

Transcriptome analysis of cassava (*Manihot esculenta* Crantz) in response to mealybug (*Phenacoccus manihoti*) infestation



Molemi Evelyn Rauwane

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, in fulfilment of the requirements for the degree of Doctor of Philosophy in the School of Molecular and Cell Biology.

Johannesburg, 2018

DECLARATION

I, Molemi Evelyn Rauwane (0507804y), am a student registered for the degree of Doctor of Philosophy (MCBG 9002) in the academic year 2018.

I hereby declare the following:

- I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the work submitted for assessment for the above degree is my own unaided work except where explicitly indicated otherwise and acknowledged.
- I have not submitted this work before for any other degree or examination at this or any other University.
- The information used in the Thesis HAS NOT been obtained by me while employed by, or working under the aegis of, any person or organisation other than the University of the Witwatersrand and the Agricultural Research Council.
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.

Signature  28th day of August 2018

ABSTRACT

Cassava mealybug (*Phenacoccus manihoti*) is one of the most damaging pests of cassava (*Manihot esculenta* Crantz) globally. While biological control of mealybugs through the use of natural predators has been successfully practiced, breeding for resistant cultivars remains an important means of control. Understanding plant responses to insect herbivory, by determining and identifying differentially expressed genes (DEGs), is a vital step towards the understanding of molecular mechanisms of defence responses in plants and the development of resistant cultivars by gene editing.

Mealybug isolates were collected from different cassava growing regions of South Africa and a detailed morphological and molecular characterization of the pests was undertaken. Morphological and molecular analysis confirmed the mealybug identity as *Phenacoccus manihoti* (Matile-Ferrero). The positively identified mealybug isolates were used to artificially infest three cassava genotypes with contrasting response to mealybug. The response of cassava genotypes AR23.1 (which has resistance to multiple diseases and to green mites), P40/1 (which has no known resistances) and 98/0581 (which has multiple pest resistances and resistance to cassava mosaic disease) was further investigated by analysing their transcriptome after mealybug infestation at 24 and 72 hours in comparison with non-infested mock tissues. A total of 301, 206 and 269 transcripts from AR23.1, P40/1 and 98/0581 were differentially expressed ($\log_2$ fold and $P \leq 0.05$) per time point following mealybug infestation. In two genotypes (AR23.1 and 98/0581), mealybug infestation resulted in a higher number of down-regulated than up-regulated genes between mealybug-infested and mock-infested tissues, while in P40/1 between infested leaves at 24 hpi and 72 hpi, more genes were up-regulated than were down-regulated. There was an increase in the number of DEGs from 24 hpi to 72 hpi suggesting an early induction of defence responses during the mealybug infestation.

Gene expression was further compared between AR23.1, P40/1 and 98/0581 to determine whether there were any genotype-specific expression patterns. Within the Gene Ontology (GO) functional classification, DEGs in the class Secondary Metabolic Processes were significantly enriched in AR23.1 in comparison with P40/1 and 98/0581. Additionally, DEGs in the class Regulation of Molecular Function were significantly enriched in P40/1, while DEGs in the classes Reproduction and Reproductive Processes, and Nutrient Reservoir, were enriched in the multi-pest resistant 98/0581 genotype compared to P40/1 and AR23.1. In the metabolic pathway analysis performed using the KEGG database, DEGs in the classes ABC Transporters and MAPK signalling pathways were present in AR23.1 and 98/0581; and not in P40/1. There were both up- and down-regulated genes that were mapped to the classes Plant-Pathogen Interactions, and Plant Hormone Signal Transduction pathways in genotypes AR23.1 and 98/0581; while in P40/1, there were only down-regulated genes classified under Plant-Pathogen Interactions pathway. The up-regulated DEGs

associated with these pathways are suggested to be involved in defence in response to *P. manihoti* feeding since these stress-associated genes, including those related to transcription factors, phytohormones and secondary metabolism were significantly induced in the AR23.1 (resistant) and 98/0581 (tolerant) genotypes, and not in the P40/1 (susceptible) genotype. Stress-associated genes such as 2-oxoglutarate / Fe (II)-dependent oxygenase superfamily protein, terpene synthase 21, heat shock family proteins and cytochrome P450 superfamily proteins were induced in AR23.1 and 98/0581 genotypes.

The results revealed a significantly different response to mealybug infestation in the three genotypes studied, with genotypes AR23.1 and 98/0581 showing a higher proportion of differentially expressed transcripts compared with the susceptible genotype P40/1 during infestation with mealybugs. Putative candidate defence-related genes that were overexpressed in the AR23.1 genotype post-infestation will be useful in future functional studies towards the control of mealybugs. These results will further form the basis for more detailed future studies on the specific role of all the differentially expressed transcripts identified. The transcripts identified in this study will also contribute significantly to the cassava EST database, and can be used to improve the annotation of the cassava genome.

RESEARCH OUTPUTS

RESEARCH ARTICLES

Molemi E. Rauwane, Damaris A. Odeny, Ian Millar, Chrissie Rey and Jasper Rees (2018). The early transcriptome analysis of cassava (*Manihot esculenta* Crantz) in response to cassava mealybugs (*Phenacoccus manihoti*) feeding. **Accepted (PLoS One).**

CONFERENCE OUTPUTS

Molemi E. Rauwane, Damaris A. Odeny, Chrissie Rey and Jasper Rees (2017). Variations in transcription profiles induced by herbivore feeding in cassava (*Manihot esculenta* Crantz). International Congress of Plant Pathology (ICPP), Boston, Massachusetts, USA, 29 July-03 August 2018.

Molemi E. Rauwane, Tinashe G. Chabikwa, Damaris A. Odeny and Jasper Rees (2017). Socio-economics of cassava production in the Mpumalanga province of South Africa. The 3rd International Conference on Global Food Security, Cape Town, South Africa, 03-06 December 2017.

Molemi E. Rauwane, Damaris A. Odeny, Chrissie Rey and Jasper Rees (2017). RNA-seq revealed differences in transcriptomes of cassava genotypes following herbivore infestation. The Fourth International Horticulture Research Conference (HORTRES2017), East Malling, UK, 16-20 July 2017.

Molemi E. Rauwane, Damaris A. Odeny, Chrissie Rey and Jasper Rees (2017). Cassava (*Manihot esculenta* Crantz) transcriptional responses as a result of mealybug (*Phenacoccus manihoti*) feeding. The Plant and Animal Genome Congress (PAGXXV), San Diego, USA, 14-18 January 2017.

Molemi E. Rauwane, Damaris A. Odeny, Chrissie Rey and Jasper Rees (2016). Time-series network of differentially expressed genes in cassava (*Manihot esculenta* Crantz) in response to mealybug (*Phenacoccus manihoti*) infestation. The 4th Joint congress of the South African Society of Bioinformatics and South African Genetics Society (SASBi-SAGS), Durban, South Africa, 20-23 September 2016.

M.E. Rauwane, D.A Odeny, M.E.C Rey and D.J.G Rees (2016). The early transcriptome analysis of cassava (*Manihot esculenta* Crantz) in response to cassava mealybugs (*Phenacoccus manihoti*) feeding. The World Congress on Root and Tuber Crops (WCRTC), Nanning, China, 18-23 January 2016.

M.E. Rauwane, N.M Vilane, D.A Odeny and T.G. Chabikwa (2013). Potential for commercialization of cassava production in the Mpumalanga province of South Africa. The 12th International Symposium of the International Society for Tropical Root and Tuber Crops-Africa Branch (ISTRC-AB), Accra, Ghana, 30 Sep - 05 Oct 2013.

M.E. Rauwane, T.G. Chabikwa, M.E.C Rey and D.A. Odeny (2012). Cassava (*Manihot esculenta* Crantz) transcriptome analysis post-infestation by cassava mealybugs (*Phenacoccus manihoti*). Global Cassava Partnership-21st century and Second Scientific Cassava Conference, Kampala, Uganda, 18-22 June 2012.

Other Research Articles

Molemi E. Rauwane, Laurette M. Marais, Oluwatosin G. Oladipo, Cornelius C. Bezuidenhout and Mark S. Maboeta (2017). Spatio-seasonal dynamics of bacteria and metal composition of platinum mine tailings dam, Rustenburg, South Africa. *South African Journal of Plant and Soil*, doi.org/10.1080/02571862.2017.1345017

ACKNOWLEDGEMENTS

I am greatly thankful to a number of people that have been supportive during the research and writing of this thesis. My supervisor and co-supervisor, Prof. Chrissie Rey and Dr. Jasper Rees for guidance and support during the study, financial support throughout the study, providing information and for carefully correcting and preparation of this thesis. Thanks for keeping me motivated; this study has equipped me with special and fascinating experience in molecular biology and bioinformatics. My mentor, Dr. Damaris Odeny, for her supervision, outstanding involvement in the planning and running of the project, advice during the experimental phase of study, analysis of results, training in bioinformatics and preparation of the thesis. You have been a great sister and friend. My colleagues, Mr. Tinashe Chabikwa, who worked with me throughout the project and Dr. Zedias Chikwambi for providing advice whenever there were problems. Dr. Charles Hefer and Dr. Keoagile Bezuidt for assistance and advice with bioinformatics analysis. My colleagues and friends in the Biotechnology Platform group for their interest in this study and their support. My sincere appreciation to the Agricultural Research Council-Professional Development Programme and Biotechnology Platform for their financial support, and training throughout the study. In addition, a special thank you to the BTP big lab candidates, especially Thuto, thank you for your outstanding support.

I would like to extend my special thanks to Prof. Felix Nweke and the CATISA (Cassava transformation in Southern Africa) group for their interest in this project and for their acceptance of me as part of them. A thank you to small-scale farmers in Mpumalanga and Mr. Jurie Steyn from the research station in Makhathini, KwaZulu Natal for their plant material and mealybug species. Mr. Ian Millar from the Plant Protection Research Institute-Biosystematics Department in Roodeplaat for his help in morphological identification of mealybugs, your expertise is really appreciated. Faculty of Sciences, School of Molecular and Cell Biology for making this study possible. Thanks to the Rauwane, Masilo and Phokungwana families for their great support, motivation, encouragement, unconditional love and financial support throughout this study, especially my sister Kagiso Rauwane, I appreciate all that you have done for me. Moreover, a thank you to my extended brother and sister, Edward Lekota and Constance Matubako.

Most of all and important, I would like to extend my unreserved gratitude to my Lord the Saviour, who chose me from nothing and showered me with His mercies to accomplish this study. In conclusion, allow me to say in Tswana ***“KEA LEBOGA”***

I can do all things through Christ who gives me strength.

Philippians 4:13

DEDICATION

In loving memory of my beloved mother:

Mathotho Pauline Rauwane

1945-2009

You were always my inspiration and may your soul rest in peace.

TABLE OF CONTENTS

DECLARATION.....	i
ABSTRACT	ii
RESEARCH OUTPUTS	iv
ACKNOWLEDGEMENTS	vi
DEDICATION	vii
TABLE OF CONTENTS	viii
LIST OF FIGURES	xi
LIST OF TABLES	xiii
LIST OF APPENDICES	xiv
LIST OF ABBREVIATIONS	xvii
CHAPTER 1	1
INTRODUCTION AND LITERATURE REVIEW	1
1.1 Cassava origin and domestication	1
1.2 Cassava characteristics	2
1.3 Cassava production and importance	3
1.3.1. Contribution to Food Security.....	4
1.3.2. Industrial starch	4
1.3.3. Bioethanol production	5
1.3.4. Other important uses of cassava.....	6
1.4 Factors affecting cassava production and utilization	6
1.4.1 Stem propagation methods.....	6
1.4.2 Diseases and arthropod pests.....	7
1.4.2.1 Cassava mosaic disease (CMD).....	7
1.4.2.2 Cassava bacterial blight disease (CBB)	8
1.4.2.3 Cassava brown streak disease (CBSD).....	8
1.4.2.4 Cassava green mites (CGM)	8
1.4.2.5 Cassava Mealybug (CMB)	9
1.5 Identification of cassava mealybugs	11
1.5.1 Morphological identification.....	11
1.5.2 Molecular identification.....	12
1.5.2.1 DNA barcoding	12
1.6 Plant-insect interactions	12
1.6.1 Direct defences.....	13
1.6.1.1 Jasmonic acid (JA)	14
1.6.1.2 Ethylene (ET)	15
1.6.1.3 Salicylic acid (SA).....	16
1.6.1.4 JA, ET and SA interactions	16
1.6.2 Indirect defences	17
1.7 Plant-herbivore resistance.....	17
1.8 Improving cassava germplasm for mealybug resistance	18
1.9 Next generation sequencing (NGS)	20
1.9.1 Roche/454 sequencing technology.....	20
1.9.2 ABI SOLiD sequencing technology.....	21

1.9.3 Illumina sequencing technology	21
1.9.5 Application of NGS in cassava breeding	23
1.9.6 Transcriptome profiling	23
2.0 MOTIVATION AND RATIONALE FOR STUDY	24
2.1 AIM	25
2.2 HYPOTHESES.....	25
2.3 SPECIFIC OBJECTIVES	25
CHAPTER 2	26
MATERIALS AND METHODS.....	26
2.1 Mealybug collection and maintenance.....	26
2.1.1 Morphological identification of mealybugs	26
2.1.2. Confirmation of mealybug identity	27
2.2 Cassava collection, multiplication and hardening	28
2.3 Mealybug infestation experiments and scoring.....	28
2.4 Total RNA isolation.....	30
2.4.1 Treatment of equipment and glassware for RNA isolation.....	30
2.4.2 RNA isolation and quantification	30
2.5. Library preparation and RNA sequencing	30
2.5.1 Methodology workflow	31
2.6 Bioinformatics analysis	32
2.6.1 Read mapping and gene expression quantification	32
2.6.2 Pathway analysis.....	33
2.7 RT-qPCR validation of DEGs	33
CHAPTER 3	35
Identification of mealybug collections from cassava in KwaZulu Natal and Mpumalanga provinces of South Africa.....	35
3.1 Introduction.....	35
3.2 Results.....	36
3.2.1 Morphological identification of mealybugs	36
3.2.2 Molecular identification of mealybugs	37
3.2.2.1 DNA isolation and PCR analysis.....	37
3.2.2.2 DNA characterization of <i>P. manihoti</i> species.....	38
3.2.2.3 Phylogenetic analysis.....	38
3.3 Discussion.....	42
CHAPTER 4	43
Transcriptomics of cassava (<i>Manihot esculenta</i> Crantz) following mealybug (<i>Phenacoccus manihoti</i>) infestation.....	43
4.1 Introduction.....	43
4.2 Results.....	44
4.2.1 Cassava phenotypic responses to mealybug infestation	44
4.2.2 Total RNA integrity of cassava leaf samples	46
4.2.3 Reads quality check and trimming	47
4.2.4 Gene expression of cassava genotypes in response to mealybug infestation	48
4.2.5 Transcriptional changes in AR23.1 in response to mealybug infestation	50
4.2.6 Transcriptional changes in P40/1 in response to mealybug infestation	55
4.2.7 Transcriptional changes in 98/0581 in response to mealybug infestation	58
4.2.8 Comparison of expression patterns between AR23.1, P40/1 and 98/0581	61
4.2.8 Network interaction of AR23.1, P40/1 and 98/0581 proteins	62
4.2.9 Validation of expressed transcripts by RT-qPCR.....	66
4.3 Discussion.....	67

CHAPTER 5	71
GENERAL DISCUSSION AND CONCLUSION	71
5.1 Introduction	71
5.2 Morphological and molecular analysis of KwaZulu Natal and Mpumalanga mealybugs in South Africa	71
5.3 Analysis of phenotypic responses of cassava to mealybug infestation	72
5.4 Potential of transcript profiling for the identification of candidate defence-related genes in response to mealybugs (<i>P. manihoti</i>).....	73
5.5 Signalling pathways involved in plant-mealybug interactions.....	75
REFERENCES	78
APPENDIX.....	99

LIST OF FIGURES

Figure 1.1: A cassava plant and roots grown and harvested in the field in Bushbuckridge (Mpumalanga Province).....	1
Figure 1.2: Global cassava production trends for a ten-year (2007-2016) period.....	3
Figure 1.3: Plant damage symptoms caused by diseases and pest infestation on cassava.	7
Figure 1.4: Cassava mealybugs (<i>Phenacoccus manihoti</i>) infesting cassava plants in South Africa.	9
Figure 1.5: Modified life cycle of <i>P. manihoti</i> as indicated by (Calatayud & Le Rü, 2006).	10
Figure 1.6: A schematic representation of direct and indirect plant mechanisms upon herbivore attack.	13
Figure 1.7: A schematic representation of signalling hormones involved in biotic stress defence response.	15
Figure 1.8: An overview of Illumina platform workflow.....	22
Figure 2.1: Cassava plants multiplied in the growthroom and greenhouse for rearing and mealybug infestation experiments.	28
Figure 2.2: Schematic representation of the experimental approach for RNA isolation and cDNA preparation for Illumina RNA-seq from biological replicates of cassava genotypes infested with mealybugs and non-infested mock.	32
Figure 3.1: Cassava mealybugs mounted and stained on a slide under the microscope.....	36
Figure 3.2: Agarose gels (1%) of mealybug (CMB) high molecular weight DNA (A) and PCR amplicons (B and C) from KwaZulu Natal and Mpumalanga provinces.	37
Figure 3.3: Classification of the South African mealybug species from cassava using 18S rRNA.	40
Figure 3.4: Classification of the South African mealybug species from cassava using 28S rRNA.	42
Figure 4.1: Mealybug-induced symptoms on the plants.	46
Figure 4.2: Agarose gel (1%) of total RNA isolated from AR23.1, P40/1 and 98/0581 cassava genotypes at 24 and 72 hours of mealybug infestation.	47
Figure 4.3: Per base sequence quality plot using FastQC of the left and right AR23.1 sequence reads after trimming.....	47
Figure 4.4: Venn diagrams of differentially expressed genes (DEGs).	50
Figure 4.5: Analysis of enriched GO terms.....	51

Figure 4.6: KEGG classification of differentially expressed genes.	53
Figure 4.7: Localization of the up-regulated plant-pathogen interaction pathway enzymes on the leaf tissues of AR23.1 post-infestation with mealybugs.	54
Figure 4.8: Localization of the down-regulated plant-pathogen interaction pathway enzymes on the leaf tissues of AR23.1 post-infestation with mealybugs.	55
Figure 4.9: Localization of the down-regulated plant-pathogen interaction pathway enzymes on the leaf tissues of P40/1 post-infestation with mealybugs.	58
Figure 4.10: Localization of the up-regulated plant-pathogen interaction pathway enzymes on the leaf tissues of 98/0581 post-infestation with mealybugs.	60
Figure 4.11: Localization of the down-regulated plant-pathogen interaction pathway enzymes on the leaf tissues of 98/0581 post-infestation with mealybugs.	61
Figure 4.12: Network representation of the interactions of the up-regulated differentially expressed genes (DEGs) in AR23.1 post-infestation with mealybugs.	64
Figure 4.13: Network representation of the interactions of the down-regulated differentially expressed genes (DEGs) in AR23.1 post-infestation with mealybugs.	65
Figure 4.14: Comparative expression levels by RNA-seq and real-time RT-qPCR of nine cassava up-regulated genes in genotype AR23.1 in response to mealybug infestation at 24 and 72 hpi calculated with reference to UBQ10 and GTPb.	67

LIST OF TABLES

Table 1.1: Current status of the cassava reference genome: <i>Manihot esculenta</i> v6.1.....	23
Table 2.1: Cassava genotypes selected for infestation and sequencing experiments.....	28
Table 2.2: Description of experimental samples.....	29
Table 2.3 Primer sequences of selected DEGs used for real-time quantitative reverse transcription-polymerase chain reaction (RT-qPCR).	34
Table 4.1: Scoring of AR23.1, P40/1 and 98/0581 cassava genotypes for phenotypic responses to cassava mealybug (CMB) infestation at 7, 14, 21, 28 and 35 dpi.....	45
Table 4.2: Summary of raw and trimmed reads generated by Illumina HiSeq 2500.	48
Table 4.3: An overview of the number of up-regulated and down-regulated genes (DEGs) identified between non-infested mock and mealybug infested leaves at 24 and 72 hours of mealybug infestation.	49

LIST OF APPENDICES

Appendix A1: Reagents, solutions and kits.....	99
Appendix A2: Solutions	99
Appendix A3: Extraction kits.....	99
Figure A1: Example of raw Sanger CMB_1.1 18S forward rRNA sequences of mealybugs (<i>P. manihoti</i>) before alignments and analysis.....	101
Figure A2: Example of raw Sanger CMB_1.1 18S reverse rRNA sequences of mealybugs (<i>P. manihoti</i>) before alignments and analysis.....	102
Figure A3: Sequence alignments of nine CMB 18S regions of species of mealybug from KwaZulu Natal and Mpumalanga in comparison with GenBank <i>P. manihoti</i> voucher species.	106
Figure A4: Multiple sequence alignments of 18S region of species of mealybug (<i>P. manihoti</i>) from KwaZulu Natal and Mpumalanga with comparison to the GenBank databases.....	112
Figure A5: Example of raw Sanger CMB_1.1 28S forward rRNA sequences of mealybugs (<i>P. manihoti</i>) before alignments and analysis.....	113
Figure A6: Example of raw Sanger CMB_1.1 28S reverse rRNA sequences of mealybugs (<i>P. manihoti</i>) before alignments and analysis.....	114
Figure A7: Sequence alignments of nine CMB 28S regions of species of mealybug from KwaZulu Natal and Mpumalanga in comparison with GenBank <i>P. manihoti</i> voucher species.	119
Figure A8: Multiple sequence alignments of 28S region of species of mealybug (<i>P. manihoti</i>) from KwaZulu Natal and Mpumalanga with comparison to the GenBank databases.....	126
Figure A9: Log/log correlations of AR23.1 leaf samples infested with mealybugs (24 and 72 hours post-infestation) compared to non-infested controls.....	127
Figure A10: Log/log correlations of P40/1 leaf samples infested with mealybugs (24 and 72 hours post-infestation) compared to non-infested controls.....	128
Figure A11: Log/log correlations of 98/0581 leaf samples infested with mealybugs (24 and 72 hours post-infestation) compared to non-infested controls.....	129
Figure A12: Localization of up-regulated plant hormone signal transduction pathway enzymes on the leaf tissues of AR23.1 post-infestation with mealybugs.....	130
Figure A13: Localization of down-regulated plant hormone signal transduction pathway enzymes on the leaf tissues of AR23.1 post-infestation with mealybugs.....	131

Figure A14: Localization of up-regulated phenylpropanoid biosynthesis pathway enzymes on the leaf tissues of AR23.1 post-infestation with mealybugs.	132
Figure A15: Localization of down-regulated phenylpropanoid biosynthesis pathway enzymes on the leaf tissues of AR23.1 post-infestation with mealybugs.	133
Figure A16: Localization of up-regulated glutathione metabolism pathway enzymes on the leaf tissues of AR23.1 post-infestation with mealybugs.	134
Figure A17: Localization of down-regulated glutathione metabolism pathway enzymes on the leaf tissues of AR23.1 post-infestation with mealybugs.	135
Figure A18: Localization of up-regulated plant hormone signal transduction pathway enzymes on the leaf tissues of P40/1 post-infestation with mealybugs.	136
Figure A19: Localization of down-regulated plant hormone signal transduction pathway enzymes on the leaf tissues of P40/1 post-infestation with mealybugs.	137
Figure A20: Localization of up-regulated phenylpropanoid biosynthesis pathway enzymes on the leaf tissues of P40/1 post-infestation with mealybugs.	138
Figure A21: Localization of down-regulated phenylpropanoid biosynthesis pathway enzymes on the leaf tissues of P40/1 post-infestation with mealybugs.	139
Figure A22: Localization of up-regulated glutathione metabolism pathway enzymes on the leaf tissues of P40/1 post-infestation with mealybugs.	140
Figure A23: Localization of down-regulated glutathione metabolism pathway enzymes on the leaf tissues of P40/1 post-infestation with mealybugs.	141
Figure A24: Localization of up-regulated plant hormone signal transduction pathway enzymes on the leaf tissues of 98/0581 post-infestation with mealybugs.	142
Figure A25: Localization of down-regulated plant hormone signal transduction pathway enzymes on the leaf tissues of 98/0581 post-infestation with mealybugs.	143
Figure A26: Localization of up-regulated phenylpropanoid biosynthesis pathway enzymes on the leaf tissues of 98/0581 post-infestation with mealybugs.	144
Figure A27: Localization of down-regulated phenylpropanoid biosynthesis pathway enzymes on the leaf tissues of 98/0581 post-infestation with mealybugs.	145
Figure A28: Localization of up-regulated glutathione metabolism pathway enzymes on the leaf tissues of 98/0581 post-infestation with mealybugs.	146

Figure A29: Localization of down-regulated glutathione metabolism pathway enzymes on the leaf tissues of 98/0581 post-infestation with mealybugs.....	147
Figure A30: Network representation of the interactions of the up-regulated differentially expressed genes (DEGs) in P40/1 post-infestation with mealybugs.	148
Figure A31: Network representation of the interactions of the down-regulated differentially expressed genes (DEGs) in P40/1 post-infestation with mealybugs.	149
Figure A32: Network representation of the interactions of the up-regulated differentially expressed genes (DEGs) in 98/0581 post-infestation with mealybugs.....	150
Figure A33: Network representation of the interactions of the down-regulated differentially expressed genes (DEGs) in 98/0581 post-infestation with mealybugs.....	151
Figure A34: Descriptive statistics of four candidate housekeeping genes (HKG) based on their crossing point (CP) values. BestKeeper index is computed together with the same descriptive parameters for four genes (UBQ10, GTPb, EF1 and Actin).	152
Figure A35: Expression stability of the four candidate housekeeping genes (UBQ10, GTPb, EF1 and Actin) in cassava leaves..	153
Table A1: A list of of raw and trimmed reads generated per time point for each biological replicate and subsequently mapping to the cassava reference genome (Mesculenta_305_v6.1).	154
Table A2: List of AR23.1 differentially expressed genes in cassava leaves in response to mealybug infestation at 24 and 72 hours post-infestation normalized against mock (non-infested).....	155
Table A3: List of P40/1 differentially expressed genes in cassava leaves in response to mealybug infestation at 24 and 72 hours post-infestation normalized against mock (non-infested).....	162
Table A4: List of 98/0581 differentially expressed genes in cassava leaves in response to mealybug infestation at 24 and 72 hours post-infestation normalized against mock (non-infested).....	167
Table A5: Gene Ontologies (GO) for AR23.1, P40/1 and 98/0581 cassava genotypes following mealybug infestation categorized in Biological process, Molecular function and Cellular component.	174
Table A6: An example sample RT-qPCR data for analysis. In this data set, housekeeping genes (UBQ10 and GTPb) and genes of interest are represented by three replicates at 24 and 72 hpi compared to mock infested.....	179

LIST OF ABBREVIATIONS

2OG	2-Oxoglutarate and Fe(II)-dependent oxygenase superfamily protein
ACMV	African cassava mosaic virus
AFLP	Amplified fragment length polymorphism
Alpha/B	Alpha/beta-hydrolases protein family
ANOVA	Analysis of variance
ARC	Agricultural Research Council
avr	Avirulence
BAC	Bacterial artificial chromosome
BC+	Bio-Cassava plus
BLAST	Basic local alignment search tool
bp	Base pair
CBB	Cassava bacterial blight disease
CBSD	Cassava brown streak disease
CC-NBS-LRR	Coiled-coil nucleotide-binding site leucine-rich repeat
cDNA	Complementary DNA
CGM	Cassava green mite
CIAT	Centro Internacional Agricultura Tropical (International Center for Tropical Agriculture)
CMB	Cassava mealybug
CMD	Cassava mosaic disease
CO ₂	Carbon dioxide
CO1	<i>Cytochrome c oxidase 1</i>
COI1	Coronatine-insensitive protein 1
C _t	Threshold cycle
CYP	Cytochrome P450 superfamily proteins
DEPC	Diethyl pyrocarbonate
DNA	Deoxyribonucleic acid
DRC	Democratic Republic of Congo
E	Efficiency
ERF1	Ethylene responsive transcription factor 1
ET	Ethylene
ETI	Effector-triggered immunity
FAO	Food and Agriculture Organization
FEC	Friable embryogenic callus
Fig	Figure

FPKM	Fragments per kilobase of exon per million mapped reads
GDSL	GDSL-like lipase/acylhydrolase superfamily protein
GH32	Glycosyl hydrolases family 32
Gibb	Gibberellin 2-oxidase 6
GO	Gene ontology
GTPb	GTP binding
hpi	Hours post-infestation
HSP20	Heat shock protein 20-like chaperones superfamily protein
IITA	International Institute of Tropical Agriculture
JA	Jasmonic acid
JAR1	Jasmonate resistant 1
kb	Kilobase
KEGG	Kyoto Encyclopaedia of Genes and Genomes
KO	KEGG ontology
KZN	KwaZulu Natal
LRR	Leucine-rich repeat
MAPKs	Mitogen activated protein kinases
Mb	Megabase
ML	Maximum likelihood
MP	Mpumalanga
MS	Murashige and Skoog medium
mt	Million tonnes
Myb	Myb domain protein family
NBS	Nucleotide-binding site
NBS-LRR	Nucleotide-binding site leucine-rich repeat
NCBI	National Center for Biotechnology Information
NGS	Next-generation sequencing
NPR1	Non-expressor of pathogenesis-related gene 1
PCR	Polymerase chain reaction
PDF1.2	Plant defensin 1.2
PIs	Plant introductions
PR gene	Pathogenesis-related gene
PTI	PAMP-triggered immunity
R gene	Resistance gene
RAPD	Random amplified polymorphic DNA
RFLP	Restriction fragment length polymorphism

RNA	Ribonucleic acid
RNA-seq	Ribonucleic acid sequencing
RT	Reverse transcription
RT-qPCR	Reverse transcription-quantitative polymerase chain reaction
TIR-NBS-LRR	Toll interleukin-like receptor nucleotide-binding site leucine-rich repeat
SA	Salicylic acid
SAB	South African Breweries
SAR	Systemic Acquired Resistance
SNP	Single nucleotide polymorphism
SOLiD	Sequencing by Oligonucleotide Ligation and Detection
SSCP	Single strand conformation polymorphism
SSR	Simple sequence repeat
UBQ10	Ubiquitin 10
VOPI	Vegetable and Ornamental Plant Institute

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1 Cassava origin and domestication

The cassava (*Manihot esculenta* Crantz) plant is a shrub that belongs to the *Euphorbiaceae* family (Figure 1.1) (Leite *et al.*, 2002; Lopez *et al.*, 2005a). The genus *Manihot* consists of 98 species native to the region from the south-western United States of America and extending southwards to Argentina (Chavarriaga-Aguirre & Halsey, 2005; Nassar, 2000; Rogers & Appan, 1973). Four of the *Manihot* species are described in Central Brazil, Mexico, Northeast Brazil and Southwest Brazil with 38, 17, 16 and 6 species respectively. Out of the known 98 species of the *Manihot* species, cassava is the only widely cultivated species that is used for food (Figure 1.1) (Nassar, 2003).



Figure 1.1: A cassava plant and roots grown and harvested in the field in Bushbuckridge (Mpumalanga Province). A: A mature cassava plant in the field ready for harvest. B: Freshly harvested cassava roots from the same cassava plant.

It has been hypothesized that domesticated cassava has unknown taxonomic origin, possibly comprising multiple *Manihot* progenitors (Rogers & Appan, 1973). However, phylogenetic and morphological studies point to a single progenitor of *Manihot esculenta* subsp. *flabellifolia* (Pohl) Ciferri (Allem, 1994; Olsen, 2004; Olsen & Schaal, 1999). Based on morphology, ecology and geography, other wild species regarded to be close to cassava are *Manihot carthaginensis* found in all countries bordering the Caribbean; *Manihot aesculifolia* found in Mexico and Central America; *Manihot grahami* found in Brazil, Paraguay, Uruguay and Argentina; *Manihot flabellifolia* found in Brazil, Paraguay, Uruguay and Argentina; and *Manihot saxicola* found in Guiana, Surinam and Venezuela (Allem, 1999, 2002). The cultivation of cassava has spread from its centre of origin to different parts of the world including Asia and Africa (Cock, 1985; Gleadow *et al.*, 2009). Cassava was brought to Asia by the European explorers in the late 18th to 19th centuries (Onwueme, 2002). India, Java and the Philippines were the first locations for the arrival of cassava

in Asia. It was brought to Africa by Portuguese sailors in the 16th century, primarily through the west coast (Congo) from where it spread towards east Africa as far as Zanzibar (Janssens, 2001). Cassava has become a major crop in sub-Saharan Africa, especially in locations with high population densities such as Nigeria, Democratic Republic of Congo (DRC) and Tanzania, (FAOSTAT, 2016 ; IITA, 1992; Jones, 1959).

