditional methods for preserving fresh roots include packing roots in moist mulch or by removing leaves two weeks prior to harvest to prolong root shelf-life to two weeks (FAO 2000). For specialized markets the roots can be peeled, chopped into chunks and frozen. However, all these methods have met with little success due to the investment required as well as convenience and availability of materials (FAO 2000; Van Oirschot 2000).

Concerted efforts aimed at addressing the PPD problem are a prerequisite to ensuring continued reliance on cassava for starch industry, and to minimize the risk of yield losses due to post-harvest deterioration. According to a FAO report (FAO 1995), increasing the storage life of cassava to a minimum of two weeks could enhance cassava utilization and reduce by 90% the PPD related constraints, thereby significantly impacting on populations dependant on this vital staple crop. Evaluation of different cassava cultivars for PPD symptoms has revealed differences in susceptibility to deterioration (Iglesias 1996), these
varietal differences could be a source of genetic material that can be used to improve the crop.

1.5.2 Molecular and biochemical understanding of PPD

The tuberous organs of cassava, sweet potato and taro are inclined to undergo mechanical wounding during harvest, storage and transport. This wounding leads to activation of different metabolic pathways in the different tubers and recovery from injury also differ significantly in these tuber crops. Natural or artificial wounding in sweet potato and yams is characterized by rapid lignification in the layers of cells close to the injury site, leading to healing (Uritani 1999), a feature found to be inadequate in cassava.

Comparative experiments by Uritani and colleagues (Uritani 1999) identified different types of polyphenolic compounds produced in the storage roots or tubers of these crops after wounding and infection. Chlorogenic acid and isochlorogenic acid were only found in sweet potato, while catechins were found in both taro and cassava. Coumarins were found in wounded cassava and sweet potato but not in taro. The predominant form of coumarin in cassava root preceding post-harvest deterioration was scopoletin (Uritani 1999). Phytoalexins were produced in taro and sweet potato as a result of microbial infection, in contrast to cassava where they were produced in the discolored region undergoing deterioration. Several key enzymes of the phenylpropanoid pathway including phenylalanine ammonia lyase (PAL), cinnamate 4-hydroxylase were induced in sweet potato (Tanaka 1974) and cassava, but the activity in cassava peaked much later (Tanaka 1983). Levels of peroxidase, which is involved in biosynthesis of lignin as well as detoxification of hydrogen peroxide, were also found to be higher in sweet potato than cassava. The high amount of polyphenols and lignin produced in sweet potato tissue and taro on wounding as opposed to cassava tissue possibly explains why wound healing in cassava does not occur. A clear understanding of the molecular mechanisms involved in cassava in response to PPD is of fundamental importance for the development of rational breeding programs (molecular markers) and transgenic strategies to improve or modulate the process. To this end, several approaches have been developed to characterize PPD development based on analysis of dramatic changes in gene expression patterns and biochemical studies on the changes of secondary metabolite accumulation during PPD.
In an attempt to characterize the cassava root transcriptome during the post-harvest period, a cDNA microarray study was performed (Reilly 2007). In this study, 72 non redundant expressed sequence tags which exhibited altered regulation during PPD were identified. Most of the up-regulated and PPD-specific expressed sequence tags were predicted to have functional annotations related to cellular processes including reactive oxygen species turnover, cell wall repair, programmed cell death, ion, water or metabolite transport, signal transduction or perception, stress response, metabolism and biosynthesis, and activation of protein synthesis.

Post-translational events such as protein cleavage, activation, and post translational modifications (PTMs) like phosphorylation mediate key regulatory events in plant responses to different stresses. It is increasingly becoming clear that, as compared to transcripts level, the protein changes integrates post-transcriptional and post-translational processing that modulates the quantity, localization and efficiency of the final cell products. Thus, proteome analysis studies have become an indispensable source of information about protein expression, splice variants, post translational modifications and protein-protein interactions for example in roots of rice (Koller 2002; Tanaka 2004; Komatsu and Konishi 2005), maize (Hochholdinger 2005).

Compared to other plants, very few studies have been undertaken using proteomics to understand the molecular and cellular functions of cassava proteins. This could be due to lack of a complete sequence genome. Cassava researchers have thus taken advantage of the advancement in DNA sequencing programs and expressed sequence tags (ESTs) libraries to develop proteome reference maps for cassava. By employing a combination of two-dimensional gel electrophoresis (2-DE) and matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS), differences in cassava fibrous and tuberous roots tissue proteins were evaluated (Sheffield 2006). A similar approach was also used identify protein markers for developmental stages of cassava somatic embryos (Baba 2008). However, the limited number of proteins identified in these two studies (237 and 83 respectively) would not allow for ascertaining the putative role of these proteins either in root development or secondary somatic embryogenesis. These two studies underscore the
potential of proteomics studies towards characterization of proteins involved in different developmental and biological processes in cassava.

Quantitative protein expression profiling has increasingly implemented stable isotopic labeling as an improvement or to complement the gel based methods. Isotope coded affinity tags (ICAT) reagents are used to covalently label cysteine residues with heavy or light hydrogen or carbon in two complex peptide samples. The labeled peptides are subsequently purified via affinity chromatography and then the samples are mixed and analyzed by LC-MS/MS (Gygi 1999). However, this technique has major limitations in that only cysteine residues are tagged and important information can be lost during the affinity purification steps. An improvement of ICAT technique involves the labeling of amine groups using a set of four or more isobaric tags. The isobaric tagging for relative and absolute quantification (iTRAQ) (Applied Biosystems) technique is superior to ICAT in that all the peptides are labeled; no affinity purification is required and it allows the analysis of up to four different samples simultaneously. This technique is beginning to receive wide application in plant proteome studies and has been used for example in studying organelle proteome in Arabidopsis (Dunkley 2006). The iTRAQ technique could be useful in surmounting some of the technical limitations encountered with 2DE for example the limitation to identification of abundant proteins, while missing out proteins of low abundance, high hydrophobicity, or extreme pl. The application of this technique to cassava would lead to detection of a large number of proteins.

1.5.3 Cassava PPD and signaling events

Cassava, like all other plants rely on various defense mechanisms for protection against injury from insects, microorganisms or mechanical wounding. Cassava is inherently predisposed to wounding due to mechanical injury that usually accompanies harvesting and handling of the storage roots (Booth 1976). Wounding presents a constant threat to plant survival because it not only physically destroys plant tissues, but also provides an entry pathway for pathogen invasion. To cope with wounding effectively, plants must prepare for pathogen attack while defending against insect predators. It is for this reason that it is hypothesized that plants may have evolved mechanisms that integrate the pathogen and wounding response.
Mechanical wounding and pathogen infection in plants arising from localized tissue damage activates defense responses not only at the site of injury (local), but that are also transmitted to other parts of undamaged tissues (systemic). These responses may involve the generation, translocation, perception and transmission of wound signals which are capable of inducing defense gene expression. Wound activated responses are directed at preparing the plant for the potential extension of wounding, healing of damaged tissues and the activation of defense mechanisms that prevent further damage by inhibiting the defense signals (Bowles 1998). These signaling pathways are mediated by many structurally different molecules including the oligopeptide systemin (Pearce 1991); molecules with hormonal activity such as jasmonic acid (JA), salicylic acid (SA), ethylene (Reymond and Farmer 1998), and hydrogen peroxide H$_2$O$_2$ (Orozco-Cardenas 1999) which further orchestrate the induction of defense responses.

Physiological deterioration in cassava parallels many features of plant wound responses since aspects of wound response are present in the harvested cassava root; however, the wound repair and subsequent down modulation of the signals necessary for restoration of integrity of damaged tissues are inadequate for wound healing. This deficiency leads to continuous cascades of wound responses that spread through the entire root. Interestingly, wounded cassava roots that remain attached to the plant are capable of normal wound repair (Mwenje 1998), suggesting that efficient wound repair mechanism of the detached root was lost during evolution, since the root serves no biological function once detached from the plant. Unlike cassava, tubers of yam, sweet potato and potato can act as propagules, since they accumulate storage proteins which can be broken down to support sprouting. The storage proteins possess biological activities consistent with a role in protecting the tubers against pests, pathogens and abiotic stresses (Shewry 2003). Cassava roots unfortunately lack true storage proteins which are necessary to mediate the stress responses associated with abiotic stresses for example wounding (Shewry 2003).

The development of PPD throughout the storage tissues of cassava after harvesting suggests the transmission of intercellular signals from the sites of damage, and apart from the involvement of ethylene and hydrogen peroxide in this signaling network, it remains widely unknown which other signaling molecules are involved and how their interplay affect PPD in
cassava. Research on the nature of mobile signals, their biochemical origin and the defense response genes they activate downstream are critical in developing a broader understanding of the PPD problem.

1.5.3.1 Reactive oxygen species production and signaling during cassava PPD

Reactive oxygen species (ROS) are continuously produced in plants as by-products of various metabolic pathways that are localized in different cellular compartments. The accumulation of partially reduced forms of oxygen such as singlet oxygen ($^{1}O_2$), superoxide anion ($O_2^-$), hydrogen peroxide ($H_2O_2$), and hydroxyl radical (HO') which are highly reactive and toxic can lead to oxidative destruction of cells (Asada 1987). Conversely in recent years, the role of ROS in control and regulation of biological processes such as growth, cell cycle, programmed cell death, hormone signaling, biotic and abiotic stress response has been identified (Mittler 2002; Mullineaux and Karpinski 2002; Overmyer 2003).

Adverse environmental conditions which can be biotic (imposed by other organisms) or abiotic (arising from an excess or deficit in the physical or chemical environment) disrupt the metabolic balance of cells resulting in enhanced accumulation of ROS (Mittler 2002). Armed with an expansive battery of enzymatic and non enzymatic antioxidant molecules (Fig 3), plants have since evolved an efficient ROS-scavenging mechanism. This mechanism maintains a tight regulation of ROS accumulation necessary for maintaining a low steady state baseline level of ROS in different cellular compartments and is necessary for protection against oxidative damage as well as signaling purposes. The major ROS-scavenging enzymes in plants include superoxide dismutase (SOD), ascorbate peroxidase (APX), catalase (CAT), glutathione peroxidase (GPX) and peroxiredoxin. In combination with the antioxidants ascorbic acid and glutathione (Noctor and Foyer 1998) these enzymes (Fig 3) provide cells with efficient machinery to detoxify $O_2^-$ and $H_2O_2$. 
Figure 3: Reactive oxygen species scavenging system in eukaryotic organisms

SODs act as the first line of defense against ROS by catalyzing the dismutation of superoxide to hydrogen peroxide and oxygen. It acts in concert with CAT to detoxify potentially dangerous H$_2$O$_2$ transforming it to water. Similarly APX utilizes ascorbic acid as an electron donor to reduce H$_2$O$_2$ with the concomitant generation of monodehydroascorbate, which is spontaneously disproportionated back to ascorbic acid and dehydroascorbate via the NADPH dependent monodehydroascorbate reductase (MDAR). Dehydroascorbate reductase (DHAR) utilizes glutathione (GSH) to reduce dehydroascorbate thereby regenerating ascorbic acid. GPX is the principle cellular enzyme capable of membrane lipid peroxidation repair and is considered as the first line of defense against oxidative membrane damage (Kuhn and Borchert 2002). Unlike other organisms, plants have multiple genes encoding SOD and APX isoforms that are specifically targeted to chloroplast, mitochondria, peroxisomes, as well as to the cytosol (Asada 1987).

On wounding, plants transiently produce reactive oxygen species (ROS) including the superoxide anion in the damaged tissue (Doke 1991), and hydrogen peroxide both locally and systemically (Orozco-Cardenas 1999). Wounding has been shown to stimulate the production of H$_2$O$_2$ in tobacco (*Nicotiana tabacum*), cassava (*Manihot esculenta* Crantz.), (Repka 1999) and sweet potato (Jih 2003) perhaps due to loss of compartmentalization.
within plant cells. The production and function of ROS in response to wounding, pathogen infection (Baron and Zambryski 1995) and postharvest senescence (Hodges and Forney 2000) has been well characterized in other plant systems but has never been investigated in detail in cassava.

The involvement of endogenous oxidative processes in PPD has been suggested by various authors based either on the observation that excluding oxygen from the cassava root tubers slows PPD (Rickard 1981; Plumbley 1991) or on the increase of activities of specific enzymes that use H$_2$O$_2$ as a substrate such as peroxidases. A rapid oxidative burst which generates the initial signals composing particularly of high levels of damaging ROS has been shown to occur in cassava storage root upon wounding (Reilly 2004). Wheatley and Schwabe (1985) suggested a possible correlation of PPD with an increase in phenolic compounds (e.g. scopoletin) and H$_2$O$_2$. In addition, it has been shown that a rapid increase of H$_2$O$_2$ occurs in cassava root over the first 24h of storage (Buschmann 2000). The generation of H$_2$O$_2$ after wounding seems to be widespread in the plant kingdom and may act as an internal chemical mediator involved in the onset of signal transduction pathways (Doke 1991). Hydrogen peroxide may function to regulate several signaling pathways. Microarray data from studies on *Arabidopsis thaliana* indicates that around 1% of the transcriptome is regulated by hydrogen peroxide (Desikan 2001). Hydrogen peroxide produced during the oxidative burst in cassava root could potentially act in a range of oxidative processes including peroxidase mediated oxidation of cellular components and cell wall cross-linking of hydroxyproline-rich glycoproteins (HRPGs) as part of the wound healing process. This signaling molecule could also lead to the induction of programmed cell death (PCD) related gene expression.

Increases in enzyme activity of both peroxidase and catalase, the two key enzymes involved in turnover of H$_2$O$_2$, have been shown to occur during PPD (Beeching 1998). In a separate study to identify genes expressed during PPD, the up regulation of genes involved in ROS turnover such as ascorbate peroxidase, catalase, secretory peroxidase, thioredoxin peroxidase, glutathione S- tranferase was demonstrated (Reilly 2007). The changes in expression patterns of the ROS regulating enzymes during PPD suggest the involvement of ROS in PPD development although no direct proof has been established. Although a complex interaction mechanism exists between the various ROS enzymes and general
oxidative stresses in plants, interaction between these pathways provides key regulatory potential for activating multiple resistance mechanisms. The over-expression of individual ROS regulating genes may typically increase oxidative stress resistance in plants and could enhance the ROS scavenging capacity in cassava storage roots thereby delaying PPD.

1.5.3.2 Ethylene and PPD in Cassava

A complex signaling network exists between plants, their pathogens and the abiotic stresses (Kende 1993). The phytohormone ethylene is a regulator of growth and development in plants and plays a key role in this network (Ecker 1995). Ethylene is synthesized from S-adenosyl methionine (SAM), which is a precursor substrate for many biochemical pathways, including polyamine biosynthesis (Ravanel 1998). SAM is synthesized from methionine by SAM synthase enzyme (Yang 1984; Kende 1993). The rate limiting step in ethylene biosynthesis is the conversion of SAM to 1-amino cyclopropane-1-carboxylic acid (ACC) via the ACC synthase enzyme (ACS). ACC is then oxidized by ACC oxidase (ACO) to ethylene (Yang 1984; Kende 1993).

Ethylene is involved in plant senescence, fruit ripening, abscission and the activation of senescence-related gene expression (Johnson 1998). It is induced upon a wide range of abiotic stresses such as wounding, flooding, chilling and desiccation or drought (O'Donnell 1996; Wang 2002). Ethylene also regulates plant responses to biotic stresses such as those induced by pathogens and enhances plant resistance against pathogen infection by the activation of the enzymatic activities of chitinase (Rakwal 2004), peroxidase, and phenylalanine ammonia lyase (Kato 2000). It is also thought to be responsible for inducing many biochemical processes that lead to programmed cell death (PCD) (Young 1997; Gunawardena 2001).

Wounding induces ethylene production in sweet potato roots within 24 hours (Sakai 1970) and was shown to precede wound lignification and wound periderm formation by 24 and 48 hours respectively (St. Amand 1989). Similarly, ethylene levels in damaged cassava roots were reported to increase after an initial lag period of 6 hours and the levels remained high over a 22 hour period (Plumbley 1981). A similar trend was observed by Hirose and co-workers whereby ethylene levels in cassava were shown to peak after 16 hour lag period
with varietal differences influencing to the rate of ethylene production (Hirose 1984). Pre-harvest pruning, which involves the elimination of the aerial part of the cassava plant, has been shown to suppress PPD development (Van Oirschot 2000). However, pre-harvest pruning has no significant influence on ethylene production following wounding and the exogenous application of ethylene on wounded roots does not affect the wound responses (Hirose 1984). All these observations have led to the speculation that ethylene may be involved albeit indirectly in signal transduction during PPD. Accelerated ethylene biosynthesis in plants is frequently associated with the induction of ACS and ACO genes which are responsible at least in part for the increased levels of enzyme activity under some conditions (McKeon 1995; Bleecker 2000; Wang 2002).

In a microarray study aimed at identifying genes expressed during cassava PPD, the ACC oxidase, was found to be up-regulated in cassava. Similarly in tomato, it has also been shown that ACC oxidase, is activated in response to wounding (Bowles 1998). The individual members of the ACS and ACCO gene families are regulated differentially in response to specific stimuli, suggesting that they play different roles in enhanced ethylene production in response to different environmental or endogenous cues (McKeon 1995; Bleecker 2000; Wang 2002). The involvement of ethylene in several agronomically and physiological important processes has contributed to its being targeted for manipulation by chemical and biotechnological methodologies (Mattoo 1991; Abeles 1992; Stearns 2003). Considerable progress has been made in identifying the mechanisms of ethylene perception, components of its signal transduction and biosynthesis pathway (Woeste 1998). However, detailed studies on the complex regulation of these pathways and of proteins involved still remains scarce and needs to be investigated.

1.5.3.3 Role of jasmonates

The plant hormone jasmonic acid (JA) is synthesized through the octadecanoid pathway from its α-linolenic acid precursor (Conconi 1996). Linoleic acid is a major fatty acid constituent of plant membranes and its application to tomato plants results in the expression of the same set of genes as the application of JA, suggesting that fatty acid release from complex membrane lipids is a critical checkpoint in JA biosynthesis (Leon 2001). The levels of endogenous jasmonates are elevated by stresses including mechanical
wounding, herbivory damage or pathogen attack. These in turn signal the induction of expression of specific jasmonate responsive genes to counteract the stress.

The wound induced synthesis of JA and its volatile derivative methyl jasmonate (MeJA) is one of the last steps in the signaling pathway which leads to activation of wound responsive genes including the induction of proteinase inhibitors (PIN) (Farmer 1992; Creelman 1997). Additionally, oligogalacturonides derived from plant cell walls upon injury have also been shown to be elicitors of wound inducible PIN gene expression in solanaceous plants (Doares 1995). In sweet potato the expression of the ipomoelin, a wound inducible protein, was shown to be enhanced by the application of methyl jasmonate (MeJA) and mechanical wounding (Imanishi 1997). However, regulation of the expression of this gene and its physiological role remain unclear.

Sporamin, a tuberous storage protein of sweet potato, is systemically expressed in leaves and stems by wound stimulation. The activation of the sporamin gene in sweet potato is similar to the wound activation of the proteinase inhibitor II gene (pin 2) in tomato, in that the pin 2 gene was up-regulated by MeJA and abscisic acid (ABA) and blocked by salicylic acid (SA) in the signal transduction pathway (Doares 1995). In addition, it has been shown that polygalacturonic acid and ABA, both known to mediate wound-induction of proteinase inhibitors in tomato and potato, induce the accumulation of sporamin mRNAs in sweet potato (Ohto 1992; Takeda 1995).

Currently there is no information regarding the role of MeJA as a possible signal leading to onset or progression of PPD. Research aimed at unraveling the major role of plant hormones such as jasmonic acid or its volatile derivative MeJA is necessary as this would shed more light into the contribution of these signaling molecules in initiating PPD.

1.5.3.4 Role of salicylic acid

The signaling molecule (SA) is crucial for local hypersensitive responses and systemic acquired resistance against many plant pathogens (Maleck 1999). SA interferes with wound-related gene expression by inhibiting the octadecanoid pathway (Pena-Cortés 1993; O'Donnell 1996). It has been reported that the defense-signaling pathways salicylic acid (SA)
and jasmonic acid (JA) can affect each other’s responses (Niki 1998). Salicylic acid was identified as a crucial signaling molecule required for the expression of plant defense responses and is important in the pathogen defense system (Klessig 1994). SA mediates the oxidative burst that leads to cell death in the hypersensitive response, and activates the pathogenesis-related (PR) genes in the systemic leaves after infection with avirulent pathogens. SA is known to have an antagonistic effect on the JA signal (Doares 1995).

SA also controls the defense system against other kinds of stresses. Treatment of mustard seedlings with exogenous SA improved their thermo-tolerance and heat acclimation (Dat 1998) while the pre-treatment of maize plants with SA has been shown to increase their chilling tolerance (Janda 1999). In Arabidopsis, SA participates in the tolerance against salt stress (Borsani 2001) leading to the expression of genes involved in developmental senescence (Morris 2000).

Despite the many studies being done to determine the role of SA in signaling pathways, none has so far shown a significant accumulation or reduction of SA during these responses indicating that SA is not the direct signal that induces these responses and further suggests that the pre-existing endogenous levels of SA may affect heat (Dat 1998), chilling (Janda 1999) and salt tolerance (Borsani 2001). It has also been shown that infection of tomato by Pseudomonas syringae pv tomato induces the accumulation of endogenous SA and enhances the wound-induced expression of phenylalanine ammonia lyase (Kohler 2002). This revelation shows that SA enhances the response of some wound-responsive genes after wounding. Although the relationships between SA and various responses have been described, the direct relationship between SA and wounding still remains obscure. Information on the involvement of SA and JA in wound response in cassava as well as in other root crops is still non-existent, necessitating the need to investigate the role these signaling molecules in stress resistance in cassava as well as in different root crops.

1.6 Conventional breeding and transgenic approaches to delay PPD

Genetic engineering can be used to complement conventional breeding methods in crop plant improvement. Transferring genes from heterologous species provides a means of selectively introducing new traits to the crop plants and increasing the gene pool beyond what can be achieved by conventional breeding approaches. Despite cassava being an
important food security crop in developing countries, it has often been neglected in plant breeding programs. Additionally, conventional breeding efforts have attempted to address many of the constraints facing cassava productivity, but with limited success. Progress has been slow due to the crop's complex genetic makeup which makes it difficult to breed efficiently.

Considerable effort has been devoted in improving cassava yield and resistance, for instance to the African cassava mosaic virus (ACMV) through conventional breeding. However, the lack of resistance genes in existing germplasm, high heterozygosity, poor flowering and the cross-pollinating nature of cassava limits the application of breeding, especially for a multigenic trait such as PPD (Jennings 2002). Additionally, a correlation found to exist between PPD and high dry matter (a desirable trait) could be difficult to separate via conventional breeding (Iglesias 1996) thereby suggesting the limitation of conventional breeding for delayed PPD. In addition, traditional breeding is usually difficult, time consuming and quite laborious. In light of these challenges to breeding for resistance to PPD, transgenic approaches look promising.

Biotechnological tools can be used to accelerate breeding programs through marker assisted selection although the mapping of quantitative trait loci (QTLs) linked to PPD in cassava crosses has not yet identified strong markers. The evaluation of all biochemical components involved in PPD and further comparison of biochemically distinct cassava cultivars may reveal basic insights into the PPD problem and thereby lead to identification of candidates which can be tested by breeding and molecular approaches. Delayed PPD has been identified in *Manihot walkerae*, a wild relative of cassava, and the development of inter-specific hybrids between cassava and *Manihot walkerae*, offers possibilities for mapping PPD related genes.

1.7 Rationale

There is little understanding of the physiological and biochemical processes that occur during PPD. At the molecular level the development of PPD is evidently a complex phenomenon involving multiple components. Cycloheximide inhibition of protein synthesis indicates PPD is an active rather than degenerative process involving changes in gene
expression and protein synthesis. Reports of analyses of genes expressed during cassava PPD demonstrated an altered regulation of genes and enzymes involved in signal transduction, reactive oxygen species (ROS) regulation, phytohormone synthesis, senescence, programmed cell death (PCD) responses, and the formation of compounds involved in the synthesis of cell wall components. While these studies have contributed towards identifying some of the key genes whose regulation indicate a critical role in cassava wound response, they were restricted to the elements of the process that are regulated at the transcript level leaving out equally important information on regulation of their gene products. Thus there is need to provide further insight into the molecular events associated with PPD and to identify novel PPD associated protein signatures in cassava root. The identification of these protein signatures is crucial in developing strategies to modulate or control the process.

In this research project, we have undertaken a proteomics profiling approach to analyze changes in total cellular protein content during PPD. Additionally, the functional role of some of the identified proteins have been characterized through enzyme assays, and their regulation at the transcript level validated through real time PCR analysis. The integration of data from the proteomics experiment with findings on changes in enzymatic patterns in addition to analysis of changes at the transcript level of the regulated gene products during PPD is discussed in this thesis and has contributed immensely towards a broader understanding of PPD problem.

Based on information from previous studies, which associate oxidative stress and ethylene biosynthesis with PPD development, transgenic approaches to delay PPD were attempted. This involved designing transformation constructs for enhancing the ROS scavenging capacity as well as constructs targeting regulation of the ethylene biosynthetic pathway in cassava roots.

**1.7.1 Aim of the thesis**

The main aim of this thesis was to investigate the molecular events underlying cassava post-harvest physiological deterioration (PPD) and to use biotechnological approaches to engineer cassava for delayed PPD.
In particular, the project attempted to address the following **key objectives**:

1. What are the protein dynamics that occur within the cassava storage root during PPD
2. Which pathways are associated with PPD in cassava
3. Determine if changes at the protein level reflect the changes occurring at the transcript level (especially for gene products involved in ROS regulation and cell wall metabolism) during PPD.
4. Determine if regulation of gene products and/or transcripts contribute to changes in the enzymatic patterns of ROS and cell wall enzymes identified.
5. Modulation of PPD through transformation of cassava with gene constructs for overexpressing ROS scavenging enzymes; Dehydroascorbate reductase (DHAR), glutathione reductase (GR), and glutathione peroxidase (GPX) as well as sense/antisense construct targeting regulation of 1-amino cyclopropane-1-carboxylic acid oxidase (ACO2) gene
Chapter 2 General Materials and Methods

2.1 Plant samples and PPD set up

2.1.1 Plant Material

Cassava storage roots (originally from Costa Rica) used in this study were purchased from Migros supermarket in Zurich, Switzerland. The roots were cleaned, and processed for PPD experiments at the Institute of Plant Sciences, ETH Zurich, Switzerland. Commercial roots of medium size, 16-18 cm minimum, without signs of mechanical damage and microbial contamination were selected.

2.1.2 PPD experiment set up

In order to mimic the events occurring during post-harvest physiological deterioration in the field, the storage roots were sliced and layered on Petri-dishes containing filter papers pre-soaked with 2 mls of distilled water (to prevent drying up) Fig 4. The sliced roots were then randomized and stored at a constant temperature of 27°C. Slices were taken and frozen in liquid nitrogen after incubation for 6h, 12h, 24h, 48h 72h and 96h. Samples of different time points were immediately frozen into liquid nitrogen then stored at -80°C for total RNA and protein extraction.

Figure 4: Cassava roots slices used for PPD experiment
2.2 Molecular Biology and Cassava Transformation Methods

2.2.1 Cassava total RNA and protein extraction

To 100 g of peeled cassava root slices, 20 mls of protein extraction buffer (20mM HEPES buffer pH 8, 5 mM MgCl and 1 tablet/50 ml buffer of EDTA-free protease inhibitor) was added and the mixture blended using a kitchen blender. The blended material was divided into two 50 ml Falcon tubes and further homogenized using Polytron homogenizer set at 7000 rpm for 1 minute. Samples were left on ice for 15 minutes, then centrifuged at 4° C, 4000 g for 20 minutes.

The supernatant from both tubes was collected and the contents from one Falcon tube centrifuged further at 12000 rpm for 20 minutes at 4° C (for protein isolation). After the second centrifugation, the supernatant was transferred to a 50 ml Falcon tube and 5 volumes of 80 % acetone added to precipitate proteins. The samples were kept at -20° C overnight. The precipitated proteins were centrifuged at 4000 g for 20 minutes at 4° C and the acetone discarded. The protein pellet was washed with 2 mls of 80 % ice cold acetone and the protein pellet allowed to dry to dry at room temperature for 10 minutes. The pellet was then re-suspended in 250 µl of 20 mM HEPES buffer pH 8, allowed to solubilize at room temperature for 10 minutes and thereafter the protein quantified.

The pellet from the first centrifugation was washed three times with 10 mls of the soluble protein extraction buffer, then incubated for 1 hour with the membrane protein extraction buffer (20 mM HEPES pH 8, 1 M NaCl, 0.1 % CHAPS, and 1 tablet/50 ml buffer of EDTA-free protease inhibitor). The extracted ‘membrane’ proteins were then desalted and concentrated to a final volume of 250ul using the Amicon Ultra-15 3kDA Ultracel membrane filters (Millipore AG, Switzerland). The proteins were quantified and kept at -20° C. The protein extraction protocol described above was used for the iTRAQ proteomics experiment.

To the supernatant from the second tube (centrifuged only once) 2 volumes of the RNA extraction buffer and 2 volumes of chloroform: isoamyl alcohol 24:1 was added and the samples incubated at 60 ° C for 15 minutes. The samples were centrifuged for 20 min at 2800 g at 4° C then the supernatant transferred and the chloroform extraction repeated. The supernatant from the second chloroform extraction was transferred to a 50 ml Falcon
tube and the RNA precipitated with 1/5 volumes of 10 M LiCl. The samples were incubated overnight at 4 °C. The samples were then centrifuged at 2800g for 20 minutes at 4° C. The supernatant was discarded and the pellet dried at room temperature for 10 minutes. The pellet from the extraction was re-suspended in 1 ml of DEPC treated water and transferred to new tubes. This was followed by the addition of 250 µl of 10 M LiCl and incubation overnight at 4° C.

The samples were centrifuged at 12,000 rpm at 4°C for 10 minutes and the supernatant discarded. The RNA pellet was re-suspended in 250 µl of DEPC-treated water. An ethanol precipitation was performed by adding 25 µl of sodium acetate pH 5.2 and 1 ml of 99 % ethanol. The samples were incubated for 1h at -20° C and then centrifuged at 12,000 rpm for 10 min, at 4° C. The pellet was dried for 10 minutes in the laminar flow hood. The pellet was re-suspended in 100 µl of DEPC-treated water, quantified and stored at -80 ° C.

Soluble proteins for two dimensional gel electrophoresis experiments were extracted from commercial cassava storage roots according to the modified method of (Cabral and Carvalho 1992). The root slices were lyophilized prior to protein extraction and the lyophilized sample used for protein extraction.

### 2.2.2 cDNA synthesis

Cassava cDNA synthesis was performed using the RevertAid™ first strand cDNA synthesis kit (Fermentas International Inc. Burlington, Canada) and the reaction done according to the manufacturers’ protocol. Briefly, 4 µg of total RNA was used for cDNA synthesis. To the RNA sample, 1 µl of oligo (dT)18, primers and 1 µl of DEPC treated water was added and the reaction incubated at 70° C for 5min. The samples were then cooled on ice and centrifuged for 1 min. The tubes were returned on ice and 4 µl of the 4X reaction buffer, 1 µl of Riboblock ribonuclease inhibitor and 2 µl of 10 mM dNTP mix added. The samples were centrifuged briefly and incubated at 37° C for 5 min. Following the incubation at 37° C, 1 µl of Revert Aid™ M-MULV reverse transcriptase was added and the samples incubated at 42° C for 1 h. The reaction was stopped by incubating the samples at 70 °C for 10 min, after which the samples were chilled on ice and the cDNA concentration determined.
2.2.3 PCR

*Arabidopsis* cDNA was used as template for amplification of genes involved in ROS regulation that would be subsequently over-expressed in cassava. These included glutathione peroxidase (GPX), Dehydroascorbate reductase, (DHAR) and Glutathione reductase (GR). Cassava cDNA was used as a template for the amplification of 1-amino cyclopropane-1-carboxylic acid oxidase 2 (ACO2) gene. The ACO2 would be cloned in antisense orientation and was chosen to target the down-regulation of ethylene biosynthesis, which has been shown to accompany PPD.

The needed DNA fragments were amplified through the polymerase chain reaction (PCR) method. For typical amplifications the following components were mixed. 2µl of cDNA template (10-50 ng), 5 µl of 10X eurobio Taq polymerase buffer, 2ul of 10 mM of each of the dNTPs, 2µl of 10 mM of each primer, 0.4 µl of EurobioTaq (5U/µl) and PCR water to a volume of 50ul. Typical program for amplification was as follows; denaturation for 45 seconds at 94 °C, followed by 30 cycles of denaturation for 15 seconds at 94 °C, annealing for 45 seconds at 56 °C and extension for 3 minutes at 68 °C. A final extension step at 72 °C for 10 min was included.
For primer combinations used to amplify target DNA sequences see table 1 below.

**Table 1: Primer combination sequences used to amplify target DNA.** Sequences in red represent the attB1 and attB2 sequences added to the gene specific primers. Sequences in black represent the gene specific primer sequence.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Source</th>
<th>attB Primers</th>
</tr>
</thead>
<tbody>
<tr>
<td>attB1_forward</td>
<td></td>
<td>GGGG ACA AGT TTG TAC AAA AAA GCA GGC T</td>
</tr>
<tr>
<td>attB2_reverse</td>
<td></td>
<td>GGGG AC CAC TTT GTA CAA GAA AGC TGG GT</td>
</tr>
<tr>
<td>Sense</td>
<td></td>
<td>Gene specific primers</td>
</tr>
<tr>
<td>DHAR_forward</td>
<td>Arabidopsis</td>
<td>AA GCA GGC TAC ATG GCT CTA GAT ATC TGC GTG AAG</td>
</tr>
<tr>
<td>DHAR_reverse</td>
<td></td>
<td>GAA AGC TGG GTC TCA CGC ATT CTT CTA TTC</td>
</tr>
<tr>
<td>GR_forward</td>
<td>Arabidopsis</td>
<td>AA GCA GGC TAC ATG GAG ATA CTT GTG ATG TCT</td>
</tr>
<tr>
<td>GR_reverse</td>
<td></td>
<td>GAA AGC TGG GTC CTA CAC CCC AGC AGC</td>
</tr>
<tr>
<td>ACCO2_forward</td>
<td>Cassava</td>
<td>AA GCA GGC TAC ATG GAG ATA CCT GTG ATG TTT</td>
</tr>
<tr>
<td>ACCO2_reverse</td>
<td></td>
<td>GAA AGC TGG GTC TTA AAC ATG GAG ACC ATG GTG GC</td>
</tr>
<tr>
<td>GPX_forward</td>
<td>Arabidopsis</td>
<td>AA GCA GGC TAC ATG GAG ATA CCT GTG ATG TTT</td>
</tr>
<tr>
<td>GPX_reverse</td>
<td></td>
<td>GAA AGC TGG GTC TTA AGA AGA GGC CTG TCC CAA C</td>
</tr>
<tr>
<td>Antisense</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DHAR_antisense_forward</td>
<td>Arabidopsis</td>
<td>AA GCA GGC TAC TCA CGC ATT CAC CTT CTA TTC</td>
</tr>
<tr>
<td>DHAR_antisense_reverse</td>
<td></td>
<td>GAA AGC TGG GTC ATG GCT CTA GAT ATC TGC GTG AAG</td>
</tr>
<tr>
<td>GR_antisense_forward</td>
<td>Arabidopsis</td>
<td>AA GCA GGC TAC CTA CAC CCC AGC AGC</td>
</tr>
<tr>
<td>GR_antisense_reverse</td>
<td></td>
<td>GAA AGC TGG GTC ATG GAG ATA CTT GTG ATG TTT</td>
</tr>
<tr>
<td>ACCO2_antisense_forward</td>
<td>Cassava</td>
<td>AA GCA GGC TAC TTA AAC ATG GAG ACC ATG GTG GC</td>
</tr>
<tr>
<td>ACCO2_antisense_reverse</td>
<td></td>
<td>GAA AGC TGG GTC ATG GAG ATA CTT GTG ATT GTG GC</td>
</tr>
<tr>
<td>GPX_antisense_forward</td>
<td>Arabidopsis</td>
<td>AA GCA GGC TAC TTA AGA AGA GGC CTG TCC CAA C</td>
</tr>
<tr>
<td>GPX_antisense_reverse</td>
<td></td>
<td>GAA AGC TGG GTC ATG GGC GAT GAA TCT CCA AAG TC</td>
</tr>
</tbody>
</table>

### 2.2.4 Agarose Gel Electrophoresis

To verify the correct size of the amplified fragments, 10 µl of the PCR products were separated on a 1.5 % w/v agarose in TAE buffer. Samples were mixed with 2 µl DNA loading buffer from Fermentas, (Fermentas International, Inc Canada) and loaded on the gel. Depending on the expected fragment size, either a 1 kb or 500 bp DNA ladder was included for sizing. Electrophoresis was carried out at 75 volts for $1-1\frac{1}{2}$ hours and the separated bands visualized using the UV transilluminator.
2.2.5 Purification of amplification fragments

The remaining PCR product from amplification of the target DNA templates were separated on a 1.5 % gel and the expected fragment excised from the gel under the UV transilluminator. Excised bands were gel purified using the Amersham purification kit (GE healthcare UK) according to the manufacturer’s instructions.

2.2.6 TOPO cloning

The amplified DNA fragments for designing the over expression constructs were cloned into TOPO cloning vector (Invitrogen life technologies, Carlsbad, California). TOPO Cloning® kit (Invitrogen life technologies, Carlsbad, California) provides a highly efficient, and fast cloning strategy for the direct insertion of Taq polymerase-amplified PCR products into a plasmid vector, thus no ligase, post-PCR procedures, or PCR primers containing specific sequences are required. The TOPO cloning reaction consisted of 4.5 µl of the gel purified PCR product, 1 µl of the dilute salt solution and 0.5 µl of the TOPO® vector and was set up according to the manufacturer’s instructions.

2.2.7 Construction of transformation vectors

The Gateway Technology with Clonase II (Invitrogen life technologies, Carlsbad, California) was adopted to transfer the cloned genes into a destination vector, which would subsequently be used to transform cassava. The cloning and recombination reactions described below were all carried out using Invitrogen kits (Invitrogen life technologies, Carlsbad, California) according to the manufacturer’s instructions. Briefly, the target genes (cloned in TOPO) were made gateway compatible by incorporating the attB1 and attB2 sites into the gene specific primers according to the Invitrogen protocol. The attB flanked PCR products were purified then combined with an attP substrate vector (pDONOR/Zeo) provided in the kit in a BP recombination reaction to generate the entry clones containing attL sites. The entry clones were propagated in competent E.coli TOP10F cells and plasmids extracted thereafter. Plasmids from the entry clones were sequenced to ascertain correct size and sequence information. The entry clones were combined with the destination vector
In a LR recombination reaction, leading to the generation of the expression clones (Fig 6). The expression clones were propagated in E.coli TOP10F.

**Figure 5**: Destination vector (modified from pCAMBIA 1301) used for generating the transformation constructs. The destination vector contains the Chloroamphenical (Cm R) resistance gene and ccdB gene which are replaced by the genes to be over-expressed in cassava.

**Figure 6**: Expression clone generated from the destination vector. ccdB gene was replaced with the genes for DAHR, GPX GR and ACC02 in sense and antisense orientation

### 2.2.8 Transformation into E. coli

Following the cloning and recombination reactions, 2 µl of the reaction products (TOPO cloning, BP or LR recombination reaction) was added to one vial of TOP10F competent cells and mixed gently and allowed to incubate for 5 min on ice. The cells were heat shocked for 30 seconds at 42 °C and returned back to ice, followed by addition the of 100 µl of SOC medium. The tubes were tightly capped and allowed to shake at 200 rpm at 37 °C for 1 hour. Selection plates (LB containing 50 µg/ml kanamycin) were warmed at 37 °C and 40 µl of 40 mg/ml X-gal in addition to 40 µl of 100 mM IPTG spread on the plates then incubated at 37 °C until ready for use. For entry clones generated from the BP reaction using pDONOR/Zeo vector, selection plates contained 50 µg/ml of Zeocin instead of kanamycin. The transformed cells (20 µl and 60 µl) were evenly spread onto the plates and the cells incubated at overnight at 37 °C.
2.2.9 Plasmid preparation

Plasmid DNA for use in confirmation of presence of inserts or for sequencing purposes was isolated using the Fermentas kit (Fermentas International, Inc Canada). Briefly, individual white colonies previously grown overnight on LB plates (transformed *E. coli*) were picked and cultured overnight at 37 °C while shaking at 200 rpm in 5 ml LB medium containing 50 μg/ml kanamycin. The mini prep plasmid DNA was obtained from a 3 ml bacterial suspension in accordance with the manufacturer’s instructions.

2.2.10 DNA and RNA yield determination

The genomic/plasmid DNA and RNA purity (measured as the ratio of absorbance at 260 and 280 nm, $A_{260/280}$) and concentration (ng/μl) were determined using the NanoDrop. The sample integrity was assessed by loading 5 μl each on a 1.5 % agarose gel electrophoresis.

2.2.11 Sequence analysis of the DNA

Sequencing was done to verify the correct sequence information of the amplified DNA or of the insert in the plasmid. Plasmid DNA material together with the specific primers were supplied to Microsynth sequencing facilities (Microsynth AG, Balgach Switzerland).

2.2.12 *Agrobacterium tumefaciens* and Cassava transformation

The expression clones (containing the genes to be transferred to cassava) were mobilized into *Agrobacterium tumefaciens* LBA4404 for transformation of cassava. *Agrobacterium*-mediated gene transfer was used to transform cassava friable embryogenic calli (FECs) according to the protocol outlined by Zhang and colleagues (Zhang 2000). Putative transgenic embryos were recovered from transformed FECs and maintained *in vitro* on regeneration media until the emergence of shoots (Zhang 2000).
2.3 Proteomics Methods

The same cassava root slices samples were used for RNA and protein extraction, the protein extraction methods used are described in section 2.2.1

2.3.1 Sample clean up

To remove excess salts, buffers and other non-protein impurities from sample preparation, the protein samples for 2D experiments were cleaned using the 2-D Clean-Up Kit (GE Healthcare/Amersham Biosciences). Briefly, 100 µl protein samples (containing 1–100 µg protein) was transferred to a 1.5 ml microcentrifuge tube and 300 µl of precipitant added. The contents were mixed by vortexing then incubated on ice for 15 minutes. After the incubation, 300 µl of co-precipitant was added to the mixture of protein and precipitant and the contents further mixed by vortexing briefly. The tubes were centrifuged in a microcentrifuge for 5 minutes at 12000 g. The supernatant was removed and 40 µl of co-precipitant added on top of the pellet. The tubes were allowed incubate on ice for 5 min, and the contents centrifuged again for 5 minutes. The supernatant was discarded and 25 µl of distilled water added to the pellet and again vortexed for 5–10 seconds. 1 ml of wash buffer (pre-chilled for at least 1 h at -20 °C) was then added to the pellet followed by the addition of 5 µl of wash additive from the kit. The tubes were vortexed for 20–30 seconds once every 10 minutes and incubated at -20 °C for 1h. After the incubation the tubes were centrifuged at 12 000 g for 5 minutes, then the supernatant was removed and the pellet allowed to dry for 5 minutes at room temperature. The protein pellet was then re-suspended in an appropriate volume of rehydration buffer (7 M urea, 2 M thiourea, 2 % CHAPS, DeStreak reagent 2 % v/v (Amersham), 2 % Pharmalyte, 10 % isopropanol, 5 % glycerol and 5 mM TBP), vortexed for 30 seconds then incubated at room temperature to fully dissolve. The protein mixture was then centrifuged at 12000 g for 5 minutes to remove any insoluble material.