1.2 Cassava characteristics

Cassava is a monoecious species, meaning both male and female flowers exist on the plant (Halsey *et al.*, 2008; Narina & Odeny, 2012). The male flowers are small (0.5 cm) and situated towards the tip of the inflorescence, while the female flowers are slightly larger and situated at the base of the inflorescence (Chavarriaga-Aguirre & Halsey, 2005). Basically, flowering can take up to 2 months, and is altered by environmental conditions (Narina & Odeny, 2012; Parkes, 2011).

Cassava is believed to have high rates of outcrossing, and interspecific crosses between *Manihot esculenta* and wild relatives are possible. The crop produces pollen grains that differ in size, approximately 130-150 µm (larger) and 90-110 µm (smaller) in diameter (Halsey *et al.*, 2008; Vieira *et al.*, 2012). Wind and insect pollination (Rogers & Appan, 1973) are reported to be common although self-pollination also occurs (Lai *et al.*, 2014). Cassava pollen loses its viability after being shed (Chavarriaga-Aguirre & Halsey, 2005) and therefore, breeders prefer to perform pollination within an hour in order to obtain successful fertilizations.

Seeds can be produced from cassava cultivars that flower. When cassava seeds are planted, it takes about 75-90 days (3 months) for them to germinate (Ceballos *et al.*, 2004). Ultimately, one seed is acquired per fruit in self-contained pollination (Jennings, 1963), however, 2-3 seeds can be acquired in trilocular ovary (Jennings & Iglesias, 2002). Seeds often have a dormancy period; therefore, newly harvested seeds require 3-6 months storage in controlled temperatures, before germination (Parkes, 2011). However, seeds at controlled temperatures may lose feasibility within one year. Propagation of cassava by seeds is therefore slow with low seed production.

The crops' stem on the other hand, is woody and tube-shaped with regular nodes and internodes (Alves, 2002). The main stems are erect with a branching pattern that is commonly dichotomous or trichotomous (Nassar *et al.*, 2008). Secondary branches occur from the first branches, and that produces other successive branching. From the branches of the stem, leaves are formed by the lamina and the petioles, which are lobed with palmated veins (Alves, 2002). Mature leaves develop between 4-5 months depending on the planting time and cultivar/genotype, and the number of lobes on mature leaves differs among cultivars. Mature leaves are smooth with each leaf surrounded by two petioles. The upper leaf surface is bright containing a

wax epidermis. The stomata are situated on the upper and lower surface of the leaves, occupying 1.4 - 31.1% of the total leaf area.

1.3 Cassava production and importance

Cassava is a major source of energy in tropical regions of Africa, Asia and Latin America (Hershey *et al.*, 2012). The starchy roots are the cassava's economic product, because they can be widely used in human food, industrial products and animal feed (Hershey *et al.*, 2012; Vieira *et al.*, 2012). Global cassava production reached a maximum of 277 million tons (mt) in 2016 (Figure 1.2), with Africa producing approximately 54% of the total world production, followed by Asia (31.1%) (FAOSTAT, 2016). According to statistics from FAOSTAT (FAOSTAT, 2016), more than 157 million tons of cassava was produced in Africa in 2016, with Nigeria producing 57.1 million tons. The other largest producers in Africa are Ghana (17.7 mt), Democratic Republic of Congo (14.6 mt), Angola (9.9 mt) and recently, Mozambique (9.1 mt).

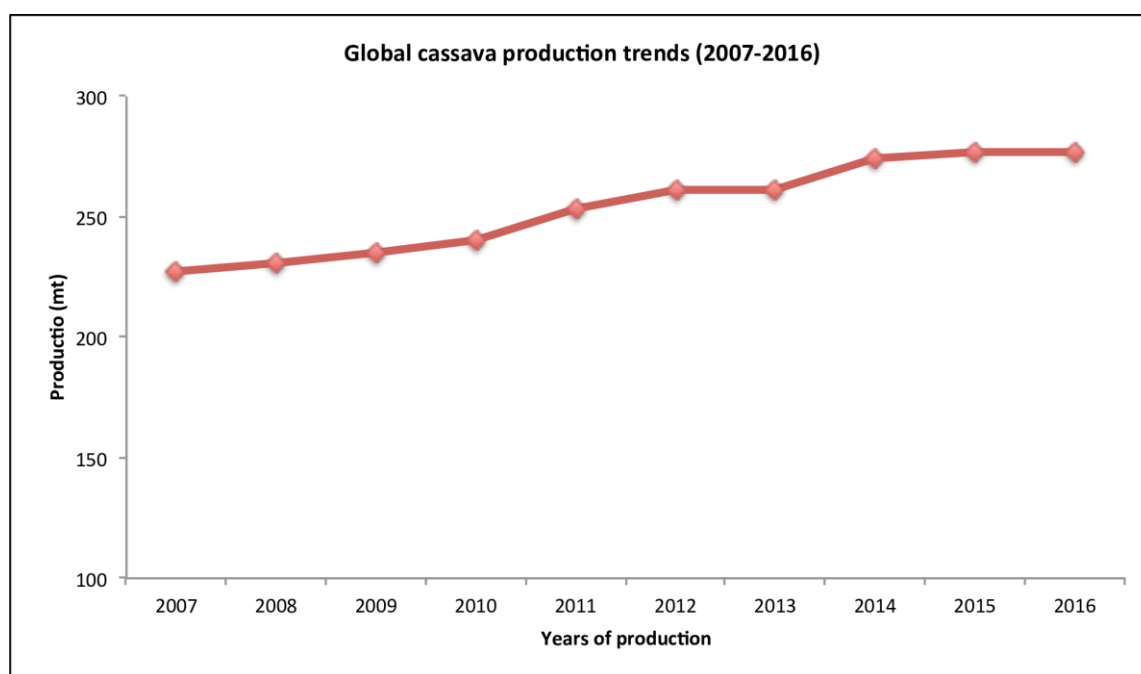


Figure 1.2: Global cassava production trends for a ten-year (2007-2016) period. The trends were measured in million tons (mt) Source: (FAOSTAT, 2016). Production data was collected from all cassava-producing areas in each country globally.

Cassava production serves different purposes in different continents. It is mainly used for food and cash income in Africa (Dahinya, 1994; Nweke *et al.*, 2002; Nweke, 1996) while in most Latin American countries used for industrial purposes (Henry & Hershey, 2002; Nassar & Ortiz, 2007). In Asia, cassava is consumed differently in different regions (Howeler, 2006). Fresh roots are produced for human consumption, chips and pellets are used for animal and poultry feed, and starch is produced for food and industry. For example, in Indonesia and India, 70% of cassava is produced for human consumption

(Onwueme, 2002). However, most cassava production in China and Vietnam is for industrial production (Jansson *et al.*, 2009) with only 10-20% being used for human consumption.

1.3.1. Contribution to Food Security

Cassava is a staple food in West and Central Africa, some parts of East Africa, and is consumed as a supplement for other staples in Southern Africa. It is widely cultivated in Africa in such a way that its annual per capita food consumption is comparable to the rest of the world (Aerni, 2006). Cassava is processed into different food products in Africa (Kenyon *et al.*, 2006). In West Africa, flour and *gari* are the most cassava-processed products that are consumed in large quantities on a daily basis (Tonukari, 2004). Besides being processed, the roots can be boiled, fried, roasted or consumed raw. For example, in Nigeria and Congo, *fufu* (roots boiled and pounded into paste) can be consumed by adding it to stew or soup. In South Africa, cassava roots may be used to substitute/compliment rice or maize meal, prepared like potato by peeling and then boiled, baked, or fried. There are also a number of cassava food-processing methods consisting of drying and pounding of the roots to produce porridge known as '*Xigema*' (Xitsonga language) (Makwarela & Rey, 2006). Cassava leaves can also be eaten as vegetables because they are high in protein, vitamin C, iron and calcium. The leaves are regarded as an important vegetable in Congo (Nweke *et al.*, 2002) and can be cooked to produce a condiment known as '*Mathapi*' in South Africa (Makwarela & Rey, 2006). Cassava leaves are also used as a substitute for other *morogo* (indigenous leafy vegetable) types (like *Amaranth* leaves) in South Africa.

1.3.2. Industrial starch

Cassava has attracted more attention recently in terms of starch because it can be cultivated on negligible land where other starch-producing crops cannot grow well (Jackson *et al.*, 2014; Wilson, 2002). Among other advantages, it supplies good quality industrial starch that is cost effective (Aerni, 2006; Sánchez *et al.*, 2017; Taiwo, 2006). Nonetheless, private sectors are not supporting cassava's starch supply because of its competitors (e.g. corn, maize, sugar cane), which are preferred by the global market. Regardless, cassava still performs better for some purposes than the other starches. It can be utilized in the pharmaceutical industry, textile industry, and used as a raw material in glue and adhesive industry, and in the cosmetics industry (Howeler, 2006; Jackson *et al.*, 2014; Tonukari, 2004; Zhang *et al.*, 2016).

Cassava starch processing was established in the 1940s in Asia, and it was the first major industrial use (Onwueme, 2002). Among the countries producing cassava starch in Asia are Malaysia, Philippines and China with Thailand being the largest producer. About half of cassava is produced for industrial use in Asia, while exports account for 27% since the start of production (Howeler, 2006; Onwueme, 2002). Additionally, in Latin American countries, starch processing from other crops is an important commodity

in local markets, however, it is not major from cassava (Henry & Hershey, 2002), but mostly used in other crops like maize and sugar cane. Small-scale farmers throughout the Latin American countries produce modified or native starch from cassava, with Brazil being the largest producer.

Africa is the largest producer of cassava, however, its success in terms of industry is still lacking. For African countries to have successful cassava industries, cost of production, processing and marketing should be reduced. This will enable African countries to advance to industrial production and to compete on the global markets. Glue, glucose syrup, ethanol and bread are among the industrial products produced in Nigeria (Dawit, 2010). Other industrial products from processed cassava are sweeteners, mosquito coils and brewing ingredients. Recently, a beer brewed in Mozambique by the South African Breweries Miller (SAB) is produced using 70% of cassava and 30% barley (Madeley, 2011). Therefore, if cassava production is optimized, it could potentially provide income to resource-poor farmers and generate foreign currency via export.

1.3.3. Bioethanol production

As a source of carbohydrates, cassava is an important alternative source of starch for ethanol production (Pingmuanglek *et al.*, 2017; Poonsrisawat *et al.*, 2017). The roots or starch extracted from the roots are the source for production (Hanif *et al.*, 2016; Sanchez & Cardona, 2008). Brazil, Thailand, China and Nigeria use cassava for ethanol production. A study on ethanol fuel production from cassava as a substitute for gasoline was performed in Nigeria (Adeniyi *et al.*, 2007). From the yield of ethanol produced, the results indicated that cassava is a feasible source of ethanol fuel that can substitute the conventional fossil fuels. Furthermore, in South Africa, ethanol was obtained from unpeeled cassava roots and cassava peels (Marx & Nquma, 2013). The highest ethanol yield was from simultaneous saccharification and fermentation, while the lowest was from direct fermentation.

In Asia, cassava is considered a great crop for production of fuel ethanol because it can be grown all year round, and its raw material is cheap (Nguyen *et al.*, 2007). A study done in China revealed that cassava-derived bio-ethanol required (6.6 tonnes) for 1 tonne of ethanol as compared to maize (3.2 tonnes) and sugarcane (13.5 tonnes) (Jansson *et al.*, 2009). In Thailand, cassava dried chips are recommended for ethanol production, however, it can also be produced from fresh roots (Silalertruksa & Gheewala, 2010), cassava stem (Zhu *et al.*, 2015), and from cassava waste (Akaracharanya *et al.*, 2011). This makes cassava an important industrial commodity in Asia. In Latin American countries, raw cassava is used for ethanol production (Ferreira-Leitao *et al.*, 2010), however, sun-dried cassava chips are preferable in order to avoid deterioration of roots before production starts.

1.3.4. Other important uses of cassava

Besides being produced for human consumption and industrial purposes, cassava can be used for animal (Balagopalan, 2002; Henderson, 2015; Trung *et al.*, 2014) and poultry feed (Chauynarong *et al.*, 2009; Morgan & Choct, 2016). Animal and poultry feed can be produced from roots, stems and leaves of cassava (Balagopalan, 2002; Henderson, 2015; Howeler, 2006; Ravindran, 1993). Leaves can be fed as forage in the dry seasons when feed is scarce. The roots can be processed into dry chips and pellets, before being fed to animals (Howeler, 2006; (Phu et al., 2016). Moreover, animal feed can be processed as silage from cassava whole plants in closed polythene bags, or produced from protein-enriched cassava by suppressed fermentation using microorganisms (Balagopalan, 2002). Processed animal and poultry feed can be used locally or exported.

1.4 Factors affecting cassava production and utilization

Although cassava cultivation has exponentially increased over the years, its production potential has been hampered by abiotic and biotic factors. The effects of each of these factors are discussed in more detail below.

1.4.1 Stem propagation methods

Propagation of cassava by stem cuttings is the method preferred by most farmers, however, the process is slow and can lead to the accumulation of viral and bacterial diseases. These constraints reduce productivity of the crop and may cause loss of superior genotypes (Nassar & Ortiz, 2007). Diseases such as cassava mosaic disease, cassava bacterial blight disease and cassava brown streak disease can be accumulated through stem cuttings from infected plants (Calvert & Thresh, 2002; Legg, 1999). To overcome the accumulation of viruses and diseases through traditional propagation methods, techniques of rapid multiplication such as micro-propagation by in culture and mini-propagation were introduced (Santana *et al.*, 2009). The development of these techniques has significantly increased the rate of multiplication of disease-free material. These innovations can also be used to produce high quality planting material. For example, in a study by Kartha & Gamborg, (1975), *in vitro* techniques by meristem culture method were used to eliminate cassava mosaic virus in Indian and Nigerian diseased plants. In addition, *in vitro* techniques were also applied in eliminating cassava brown streak virus (Wasswa *et al.*, 2010) and cassava mosaic virus (Nkaa *et al.*, 2013; 2014) in cassava. On the other hand, cassava seeds, which are mostly used in breeding programs depending on the cultivar, are a promising option to eliminate constraints, particularly diseases associated with vegetative propagation.

1.4.2 Diseases and arthropod pests

Diseases and pests reduce cassava yields, especially in Africa where cassava is an important food crop (Mahungu *et al.*, 2009; Nassar & Ortiz, 2007). Major cassava diseases include cassava mosaic disease (CMD; Fig1.3A), cassava bacterial blight (CBB) and cassava brown streak disease (CBSD) (Anjanappa *et al.*, 2016), while pests are whiteflies, cassava green mites (CGM) and cassava mealybugs (CMB; Fig 1.3 B) (Hillocks *et al.*, 2002). These constraints will further be discussed in detail.



Figure 1.3: Plant damage symptoms caused by diseases and pest infestation on cassava. A show the virus symptoms of cassava mosaic disease (CMD) on the leaves of the plant in the field in Bushbuckridge (Mpumalanga). B shows the cassava leaves and the stem infested with mealybugs (CMB) in a glasshouse at ARC-VOPI, Roodeplaat. Some leaves are curled and the stem is distorted due to infestation.

1.4.2.1 Cassava mosaic disease (CMD)

Cassava mosaic disease is one of the important biotic constraints of cassava and is caused by the whitefly-transmitted *Geminivirus* species (Allie *et al.*, 2014; Calvert & Thresh, 2002; Otim-Nape *et al.*, 1994; Thresh & Cooter, 2005). It is widespread in Africa and species that cause this disease include African cassava mosaic virus, East African cassava mosaic virus and South African cassava mosaic virus. The disease was first identified in Democratic Republic of Congo (DRC) vectored by whitefly of genus *Bemisia* (Legg, 1999; Legg *et al.*, 2006), and became a threat to cassava production in Africa. Whiteflies use their stylets to pierce the plant tissues and inject toxic saliva during sucking of the phloem elements into the plant. Symptoms of infection by CMD to cassava plants include chlorotic mosaic symptoms, leaf deformation and stunting. Infection with CMD can cause between 20-60% of storage yield reductions, and sometimes causes total crop failures. Farmers use insecticides to control the spread of whiteflies from transmitting the virus to other plants (Calvert & Thresh, 2002). However, the use of resistant varieties is the control measure that is being given more attention, due to their ability to resist infection by the virus. Any diseased plants that occur in the field are also removed as a control measure.

1.4.2.2 Cassava bacterial blight disease (CBB)

Cassava bacterial blight disease is the most economically important bacterial disease of cassava in cassava growing regions (Lopez *et al.*, 2005b; López & Bernal, 2012). It was introduced to Africa in the 1970s, caused by *Xanthomonas axonopodis* pv. *manihotis* (Boher & Verdier, 2009; Wydra & Verdier, 2002). Dissemination of the disease is mainly by infected planting material (Jorge *et al.*, 2000), which result in wilting and defoliation of infected leaves, and could lead to death of plants in susceptible cultivars. Where there are favourable conditions for the spread of the pathogen, complete yield losses can also be experienced by CBB infection. Control measures include the use of disease-free planting material, and disease resistance screening of the genotypes (Fokunang *et al.*, 2000).

1.4.2.3 Cassava brown streak disease (CBSD)

Cassava brown streak disease is a cassava disease caused by cassava brown streak virus (CBSV) of the genus *Ipomovirus* (Hillocks & Jennings, 2003; Mbanzibwa *et al.*, 2011), which is transmitted by whiteflies (Mbanzibwa *et al.*, 2009). The disease was first described in Tanzania (Maruthi *et al.*, 2005), and since then, it has been found in all cassava growing regions, while it became endemic to East African coastal regions. The disease can be disseminated by infected planting material (Calvert & Thresh, 2002). Infected plants with CBSD show stem and root necrosis which could lead to 70% loss of root weight in susceptible cultivars (Wasswa *et al.*, 2010). Natural measures to prevent the spread among plants is still lacking for control of the disease. Sources of resistance to CBSD are also assessed (Patil *et al.*, 2011), together with *in vitro* techniques for elimination of the disease.

1.4.2.4 Cassava green mites (CGM)

Cassava green mites are important cassava pests that mostly occur during the dry season. They were first found infesting cassava by late 1971 in Uganda, East Africa (Gutierrez *et al.*, 1988). Since the pest's introduction in Africa, it has spread to the entire cassava belt causing serious damage to cassava production (Mutisya *et al.*, 2015b; Yaninek & Hanna, 2003). Reduction in leaf size, defoliation and candlestick-like appearance of young shoots are observed in CGM infested plants. Different methods can be used for control of CGM. Due to the unsuccessful use of pesticides to control the pest, researchers at International Institute of Tropical Agriculture (IITA) then established biological control of CGM. CGM's true identity was established and its behaviour within the cassava ecosystem and damage to the cassava plants was studied. IITA researchers then introduced and established predator mites *Typhlodromalus aripo* and *Typhlodromalus manihoti*, and later *Neozygites tanajoae* for biological control of CGM. The mites have been used since the 1990s as control measures in Africa (Aidoo *et al.*, 2016; Mutisya *et al.*, 2015a).

1.4.2.5 Cassava Mealybug (CMB)

Mealybugs are major pests infesting a wide range of crops and ornamental plants worldwide (Miller *et al.*, 2005; Miller *et al.*, 2002; Miller & Rossman, 1995). These phloem-sap sucking insects have been studied intensively for decades because of the economic losses they cause to agriculture, through direct damage to crops, virus transmission and decreases in yield and quality (Daane *et al.*, 2012; Kuniyuki *et al.*, 2006; Meyer *et al.*, 2008; Nakaune *et al.*, 2008; Sforza *et al.*, 2003). Cassava mealybug (*P. manihoti*) (Figure 1.4) is a polyphagous insect feeding on cassava leaves and stems (Calatayud *et al.*, 1994). It multiplies explosively in favourable conditions and can quickly devastate cassava crops. Upon introduction into Africa, the mealybug spread quickly throughout all cassava producing regions, threatening food security of over 500 million people (Bellotti *et al.*, 1999; Neuenschwander, 2001).



Figure 1.4: Cassava mealybugs (*Phenacoccus manihoti*) infesting cassava plants in South Africa. Mealybugs collected in a petri dish (1); mealybugs infesting cassava in a glasshouse at ARC-VOPI (2, 3 and 4).

Phenacoccus manihoti is a parthenogenetic species that reproduces without a mate. It has four developmental instars during its life cycle. Each adult female can lay up to 500 eggs in an ovisac under

favourable conditions. The first instar is the most mobile stage (Calatayud & Le Rü, 2006; Parsa *et al.*, 2012) and the entire life cycle from egg to adult takes about 21 days (Figure 1.5).

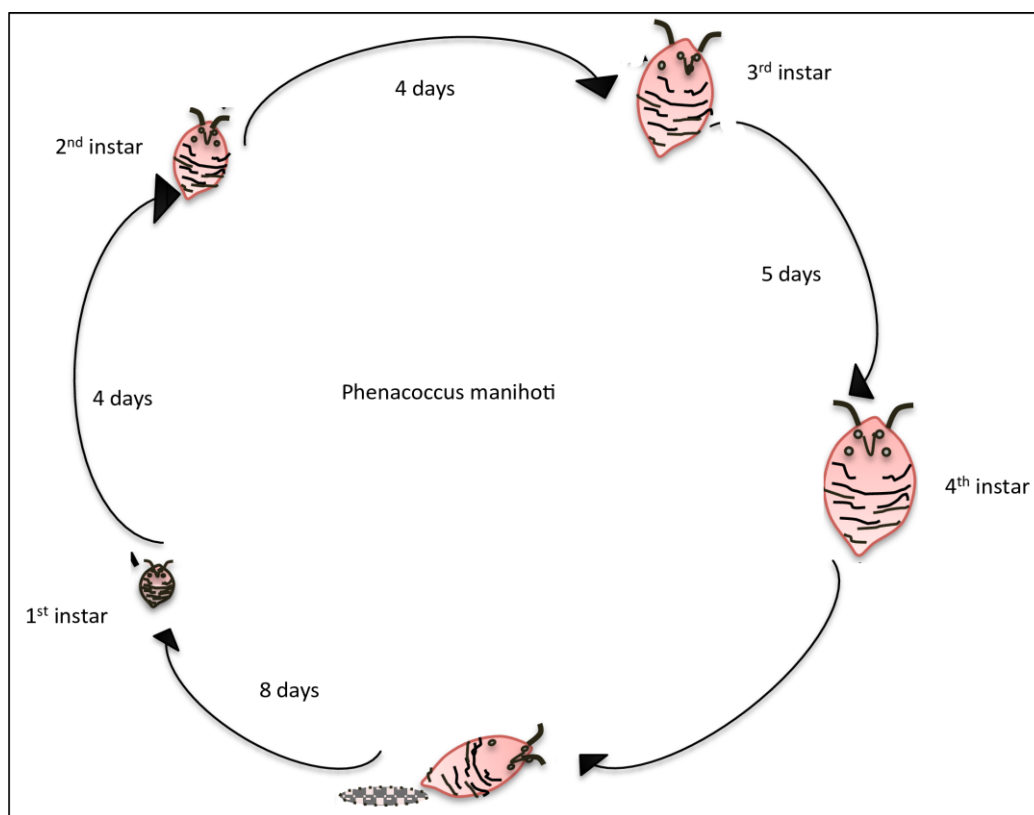


Figure 1.5: Modified life cycle of *P. manihoti* as indicated by (Calatayud & Le Rü, 2006). The developmental stages start from egg hatching until the adult stage, which takes 21 days.

Cassava mealybugs are specialist pests that prefer cassava and other *Manihot* species as their host for feeding. The mealybugs use stylets to suck the phloem sap from the plant tissues, and inject saliva into the plants (Le Ru & Calatayud, 1994). Infestation of plants results in leaf yellowing, leaf necrosis and defoliation, deformation of the terminal shoots, distorted stems, and plants may die in heavy infestations (Calatayud & Le Rü, 2006; Parsa *et al.*, 2012). Mealybug infestations may result in leaf loss and poor quality planting material for stem propagation. In addition, cassava mealybugs also secrete honeydew, which can be used to attract ants that protect the mealybugs from predators.

Besides the use of insecticides, biological control of CMB was implemented in the 1980s by the IITA research team in Africa and has been practiced with some success (Nassar & Ortiz, 2007; Neuenschwander, 2001). The researchers matched the CMB species infesting the cassava crop in Africa in South America (place of origin), and further identified the natural enemies (parasitoid) of the same species in South America. After rearing the natural enemies in the lab at IITA, they released the predatory wasps, *Anagyrus*

lopezi, and also developed a simulation model for plant/pest/predator-interactions, leading to a successful CMB biological control. Another means of controlling CMB is through establishment of resistant varieties.

Overall, the best strategy to offset the destructive effects of diseases and pests is by breeding for resistance (Akinbo *et al.*, 2012; Carabalí *et al.*, 2010a; 2010b; Morante *et al.*, 2004). To achieve this, cassava germplasm resistant to diseases and pests was established in the 1970s in Mozambique (Hillocks, 2002). They introduced the germplasm through subsequent breeding programmes that have since generated high yielding disease resistant genotypes. Most of the improved cultivars incorporating resistance to CMD and CBB are available at the IITA, Nigeria (Hillocks, 2002). Several clones that combine resistance to CGM, CMB and other diseases and pests have also been developed.

1.5 Identification of cassava mealybugs

Mealybugs can be identified using traditional morphological methods and molecular methods. The degree of morphological similarity between different species of mealybugs makes it difficult to correctly identify them and can limit their study and management (Malausa *et al.*, 2011). Therefore, molecular methods can be used to complement morphological methods. Both morphological and molecular identification methods will be discussed here in detail.

1.5.1 Morphological identification

Mealybugs are traditionally identified by subtle morphological differences between adult females (Millar, 2002). Identification depends on slide-mountings and thorough examination of tiny features under a microscope. Careful handling of the specimen is required for successful morphological identification due to the delicate nature of the process (Millar, 2002). Any damage during collection, preservation, transference or preparation for slide mounting could result in a lack of the characteristics necessary for identification to species level (Malausa *et al.*, 2011). Therefore, morphological identification requires a high level of expertise and cannot always distinguish between closely related taxa (Abd-Rabou *et al.*, 2012; Malausa *et al.*, 2011). Moreover, the taxonomic morphological-based identification method is laborious and only the morphology of adult females is used for species identification (Cook *et al.*, 2010). If plants in the field were infested with immature mealybugs, immediate identification would be impossible based on only morphology. Despite the delicate nature of morphological identification, the method has been used to successfully identify mealybug genera in South Africa (Millar, 2002). Other studies have combined the use of both morphological and molecular identification methods (da Silva *et al.*, 2014; Sethusa *et al.*, 2014; Wetten *et al.*, 2016). It is therefore recommended to use both morphological and molecular methods for more reliable identification of insect species of different genera.

1.5.2 Molecular identification

An alternative method to traditional morphological identification of mealybugs is through molecular methods (Beuning *et al.*, 1999; da Silva *et al.*, 2014; Wetten *et al.*, 2016). Molecular identification methods such as DNA fragment mapping (Alamalakala *et al.*, 2009; Cavalieri *et al.*, 2008; Rosas-Garcia *et al.*, 2010), DNA conformation polymorphism, (Sharpe *et al.*, 1999; Ren, 2000), and random amplification of polymorphic DNA (Demontis *et al.*, 2007) have been extensively applied in identifying different insect species groups. However, recently new tools such as barcoding of the mitochondrial DNA have been introduced that are more powerful than the previous methods (Abd-Rabou *et al.*, 2012; Park *et al.*, 2010a; Sethusa *et al.*, 2014; Sperling *et al.*, 1999).

1.5.2.1 DNA barcoding

DNA barcoding is referred to as a taxon identification method using a DNA region that is uniform (Ratnasingham & Hebert, 2007; Valentini *et al.*, 2009). The sequence data obtained is analysed with the intent of finding a genetic region of suitable variability for species identification (Kress & Erickson, 2008). DNA barcoding can be used to identify unknown specimens or species that have been previously described and to enhance the discovery of new species (Moritz & Cicero, 2004; Rach *et al.*, 2008). For example, the use of DNA barcoding has been demonstrated by (Abd-Rabou *et al.*, 2012) who identified 17 species of mealybug infesting crops and ornamental plants in Egypt using a combination of three DNA markers (Large subunit of the 28S ribosomal RNA: 28S-D2, *cytochrome c oxidase 1*: CO1, and *internal transcribed spacer 2*: ITS2) as well as morphological examination. Rung *et al.* (2008) also used two molecular markers, CO1 and *elongation factor 1* (EF1 α) to ascertain the identity of *Planococcus citri* and *Planococcus minor*, thus affirming DNA barcoding as a reliable method for the identification of species.

Overall, morphological and molecular techniques are both important and relevant for identification of mealybugs. However, it should be noted that molecular characterization has other advantages: it does not need a complete specimen, it is less time-consuming, and does not need specialist taxonomic expertise for identification.

1.6 Plant-insect interactions

From over 350 million years ago, plants and insects continue to battle against each other (Gatehouse, 2002; Wu & Baldwin, 2010). Insects have advanced to being able to recognize and detect their host plant for feeding using natural signals during their coevolution. As immobile organisms, plants encounter several species of herbivorous insects chewing on their tissues, sucking up their cell content and feeding on their vascular sap (Korth, 2003; Van Poecke, 2007; Wei *et al.*, 2007). An abundance of direct (Arimura *et al.*, 2005; Chen, 2008; Zhou *et al.*, 2011) and indirect defence (Dudareva *et al.*, 2013; Kessler & Baldwin,

2001, 2002; Kessler *et al.*, 2004; Stam *et al.*, 2014) mechanisms (Figure 1.6) have evolved to protect plants from pathogens and herbivores (Broekgaarden *et al.*, 2011). These mechanisms can be constitutively present or induced by herbivory stimuli. Accumulation of constitutive defences depends on the type of species infesting the plant; whilst with induced defences, characteristics are shared between the plants (Gatehouse, 2002).

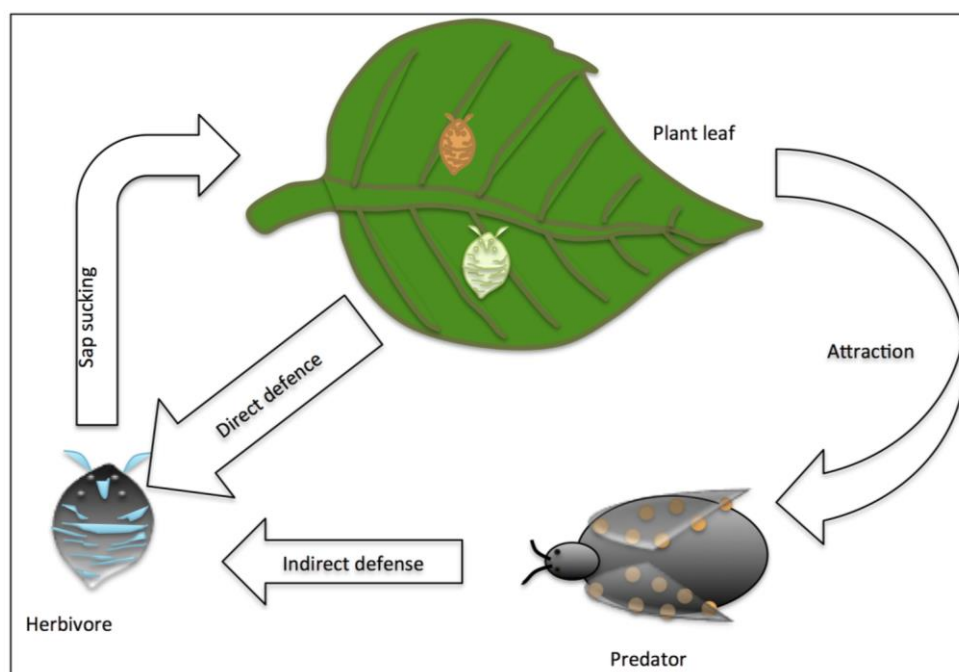


Figure 1.6: A schematic representation of direct and indirect plant mechanisms upon herbivore attack.