2.3.2 Protein quantification

Total protein concentration for 2D-gel experiments was quantified using the 2D quant kit (GE healthcare, UK). Bovine serum albumin (BSA) protein standards containing between 0 and 50 µg of protein, and 3ul of each of the samples to be analyzed were transferred to
labeled 2 ml Eppendorf tubes and 500 µl of precipitant added. The contents were vortexed and incubated for 3 minutes at room temperature. 500 µl of the co-precipitant was then added and the contents mixed briefly then centrifuged for 5 minutes at 10000 rpm at room temperature. 100 µl of copper solution and 400 µl of water were then added to the pellet and the content vortexed briefly. 1 ml of the color reagent was then added and the samples incubated at room temperature for 15 minutes. The absorbance of each sample and standard was read at 480 nm using water as the reference.

A standard curve was generated by plotting the absorbance of the standards against the quantity of protein and this was then used to determine protein concentration of the samples.

2.3.3 Determination of protein concentration by Bradford Assay

Since some of the chemicals in the 2D quant kit (GE healthcare, UK) used in protein quantification may interfere with iTRAQ labeling, the proteins used for the iTRAQ experiment were quantified using the Bradford assay (Bradford 1976) with minor modifications. The commercial protein assay reagent (5X Roti-Nanoquant) was diluted to a working solution (1:5) with distilled water. Standards of known concentrations in the range of 0-50 µg protein (bovine serum albumin BSA) were prepared and 5 µl of the protein standards and of the samples were added to 1 ml of the diluted assay reagent. The samples were mixed by vortexing briefly and allowed to incubate for 15 minutes. Absorbance of the samples was then measured at 595 nm against a blank containing no proteins, and the readings noted. A standard curve of absorbance versus micrograms of protein was prepared and the concentration of protein in the samples determined from the curve.

2.3.4 Protein–cyanine dye labelling

Protein labelling was performed using the CyDyes DIGE Fluors developed for fluorescence 2-D DIGE technology (GE Healthcare/Amersham Biosciences) according to the manufacturer’s recommended protocol. Briefly the Cy2, Cy3, and Cy5 N-hydroxysuccinamide (NHS) esters were freshly dissolved in anhydrous N,N-dimethylformamide (DMF, 99.8%, Sigma-Aldrich Chemie GmbH, Taufkirchen, Germany) and centrifuged at 12 000 g, for 5 min at room temperature. In each case 50 µg of protein (pH 8) was labeled with 400 pmol of amine-reactive cyanine dyes. The labelling mixture was incubated on ice in the dark for 30 min and
centrifuged at 12,000 g, for 5 min at room temperature. The reaction was terminated by adding 10 µl of 10 mM lysine (Sigma-Aldrich), which reacts with the remaining free NHS esters of the cyanine dyes, and incubated on ice for 10 min. Each of the labeled protein samples were mixed and an equal volume of DIGE 2X sample buffer (7 M Urea, 2 M Thiourea, 3 % CHAPS and 2 % v/v Pharmalyte pH 3-10) were added prior to IEF. The experimental design shown in Table 2 was used for labelling cassava proteins. Proteins from the 0h, 24h and 48h PPD time points each with four biological replicates were labelled using Cy2, Cy3 and Cy5. Internal standard containing all the protein samples were labelled with Cy2. Sample replicates from 0h, 24h and 48h PPD were labelled with both Cy3 and Cy5 in a 6 gel experiment (Table 2).

Table 2: Cassava DIGE labelling design

<table>
<thead>
<tr>
<th>Gel No</th>
<th>Cy2 standard</th>
<th>Cy3</th>
<th>Cy5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50ug of (0h + 24h + 48h PPD)</td>
<td>50ug of 0h</td>
<td>50ug of 24h</td>
</tr>
<tr>
<td>2</td>
<td>50ug of (0h + 24h + 48h PPD)</td>
<td>50ug of 24h</td>
<td>50ug of 0h</td>
</tr>
<tr>
<td>3</td>
<td>50ug of (0h + 24h + 48h PPD)</td>
<td>50ug of 0h</td>
<td>50ug of 48h</td>
</tr>
<tr>
<td>4</td>
<td>50ug of (0h + 24h + 48h PPD)</td>
<td>50ug of 48h</td>
<td>50ug of 0h</td>
</tr>
<tr>
<td>5</td>
<td>50ug of (0h + 24h + 48h PPD)</td>
<td>50ug of 24h</td>
<td>50ug of 48h</td>
</tr>
<tr>
<td>6</td>
<td>50ug of (0h + 24h + 48h PPD)</td>
<td>50ug of 48h</td>
<td>50ug of 24h</td>
</tr>
</tbody>
</table>

2.3.5 Rehydration of strips and Running the IEF

For normal 2D-gel 24 cm IPG Strips, 450 µl rehydration buffer (7 M urea, 2 M thiourea, 2 % CHAPS, DeStreak reagent 2 % v/v (Amersham), 2 % Pharmalyte, 10 % isopropanol, 5 % glycerol and 5 mM TBP), was pipetted into the slots of the reswelling tray. The strips were prepared by first removing the protective cover using forceps starting at the acidic end. They were then positioned in the solution in the reswelling tray then lowered gently with the gel side down and overlaid with 3 ml cover fluid. The lid of the reswelling tray was covered and the strips allowed to rehydrate overnight at room temperature (10-20 h). For the DIGE experiment, 24 cm IPG strips with a non linear gradient (pH 3-10) were rehydrated using 450 µl of rehydration solution (DIGE labeling buffer: mixed with the 2X sample buffer.
on a 1:1 ration was used). The DIGE labeling buffer contained 7 M Urea, 2 M Thiourea, 3 % CHAPs and 30 mM Tris, while the 2 X sample buffer contained 7 M Urea, 2 M Thiourea, 3 % CHAPS and 2 % v/v Pharmalyte pH 3-10.

Filter papers of dimensions 5 x 5 mm were cut and dampened with water. After rehydration, the strips were removed from the reswelling tray then placed in the IPGphor cup loading strip holder with the positive end on the right, and gel side up. The damp filter paper wicks were placed at each end of the strip, sticking out slightly. The electrodes were placed over the wicks and the loading cup placed over the strip at the positive end, next to the electrode. The strips were covered with 3 ml of cover fluid and 100 µg of the protein sample previously dissolved in the rehydration buffer, loaded onto the strip through the cup. For CyDye labelled samples used in the DIGE experiment, 150 µg of protein was loaded. The strip holder was placed on the Ettan IPG phor platform (pointed end at the anode) and covered with a clear plastic cover. The IPG phor lid was then closed and the 1st dimension electrophoresis set up. The IEF parameters used for the 24 cm strip pH 3-10 is shown in table 3.

Table 3: The IEF parameters used for focusing for the 24 cm strip pH 3-10

<table>
<thead>
<tr>
<th>Step</th>
<th>Voltage</th>
<th>Type</th>
<th>Time (hours)</th>
<th>Total kVh</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>150V</td>
<td>Step/Hold</td>
<td>3:00</td>
<td>0.225</td>
</tr>
<tr>
<td>2</td>
<td>300V</td>
<td>Step/Hold</td>
<td>3:00</td>
<td>1.125</td>
</tr>
<tr>
<td>3</td>
<td>1000V</td>
<td>Gradient</td>
<td>5:00</td>
<td>5.025</td>
</tr>
<tr>
<td>4</td>
<td>8000V</td>
<td>Gradient</td>
<td>3:00</td>
<td>18.5</td>
</tr>
<tr>
<td>5</td>
<td>8000V</td>
<td>Step/Hold</td>
<td>6:00</td>
<td>66.5</td>
</tr>
<tr>
<td>6</td>
<td>500V</td>
<td>Step/Hold</td>
<td>9:00</td>
<td>- - -</td>
</tr>
</tbody>
</table>

Focusing was done overnight and the paper wicks changed occasionally to remove the excess salts. At the end of the run the total volt hours was noted. The above IEF parameters for 24 cm IPG strips pH 3-10 was also used for DIGE gels.
2.3.6 Casting and running of 2DE

Second dimension gel electrophoresis was performed on 12.5 % SDS - polyacrylamide gel prepared using the Ettan Daltll System Gel Caster (GE healthcare UK). The gel caster was tilted backwards and blank cassettes placed at the back of the gel caster. Glass plates were placed in front. Printed labels were then inserted between the glass plates (lower left corner) before inserting the glass plates in the caster box. The separator sheet was placed in the caster, then followed by the gel cassette and again another sheet followed by gel cassette or filler block until the caster was full. All empty space was filled in with thicker separator sheets. The front plate of the gel caster was fixed and the gel caster tilted forward to the upright position. The gel solution was then prepared, stirred briefly mix until it was completely dissolved and the solution was then degassed. Ammonium persulfate APS was added gently and the gel solution mixed until the APS was completely and evenly distributed. Using the peristaltic pump, the gel solution was poured on to the gel caster. 1 ml of 0.1% SDS was then layered on top of each gel (not over blanks) to ensure smooth edges of the gels. The gels were allowed to polymerize for 2h at room temperature. At the end of the 1st dimension electrophoresis, the strips were equilibrated in 4 ml SDS-equilibration buffer containing 10 mM DTT in 25 mM ammonium bicarbonate pH 8, for 15 minutes with shaking at 100 rpm. The strips were then transferred to 4 ml SDS-equilibration buffer containing 50mM IAA in 25 mM ammonium bicarbonate pH 8, for 15 minutes with shaking as above. During this time 0.1 % agarose sealing solution was prepared and kept at 58°C. The strips were then rinsed with distilled water and placed between the glass plates of the gel cassette using forceps, the marked part (e.g., pH 3-10) on the left. The strips were overlaid with 0.5 % agarose sealing solution and the agarose allowed to set.

The Ettan DALT II System bottom tank was filled with 7.5 l of 1x SDS electrophoresis buffer (25 mM Tris-HCl pH 8.3, 192 mM glycine, 0.1 % SDS). The required numbers of fillers were placed into the tank slots by first pre-wetting them in the SDS Electrophoresis buffer. Similarly gels were first submerged in the SDS Electrophoresis buffer, and then placed in the slots in the tank. The tank was then filled with approximately 1.9 l of 2x SDS electrophoresis buffer (contains twice the constituents of the 1x SDS electrophoresis buffer). The 2x buffer prevents glycine depletion and SDS depletion during the run. The lid was closed and the proteins resolved using the following parameters:
For 1 mm thick gels:

<table>
<thead>
<tr>
<th>Constant Power (Watts)</th>
<th>Overnight Run (milliAmperes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1: 2.0 W/gel, ie. 30 W for 12 gels, 45 min</td>
<td>7 mA/gel</td>
</tr>
<tr>
<td>Step 2: 17 W/gel: 180 W* for 12 gels, 4h</td>
<td>15 mA/gel</td>
</tr>
<tr>
<td>Step 3: 1 W/gel as a ‘hold’.</td>
<td>1 mA/gel</td>
</tr>
</tbody>
</table>

The gels were allowed to run until the blue tracking dye reached the bottom of the gel, after which the electrophoresis was stopped. The gel cassettes were then removed using metal holders and placed in plastic trays for staining.

2.3.7 Fixation, staining and destaining of the gels

The gels were fixed and stained with 1 μM ruthenium tris (bathophenanthrolate sulfonate) or RuBP according to the method of Lamanda, 2004. This method is less expensive, and twice as sensitive as SYPRO Ruby. Briefly, the gels were fixed overnight in fixing solution containing 30 % ethanol, 10 % acetic acid, then rinsed 3 times for 30 minutes each in 20 % ethanol. The gels were then incubated in 1 mM RuBP solution for 6h in the dark, and thereafter equilibrated twice for 10 minutes in water. After the equilibration, de-staining was done in 40 % ethanol, 10 % acetic acid for 16 h. The gels were again equilibrated twice in water for 10 minutes, then scanned using the Typhoon 9400 (GE healthcare UK).

2.3.8 Gel imaging scanning and analysis

The ruthenium stained 2-DE gels were scanned using the Typhoon 9400 and exported as TIFF files to the gel image analysis software Proteomweaver 3.0 (Definiens, Munich, Germany). Image analysis was done according to the proteome weaver manual.

For DIGE gels, the low fluorescence plates were used hence scanning was done while the gels were still assembled within the glass plates. Scanning was done according to the Ettan DIGE User Manual (18-1164-40 Edition AA). Briefly, the gels were placed on to scanner and the number of scan channels entered. Three channels (for the three different dyes) were selected for the DIGE gels, followed by selection of appropriate emission filters. The PMT voltage and sensitivity for each scan wavelength were then selected and scanning performed. The files were saved using the recommended DIGE file naming format. After
fluorescence imaging the DIGE gels were fixed and prepared for the second staining by fixing in 10 % methanol, 7 % acetic acid for 30 minutes twice then stained using the SYPRO® Ruby Protein Stain (BIO-RAD) for 4 hours. After staining the gels were rinsed twice in 10 % methanol, 7 % acetic acid for 30 minutes. All the steps were carried out in the dark on an orbital shaker with gentle agitation at 100 rpm. After staining, the gels were visualized using the typhoon 9400 scanner, and scanned at 100 microns/pixel and saved as gel file images, which were later converted to TIF file images.

2.3.9 Spot Picking and In-gel protein digestion

Spot excision was performed using the GelPix (Spot Picking Robot) system. This is a high-throughput protein spot excision system designed to image, analyze and excise stained protein spots from 2-D gels. The gels were equilibrated in distilled water for 15-30 minutes then imaged using an on-board camera. The camera images were viewed side-by-side with the typhoon scanned images to see if scaling and warping was done well. The images were then analyzed by Proteomweaver software, which generated a user defined protein target excision list. The selected protein spots were then excised using an eight channel excision head and loaded onto a Greiner 96-well MTP plate. Excised spots were reduced and alkylated then digested with 4ul of trypsin solution (7.5 ng/µl trypsin in Tris- HCl buffer pH 8.3) for 6h at 37º C. The digestion was performed using the Tecan Pro Team 150 robot.

2.3.10 Peptide desalting and MALDI plate spotting

The protein digests were desalted using the reversed-phase µ-C18 ZipTips. 10 µl of wetting solution (80 % ACN; 0.1 % TFA) was aspirated into the Zip tips and dispensed to waste. This was repeated once followed by aspiration of the wash solution (0.1 % TFA), also dispensed to waste. This process was repeated once. The peptides were bound to ZipTips by fully depressing the pipette plunger to dead stop and aspirating and dispensing the sample 5 to 10 cycles for maximum binding of peptides. The wash solution was then aspirated and dispensed to waste, twice. Approximately 0.7 µl of MALDI matrix solution (3 to 5 mg of cyano-4-hydroxycinnamic acid in 60 % ACN/0.1 % TFA per ml) was then pipette on to the target plate by using a regular pipette tip. The matrix solution was aspirated and dispensed up and down through the ZipTip three times, then dispensed onto the MALDI plate. Desalting of the peptide digests and spotting was done using the Tecan robot.
2.3.11 MALDI analysis of protein spots and database searches

The mass spectrometric measurements were performed on a 4700 MALDI-TOF/TOF mass spectrometer. The MS/MS spectra were searched against cassava EST database (downloaded in 2005, and contained 300,000 ESTs). The mass tolerance was set to 35 ppm, with toleration of one missed cleavage. The proteins were regarded as identified according to the significance criteria of the search program.

2.3.12 Reduction, denaturation and blocking of cysteine residues

Proteins from cassava at different PPD time points were extracted according to method described in section 2.2.1 and used for iTRAQ labeling experiment. Proteins were dissolved in 100 mM triethylammonium bicarbonate buffer at pH 8.5. Protein reduction, denaturation, and blocking of cysteine residues was carried out according to the iTRAQ™ reagent protocol using the reagents provided (Biosystems). Briefly, 100 µg of each sample was treated with 2 µL of reducing agent (tris(2-carboxyethyl) phosphine (TCEP)) at 60° C for 1 h and alkylated with 1 µL of cysteine blocking reagent, methyl methanethiosulfonate (MMTS) for 10 minutes at room temperature.

2.3.13 Protein Digestion

After blocking the cysteine residues, the protein samples were digested using sequencing grade trypsin (20 µg Promega) at a concentration corresponding to 5 % of protein by weight for 16 h at 37 °C. Peptides from each sample in a final volume of 50 µL were labeled with one of the four iTRAQ reagents at room temperature.

2.3.14 Labeling the protein digests with iTRAQ Reagents

Labeling of the samples (100 µg) with the 4-plex iTRAQ reagents (Applied Biosystems, Foster City, CA, USA) was performed according to the manufacturer’s recommendations. Protein extracts (soluble and membrane fraction) from 0h, 6h, 12h, 24 h (early PPD root slices ) and 0h, 48h, 72h and 96h (late PPD root slices) were labeled with iTRAQ reagents having molecular weights 114, 115, 116, and 117 Da respectively. Briefly, after centrifuging the peptide mixtures were resuspended in 30 µL of 0.5 M triethylammonium bicarbonate (TEAB), pH 8.5. The appropriate iTRAQ reagent (dissolved in 70 µL of ethanol) was added to
each of the samples and allowed to react for 1h at room temperature. The entire experiment was performed twice (biological replicates). The general workflow of the iTRAQ experiment is shown in figure 7. The same set-up was used for the late hour PPD time points. After labeling, all the 4 samples (4 PPD time points) were combined, concentrated and acidified to pH 3 using phosphoric acid.

Figure 7: General workflow of the iTRAQ experiment. The workflow shown above was used for the early PPD time points, the same was employed for the late PPD time points whereby 0h, 48h, 72h and 96h samples were labeled with 114, 115, 116, and 117 isobaric tags.

2.3.15 Strong cation exchange separation of peptides (HPLC)

After labeling tryptic digests from the different samples, the peptides were diluted with buffer A (10 mM potassium phosphate buffer, pH 2.85, 25 % ACN) in the ratio 1:3. The peptides were then mixed and fractionated on strong cation exchange chromatography on Poly-Sulfoethyl A column (PolyLC, Columbia, MD) (100 × 2.1 mm, 5 μm particles with 300 Å pores) using an Agilent 1100 HPLC system connected to a Probot fraction collector. Twenty-four SCX fractions (0.6 ml) were collected from a 0-350 mM KCl gradient in the presence of solvent A for 90 min at a flow rate of 0.3 ml/min. Solvent B contained 10 mM potassium
phosphate buffer, 350 mM KCl, pH 2.85, 25 % acetonitrile. Absorbance at 214 nm was used to monitor the eluent. Samples were dried by a speedvac prior to MALDI-TOF/TOF/MS analysis. The cation exchange fractions were split into two aliquots, with one aliquot initially analyzed by MALDI-TOF-TOF MS to ascertain the amount of peptides in the fraction and to check if labelling proceeded successfully.

2.3.16 Nano-LC separation and MALDI target spotting of tryptic peptides

Peptide separation was performed on an Ultimate chromatography system (Dionex - LC Packings, Sunnyvale, CA) equipped with a Probot MALDI spotting device. 5 µL of the samples were injected by using a Famos autosampler (Dionex - LC Packings) and loaded directly onto a 75 µm x 150 mm reversed-phase column (PepMap 100, 3 µm; Dionex - LC Packings). Peptides were eluted at a flow rate of 300 nL/min by using the following gradient: 0-10 min, 0% solvent B; 10-105 min, 0-50 % solvent B; and 105-115 minutes, 50-100 % solvent B. Solvent A contained 0.1 % TFA in 95:5 water/acetonitrile, and solvent B contained 0.1 % TFA in 20:80 water/acetonitrile. For MALDI analysis, the column effluent was directly mixed with MALDI matrix (3 mg/mL α-cyano-4-hydroxycinnamic acid in 70 % acetonitrile/0.1 % TFA) at a flow rate of 1.1 µL/min via a µ-Tee fitting. Fractions were automatically deposited every 10 s onto a MALDI target plate (Applied Biosystems/MDS Science, Foster City, CA) using a Probot micro fraction collector. A total of 416 spots were collected from each HPLC run.

2.3.17 MALDI-TOF/TOF mass spectrometry

MALDI plates were analyzed on a 4800 MALDI TOF/TOF system (Applied Biosystems) equipped with a Nd:YAG laser operating at 200 Hz. All mass spectra were recorded in positive reflector mode and generated by accumulating data from 800 laser shots. First, MS spectra were recorded from peptide standards on each of the six calibration spots, and the default calibration parameters were updated. Second, MS spectra were recorded for all sample spots on the MALDI target plate (416 spots per sample, 4 samples per plate). The MS spectra were recalibrated internally based on the ion signal of neurotensin peptide (Sigma, St. Louis, MO).

Spectral peaks that met the threshold criteria and were not on the exclusion list were included in the acquisition list for the MS/MS spectra. The following threshold criteria and
settings were used: Mass range: 800 to 4000 Da; minimum signal-to-noise (S/N) for MS/MS acquisition: 100; maximum number of peaks/spot: 8. Peptide CID was performed at a collision energy of 1 kV and a collision gas pressure of approximately $2.5 \times 10^{-6}$ Torr. During MS/MS data acquisition, a method with a stop condition was used. In this method, a minimum of 1000 shots (20 sub-spectra accumulated from 50 laser shots each) and a maximum of 2000 shots (40 sub-spectra) were allowed for each spectrum. The accumulation of additional laser shots was halted whenever at least 6 ion signals with a S/N of at least 60 were present in the accumulated MS/MS spectrum, in the region above $m/z$ 200.

2.3.18 Peptide and protein identification by database searching

GPS (Global Proteome Server) Explorer Software (Applied Biosystems) was used for submitting data acquired with the MALDI-TOF/TOF mass spectrometer for database searching. The MS and MS/MS data were searched using Mascot version 2.1.0 (Matrix Science, London, UK) as the search engine. All searches were performed against a Viridiplantae database (database released on April 2, 2008; contained 495180 protein sequences). The following search settings were used: maximum missed cleavages: 1; maximum number of signals per spectrum: 55; peptide mass tolerance: 25 ppm MS/MS tolerance: 0.2 to 0.3 Da. ITRAQ labeling of lysine and of the N-terminal amino group of peptides and methyl methanthiosulfonate (MMTS) derivatization of cysteine were specified as fixed modifications.

2.3.19 Statistical analysis of the iTRAQ data

Following the analysis of each reaction using GPS explorer, the control (0h PPD) sample was selected as the denominator for calculating the fold change of each protein. A ratio of one was assigned to the time point chosen as the denominator in each case. In this way, a reading was obtained for each of the four time points included in every reaction. A combined forward and reverse sequence database (Viridiplantae) was also searched to estimate the rate of false positive protein discovery and the method according to Käll and colleagues (Kall 2008) was used to calculate the false discovery rate (FDR scores).
The proteins identified were mapped to putative orthologs in Arabidopsis thaliana (NCBI, NC_003070.5; GI:42592260; 04-NOV-2005) by aligning each sequence from the eight datasets against all Arabidopsis amino-acid sequences (Smith-Waterman, significance threshold of 1e-20). If there was more than one significant alignment, the highest scoring one was selected. The highest scoring sequence in Arabidopsis is likely to be orthologous, though not guaranteed (Dessimoz et al. 2006).

2.4 Transcript analysis

2.4.1 Bioinformatics

The amino acid sequence for a selection of proteins identified in the iTRAQ experiment were blasted (tblastn) (http://blast.ncbi.nlm.nih.gov/Blast.cgi) against all Manihot esculenta ESTs retrieved from the public database in order to capture all the Manihot esculenta ESTs matching the peptide sequences originating from the proteomics experiment. The obtained cassava EST matches from NCBI (National Center for Biotechnology Information NCBI, Bethesda, MD, USA) were then assembled into contigs and singletons using the CAP3 sequence assembly software (http://deepc2.psi.iastate.edu/aat/cap/cap.html) and the encoded proteins predicted through a translation search tool (http://www.expasy.ch/tools/dna.html). Sequence comparison and cluster analysis was then done between the cassava predicted proteins (originating from the assembled contigs/singletons) and the Arabidopsis orthologs originating from blast searches performed against Arabidopsis thaliana TAIR8 protein database (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_clustalw.html). The combined data was summarized in a phylogenetic tree using the software from ETH, Switzerland (http://www.cbrg.ethz.ch/services/AllAll). Sequences of the assembled contigs were all aligned against each other using the MULTALIN tool, and the output used to design real time PCR primers (within non-conserved regions) that could discriminate between the different contigs and singletons.

2.4.2 Primers for real time PCR
The primers used for real time PCR are described in table 4.
Table 4: Primers used for real time PCR

<table>
<thead>
<tr>
<th>Reference genes</th>
<th>Forward primer</th>
<th>Reverse primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein phosphatase 2A (PP2A)</td>
<td>TGCAAGGCTCACACTTTCATC</td>
<td>CTGAGCGTAAAGCAGGGAAG</td>
</tr>
<tr>
<td>Cu/Zn SOD 1.1</td>
<td>GGGCCACACTTTAACCCTTC</td>
<td>AACAACAAACTGCCCCCTCCTATG</td>
</tr>
<tr>
<td>Cu/Zn SOD 1.2</td>
<td>CAACCTGGTGGAAGAGAGAGAG</td>
<td>CCCTTCTACAATGGAATGC</td>
</tr>
<tr>
<td>Cu/Zn SOD 2</td>
<td>GTCCCACACACTGGAGATGTT</td>
<td>ACCCCATCAGCATTAGCAAC</td>
</tr>
<tr>
<td>Cu/Zn SOD 3</td>
<td>CAATGGGGCTTACCATATCAC</td>
<td>CAAAATGTGCCCAAAATCC</td>
</tr>
<tr>
<td>APX 1.1</td>
<td>TGACCAGGATATGTTGCTTC</td>
<td>AAGGCCCTTCTCTCTCCAC</td>
</tr>
<tr>
<td>APX 1.2</td>
<td>GTGAGACGAACACTGGAGGTC</td>
<td>TGGGTGAAGAATGGAGATACAG</td>
</tr>
<tr>
<td>Beta 1,3-glucanase</td>
<td>CTCAATTCTGCTGTCTGAGGAG</td>
<td>CCAGCAATGGGAGATGTTGTC</td>
</tr>
<tr>
<td>Pectin methyl esterase 2</td>
<td>AGCAGAAACGTTGTTGATGG</td>
<td>ACCGCTGTGTTGGAAG</td>
</tr>
</tbody>
</table>

2.4.3 Real time PCR (qPCR)

RNA extraction and cDNA synthesis was done in accordance with the protocols described in section 2.2.1 and 2.2.2. Real time PCR reactions were done on the 7500 Fast Real Time PCR System (Applied biosystems, Foster City, CA) using the SDS software. A typical reaction mixture contained 10 µl SYBR Master Mix, 1µl each of both forward and reverse primers, 7µl of distilled water and 1 µl of cDNA template in a total volume of 20 µl. PCR thermal cycles used were as follows: Initial denaturation step for 20 seconds at 95 °C, followed by 45 cycles of denaturation for 3 seconds at 95 °C, annealing for 15 seconds at 60 °C and extension for 30 seconds at 72°C. For data analysis, the default setting of the 7500 fast System SDS software was used. The reactions were performed in duplicates and the results thereafter averaged. The 2Ct method (Livak and Schmittgen 2001) was used to perform relative quantification, using Manihot esculenta protein phosphatase 2A (PP2A) as the reference gene for mRNA expression.
2.5 Biochemical assays

2.5.1 ROS enzyme assays

2.5.1.1 Superoxide dismutase

SOD activity was determined spectrophotometrically in an indirect assay method based on xanthine oxidase and a novel color reagent using the SOD assay kit, Sigma -19160 (Sigma, St. Louis, MO) according to manufacturer’s instructions.

2.5.1.2 Ascorbate peroxidase

The ascorbate peroxidase reaction mixture (1 ml) contained 50 mM sodium acetate buffer (pH 6), 15 mM guaiacol, 40 µl of 240 mM H$_2$O$_2$ and 10 µg of enzyme extract. The absorbance was noted as increase in absorbance at 470 nm against the blank (lacking H$_2$O$_2$) and the rate of reaction compared with readings obtained from a standard curve derived from pure enzyme horse radish peroxidase (HRP). The enzyme activity was expressed in U mg$^{-1}$ protein (U = change in 0.1 absorbance min$^{-1}$ mg$^{-1}$ protein).

2.5.2 Cell wall enzyme assays

2.5.2.1 Gel diffusion assay for pectin methyl esterase

Gel diffusion assay was performed on the protein extracts to evaluate the changes in PME activity during PPD. The final gel matrix consisted 20 mM Tris HCl pH 7.6, 160 mM NaCl, 1% pectin, (citrus pectin) with 0.75 % agarose as a solidifying agent. The solution was heated to melt the agarose and allowed to cool then poured onto plates. Wells were made on the solidified matrix using the yellow pipette tips, 10 µg of the protein extracts from the different PPD time points were pipetted into the wells and the plates incubated overnight at 37° C. Negative control was carried out with boiled enzyme, while for the positive control 5µg of the pure commercial enzyme, pectin methyl esterase (PME) from orange peel (Sigma) was used. PME activity was visualized by staining with 0.05 % ruthenium red solution.
2.5.2.2 Pectin methyl esterase assay

PME activity leads to medium acidification in the presence of esterified pectin. The acidification was monitored using the pH indicator methyl red based on the protocol adapted from Hagerman and Austin (1986). In brief, esterified pectin from citrus (Sigma P-9561, 1 % in 20 mM Tris-HCl, 5 mM EDTA, 160 mM NaCl at pH 7.6) was incubated with 0.002 % methyl red and 4 µg of cassava membrane proteins. The assay was done on samples from three technical replicates. The absorbance was measured every 5 minutes at 416 nm and 517 nm. Different concentrations of the commercial enzyme pectin methyl esterase (PME) from orange peel (Sigma) were used to prepare a standard curve for activity measurements.

2.5.2.3 Beta 1, 3-glucanase

Glucanase activity was assayed by monitoring the amount of glucose released from laminarin (Laminaria digitata, Sigma) as substrate (Miller 1959). Soluble protein extracts were incubated with 1 mg/ml of laminarin dissolved in 50 mM potassium acetate buffer pH 5 at 40 °C for 30 min. Assay was done on extracts from three technical replicates. The reaction mixture was terminated by boiling for 5 minutes. The amount of glucose content was then measured spectrophotometrically at 540 nm using the glucose oxidase method (glucose oxidase kit, sigma-GAGO20). D-Glucose at concentrations from 0 to 50 µg/ml was used to construct a standard curve. One unit of beta-1, 3-glucanase activity was defined as the amount of the enzyme that released reducing saccharides equivalent to 1 µmol reducing sugars under the assay conditions.
Chapter 3 Results

3.1 Proteomic Results

3.1.1 Analysis of protein changes during cassava PPD using 2D gels

Proteomics has been used extensively to investigate the protein expression pattern under several different abiotic stresses. However, there has been no proteomic study conducted so far to determine stress responses in cassava during PPD. Proteomic techniques and tools currently available can aid effectively in understanding the molecular events and improve the comprehensive knowledge of the complexity of the networks underlying this stress condition in cassava. The purpose of this study was to apply 2D gel proteomic approach to identify proteins and genes that are differentially regulated during PPD.

We investigated changes in cassava root proteins during early PPD. Soluble proteins (0h and 12h) were extracted and analyzed by 2DE. A high resolution of 2-DE gel pattern in an isoelectric point (pI) range between 3 and 10 was detected by ruthenium staining. Approximately 800 - 900 proteins spots were reproducibly detected on each gel (Fig 8A and B). Differences in the intensity of protein spots between the control (0h) and early PPD time point (12h) were compared using the proteome weaver software.

Proteins spots showing at least a 1.5-fold difference in abundance between control and 12h time point or that were unique to the control sample (p<0.05) were selected, excised from the preparative gels and analyzed by MALDI-TOF/TOF (Fig 8C and D). A 1.5-fold threshold value was chosen in order to focus protein identification efforts to the most responsive proteins and for consistency with previous proteomic experiments. Quantitative image analysis, revealed a total of 74 protein spots which exhibited a significant change in expression in the tested PPD time points. Of these, 25 protein spots were visible only in the control sample (0h PPD) (Fig 8C), while 49 spots were down regulated after 12 h PPD while (Fig 8D).
Figure 8: 2D analysis of cassava root proteins during PPD. Comparison of protein patterns between 0h PPD (A) and 12h PPD (B), unique protein spots (C) and down regulated spots (D) identified, picked and analyzed.

Figure 9A and B shows the magnified view of two representative protein spots shown to change significantly during PPD. As shown, differences in the intensity of the representative protein spots 1242 and 507 between 0h and 12h were clearly discerned. The protein spot circled 1242 in the two replicate gels (Fig 9A) was predominantly seen in the 0h time point and was missed out in the 12h time point. It was noted that several other proteins were absent in the 12h PPD sample but present in the control. The presence of these protein spots only in the 0h time point (control) may be suggestive of qualitative changes in protein profiles that accompany PPD.
3.1.2 Proteins regulated 12hs after PPD, analysis by normal 2D gels

In order to determine the significant protein scores, we performed the Mascot Search within the cassava EST database using probability based Molecular Weight Search score (MOWSE). Protein scores greater than 63 were considered significant (p <0.05). Using this approach, we identified spots 1242 (that was unique to the 0h time point) and 507 that was significantly down-regulated after 12h (refer to figure 9A and B above) as thioredoxin peroxidase and pectin methylesterase respectively.

Further to this, of the 74 spots analyzed by tandem MS (MS/MS), we identified 33 candidate proteins by database comparisons. 12 out of 33 proteins were positively identified and were only visible in the control sample (Table 5) while the remaining 21 proteins were shown to be down-regulated (Table 6).
<table>
<thead>
<tr>
<th>Spot ID</th>
<th>Accession No.</th>
<th>Protein Name</th>
<th>pi</th>
<th>Mw</th>
<th>Protein score</th>
<th>Peptides</th>
</tr>
</thead>
<tbody>
<tr>
<td>1416</td>
<td>gi56921781</td>
<td>low Mw heat shock protein gene</td>
<td>8.18</td>
<td>17.67</td>
<td>135</td>
<td>5</td>
</tr>
<tr>
<td>1233</td>
<td>gi56924258</td>
<td>Actin</td>
<td>8.89</td>
<td>10.74</td>
<td>136</td>
<td>6</td>
</tr>
<tr>
<td>1417</td>
<td>gi56925102</td>
<td>cinnamyl alcohol dehydrogenase-like protein</td>
<td>7.22</td>
<td>16.21</td>
<td>252</td>
<td>4</td>
</tr>
<tr>
<td>1153</td>
<td>gi56921781</td>
<td>low MW heat shock protein gene</td>
<td>8.18</td>
<td>17.67</td>
<td>78</td>
<td>6</td>
</tr>
<tr>
<td>1258</td>
<td>gi56927724</td>
<td>malic enzyme</td>
<td>8.07</td>
<td>14.79</td>
<td>69</td>
<td>5</td>
</tr>
<tr>
<td>1159</td>
<td>gi56918749</td>
<td>allene oxide cyclase C4</td>
<td>9.44</td>
<td>23.58</td>
<td>143</td>
<td>8</td>
</tr>
<tr>
<td>1242</td>
<td>gi56917946</td>
<td>thioredoxin peroxidase</td>
<td>5.34</td>
<td>13.79</td>
<td>145</td>
<td>4</td>
</tr>
<tr>
<td>1297</td>
<td>gi56927357</td>
<td>17.4 kD class I small heat shock protein</td>
<td>7.22</td>
<td>16.21</td>
<td>169</td>
<td>4</td>
</tr>
<tr>
<td>1196</td>
<td>gi67208020</td>
<td>hypothetical protein</td>
<td>9.21</td>
<td>27.01</td>
<td>125</td>
<td>9</td>
</tr>
<tr>
<td>1478</td>
<td>gi67215572</td>
<td>17.7 kDa class I small heat shock protein</td>
<td>9.46</td>
<td>15.80</td>
<td>300</td>
<td>11</td>
</tr>
</tbody>
</table>
In addition to the above described proteins; thioredoxin peroxidase and pectin methylesterase, a number of other interesting candidate proteins linked to ROS regulation and stress defense mentioned in table 5 and 6 above require special highlights. These include; alcohol dehydrogenase-like protein, allene oxide cyclase, malic enzyme, ascorbate peroxidase 2, ascorbate peroxidase 3, dehydrin and lactoylglutathione lyase. Several small heat shock proteins were also identified. The identification of proteins involved in redox regulation, malate metabolism, jasmonate biosynthesis and cell wall metabolism underscores the potential role of these pathways in PPD development.

Table 6: Representative proteins down regulated 12hs after PPD

<table>
<thead>
<tr>
<th>Spot ID</th>
<th>Accession No.</th>
<th>Protein Name</th>
<th>pl</th>
<th>Mw (kDa)</th>
<th>Protein score</th>
<th>Peptides</th>
</tr>
</thead>
<tbody>
<tr>
<td>818</td>
<td>gi56922467</td>
<td>Elongation factor 1-beta</td>
<td>4.74</td>
<td>19.21</td>
<td>118</td>
<td>4</td>
</tr>
<tr>
<td>591</td>
<td>gi56926827</td>
<td>Elongation factor 1B alpha-subunit 2</td>
<td>4.81</td>
<td>23.25</td>
<td>141</td>
<td>7</td>
</tr>
<tr>
<td>960</td>
<td>gi56920161</td>
<td>Ascorbate peroxidase APX2</td>
<td>9.16</td>
<td>19.53</td>
<td>174</td>
<td>7</td>
</tr>
<tr>
<td>760</td>
<td>gi67216388</td>
<td>Glyceraldehyde 3-phosphate dehydrogenase</td>
<td>9.04</td>
<td>24.75</td>
<td>160</td>
<td>6</td>
</tr>
<tr>
<td>533</td>
<td>gi56928361</td>
<td>26.5 kDa class I small heat shock protein</td>
<td>9.79</td>
<td>20.24</td>
<td>88</td>
<td>4</td>
</tr>
<tr>
<td>991</td>
<td>gi67212205</td>
<td>Glyceraldehyde phosphate dehydrogenase</td>
<td>9.82</td>
<td>87.67</td>
<td>160</td>
<td>5</td>
</tr>
<tr>
<td>507</td>
<td>gi56920912</td>
<td>Putative pectin methylesterase</td>
<td>8.85</td>
<td>19.91</td>
<td>204</td>
<td>7</td>
</tr>
<tr>
<td>378</td>
<td>gi56918960</td>
<td>Ascorbate peroxidase APX3</td>
<td>5.76</td>
<td>17.32</td>
<td>204</td>
<td>7</td>
</tr>
<tr>
<td>488</td>
<td>gi56923016</td>
<td>Alpha-1,4-glucan phosphorylase</td>
<td>4.82</td>
<td>18.79</td>
<td>86</td>
<td>5</td>
</tr>
<tr>
<td>290</td>
<td>gi56919015</td>
<td>Putative lactoylglutathione lyase</td>
<td>5.2</td>
<td>23.49</td>
<td>114</td>
<td>8</td>
</tr>
<tr>
<td>919</td>
<td>gi56919325</td>
<td>Fructose-bisphosphate aldolase-like protein</td>
<td>8.39</td>
<td>20.04</td>
<td>184</td>
<td>8</td>
</tr>
<tr>
<td>354</td>
<td>gi56924258</td>
<td>Actin 2</td>
<td>8.89</td>
<td>10.74</td>
<td>142</td>
<td>5</td>
</tr>
<tr>
<td>244</td>
<td>gi67216388</td>
<td>Glyceraldehyde 3-phosphate dehydrogenase</td>
<td>9.04</td>
<td>24.74</td>
<td>130</td>
<td>8</td>
</tr>
<tr>
<td>327</td>
<td>gi56925373</td>
<td>Dehydrin</td>
<td>5.92</td>
<td>15.55</td>
<td>139</td>
<td>8</td>
</tr>
<tr>
<td>235</td>
<td>gi67208738</td>
<td>Chloroplast small heat shock protein</td>
<td>9.85</td>
<td>13.35</td>
<td>87</td>
<td>4</td>
</tr>
</tbody>
</table>
3.1.3 Analysis of protein changes during cassava PPD using DIGE gels

Due to the inherent limitations in separation capacity of 2D gels, several technical developments have been introduced to improve sensitivity, linearity, and reproducibility. The progress in 2-DE improvement includes gels with narrower pH gradients, staining procedures with improved dyes based either on covalent or non-covalent labeling of proteins (Patton 2000; Patton 2002). A novel approach in 2D gel based quantitative proteomics is the application of fluorescent cyanine dyes (Cy2, Cy3, Cy5) to label proteins before they are separated on a 2D gel (Unlu 1997; Unlu 1999). The characteristic nature of these labels allows the analysis of up to three pools of protein samples simultaneously on a single 2D gel. This approach eliminates to a great extent the gel-to-gel variation, which is the main limitation of 2D gel electrophoresis. In a standard protocol, two of the dyes (typically Cy3 and Cy5) are used to label two different pools of protein samples, while the third label (Cy2) is used to label an internal standard that consists of equal amounts of the two pools. This internal standard allows a correction for further experimental errors, thereby distinguishing biological from experimental variation (Alban 2003).

To detect proteins variably expressed in the root after 24 and 48h PPD, 2D DIGE patterns from the control (0h) and two PPD (24h, and 48h) time points were compared. The DIGE labeling design described in section 2.3.4 and table 2 was employed in this experiment, resulting in 150 µg of protein in each gel. Since there were two test samples (24h and 48h PPD), and the control (0h PPD), cross-labeling with the Cy5 and Cy3 dye was employed for all the samples (refer to table 2).

A representative image from a 2D DIGE gel of the 24h PPD sample is shown in Fig 10A. At least two gels representing two technical replicates were used for image analysis and an average of 600-700 spots could be reproducibly detected in each sypro stained gels (Fig 10A). Selected channels correspond to mixed sample IPS (Cy2, Fig 10B), 0h control (Cy5, Fig 10C), and 24h PPD (Cy3, Fig 10D).
3.1.4 Proteins found to be regulated 24h and 48h after PPD in the DIGE experiment

Comparative analysis using the proteomeweaver software revealed a total of 60 protein spots which showed a more than 1.5 fold (p < 0.05) difference in spot intensities in the 24h and 48h samples. In comparison to controls, 19 out of 60 spots increased in intensity and the remaining 41 spots showed decreased intensity. All the 60 spots were chosen and digested as described in section 2.3.9.
MALDI-TOF/TOF analysis was carried on the excised spots resulting in 34 protein identifications with a significant match (protein score above 63 was considered significant p <0.05). Of the identified protein spots, 10 were up-regulated (Table 7) while 24 proteins were down-regulated (Table 8). All the identified proteins showed at least a 1.5 fold change (p <0.05) compared to the control sample.