1.6.1 Direct defences

Direct defences may prevent the insect from feeding (trichomes, thorns, thick leaves) or produce toxic chemicals that include primary and secondary metabolites (Figure 1.6) that destroy or impede the development of the insect (Kessler & Baldwin, 2001; War *et al.*, 2012; Zhou *et al.*, 2011). In tomato Duffey & Stout, (1996) and tobacco (Steppuhn & Baldwin, 2007), the presence of secondary metabolites has been demonstrated to discourage feeding by herbivores. *Medicago truncatula* plants deterred aphids from settling and feeding on the plants within six hours of release in a study by (Klingler *et al.*, 2005). In cassava, both grasshopper (*Zonocerus variegatus*) and burrowing bug (*Cyrtomenus bergi*), appear to be deterred from feeding by cyanogen in leaves and roots (Bellotti & Riis, 1994; Riis *et al.*, 2003). Furthermore, artificial infestation in a wild species of cassava, *Jatropha gossypifolia* (Calatayud & Le Rü, 2006) with cassava mealybug (*Phenacoccus herreni*) ended in 100% mealybug mortality 48 hours, due to the release of the cyanogen in the leaves.

One of the most significant elements of pest defence in plants is a direct induced defence response (Howe & Jander, 2008). These defences are pest-specific and they have the advantage of reducing the metabolic costs to the host plant (Chen, 2008; Stam *et al.*, 2014). The cross-talk between the signal pathways results in the release of volatiles that further activate the expression of defence-related genes. To date, activation of plant defence signals by pierce-sucking herbivore feeding has been reported to be similar to those in response to pathogen infection (Artico *et al.*, 2014; De Vos *et al.*, 2005; Gutsche *et al.*, 2009; Lorenzo & Solano, 2005). Examples of these studies include the pathogen *Pseudomonas syringae* pv. *tomato* infection in *Arabidopsis thaliana* (De Vos *et al.*, 2005); green peach aphid (*Myzus persicae*) feeding in *Arabidopsis thaliana* (De Vos *et al.*, 2005; Moran & Thompson, 2001); Russian wheat aphid (*Diuraphis noxia*) feeding in wheat and barley (Gutsche *et al.*, 2009; Liu *et al.*, 2005); potato aphid (*Macrosiphum euphorbiae*) feeding in tomato (Cooper & Goggin, 2005), silverleaf whitefly (*B. tabaci* type B) feeding in *Arabidopsis thaliana* (Kempema *et al.*, 2007), greenbug (*Schizaphis graminum*) feeding in sorghum (Park *et al.*, 2006), and the grapevine mealybug, *Planococcus ficus* feeding in grapevines (Timm & Reineke, 2014) and the cotton mealybug, *Phenacoccus solenopsis* in cotton (Zhang *et al.*, 2011b). Signal molecules such as jasmonic acid (JA), salicylic acid (SA) and ethylene phytohormones are the ones induced upon biotic stress, and abscisic acid is induced upon abiotic stress (Gutzat & Scheid, 2012). These hormones have also been associated with plant-herbivore/pathogen interactions, which are triggered upon attack by insects and pathogens (Allie *et al.*, 2014; Artico *et al.*, 2014; De Vos *et al.*, 2005; 2006; Gao *et al.*, 2016; Li *et al.*, 2008; Utsumi *et al.*, 2016).

1.6.1.1 Jasmonic acid (JA)

Jasmonic acid signalling is involved in diverse processes such as root growth and tuber formation, as well as in biosynthesis of systemic acquired resistance (SAR) (Grant & Lamb, 2006; Ryals *et al.*, 1996) Hilfiker *et al.*, 2014; Luna *et al.*, 2012). It plays a major role in induced direct plant defence against insects and pathogens (Bari & Jones, 2009; Halitschke & Baldwin, 2004; Wang & Wu, 2013). This phytohormone is reported to be synthesized by the oxylipin pathway, and can interact with other phytohormones to facilitate transcriptional and metabolic realignments in plants after herbivore attack (Verhage *et al.*, 2010). Halitschke *et al.* (2001) indicated that the accumulation of JA appeared to be induced by the oral secretions of a specialist herbivore *Manduca sexta* in *Nicotiana attenuata*. Moreover, the expression of SA and JA-responsive genes were induced by green peach aphid (*Myzus persicae*) feeding in *Arabidopsis thaliana* (Moran & Thompson, 2001). Similarly to aphids, mealybug feeding does induce JA signal molecules in different plant species (Dubey *et al.*, 2013).

Jasmonic acid signalling is regulated by many components, however, three main components namely: Coronatine-Insensitive1 protein (COI1), Jasmonate Resistant 1 (JAR1) and MYC2 transcription factor play the major roles in its regulation (Figure 1.7). COI1 is an F-box protein that is part of the E3 ubiquitin-ligase

skip-cullin-F-box (SCF) complex (Chung *et al.*, 2009; Song *et al.*, 2014; Xie *et al.*, 1998; Xu *et al.*, 2002; Yan *et al.*, 2009). JAR1 component encodes a JA amino acid synthetase involved in conjugating JA-Ile (Staswick & Tiryaki, 2004), while MYC2 encodes the transcription factor involved in transcriptional regulation of JA-responsive genes (Kazan & Manners, 2013; Lorenzo *et al.*, 2004; Schweizer *et al.*, 2013; Staswick & Tiryaki, 2004). These three components have been indicated to play a major JA regulation upon plant insect/pathogen attack. For example, in a study by Huang *et al.* (2015), genes that were involved in JA biosynthesis and signalling pathways (MYC2 transcription factors and JAR) were activated upon feeding by cotton bollworm (*Helicoverpa armigera*) in cotton plants. Moreover, JAR1 was activated in response to infestation by the diamond back moth (*Plutella xylostella*) in *Arabidopsis thaliana* (Ehrling *et al.*, 2008), providing support of the important role of JA in response to biotic and abiotic stresses.

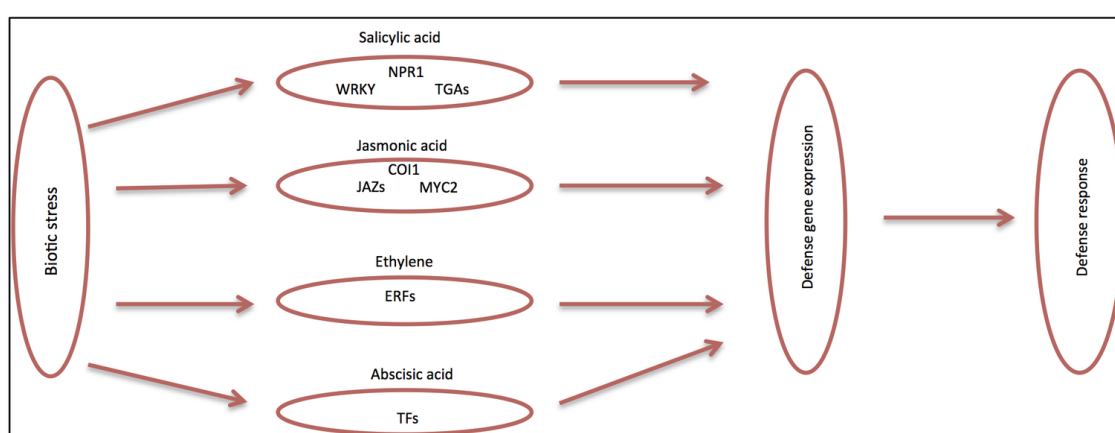


Figure 1.7: A schematic representation of signalling hormones involved in biotic stress defence response.

1.6.1.2 Ethylene (ET)

Ethylene is a gaseous hormone that plays a role in regulating growth and development in plants (Chang *et al.*, 2013; Guo & Ecker, 2004; McManus, 2012; Sakai *et al.*, 1998; Von Dahl & Baldwin, 2007), including leaf senescence and abscission, programmed cell death, and response to stress and pathogen and herbivore attack (Müller & Munné-Bosch, 2015; Ouaked *et al.*, 2003). Both abiotic and biotic stresses regulate ethylene signalling and establishment of SAR (Fu & Dong, 2013). The first observation of ethylene burst due to insect attack was reported by Williamson (1950), and since then, more studies have demonstrated ethylene burst in response to insect attack. In tobacco (*Nicotiana attenuata*), an attack by the specialist herbivore *Manduca sexta* resulted in a dramatic release of ethylene (Winz & Baldwin, 2001). The release of ethylene was also activated in cotton plants in response to infestation by cotton bollworm (*Helicoverpa armigera*) in a study by Huang *et al.* (2015). During insect feeding and wounding, ethylene biosynthesis is regulated by kinases such as mitogen-activated protein kinases (MAPKs) and ethylene response factors

(ERFs; Figure 1.7). Maffei *et al.* (2007) indicated that MAPKs play a role in signalling of abiotic and biotic stresses, pathogens and plant hormones in plants. For instance, in a study by Seo *et al.* (1999), transcripts in tobacco encoding MAPKs began to accumulate only a few minutes after mechanical wounding, representing the presence of ethylene. Furthermore, upon infestation by diamondback moth (*Plutella xylostella*) in *Arabidopsis thaliana*, transcripts encoding MAPKs were also activated (Ehlting *et al.*, 2008), supporting the role of ethylene signalling in response to insect feeding and wounding.

1.6.1.3 Salicylic acid (SA)

Salicylic acid signalling plays an important role in regulating plant responses against biotrophic pathogens as well as phloem-feeding insects (Glazebrook, 2005; Pieterse *et al.*, 2012). It also plays a role in the establishment of SAR and shown through genetic studies (Grant & Lamb, 2006; Ryals *et al.*, 1996; Hilfiker *et al.*, 2014; Luna *et al.*, 2012). SA signalling involves an important regulatory component, *nonexpresser of pathogenesis-related gene 1* (*NPR1*, Figure 1.7). These components interact with TGA transcription factors and are involved in activation of SA-responsive pathogenesis-related (*PR*) genes. In a study by Zhang *et al.* (2015a), feeding by cotton mealybug *Phenacoccus solenopsis* induced the expression of *PR* genes (*PR-1a* and *Glu A*) in tomato. In addition, *PR* genes were differentially expressed during inoculation by *Alternaria solani* in tomato (Upadhyay *et al.*, 2014). On the other hand, *PR* genes associated with SA-mediated signalling were induced by aphid feeding in sorghum (Park *et al.*, 2006) and tomato (de Ilarduya *et al.*, 2003). Expression of these *PR* genes responsive to SA signalling is therefore used as a marker for SA-dependent systemic acquired resistance.

1.6.1.4 JA, ET and SA interactions

Jasmonic acid, ET and SA are involved in plant defence by regulating plant defence responses against biotic and abiotic stresses (Bari & Jones, 2009; Caarls *et al.*, 2015; Dong, 1998; Zhang *et al.*, 2013). In some studies (Beckers & Spoel, 2006; Kunkel & Brooks, 2002; Pieterse *et al.*, 2009; Spoel *et al.*, 2007; Zhang *et al.*, 2013), synergistic interactions between JA, ET and SA have been reported. For instance, Spoel *et al.* (2007) indicated that *NPR1* plays an important role in interactions between SA and JA in *Arabidopsis thaliana* plants. Transcription factors MYC2 and ERF1 appeared to function as important nodes that integrate signals from JA and ET pathways upon herbivory by *Pieris rapae* in *Arabidopsis thaliana* (De Vos *et al.*, 2005; 2006). Furthermore, an interaction between JA and ET mediated defence signalling was regulated by transcription factor MYC2 in *Arabidopsis thaliana* in response to biotic stress (Anderson *et al.*, 2004). Petersen *et al.* (2000) reported the MAPK4 transcription factors as a negative regulator of SA signalling and a positive regulator of JA signalling in *Arabidopsis thaliana*. Moreover, in a study by Li *et al.* (2008), the expression of WRKY70 in *Arabidopsis thaliana* caused enhanced expression of SA-responsive *PR* genes and concomitantly suppressed methyl-jasmonate-induced expression of JA-

responsive marker gene *Plant defensin 1.2* (*PDF1.2*). Some of these regulatory proteins are also known to play a role in shaping the outcome of *Arabidopsis thaliana* defence response.

1.6.2 Indirect defences

In indirect defences (green leaf volatiles, volatile organic compounds, and extra floral nectars), the plant may employ mechanisms such as attractants of predators and parasitoids (Figure 1.6) of herbivores in facilitating the control of herbivores (Agrawal & Rutter, 1998; Krimmel & Pearse, 2013), or supply food and shelter to increase the effectiveness of natural enemies (War *et al.*, 2012). Studies with sap-sucking insects on *Arabidopsis thaliana* (Kuśnierczyk *et al.*, 2008; Zhang *et al.*, 2013) and cotton (Hegde *et al.*, 2011) have revealed that feeding by sap-sucking insects results in the release of parasitoid attractants that can assist in the control of herbivorous pests, including cassava pests. For example, in a study by Zhang *et al.* (2013), β -myrcene was released from *Arabidopsis thaliana* following feeding by whitefly (*Bemisia tabacci*), which resulted in the attraction of the whitefly parasitoid *Encasia formosa*. In addition, a blend of volatiles released by cotton in response to *Aphis gossypii* could trigger the antennal responses of *Chrysoperla lucasina*, a predatory lacewing to cotton aphid (Hegde *et al.*, 2011). Furthermore, the release of cyanogen glycosides from wild cassava leaves infested with mealybug (*Phenacoccus herreni*) appeared to attract female *Anagyrus lopezi* wasp, which is the predator to cassava mealybugs (Nadel & Alphen, 1987).

1.7 Plant-herbivore resistance

Plant defence responses are generally activated at the site of infection or herbivore feeding (Biere & Goverse, 2016; Klauser *et al.*, 2015). For plants to specifically recognize insect or pathogen attack (Biere & Goverse, 2016; Hammond-Kosack & Jones, 1997; Pan *et al.*, 2000), and the corresponding avirulence (*avr*) gene in the pathogen, they require dominant or semi-dominant resistance (*R*) gene alleles. During pathogen attack, the plant *R* protein interacts with the *avr* effectors and the defence responses become stimulated. Functional plant *R* genes that have been isolated so far encode resistance to bacterial, viral, fungal, oomycete, nematodes and insect pests (Dangl & Jones, 2001; Jones & Dangl, 2006). The vast majority of the *R* genes cloned belong to the nucleotide-binding site leucine-rich repeat, extracellular leucine-rich repeat, or leucine-rich repeat kinase superfamilies (Marone *et al.*, 2013; McDowell & Woffenden, 2003; Michelmore & Meyers, 1998; Petit-Houdenot & Fudal, 2017). These superfamilies were identified using map-based cloning techniques in tomato and *Arabidopsis thaliana*.

In plants, several *R* genes that confer resistance to insect herbivores have been identified and they belong to the NBS-LRR domains (Du *et al.*, 2009; Hogenhout & Bos, 2011; Rossi *et al.*, 1998). These *R* genes include the *Vat* gene that confers resistance to cotton aphid *Aphis gossypii* in melon and cotton (Dogimont

et al., 2010; Klingler *et al.*, 2001); *Bph* genes that confer resistance to the brown planthopper *Nilaparvata lugens* in rice (*Oriza sativa*) (Du *et al.*, 2009); and the *Mi-1.2* gene that confers resistance to potato aphid *Macrosiphum euphorbiae* in tomato (Rossi *et al.*, 1998). Furthermore, other classes of R-genes such as the *Rag* gene that confers resistance to the soybean aphid *Aphis glycine* (Kim *et al.*, 2010a; 2010b; Zhang *et al.*, 2009); the *AIN* (*Acyrtosiphon-induced necrosis*) gene that confers resistance to bluegreen aphid *Acyrtosiphon kondii* in *Medicago truncatula*, the *Dn* (*Diuraphis noxia*) genes and *H* genes that confers resistance to hessian fly *Mayetiola destructor* and Russian wheat aphid *Diuraphis noxia* in wheat have also been reported (Dweikat *et al.*, 1997; Klingler *et al.*, 2009; Liu *et al.*, 2005). As yet, no genes that confer resistance to cassava mealybugs or whitefly have been reported. Only *CMD* genes that confer resistance to cassava mosaic disease has been reported from East African cultivars (Akano *et al.*, 2002; Fregene *et al.*, 1997), and *CMD* tolerance in the African landrace TME3 has also been described (Bengyella *et al.*, 2015; Louis & Rey, 2015). With the use of current next generation sequencing tools, plant-insect interactions in cassava can be unravelled to facilitate the discovery of genes involved in multiple pest resistance.

1.8 Improving cassava germplasm for mealybug resistance

Cassava improvement involves breeding for high yield, drought tolerance, and resistance to diseases and pests (Mahungu *et al.*, 2009), and a range of breeding technologies is improving for crop improvement. Conventional breeding using recurrent selection methods (Ceballos *et al.*, 2004) was considered, however, it was discovered that more efforts would be required to introduce traits into different cultivars without altering characteristics of the parent genotype. In cassava, many factors affect its conventional breeding (Jennings & Iglesias, 2002). These factors include high levels of genetic heterozygosity, variable flowering patterns, and poor seed set and germination (Fregene *et al.*, 1997; Jennings & Iglesias, 2002).

Conventional breeding for resistance to diseases and pests has been successfully practiced in cassava (Morales, 2001; Nassar & Ortiz, 2010), however, it takes decades to achieve the products. Breeding for resistance to African Cassava Mosaic Virus (ACMV), transmitted by whiteflies (*Bemisia tabaci*) started in the 1920s in East Africa (Morales, 2001). ACMV tolerant cultivars were identified within the collected germplasm, and crosses were made between the tolerant cultivars, however, high levels of resistance was not achieved. For high levels of resistance to be achieved, interspecific crosses were further made using wild *Manihot* species such as *M. glaziovii*, and then backcrossed again. It took three generations of backcrosses to finally get improved ACMV improved cultivars. In addition, breeding for resistance to mosaic and brown streak viruses was also carried out at the East African Research Institute, in Tanzania (Nichols, 1947). The method used entailed intervarietal hybridization with *M. utilissima* and interspecific hybridization, which included seed production, propagation and field trials of virus resistance. Backcrossing from wild *Manihot* to cassava was performed, until to the third backcross. Virus resistant clones were generated and tested in the field with *M. utilissima* and interspecific hybrids with high levels

of resistance. To overcome the expensive and slow process of traditional conventional breeding methods, new molecular breeding techniques were adapted for use in cassava.

Since molecular breeding methods were introduced, cassava breeding for resistance to diseases and pests has improved. Molecular breeding of cassava has been successful due to major contributions made by technologies in the area of tissue culture, genetic transformation and high throughput next generation sequencing techniques (Ceballos *et al.*, 2004; Lapidot & Friedmann, 2002; Liu *et al.*, 2011; Mardis, 2008a, 2008b; Sharma & Ortiz, 2000; Taylor *et al.*, 2004; Zhang *et al.*, 2005). A genetic linkage map of cassava has provided an initial tool for the genetic analysis of important traits (Fregene *et al.*, 1997). Several genetic markers have since been generated using simple sequence repeats (SSR), random amplified polymorphism DNA (RAPD) and isoenzyme markers, making the genetic linkage map denser (Fregene *et al.*, 1997; Raji *et al.*, 2009). Single nucleotide polymorphism (SNP) markers have also been constructed to make the genetic map denser (Rabbi *et al.*, 2014; Soto *et al.*, 2015).

A Bacterial Artificial Chromosome (BAC) library was then constructed in order to accelerate the cloning of resistance genes. The BAC library was constructed from the clones TMS 30001 while a second one was later made from M Ecu-72 (Tomkins *et al.*, 2004). Apart from the two BAC libraries constructed, a third library was constructed from a TME3 variety. This new variety has a new source of CMD resistance that was attributed to a single dominant gene designated CMD2 (Akano *et al.*, 2002).

Genetic engineering was later introduced to cassava; however, it was considered recalcitrant due to selection problems (Fregene & Puonti-Kaerlas, 2002). Cotyledons or embryogenic cultures have been used as target tissues in many protocols and transformed by particle bombardment or *Agrobacterium*-mediated procedures (Liu *et al.*, 2011). Therefore, the use of embryogenic tissues (i.e. friable embryogenic callus, FEC) in combination with *Agrobacterium*-mediated transformation became the favoured method of cassava transformation because of its higher efficiency compared to the cotyledon-based protocol (Abhary *et al.*, 2011; Bull *et al.*, 2009; Chauhan *et al.*, 2015; Nyaboga *et al.*, 2015). Genetically engineered (transgenic) plants (Bi *et al.*, 2010; Taylor *et al.*, 2004) with additional genetic properties were then produced with these methods, making a contribution to integrated pest management programs.

Molecular breeding for resistance against pests of economical importance through the insertion of different genes was later introduced. *Bacillus thuringiensis* (*Bt*) genes, which produces several insect toxins (Swadener, 1994), are commonly used. Transgenic plants that include *Bt* genes have been generated in *Zea mays* (Koziel *et al.*, 1993), *Glycine max* (Stewart Jr *et al.*, 1996), *Nicotiana tabacum* (De Cosa *et al.*, 2001) and in cotton (*Gossypium hirsutum*) (Lu *et al.*, 2012), and many others. For example, *Bt* cotton that consists of the insecticidal protein Cry expresses resistance to lepidopteran species and has proved to be resistant to

other insect pests in addition to lepidopteran pests (Lu *et al.*, 2012). These studies opened a platform to breeding for multiple pest resistance in cassava. Furthermore, genetically modified (GM) approaches were also adapted for development of genetically engineered cassava through the BioCassava Plus (BC+) programme (Sayre *et al.*, 2011). Genetically engineered cassava with increased nutrients and additional traits (increased level of shelf life, resistance to viral diseases) were addressed, with reports from field trials in Puerto Rico and Nigeria. In addition, more advanced technologies like NGS and SNP genotyping were also introduced to molecular breeding of cassava (Allie *et al.*, 2014; An *et al.*, 2012; Ballén-Taborda *et al.*, 2013; Owiti *et al.*, 2011; Utsumi *et al.*, 2012; Wang *et al.*, 2014).

1.9 Next generation sequencing (NGS)

The introduction of NGS technologies (massively parallel sequencing) has transformed genomic and transcriptomic methods (Davey *et al.*, 2011; Mardis, 2008b; Metzker, 2010; Morozova *et al.*, 2009; Morozova & Marra, 2008). These technologies depend on the several procedures that are grouped together as template preparation, sequencing and imaging, and data analysis. The advantage of NGS technologies is their capabilities to generate a massive amount of data that is less costly (Metzker, 2010), compared to the ABI-Sanger sequencing. Presently, commercial platforms of sequencing systems that are available for high throughput of genomics and transcriptomics research include Roche/454, ABI Sequencing by Oligonucleotide Ligation and Detection (SOLiD) and Illumina (Davey *et al.*, 2011; Mardis, 2008b; Metzker, 2010; Morozova *et al.*, 2009; Morozova & Marra, 2008).

1.9.1 Roche/454 sequencing technology

Roche/454 sequencing technology was the first next generation sequencing technology that was commercially introduced into the market in 2004 after ABI-Sanger sequencing (Morozova & Marra, 2008; Rothberg & Leamon, 2008). This technology uses a pyrosequencing approach, which involves massively parallel sequencing by synthesis of long reads and produces average read length of 400bp- 1000bp (Hall, 2007; Mardis, 2008a, 2008b; Wicker *et al.*, 2009). Pyrosequencing utilizes DNA polymerase incorporating the nucleotides resulting in the release of pyrophosphate molecules. The release of phosphate molecules then recruits a sequence of downstream responses that eventually generate light by the enzyme luciferase (Mardis, 2008a, 2008b). This technology has been employed in genomics and transcriptomics research including in humans (Qin *et al.*, 2010; Worthey *et al.*, 2011), animals (Cahais *et al.*, 2012; Hou *et al.*, 2011) and plants (Varshney *et al.*, 2012; Wicker *et al.*, 2009). For example, in plants, a barley complex genome was sequenced using the Roche/454 technology (Wicker *et al.*, 2009). The technology provided much more coverage of the BAC clones than in ABI-Sanger sequencing. Cassava was also sequenced and assembled using the 454 FLX Titanium sequencing technology (Prochnik *et al.*, 2012). However, the sequencing

platform was discontinued due to the introduction of the new and faster platforms. Since the sequencing of the cassava draft genome, more genomic and transcriptomic studies have been made possible.

1.9.2 ABI SOLiD sequencing technology

The ABI SOLiD system, which was first introduced in 2007, is a program that is based upon the hybridization ligation process (Ansorge, 2009; Mardis, 2008a, 2008b; Morozova & Marra, 2008; Shearer *et al.*, 2011). The system process includes oligo adaptor ligation of DNA fragments, attaching DNA fragments to the tiny magnetic beads, and amplification of the DNA on the beads by emulsion PCR. Within the sequencer, lies a glass slide placed into the fluidics cassette. The beads are then attached to the surface of the glass slide after amplification, and sequenced. The SOLiD system produces read length of up to 75 bp and can construct about 320 Gb of sequences per run. This platform has been applied in plant genomics such as in a study by (Ashelford *et al.*, 2011) a novel circadian clock mutation was revealed by re-sequencing a full genome in *Arabidopsis thaliana* using both ABI SOLiD and Roche 454 NGS technologies. In addition, basal defence and cell wall associated genes of susceptible and tolerant cassava landraces infested with South African cassava mosaic virus were identified by transcriptome analysis using ABI SOLiD technologies (Allie *et al.*, 2014). This technology has proven to unravel genomics and transcriptomics problems, however, it has limiting factors such as being slower, with a time-consuming method for library preparation as compared to the other sequencing technologies (Zhang *et al.*, 2011).

1.9.3 Illumina sequencing technology

The Illumina approach is still the most used system for sequencing since its introduction to the market in 2006 (Mardis, 2008a, 2008b; Zhang *et al.*, 2011). Illumina platforms include MiSeq, HiSeq series and the NextSeq 500 (Buermans & Den Dunnen, 2014; Reuter *et al.*, 2015). The MiSeq is designed for sequencing small genomes, while the HiSeq series are designed for high throughput sequencing applications. In addition, the NextSeq 500 is designed for human genome sequencing. In the present study, the HiSeq 2500 was used for transcriptome analysis. This approach uses bridge amplification for cluster generation and sequencing by synthesis of reads of up to 300 bp in length (Figure 1.8). The system uses a combination of four nucleotides, each classified with a unique fluorophore that are added concurrently to the flow cell channels along with DNA polymerase for integrating into oligo-primed cluster fragments (Egan *et al.*, 2012; Mardis, 2008a; Zhang *et al.*, 2011). Cluster strands are then extended by bridge amplification and sequenced. The HiSeq 2500 is able to produce paired-end reads of 125 bp in length; however, recent sequencer models (NextSeq 500 and HiSeq X Ten) produce even longer reads of about 150 bp in length. To date, this platform is used by most studies including in plant genomics and transcriptomics (Coppola *et al.*, 2017; Egan *et al.*, 2012; Huang *et al.*, 2009; Wu *et al.*, 2013; Yang *et al.*, 2015).

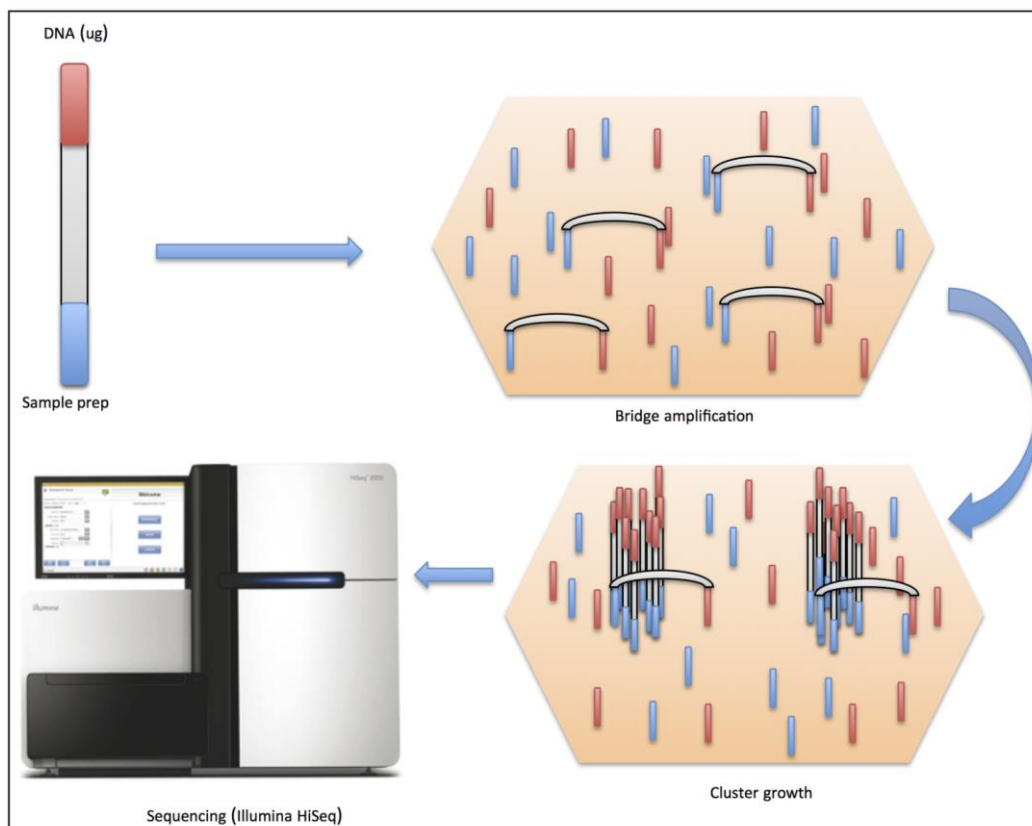


Figure 1.8: An overview of Illumina/Solexa platform workflow. RNA was purified by the removal of rRNA using magnetic kit to produce enriched mRNA. The library is prepared from enriched mRNA, which is fragmented, adapter ligated and PCR amplified to enrich purified cDNA template. cDNA library is added to the flow cell for bridge amplification. Fragmented clusters are subjected to denaturation, annealed with the sequencing primer and sequenced using an Illumina instrument to produce 125 bp paired-end sequencing reads (Mardis 2008a, 200b).

1.9.4 Single molecule sequencing

Although the first and second-generation sequencing platforms are still used in the agricultural industry, there are new third generation sequencing platforms that have recently been introduced (Greninger *et al.*, 2015; Lu *et al.*, 2016; Schadt *et al.*, 2010; Rhoads & Au, 2015; Schneider & Dekker, 2012). These platforms use the single-molecule sequencing by degradation or direct physical inspection of the DNA molecule, with higher throughput, longer read lengths, smallest amount of starting material, low cost and faster turnaround time (Greninger *et al.*, 2015; Schadt *et al.*, 2010; Schneider & Dekker, 2012), compared to the first and second-generation sequencing platforms. The platforms include Pacific BioSciences (Rhoads & Au, 2015) and Oxford Nanopore MinION (Lu *et al.*, 2016) third generation sequencing platforms.

Compared to the second-generation sequencing platforms, Pacific BioSciences (PacBio) and Oxford nanopore platforms have an advantage of producing longer read lengths, however, the reads have high error rate. Circular consensus sequencing can be done to reduce the error rates (Larsen *et al.*, 2014), thereby

overcoming the repetitive regions of the second-generation sequencing. Various studies have demonstrated the potential of the nanopore sequencing, including in the assembly of the genome of *Arabidopsis thaliana* and *Drosophila melanogaster* (Berlin *et al.*, 2014). In spite of that, the first and second-generation sequencing platforms have led the way in transforming the genomics field, opening a platform for the development of the current third generation sequencing, and they are still used.

1.9.5 Application of NGS in cassava breeding

DNA-based molecular markers are useful for detecting the genetic variation available in germplasm collections and breeding lines. However, with the use of next generation sequencing technologies, it is possible to resequence the entire plant genome and the entire transcriptome more efficiently at a greater depth than before (Varshney *et al.*, 2009). Genome and transcriptome sequence data of model plants and other crop species such as *Populus trichocarpa* (Tuskan *et al.*, 2006), *Medicago truncatula* (Tang *et al.*, 2014), *Zea mays* (Schnable *et al.*, 2009) and *Arabidopsis thaliana* (Meyerowitz & Somerville, 1994) have been sequenced using the NGS technologies. These platforms offer high throughput with reduced sequencing costs as compared to Sanger sequencing (Varshney *et al.*, 2009). Cassava is one of the crop species that has been sequenced using NGS platforms, Roche/454; and has recently been re-annotated (Bredeson *et al.*, 2016). The sequencing and assembly of the cassava genome (Table 1.1) is invaluable in the application of molecular technologies towards cassava improvement, and this include transcriptome profiling of genes associated with resistance to herbivorous pests.

Table 1.1: Current status of the cassava reference genome: *Manihot esculenta* v6.1.

Database title	Existing data
Genome size (approximately)	582.25 Mb (~15 % gaps)
Contig	40 044
Scaffolds	2 001
Chromosomes	18 (not yet anchored on chromosome)
Loci	30 033
Alternate transcripts	8 348

1.9.6 Transcriptome profiling

A transcriptome is defined as set of all coding and non-coding RNA molecules produced in a population of cells (Wise *et al.*, 2007). Analysis of a transcriptome involves characterization of all transcriptional activity without prior knowledge of the genome. Advances in sequencing enable rapid profiling and deep investigation of the whole or a portion of a transcriptome (Wise *et al.*, 2007). Transcriptome studies can be

achieved using genome-guided tools (Kim *et al.*, 2013; Trapnell *et al.*, 2012) or *de novo* assembly tools (Birol *et al.*, 2009; Chen *et al.*, 2010; Grabherr *et al.*, 2011; Haas *et al.*, 2013; Robertson *et al.*, 2010).

In most plant species, microarrays have been the most used technology for transcriptome profiling (Meyer *et al.*, 2008; Meyers *et al.*, 2004; Park & Ryu, 2014; Studham & MacIntosh, 2013). Microarrays were the first platform to be used for expression profiling since its emergence (Schena *et al.*, 1995). This platform can be used to study the expression of genes simultaneously, however, it only measures expression levels of genes for which a probe is available (Rensink & Buell, 2005). A number of expressed profiles have been studied using this technique in plant defence against insect pests (Couldridge *et al.*, 2007; Ehrling *et al.*, 2008; Gutsche *et al.*, 2009). Other technologies that do not utilize probes such as Illumina/Solexa, ABI/SOLiD, and 454/Roche have been introduced for expression profiling purposes, even through *de novo* sequencing (Mardis, 2008a, 2008b; Zhang *et al.*, 2011). The availability of technologies like these has proven the ability to measure the expression of thousands of genes in plant species (Chen *et al.*, 2012; Wall *et al.*, 2009; Yang *et al.*, 2012) in response to both biotic and abiotic stresses, which has presented scientists with new opportunities.

The availability of a complete genome of *Arabidopsis* made it the first plant for which transcriptional profiling platforms were developed (Meyer *et al.*, 2008; Meyers *et al.*, 2004). Sequences of other crop species such as poplar (Tuskan *et al.*, 2006), *Medicago truncatula* (Cannon *et al.*, 2005), tomato (Tomato Genome Consortium, 2012) and maize (Schnable *et al.*, 2009) are also available. Only few transcriptome profiling studies have been performed in cassava (Allie *et al.*, 2014; Ballén-Taborda *et al.*, 2013; Bohorquez *et al.*, 2012; Fregene *et al.*, 2004; Lopez *et al.*, 2005b; Owiti *et al.*, 2011; Utsumi *et al.*, 2012; Wang *et al.*, 2014), however, stress-associated genes with biological functions were identified. To date, genes associated with pest resistance in cassava have not been reported. This study will address part of these concerns by identifying putative candidate defence genes involved in cassava mealybug resistance by RNA-seq transcriptome profiling.

2.0 MOTIVATION AND RATIONALE FOR STUDY

Cassava is a woody perennial plant that feeds over 850 million people living throughout the tropics. It is mainly cultivated for its large starchy roots but the leaves are also consumed as a vegetable, especially in Africa. In South Africa, cassava is grown in Limpopo, Mpumalanga and KwaZulu-Natal provinces as a secondary crop. In the neighbouring countries of Swaziland and Mozambique, cassava is a staple food crop and is also used for industrial starch production. Cassava has also gained importance as an industrial starch in the pharmaceutical, textile and cosmetics industry. Furthermore, it is also used in bio-ethanol production and has been regarded as one of the most important crops in bio-ethanol production. However, the pests of cassava (*P. manihoti*) can devastate the crop resulting in up to 80% yield losses. Several approaches such

as biological control using parasitoids have been implemented with varying degrees of success in central and western Africa, however, cultivating resistant varieties remains the most sustainable option for managing the pests.

Cassava breeding over the years mainly focused on high yields and low fibre content with very little emphasis on pest resistance (Morante *et al.*, 2004). The significance of yield reduction due to pests has led to the inclusion of breeding for resistance to pests as a major objective in cassava breeding programs. Basic techniques used in plant pathology and genetics were used to decipher the different mechanisms in response to pathogen and insect attacks (Ceballos *et al.*, 2004; Fregene *et al.*, 1997; Liu *et al.*, 2011; Nassar & Ortiz, 2010). However, with the recent advances in NGS sequencing (second and third generation), molecular breeding of cassava for pest resistance can be more effective. Understanding the interactions between the insect pest (cassava mealybugs) and the plant host (cassava) will be important for elucidating the mechanism and nature of cassava resistance to mealybug. The subsequent identification and development of cassava genotypes with resistance to mealybug will go a long way in contributing to sustainable cassava production in sub-Saharan Africa.

2.1 AIM

The aim of the project was to identify genes and molecular pathways associated with cassava response to mealybug (*Phenacoccus manihoti*) through transcriptome profiling of cassava genotypes infested with mealybugs.

2.2 HYPOTHESES

- i. Barcoding genes correctly identify the various species of mealybugs present in South Africa.
- ii. Different cassava genotypes respond differently to mealybug infestation and feeding.
- iii. Cassava plants respond to mealybug infestation by induction or suppression of different classes of transcripts and gene pathways depending on their level of resistance and/or susceptibility.

2.3 SPECIFIC OBJECTIVES

- i. To characterize mealybug collections from different cassava-growing provinces in South Africa using barcoding techniques.
- ii. To phenotypically screen diverse cassava genotypes for their response to cassava mealybug infestation.
- iii. To identify candidate genes and determine gene expression pathways in response to mealybug infestation in resistant and susceptible cassava genotypes.

CHAPTER 2

MATERIALS AND METHODS

2.1 Mealybug collection and maintenance

Colonies of cassava mealybugs were collected from cassava plants at Makhathini (27.3953° S, 32.1764° E; KwaZulu Natal province), Malelane (25.4861° S, 31.5484° E) and Bushbuckridge (24.8398° S, 31.0464° E; Mpumalanga province). The mealybugs were collected randomly from cassava growing farms by harvesting branches from mealybug infested plants. Branches with mealybugs were then placed on cassava plants in pots in separate cages to allow the pests to feed. The collected mealybugs were transported to the Agricultural Research Council–Vegetable and Ornamental Plants Institute (ARC-VOPI) and maintained as independent colonies on cassava plants under greenhouse conditions at 27 °C day and night temperature. A sample of each mealybug collection was further preserved in 100% ethanol and stored at room temperature for morphological and molecular identification. The South African Authority (South African National Biodiversity Institute) does not require any permission for scouting and/or collection of pests of any kind either from open fields or from farmers' fields. The current study did not involve endangered or protected species.

2.1.1 Morphological identification of mealybugs

All solutions, reagents, chemicals and kits used in this study are listed in Appendix A1-A3. From the field samples (branches with multiple mealybugs) collected from several farms in two provinces, 15 adult female mealybug (five specimens per site: Makhathini, Malelane and Bushbuckridge) were prepared for morphological identification according to the protocol of (Williams & de Willink, 1992) with slight modification as follows: Mealybug specimens were incised with a needle in the middle of the dorsum of the thorax and incubated in 10% potassium hydroxide overnight. Specimens were heated in a tube in boiling water until the body contents were pale and translucent, and then soaked in distilled water for three hours. The specimens were then rinsed with 96% ethanol and incubated for a further 10 minutes in the same solution before transferring to carbol xylene for another 10 minutes to dissolve any residual body wax. Specimens were placed in 96% alcohol again to remove the carbol xylene, and stained in acid fuchsin for 1 h before dehydrating in acid alcohol for 30 minutes. The dried specimens were placed on a slide, covered with a drop of clove oil for 20 min, and then mounted on microscope slides with Canada balsam for identification. Mealybugs were identified according to species keys described by (Williams & de Willink, 1992).

2.1.2. Confirmation of mealybug identity

Total DNA was extracted from individual intact specimens (three specimens from three different field sites) using the blood and body fluid protocol of the DNeasy DNA extraction kit (Qiagen, Valencia, CA), with modifications according to Sethusa *et al.* (2014). DNA concentrations were quantified using a Qubit® fluorometer (Invitrogen/Life Technologies Inc., Carlsbad, CA, US) and verified on a 1% agarose gel in 1xTAE buffer at 10 V/cm for 30 min using a BioRad Geldoc system (BioRad laboratories Inc., CA). DNA was stored at 4 °C for PCR amplification. Barcoding primers of the 28S ribosomal RNA (rRNA) region, s3660 (F- 5'GAGAGTTMAASAGTACGTGAAAC3') (Dowton & Austin, 1998) and 28S (R- 5'TCGGAAGGAACCAGCTACTA3') (Whiting *et al.*, 1997); 18S rRNA region, 2660 (F- 5'CTGGTTGATCCTGCCAGTAG3') (Tautz *et al.*, 1988) and 18S-B (R- 5'CCGCGGCTGCTGGCACCAGA3') (von Dohlen & Moran, 1995); and the CO1 region, PcoF1 (5'CCTTCAACTAATCATAAAAATATYAG3') and LepR1- (5'TAAACTTCTGGATGTCCAAAAAATCA3') (Park *et al.*, 2010b) were used to amplify genomic regions (18S and 28S rRNA; and CO1 region) in mealybug samples. Gene fragments were amplified using a touchdown protocol that included initial denaturation at 94 °C for 5 min followed by 14 cycles of 94 °C for 30 sec, 56 °C for 30 sec and 72 °C for 2 min, with 1.0 °C decreasing temperature from 56 to 42 °C for 30 sec every cycle. The remaining 18 cycles were performed at 94 °C for 30 sec, 42 °C for 30 sec and 72 °C for 2 min, with the final elongation at 72 °C for 7 min. PCR products were purified using the QIAquick PCR purification kit (Qiagen, Valencia, CA). After PCR amplification, the products were verified on a 1% agarose gel electrophoresed in 1xTAE buffer using the BioRad Geldoc system (BioRad laboratories Inc., CA).

Sanger sequencing of PCR products was outsourced to Inqaba Biotec (Inqaba Biotechnical Industries, Gauteng, South Africa). Resultant sequences were assembled using Pregap4 and Gap4 programs within the molecular biology software STADEN package (Staden *et al.*, 1998). Consensus sequences were subjected to BLASTn on NCBI (www.ncbi.nlm.nih.gov/BLAST) and EMBL (<https://www.ebi.ac.uk/Tools/sss/ncbiblast/nucleotide.html>) to retrieve known homologous sequences as species identification references. CLC Bio Genomics workbench v9.0 was used for multiple sequence alignment and the molecular evolutionary genetics analysis package (MEGA v.6.0) for phylogenetic analysis. Non-parametric maximum likelihood (ML) Tamura-Nei bootstrapping model with heuristic searches of 10,000 replications was performed to assess branch support in phylogenetic trees generated. Under the Tamura-Nei model, branch lengths and substitution rate parameters are optimized to fit the data submitted to construct trees. The method is statistically well founded and leads to higher accuracy when constructing trees. Values of 95% or larger were regarded as evidence that the groupings were of the same species, while values of 70% - 94% were regarded as groupings of a related species.

2.2 Cassava collection, multiplication and hardening

Cassava genotypes AR23.1 (CIAT), P40/1 (South Africa) and 98/0581 (IITA) were sourced from ARC-VOPI (Table 2.1) and multiplied through tissue culture in Murashige and Skoog (MS) medium (Murashige & Skoog, 1962) in a growth room at 21 °C (16 h light and 8 h darkness) until the plants developed roots and leaves for 7 weeks (Figure 2.2). *In vitro* plantlets were later transferred to 20 cm diameter pots under greenhouse conditions at 27 ± 2 °C day and night temperature, grown in a medium mix of 60% topsoil, 20% sand and 20% vermiculite (Figure 2.1). Plants were watered once daily, and supplemented with nutrients (Multifeed, Makro, SA) for a period of seven weeks.

Table 2.1: Cassava genotypes selected for infestation and sequencing experiments.

Genotype	Origin	Source	Other information
AR23.1	CIAT, Colombia	ARC-VOPI	Stay green, grow fast in vitro and in greenhouse, multiple disease and green mite resistance
P40/1	ARC, South Africa	ARC-VOPI	No known resistance
98/0581	IITA, Nigeria	ARC-VOPI	Multiple pest resistance, CMD resistance

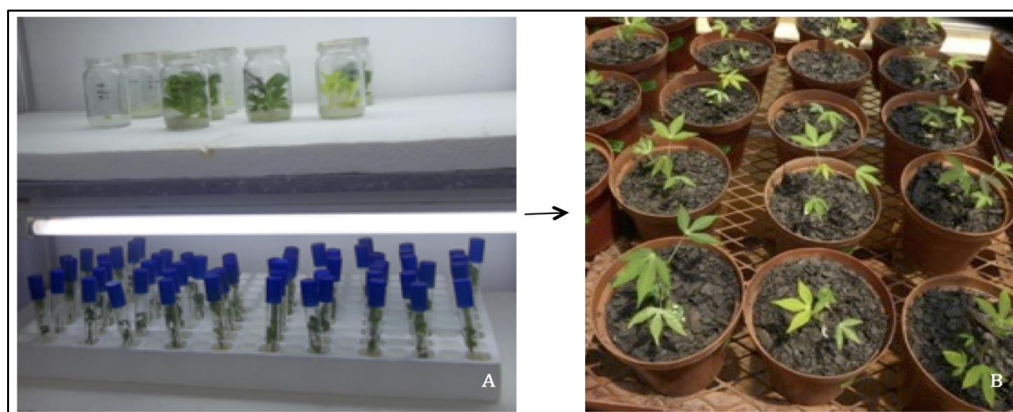


Figure 2.1: Cassava plants multiplied in the growthroom and greenhouse for rearing and mealybug infestation experiments. Plants were multiplied under controlled temperature. A: *In vitro* plants in a growth room in bottles and tubes at 21 °C (16 h light and 8 h darkness). B: Greenhouse plants hardened in a soil mix at 27 ± 2 °C day and night temperature.

2.3 Mealybug infestation experiments and scoring

Plant leaves that were seven weeks old were infested with the third stage instar mealybugs. Nine plants were divided into groups of three, each representing a biological replicate (for each genotype and each time

point). Plants were prepared separately for both phenotypic evaluation and RNA-sequencing under the same conditions. Two fully expanded leaves (directly below the apex) were infested with 15 mealybugs by gently brushing the insects on the leaves using a fine paintbrush. Pots with plants infested with mealybugs were covered with plastic bags to protect the mealybugs from moving from one plant to the other. Non-infested controls were plants from which leaf tissues were harvested at zero-time point prior to infestation. Mealybugs were removed from the infested leaves using a paintbrush before harvesting plant tissues for RNA-sequencing. Leaves infested with mealybugs or mock-infested (leaves were wiped with sterile distilled water using a sterile paintbrush) were harvested at 24 and 72 hours of mealybug infestation (Table 2.2).

Table 2.2: Description of experimental samples.

Tissue	Description	Genotype	RNA-seq dataset
Non-Infested tissue harvested at 24 h	Non-infested mock	AR23.1	AR23-24-mock
Infested tissue harvested at 24 hpi	Experimental tissue	AR23.1	AR23-24-inf
Non-Infested tissue harvested at 72 h	Non-infested mock	AR23.1	AR23-72-mock
Infested tissue harvested at 72 hpi	Experimental tissue	AR23.1	AR23-72-inf
Non-Infested tissue harvested at 24 h	Non-infested mock	P40/1	P40-24-mock
Infested tissue harvested at 24 hpi	Experimental tissue	P40/1	P40-24-inf
Non-Infested tissue harvested at 72 h	Non-infested mock	P40/1	P40-72-mock
Infested tissue harvested at 72 hpi	Experimental tissue	P40/1	P40-27-inf
Non-Infested tissue harvested at 24 h	Non-infested mock	98/0581	0581-24-mock
Infested tissue harvested at 24 hpi	Experimental tissue	98/0581	0581-24-inf
Non-Infested tissue harvested at 72 h	Non-infested mock	98/0581	0581-72-mock
Infested tissue harvested at 72 hpi	Experimental tissue	98/0581	0581-27-inf

Leaf tissues were pooled from each plant per replicate (6 leaves per biological replicate), snap frozen in liquid nitrogen and stored at -80 °C. The plants were further monitored weekly post-infestation (7, 14, 21, 28 and 35 dpi) to assess and record the extent of damage and the number of mealybugs per surviving plant.

Screening and scoring were done according to the method of the Collaborative Study of Cassava in Africa (Nweke *et al.*, 1999) on the 7th, 14th, 21st, 28th and 35th days post-infestation by comparing mealybug-infested with mock-infested leaves. Symptom severity was evaluated per plant on a scale of 0 (resistant, no observable symptoms) to 4 (a candlestick-like appearance, internode length reduced, curved/complete defoliation of the young portion of shoots, ≥ 200 mealybugs and severe symptoms). Plants that had ≥ 100 and ≤ 200 mealybugs with moderate symptoms were scored as moderately susceptible (score 3) while those with ≥ 50 and ≤ 100 mealybugs were scored as moderately resistant/tolerant (score 2). Plants with less than 50 mealybugs, and slight symptoms were scored as partially resistant (score 1).

2.4 Total RNA isolation

2.4.1 Treatment of equipment and glassware for RNA isolation

Pestles and mortars, spatulas and tweezers that were used for RNA preparations and isolations were treated with DEPC in water, then autoclaved at 121 °C for 15 minutes to sterilize. After autoclaving, the instruments and glassware were subjected to 80 °C dry heat to dry before being utilized. Equipment such as gel preparation trays, combs, gel running tank and gel doc were also treated with DEPC water before being used.

2.4.2 RNA isolation and quantification

Leaf tissue (100 mg) from infested and non-infested (mock) plants from each biological replicate and each time point were ground in liquid nitrogen. In each biological replicate (at each time point), three plants were used and from each plant, two leaves (directly below the apex) were pooled. Biological variation was represented by 6 leaves per biological replicate. RNA was isolated using the RNeasy Plant Mini kit (Qiagen, Valencia, CA) according to the manufacturer's instructions. RNA was digested with DNase using the RNase-Free DNase kit (Qiagen, Valencia, CA), subsequently eluted in 50 µl RNase-free water, and stored at -80 °C until ready for use. RNA concentration was quantified using a Qubit® fluorometer. Agarose gel electrophoresis was used to view RNA quality, the electrophoresis buffer was 1xTBE and electrophoresis was at 10 V/cm for 30 min using a BioRad Geldoc system (BioRad laboratories Inc., CA). A RiboRuler RNA ladder was used as molecular size marker (RiboRuler, Inqaba Biotec, SA). Ethidium bromide stained images were captured on a BioRad Gel/Chemi doc (BioRad laboratories Inc., CA) and BioRad image lab software 5.1. After quantifying concentrations using a Qubit® fluorometer, all RNA samples were adjusted to the same concentration for reverse-transcription reactions and sequencing.

2.5. Library preparation and RNA sequencing

Total RNA was purified by the removal of rRNA using oligo dT beads to produce enriched mRNA, and was prepared for sequencing using a TruSeq Stranded mRNA Sample Preparation Kit (Illumina, San Diego, CA, USA). First strand cDNA was synthesized using reverse transcriptase and random hexamers followed by the second strand synthesis. The cDNA fragments were 3' adenylated followed by adapter ligation with RNA adapter indexes. PCR amplification was performed to enrich purified cDNA template. The cDNA concentration of the library was determined using a Qubit® fluorometer according to the manufacturer's instructions, and the library was further validated on a 1% 1xTAE agarose gel. Library dilutions were prepared and loaded on a cBOT system (Illumina, San Diego, CA, USA) for cluster generation and sequenced using an Illumina HiSeq 2500. Each biological replicate (3 per time point) was sequenced separately for each genotype AR23.1, P40/1 and 98/0581 at 24 and 72 hpi as well as for mock-infested. An RNA-seq workflow for the methods is shown in Figure 2.2.

2.5.1 Methodology workflow

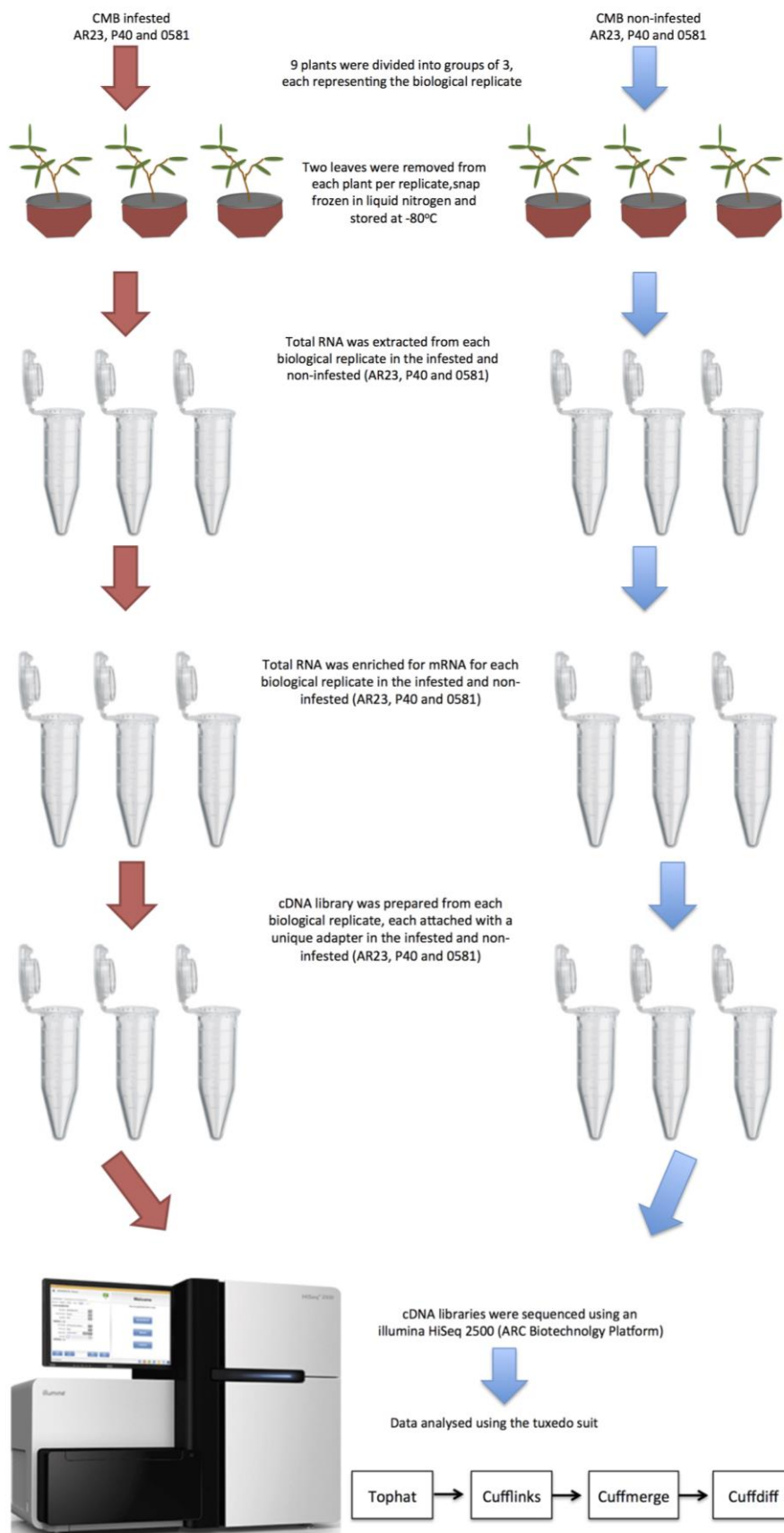


Figure 2.2: Schematic representation of the experimental approach for RNA isolation and cDNA preparation for Illumina RNA-seq from biological replicates of cassava genotypes infested with mealybugs and non-infested mock. This figure represents the workflow for one biological replicate from one genotype at one time point. This method was repeated for all the genotypes at different time-points on leaves infested with mealybugs compared to non-infested mock.

2.6 Bioinformatics analysis

2.6.1 Read mapping and gene expression quantification

The quality of sequencing data was assessed using the FastQC v0.10.1 tools (Bianchi *et al.*, 2013). Sequences were trimmed for quality and adaptors removed using Trimmomatic v0.30 software (Bolger *et al.*, 2014). The sequences used for further analysis had a Phred score with a confidence level above 20. Trimmed transcript sequences were mapped to the annotated cassava reference genome (Mesculenta_305_v6.1) and quantified using Tophat/Cufflinks (Birol *et al.*, 2009) v2.0.11 pipeline (Kim *et al.*, 2013; Trapnell *et al.*, 2012). Differential gene expression in response to mealybug infestation was analysed using the Cuffdiff tool (Trapnell *et al.*, 2012). Gene expression counts were normalized using fragments per kilobase of exon per million mapped reads (FPKM). Gene expression data from the biological replicates of each genotype was compared to determine the positive correlation within the same time-point using MS Excel (Appendix: Figure A9-A11). Since the three biological replicates had no significant variation among them, data from three biological replicates were pooled together per time-point for each genotype before analysis of differentially expressed genes. Differentially expressed genes (DEGs) were determined with a $\log_2$ fold change ≥ 2.0 cut-off and an absolute *P*-value of ≤ 0.05 (Appendix: Table A2-A4). Differentially expressed genes were assigned functions using the *Mesculenta_305_v6.1* annotations, according to their gene ID, and compared to public databases (NCBI BLASTx) and literature from RNA-seq analysis post-infestation with herbivores (Artico *et al.*, 2014; Ehling *et al.*, 2008; Huang *et al.*, 2015; Park *et al.*, 2006; Wei *et al.*, 2013). The raw sequence data was submitted to the NCBI Small Read Archive (accession number SRP058748).

Gene ontology (GO) analysis was conducted on the differentially expressed genes using InterProScan (<http://www.ebi.ac.uk/Tools/pfa/iprscan/>) with default parameters. The GO terms associated with each DEG were then obtained to describe biological processes, molecular functions and cellular components. InterProScan output file was input into BGI WEGO program and GO annotations were plotted (<http://wego.genomics.org.cn>) (Ye *et al.*, 2006).

2.6.2 Pathway analysis

Pathway analysis was performed using the Kyoto Encyclopedia of Genes and Genomes (KEGG) annotation service (Kanehisa *et al.*, 2004; Moriya *et al.*, 2007). KEGG Ontology (KO) terms were assigned to DEGs using the *Manihot esculenta* annotation file (www.phytozome.net). Genes that were assigned KO terms were then mapped to KEGG pathways and significantly enriched terms were further identified in comparison with the castor bean (*Ricinus communis*) genome background.

2.7 RT-qPCR validation of DEGs

Nine (9) candidate defence-related genes (Table 2) were selected from the DEGs for validation using quantitative PCR. Three biological replicates per time point were used for RT-qPCR confirmation analysis. In each biological replicate (at each time point), three plants were used and from each plant, 2 leaves (directly below the apex) were pooled. For each pooled sample there were three technical replicates. Therefore, biological variation was represented by 6 leaves per biological replicate. Primer pairs were designed using CLC Bio Genomics Workbench v9.0, with the amplicon length less than 200 bp and melting temperature between 50-60 °C. Ubiquitin 10 (UBQ10) and GTP binding (GTPb) (Moreno *et al.*, 2011) were used as housekeeping genes for the normalization of the RT-qPCR data.

Table 2.3 Primer sequences of selected DEGs used for real-time quantitative reverse transcription-polymerase chain reaction (RT-qPCR).

Genes	Primer names	Forward (5'-3')	Reverse (5'-3')	Product size
1. <i>Ubiquitin 10</i>	UBQ10	tgcacatcgttctccgattg	gcgaagatcagtcgttggtgc	166 bp
2. <i>GTP binding</i>	GTPb	ggtgccttcttttgcgtttct	gcaatttgatccggtttccat	114 bp
3. <i>2-Oxyglutarate (2OG) and FE(II)-dependent oxygenase superfamily protein</i>	2OG-FE(II)	ggagttgaggttgctgaaga	ggaaggaggagaagaagaaga	146 bp
4. <i>HSP20-like chaperones superfamily protein</i>	HSP20	aaaaacccacaagcccaa	ctctcactttcttctcatcttc	190 bp
5. <i>Myb domain protein 4</i>	MYB	ggtgaggggtggtttaagt	ctctagatttgcgtttgatg	161 bp
6. <i>Glycosyl hydrolases family 32</i>	GH32	gattaggcgttgaagaag	gaacaggagctgaagcta	145 bp
7. <i>Cytochrome P450, family 94, subfamily B, polypeptide 3</i>	CYP450	actcccacttatcctacac	caacgaaaccacaaccaa	120 bp
8. <i>Alpha/beta-hydrolases protein family</i>	Alpha/B	cactctctcttctacttcttct	ttttcccttgtggccctt	179 bp
9. <i>GDSL-like Lipase/Acylhydrolase family protein</i>	GDSL	ggttttgtaatggagtgg	tattttgggtgtgagagg	176 bp
10. <i>Terpene synthase 21</i>	Terp	tcttcttcttctcttctctc	tccccacttacaatcccc	141 bp
11. <i>Gibberelin 2-oxidase 6</i>	Gibb	tgcacatcgttctccgattg	catggatcagtcgttggtgc	150 bp

Reverse transcription (RT) of RNA was achieved using iScript Advanced first strand cDNA synthesis kit (Bio-Rad Laboratories Inc., Hercules, CA, USA) in a total reaction volume of 20 µl containing 1 µg of total RNA according to the manufacturer's instructions. Complementary DNA (cDNA) concentrations were quantified using a Qubit® fluorometer, and equal concentrations (40ng) were prepared for RT-qPCR reactions. Quantitative RT-qPCRs were performed in a Roche LightCycler® 96 System (Roche Applied Science, US) using the KAPA SYBR FAST Universal qPCR kit (KAPA Biosystems, SA) according to the manufacturer's instructions. Thermal cycling conditions were as follows: 94°C for 10 min, followed by 40 cycles of 94°C for 30 sec, 50-60°C for 30 sec (depending on the annealing temperature of each primer set), and 72°C for 30 sec. Each qPCR reaction was performed in triplicate with negative controls. Fluorescence values were acquired at the end of the elongation phase. A melting curve analysis was performed after the final PCR cycle to verify primer specificity. Transcript levels were estimated using the relative quantification method, as described by Livak & Schmittgen, (2001) and Pfaffl, (2004) by calculating the ratio of relative expression from the efficiency (E) and threshold cycle (C_t) values of an unknown sample versus a control sample. The stability of the reference genes across all samples was determined using BestKeeper statistical algorithm (Pfaffl *et al.*, 2004).

CHAPTER 3

Identification of mealybug collections from cassava in KwaZulu Natal and Mpumalanga provinces of South Africa.

AIM – To identify and confirm mealybug species collected from different cassava-growing provinces in South Africa using molecular techniques.

3.1 Introduction

Mealybugs are small sap-sucking herbivorous insects belonging to the Pseudococcidae family (Hardy *et al.*, 2008; Park *et al.*, 2010a). They are serious agricultural insects that cause damage to food crops and ornamental plants (Miller *et al.*, 2005). These insects can be generalized or specialized depending on the kind of host they feed on. Generalist herbivorous insects feed on different plant species of different families, whilst specialists feed on only one or a few plant species of the same family (Ali & Agrawal, 2012; Mello & Silva-Filho, 2002). In cassava, two mealybug species of the genus *Phenacoccus* have been identified to be specialized on cassava as host, namely: *Phenacoccus herreni* Cox and Williams and *Phenacoccus manihoti* Matile-Ferrero (Bertschy *et al.*, 1997; Calatayud & Le Rü, 2006). *P. herreni* is known to occur only in South America, while *P. manihoti* is labelled an African species (Herren & Neuenschwander, 1991).