Table 7: A selection of proteins up-regulated in both 24h and 48h PPD compared to 0h from DIGE experiment

<table>
<thead>
<tr>
<th>Spot ID</th>
<th>Accession No.</th>
<th>Protein Name</th>
<th>pI</th>
<th>Mw kDa</th>
<th>Protein Score</th>
<th>Peptides</th>
</tr>
</thead>
<tbody>
<tr>
<td>1755</td>
<td>gi2660770</td>
<td>Cytosolic heat shock 70 protein</td>
<td>5.09</td>
<td>71.85</td>
<td>98.6</td>
<td>14</td>
</tr>
<tr>
<td>1583</td>
<td>gi33342046</td>
<td>Glyceraldehyde-3-phosphate dehydrogenase</td>
<td>8.44</td>
<td>24.07</td>
<td>280</td>
<td>7</td>
</tr>
<tr>
<td>1716</td>
<td>gi84453208</td>
<td>Putative cytosolic factor</td>
<td>4.72</td>
<td>67.82</td>
<td>96.5</td>
<td>4</td>
</tr>
<tr>
<td>1575</td>
<td>gi108707277</td>
<td>Ubiquitin, putative, expressed</td>
<td>9.56</td>
<td>11.44</td>
<td>47.3</td>
<td>6</td>
</tr>
<tr>
<td>1786</td>
<td>gi115474601</td>
<td>Os08g0126300</td>
<td>6.61</td>
<td>36.56</td>
<td>68.9</td>
<td>5</td>
</tr>
<tr>
<td>1656</td>
<td>gi84453208</td>
<td>Putative cytosolic factor</td>
<td>4.72</td>
<td>67.82</td>
<td>106</td>
<td>5</td>
</tr>
<tr>
<td>1512</td>
<td>gi34099812</td>
<td>Glyceraldehyde-3-phosphate dehydrogenase</td>
<td>6.4</td>
<td>31.9</td>
<td>189</td>
<td>6</td>
</tr>
<tr>
<td>1551</td>
<td>gi3243234</td>
<td>Isoflavone reductase related protein</td>
<td>6.02</td>
<td>33.8</td>
<td>139</td>
<td>4</td>
</tr>
<tr>
<td>1673</td>
<td>gi114795078</td>
<td>Glutathione S-transferase</td>
<td>5.42</td>
<td>23.63</td>
<td>53.3</td>
<td>4</td>
</tr>
<tr>
<td>1657</td>
<td>gi115441855</td>
<td>Os01g0918300</td>
<td>4.95</td>
<td>11.03</td>
<td>73.9</td>
<td>3</td>
</tr>
</tbody>
</table>
Table 8: A selection of proteins down-regulated in both 24h and 48h PPD compared to 0h from DIGE experiment

<table>
<thead>
<tr>
<th>Spot ID</th>
<th>Accession No.</th>
<th>Protein Name</th>
<th>pI</th>
<th>Mw kDa</th>
<th>Score</th>
<th>Peptide</th>
</tr>
</thead>
<tbody>
<tr>
<td>1715</td>
<td>gi1143427</td>
<td>Heat shock protein 70</td>
<td>5.15</td>
<td>75.48</td>
<td>601</td>
<td>29</td>
</tr>
<tr>
<td>1523</td>
<td>gi4139264</td>
<td>Actin</td>
<td>5.29</td>
<td>41.89</td>
<td>554</td>
<td>22</td>
</tr>
<tr>
<td>1618</td>
<td>gi758362</td>
<td>Adenosinetriphosphatase</td>
<td>6.02</td>
<td>55.79</td>
<td>448</td>
<td>25</td>
</tr>
<tr>
<td>1593</td>
<td>gi18831</td>
<td>Mitochondrial ATP synthase beta-subunit</td>
<td>5.95</td>
<td>60.33</td>
<td>342</td>
<td>16</td>
</tr>
<tr>
<td>1709</td>
<td>gi3377794</td>
<td>Proteosome IOTA subunit</td>
<td>5.83</td>
<td>27.49</td>
<td>196</td>
<td>9</td>
</tr>
<tr>
<td>1534</td>
<td>gi1771261</td>
<td>1,4-alpha-glucan branching enzyme</td>
<td>5.38</td>
<td>97.02</td>
<td>184</td>
<td>12</td>
</tr>
<tr>
<td>1690</td>
<td>gi21592744</td>
<td>Transaldolase-like protein</td>
<td>6.08</td>
<td>47.95</td>
<td>183</td>
<td>8</td>
</tr>
<tr>
<td>1767</td>
<td>gi113579941</td>
<td>Os05g0574500</td>
<td>6.66</td>
<td>25.36</td>
<td>180</td>
<td>10</td>
</tr>
<tr>
<td>1548</td>
<td>gi37928995</td>
<td>Cytosolic malate dehydrogenase</td>
<td>6.6</td>
<td>24.56</td>
<td>180</td>
<td>7</td>
</tr>
<tr>
<td>1540</td>
<td>gi1771261</td>
<td>1,4-alpha-glucan branching enzyme</td>
<td>5.38</td>
<td>97.022</td>
<td>172</td>
<td>12</td>
</tr>
<tr>
<td>1892</td>
<td>gi62526587</td>
<td>Ascorbate peroxidase APX2</td>
<td>5.31</td>
<td>27.77</td>
<td>138</td>
<td>6</td>
</tr>
<tr>
<td>1637</td>
<td>gi135406</td>
<td>Tubulin alpha-3-alpha-5 chain</td>
<td>4.95</td>
<td>50.25</td>
<td>123</td>
<td>7</td>
</tr>
<tr>
<td>1633</td>
<td>gi16930753</td>
<td>small heat shock protein</td>
<td>5.82</td>
<td>17.94</td>
<td>116</td>
<td>6</td>
</tr>
<tr>
<td>1780</td>
<td>gi62526587</td>
<td>Ascorbate peroxidase APX2</td>
<td>5.31</td>
<td>27.77</td>
<td>106</td>
<td>5</td>
</tr>
<tr>
<td>1576</td>
<td>gi90820120</td>
<td>UDP-glucose pyrophosphorylase</td>
<td>6.69</td>
<td>52.2</td>
<td>95.6</td>
<td>6</td>
</tr>
<tr>
<td>1566</td>
<td>gi90820120</td>
<td>UDP-glucose pyrophosphorylase</td>
<td>6.69</td>
<td>52.2</td>
<td>92.1</td>
<td>7</td>
</tr>
<tr>
<td>1613</td>
<td>gi9716271</td>
<td>Putative pectin methylesterase</td>
<td>8.24</td>
<td>63.33</td>
<td>91.6</td>
<td>6</td>
</tr>
<tr>
<td>1569</td>
<td>gi1771261</td>
<td>1,4-alpha-glucan branching enzyme</td>
<td>5.38</td>
<td>97.02</td>
<td>87.5</td>
<td>7</td>
</tr>
<tr>
<td>1745</td>
<td>gi34595979</td>
<td>Betaine aldehyde dehydrogenase</td>
<td>5.34</td>
<td>55.42</td>
<td>85.1</td>
<td>5</td>
</tr>
<tr>
<td>1612</td>
<td>gi1771261</td>
<td>1,4-alpha-glucan branching enzyme</td>
<td>5.38</td>
<td>97.02</td>
<td>76.2</td>
<td>7</td>
</tr>
<tr>
<td>1861</td>
<td>gi4928460</td>
<td>Thioredoxin peroxidase</td>
<td>4.82</td>
<td>13.95</td>
<td>70.1</td>
<td>5</td>
</tr>
<tr>
<td>1702</td>
<td>gi56784992</td>
<td>ATP synthase beta subunit</td>
<td>5.26</td>
<td>45.26</td>
<td>68.2</td>
<td>8</td>
</tr>
</tbody>
</table>
The up-regulated proteins identified requiring special mention included glutathione S transferase (GST), ubiquitin and isoflavone reductase (refer to table 7). More proteins were significantly down-regulated 24 and 48h after PPD than up-regulated. The down-regulated proteins included actin, malate dehydrogenase, proteosome subunit protein, small heat shock proteins, 1,4-glucan branching enzyme, and thioredoxin peroxidase among others (refer to table 8).

Two isoforms of the same protein were identified in adjacent spots; for example ascorbate peroxidase (refer to table 8, spots ID: 1780 and 1892 in bold), and several isoforms of the same protein could be identified for different protein spots, for example alpha glucan starch branching enzyme (refer to table 8, spots ID: 1534, 1540, 1569 and 1612 in bold), often these changes were in the same direction, but sometimes the altered expression was in the opposite direction. These adjacent spots identified as the same protein likely contained different post-translational modification (PTM) forms of the protein, whose relative abundance was altered during PPD.

In Figure 11, the intensity of the two spots circled (1534 and 1540) was shown to decrease in both 24h sample and 48h sample. These protein spots were clearly down-regulated in the 24 and 48h PPD samples. Both the protein spots were identified to be 1,4-glucan branching enzyme. The down-regulation of 1,4-glucan branching enzyme during PPD was also demonstrated in the normal 2D experiments, whereby at 12h PPD the protein abundance decreased significantly compared to the control sample (refer to table 6).
3.1.5 Classification of regulated proteins into pathways by Mapman

For proteins that were shown to be regulated, the *Arabidopsis thaliana* orthologs were obtained through BLAST (Basic Local Alignment Search Tool) searches. The physiological role of the proteins that were significantly induced or repressed in all the tested PPD time points compared to control were visualized with the software Map-Man (Thimm 2004). Map man bin classification of the best *Arabidopsis* sequence ortholog was used to assign proteins to specific pathways. The average change in protein expression in the tested PPD time points compared to control conditions for cassava is depicted in a Mapman graph (Fig. 12) to give an overview of the general regulation pattern of proteins/enzymes which change during PPD and the biochemical pathways in which they belong.
Figure 12: Overview of cassava proteins differentially expressed during PPD. The proteins are classified into functional bins according to Mapman. Red boxes indicate up regulated and blue boxes down-regulated genes. The grey boxes show proteins whose abundance does not change.

Although only a small number of proteins were positively identified after MS analysis, Mapman analysis software revealed the pronounced down-regulation of proteins belonging to biotic and abiotic stress regulation, cellular organization and redox regulation pathways, while enzymes involved in protein degradation were significantly up regulated. In addition some proteins of unknown function or with no known ontology were also found to be down-regulated. The characterization of these unknown proteins could lead to the identification of other novel proteins that play a significant role in initiating PPD.
3.1.6 Analysis of protein changes during PPD by iTRAQ

The iTRAQ strategy involves labeling peptides from different samples using isobaric tags (i.e., same mass, different tags) that fragment into different reporter ions upon collision induced dissociation. Peak area ratios of the reporter ions are subsequently used to assess the relative abundance of the peptides, and ultimately the proteins from which they are derived. An advantage of this approach is that it enables simultaneous identification and quantification of the same protein from different samples using MS/MS. Further, this technology is amenable to multiplexing for analysis of more than two samples in parallel, and like some of the shotgun methods, it also permits multiple independent measures to be made for the relative abundance of any given protein. Among the challenges associated with this and other quantitative shotgun approaches are the possibility that a particular proteolytic peptide may be derived from different proteins and that the MS/MS spectra from multiple precursors from a given protein may demonstrate significantly different quantifications for that protein (for technical or other reasons).

Conventional gel-based methods for determining protein expression patterns are insufficient for large scale proteomic applications. They are also limited to identification of abundant proteins, thereby missing out proteins of low abundance, high hydrophobicity, extreme pI or high Mr (López 2007). To overcome these shortcomings and to complement data obtained from 2D PAGE analysis, we used iTRAQ analysis (isobaric tags for relative and absolute quantification) to profile protein expression differences and quantify their relative abundance during PPD in cassava.

Cassava storage root protein samples were extracted from two biological replicates of a time course study and subjected to iTRAQ (isobaric tags for relative and absolute quantification) analysis. The expression profiles of soluble as well as membrane proteins from cassava root at early PPD (0h, 6h 12h and 24h) and late PPD (0h, 48h, 72h and 96h) (refer to section 2.3.14 and Fig 7) were analyzed in two independent biological experiments. These biological replicates are referred to henceforth as experiment 1 and experiment 2.
In total 8,960 proteins (confidence cut-off > 90 %) were identified from the two biological replicates, (5,665 proteins from experiment 1 and 3,295 proteins from experiment 2) based on searches against the *Viridiplantae* database (containing 495,180 protein sequences). Performing the search against a decoy (reversed) database allowed estimation of a false discovery level of 3.22 % and 2.57 % at 95 % confidence interval, for representative datasets from experiment 1 and 2 respectively.

Due to redundancy in the protein identities, the *Arabidopsis thaliana* orthologs of the identified proteins were determined and this led to the generation of orthologs in *Arabidopsis* for which 1,110 ATGs numbers (corresponding to unique proteins) could be assigned. There were a total of 711 proteins for which *Arabidopsis* orthologs could not be obtained. These proteins did not exhibit a significant match with any *Arabidopsis* protein, implying that they either do not exist in *Arabidopsis*, hence may be unique to cassava or that their polypeptide sequences may not be conserved within the plant kingdom. All plant genomes may not just be variants of the *Arabidopsis* gene set but may include a wide variety of genes that have no orthologs in *Arabidopsis*. Improved gene prediction algorithms and more refined genome annotation based on experimental characterization of genes believed to lack orthologs between pairs of species will be needed to resolve this issue.

The proteins identified were also aligned against each other thus leading to the generation of protein groups representing unique proteins (non-redundant dataset) and the 8,960 proteins could be collapsed into 1387 individual protein groups. However, further downstream analysis of the identified proteins was done using the *Arabidopsis* orthologs generated.

Comparison of unique proteins (with known *Arabidopsis* orthologs) from the two experiments revealed an intersection of 496 (45 %) proteins between the two experiments (Fig 13). This comparison was done in order to assess the reproducibility of the experimental replicates. Although a higher percentage of intersection would be expected from the two experiments, this was not the case and this could be attributed to high variability in terms of protein identifications from the two experiments and fewer proteins being identified in the second experiment compared to the first one.
Comparison was also done between proteins found in the soluble fraction and the membrane fraction in the two experiments. A total of 800 proteins were identified from the soluble fraction of both experiment 1 and 2, while 752 unique proteins were identified from the membrane fraction of both experiments, of these 442 proteins (40 %) were common to both fractions (Fig 14). Although only about 48 (4.4 %) more proteins were present in the soluble fraction compared to the membrane fraction, the relatively low overlap (40 %) between the different fractions could be associated with different proteins being enriched for in the two fractions as a result of the different extraction procedures used.

Figure 14: Overlap between protein from soluble and membrane proteins in both experiments
In order to ascertain if early and late PPD could be characterized by differences in proteins identified in the two stages of PPD, a comparison between the proteins identified from the early PPD time points and late PPD time points was determined from the two experiments. The number of proteins originating from the early PPD time points in both experiments was 885 while proteins that were found in the late PPD time points in both experiments was 838. A total of 613 (55 %) proteins were found to be common to both conditions (Fig 15). This suggests that more than half of the proteins identified in the entire experiment could be found in both early and late PPD stages. However, differences in the identity of proteins that characterize the two PPD stages could still not be excluded. A total of 94 proteins were found in both iTRAQ experimental replicates, all the PPD time points tested (early and late PPD) and in both the fractions (soluble and membrane), suggesting the possibility that these proteins exist in high copy number within the cassava storage root.

![Venn diagram showing overlap between early and late PPD proteins](image)

**Figure 15: Overlap between protein from early and late PPD time points in both experiments**

### 3.1.7 Determination of regulation cut-off for the iTRAQ identified proteins

Prior to the set up of the global iTRAQ experiment to determine changes in protein profiles of cassava storage root during PPD, preliminary iTRAQ experiments were conducted with the sole aim of establishing the threshold parameters that would be used to classify proteins as regulated. In two independent preliminary experiments, the reproducibility of technical replicates (similar cassava root material, but protein extraction, digestion and labelling done separately) and biological replicates (different cassava root material, and protein extraction,
digestion and labeling also done separately) was evaluated. In this experiment, the soluble and membrane enriched protein fractions from the technical and biological replicate samples were labeled with 4 iTRAQ labels. Two labels were used for the technical replicates and the other two labels for the biological replicates. The soluble and membrane fraction experiments were conducted separately. The fold change of the ratio of intensity of iTRAQ reporter ions in an ideal case would then be expected to be zero, since no treatment was introduced to the roots. Since large changes in protein abundance would not be expected in this experiment, a 2 fold change cut-off was set as the threshold for considering a protein to be regulated and applying this on the preliminary experiment data then allowed the determination of an acceptable confidence interval with which to work.

As pertains to the technical replicates, by choosing a 2 fold cut-off, 22 out of 637 (3.5 %) soluble proteins were found to change in abundance although ideally no change in abundance would be expected. By applying the same cut off for the membrane proteins, 5 out of 600 (0.8%) proteins changed in abundance. This implies that more than 95 % of the proteins identified in the technical replicates from both fractions did not show any significant change in abundance (Table 9). Taken together, this would mean that less than 4 % of all quantified proteins (Table 9) would be classified as regulated, (false positives) although ideally they originate from technical replicates and therefore would not be expected to show any change in abundance.

With respect to the biological replicates, 142 out of 637 (22 %) of the proteins in the soluble fraction changed in abundance, while 44 out of 600 (7.3 %) of proteins in the membrane fraction were shown to change in abundance when a 2 fold cut off was applied. The number of protein whose abundance changed in the biological replicates was clearly high in both fractions. These results show that there is a high level of variation within biological replicates as compared with technical replicates as would be expected. In the analysis of protein changes during PPD, the different time points would be considered as technical replicates, hence the 2 fold cut-off would still be applicable since as mentioned in the above paragraph, less than 4 % of all quantified proteins would be classified as regulated, although they should not change in abundance.
Table 9: Percentage numbers of proteins that show a 2 fold change in the soluble and membrane protein fractions

<table>
<thead>
<tr>
<th>Fraction</th>
<th>Total number of proteins identified</th>
<th>No of proteins that show a ±2 fold change</th>
<th>% regulated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Membrane (Technical replicates)</td>
<td>600</td>
<td>5</td>
<td>0.8</td>
</tr>
<tr>
<td>Soluble (Technical replicates)</td>
<td>637</td>
<td>22</td>
<td>3.5</td>
</tr>
<tr>
<td>Membrane (Biological replicates)</td>
<td>600</td>
<td>44</td>
<td>7.3</td>
</tr>
<tr>
<td>Soluble (Biological replicates)</td>
<td>637</td>
<td>142</td>
<td>22</td>
</tr>
</tbody>
</table>

3.1.8 Regulated proteins from iTRAQ experiment

Following our preliminary iTRAQ experiments described in section 3.1.7 above, the fold changes were calculated from the ratio of intensity of iTRAQ reporter ions obtained for 0 h PPD derived peptides (reference sample) compared to those derived from 6, 12, 24h PPD in the early PPD data set and 0 h derived peptides compared to those derived from 48, 72 and 96 h for late PPD data set. As defined before in section 3.1.7, we used the 2-fold cutoff to designate up or down regulated proteins, the expression level of 444 proteins from experiment 1 and 452 proteins from experiment 2 (soluble fraction) was found to change significantly (Table 10). Similarly, 527 proteins from experiment 1 and 339 proteins from experiment 2 (membrane fraction) showed a significant change in abundance (Table 11).
### Table 10: Number of soluble proteins found to be regulated

<table>
<thead>
<tr>
<th>Soluble Proteins</th>
<th>Experiment 1</th>
<th>Experiment 2</th>
<th>Overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Up regulated</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early PPD</td>
<td>33</td>
<td>133</td>
<td>14</td>
</tr>
<tr>
<td>Late PPD</td>
<td>198</td>
<td>161</td>
<td>56</td>
</tr>
<tr>
<td><strong>Down regulated</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early PPD</td>
<td>63</td>
<td>84</td>
<td>17</td>
</tr>
<tr>
<td>Late PPD</td>
<td>150</td>
<td>74</td>
<td>40</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>444</td>
<td>452</td>
<td>127</td>
</tr>
</tbody>
</table>

### Table 11: Number of membrane proteins found to be regulated

<table>
<thead>
<tr>
<th>Pellet fraction proteins</th>
<th>Experiment 1</th>
<th>Experiment 2</th>
<th>Overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Up regulated</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early PPD</td>
<td>72</td>
<td>35</td>
<td>5</td>
</tr>
<tr>
<td>Late PPD</td>
<td>132</td>
<td>145</td>
<td>50</td>
</tr>
<tr>
<td><strong>Down regulated</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early PPD</td>
<td>137</td>
<td>57</td>
<td>13</td>
</tr>
<tr>
<td>Late PPD</td>
<td>186</td>
<td>102</td>
<td>43</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>527</td>
<td>339</td>
<td>111</td>
</tr>
</tbody>
</table>

Taking all the regulated proteins from both soluble and membrane fractions, and considering the two different replicate experiments independently, we observed that 435 proteins were up-regulated in experiment 1, while 474 proteins were up-regulated in experiment 2. Similarly, 536 proteins were down-regulated in experiment 1, compared to 317 proteins in experiment 2. It is worth mentioning here that the total number of up or down-regulated proteins in both the experiments included those proteins that overlapped.
between early and late PPD time points as well as between the soluble and membrane fractions. Hence the totals would ideally be less than what is highlighted above if the overlapping proteins were included only once. We noted that 219 more proteins were down-regulated in experiment 2 compared to experiment 1 and 39 more proteins were up-regulated in experiment 1 compared to experiment 2.