Mealybugs are known to transmit viruses to plants causing long term damage (Douglas & Krüger, 2008; Millar *et al.*, 2002). These viruses include grapevine virus A and B in grapevine among other viruses. However, no known viruses have been reported that have been transmitted by cassava mealybugs in cassava. These mealybugs mainly colonize the *Manihot* species (Calatayud & Le Rü, 2006) by carefully piercing the plant with stylet and injecting the toxic saliva into a plant while sucking the sap from tissues (Bellotti *et al.*, 1999). This results in leaf yellowing, curling and cabbage-like malformation, leaf necrosis, stem distortion, shoot death and may even kill the plant if the insects are not controlled (Bellotti, 2002; Parsa *et al.*, 2012).

Cassava mealybug is one of the major insect pests of cassava and has been reported to cause a global yield reduction of about 30-80% (Bellotti, 2002). In South Africa, there are no governmental reports on the incidences of mealybug damage on cassava plants. However, the small-scale farmers in Mpumalanga (lowveld) and KwaZulu Natal provinces have reported (no formal reports) distortion of the tubers by mealybugs over the past few years. Therefore, identification of mealybug species that damage cassava in South Africa is important for control purposes.

Mealybugs can be identified traditionally by morphological differences or molecular methods. Traditional identifications rely on slide-mountings of female adult specimens and inspection under the microscope (Malausa *et al.*, 2011; Millar, 2002). However, morphological identification cannot always distinguish between closely related taxa. To supplement the traditional morphological identification methods, molecular methods can be used as complement method (Beuning *et al.*, 1999; Sethusa *et al.*, 2014; Weeto *et al.*, 2004). In this study, we used combined morphological and molecular methods to confirm mealybug species collected from KwaZulu Natal (KZN) and Mpumalanga (MP) provinces.

3.2 Results

3.2.1 Morphological identification of mealybugs

Morphological examination and comparison with the species key confirmed the identification of only one species of the *Phenacoccus* genus, *Phenacoccus manihoti* (Matile-Ferrero). Important diagnostic features, such as the position of the pores, ducts and setae on the dorsum (Figure 3.1), matched the distribution patterns of the South African structures as depicted in the Williams & de Willink, (1992) illustration, confirming the morphological identity of mealybugs in South Africa.

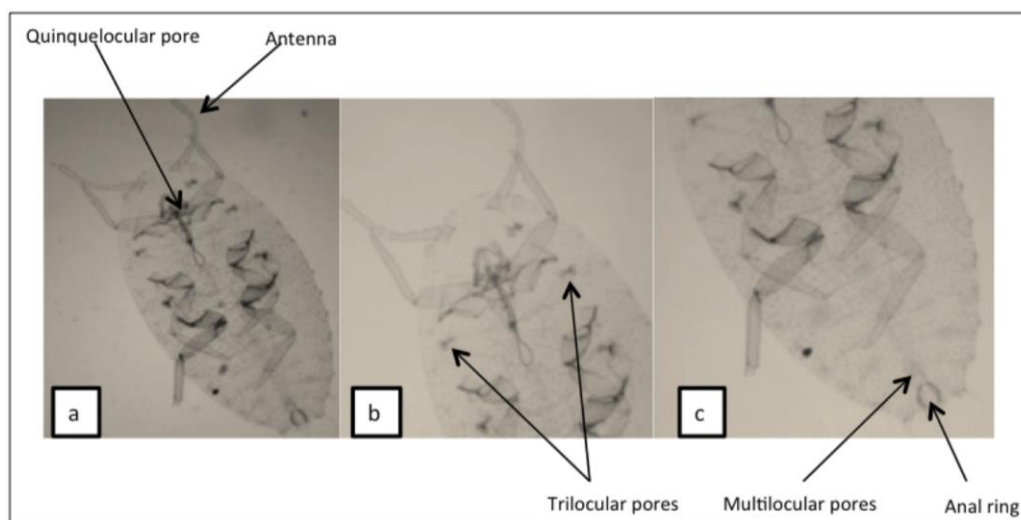


Figure 3.1: Cassava mealybugs mounted and stained on a slide under the microscope. This figure represents one species of the mealybug. Morphological analysis was done on all the mealybug specimens submitted and were similar in morphology. Figure 3.1 a, b and c shows mature female mealybug with body features. These features are used to distinguish the species of the mealybug from the other *Phenacoccus* species family. Only one specimen was used to show the features of the adult mealybug species.

The following morphological features were observed during the analysis of the mealybug specimens:

- i. Multilocular disc pores were situated near the margins of the dorsum, but absent on the dorsum.
- ii. Quinquelocular pores were present on the venter.
- iii. Specimens had 18 pairs of cerrari and a circulus.
- iv. Specimens had two tubular ducts, without translucent pores from the hind tibiae.
- v. Multilocular disc pores were present with tubular ducts on the dorsal margins.
- vi. Quinquelocular pores anterior to clypeolabral shield were 38 in number.

3.2.2 Molecular identification of mealybugs

3.2.2.1 DNA isolation and PCR analysis

DNA was isolated from mealybug samples and 20-30 ng/ μ l was obtained per sample (Figure 3.2 A). DNA concentrations were quantified with a Qubit® fluorometer and viewed on an agarose gel, and then subjected to PCR to amplify regions from the 18S and 28S rRNA and *COI* genes as explained in Chapter 2.1.2. The amplification of 18S and 28S rRNA regions resulted in the expected PCR products, which were sequenced (Figure 3.2 B and 3.2 C). However, the *COI* region could not amplify on all the mealybug samples from KZN and MP. The *COI* region was amplified repeatedly without success.

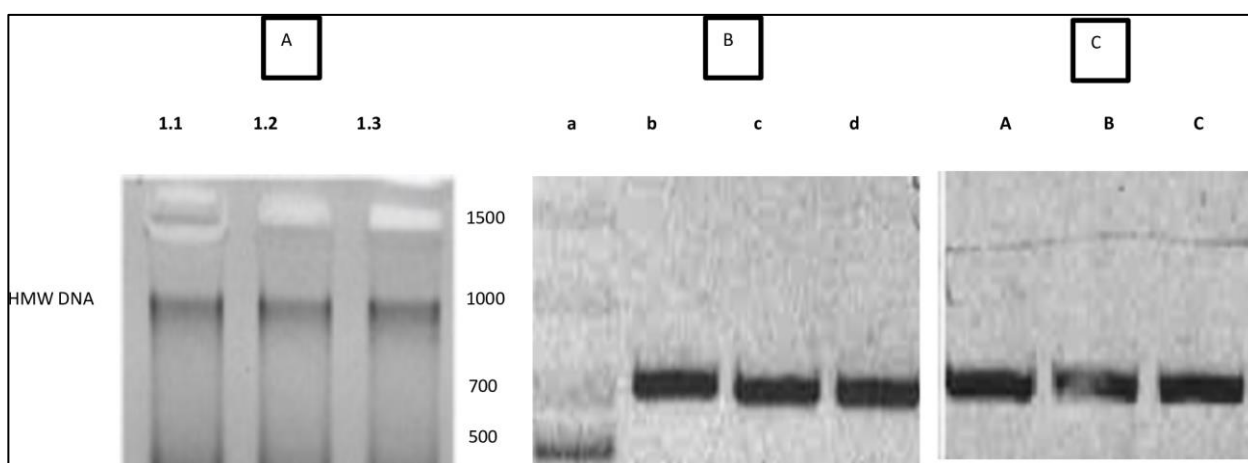


Figure 3.2: Agarose gels (1%) of mealybug (CMB) high molecular weight DNA (A) and PCR amplicons (B and C) from KwaZulu Natal and Mpumalanga provinces. Picture 3.2 A, Lanes 1.1 and 1.2: mealybug DNA samples from MP; and lane 1.3: sample from KZN. Picture 3.2 B, lane a - 1kb molecular ladder; lane b and c (MP samples), d (KZN sample) are PCR amplicons using the 18S primers. Picture 3.2 C, lane A and B (MP samples), C (KZN sample) are PCR amplicons using the 28S primers.

3.2.2.2 DNA characterization of *P. manihoti* species

Nine 18S and nine 28S raw Sanger sequences were obtained from Inqaba Biotec in this study (Appendix, Figure A1, A2, A5 and A6). Sequences received were reverse complimented and subjected to BLASTn searches on NCBI and EMBL. CMB 18S and 28S were aligned separately with the *P. manihoti* voucher sequences to estimate the similarities between the samples (Appendix, Figure A3 and A7), and further aligned with the GenBank datasets using CLC Bio Genomics workbench v9.0 (Appendix, Figure A4 and A8). All nine species were genetically identical, showing a similar population of *P. manihoti* without differences for 18S (Appendix: Figure A3) and 28S, except for the CMB 28S 1.2 and 1.3 (Appendix: Figure A7), which had a few different bases and gaps compared to the other sequences. However, the phylogenetic analysis identified them as *P. manihoti*, and not any other related species.

3.2.2.3 Phylogenetic analysis

The alignments of the nine species from KZN and MP were utilized to construct phylogenetic trees using maximum likelihood method. Phylogeny was limiting to the family Pseudococcidae, with 97% similarity. Phylogenetic analysis shows the family relationships between the species identified. Phylogenetic analysis confirmed the morphological identity of the sampled mealybug species as *P. manihoti*. Both the 18S and 28S rRNA generated sequences of the nine individual mealybug samples clustered with known *Phenacoccus manihoti* sequences (JQ651008.1, JQ651255.1, JQ651256.1, JQ651257.1 and JQ651258.1 vouchers; Sethusa *et al.*, 2014) that had been retrieved from the Genbank with 99% (Figure 3.3 and 3.4) similarity respectively. *Phenacoccus madeirensis* (Madeira mealybug) was shown to be the closest relative to *P. manihoti* in both 18S and 28S rRNA analysis (Figures 3.3 and 3.4). Consensus sequences were submitted to the NCBI GenBank accession numbers KY271370-KY271372 and MG906880- MG906885 (18S sequences); and KY271373-KY271375 and MG910449- MG910454 (28S sequences).

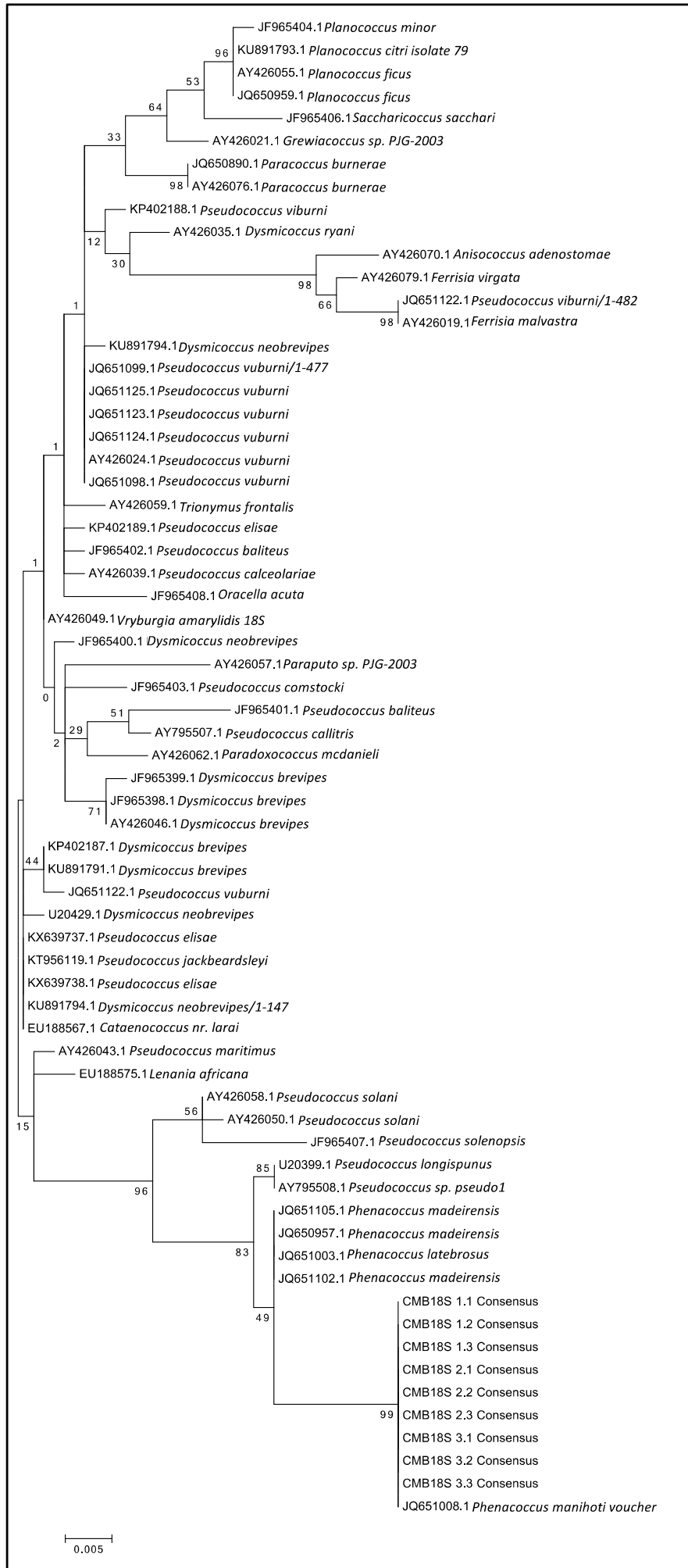


Figure 3.3: Classification of the South African mealybug species from cassava using 18S rRNA.

Phylogenetic tree of *Phenacoccus* species derived from the maximum likelihood analysis of 18S rRNA sequences. Mealybug samples collected in the current study clustered with *Phenacoccus manihoti* 18S RNA sequence (Genbank accession number JQ651008.1 voucher) confirming their identity. Bootstrap values are calculated using non-parametric maximum likelihood (ML) Tamura-Nei bootstrapping model from 10 000 replicates.

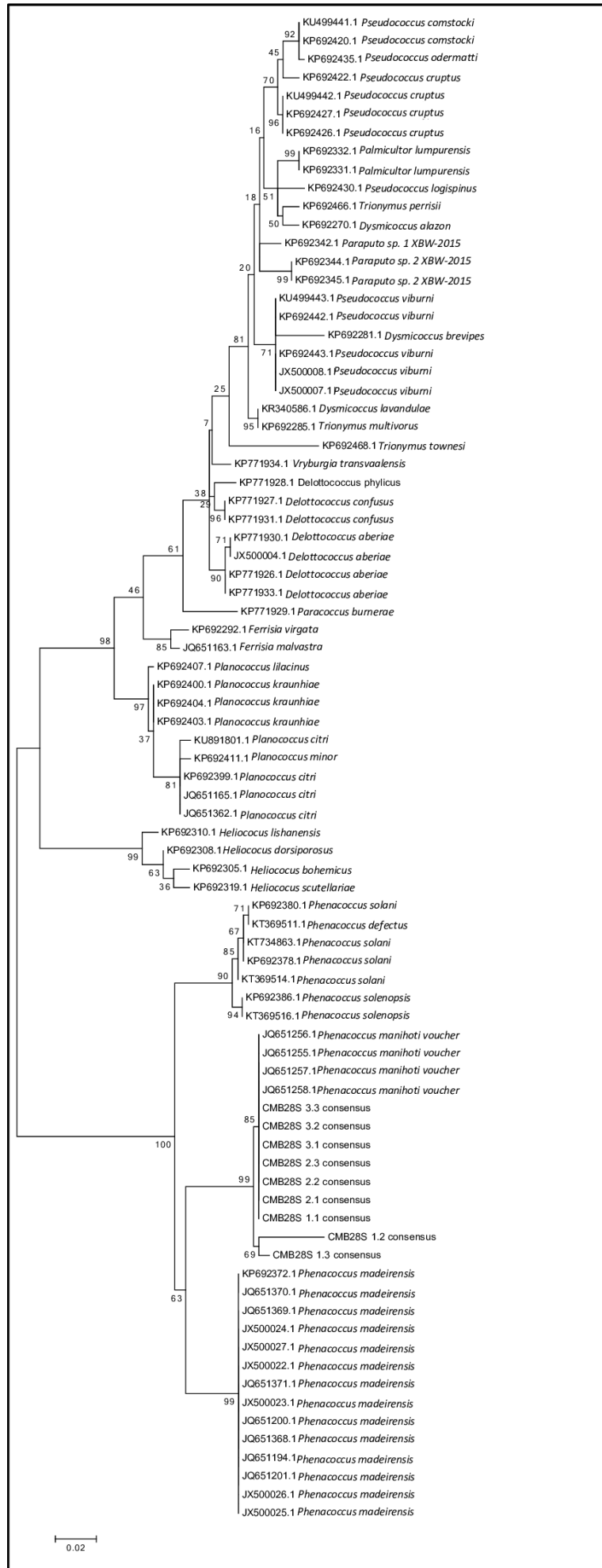


Figure 3.4: Classification of the South African mealybug species from cassava using 28S rRNA.

Phylogenetic tree of *Phenacoccus* species derived from the maximum likelihood analysis of 28S rRNA sequences. Mealybug species here also clustered with *Phenacoccus manihoti* when 28S RNA sequences (Genbank accession number JQ651255.1, JQ651256.1, JQ651257.1, and JQ651257.8 vouchers) were used to draw the phylogenetic tree. Bootstrap values are calculated using non-parametric maximum likelihood (ML) Tamura-Nei bootstrapping model from 10 000 replicates.

3.3 Discussion

The purpose of this chapter was to characterize and confirm the identity of mealybugs infesting cassava plants in KZN and MP of South Africa using morphological and molecular methods. We present here a study in which the morphological and molecular identification of cassava mealybug species in these locations was confirmed for the first time in South Africa (SA). Previous cassava mealybug studies in South Africa have relied on morphological identification methods, which are often laborious and dependent on the female adult (Millar, 2002). However, a combination of two methods for identification of species complements each other. In this study, the taxonomic identifications of the morphological characterization were consistent with the DNA analyses, allowing the identification of *P. manihoti* species from KZN and MP. Important to note is that these two regions are the main cassava producing areas in this country. It should therefore be emphasized that, having only one mealybug species makes each region feasible to deploy control methods. Therefore, the current study serves as a baseline to conduct a more thorough mealybug sampling strategy in KZN and MP, as well as in all cassava growing regions in South Africa and the rest of the African continent.

Overall, this chapter shows the genetic relationship between the species from KZN and MP with the *P. manihoti* voucher species. Furthermore, the phylogenetic analysis also confirms that the species is really *P. manihoti*, and not any other related characterized mealybug species. Both morphological and molecular methods were reliable in confirming the mealybug species collected from KZN and MP as *P. manihoti*. This study provides validation to the existence of the cassava mealybug species identified in the databases, and also confirms that this is not the other cassava mealybug species, *P. herreni*.

CHAPTER 4

Transcriptomics of cassava (*Manihot esculenta* Crantz) following mealybug (*Phenacoccus manihoti*) infestation

AIMS – To phenotypically screen diverse cassava genotypes for their response to cassava mealybug infestation, identify candidate genes and determine gene expression pathways in response to mealybug infestation in resistant and susceptible cassava genotypes.

4.1 Introduction

Cassava as a food security and industrial crop in Africa and globally is currently at risk due to abiotic and biotic stresses. Cassava mealybug (*P. manihoti*) (Millar, 2002; Sethusa *et al.*, 2014), one of the pests of cassava, have been reported to occur in South Africa, however, no information has been formally reported on the damage caused by this pest. Regardless of the success in biological control of cassava mealybug, growing resistant cultivars remains the most sustainable option for managing the pests. More recently, the introduction and availability of genomics and transcriptomics approaches have led to better understanding of pest resistance mechanisms in plants in general (Artico *et al.*, 2014; Broekgaarden *et al.*, 2008; De Vos *et al.*, 2005; Timm & Reineke, 2014).

New developments in sequencing technologies (NGS) have allowed the production of less expensive sequence data to analyse genomics and transcriptomics questions. These NGS technologies are faster and better than Sanger sequencing (Davey *et al.*, 2011; Mardis, 2008a, 2008b; Metzker, 2010; Morozova *et al.*, 2009; Morozova & Marra, 2008). Candidate genes involved in herbivore feeding and wounding can be identified by studying the transcriptional profiles of non-model plants in response to herbivore attack. Although no transcriptomics studies have been reported for response to cassava mealybug using NGS technologies, a recent study in cassava successfully utilized the ABI SOLiD NGS platform to analyse the transcriptome of cassava landraces infected with South African cassava mosaic virus (Allie *et al.*, 2014). In addition, a transcriptome profile of low temperature-treated cassava apical shoots demonstrated dynamic responses to the abiotic cold stress treatment (An *et al.*, 2012) resulting in the expression of MYB and WRKY transcription factors; CYP450 family proteins and MAP kinases, which are known to play a defence role in response to both abiotic and biotic stress. The expression of the aforementioned genes has also been reported during pathogen and insect attack (Artico *et al.*, 2014; De Vos *et al.*, 2005; Ehling *et al.*, 2008; Timm & Reineke, 2014) suggesting their involvement in plant defence.

This chapter aimed at evaluating cassava genotypes for the phenotypic impact of mealybug infestation at 35 days interval, and employing NGS technologies in order to identify transcripts and gene pathways that

play a role in early response (24-72 hours) to mealybug (*P. manihoti*) infestation. With the availability of the cassava genome (Bredeson *et al.*, 2016; Prochnik *et al.*, 2012), novel genes for cassava mealybug resistance manipulation were identified. Comparisons of differentially expressed genes between the different time-points (non-infested mock and infested at 24 and 72 hpi) per genotype in response to mealybug infestation/feeding identified both up-regulated and down-regulated transcripts. These comparisons allowed for identification of putative candidate defence-related genes involved in *P. manihoti* infesting cassava in South Africa. These findings will be useful in understanding cassava-plant-herbivore interactions and in the development of new resistant/tolerant cassava cultivars.

4.2 Results

4.2.1 Cassava phenotypic responses to mealybug infestation

From the phenotypic evaluation, AR23.1 was regarded as partially resistant, P40/1 as susceptible and 98/0581 as tolerant (Table 4.1). Leaves of infested AR23.1 genotype were asymptomatic up to 14 days post-infestation (dpi) (Figures 4.1a and 4.1b) but subsequently showed mild morphological changes around 21 (Figure 4.1c) and 28 dpi (Figure 4.1d). In contrast, P40/1 showed distinct symptoms with extensive wilting and chlorosis by 21 dpi (Figure 4.1g). While the AR23.1 plants continued to regenerate new leaves despite active colonization by the insects, P40/1 plants did not regenerate any new leaves and subsequently wilted and died (Figure 4.1h). Genotype 98/0581, on the other hand, showed tolerance to insect infestation with signs of folding and a slight change in leaf colour (Table 4.1). However, the plants continued to develop new leaves even with high populations of mealybugs.

Table 4.1: Scoring of AR23.1, P40/1 and 98/0581 cassava genotypes for phenotypic responses to cassava mealybug (CMB) infestation at 7, 14, 21, 28 and 35 dpi.

Days post-infestation (dpi)	AR23.1		P40/1		98/0581	
	CMB counts	Symptoms	CMB counts	Symptoms	CMB counts	Symptoms
7 dpi	< 50 CMB	No visible symptoms	> 50 CMB	Slight insect nodules on leaves	> 50 CMB	Slight insect nodules on leaves
14 dpi	< 50 CMB	Plant green with slight insect nodules	± 100 CMB	Moderate insect nodules	± 100 CMB	Slight insect nodules
21 dpi	± 50 CMB	Moderate insect nodules and slight change of leaf colour	> 150 CMB	Change in leaf colour	> 100 CMB	Moderate insect nodules on leaves
28 dpi	≥ 50 CMB	Dying of old damaged leaves	> 200 CMB	Curled leaves, severe reduction of leaf area, Plant dying	± 150 CMB	Curling of leaves, slight reduction of leaf tissue
35 dpi	< 50 CMB	Slight loss of leaf tissue	> 200 CMB	Death of a plant	± 200 CMB	Moderate reduction of leaf tissue, and developing new leaves
Score	Score 1- Partially Resistant		Score 4-Susceptible		Score 2- Tolerant	

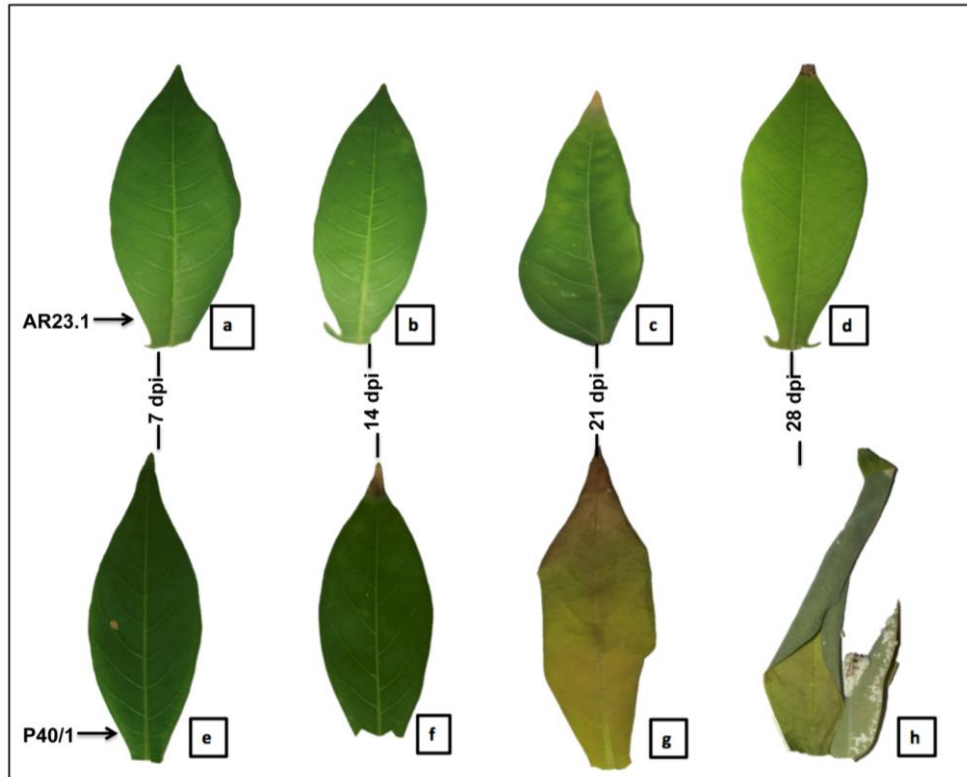


Figure 4.1: Mealybug-induced symptoms on the plants. Cassava genotypes AR23.1 (CIAT) (a-d) and P40/1 (South Africa) (e-h) plants infested with mealybugs (*P. manihoti*) in a greenhouse at ARC-VOPI, South Africa (R753 25° 59" S 28° 35" E). A clear difference on the effect of mealybug across the two genotypes is observed from 21 days post-infestation (dpi). The only observable symptom on AR23.1 leaves at 28 dpi is yellowing on the tips (Figure 4.1d) while P40/1 leaves begin to become soft and show signs of yellowing and curling at the edges at the same time point (Figure 4.1h).

4.2.2 Total RNA integrity of cassava leaf samples

Total RNA was quantified and viewed on an agarose gel electrophoresis (Figure 4.2) as indicated in Chapter 2.4.2. It is evident that the RNA was of a consistent good quality and that the yield was considerably high. RNA concentrations varied between 49.5 ng/μl and 320.2 ng/μl. The RNA integrity was thus classified as excellent and suitable for PCR-based sequencing analyses.

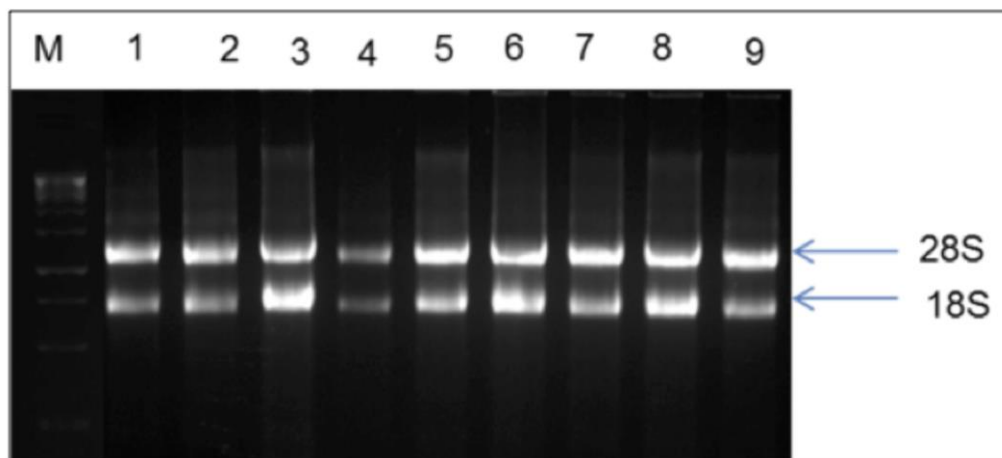


Figure 4.2: Agarose gel (1%) of total RNA isolated from AR23.1, P40/1 and 98/0581 cassava genotypes at 24 and 72 hours of mealybug infestation. Lane M is RiboRuler High Range RNA ladder. Lane 1, 2, 3 represents AR23.1 RNA samples; Lane 4, 5, 6 represents P40/1 RNA samples; while Lane 7, 8, 9 represents 98/0581 RNA samples at 24 and 72 hours of mealybug infestation.

4.2.3 Reads quality check and trimming

Sequenced reads were subjected to FastQC tools to check quality of the reads generated before trimming (Figure 4.3). The plot shows highlights of per base sequence quality of read after sequencing, and trimming with Trimmomatic v0.30 tools. Over 85% of the reads were maintained after trimming, reflecting that the sequences generated were of very high quality for further analysis.

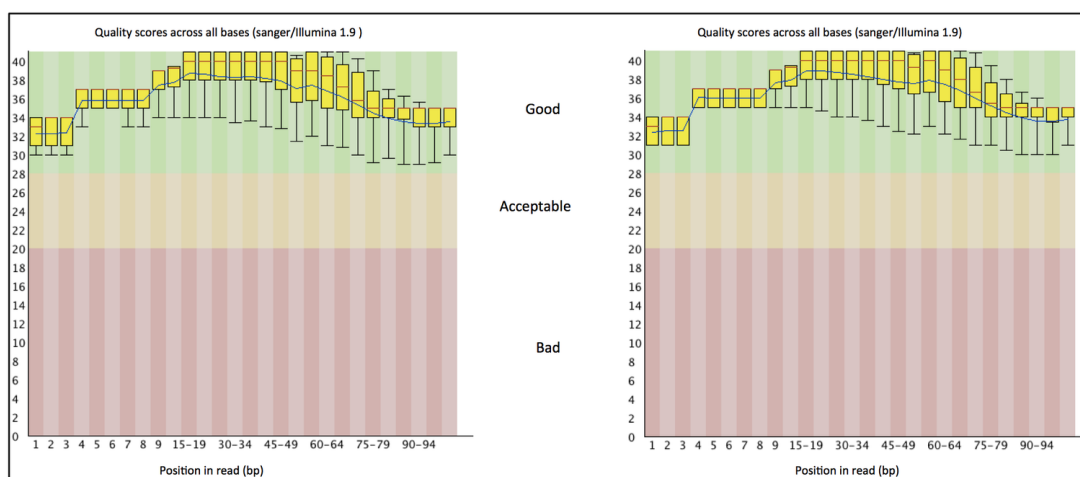


Figure 4.3: Per base sequence quality plot using FastQC of the left and right AR23.1 sequence reads after trimming. The green area indicates the good reads suitable for further analysis and data representation, orange is acceptable reads, and the pink area is the bad unacceptable reads.

4.2.4 Gene expression of cassava genotypes in response to mealybug infestation

A detailed account of all sequencing reads generated per biological replicate is provided in Appendix: Table A1. High quality reads were generated, resulting in the retention of at least 85% of the sequences after trimming (Appendix: Table A1). There was also good correlation between reads from the different biological replicates of each genotype, with correlation values ranging from $R^2 = 0.991 - 0.999$ (Appendix: Figure A9) for AR23.1, $R^2 = 0.998 - 0.999$ (Appendix: Figure A10) for P40/1 to $R^2 = 0.987 - 0.998$ for 98/0581 (Appendix: Figure A11). Given the significant correlation across replicates, the reads were pooled from the three biological replicates at each time point for each sample before mapping reads to the reference genome. Table 4.2 shows a summary of pooled reads per time point for each genotype before and after trimming. On average, 80% of the trimmed reads could be mapped to the cassava reference genome (Mesculenta_305_v6.1) (Table 4.2) using Tophat/Cufflinks (Kim *et al.*, 2013; Trapnell *et al.*, 2012).

Table 4.2: Summary of raw and trimmed reads generated by Illumina HiSeq 2500. Reads are presented per time point after pooling across the three biological replicates of AR23.1 (resistant), P40/1 (susceptible) and 98/0581 (tolerant); and subsequently mapping to the cassava reference genome (Mesculenta_305_v6.1).

RNA-seq dataset	No. of raw reads	Trimmed reads	% Trimmed reads	Mapped reads	Mapped reads/million sequence reads
AR23-24-mock	81,490,264	78,829,123	96.7	70,228,998	890,901
AR23-24-inf	85,904,324	82,499,207	96.0	68,587,064	831,366
AR23-72-mock	120,666,228	112,979,340	93.6	85,831,373	759,708
AR23-72-inf	116,142,504	104,578,982	90.0	90,357,903	864,015
P40-24-mock	88,204,836	85,273,331	96.6	76,921,462	902,057
P40-24-inf	84,905,130	79,802,477	94.0	66,220,010	829,798
P40-72-mock	98,680,132	91,365,796	93.3	74,305,354	813,273
P40-27-inf	95,190,544	90,807,902	95.4	77,618,821	854,758
0581-24-mock	53,359,502	40,762,232	94.0	34,202,461	839,072
0581-24-inf	47,951,444	43,628,412	91.0	36,126,443	828,048
0581-72-mock	79,116,061	63,315,061	91.6	55,392,669	874,873
0581-27-inf	78,674,145	59,668,658	90.8	54,547,181	914,168

A total of 301, 206 and 269 transcripts from AR23.1, P40/1 and 98/0581 were differentially expressed ($\log_2$ fold change and $P \leq 0.05$) at various time points following mealybug infestation (Table 4.3). In the AR23.1, P40/1 and 98/0581 genotypes, mealybug infestation resulted in a higher number of down-regulated than up-regulated genes (Table 4.3), except in P40/1 between infested leaves at 24 hpi and 72 hpi, where up-regulated genes were more frequent than the down-regulated genes (Table 4.3). Comparative analysis revealed that most of the DEGs observed were uniquely expressed at specific time points and that there were less common DEGs across different time points in all the three genotypes (Figure 4.4). Of all the

DEGs identified, only 4, 1 and 1 were common across all time points for genotype AR23.1, P40/1 and 98/051 respectively (Figure 4.4). There was a steady increase in the number of DEGs with increase in time post-infestation (Table 4.3) suggesting an increase in mealybug defence response with progress in infestation.

Table 4.3: An overview of the number of up-regulated and down-regulated genes (DEGs) identified between non-infested mock and mealybug infested leaves at 24 and 72 hours of mealybug infestation. In P40/1 the number of down-regulated DEGs between the infested tissues at 24 hpi versus 72 hpi was lower compared to up-regulated.

Dataset 1	Dataset 2	Time-points	Up-regulated	Down-regulated	Total
AR23-24-mock	AR23-24-inf	24 h versus 24 hpi	38 (41%)	54 (59%)	92
AR23-72-mock	AR23-72-inf	72 h versus 72 hpi	42 (42%)	57 (58%)	99
AR23-24-inf	AR23-72-inf	24 hpi versus 72 hpi	33 (30%)	77 (70%)	110
		Total	113 (38%)	188 (62%)	301
P40-24-mock	P40-24-inf	24 h versus 24 hpi	31 (46%)	36 (54%)	67
P40-72-mock	P40-72-inf	72 h versus 72 hpi	27 (39%)	43 (61%)	70
P40-24-inf	P40-72-inf	24 hpi versus 72 hpi	38 (55%)	31 (45%)	69
		Total	96 (47%)	110 (53%)	206
0581-24-mock	0581-24-inf	24 h versus 24 hpi	34 (42%)	47 (58%)	67
0581-72-mock	0581-72-inf	72 h versus 72 hpi	38 (42.2%)	52 (57.8%)	70
0581-24-inf	0581-72-inf	24 hpi versus 72 hpi	33 (34%)	65 (66%)	69
		Total	105 (39%)	164 (61%)	269

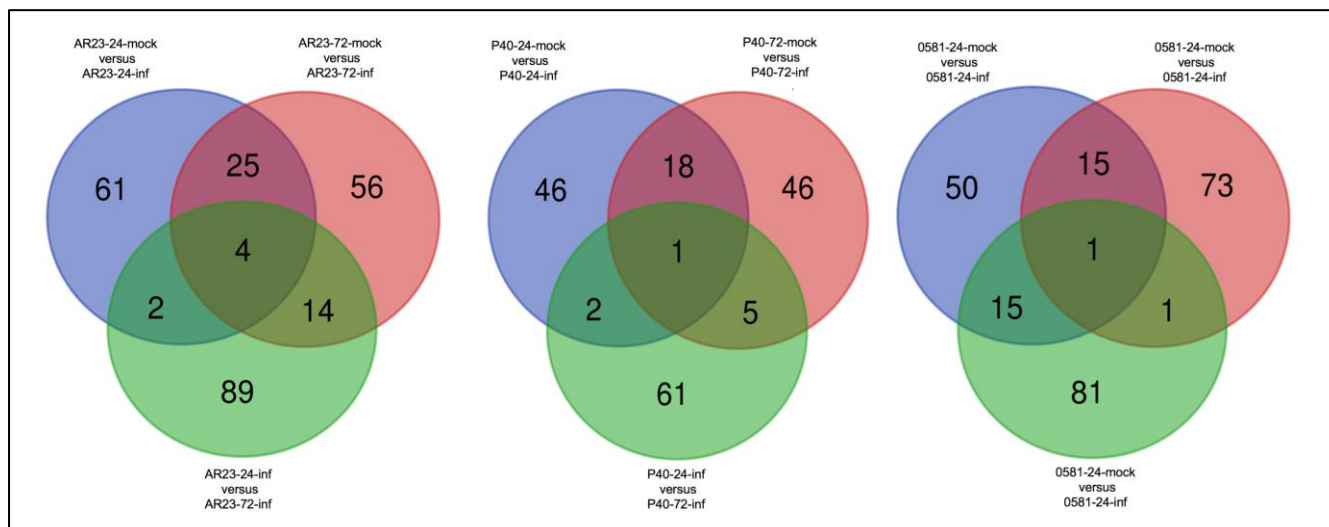


Figure 4.4: Venn diagrams of differentially expressed genes (DEGs). A comparison of DEGs that were common across different time points (24 and 72 hours compared to non-infested mock) in AR23.1 and P40/1 post-infestation with *P. manihoti*. The Venn diagrams represent an average across all three biological replicates at respective time points. Only transcripts with a minimum of log₂ fold change and $P \leq 0.05$ were included in the data.

4.2.5 Transcriptional changes in AR23.1 in response to mealybug infestation

Out of the 301 differentially expressed genes (DEGs) identified (Table 4.2) in AR23.1, 242 were unique (Appendix: Table A2). GO terms were assigned for 132 out of the 242 unique DEGs (55%) for describing biological processes, cellular components and molecular functions (Figure 4.5). Under the GO functional categorization, DE genes were highly represented in the cellular process, metabolic process, response to stimulus and response to stress in the biological process category. Binding and catalytic functions were over-represented within the molecular function category while the most abundant DE genes under cellular component category were the cell and cell part (Figure 4.5).

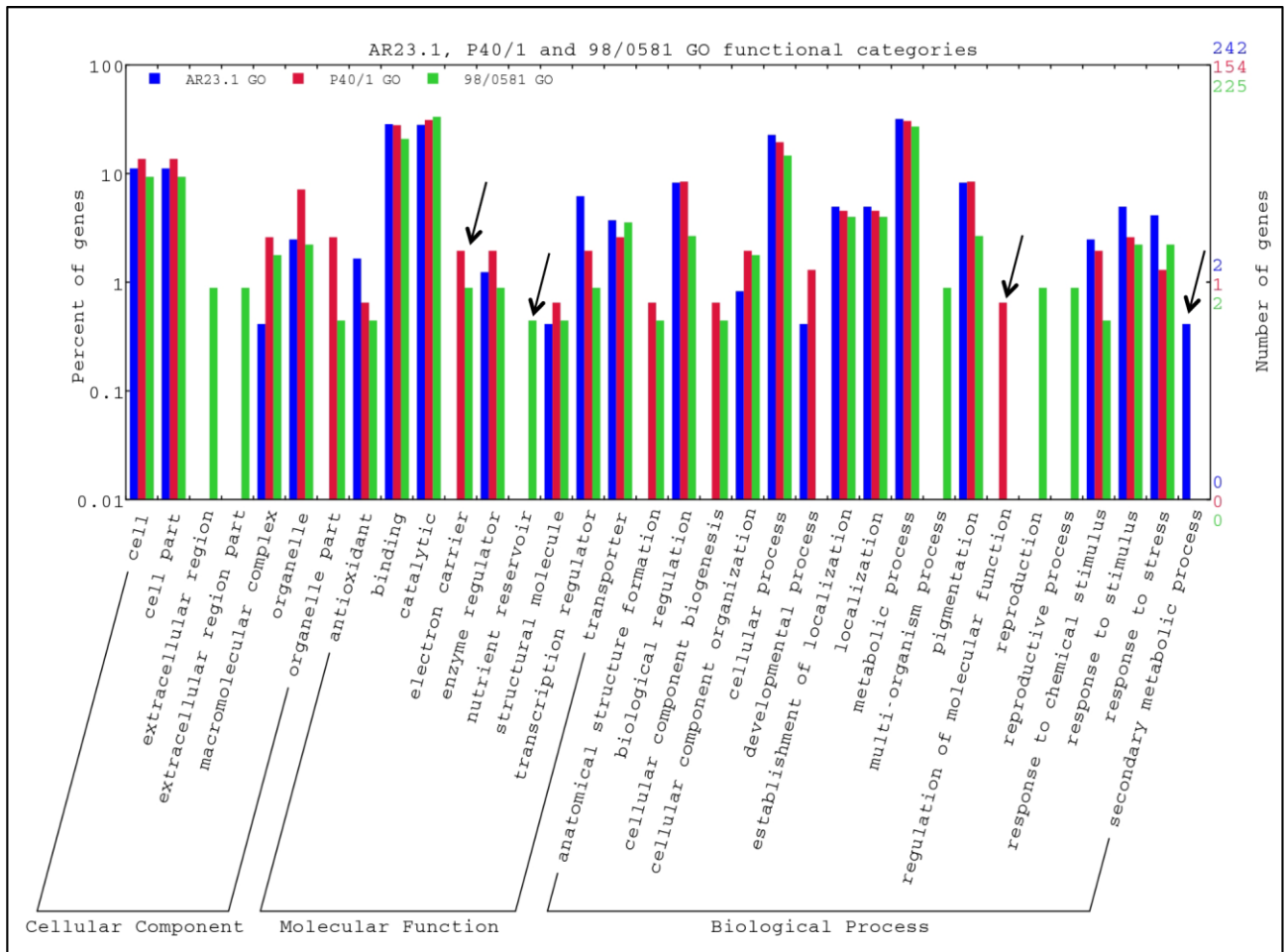


Figure 4.5: Analysis of enriched GO terms. Functional categorization of all unique differentially expressed genes based on biological processes, cellular components and molecular functions for AR23.1, P40/1 and 98/0581 genotypes. Significant differences observed between the DEGs of the two genotypes are highlighted using the pointing arrows.

An abundance of genes encoding pathogenesis-related (PR) proteins was observed among the differentially expressed genes in AR23.1, most of which were classified as resistance (R) receptor like detection and signalling proteins (Appendix: Table A2). Two Leucine-rich repeat receptor-like protein kinase (LRR-PK) family proteins (Manes.03G048300, Manes.01G142200) that were up-regulated within 24 hours of mealybug infestation, and a third one (Manes.01G129500) that was up-regulated at 72 hours of mealybug infestation (Appendix: Table A2) were also detected. Most of the PR proteins appeared to be induced within 24 hours of mealybug infestation including an HSP20-like chaperone superfamily protein (Manes.02G124600), and MYB domain proteins 4 (Manes.01G147500) and 55 (Manes.01G235800). There was evidence of reduced expression by 72 hours of mealybug infestation demonstrating a decreased expression of these proteins after the first 24 hours with mealybugs.

However, both up- and down-regulation of some of the pathogenesis-related proteins were observed at 24 hours of mealybug infestation including WRKY DNA-binding protein 56 (Manes.01G252200) and WRKY

family transcription factor (Manes.01G273700). Other known insect-defence related proteins, such as Cytochrome P450 family 339 genes (Manes.01G200000, Manes.15G189500, Manes.15G154400, Manes.01G178000, Manes.01G264600, Manes.10G146700, Manes.03G138800, Manes.08G047000, Manes.02G071200) as well as Kunitz trypsin inhibitors (Manes.01G126300), were either up- or down-regulated at both 24 and 72 hours of mealybug infestation (Appendix: Table A2).

Metabolic pathway analysis, using the KEGG database, revealed that 73 out of 244 unique DEGs (30%) identified from AR23.1 mapped to 14 KEGG pathways (Figure 4.6A). There were more down-regulated than up-regulated genes in almost all of the KEGG pathways mapped (Figure 4.6A), although a high percentage of genes could not be mapped due to a lack of KO terms. There were both up- and down-regulated genes mapped to the plant-pathogen interactions (Figure 4.7 and 4.8) and plant hormone signal transduction pathway (Appendix: Figure A12 and A13), phenylpropanoid biosynthesis (Appendix: Figure A14 and A15) and glutathione metabolism (Appendix: Figure A16 and A17).

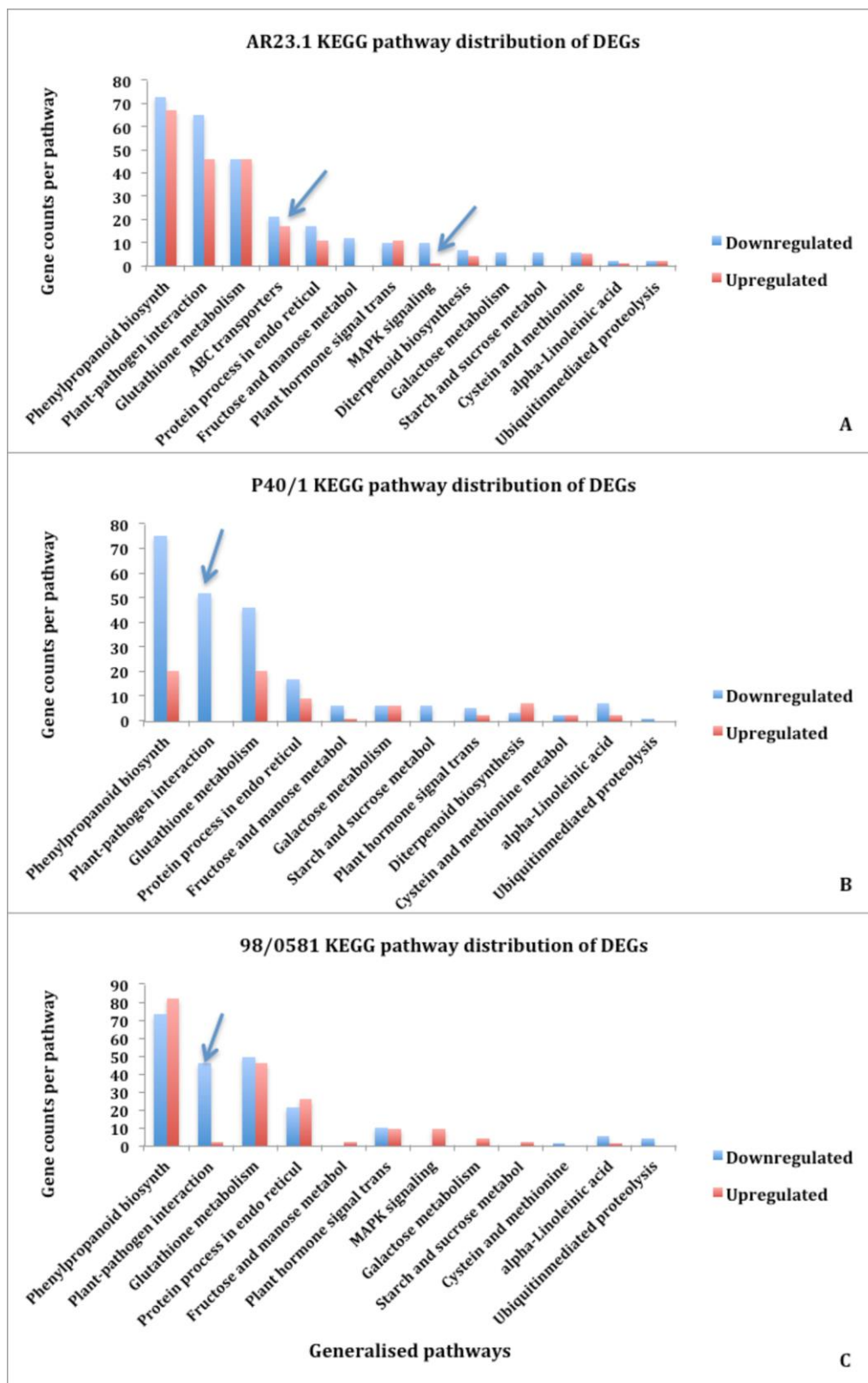


Figure 4.6: KEGG classification of differentially expressed genes. The KEGG pathway distribution of DEGs in AR23.1 (A), P40/1 (B) and 98/0581 (C) in response to cassava mealybug infection using castor bean (*Ricinus communis*) as the reference. The graph represents the 14 most enriched pathways, and the number of genes per pathway. Significant differences observed between the three genotypes are pointed with arrows.

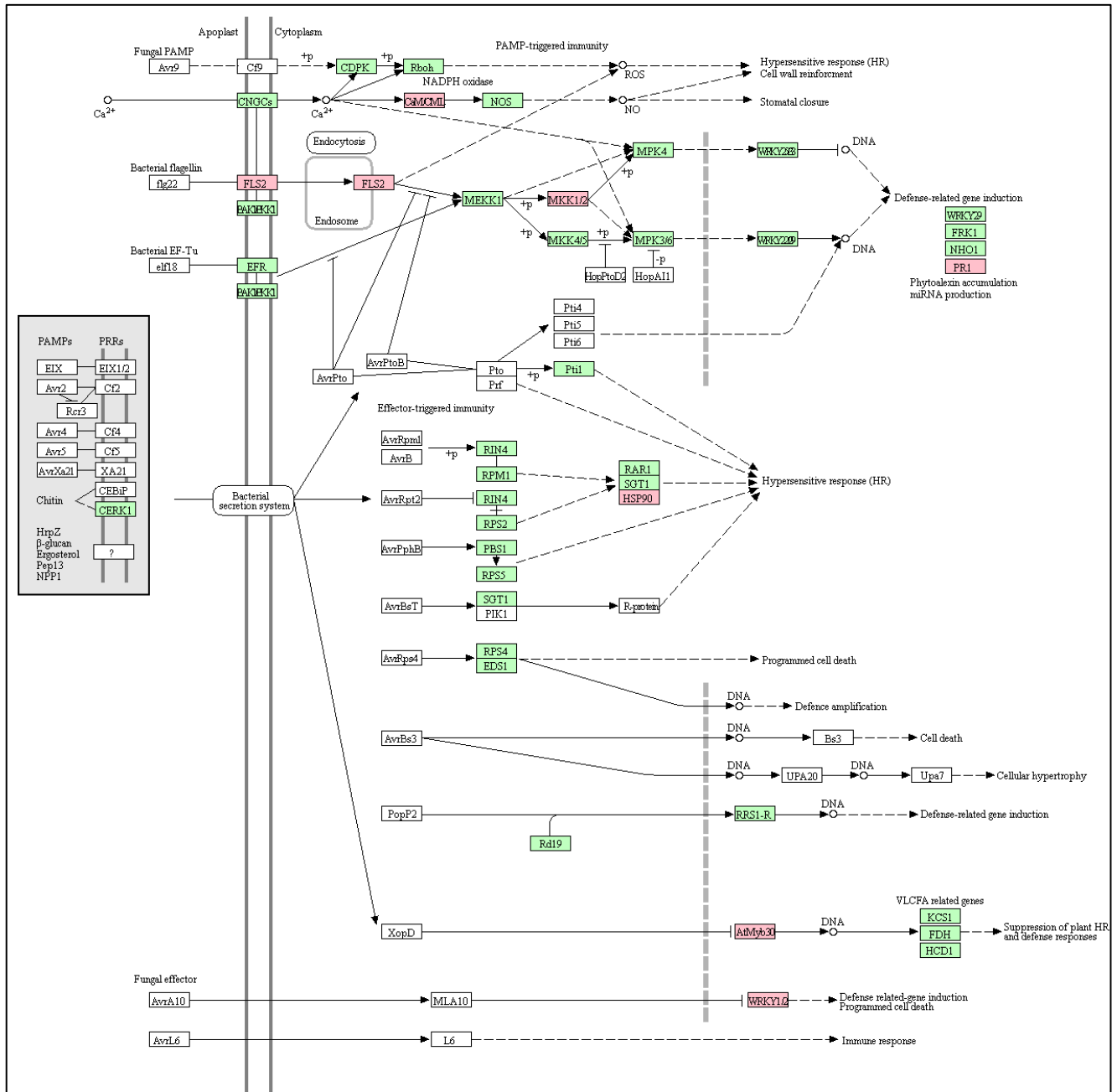


Figure 4.8: Localization of the down-regulated plant-pathogen interaction pathway enzymes on the leaf tissues of AR23.1 post-infestation with mealybugs. Boxes highlighted in green represent genes that are present in the castor bean genome (*Ricinus communis*; reference genome used for KEGG maps), while boxes highlighted in pink are the down-regulated DEGs in the leaf tissues post mealybug infestation. Boxes that are not highlighted are the genes that are not present in the castor bean genome.

4.2.6 Transcriptional changes in P40/1 in response to mealybug infestation

Out of the 206 DEGs identified (Table 4.2) in P40/1, 154 were unique (Appendix: Table A3). GO terms for 86 out of the 154 unique DEGs (56%) were assigned to biological processes, cellular components and molecular functions (Figure 4.5). GO functional categorization was over-represented in the cellular and metabolic process, and response to stress in the biological process category, binding and catalytic in the

molecular function category, and cell and cell part in the cellular component category (Figure 4.5). However, genes under regulation of molecular function, electron carrier and organelle part categories were specifically enriched in the P40/1 genotype in comparison with AR23.1.

Similar to the genotype AR23.1, we observed the up-regulation of the Leucine-rich repeat receptor-like protein kinase family protein (Manes.02G150200) and leucine-rich repeat transmembrane protein kinase family protein (Manes.03G048300) within 24 hours of mealybug infestation. Other pathogenesis related (PR) proteins such as IAA-leucine resistant (ILR)-like 2 (Manes.01G196200) were only up-regulated at 72 hours of mealybug infestation. The same disease resistance gene (Manes.02G205700) belonging to the TIR-NBS-LRR class of *R* genes, that was found to be down-regulated at 72 hours of mealybug infestation in a resistant genotype AR23.1 (Appendix: Table A2), was also down-regulated in susceptible genotype P40/1 (Appendix: Table A3) at the same time point. Moreover, two other disease resistance family proteins (TIR NBS LRR; Manes.11G156500, Manes.14G165100) were also down-regulated in mealybug-infested tissues harvested at 24 hours compared with mealybug infested tissues harvested at 72 hours.

Most of the pathogenesis-related proteins in P40/1 were down-regulated, except for two HSP20-like chaperone-superfamily proteins (Manes.02G124800, Manes.02G124700), which were highly up-regulated in mealybug infested tissues harvested at 24 hours compared with mealybug infested tissues at 72 hours (Appendix: Table A3). These two HSP20-like chaperone-superfamily proteins were different from those that were up-regulated in AR23.1 above. Up-regulation of a WRKY DNA-binding protein 27 (Manes.02G017100) was also observed when mealybug infested tissues harvested at 24 hours were compared with infested tissues at 72 hours. This observation suggested a later (after 24 hours) induction of the two HSP20-like chaperones superfamily proteins (Manes.02G124800, Manes.02G124700) and WRKY DNA-binding protein 27 (Manes.02G017100).

The up-regulation of UDP-glucosyl transferase 74B1 (Manes.01G196500) was observed when tissues harvested at 24 hours were compared with those harvested at 72 hours of mealybug infestation, suggesting a later induction of this gene during mealybug feeding. The same gene was found to be up-regulated when mealybug infested tissues harvested at 24 hours were compared with non-infested mock tissues at the same time further confirming the late induction of this gene. UDP-glucosyl transferases have been reported to be responsible for the last step in the biosynthesis of linamarin and lotaustralin, which are the major cyanogenic glucosides in cassava (Jørgensen *et al.*, 2011).

KO terms were assigned to 52 out of the 206 (25.2%) DEGs identified, and were subsequently mapped to 12 KEGG pathways (Figure 4.6B). DE genes that were mapped to the plant-pathogen interactions pathway were down-regulated (Figure 4.9) while those that were mapped to the plant hormone signal transduction

pathways were both up- (MAP kinase substrate 1, Auxin efflux carrier 17 family protein, *Jamonate-zim-domain* protein 8, Highly ABA-393 induced *PP2C* gene 3) and down-regulated [(pathogenesis-related thaumatin superfamily protein, *Pathogenesis-related* 4) (Appendix: Figure A18 and A19)]. It is important to note that due to lack of representative KO terms, not all genes were mapped and therefore the output from KEGG analysis would not be representative of all the DEGs.

process, nutrient reservoir, and multi-organism process were only enriched in the 98/0581 genotype (Figure 4.5).

A disease resistance family protein/LRR family protein (Manes.10G129500) was up-regulated between infested tissues at 24 hours of mealybug infestation compared to non-infested mock, while a similar gene [(disease resistance family protein/LRR family protein) (Manes.07G042900)] was up-regulated at 72 hours of mealybug infestation compared to non-infested mock. Two LRR receptor-like protein kinase family proteins were also found to be down-regulated when infested tissues at 24 hours (Manes.06G012500) and 72 hours of mealybug infestation (Manes.11G118000) were compared to non-infested mock tissues. Both the *R* genes that were observed at 24 hours of mealybug infestation were not significant at 72 hours of mealybug infestation, and vice versa.

Most of the pathogenesis-related genes were down-regulated, except for the mitogen activated protein kinases that were up-regulated at 24 hours of mealybug infestation compared to non-infested mock (Manes.05G155300), 72 hours of mealybug infestation compared to non-infested mock (Manes.05G031200), and between infested tissues at 24 and 72 hours of mealybug infestation (Manes.18G020900). Several down-regulated pathogenesis-related genes included WRKY DNA-binding protein 70 (Manes.08G170600) and the myb domain proteins (Manes.04G140800, Manes.05G020800, Manes.16G096700) that were up- or down-regulated between the time-points (Appendix: Table A4). In addition, cytochrome P450 and heat shock proteins were also up- or down-regulated between the time-points (Appendix: Table A4).

KO terms were assigned to 72 out of 225 DEGs identified, which mapped to 12 KEGG pathways (Figure 4.6C). Up- and down-regulated genes significantly mapped to plant-pathogen interactions (Figure 4.10 and 4.11), plant hormone signal transduction (Appendix: Figure A24 and A25), phenylpropanoid biosynthesis (Appendix: Figure A26 and A27) and glutathione metabolism (Appendix: Figure A28 and A29).

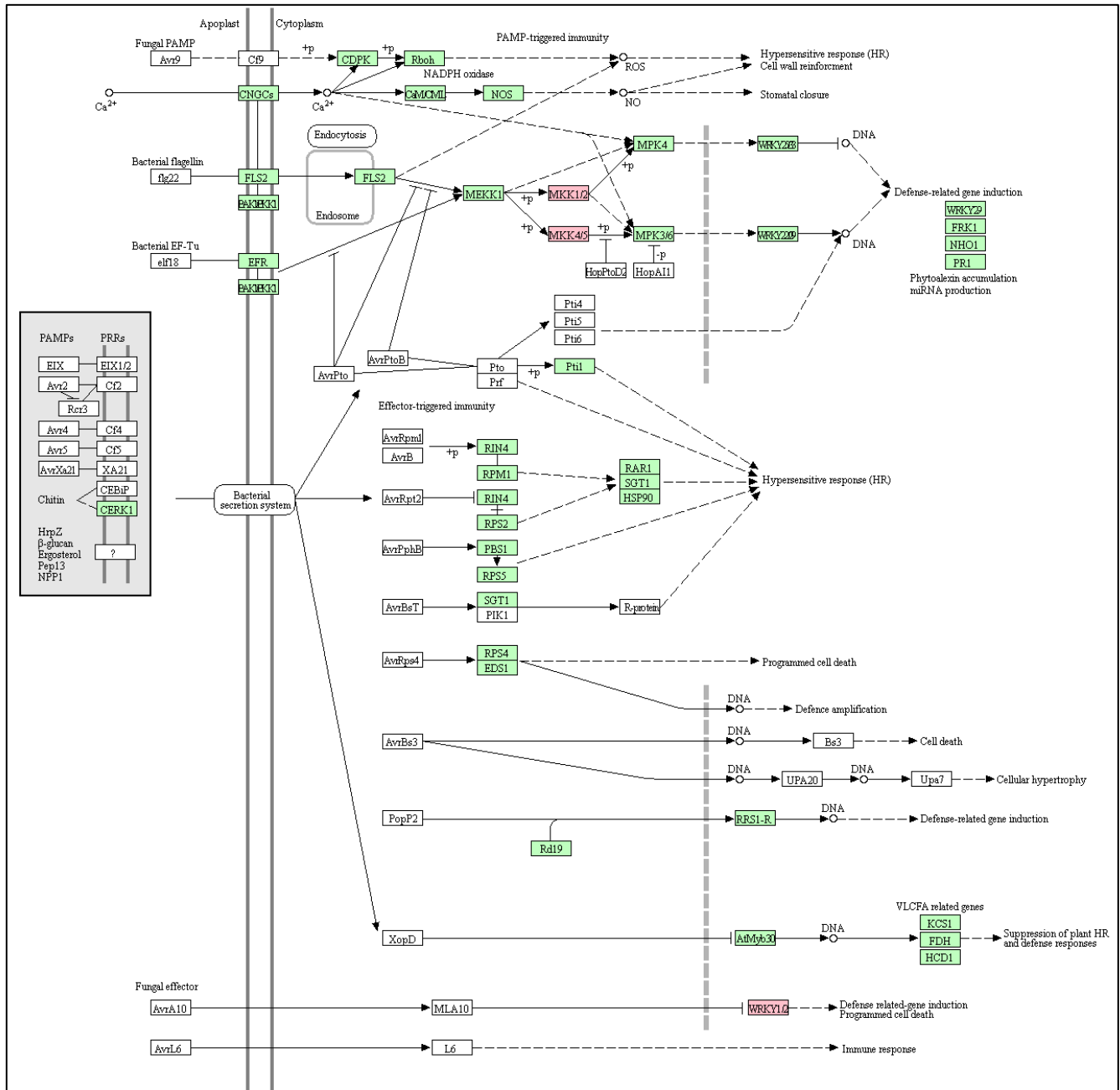


Figure 4.10: Localization of the up-regulated plant-pathogen interaction pathway enzymes on the leaf tissues of 98/0581 post-infestation with mealybugs. Boxes highlighted in green represent genes that are present in the castor bean genome (*Ricinus communis*; reference genome used for KEGG maps), while boxes highlighted in pink are the down-regulated DEGs in the leaf tissues post mealybug infestation. Boxes that are not highlighted are the genes that are not present in the castor bean genome.