Of the proteins shown to be up-regulated in experiment 1 and 2, 125 proteins were clearly up-regulated in both experiments. A selection of some of the up-regulated proteins and the corresponding sample identities (soluble or membrane fraction, and early or late PPD) are shown in table 12. Most of the proteins found to be regulated in both experiments also had a high protein score (p-value < 0.05). Of the 536 proteins down-regulated in experiment 1 and 317 proteins down-regulated in experiment 2, 113 proteins were shown to overlap between the two experiments. A selection of the proteins shown to be significantly down-regulated in both experiments is shown in table 13.
Table 12: A selection of proteins found to be up-regulated in both iTRAQ experiments

<table>
<thead>
<tr>
<th>Tag_Exp ID</th>
<th>Protein Name</th>
<th>Accession Number</th>
<th>Peptide Count</th>
<th>Best Ion Score</th>
<th>C.I. %</th>
<th>log:115/114</th>
<th>log 116:114</th>
<th>log 117:114</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp1_memb_late</td>
<td>Linamarase - <em>Manihot esculenta</em></td>
<td>Q41172</td>
<td>Q41172_MANES</td>
<td>1</td>
<td>99.99</td>
<td>2.722539</td>
<td>1.928995</td>
<td>2.307117</td>
</tr>
<tr>
<td>Exp1_memb_late</td>
<td>Xyloglucan endotransglucosylase/hydrolase precursor</td>
<td>Q6YDN9</td>
<td>XTH_BRAOB</td>
<td>1</td>
<td>100</td>
<td>2.542331</td>
<td>2.129108</td>
<td>1.644873</td>
</tr>
<tr>
<td>Exp1_memb_late</td>
<td>Cellulose synthase - <em>Populus tremula x Populus</em></td>
<td>Q6J8X1</td>
<td>Q6J8X1_9ROSI</td>
<td>1</td>
<td>95.853</td>
<td>2.348907</td>
<td>1.838446</td>
<td>1.601967</td>
</tr>
<tr>
<td>Exp1_sol_early</td>
<td>Phenylalanine ammonia-lyase 2 - <em>Manihot esculenta</em></td>
<td>Q94F89</td>
<td>Q94F89_MANES</td>
<td>3</td>
<td>99.99</td>
<td>0.227446</td>
<td>1.883884</td>
<td>1.901733</td>
</tr>
<tr>
<td>Exp1_sol_early</td>
<td>Superoxide dismutase [Cu-Zn] - <em>Manihot esculenta</em></td>
<td>Q58ZE5</td>
<td>Q58ZE5_MANES</td>
<td>2</td>
<td>100</td>
<td>0.400704</td>
<td>2.044439</td>
<td>0.85762</td>
</tr>
<tr>
<td>Exp1_sol_early</td>
<td>Glutamic acid-rich protein - <em>Manihot esculenta</em></td>
<td>Q6XQ13</td>
<td>Q6XQ13_MANES</td>
<td>5</td>
<td>100</td>
<td>0.690788</td>
<td>1.352276</td>
<td>0.620524</td>
</tr>
<tr>
<td>Exp1_sol_late</td>
<td>Hydroxycinnamoyl transferase - <em>Nicotiana tabacum</em></td>
<td>Q8GSM7</td>
<td>Q8GSM7_TOBAC</td>
<td>1</td>
<td>99.85</td>
<td>0.086482</td>
<td>1.162425</td>
<td>1.614315</td>
</tr>
<tr>
<td>Exp1_memb_late</td>
<td>Putative pectin methylesterase precursor - <em>Populus tremula</em></td>
<td>Q9FY03</td>
<td>Q9FY03_9ROSI</td>
<td>2</td>
<td>100</td>
<td>1.158967</td>
<td>1.327339</td>
<td>1.567962</td>
</tr>
<tr>
<td>Exp1_sol_early</td>
<td>Beta-1,3-glucanase (Fragment) - <em>Manihot esculenta</em></td>
<td>A0SVL9</td>
<td>A0SVL9_MANES</td>
<td>2</td>
<td>99.99</td>
<td>0.358123</td>
<td>1.169332</td>
<td>0.752395</td>
</tr>
<tr>
<td>Exp1_memb_late</td>
<td>Putative uncharacterized protein - <em>Populus trichocarpa</em></td>
<td>A9PHH6</td>
<td>A9PHH6_POPTR</td>
<td>1</td>
<td>99.91</td>
<td>2.199946</td>
<td>1.773284</td>
<td>0.562967</td>
</tr>
</tbody>
</table>


**Table 13: A selection of proteins identified as down-regulated in both iTRAQ experiments**

<table>
<thead>
<tr>
<th>Experiment identity</th>
<th>Protein Name</th>
<th>Accession Number</th>
<th>Peptide Count</th>
<th>Best Ion Score</th>
<th>log:115/114</th>
<th>log 116:114</th>
<th>log 117:114</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp1_memb_late</td>
<td>Putative glycine-rich RNA-binding protein - <em>Oryza sativa</em></td>
<td>Q6YN51</td>
<td>Q6YN51_PRUAV</td>
<td>2</td>
<td>100</td>
<td>-0.25475</td>
<td>-1.32952</td>
</tr>
<tr>
<td>Exp1_sol_late</td>
<td>Sulfite reductase (Fragment) - <em>Allium cepa</em></td>
<td>Q5IE29</td>
<td>Q5IE29_ALLCE</td>
<td>1</td>
<td>100</td>
<td>-2.19622</td>
<td>-4.48975</td>
</tr>
<tr>
<td>Exp1_sol_late</td>
<td>Chloroplast allene oxide cyclase precursor -</td>
<td>Q2PP43</td>
<td>Q2PP43_IPONI</td>
<td>3</td>
<td>99.99</td>
<td>-1.60768</td>
<td>-2.52078</td>
</tr>
<tr>
<td>Exp1_memb_early</td>
<td>Putative uncharacterized protein -</td>
<td>A3BYR8</td>
<td>A3BYR8_ORYSJ</td>
<td>1</td>
<td>99.24</td>
<td>-1.75617</td>
<td>-1.53764</td>
</tr>
<tr>
<td>Exp1_sol_early</td>
<td>Ascorbate peroxidase - <em>Manihot esculenta</em></td>
<td>Q52QX1</td>
<td>Q52QX1_MANES</td>
<td>9</td>
<td>100</td>
<td>0.43591</td>
<td>-0.04779</td>
</tr>
<tr>
<td>Exp1_memb_late</td>
<td>Dehydroascorbate reductase (Fragment) - <em>Manihot esculenta</em></td>
<td>Q1G149</td>
<td>Q1G149_SOLTU</td>
<td>1</td>
<td>99.99</td>
<td>-1.48791</td>
<td>-1.77378</td>
</tr>
<tr>
<td>Exp1_memb_late</td>
<td>Peroxidase (Fragment) - <em>Manihot esculenta</em></td>
<td>Q94IP4</td>
<td>Q94IP4_MANES</td>
<td>9</td>
<td>100</td>
<td>-1.48308</td>
<td>-1.16835</td>
</tr>
<tr>
<td>Exp1_memb_early</td>
<td>Cytosolic class I small heat shock protein 2A</td>
<td>Q53E40</td>
<td>Q53E40_TOBAC</td>
<td>4</td>
<td>99.99</td>
<td>-1.19633</td>
<td>-1.80505</td>
</tr>
<tr>
<td>Exp1_memb_late</td>
<td>Aquaporin - <em>Prunus persica</em></td>
<td>A5A8K8</td>
<td>A5A8K8_PRUPE</td>
<td>1</td>
<td>99.76</td>
<td>-2.52659</td>
<td>-2.3656</td>
</tr>
<tr>
<td>Exp1_memb_late</td>
<td>Putative uncharacterized protein - <em>Populus</em></td>
<td>A9P9Z5</td>
<td>A9P9Z5_POPTR</td>
<td>2</td>
<td>100</td>
<td>-0.24532</td>
<td>-1.29396</td>
</tr>
</tbody>
</table>
The proteins up-regulated during PPD were mainly associated with cell wall metabolism, secondary metabolite biosynthesis, defense/stress responses, and metabolic “housekeeping” functions. Down-regulated proteins identified were mostly associated with DNA modification, signaling, protein synthesis/processing and oxidative stress regulation. When compared to DIGE approach, iTRAQ technique identified up to 4 times more differentially expressed proteins, with at least 238 proteins identified from iTRAQ experiment compared to 60 protein spots from the gel-based method. Most of the proteins identified by DIGE experiment were also identified by iTRAQ, and some of the common identifications showed similar regulation between DIGE and iTRAQ, for example ascorbate peroxidase (refer to table 8, spot number 1892 and table 13) and small heat shock proteins (refer to table 8, spot number 1633 and table 13) which were down-regulated in both experiments.
As depicted in the previous section, we identified a wide range of regulated proteins using 2D-gels and DIGE approach (refer to section 3.1.2 and 3.1.4), and iTRAQ approach (refer to section 3.1.6). Using the TAIR8 database Gene Ontology (GO) annotation (Berardini, Mundodi et al. 2004), we evaluated the distribution of the regulated proteins identified by iTRAQ into different biological functions. We then compared the outcome with whole Arabidopsis thaliana genome annotation to determine which categories were under or over-represented. The analysis revealed an over-representation of proteins with functional annotation to protein metabolism, other metabolic processes, response to stress, response to abiotic and biotic stimuli, (Fig 16) and an under representation of proteins involved in biological processes such as transcriptional regulation for the up regulated proteins.

Similarly the down-regulated proteins were enriched with proteins with functional categories corresponding to protein metabolism, other cellular responses, response to stress and response to abiotic and biotic stimuli, and under-represented in the functional categories corresponding to transcription (Fig 17).
Figure 16: GO categories of the up regulated proteins from the iTRAQ experiment.

The categories were compared with GO categories generated from *Arabidopsis* whole genome annotation.
Figure 17: GO categories of all the down-regulated proteins from the iTRAQ experiment

The categories were compared with GO categories generated from *Arabidopsis* whole genome annotation.
As pointed in section 1.7.1, one of the aims of this work was to measure changes in protein levels linked to ROS regulation and cell wall metabolism. Figure 18 A-D shows the PPD protein abundance pattern of the key proteins selected for downstream processing. Superoxide dismutase and ascorbate peroxidase are involved in ROS regulation, while beta-1,3-glucanase and pectin methyl esterase have roles associated with cell wall metabolism. There was consistency in the abundance pattern for the selected proteins in the early PPD time points for the replicate experiments; however this trend was inconsistent in the later PPD time points of the two biological replicates. The consistency in protein abundance during early PPD in the two experiments coincides with primary PPD which is a physiological and developmental process, and starts from 24h to 48h after harvest depending on the cassava genotype. The protein abundance in the late PPD time points differed in the two experimental samples and this could be because the late time points represent secondary PPD which is the combination of primary PPD and deterioration due to micro-organisms such as bacteria and fungi.

Figure 18: Histograms showing the abundance ratio of the regulated proteins selected for downstream analysis
3.2 Transcript analysis

3.2.1 Comparison of regulated cassava gene products with their Arabidopsis orthologs

As mentioned before, the protein candidates found to be regulated using the iTRAQ profiling experiment and selected for further transcript analysis included, superoxide dismutase, ascorbate peroxidase beta 1,3-glucanase and pectin methyl esterase. We know most of these proteins are members of gene families and therefore it was fundamentally important to design real-time primers that could discriminate between the different family members or isozymes. This information would allow better interpretation of the transcript analysis data. Since the cassava genome is not sequenced, the available cassava EST sequences corresponding to these proteins were compared with their Arabidopsis orthologs and the outcome from this comparison formed a basis for the design of the real time PCR primers used to analyze transcript activity during PPD.

A total of 156 ESTs corresponding to Manihot esculenta Ascorbate peroxidases (MeAPXs) and 35 ESTS corresponding to Manihot esculenta Cu-Zn superoxide dismutases (MeCSDs) were obtained from the public database (NCBI, National Center for Biotechnology Information, Bethesda, MD, USA). All the ESTs corresponding to APXs were assembled into 5 contigs and 2 singletons (Fig 19), while those ESTs from CSDs could be assembled into 4 contigs. A phylogenetic tree was created using a BLAST pairwise program, based on the alignment of predicted amino acid sequences of Arabidopsis, and the corresponding cassava ESTs (Fig 19 and Fig 20).
Figure 19: APX gene family in cassava. The numbers in the figure show the scale for genetic distances. Peroxisomal, cytosolic, and chloroplastic are marked as Per., Cyt., and Chl. respectively.

With respect to APX (fig 19), 2 of the 5 contigs were predicted to be peroxisomal (named as Me APX 3.1 and Me APX 3.2, Fig 19 cluster designated Per) since they clustered more closely with the *Arabidopsis thaliana* At APX3 which is peroxisomal. Another two (named as Me APX 1.1 and Me APX 1.2 identified by iTRAQ experiment) were predicted to be cytosolic (Fig 19, cluster designated Cyt) and clustered more closely with the *Arabidopsis thaliana* At APX 1 and At APX 2. The last contig named as ctg5-MechIAPX was predicted to be chloroplast localized based on its clustering with At sAPx and At tAPX, which are known to be stromal and thylakoid localized, respectively (Fig 19 cluster designated Chl). For the singletons, one of them named as sgt1-Me APX2 was predicted to be cytosolic since it clustered more closely to the cytosolic At APX 2, while the other named as stg2-MeAPX 6 clustered more closely with At APX 6. There is still no clear information regarding the cellular localization for At APX6.
We also pointed out above that the 35 cassava ESTs corresponding to CSDs were assembled into 4 contigs (Fig 20), of which two were designated as ctg1-Me CSD 1.1 and ctg2-Me CSD 1.2 (identified by iTRAQ experiment) and were predicted to be cytosolic (Fig 20, cluster designated Cyt) as they clustered closely with the cytosolic *Arabidopsis thaliana* At CSD 1. Of the remaining two contigs, one ctg3-MeCSD 3 was predicted to be localized within the peroxisome (Fig 20 cluster designated Per) and the other one, ctg 4-CSD 2 to the chloroplast (Fig 20, cluster designated Chl).

**Figure 20: Cu-Zn SOD (CSD) gene family in cassava.** The numbers in the figure show the scale for genetic distances. Peroxisomal, cytosolic, and chloroplastic are marked as Per., Cyt., and Chl. respectively.

At the time of writing this thesis, two contigs MeAPX 1.2 for APX and MeCSD1.2 for CSD were being over-expressed in cassava roots under the root specific patatin promoter.

For beta-1,3-glucanase, a total of 8 ESTs that were identical to the *Manihot esculenta* beta-1,3-glucanase identified by iTRAQ experiment were assembled into a single contig corresponding to a cDNA sequence of 1263 bp. The closest homologs to this contig were
found in *Prunus persica*, *Populus trichocarpa* and *Arabidopsis thaliana*, as identified by BLASTP at NCBI. The beta-1,3-glucanase cDNA sequence was then used to design real time PCR primers.

As there is a large number of cassava PME ESTs, it was difficult to assemble all the ESTs into contigs and singletons. Therefore only those ESTs with significant matches to the *Arabidopsis* pectinesterase-2, (identified by the iTRAQ experiment), were assembled leading to generation of a partial contig corresponding to the C-terminal end of the *Manihot esculenta* PME2. The generated partial contig was used to design primers used in real PCR.

### 3.2.2 Quantitative PCR (qPCT) analysis of CSD and APX genes

As already emphasized, part of the objective of this work was to evaluate the correlation between protein expression levels and their corresponding mRNA transcripts. As a consequence we performed real time PCR analyses on a selection of gene products (highlighted in section 3.2.1) that displayed changes in abundance in the iTRAQ experiment (Fig 18 A-D). Significant accumulation differences that were initially detected on the protein level were confirmed on the mRNA level for some of the genes analyzed.

#### 3.2.2.1 Transcript expression analysis of APX during PPD

To test whether the change in abundance of APX protein observed from the iTRAQ experiment correlated with the expression of any given APX gene, the mRNA levels of three known APX genes were investigated. We have shown in figure 19 the subcellular localization of the APX isozymes. For this section, we studied the expression levels of the cytosolic APX isozymes; APX1 and APX2. The expression levels of cytosolic APX isozymes, APX 1.1 and APX 1.2 (Fig 21 A and B) were both shown to increase during early PPD between 6 and 12 hours, after which the transcript activity decreased and remained relatively stable after 48h PPD.

Although considered isozymes the trend observed in the pattern of activity for Me APX 1.1 and APX 1.2 was not very different. In both cases an up regulation was detected in the 6h and 12h PPD time points, suggesting that they could be copies of the same gene.
Linking the APX1.1 and APX1.2 protein regulation/expression data from iTRAQ experiments (refer to figure 18B) with their corresponding mRNA expression data shown above, we observe a clear correlation pattern in the APX1.1 and APX1.2 during the early PPD (0-24h). However, during the late PPD, we note that APX1.2 shows significant correlation in protein abundance with the transcript levels compared to the APX1.1 (Fig 18B to Fig 21B vs Fig 18B to Fig 21A above).

### 3.2.2.2 Transcript expression analysis of CSD during PPD

The expression patterns of all the 4 *Manihot esculenta* CSD genes (CSD 1.1, CSD 1.2, CSD2 and CSD 3) described in figure 20 was monitored by qPCR. Three of the four analysed CSDs; cytosolic CSD1.1, CSD1.2 and chloroplast-localised CSD 2 showed similar expression patterns in which there was a striking up-regulation during early PPD 6-12h. Their expression levels then dropped at 24h after PPD and thereafter remained relatively stable up to the 96h of late PPD (Fig 22A, B and Fig 23A).
**Figure 22**: Changes in transcript levels of the CSD1.1 and CSD 1.2 isoenzymes during PPD. Data are means ± standard error of the mean from three replicates.

In relation to the CSD 1.1, CSD 1.2 and CSD2, the mRNA of the peroxisomal CSD3 was comparatively less expressed at early PPD (6-12h) with levels remaining relatively low and stable after 24h to the late PPD of 96h (Fig 23B).

Due to the differences in sub-cellular localization of the CSD isozymes, we expected differences in the expression pattern but this was not the case as the chloroplast CSD2 had a similar pattern to the cytosolic CSD1.1 and 1.2. Although CSD3 mRNA expression levels from early to late PPD were comparatively low, all the four CSD isozymes had a somewhat similar expression pattern. They all started off with higher levels in the early PPD (6-12h) and gradually leveled off towards the 96h of late PPD. CSD isozymes have been associated with ROS scavenging. We hypothesise the induction of the cytosolic CSD isozymes during early PPD could be in response to increased ROS generation during this period. The induction of CSD isozymes during early PPD was also reflected in the proteomics experiment as demonstrated by the iTRAQ data (refer to figure 18 A).
3.2.3 Quantitative PCR (qPCR) analysis of genes involved in cell wall metabolism

3.2.3.1 Beta 1, 3-glucanase

We have previously stated that all the 8 ESTs of cassava beta-1,3-glucanase were assembled into a single contig that had closest homology to the beta-1,3-glucanase enzymes in other species. This is a key enzyme involved in cell wall metabolism. We have also shown in our iTRAQ experiment that this enzyme is regulated during PPD. As a logical follow-up to these observations, we studied the mRNA expression levels of this enzyme. Our data show a significant increase in the enzyme at the 12h of early PPD followed by a sudden drop at the 24h after which the levels remain low and stable up to the 96h of late PPD. (Fig 24). This clearly demonstrates that for PPD certain key genes are activated during the early phase of the deterioration and their function is limited in the later phases of PPD.

The transcript data correlated well with the iTRAQ data as there was a striking increase in transcriptional activity of beta 1,3-glucanase from 0-12h PPD which was in line with a comparative increase in protein abundance from 6-12h PPD (refer to fig 18C).

Figure 23: Changes in transcript levels of the CSD2 and CSD3 isoenzymes during PPD. Data are means ± standard error of the mean from three replicates.
3.2.3.2 Pectin methyl esterase

Another cell wall metabolism enzyme of interest is PME. Our iTRAQ data (refer to fig 18D) shows this protein to be regulated during the entire spectrum of PPD (0-96h). Like in the previous cases, we also analyzed the mRNA expression pattern of PME as a framework to understand whether the protein changes observed are linked to the corresponding mRNA. We found PME increased in the early PPD with the highest peak at the 24h followed immediately by a slight drop at the 48h. Levels then remained fairly stable up to the 96h (Fig 25). Although the iTRAQ levels increased gradually from the early to late PPD (refer to fig 18D), the transcript and iTRAQ data showed higher levels in both early and late PPD.
Figure 25: Changes in transcript level of pectin methyl esterase during cassava PPD. Data are means ± standard error of the mean from three technical replicates.
3.3 Enzyme assays

To complement the iTRAQ experiment data, enzymatic assays for a selection of the proteins identified as regulated by iTRAQ and qPCR was performed. The activities of two soluble proteins involved in ROS regulation (APX and SOD) as well two cell wall proteins (PME and beta 1, 3-glucanase) were assayed.

3.3.1 Changes in activity of APX during PPD

APX activity was monitored as described in section 2.5.1.2. Our data shows a gradual increase in the activity of APX from 0h to the 96h of late PPD (Fig 26). Comparing the two phases, APX had a remarkable activity in the late PPD phase (72-96h) compared to the early PPD phase (6-12h). This is in contrast to the mRNA levels (refer to figures 21 A and B) and protein abundance (refer to figure 18B). In these two cases, the cytosolic APX1.1 and APX1.2 mRNA levels and protein abundance were higher in the early PPD. This discrepancy could be as a result of measurement of whole (crude) versus the specific isozyme. It is worth noting that in the measurements of APX activity, protein extracts integrate the activities of the different APX isoenzymes. Soluble protein extracts used in the activity assay most likely contained cytosolic, microsomal, and stromal APX isoforms which could display different patterns of activity compared to the cytosolic isoforms analysed in the iTRAQ experiment and transcript analysis. This could help explain the differences observed in the APX enzyme activity compared to the transcript or protein abundance data.

![APX activity](Image)

**Figure 26: Changes in total APX activity during PPD.** Activity is relative to 0hs. Data are means ± standard error of the mean from three replicates.
Although the steady-state mRNA levels of cytosolic APX 1.1 and APX 1.2 increased only during 6-12h PPD (refer to fig 21), and the protein abundance remained relatively stable over these time points, the activity of APX isozymes increased markedly throughout the PPD time points tested suggesting that complex post translational mechanisms accompany PPD in cassava.

3.3.2 Changes in activity of SOD during PPD

The superoxide dismutase activity in the samples was determined in an indirect assay method based on xanthine oxidase and a novel color reagent using the SOD assay kit from sigma (S-19160). We observed two distinct patterns in the activity of SOD between 0-96h (Fig 27). In case 1, there was a gradual increase in SOD activity from 0 to 12 h followed by a drop in the 24h. In case 2, the activity picked up at 48h and then gradually increased with peak levels registered at 96h. Putting this result into context with the corresponding transcript and iTRAQ data, we observed a similar pattern in the mRNA regulation and protein expression in early PPD. All the three revealed a gradual increase up to 12h followed by a decline in the 24h. However after 24h the patterns differed. With an exception at the 72h of exp 1 of iTRAQ (refer to 18A), we observed that the mRNA and protein levels remained generally low and stable up to 96h while for enzyme assay, the activity gradually increased as stated above.

**Figure 27: Changes in total SOD activity during PPD.** Activity is relative to 0hs. Data are means ± standard error of the mean from three replicates.
There are three classes of SOD variants depending on their metal cofactors namely, manganese-containing SOD (Mn-SOD; mitochondrial and peroxisomal), iron-containing SOD (Fe-SOD; chloroplastic), and copper and zinc containing SOD (Cu-Zn SOD; cytosolic, chloroplastic and apoplastic). The SOD assay kit used in the assay had one limitation of not being able to readily distinguish between the activities of any of the three forms of SODs but rather gave the total activity of all the isoforms. This limitation could partly contribute to the slight inconsistency between enzyme activity assay and the mRNA and iTRAQ data.

3.3.3 Change in activity of Beta 1, 3-glucanase during PPD

Changes in cell wall development may involve changes in the level of expression of genes coding for proteins (cell wall-related enzymes) required for cell wall loosening and cellulose synthesis. Marked changes in the pattern of translatable mRNAs is associated with the transition from primary to secondary wall synthesis (Delmer et al., 1985). Therefore, regulation between cell elongation and cellulose deposition during development may occur at many levels with the increase and decrease in protein (enzyme) levels. Herein, we measured the activity of beta 1,3-glucanase, and pectin methyl esterase in cassava root during PPD.