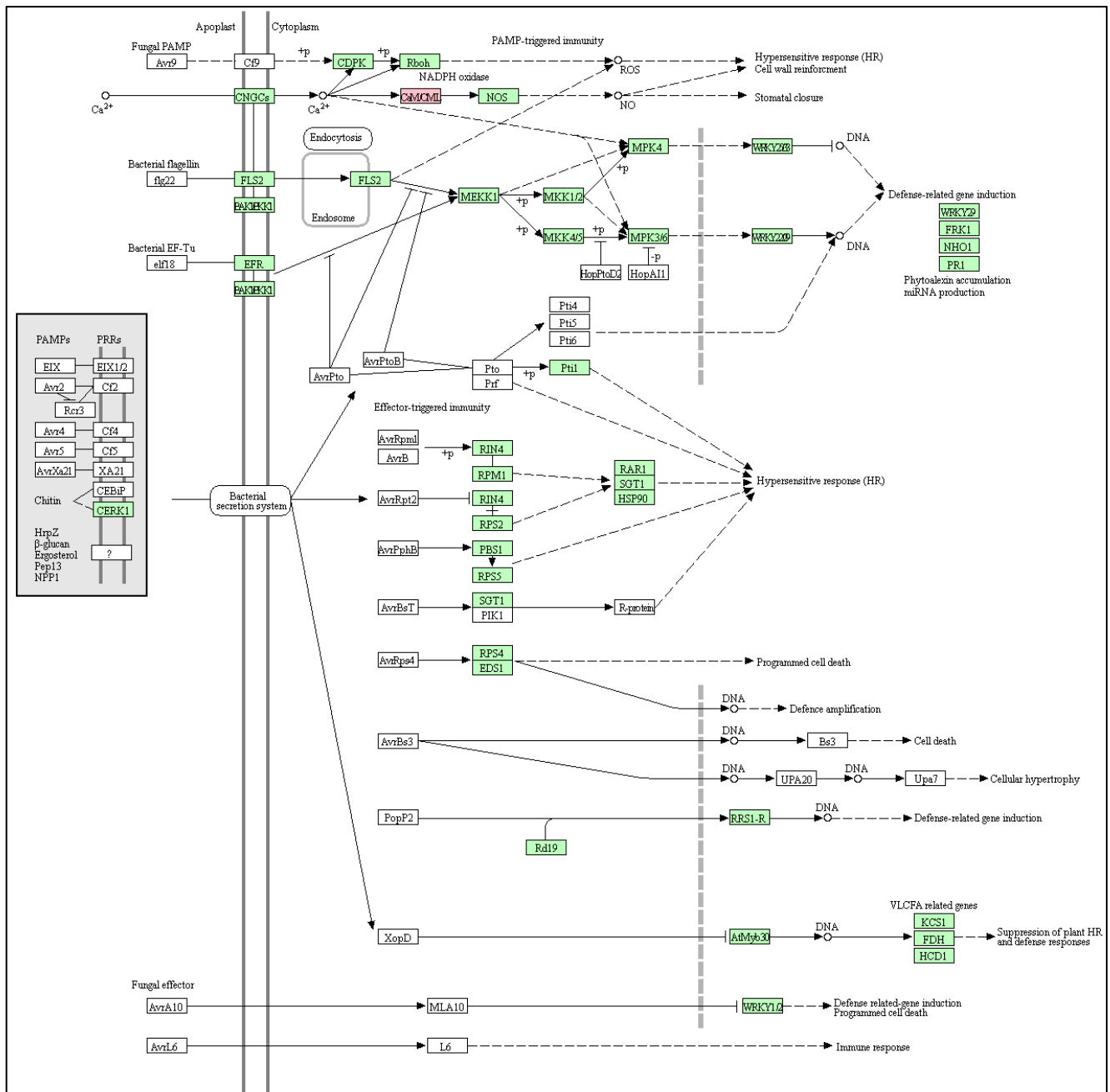


Figure 4.11: Localization of the down-regulated plant-pathogen interaction pathway enzymes on the leaf tissues of 98/0581 post-infestation with mealybugs. Boxes highlighted in green represent genes that are present in the castor bean genome (*Ricinus communis*; reference genome used for KEGG maps), while boxes highlighted in pink are the down-regulated DEGs in the leaf tissues post mealybug infestation. Boxes that are not highlighted are the genes that are not present in the castor bean genome.

4.2.8 Comparison of expression patterns between AR23.1, P40/1 and 98/0581

Gene expression was further compared between AR23.1, P40/1 and 98/0581 to determine whether there were any genotype-specific expression patterns. Within the GO functional classification, genes under secondary metabolic process were significantly enriched in AR23.1 in comparison with P40/1 and 98/0581

(Figure 4.5). On the other hand, genes under regulation of molecular function were significantly enriched in P40/1, while genes under reproduction and reproductive process and nutrient reservoir, were enriched in the 98/0581 genotype compared to P40/1 and AR23.1 (Figure 4.5). Under the metabolic pathway analysis using the KEGG database, genes under ABC transporters and MAPK signalling pathways, both of which have been reported to be involved in plant defence, were present in AR23.1 and 98/0581; and not in P40/1 (Figure 4.6). There were both up- and down-regulated genes that were mapped to the plant-pathogen interactions (Figures 4.7, 4.8, 4.10 and 4.11) and plant hormone signal transduction pathway (Appendix: Figures A12, A13, A24 and A25) in genotype AR23.1 and 98/0581, while in P40/1, there were no up-regulated genes classified under plant-pathogen interactions pathway (Figure 4.9). It must be noted that the genes shown in Figures 4.7 - 4.11; and Appendix: Figures A12-A29 are an under-representation of all plant-pathogen interactions, plant-hormone signal transduction, phenylpropanoid biosynthesis and glutathione metabolism genes that were found to be differentially expressed in AR23.1, P40/1 and 98/0581 due to lack of representative KO terms.

Four genes [Gibberellin 2-oxidase 6 (Manes.01G231100), HSP20-like chaperones superfamily protein (Manes.02G124600), alpha/beta-Hydrolases superfamily protein (Manes.01G274500), and GDSL-like Lipase/Acylhydrolase superfamily protein (Manes.02G189200),] that were up-regulated in AR23.1 were down-regulated in P40/1 at 24 hours of mealybug infestation. Interestingly, a similar HSP20-like chaperones superfamily protein (Manes.02G124600) was also up-regulated (+0.58 log₂ fold) in the 98/0581 genotype at the same time-point, together with the alpha/beta-Hydrolases superfamily protein (Manes.09G040500), and GDSL-like Lipase/Acylhydrolase superfamily protein [(Manes.13G130400) (Appendix: Table A4)]. A “pectin lyase-like superfamily” protein (Manes.02G092000) that was up-regulated (+1.73) in P40/1 at 72 hours of mealybug infestation was down-regulated in AR23.1 (-1.51) although a different “pectin lyase-like superfamily” protein (Manes.01G231200) was up-regulated in AR23.1 at the same time period (Appendix: Tables A2 and A3). The up-regulation of a “Calmodulin like 23” (Manes.02G075400) in AR23.1 and a down-regulation of the same family protein [“Calmodulin like 23” (Manes.01G116700)] was observed in P40/1 (Appendix: Tables A2 and A3). The UDP-glucosyl transferase was also down-regulated in AR23.1 (Manes.02G054300) and 98/0581 (Manes.10G097000, Manes.06G086200, Manes.10G097000), whereas in P40/1, UDP-glucosyl transferase 74B1” (Manes.01G196500) was up-regulated.

4.2.8 Network interaction of AR23.1, P40/1 and 98/0581 proteins

Interacting genes/proteins were further estimated between the up- and down-regulated DEGs in AR23.1 (Figure 4.12 and 4.13), P40/1 (Appendix: Figure A30 and A31) and 98/0581 (Appendix: Figure A32 and A33) genotypes using a Search Tool for Retrieval of Interacting Genes/Proteins (STRING) tools (Jensen *et al.*, 2009; Szklarczyk *et al.*, 2011). The analysis was specifically focused on the defence response genes

(highlighted in red) post mealybug feeding (Figure 4.12 and 4.13). Only a low level of network interaction was observed between the up- and down-regulated DEGs tested in the AR23.1 genotype, with 13 defence response genes being up-regulated and 11 down-regulated. Defence response genes (highlighted in red) such as cytochrome P450, family 94, subfamily B, polypeptide 3 (*CYP94B3*), putative WRKY transcription factor 53 (WRKY53), ethylene-responsive transcription factor 1A (*ERF-1A*), BON association protein 2 (BAP2) and glutathione S-transferase (*GSTF6*) were among the up-regulated, while MLO-like family protein (*MLO8*), FERONIA (*FER*), glutathione S-transferases (*GSTF9* and *GSTF11*), S-adenosyl-L-methionine-dependent methyltransferase-like protein (*BSMT1*) and putative WRKY transcription factor 72 (WRKY72) were down-regulated. Although the numbers of network interactions were low amongst defence-related proteins, there was, for example, co-expression of WRKY53 and ERF-1 transcription factor, *MYC2*, and *CYP94B3* (Figure 4.12). Moreover, there were no interacting proteins observed between the down-regulated defence response genes (highlighted in red), except between *MLO8* and *FER* genes (Figure 4.13). Defence response genes appeared to interact mainly with other classes of genes that are not necessarily of the defence response genes.

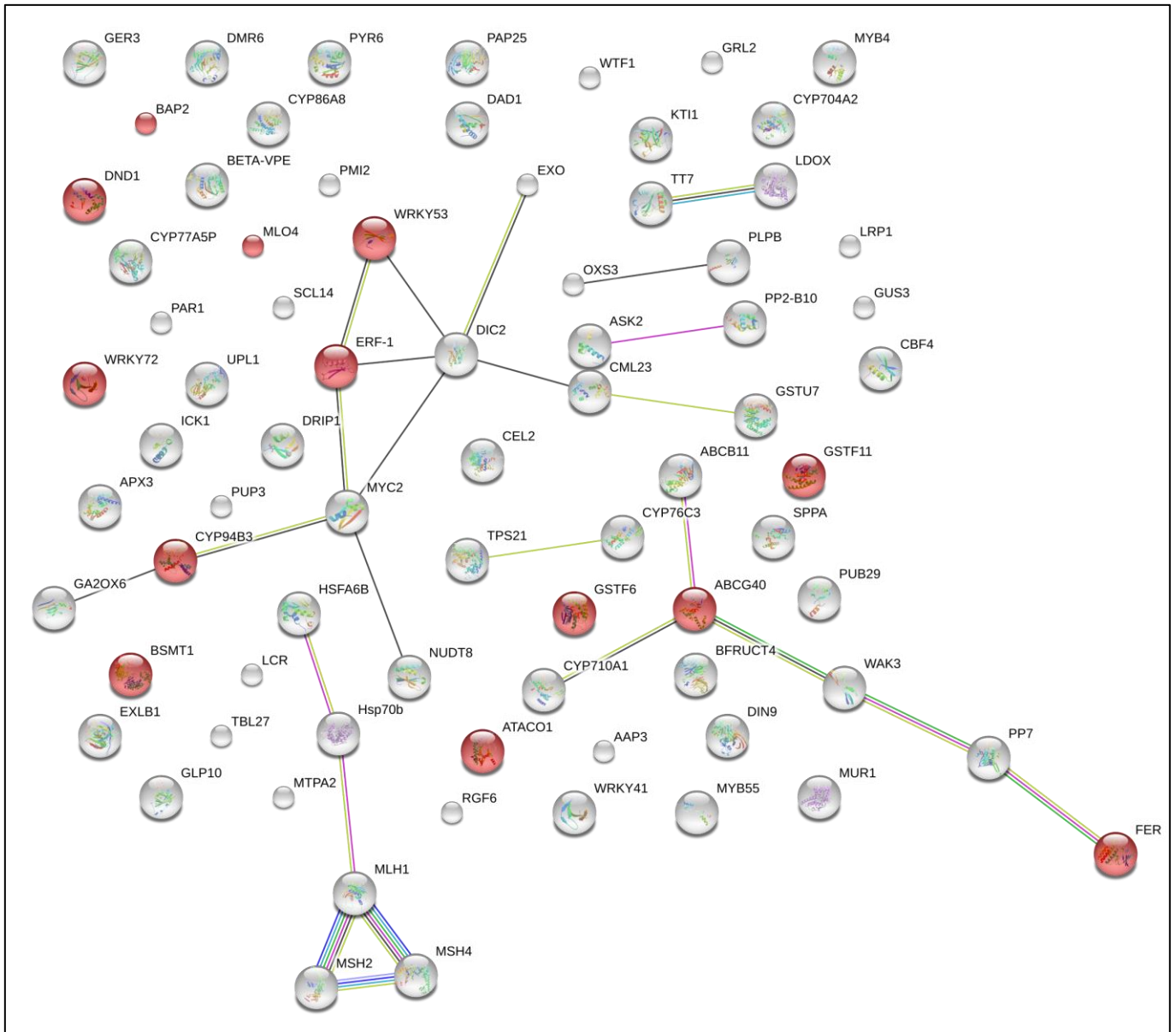


Figure 4.12: Network representation of the interactions of the up-regulated differentially expressed genes (DEGs) in AR23.1 post-infestation with mealybugs. The *Arabidopsis thaliana* genome was used to infer interactions in cassava. Different lines and colours identify types of evidence of interactions. Green line represents gene neighbourhood, red: gene fusion, light blue: protein homology, dark blue: gene occurrence, black: co-expression, and pink is the experimentally determined genes. DEGs, in the leaf tissues post mealybug infestation (highlighted in red) represent defence response genes.

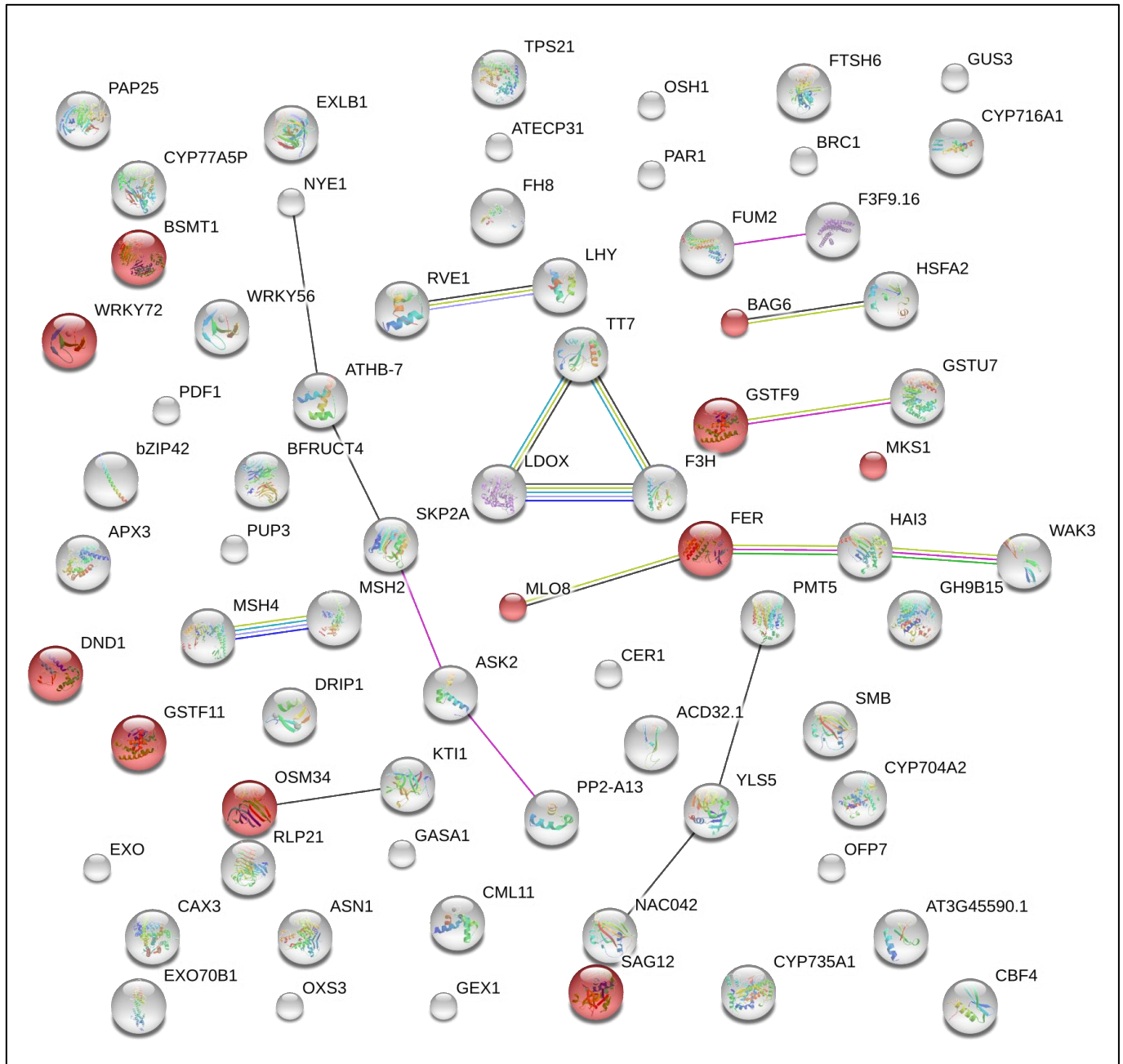


Figure 4.13: Network representation of the interactions of the down-regulated differentially expressed genes (DEGs) in AR23.1 post-infestation with mealybugs. The *Arabidopsis thaliana* genome was used to infer interactions in cassava. Different lines and colours identify types of evidence of interactions. Green line represents gene neighbourhood, red: gene fusion, light blue: protein homology, dark blue: gene occurrence, black: co-expression, and pink is the experimentally determined genes. DEGs, in the leaf tissues post mealybug infestation (highlighted in red) represent defence response genes.

In the P40/1 genotype, eight defence response genes (highlighted in red) were up-regulated (Appendix: Figure A30) and there was no interaction between these genes. Each gene was a standalone or interacting with other classes of genes. On the other hand, only one defence response gene, 3-Ketoacyl-CoA synthase protein (*KCS1*), was down-regulated (Appendix: Figure A31). The 98/0581 genotype showed 17 defence

response genes (highlighted in red) that were up-regulated. These genes included the heat shock 70kDa PROTEIN (HSC70-1), cytochrome P450, family 82, subfamily C, polypeptide 4 (*CYP82C4*), putative transcription factor WRKY40, ethylene-responsive transcription factors (ERF-1 and ERF-6) and Osmotin 34 (*OSM34*) among others (Appendix: Figure A32). A high interaction was observed between the genes such as HSC70-1 and glutathione S-transferase (*GSTT1*), cytochrome P450, family 81, subfamily D, polypeptide 8 (*CYP81D8*), heat shock family proteins (HSP21 and HSP17.61I) and Bcl-2-associated athanogene 6 (*BAG6*). For example, *HSC70-1* and *GSTT1* were interacting in terms of co-expression and experimental data, while *HSP17.61I* and *HSP21* were interacting in terms of co-expression and cooccurrence across the genome (Appendix: Figure A32). Co-expression interactions seemed to be most frequently observed between the defence response genes, however, other interactions were observed between these genes and other classes. In addition, nine defence response genes were down-regulated, including ethylene-responsive transcription factor 5 (ERF5), senescence-associated gene 12 (*SAG12*), BAK1-interacting receptor-like kinase 1 (*BIR1*) and *GST11* among others (Appendix: Figure A33). There were few interactions between these down-regulated defence response genes in the 98/0581 genotype.

4.2.9 Validation of expressed transcripts by RT-qPCR

The expression of nine candidate mealybug defence-related genes (Table 2.3) was validated using qRT-PCR with UBQ10 and GTPb as housekeeping genes. The validated candidate DEGs were selected based on their likely defence role in response to herbivore infestation as reported in other plant systems (Coppola *et al.*, 2013; De Vos *et al.*, 2005; Timm & Reineke, 2014; Wei *et al.*, 2013). The analysis of the data provided the log₂ ratios between time points in relation to both housekeeping genes UBQ10 and GTPb (Figure 4.14). Consistent with RNA-seq data, 2OG-FE(II) gene was among the most highly expressed at 24 hours and 72 hours of mealybug infestation, followed by CYP450, GH32 and Alpha/B family genes with reference to UBQ10 and GTPb. Although the two procedures (RNA-seq and qPCR) are very different, they could independently confirm the general trend of expression of the selected genes upon mealybug infestation (Figure 4.14).

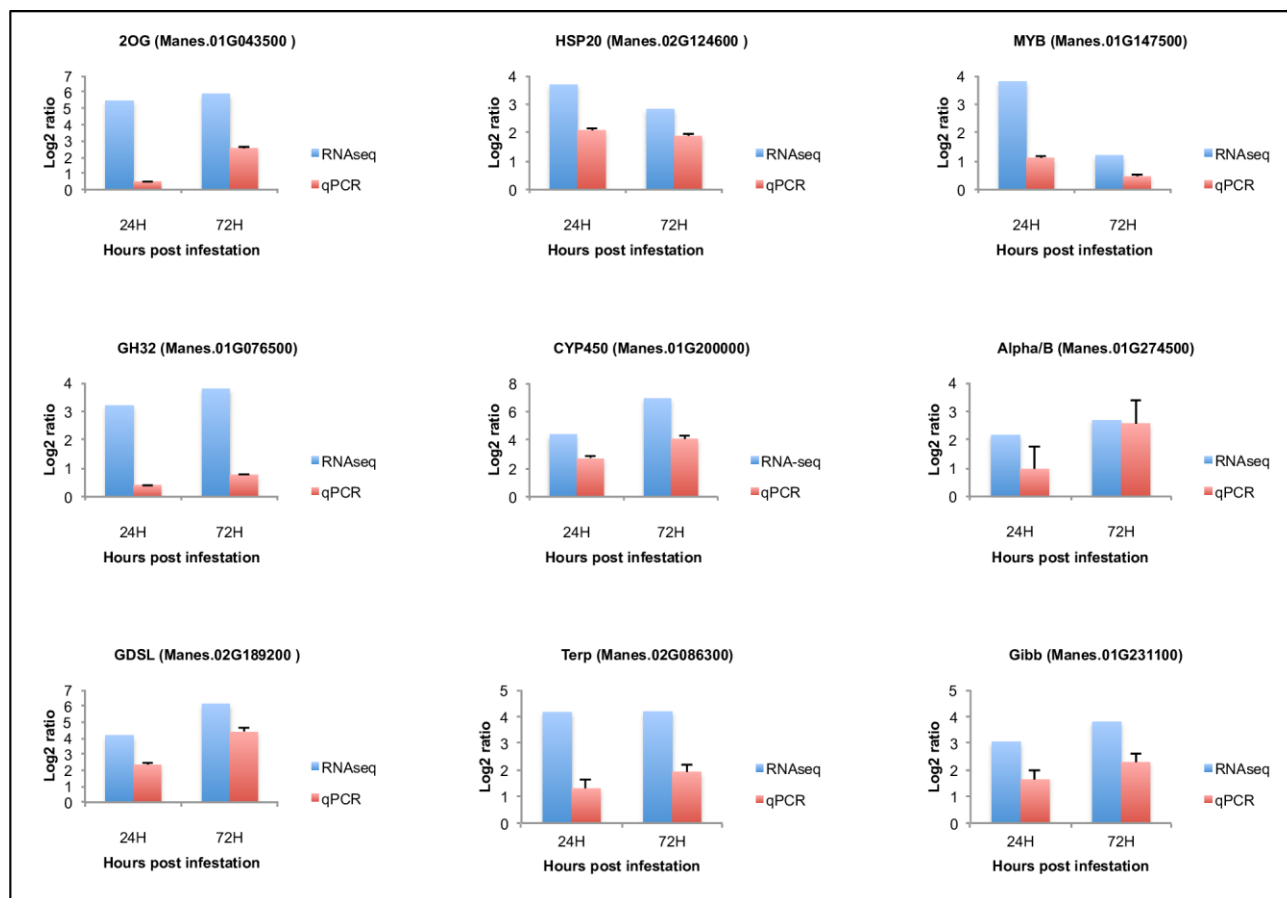


Figure 4.14: Comparative expression levels by RNA-seq and real-time RT-qPCR of nine cassava up-regulated genes in genotype AR23.1 in response to mealybug infestation at 24 and 72 hpi calculated with reference to UBQ10 and GTPb. Error bars on each qPCR column indicate the standard error (SE) from three biological replicates.

4.3 Discussion

The study established different response patterns to *P. manihoti* in the three genotypes, AR23.1, P40/1 and 98/0581. The findings point to partial resistance to *P. manihoti* in genotype AR23.1, susceptibility in P40/1 and tolerance in 98/0581 based on the observed phenotype after mealybug infestation (Table 4.1 and Figure 4.1). There was no significant damage to genotype AR23.1 other than minor changes in leaf colour and slight loss of leaf tissue by the 28 dpi, compared to the susceptible genotype P40/1 that wilted and died by 35-day post infection with mealybugs. The tolerant genotype, on the other hand, showed moderate symptoms (Table 4.1) while still developing new leaves even with high mealybug infestation. Resistance in genotype AR23.1 was hallmarked by minimal colonization by *P. manihoti* and a relatively long-term interaction (35 dpi) with mealybugs without showing severe symptoms of infestation/feeding. A similar observation whereby resistant cereals showed mild leaf chlorosis symptoms from aphid feeding has been reported in barley (*Hordeum vulgare*) and wheat (*Triticum aestivum*) with aphids (*Diuraphis noxia*) at nine days after infestation (Ni *et al.*, 2001). In addition, aphid (*Aphis glycine*) feeding in soybean (*Glycine max*) showed little or no visible symptoms in resistant varieties and partially resistant plant introductions at 28

days after infestation with aphids (Mian *et al.*, 2008). Moreover, cotton HR resistant cultivar was not visibly damaged in response to whitefly *Bemisia tabacci* infestation, whereas whiteflies caused obvious damage to the susceptible ZS cotton cultivar for one month (Li *et al.*, 2016). The population of the whiteflies was consistently higher in the susceptible cotton cultivar compared to the resistant cultivar over the three-year period. In another study by Klingler *et al.* (2009), a phenotypic evaluation of *Medicago truncatula* lines A20 (resistant) and A17 (susceptible) at 19 days after infestation with bluegreen aphid (*Acyrtosiphon kondio*) was demonstrated. Line A17 had developed damage symptoms, while line A20 showed no obvious results. The findings presented in this study suggest that *P. manihoti* may prefer to feed and reproduce on susceptible genotypes than on resistant genotype, however, this observation needs to be further examined at a larger scale. The observation from the results in this study further showed that tolerant and resistant genotypes have the ability to survive feeding by *P. manihoti* infestation.

Based on the phenotypic evaluation in this study, *P. manihoti* multiplication on the resistant and tolerant genotypes did not affect the survival of the cassava plants, compared to the susceptible genotype. However, as time progressed, the survival of the pest on the resistant genotype was reduced. Thus, the ability of AR23.1 to minimize *P. manihoti* colonization by 35 dpi could be the result of the release of cyanogenic glycoside toxins (Bellotti & Riis, 1994; Riis *et al.*, 2003; Steppuhn & Baldwin, 2007) deterring herbivorous pests from feeding on the host plants.

The response of AR23.1, P40/1 and 98/0581 genotypes to mealybug infestation was further investigated using high throughput RNA-sequencing technology, which allowed a detailed identification of differentially expressed genes. Differentially expressed transcripts at 24 and 72 hours of mealybug infestation were compared between a resistant (AR23.1), green mite-susceptible (P40/1) and a tolerant (98/0581) cassava genotype from Colombia, South Africa and Nigeria, respectively. Time points were selected for the analysis of transcripts at 24 and 72 hours of mealybug infestation based on previous transcriptomic studies (Coppola *et al.*, 2013; De Vos *et al.*, 2005; Timm & Reineke, 2014; Zhang *et al.*, 2015a). However, transcriptional changes have been observed as early as 1 hpi with diamondback moth (*Plutella xylostella*) in *Barbarea vulgaris* (Wei *et al.*, 2013; Zhang *et al.*, 2015b) and *Arabidopsis thaliana* (Ehlting *et al.*, 2008); and green peach aphids (*Myzus persicae*) in *Arabidopsis thaliana* (Couldridge *et al.*, 2007). In the case of mealybugs, grapevines has shown transcriptional changes at 6 hpi after feeding by grapevine mealybug *Planococcus ficus*; and as late as 96 hours (Timm & Reineke, 2014) and 120 hours of mealybug infestation in tomato (Zhang *et al.*, 2015a). A higher number (54, 36 and 47 for AR23.1, P40/1 and 98/0581, respectively) of down-regulated DEGs at 24 hours of mealybug infestation compared with up-regulated DEGs (38, 31 and 34 in AR23.1, P40/1 and 98/0581, respectively) in all three genotypes was observed (Table 4.2), demonstrating initial suppression of DEGs in response to mealybug feeding. Other studies have reported similar results in *Arabidopsis thaliana* (De Vos *et al.*, 2005), *Solanum lycopersicum*

(Coppola *et al.*, 2013), *Gossypium hirsutum* (Dubey *et al.*, 2013) and *Barbarea vulgaris* (Wei *et al.*, 2013), where initial suppression of DEGs in response to herbivore attack were observed. The suppression of DEGs may be a change in the metabolic status of the plant in response to stress, and has been reported in the interaction between sap-sucking insects (*Bemisia tabacci* and *Aphis gossypii*) with cotton (*Gossypium hirsutum*) (Dubey *et al.*, 2013). In contrast, a higher number of DEGs were up-regulated compared to down-regulated in response to mealybug feeding, and only one study has been reported (Timm & Reineke, 2014). However, most of the defence-related DEGs appeared to have been induced at 24 hours of mealybug infestation in AR23.1 and 98/0581 but not in the susceptible genotype P40/1, consistent with other studies with different insects in *Arabidopsis thaliana* (Ehltling *et al.*, 2008), *Glycine max* (Li *et al.*, 2008), *Medicago truncatula* (Gao *et al.*, 2010), *Brassica oleracea* (Broekgaarden *et al.*, 2007) and *Vitis vinifera* (Timm & Reineke, 2014).

Although more detailed studies will be required to establish the true mechanism of tolerance/resistance observed in AR23.1, it was possible to identify the induction of specific leucine-rich repeat (LRR) TIR-NBS-LRR and LRR receptor-like protein kinase (PK) resistance family proteins (Gururani *et al.*, 2012) in this study. TIR-NBS-LRR are involved in activation of defence signalling (Century *et al.*, 1997; Marone *et al.*, 2013; Parker *et al.*, 1996), while LRR-PK resistance family proteins are involved in regulation of developmental and defence-related processes (Halter *et al.*, 2014; Liebrand *et al.*, 2013; Torii, 2004). These R proteins were up-regulated in AR23.1 and were either absent or down-regulated in P40/1 (Appendix: Tables A2 and A3), suggesting that the R proteins may be involved in recognizing the presence of the insect and further, be playing a role in defence against mealybug feeding. Furthermore, some of these R proteins were also up-regulated in the 98/0581 genotype (Appendix: Table A4) compared to P40/1. Other defence-associated genes such as MYB domain transcription factors, HSP20-like chaperone superfamily proteins and calmodulin-like proteins were also differentially expressed in AR23.1 and 98/0581 and not in P40/1. Many of these induced defence-related proteins have also been observed in several other plant-insect and plant-pathogen interaction studies (Allie *et al.*, 2014; Artico *et al.*, 2014; Bricchi *et al.*, 2012; De Vos *et al.*, 2005; Timm & Reineke, 2014; Utsumi *et al.*, 2016; Wei *et al.*, 2013), supporting their role in defence against *P. manihoti*.