Activity of beta-1,3-glucanase was detected in the soluble fraction of the protein extracts from cassava using the assay described in section 2.5.2.3. Two distinct patterns in the enzyme activity were observed (Fig 28). In the first one, the activity increased gradually during the early PPD time point, 6-12h, and then slightly decreased at the start of the late time point, 48h. In the second one, the activity picked up after the 48h and gradually increased up to the 96h. Time points 24h and 96h of early and late PPD respectively registered the highest activities. Relating the activity assay to the transcript and iTRAQ data, we noticed similarities in the early PPD. Here, the enzyme activity, mRNA and protein expression increased gradually from 0-12h but at the 24h, the mRNA and protein expression dropped. Unlike in mRNA where the levels remained low and stable between 24-96h, for the enzyme assay and iTRAQ data they shared a similar pattern from 48-96h since the activity and protein levels gradually increased.
Figure 28: Beta 1,3-glucanase activity in cassava roots during PPD. Activity is relative to 0h. Mean ± standard error of the mean from three technical replicates.

3.3.4 Change in activity of pectin methyl esterase

Proteins extracted from the pellet fraction produced in line with the extraction procedure described in section 2.2.1 were analyzed for their pectin methyl esterase activity (PME, EC 3.1.1.11). PME activity was assayed qualitatively using the gel diffusion assay and quantitatively based on colorimetric measurements. This assay is based on the increased binding of ruthenium red to pectin as the number of methyl esters attached to the pectin decreases. 10 µg of membrane protein from each of the PPD time points was loaded into the sample wells formed on the agarose matrix. Commercial pectin esterase from orange peel was used as a positive control (5 µg of enzyme loaded) and the boiled enzyme was used as a negative control. From the gel diffusion assay, the activity of PME visualized as dark ring formation (stained zones) around the wells containing protein extract could be observed in all the PPD samples tested (Fig 29).
Figure 29: Gel diffusion assay revealing PME activity during PPD. Wells were loaded with 5 ug of protein extracts from early, 0-24h (A) and late 0-96h PPD (B). White arrow shows the positive control and black arrow the negative control.

The intensity and size of the stained zones from extracts originating from the different time points were generally similar hence differences in PME activity in the extracts could not be ascertained. To confirm the qualitative data obtained from the gel diffusion assay, spectrophotometric monitoring of the color change of a pH indicator in PME catalyzed reaction was performed. PME causes hydrolysis of ester bonds in the substrate, pectin, causing formation of acid groups which lowers the pH. The decrease in pH is then monitored by the indicator dye (methyl red) which changes color from yellow to red. This method is specific for pectin-degrading esterases because the natural substrate of PME is used.

Acidification due to PME activity turned the solution colour from yellow to red and the A517/A416 ratio increased with time to reach a plateau. The slope of the initial linear part of the curve was used to compare the PME activities of the protein from the different PPD time point extracts. PME activity increased gradually during early PPD (0-12h) and slightly reduced at 24h, thereafter staying high and stable throughout the late PPD time points (Fig 30).
The enzyme activity data was in agreement with the transcript data whereby an increase in transcript level was noted both in the early PPD time points, slightly decreasing at 24h PPD then remaining high throughout the late PPD time points. The PME protein abundance was shown to increase gradually from early to late PPD, although the increase was significant only in the late PPD time points between 48 and 96h PPD (refer to figure 18D).
3.4 Cassava transformation

The design of appropriate genetic engineering strategies for the modulation of PPD in cassava demands an increased knowledge and understanding of all the components central to the PPD process as well as their control points. Biochemical dissection of post-harvest deterioration in cassava storage roots and cloning of genes involved in these pathways is providing important new information as to the underlying processes (Reilly et al., 2004). However, the tools for designing genetic constructs to explore the nature of the PPD response and even to approach modulating this response in transgenic cassava have been lacking. Molecular research into cassava PPD has been very limited and it was only in the last couple of years that gene expression profiles through microarray studies have offered a better insight into events that take place during PPD.

Our study provided additional tangible evidence for the importance of reactive oxygen species (ROS) and enzymes and compounds that modulate them in PPD response. This led to the generation of a list of PPD-responsive proteins/genes both known and novel that can be tested through the direct manipulation of cassava’s genetic makeup. Several of the enzymes involved in ROS scavenging were found to be regulated in our study. These included enzymes like SOD, APX GST, peroxidases, GPX, DHAR and GR. Transgenic plants with altered activities of ROS regulating enzymes have been generated and have helped to shed light on the contribution of these enzymes to oxidative stress responses. It has been demonstrated that the over-expression of some of these genes individually or simultaneously can confer tolerance to oxidative stress in other species. For example, transgenic plants over-expressing GR have higher ascorbate contents and improved tolerance to oxidative stress (Foyer et al., 1991; Foyer et al., 1995) while reduced GR activity has been reported to increase stress sensitivity (Aono et al., 1995).

DHAR activity is enhanced in response to various environmental stresses (Urano et al., 2000) and DHAR over-expressing tobacco plants have been shown to increase tolerance to oxidative stress derived from various sources (Ahn et al., 1999; Kwon et al., 2000). Additionally, Roxas and colleagues showed enhanced seed germination and seedling growth under stressful condition by expressing plant glutathione-S-transferase/glutathione
peroxidase (GST/GPX) (Roxas et al., 1997, 2000). In this transgenic plant, increased glutathione-dependent peroxide scavenging and alterations in glutathione and ascorbate metabolism lead to reduced oxidative damage thereby conferring enhanced stress tolerance. These reports clearly indicate that alterations in the expression of enzymes involved in ROS-scavenging can have significant metabolic effects and provide the hope that this strategy can be used to develop crop plants with increased stress tolerance.

As reviewed in section 1.5.3.2, ethylene is a pivotal signaling molecule whose biosynthesis is induced by a variety of abiotic and biotic stresses. Since the increase in ethylene production during PPD has been experimentally demonstrated, we hypothesize that the modulation of aspects of its synthesis or perception, would lead to the production of cassava plants that are more robust than their non-transformed counterparts in the face of for example wounding stress which leads to PPD. Down-regulation of ethylene biosynthetic genes in cassava should therefore have a positive effect on PPD tolerance.

The ultimate aim for all plant genetic engineering programs is the development of an efficient genotype-independent transformation and tissue culture system that allow selection and regeneration of transgenic plants. However this remains elusive due to varying in vitro morphogenic response between even closely related cultivars. Cassava is notoriously recalcitrant to transformation and regeneration approaches and only after nearly a decade of unsuccessful attempts, four groups eventually produced the first transgenic cassava plants in 1996 using either embryogenic suspensions (Schöpke et al., 1996; Raemakers et al., 1996) or cotyledons from somatic embryos, from which plants were regenerated through organogenesis (Li et al., 1996) or through embryogenesis (Sarria et al., 1995). The embryogenic suspensions were transformed via particle bombardment, while the cotyledons were transformed via infection with Agrobacterium tumefaciens.

Several transformation systems have since been developed for cassava (Puonti-Kaerlas, 1998) and the technology has been applied to confer useful traits to cassava root (Taylor et al., 2001b; Zhang et al., 2003) including features related to PPD (Taylor et al., 2001a) however, the technology remains high risk, is restricted to only a few advanced cassava
research laboratories in the developed world and is yet to be transferred to the developing countries.

The use of gene transfer technology provides the ability to modify the levels of selected enzymes in plants and may help unravel the functional role of these enzymes in plant cells. Therefore, building on this information, efforts directed at for example enhancing the capacity for ROS scavenging as well as regulation of ethylene biosynthesis serves as logical starting points for PPD modulation.

Understanding the role of ROS in PPD was a collaborative venture between the University of Bath, UK and ETH, Zurich. Our ROS transformation studies focused on dehydroascorbate reductase isozyme 2 (DHAR2) and glutathione peroxidase isozyme 2 (GPX2). As stressed already, the cassava genome has not been sequenced hence we adopted the Arabidopsis plant as our model, by sourcing these genes from Arabidopsis. At the time of writing this thesis cassava was being transformed to express altered activities of some of the aforementioned genes. The cDNAs encoding Arabidopsis dehydroascorbate reductase, (DHAR2), and Arabidopsis glutathione peroxidase (GPX2) were isolated from Arabidopsis (Fig 31A and B) and cloned in appropriate transformation vectors.

The 1-amino cyclopropane-1-carboxylic acid oxidase enzyme (ACO2) is a key enzyme involved in the final stage of ethylene biosynthesis. The cassava ACCO2 cDNA sequence was isolated (Fig 31C), cloned in appropriate transformation vector and used as a target sequence for reduction of ethylene production in cassava roots. The antisense and sense ACO2 constructs were generated to test the impact of ethylene production and/or reduction on the PPD process.
Figure 31: PCR fragments amplified from Arabidopsis and cassava cDNA. A. *Arabidopsis* DHAR (641 bp), B. *Arabidopsis* GPX (503 bp) and C. cassava ACO2 (920 bp). Bands corresponding to 750 bp, 500 bp and 2 kb are indicated.

The gateway recombination system was used for construction of the vectors and all the genes were placed under the regulation of tuber-specific patatin promoter. Transformation vectors (sense and antisense) targeting the above ROS enzymes as well as ACO2 (ethylene biosynthesis) were constructed and used for cassava transformation.

As briefly pointed out in section 2.2.12, the cassava transformation system employed was reliant on production of embryogenic tissues from *in vitro* axillary bud meristematic tissue used as explants. Friable embryogenic calli (FECs) were generated from the somatic embryogenic structures. The FECs produced developed into mature somatic embryos, which then formed cotyledons from which *in vitro* shoots were regenerated. The entire process from generation of embryogenic tissues to emergence of shoots took approximately 8-9 months for a single FEC cell line batch.

A total of 12 different FEC lines were produced, and 11 transformation events attempted during the first two years of the study. However from all these transformations attempted, the putative transgenic plants obtained tested negative on rooting test, a clear illustration of its high risk as earlier mentioned. The rooting test is a simple initial screen for transgenic plants and is based on the fact that transgenic plants have a selective advantage over the
non transgenic plants when cultured on media containing the selection antibiotic. This is because the plants are transformed with constructs that also contain genes conferring resistance to the antibiotic used for selection (hygromycin in our case). From the rooting test experiments, we ascertained that all the putative transgenic plants generated were escapes.

After prolonged and pain-staking trouble shooting, modifications were done on the transformation protocol by Zhang and colleagues (Zhang 2000) and this led to the generation of new putative cassava transgenic plants. Patatin driven sense expression constructs for cassava ACO2 as well as *Arabidopsis* DHAR and GPX were used to transform cassava. Putative transgenic plants have been generated from these constructs (Fig 32 A-C)

![Figure 32: Putative cassava transgenic plants. Panel A, transformed with cassava ACO2, panel B, transformed with *Arabidopsis* GPX sense and panel C *Arabidopsis* DHAR sense](image)

At the time of writing this thesis, a total of 41 transformed shoots from two different FEC lines had been obtained from *Arabidopsis* GPX sense construct, 1 shoot from *Arabidopsis* DHAR construct and 1 shoot from cassava ACO2 constructs. More shoots were coming up from these transformations.
Figure 33: Transgenic cassava shoots culture on CBM media containing 10 μg/ml hygromycin. Panel A, shoots from *Arabidopsis* GPX sense transformation and panel B, shoot from wild type plant (TMS60444).

In a preliminary rooting test, 18 shoots obtained from *Arabidopsis* GPX sense construct were screened and 14 shoots were found to form roots (Fig 33 A) on the rooting media (containing hygromycin) compared to the wild type cassava which could not form roots on the rooting media (Fig 33 B). This preliminary test shows that the generated plants are transgenic, however further molecular screening experiments will have to be conducted to determine successful transgene integration and also to determine the copy numbers of integrated genes.
Chapter 4 General Discussion

In section 1.7, we provided the rationale behind this thesis. We argued that there is little understanding on the physiological and biochemical processes that occur during cassava PPD and that the process itself is a complex phenomenon. In order to shed more light on this problem and delve deeper into the complexity of the process, we set out several aims as outlined in section 1.7.1. with an overall goal of investigating the molecular events that underlie cassava PPD, identification of potential PPD markers to be used as targets for biotransformation and hence modulation of the process, or to be tested in breeding programs.

It is widely accepted that PPD is associated with protein dynamics. It is these changes in protein levels, their associated mRNA changes and the subsequent impact on the biochemical reactions they catalyze during PPD that have attracted a great deal of interest and was the primary focus of this work. Comparative proteomic analysis of commercial cassava roots undergoing PPD was adopted to identify PPD-responsive proteins or primary targets and possible pathways that are potentially involved in PPD development. The alterations in the proteome of cassava storage roots during early and late PPD was monitored by normal 2D gels, 2DE DIGE and iTRAQ approaches. In addition to the proteomic analysis, transcript activity of some of the identified gene products and enzyme activity assays were monitored. Physiological deterioration of the roots was monitored 0-96h after slicing of the commercially bought cassava storage roots. Deterioration symptoms were observed within 24h after wounding/slicing using the PPD induction procedure described in the section 2.1.2 of the study. The study identified novel proteins regulated during early and late PPD and also confirmed the regulation of some known PPD-associated proteins.

The first part of our work, the 2D proteomics study (section 3.1.1) was to identify proteins that are differentially regulated during early PPD development in cassava roots. Using the 2D-PAGE approach we compared differences in protein profiles between 0h and 12h after wounding. We observed that although the gel protein patterns between the two time points were somewhat similar, there were some differences in the proteins identified. Some protein spots were only detected in the 0h time point but were absent in the 12 hour PPD
time point. The following proteins; small heat shock proteins, alcohol dehydrogenase, malic enzyme, allene oxide cyclase, thioredoxin peroxidase and one hypothetical protein were detected in the 0 h PPD time point but were not present in the 12h PPD time point gels suggesting that qualitative changes in protein profiles may contribute to PPD development. This was not a peculiar observation as most of these proteins have been shown to be involved in many plant stress responses (Kato-Noguchi 2001).

The precise functional role of small heat shock proteins remains unclear, but some studies have shown small heat shock proteins to act as chaperones, where their accumulation has been correlated with stress tolerance, including tolerance to heat stress, desiccation and dehydration (Wang, Vinocur et al. 2004). They have also been shown to decrease the intracellular level of reactive oxygen species (Heckathorn 2002). As shown in Table 5 and 6, we identified isoforms of small heat shock proteins. Specific isoforms were unique to the 0h PPD time point but other isoforms were down regulated in the 12h time point for example the chloroplast isoform. Further characterization of cassava root proteome for changes associated with PPD development using both 2DE-DIGE and iTRAQ confirmed the down regulation of small heat shock proteins. Our protein datasets demonstrated the down-regulation of small heat shock proteins during PPD. These observations clearly suggest a distinctive but yet an unidentified role of the small heat shock proteins during cassava PPD.

As already emphasized throughout the thesis, wounding is one of the most severe plant stresses and is associated with dramatic changes in gene expression and protein synthesis (Carrera 1988; Dammann 1997). Wounding is also considered as a potential trigger for post-harvest deterioration which is an inherent problem in cassava since wounding and/or mechanical damage of cassava tuberous roots cannot be prevented during harvesting. Wounding stress has been shown to increase the activity of alcohol dehydrogenase (ADH) in maize and lettuce seedlings rapidly during the first 12h and more gradually thereafter (Kato-Noguchi 2001). In cassava, we found that ADH was one of the proteins that was only detected in the 0h and not found in the 12h after PPD. For this reason, we could not precisely establish the precise regulation pattern of ADH leaving behind the assumption that it is down-regulated. A similar observation was made with the cytosolic NADP malic enzyme. This enzyme has been reported to be induced upon wounding in bean (Phaseolus vulgaris)
(Schaaf 1995), possibly as a component of defense process. However, our results failed to confirm this in cassava. As mentioned above, the enzyme was only seen in the 0h and absent in the 12h PPD.

A well known process that occurs during PPD and which was the subject of our work is the role of ROS (see section 1.5.3.1 for detailed reaction mechanisms). To prevent oxidative damage brought about by biotic or abiotic stresses, or by products of aerobic metabolism, plants have evolved mechanisms to regulate the production of ROS. In the gel based proteomics study (2D gels and DIGE), we identified several oxidative stress related proteins whose expression changed significantly during PPD. One example of such protein is the APX. In plants APX is the key enzyme involved in hydrogen peroxide detoxification and hence protects plant cells from oxidative stress (Shigeoka, Ishikawa et al. 2002). Apart from the protein changes, the transcription of these genes has been shown to be up-regulated following wounding, blast pathogen infection and exposure to an array of signaling molecules including salicylic acid, ethylene, abscisic acid and hydrogen peroxide in rice (Agrawal 2003).

In both the gel based proteomics experiments ascorbate peroxidase (APX) isoforms were identified. APX 2 and APX 3 (later confirmed to be APX 1.2) were shown to be down-regulated 12h after PPD and likewise in the DIGE experiment, APX 2 was found to be down regulated in both 24 and 48h PPD time points. Interestingly, in the iTRAQ experiment the abundance of APX 1.2 was also shown to decrease after 24h and this was consistent with the observations from the gel based experiments. The down-regulation of APX has also been observed in tobacco and barley plants with different conditions of H$_2$O$_2$ stress (de Pinto et al., 2006; Fath et al., 2001). Although APX is an important H$_2$O$_2$-scavenging enzyme located in the cytosol, chloroplasts, mitochondria, and peroxisomes of higher plants, the decrease of APX abundance especially in the later PPD phase is likely to contribute to an accumulation of H$_2$O$_2$ and the eventual death of plant cells. This could be a potential trigger for PCD during cassava deterioration.

We employed the Mapman software to summarize the protein data from the DIGE experiment. We acknowledge that although the pathway analysis by Mapman could give
meaningful results, a disadvantage with this software is that, comparisons are made based on data heavily biased towards *Arabidopsis thaliana* as a model organism and as a consequence this pathway analysis may not mirror precisely what goes on in other species. For example, the involvement of a particular protein in cellular functions, molecular pathways and biological processes in non-model species such as cassava may differ from that in a model species. It is this biasness towards a specific model organism that presents a significant disadvantage for work with non-model species.

Although the 2D-gel based approach is a good starting point for understanding the root proteome responses of cassava to PPD, its application to identification of stress-induced low-abundant or rare proteins is limited. Additionally, the labor-intensive image analysis process requiring gel matching and manual removal of artifacts, the difficulty in electrophoretic separation of large and hydrophobic proteins in the first dimension (Lopez 2007) as well as under representation of extremely acidic and basic proteins technically limits its widespread application.

To overcome the obstacles highlighted above associated with gel-based proteomics; we used a much more superior technique, the iTRAQ (isobaric tagging for relative and absolute quantification). This technique allows the possibility of simultaneously identifying and quantifying multiple samples. Using this approach, it was possible to systematically screen a large number of proteins expressed during PPD in cassava storage roots. Proteins exist in either soluble or membrane-bound fractions therefore, in principle by including both soluble proteins and ‘membrane-bound’ proteins from the pellet obtained after extraction of the soluble proteins in the experiment, it implies that virtually all modulated proteins are captured. Prior to setting up of the PPD experiment, a preliminary experiment that compared protein profiles of cassava roots in two technical and two biological replicates was conducted in order to determine the regulation cut-offs that would be used for the iTRAQ experiment. Based on the data obtained from the preliminary experiment, a cut-off of 2-fold up or down regulation of the iTRAQ ratios was chosen and fixed for determination of regulated proteins.
The iTRAQ approach led to the identification of 1387 unique proteins from cassava storage root and a large number of membrane proteins as well as low abundance proteins such as transcription factors were identified. A significantly higher number of regulated proteins could be detected with the iTRAQ experiment compared to the 2D experiments. While iTRAQ labeling coupled with MALDI-TOF/TOF MS proved overall to be superior to the DIGE approach in sensitivity and quantitation, consistent detection of predicted protein sequences between the biological replicate samples from cassava root was limited. Consistency in trends in the ratios observed in the different PPD time points for those proteins that were detected in both the biological replicates was further limited. Inconsistent ratiometric data for some of the proteins detected in both biological replicates may represent differences in expression due, for example, to the use of different starting material (different root samples) and not technical variability. Despite this, the limited replicable detection of proteins among biological samples for technical reasons is common with liquid chromatography and MS-based proteomics (Thelen 2007) and may arise from variation in preparation of total proteins and/or iTRAQ labeled peptides from sample to sample.

In addition to the apparent variations arising from sample preparations, are constraints on detection imposed by the mass spectrometer, both with respect to the dynamic range of the instrument and the nature of selections of peptides in the first MS by the MS software for export to the collision cell prior to amino acid detection via the second MS. Data presented in this thesis from the iTRAQ experiment show that about half of the cassava root proteins were detected in both biological replicates suggesting the probability that most of the proteins detected in any given total protein sample using a shotgun approach are the highly abundant, housekeeping-type proteins.

Previous proteomics studies have characterized the cassava proteins expressed during somatic embryogenesis (Baba et al., 2008) and also compared the cassava protein profiles from fibrous and storage roots (Sheffield et al., 2006). By employing high resolution 2-DE, Sheffield and colleagues resolved proteins from cassava fibrous and tuberous roots and with this approach obtained 237 protein identifications (Sheffield 2006) while 83 proteins were identified by Baba and colleagues in cotyledons of cassava somatic embryos undergoing
secondary somatic embryogenesis (Baba et al., 2008). However both these studies presented less protein identifications compared to our iTRAQ approach in which 1387 proteins were identified. This is a great improvement in the already existing methods of cassava protein identification.

Despite the improved protein identification through iTRAQ, one major set-back is the lack of a complete annotated sequence of the cassava genome, thus most of the proteins were identified according to their homology with other proteins based on Blast searches. However, the genomics resources and functional genomics tools developed in Arabidopsis can be harnessed to address fundamental questions aimed at understanding cassava PPD. We therefore classified the identified regulated proteins according to their biological functions on the basis of their Gene Ontology (GO) annotation (Berardini, Mundodi et al. 2004). The gene ontology identity was determined by comparing the regulated proteins identified in cassava with their Arabidopsis thaliana whole genome annotation. As we have continuously pointed out, a number of important processes occur during PPD, amongst these are stress response due to wounding, oxygen response, signaling and the deposition of several phenolic compounds. Several proteins involved in the co-ordination and/or regulation of these processes were depicted in the GO categories that were found to be over-represented compared to all the proteins identified. The regulated protein dataset was found to be enriched with proteins that have functions related to translation, jasmonic acid biosynthesis, defense response, oxidative stress response, response to mechanical stimuli, stress response, response to wounding, and response to hydrogen peroxide.

It is widely known that a complex signaling network exists between plants and abiotic stresses such as wounding. One molecule known to participate in such cell signaling pathway is jasmonic acid (see section 1.5.3.3). Jasmonic acid forms an integral part of the general signal transduction system thus regulating inducible defense genes in plants. Together with its methyl ester derivative, methyl jasmonate, they are natural hormonal regulators that induce plant protease inhibitors in response to wounding and pathogen attack (Farmer and Ryan 1990). Strikingly, several proteins involved in jasmonic acid biosynthetic pathway were induced during PPD as demonstrated by the iTRAQ experiment. These proteins include; 12-oxophytodienoate reductase and lipoxygenase. Induction of the
expression of 12-oxophytodienoate reductase has been previously demonstrated in tomato in response to wounding (Parchmann 1997). On the other hand, we also found down-regulation of two other proteins that are known to play important roles in this biosynthetic pathway. These were; allene oxide cyclase and 4 coumarate CoA ligase family proteins. The JA derivative methyl jasmonate has been shown to initiate *de novo* transcription of genes, such as phenylalanine ammonia lyase, an enzyme known to be involved in the chemical defense mechanisms of plants (Farmer and Ryan 1990). Indeed phenylalanine ammonia lyase protein was found to be highly up-regulated during PPD. These data demonstrate the fundamental role of jasmonic acid and its derivative, methyl jasmonate in the intracellular signaling cascade that leads to the direct or indirect activation of genes involved in defensive secondary metabolism as was shown with the regulation pattern of phenylalanine ammonia lyase.

As displayed in section 1.5.3 of the thesis, during harvesting, and/or transportation, cassava is subjected to mechanical stress that may lead to injury of the root surface cell layers. This introduces wounds that cut though cells and leave previously intact cells of internal tissues exposed. These postharvest and processing operations are traumatic for the cells proximal to the damage site and induce a complex series of molecular events aimed at repairing the damage caused to the tissue (Surjadinata and Cisneros-Zevallos, 2003). Wounding in turn switches on numerous signal transduction pathways. Based on this inference, the displayed modulation of proteins associated with jasmonic acid biosynthesis during PPD could be in response to wounding. This observation is further supported by the over representation of proteins belonging to the GO categories corresponding to defense response, stress response and wounding response during PPD.

While there is no other comparable proteomics study done on cassava roots to identify PPD-responsive proteins, the microarray study by Reilly and colleagues (Reilly et al., 2007) to identify the genes expressed during cassava PPD offers a good comparative advantage. This study revealed the significant regulation of 72 genes during PPD. Of these, 63 genes were induced while 9 were down-regulated. In comparison to this study, our proteomics study revealed the induction of 125 proteins during PPD while 113 proteins were down-regulated. This too is a significant improvement in terms of the number of PPD-associated regulated
gene products identified. The classification of the regulated genes from the microarray experiment into different metabolic pathways allows us easy comparison with our regulated proteins which were categorized on the basis of their different biological functions. Several of the regulated proteins were common to both studies. These included; xyloglucan endotranseglycosylase (XET), APX, sulfite reductase, thioredoxin peroxidase, aquaporin, dehydrin, cysteine protease, ACO2, and small heat shock proteins. Although these proteins were identified in both studies, the results obtained in the microarray study were however difficult to compare with our proteomics study. This is because of the use of different experimental parameters (regulation cut-offs) and PPD induction set up (wounding, root storage temperature) and cultivars used. Commercially bought cassava roots from Costa rica were used in our iTRAQ study compared to the greenhouse grown roots (cultivar CM2177-2) used in the microarray experiment. Although there were differences in the regulation patterns of the genes and proteins commonly identified in two approaches, the metabolic processes in which the proteins or genes are involved are similar. The main metabolic systems commonly identified by both the microarray and proteomics studies were the antioxidant response mechanism, stress-defense mechanisms, other cellular metabolism, protein metabolism, and response to biotic/abiotic stress stimuli.

Using the iTRAQ approach, we also identified several unique oxidative stress response proteins whose expression has been shown to increase during PPD (Reilly, Bernal et al. 2007). These included proteins such as APX, Glutathione-S-transferase (GST) and secretory peroxidase. GST, Cu-Zn superoxide dismutase (CSD1.2) and secretory peroxidase were up regulated during PPD. APX 1.2 was shown to be stable during early PPD and decreased considerably in the late PPD time points. In most plant species, cytosolic forms of APX are the most responsive to environmental stresses and the up-regulation of ascorbate peroxidase (APX) genes has been shown to accompany wounding, blast pathogen infection, exposure to diverse signaling substances including ethylene, abscisic acid and hydrogen peroxide in rice (Agrawal 2003). Contrary to these observations, the protein abundance of APX 1.2 was shown to increase gradually during early PPD and decreased considerably in the late PPD time points. This increase was probably insufficient in magnitude to counteract the oxidative burst that develops rapidly upon wounding (Reilly et al., 2004). Several other proteins previously implicated in the response of plants to oxidative stress were also found.
in our regulated protein dataset. These included mitochondrial glyceraldehyde-3-phosphate dehydrogenase (Sweetlove, Heazlewood et al. 2002), cytosolic isoflavone reductase (Babiychuk, Kushnir et al. 1995), and allene oxide cyclase (Stenzel, Hause et al. 2003).

Our proteomics data also agrees with and confirms data from directed and random sequencing of cDNA clones corresponding to genes that are expressed during PPD (Huang et al., 2001). Interestingly, several proteins with functional roles in signaling (e.g. calmodulin, ACC oxidases) or cell wall strengthening (e.g. cinnamoyl coenzyme A reductase and peroxidase), programmed cell death (e.g. 26S proteosome), antimicrobial activity (e.g. proteases and β-1,3-glucanases) and biosynthesis of secondary metabolites (e.g. phenylalanine ammonia lyase) were shown to be regulated during PPD by both studies. Other regulated proteins identified were also found to have no known function. This could be due to non existence of similar proteins in other plant species or perhaps because these proteins possess special functional motifs not found in other plants.

Experimental evidence exists that displays the interplay of multiple pathways including ROS regulation, signal transduction, cell wall metabolism, secondary metabolite biosynthesis and protein biogenesis during PPD development (Reilly et al., 2004). However, a mechanism that depicts a coherent sequence of events of the above pathways culminating in PPD is still lacking. Our protein data suggest that post-harvest deterioration of cassava roots induces an oxidative challenge that is efficiently taken up by the proteomic and metabolic responses of the roots. The oxidation of scopoletin by peroxidases is known to give rise to a bluish product of unknown structure similar to blue/black vascular streaking observed during PPD (Wheatley and Schwabe, 1985), this may suggest that peroxidase-mediated oxidation of scopoletin accompanies PPD. Based on this information as well as other reports, a tentative model for some of the processes that occur in the cassava storage root during PPD has been proposed (Reilly et al., 2004). This model speculates that the wound-induced oxidative burst in injured cassava storage roots lead to enhanced production of ROS which act as signals or which cause membrane breakdown thereby producing elicitors that alter the down-stream gene expression. The overall impact is observed in the altered expression of PCD-related genes, antioxidant regulatory genes and defense response genes. Proteins which form important components of this model were identified in our iTRAQ study.
As pertaining to PCD, cysteine protease inhibitor which is a modulator of PCD in plants was down-regulated, which may suggest that protease activity in the cells was enhanced thereby leading to PCD. Some of the antioxidant regulatory genes including DHAR, and thioredoxin peroxidase were also shown to be down-regulated during PPD suggesting a compromised ROS scavenging capacity within the roots. An up-regulation of proteins related to defense response and wound healing particularly phenylalanine ammonia lyase, beta 1,3 glucanase, was observed. The iTRAQ results thus further confirms the interplay of the different pathways in PPD development and supports the proposed model for molecular events that accompany PPD. Other important pathways which would fit into the model include signaling pathway and cell wall metabolism. Proteins involved in signaling including calmodulin, aquaporins, AC02 were confirmed to be regulated.

A large subset of cell wall modifying proteins was shown to be up regulated during PPD. This included proteins such as linamarase, PME, XET, hydroxycinnamoyl transferase and cellulose synthase. Our observations thus confirm and extend previous findings regarding the potential pathways implicated in PPD (Beeching, 2001; Han et al., 2001; Huang et al., 2001; Reilly et al., 2001; Reilly et al., 2004) and are in agreement with the proposed model (Reilly et al., 2004). These findings will be informative in elucidating how the target proteins may be modulated during PPD.

Among all identified proteins, CSD, APX, PME and beta 1,3 glucanase were selected in order to investigate their expression patterns at the mRNA level, and enzyme activity during PPD. These proteins were selected for evaluation of their gene expression profiles so as to determine if the protein expression patterns observed corresponded to the transcript activity. These results would confirm the functional role of the identified proteins during PPD.

All the above proteins exhibited similar expression patterns between their mRNA and protein levels during both early and late PPD however, the mRNA levels of APX and beta 1,3 glucanase did not match their protein levels at late PPD. The result of the present investigation is supported by the previously established concept that transcription patterns
are not directly concomitant with protein expression levels (Gygi et al., 1999). Discordance between mRNA levels and protein abundance is likely because of the problems relating to the existence of multigene families, or due to differential turnover rates as well as different stabilities of RNA and protein (Bae et al., 2003) especially when dynamic processes such as stress response are analyzed.

The measurement of the transcript activity of APX 1.1 and APX 1.2 revealed an induction during the early PPD time points followed by decreasing levels after 24h. APX1 transcript expression has been shown to be induced mainly due to combined heat and drought stress (Rizhsky, Liang et al. 2004). However, the APX mRNA induction may also be found in many plant tissues even in the absence of any stress (Pnueli, Liang et al. 2003). The activity of APX2 is inducible mainly by extreme light, heat stress or wounding (Panchuk, Volkov et al. 2002; Chang, Ball et al. 2004; Mullineaux, Karpinski et al. 2006).

Since changes in gene expression do not always reflect changes in the protein abundance, the activity of APX was measured in the protein extracts used for the ITRAQ experiments. The total activity of APX was observed to increase throughout PPD time points although the transcript levels only increased during early PPD. This observation suggests that perhaps more than one isozyme of APX is activated during PPD development in cassava. A similar observation was reported by Lopez (1996) in which protein activity and not the mRNA expression was shown to accompany the salt-stress-induced increase in ascorbate peroxidase activity in radish that also was not paralleled by differences in mRNA levels. Further analysis of APX isoforms by native gel electrophoresis would provide evidence for the contribution of each cellular APX isoform to the observed increase in total activity. Many research groups have reported that the expression of the APX gene, APX protein, and APX activity can be altered by many biotic and abiotic stresses (Mittler and Zilinskas 1994; D’Arcy-Lameta, Ferrari-Iliou et al. 2006).

Cytosolic Cu/Zn-SOD (CSD) is considered to be a general stress response enzyme (Bowler et al., 1992). In the study presented in this thesis both CSD 1 and CSD 2 were found to be up-regulated (transcript level) during early PPD, however the expression level of CSD 3 remained relatively stable. Similarly, the protein abundance of CSD 1.2 was significantly high
especially during early PPD and total activity of CSDs in general was found to increase from early to late PPD. These results show that members of the CSD family are differentially regulated depending on the stress being encountered. CSD 1 and CSD 2 in particular are up-regulated in response to stress that generates ROS, including O$_3$, UV-B and high light levels (Kliebenstein, Monde et al. 1998).

The alignment of Cu/Zn SOD amino acid sequences indicate that a high homology of CSDs exist among plant species with some features that can differentiate chloroplastic from cytosolic Cu/Zn SODs. The intron number and position in the CSD genes are highly conserved in plants (Fink and Scandalios 2002) and this can allow comparison of CSD members from different plant species especially plants with unsequenced genomes. This allows the prediction of possible intron regions for example in the case of cassava CSDs and is of fundamental importance in designing primers for studying CSD gene expression patterns during different plant stresses.

Plant cell walls constitute a physical barrier between the environment and the internal contents of the cells. Modifications on the cell wall structure are often associated with plant defense responses (Vorwerk, Somerville et al. 2004). The changes that accompany cell development may involve changes in the level of expression of genes coding for proteins (cell wall-related enzymes) required for cell wall loosening and cellulose synthesis. These enzymes have a major role in cell wall loosening leading to generation of cell wall fragments such as oligogalacturonides (obtained from the cell wall pectin degradation) and or oligosaccharides (products of cellulose degradation). The cell wall fragments, collectively referred to as elicitors, could act as signaling compounds through the induction of jasmonic acid biosynthesis. Several candidate proteins that have been implicated in cell wall loosening, including β-glucanases (Thomas, Inouhe et al. 2000) and xyloglucan endotransglycosylases (Fry 1995), were identified in the membrane protein fractions and most of these were shown to be up-regulated during PPD.

One of the key cell wall enzymes we studied is PME. Increase in transcript as well as enzyme activity of pectin methyl esterase was observed to accompany PPD with the induction becoming more pronounced during the late PPD time points. In comparison to transcript
and enzyme activity, increase in protein abundance became significant during late PPD time points. The data clearly shows an increase in the activity of PME during cassava PPD. Similar observations have been reported whereby several PME transcripts have been shown to be regulated by wounding and ethylene (De Paepe, Vuylsteke et al. 2004). Although there has been no previous study on cell wall-degrading enzymes and their role in PPD advancement, we speculate that the loss of membrane integrity and leakage between cellular compartments that accompanies PPD could result in the activation of PME.

Increased activities of PMEs results in decrease in the degree of methylesterification of the pectins in the cell wall and middle lamella and allows more calcium cross-linking between calcium molecules, increasing firmness for example in fruits (Pilnik 1991). PME activity can also contribute to cell wall loosening and degradation by promoting the activity of cell wall hydrolases such as endopolygalacturonases (Micheli 2001; Willats, McCartney et al. 2001). Thus, PMEs can regulate the texture and mechanical properties of plant cell walls in opposite ways. PMEs catalyse the C-6 demethylation of homogalacturonan within plant cell walls. In this process, methanol (MeOH) is released and negatively charged carboxyl groups are created which can associate with other homogalacturonan chains by calcium cross-links, making the cell wall more rigid. However, there is no previous study implicating PME activation with induction of signal cascades that would bring about stress response in plants or eventually lead to wound healing.

Another key cell wall enzyme we studied is β-1,3-Glucanases. β-1,3-Glucanases belong to a large gene family that has been subdivided into at least three categories depending on their primary structure (Neale, Wahleithner et al. 1990). They have been well characterized in herbaceous plants both at the molecular and physiological level due to their divergent role in mediating defense responses following pathogen infection, wounding, ethylene treatment as well as chemical stress (Simmons, Litts et al. 1992; Krishnaveni 1999)

The induction of the activity of β-1,3-glucanase as observed in cassava during PPD is mirrored in several previous reports depicting the induction of pathogenesis-related (PR) proteins including β-1,3-glucanase by abiotic stresses such as low temperature (Hon, Griffith et al. 1995). In winter rye, β-1,3-glucanase was shown to be induced in response to low
temperature (Hon, Griffith et al. 1995), and in tomato the induction of class III β-1,3-glucanase was implicated with postharvest chilling-induced cell damage (Sanchez-Ballesta 2006).

The proteomics, protein activity and transcript analysis tools explored largely in this research study to understand the molecular basis of PPD is expected to bridge the gap between acquiring new knowledge and implementing that knowledge in designing transgenic strategies to mitigate the rapid degeneration of cassava roots, which presently restricts full development of the crop. Following our detailed proteomic analysis of cassava during PPD, there appears to be an urgent need to sustain the content of the protein levels as well as diminish the deterioration processes following wounding. This process can be achieved through genetic manipulation as attempted in our transformation study.

Although cassava is an important food security crop in developing countries, its transformation and regeneration presents a great challenge to the research community. The lack of regeneration methods for transgenic cassava plants continues to hinder the use of transformation technology in cassava improvement. The lengthy transformation and regeneration process compels the need for a straightforward protocol that can be used across the research community.

As already captured in section 3.4 of the thesis, during the period of this study several attempts were made to transform cassava using published protocols, however all attempts proved futile. Typically, the friable embryogenic calli (FEC) used in cassava transformation are generated during 3-4 months of sterile culture of axillary bud meristematic tissue, which give rise to somatic embryos from which FECs emerge. The amount of time required to generate transgenic cassava plants from the axillary bud meristematic tissues, through somatic embryos and FECs to in vitro plantlets is approximately 8-9 months.

Regeneration of transgenic plants from the transformed tissues (FECs) takes approximately 4-5 months and involves a series of sub-culturing steps in appropriate regeneration media. In our case, successful regeneration could not be achieved in most of the transformation events attempted, and since there is usually no way of ascertaining whether the
transformed cells will eventually regenerate into transgenic plants, this can be an extremely frustrating exercise when the cells fail to form plants.

The transformation protocol by Zhang and colleagues (Zhang 2000) employed in our transformation experiments includes a liquid media culture step prior to and after FEC Agro infection (transformation). The cells have to be replenished with a new supply of fresh media every two days for 4 weeks. This prolonged culture of FECs in the liquid culture step contributes to contamination of the transformed cells therefore further reducing the regeneration efficiency of the FECs as an enormous amount material is lost in the process. Additionally, we observed that during the liquid culture steps after transformation of FECs, their morphology changes from clusters of organized cells to single calli round shaped structures. We speculate that the change in FEC morphology is brought about by the liquid culture step and this reduces their sensitivity to Agro-infection and subsequently prevents their successful regeneration into transgenic plants.

In cases where we obtained putative transgenic plants, we found that all the plants regenerated were escapes. A simple initial positive screen for transgenic plants is usually done when transgenic shoots are obtained from the transformed cells. The fact that this screening can only been done on putative transgenic plants developed over a period of 9 months further attest to the high risk nature of cassava transformation technology.

Due to the setbacks mentioned above, there was need to assess all the transformation and regeneration protocols used previously to transform cassava. One of the key areas considered for improvement was the antibiotic concentrations used for selection, and for eliminating Agrobacterium after transformation. As opposed to using a single dose of antibiotic concentration throughout the regeneration process, a gradual increase in the concentration of the selection antibiotic used during regeneration of transformed FECs was found to improve the regeneration capacity of the transformed cells. It became increasingly clear that the transformation and regeneration capacity of the FEC lines used in transformation varied greatly from batch to batch. Some FEC lines took relatively long to regenerate or failed to regenerate completely when cultured on the appropriate regeneration media, suggesting that the transformation and regeneration efficiency
depends on the quality of the FECs. A monthly induction of new FECs lines was initiated to ensure a continuous supply of high quality FECs. Failure of the FECs to regenerate into transgenic plants was also partially attributed to the prolonged culture of transformed FECs on a liquid media. As we have mentioned before, this media was found to induce morphological changes in the FECs structure which probably hindered differentiation and hence development of the transformed cells. We presuppose that the exclusion of this liquid step from cassava transformation will greatly improve the regeneration and transformation efficiency of cassava, and most importantly cutback the amount of time invested on generating transgenic plants.

By critically assessing the above described key areas and incorporating the highlighted changes into the protocol by Zhang and others (Zhang 2000) we developed a more robust protocol for cassava transformation. We exploited this new method and successfully transformed cassava with the *Arabidopsis* GPX, *Arabidopsis* DHAR and cassava ACO2 constructs.

The protocol developed for cassava transformation is easy to handle and clusters of FECs can be simultaneously transformed within a reasonable time. The development of somatic embryos from the transformed FECs is relatively fast. Most importantly, with the improved protocol, the transfer of the transformation technology to other laboratories should be relatively easy as it does not depend on the availability of technically advanced or expensive equipment.
Conclusion

PPD is responsible for the loss of root quality in cassava resulting in reduced consumption and exploitation for commercial purposes. This poses a significant risk to a majority of the population in developing countries that rely on cassava for food. The association between PPD and protein changes is absolute but there is little progress made in the tools for protein identification.

One key hallmark of the thesis was the development of the iTRAQ tool to analyze cassava proteins. This is the first global quantitative proteomics profiling experiment carried out in the cassava root during PPD. Although the current depth of coverage for cassava proteome is still significantly less than in other plants, this approach permitted the generation of a more comprehensive map of proteins within the cassava root. The technique also captured some hitherto unidentified proteins regulated during PPD like pectin methyl esterase, linamarase, hydroxycinnamoyl transferase, glutamic acid rich protein and also confirmed others such as phenylalanine ammonia lyase, ascorbate peroxidase, glutathione S transferase, beta-1,3-glucanase and peroxidase.

This is an important milestone in cassava research as the increased repertoire of proteins will accelerate the generation of a cassava root proteome. This information will further lead to a better understanding of the protein function and metabolic pathways involved in PPD. Utilizing this acquired information, we performed a series of transcriptional analysis and enzymatic assays and this led to the identification of particular proteins that can be applied as potential PPD-molecular markers in cassava.

We also found a selection of these genes to be under tight regulatory patterns and this could be as a result of polymorphisms on the promoter regions of the genes. This new information is important as the isolation and analysis of their promoter regions can ascertain whether they are wound/PPD inducible.
With the improved identification of proteins, confirmation of their expression profiles and a better understanding of their reaction processes during PPD, in addition to the development of an improved transformation system, we now have a platform for the molecular cloning of corresponding genes. This forms the basis for effective engineering strategies to improve cassava and help fight food insecurity.
References


Proceedings of the validation forum of the global cassava development strategy. Rome, 26 - 28 April 2000


Thimm, O., Blaesing, O., Gibon Y et al., (2004). "MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes."