Cytochrome P450 family proteins are involved in the metabolism of plant allelochemicals which are induced as a result of abiotic and biotic stresses (An *et al.*, 2012; Artico *et al.*, 2014; Danielson, 2002; Johnson *et al.*, 2014; Wei *et al.*, 2013). Allelochemicals are secondary plant metabolites that can stimulate or inhibit growth or behaviour of the invading organism either negatively or positively. Allelochemicals with a negative allelopathic effects have been reported to play a role in defence against herbivory, through toxicity and thus repelling the insect. Cytochrome P450 family proteins were up- or down-regulated in the three genotypes, suggesting that they may not necessarily associate with plant defence in response to

mealybug feeding. A CYP94B3 protein (Manes.01G200000) was up-regulated in the resistant AR23.1 genotype at 24 and 72 hours of mealybug infestation (Appendix: Table A2 and A4), but not in the tolerant 98/0581 and susceptible P40/1 genotypes. Furthermore, CYP86A8 protein (Manes.01G059800) was up-regulated in the resistant and tolerant genotypes at 24 hours of mealybug infestation, and not in the susceptible genotype (Appendix: Table A2 and A4). CYP94B3 plays a role in catabolism and activation of plant hormone jasmonyl-L-Isoleucine (Koo *et al.*, 2011). CYP94B3 encodes jasmonoyl-L-isoleucine 12-hydroxylase (Kitaoka *et al.*, 2011), therefore, the expression of this gene in response to mealybug feeding suggests specific activity towards jasmonoyl-L-isoleucine. These findings implicate a role for CYP94B3 in cassava stress response following mealybug feeding only for the AR23.1 genotype. CYP86A8 on the other hand is involved in the biosynthesis of cutin in the epidermis (Duan & Schuler, 2005; Mizutani & Ohta, 2010), thereby the role of this gene in response to herbivore feeding is still unknown. However, these specific family proteins (CYP86 and CYP94) collectively play potential roles in the prevention of accumulation of toxic levels of free fatty acids (Duan & Schuler, 2005; Bak *et al.*, 2011). In addition, CYP82C2 (Manes.05G167500) was also up-regulated in the tolerant 98/0581 genotype and down-regulated in the susceptible P40/1 genotype, but was not differentially expressed in the resistant genotype AR23.1. This gene is involved in the biosynthetic pathway of 4-hydroxyindole-3-carboxyl nitrile, a cyanogenic metabolite required for inducible pathogen defence (Rajnaik *et al.*, 2015). The expression of this gene in the tolerant genotype and not in a susceptible genotype may be implicating its role in defence against *P. manihoti*. Although CYP82C2 has not been reported to play a role in response to herbivore attack, it has been implicated in the biosynthesis of novel pathogen-related secondary metabolites (Rajnaik *et al.*, 2015), as well as in elevating the response of plant defence genes and increased resistance to *Botrytis cinera* (Liu *et al.*, 2010). Since the CYP450 family proteins have a wide diversity in plants, more comprehensive experiments will be needed to determine their specific roles in response to *P. manihoti* feeding in cassava.

The KEGG database revealed the enrichment of genes under ABC transporters and MAPK signalling pathways specifically in AR23.1 and 98/0581, and not P40/1. ABC transporters play a role in organ growth, plant nutrition, plant development and in response to abiotic and biotic stresses (Kang *et al.*, 2011), while MAPKs play a role in signalling of a variety of abiotic and biotic stress (Pitzschke *et al.*, 2009). Elicitation of transporters and signalling molecules in the resistant AR23.1 and tolerant 98/0581 point to an induced defence response to mealybug feeding not observed in the susceptible genotype P40/1. The key mealybug-induced changes and DEGs involved in signalling and secondary metabolites have been summarised in this discussion in order to demonstrate the interactions between cassava plants and the pest.

CHAPTER 5

GENERAL DISCUSSION AND CONCLUSION

5.1 Introduction

Cultivation of cassava in Africa has been hampered by several biotic and abiotic factors resulting in major yield reductions. One of the major pests of the cassava crop is cassava mealybug (*Phenacoccus manihoti*), which is increasingly threatening the production of cassava in sub-Saharan Africa. Since the mealybug was first reported in Africa in the 1970s (Neuenschwander, 2001), it has become the major pest in most cassava growing regions. A transcriptome profiling study was undertaken to identify candidate defence-related genes that are involved in response to cassava mealybug in cassava. Previous studies in a variety of systems reported differentially expressed genes in response to herbivore attack (Bohorquez *et al.*, 2012; Coppola *et al.*, 2013; De Vos *et al.*, 2005; Timm & Reineke, 2014; Wei *et al.*, 2013; Zhang *et al.*, 2015a), however, only one study has been undertaken in cassava (Bohorquez *et al.*, 2012). This thesis reports, for the first time, transcriptome responses to feeding by cassava mealybug (*P. manihoti*) on cassava. This analysis will positively impact in the development of cassava mealybug resistant genotypes in future.

5.2 Morphological and molecular analysis of KwaZulu Natal and Mpumalanga mealybugs in South Africa

Controlling pests of economic importance involves identification of species, which can be done using both morphological and molecular methods. This study utilized both morphological and molecular methods to confirm the identity of mealybugs infesting cassava plants in KwaZulu Natal and Mpumalanga provinces of South Africa. Although morphological methods are slow and inefficient, previous studies relied on these methods for identification of mealybugs (Millar, 2002; Nagrare *et al.*, 2009; Watson & Kubiriba, 2005). These studies provide a confirmation that they are still reliable methods for species identification.

Molecular methods, on the other hand, are preferred because they are faster and any morphological stage of the mealybug can be used for species identification. The 18S and 28S rRNA genes investigated confirmed the identity of the species from KwaZulu Natal and Mpumalanga as *P. manihoti* in the present study. Other studies, such as Park *et al.* (2010a), preferred the use of molecular methods for identification of mealybugs infesting Korean pears, because morphological methods were found to be time-consuming and led to delays in shipment of fruits. Sethusa *et al.* (2014), however, reported the use of both morphological and molecular methods in identifying scale insects of economic importance, and molecular methods were confirmed to be more reliable than the morphological methods. Furthermore, this study has demonstrated the advantages of molecular methods for mealybug species identification compared to morphological methods. These results have proven that the concomitant usage of morphological and

molecular techniques are important and relevant for mealybug identification as the two methods complement or validate each other, thereby supporting the identification of the mealybugs in the present study, as confirmation of the *P. manihoti* species present in KwaZulu Natal and Mpumalanga province of South Africa.

5.3 Analysis of phenotypic responses of cassava to mealybug infestation

The phenotypic characteristics of different cassava genotypes were evaluated post-infestation with mealybugs, and varying responses of the three genotypes to mealybug infestation were observed. Resistance to mealybugs was identified in genotype AR23.1, which showed significantly higher levels of resistance than the 98/0581 and P40/1 when evaluated for a period of 5 weeks under infestation with mealybugs in the greenhouse. Natural variation for pest resistance has been observed in several plants in reaction to several pests including *Medicago truncatula* infested by aphids (Kamphuis *et al.*, 2012; Klingler *et al.*, 2009), *Arabidopsis thaliana* infested by diamondback moth (Handley *et al.*, 2005) and other insect herbivory systems (Bidart-Bouzar and Kliebenstein, 2008). In cassava, natural variation was observed in wild *Manihot tristis* (Carabalí *et al.*, 2010a; 2010b) and *Manihot esculenta* sub spp *flabelifolia* (Morante *et al.*, 2004). Given the lack of literature on breeding for resistance to mealybug in cassava, the current study will form an important reference point for relevant future studies in cassava, especially in the AR23.1 genotype.

Interestingly, with genotype AR23.1 exhibiting some resistance to mealybug, the results support the reports to date that plant responses to insect pests in general trigger a common cohort of defence-related genes (Broekgaarden *et al.*, 2007; Ehrling *et al.*, 2008; Li *et al.*, 2008; Li *et al.*, 2016; Timm & Reineke, 2014; Zhang *et al.*, 2015b). The results further suggest that *P. manihoti* feeding could have repressed the growth development of the P40/1 susceptible genotype, compared to the resistant AR23.1 and the tolerant 98/0581. Compared to our results, phenotypic evaluation of different host plants in response to insect herbivores has been reported with less damage on resistant lines compared to susceptible lines (Klingler *et al.*, 2009; Li *et al.*, 2016; Mian *et al.*, 2008; Ni *et al.*, 2001), supporting the results observed in the present study. Additionally, partial resistance in cultivated cassava has been shown (Tertuliano *et al.*, 1993). However, total resistance resulting in absence of mealybugs has not been identified in cultivated cassava, only in wild *Manihot* species. Thus, natural variation in resistance towards mealybugs can be exploited to provide sustainable crop protection against these insect herbivores.

5.4 Potential of transcript profiling for the identification of candidate defence-related genes in response to mealybugs (*P. manihoti*)

To determine if the cassava mealybug *P. manihoti* induced a cohort of specific defence-related genes, two resistant/tolerant cassava genotypes (AR23.1, 98/0581) were compared with the susceptible genotype P40/1. Mealybug-induced responses were compared between cassava genotypes at different time-points. A comparison of transcripts up- and down-regulated between the time-points provided insights into molecular processes involved in cassava-mealybug interactions. All three genotypes showed an increase in number of DEGs as infestation progressed, however, more down-regulated genes were observed between the time-points compared to up-regulated (Table 4.2). These findings were consistent with other studies with different insects infested onto different host plants (Coppola *et al.*, 2013; De Vos *et al.*, 2005; Dubey *et al.*, 2013; Wei *et al.*, 2013). Moreover, important evidence was also provided through the GO analysis of specific processes involved in transcriptome changes in cassava leaves in response to *P. manihoti* feeding. Gene ontology (GO) annotation of DEGs expressed in response to mealybug feeding revealed differential responses to stress, regulation of biological processes, cell and binding categories between the resistant (AR23.1), susceptible (P40/1) and tolerant (98/0581) genotypes (Figure 4.5). These functions have also been reported in other plants in response to both biotic and abiotic stress (An *et al.*, 2012; Wei *et al.*, 2013), supporting their role in response to mealybug feeding. Similar to our findings, a study by Wei *et al.* (2013) also demonstrated induced responses to insect feeding that resulted in the synthesis, regulation and/or transportation of anti-insect secondary metabolites and wound healing.

Insect feeding has been shown to result in both constitutive and induced defence mechanisms, which lead to activation of signal cascades and release of volatiles for protection (Kessler & Baldwin, 2002). The genotypes AR23.1 and 98/0581 showed a high level of enrichment for molecular and signal transducers under GO term classification (Figure 4.5). The role of signal transducers such as jasmonates in defence against insects has been extensively reported (Artico *et al.*, 2014; Broekgaarden *et al.*, 2008; De Vos *et al.*, 2005; Huang *et al.*, 2015). The enrichment of signal transducers in AR23.1 and 98/0581, and absence of the same class in P40/1 (Appendix: Tables A2-A4) strongly suggests a direct role in defence mechanisms in AR23.1 and 98/0581 against cassava mealybug. The specific up-regulation of defence-related genes in AR23.1 and 98/0581 (Appendix: Tables A2 and A4) and absence in P40/1 (Appendix: Tables A3) further supports this hypothesis, and also supports the conclusion that green mite AR23.1 and multiple pest-resistant 98/0581 genotypes could share similar defence responses in general.

In this study, gibberellin 2-oxidase 6 (Manes.01G231100), HSP20-like chaperones superfamily protein (Manes.02G124600), alpha/beta-Hydrolases superfamily protein (Manes.01G274500) and GDSL-like Lipase/acylhydrolase superfamily protein (Manes.02G189200) were up-regulated in AR23.1 but down-regulated in P40/1 at 24 h after mealybug infestation (Appendix: Tables A2 and A3), implicating these

genes in the biotic stress response to *P. manihoti* infestation (Beneventi *et al.*, 2013; Gao *et al.*, 2016; Gao *et al.*, 2010; Kumar *et al.*, 2015; Li *et al.*, 2016; Lo *et al.*, 2016; Utsumi *et al.*, 2016; Wang *et al.*, 2016). The same family proteins of the GDSL-like Lipase/acylhydrolase superfamily protein (Manes.013G130400) and the alpha/beta-Hydrolases superfamily protein (Manes.09G040500) were also up-regulated in the tolerant genotype 98/0581 (Appendix: Tables A4). Alpha/beta-hydrolase superfamily proteins function as hydrolases, hormone precursors or transporters and chaperones of other proteins (Lenfant *et al.*, 2013), while the GDSL-like lipase/acylhydrolase superfamily proteins play a role in biotic stress response by breaking down fatty acids, lipids and isoprenoids in plants (Rawat *et al.*, 2012). The observation that hydrolases and lipase/acylhydrolases are both up-regulated in resistant and tolerant cassava genotypes, but not susceptible, suggests that these are general abiotic stress responses, but were not as effective in tolerance, as mealybug damage was more severe in tolerant compared to resistant genotypes.

Gibberellin 2-oxidase family proteins are involved in determining the bioactive gibberellin content in plants which may result in dwarfism if reduced (Hu *et al.*, 2017; Zhou *et al.*, 2012), as well as in plant growth and development (Huttly & Phillips, 1995; Synek *et al.*, 2006). Overexpression of gibberellin 2-oxidase 6 had a great impact on vegetative and reproductive growth of transgenic physic nut (*Jatropha curcas*). In the present study, the up-regulation of the gibberellin 2-oxidase 6 genes in the resistant genotypes could be playing a role in sustaining plant growth and development following mealybug attack. On the other hand, the HSP20-like chaperone superfamily proteins are involved in the stability and accumulation of *R* genes, and in the coordination of the entire defence signalling (Lopes-Caitar *et al.* 2013). HSP20 genes were expressed in soybean genotypes inoculated by *Meloidogyne javanica* (Lopes-Caitar *et al.* 2013), and grapevines infested by *Planococcus ficus* (Timm & Reineke, 2014), which were similar to the genes expressed in response to abiotic stress. Similar results have been observed in the present study whereby HSP20 chaperones were expressed in response to mealybug feeding, suggesting that some of the HSP20 may exhibit functions that are unrelated to heat shock, such as biotic stress response. The diversity of functions within the HSP20 family has also been reported in other studies in response to herbivore feeding such as soybean (Prochaska *et al.*, 2015; Timm & Reineke, 2014).

Overall, the afore-mentioned genes have been associated with both abiotic and biotic stress responses in various crops including *Arabidopsis thaliana* (Bricchi *et al.*, 2012; Su & Li, 2008), cotton (Kumar *et al.*, 2015), soybean (Beneventi *et al.*, 2013), rice (Li *et al.*, 2016; Lo *et al.*, 2016; Rawat *et al.*, 2012) chickpea (Hiremath *et al.*, 2011) and cassava (Utsumi *et al.*, 2016).

Our results show a high level of induction of 2-oxoglutarate (2OG) and Fe (II)-dependent oxygenase superfamily proteins and MYB domain transcription factor in AR23.1, and not in P40/1. In addition, the same gene families were either up- or down-regulated in the 98/0581 genotype. The 2-oxoglutarate (2OG)

and Fe (II)-dependent oxygenase superfamily proteins have been reported to be involved in plant defence (Kawai *et al.*, 2014) by catalysing the formation of plant hormones while the MYB domain family proteins are involved in regulation of transcription factors in plants (Ambawat *et al.*, 2013; Roy, 2016). Previous studies have demonstrated up-regulation of these before-mentioned family genes in response to herbivore infestation (Bricchi *et al.*, 2012; Broekgaarden *et al.*, 2008; Coppola *et al.*, 2013; Timm & Reineke, 2014), suggesting that these are associated with insect feeding response. For instance, Coppola *et al.* (2013) and Artico *et al.* (2014) reported up-regulation of genes such as MYB domain transcription factor and 2-oxoglutarate (2OG) and Fe (II)-dependent oxygenase superfamily protein post-infestation by *Macrosiphum euphorbiae* in tomato and *Anthonomus grandis* in cotton, respectively, as well as in a study by Timm & Reineke (2014), in response to *Planococcus ficus* infestation in grapevines.

It was also interesting to note that different members of the same gene families can be found to be either induced and suppressed at the same time-points in the same genotype. For example, heat shock proteins, gibberellin-regulated proteins and cytochrome P450 family proteins were both up- and down-regulated at 24 and 72 hpi compared to mock in the partially resistant AR23.1 and 98/0581 genotypes (Appendix: Table A2 and A4). A similar trend was observed in a study by Broekgaarden *et al.* (2008) whereby members of protein families were reported to be induced or repressed in cabbage cultivars Rivera and Christmas Drumhead post-infestation by aphid *Brevicoryne brassicae*. These results point to diversity in the functions of the same gene family in response to herbivorous insects' feeding.

5.5 Signalling pathways involved in plant-mealybug interactions

Although we observed a significant number of DEGs, most of them were not assigned to specific KEGG Orthology (KO) terms which is partly due to the incomplete annotation of the cassava genome, making it difficult to use the KEGG database and establish complete pathway analyses. Nevertheless, the DEGs that were assigned KO terms in the present study demonstrated an abundance of genes playing roles in plant hormone signal transduction and plant-pathogen interactions. The plant hormone signal transduction pathways rely on signalling molecules such as abscisic acid ABA, JA, ET, and SA which are elicited in response to insect interactions (De Vos *et al.*, 2005). Mealybugs appeared to elicit a defence reaction based on the cross-communication of different hormone-related signalling pathways. In the present study, *P. manihoti* feeding suppressed more defence genes compared with induced genes in the susceptible P40/1, tolerant 98/0581 and partially resistant AR23.1 cassava genotypes. Many genes involved in JA, SA, ET and GA (Appendix: Tables A2-A4) biosynthesis were down-regulated post-infestation by mealybugs at different time-points. The suppression of JA-regulated genes and defence metabolites in tomato plants infested with *Phenacoccus solenopsis* was also reported in a recent study by Zhang *et al.* (2015a), demonstrating the intervention of JA pathway in defence response to *Phenacoccus solenopsis* feeding. Several studies in the past have confirmed the induction of signal transduction pathways and primary

induced defence responses upon pest infestation in different plants including *Arabidopsis thaliana* (Bricchi *et al.*, 2012), tomato *Solanum lycopersicum* (Coppola *et al.*, 2013), soybean (Li *et al.*, 2008) and *Medicago truncatula* (Klingler *et al.*, 2005).

In addition to PAMP-triggered immunity (PTI), effector-triggered immunity (ETI), represented by differentially expressed calcium binding and calmodulin proteins, heat shock proteins and transcription factors such as those of the MAPK and WRKY families among others (Appendix: Tables A2-A4) were observed in AR23.1, P40/1 and 98/0581 genotypes. ETI is also implicated, from other transcriptome studies, in herbivore infestation and pathogen infection (Artico *et al.*, 2014; Bricchi *et al.*, 2012; Coppola *et al.*, 2013; Reymond *et al.*, 2000; Tao *et al.*, 2003; Venu *et al.*, 2010; Zagrobelny *et al.*, 2009) as well as in other abiotic stress-related studies such as drought, cold stress and heat treatments (An *et al.*, 2012; Johnson *et al.*, 2014). Collectively, this and other studies suggest that the response to insect feeding or damage in plants could be highly conserved across different plant hosts and different insect species, and that commonalities in molecular responses with other abiotic stresses also occurs.

Cassava gene interaction networks were further generated to identify functional associations between the proteins that jointly contribute to the same functional process in the 3 genotypes, using the STRING web resource. In this study, we focused on the defence response genes (Figure 4.12 and 4.13; Appendix: Figure A24-27). Defence response genes (highlighted in red) showed satisfactory network interactions in terms of co-expression in all the genotypes (Figure 4.12 and 4.13; Appendix: Figure A24-27). The co-expression network distinguishes genes that display an interrelated expression pattern across a variety of samples (Jensen *et al.*, 2009; van Dam *et al.*, 2017). Furthermore, the interactions demonstrated the shortest path between genes, however, these genes have been well documented to function in plant defence in response to herbivorous insects (Artico *et al.*, 2014; Bricchi *et al.*, 2012; Coppola *et al.*, 2013; De Vos *et al.*, 2005; Timm & Reineke, 2014; Zhang *et al.*, 2015a). Other interactions including co-occurrence across the genome were also observed, however, not common in all the genotypes. Interacting networks were further observed between defence response genes and genes from other biological and molecular function categories. The gene network representation (Chapter 4.2.8) was in agreement with the KEGG and GO analysis, and also showed a clear representation of defence gene associations within each genotype.

In conclusion, transcript profiling of AR23.1, P40/1 and 98/0581 during the interaction with *P. manihoti* in the present study revealed that mealybug feeding induces transcriptome reprogramming associated with basal defence and abiotic stress hallmarks in cassava at a very early stage of infestation and feeding. However, the link between these early hormone signalling responses and downstream biochemical responses to tolerance/resistance to feeding needs to be established. In cassava, fewer studies have been published on transcriptomics (Allie *et al.*, 2014; Ballén-Taborda *et al.*, 2013; Bohorquez *et al.*, 2012;

Fregene *et al.*, 2004; Lopez *et al.*, 2005b; Owiti *et al.*, 2011; Utsumi *et al.*, 2016; Utsumi *et al.*, 2012; Wang *et al.*, 2014), and only one in response to herbivore attack (Bohorquez *et al.*, 2012). Reports demonstrating the involvement of secondary metabolites in deterring insect feeding (Duffey & Stout, 1996; Steppuhn & Baldwin, 2007) will require a metabolomics study to resolve this theory. The preliminary results provide strong evidence for the induction of plant disease resistance genes, hormonal signal transduction and plant-pathogen interaction pathways suggesting a role for these genes and pathways in response to mealybug feeding. These results form the basis for more detailed future studies on the specific role of some of the differentially expressed genes identified. The transcripts generated from this study will also contribute further to the cassava genome annotation.

The molecular mechanisms involved in response to the cassava-mealybug interaction of the 3 genotypes were characterised in this study. Although there has been extensive progress in understanding the plant-herbivore interactions in some plants, more research still needs to be undertaken in cassava in order to validate candidate genes for use in the development of mealybug resistant varieties. Further research to facilitate the development of new pest-resistant cassava cultivars would then be possible either through genomic selection strategies in breeding programs, or alternatively using gene-editing methods such as CRISPR/Cas9 to manipulate the function of genes including those identified in this study. The development of cassava genotypes with resistance to mealybug will be important in achieving sustainable cassava production.

REFERENCES

- Abd-Rabou, S., Shalaby, H., Germain, J.-F., Ris, N., Kreiter, P., & Malausa, T. (2012). Identification of mealybug pest species (Hemiptera: Pseudococcidae) in Egypt and France, using a DNA barcoding approach. *Bulletin of Entomological Research*, **102**, 515-523.
- Abhary, M., Siritunga, D., Stevens, G., Taylor, N. J., & Fauquet, C. M. (2011). Transgenic biofortification of the starchy staple cassava (*Manihot esculenta*) generates a novel sink for protein. *PLoS One*, **6**, 16256-16264.
- Adeniyi, O., Kovo, A., Abdulkareem, A., & Chukwudozie, C. (2007). Ethanol fuel production from cassava as a substitute for gasoline. *Journal of Dispersion Science and Technology*, **28**, 501-504.
- Adejare, G. O., & Coutts, R. H. A. (1981). Eradication of cassava mosaic disease from Nigerian cassava clones by meristem-tip culture. *Plant cell, Tissue and Organ Culture*, **1**, 25-32.
- Aerni, P. (2006). Mobilizing science and technology for development: The case of the Cassava Biotechnology Network (CBN). *AgBioForum*, **9**, 1-14.
- Aidoo, R., Osekere, E. A., Logah, V., & Bakang, J.-E. A. (2016). Economic benefits of biological control of cassava green mite (CGM) in Ghana. *Journal of Development and Agricultural Economics*, **8**, 172-185.
- Akano, A., Dixon, A., Mba, C., Barrera, E., & Fregene, M. (2002). Genetic mapping of a dominant gene conferring resistance to cassava mosaic disease. *Theoretical and Applied Genetics*, **105**, 521-525.
- Akaracharanya, A., Kesornsit, J., Leepipatpiboon, N., Srinorakutara, T., Kitpreechavanich, V., & Tolieng, V. (2011). Evaluation of the waste from cassava starch production as a substrate for ethanol fermentation by *Saccharomyces cerevisiae*. *Annals of Microbiology*, **61**, 431-436.
- Akinbo, O., Labuschagne, M., & Fregene, M. (2012). Introgression of whitefly (*Aleurotrachelus socialis*) resistance gene from F1 inter-specific hybrids into commercial cassava. *Euphytica*, **183**, 19-26.
- Alamalakala, L., Foster, J. E. & Skoda, S. R. (2009). Amplified fragment length polymorphisms used for inter- and intraspecific differentiation of screwworms (Diptera: Calliphoridae). *Bulletin of Entomological Research*, **99**, 139-149.
- Ali, J. G., & Agrawal, A. A. (2012). Specialist versus generalist insect herbivores and plant defense. *Trends in Plant Science*, **17**, 293-302.
- Allem, A. C. (1994). The origin of *Manihot esculenta* Crantz (Euphorbiaceae). *Genetic Resources and Crop Evolution*, **41**, 133-150.
- Allem, A. C. (1999). The closest wild relatives of cassava (*Manihot esculenta* Crantz). *Euphytica*, **107**, 123-133.
- Allem, A. C. (2002). The origins and taxonomy of cassava. In: Hillocks, R.J., Thresh, J.M. and Bellotti, A.C. (eds). *Cassava: Biology, Production and Utilization*, CABI Publishing, CIAT, Cali, Colombia and University of Greenwich, Kent, UK, pp 1-16.
- Allie, F., Pierce, E. J., Okoniewski, M. J., & Rey, C. (2014). Transcriptional analysis of South African cassava mosaic virus-infected susceptible and tolerant landraces of cassava highlights differences in resistance, basal defense and cell wall associated genes during infection. *BMC Genomics*, **15**, 1006-1035.
- Alves, A. A. C. (2002). Cassava botany and physiology. In: Hillocks, R.J., Thresh, J.M. and Bellotti, A.C. (eds). *Cassava: Biology, Production and Utilization*, CABI Publishing, CIAT, Cali, Colombia and University of Greenwich, Kent, UK, pp 67-89.
- Ambawat, S., Sharma, P., Yadav, N. R., & Yadav, R. C. (2013). MYB transcription factor genes as regulators for plant responses: an overview. *Physiology and Molecular Biology of Plants*, **19**, 307-321.
- An, D., Yang, J., & Zhang, P. (2012). Transcriptome profiling of low temperature-treated cassava apical shoots showed dynamic responses of tropical plant to cold stress. *BMC Genomics*, **13**, 64-87.
- Anderson, J. P., Badruzsaufari, E., Schenk, P. M., Manners, J. M., Desmond, O. J., Ehlert, C., Maclean, D. J., Ebert, P. R., & Kazan, K. (2004). Antagonistic interaction between abscisic acid and jasmonate-ethylene signaling pathways modulates defense gene expression and disease resistance in *Arabidopsis*. *The Plant Cell*, **16**, 3460-3479.

- Anjanappa, R. B., Mehta, D., Maruthi, M., Kanju, E., Gruissem, W., & Vanderschuren, H. (2016). Characterization of Brown Streak Virus–Resistant Cassava. *Molecular Plant-Microbe Interactions*, **29**, 527-534.
- Ansorge, W. J. (2009). Next-generation DNA sequencing techniques. *New Biotechnology*, **25**, 195-203.
- Arimura, G.-i., Kost, C., & Boland, W. (2005). Herbivore-induced, indirect plant defences. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids*, **1734**, 91-111.
- Artico, S., Ribeiro-Alves, M., Oliveira-Neto, O. B., de Macedo, L. L., Silveira, S., Grossi-de-Sa, M. F., Martinelli, A. P., & Alves-Ferreira, M. (2014). Transcriptome analysis of *Gossypium hirsutum* flower buds infested by cotton boll weevil (*Anthonomus grandis*) larvae. *BMC Genomics*, **15**, 854-877.
- Ashelford, K., Eriksson, M. E., Allen, C. M., D'Amore, R., Johansson, M., Gould, P., Kay, S., Millar, A. J., Hall, N., & Hall, A. (2011). Full genome re-sequencing reveals a novel circadian clock mutation in *Arabidopsis*. *Genome Biology*, **12**, 28-39.
- Bak, S., Beisson, F., Bishop, G., Hamberger, B., & Höfer, R. (2011). Cytochromes P450. *BioOne*, **9**, 1-56.
- Balogopalan, C. (2002). Cassava utilization in food, feed and industry. In: Hillocks, R.J., Thresh, J.M. & Bellotti, A.C. (eds). *Cassava: Biology, Production and Utilization*, Central Tuber Crops Research Institute, Sreekariyam, Kerala, India, pp 301-318.
- Ballén-Taborda, C., Plata, G., Ayling, S., Rodríguez-Zapata, F., Becerra Lopez-Lavalle, L. A., Duitama, J., & Tohme, J. (2013). Identification of cassava MicroRNAs under abiotic stress. *International Journal of Genomics*, **2013**, 1-10.
- Bari, R., & Jones, J. D. (2009). Role of plant hormones in plant defence responses. *Plant Molecular Biology*, **69**, 473-488.
- Beckers, G., & Spoel, S. (2006). Fine-tuning plant defence signalling: salicylate versus jasmonate. *Plant Biology*, **8**, 1-10.
- Bellotti, A. C. (2002). Arthropod pests. In: Hillocks, R.J., Thresh, J.M. and Bellotti, A.C. (eds). *Cassava: Biology, Production and Utilization*, CABI Publishing, CIAT, Cali, Colombia and University of Greenwich, Kent, UK, pp 209-235.
- Bellotti, A., Herrera, B.V., & Hyman, G. (2012). Cassava production and pest management: present and future potential threats in a changing environment. *Tropical Plant Biology*, **5**, 39-72.
- Bellotti, A. C., & Riis, L. (1994). Cassava cyanogenic potential and resistance to pests and diseases. In: *International Workshop on Cassava Safety*, Wageningen, The Netherlands, pp 375.
- Bellotti, A. C., Smith, L., & Lapointe, S. L. (1999). Recent advances in cassava pest management. *Annual Review of Entomology*, **44**, 343-370.
- Beneventi, M. A., da Silva, O. B., de Sá, M. E. L., Firmino, A. A. P., de Amorim, R. M. S., Albuquerque, É. V. S., da Silva, M. C. M., da Silva, J. P., de Araújo Campos, M., & Lopes, M. J. C. (2013). Transcription profile of soybean-root-knot nematode interaction reveals a key role of phytohormones in the resistance reaction. *BMC Genomics*, **14**, 322-338.
- Bengyella, L., Waikhom, S. D., Allie, F., & Rey, C. (2015). Virus tolerance and recovery from viral induced-symptoms in plants are associated with transcriptome reprogramming. *Plant Molecular Biology*, **89**, 243-252.
- Bernoux, M., Ve, T., Williams, S., Warren, C., Hatters, D., Valkov, E., Zhang, X., Ellis, J. G., Kobe, B., & Dodds, P. N. (2011). Structural and functional analysis of a plant resistance protein TIR domain reveals interfaces for self-association, signaling, and autoregulation. *Cell Host Microbe*, **9**, 200-211.
- Bertschy, C., Turlings, T. C., Bellotti, A. C., & Dorn, S. (1997). Chemically-mediated attraction of three parasitoid species to mealybug-infested cassava leaves. *Florida Entomologist*, **80** 383-395.
- Beuning, L., Murphy, P., Wu, E., Batchelor, T., & Morris, B. (1999). Molecular-based approach to the differentiation of mealybug (Hemiptera: Pseudococcidae) species. *Journal of Economic Entomology*, **92**, 463-472.
- Bi, H., Aileni, M., & Zhang, P. (2010). Evaluation of cassava varieties for cassava mosaic disease resistance jointly by agro-inoculation screening and molecular markers. *African Journal of Biotechnology*, **4**, 330-338.
- Bianchi, V., Colantoni, A., Calderone, A., Ausiello, G., Ferre, F., & Helmer-Citterich, M. (2013). DBATE: database of alternative transcripts expression. *Database*, **2013**, 50-59.

- Bidart-Bouzat, M. G., & Kliebenstein, D. J. (2008). Differential levels of insect herbivory in the field associated with genotypic variation in glucosinolates in *Arabidopsis thaliana*. *Journal of Chemical Ecology*, **34**, 1026-1037.
- Biere, A., & Govere, A. (2016). Plant-mediated systemic interactions between pathogens, parasitic nematodes, and herbivores above-and belowground. *Annual Review of Phytopathology*, **54**, 499-527.
- Birol, I., Jackman, S. D., Nielsen, C. B., Qian, J. Q., Varhol, R., Stazyk, G., Morin, R. D., Zhao, Y., Hirst, M., & Schein, J. E. (2009). *De novo* transcriptome assembly with ABySS. *Bioinformatics*, **25**, 2872-2877.
- Boher, B., & Verdier, V. (2009). Cassava bacterial blight in Africa: The state of knowledge and implications for designing control strategies. *African Crop Science Journal*, **2**, 505-509.
- Bohorquez, A., Becerra López-Lavalle, L. A., Bellotti, A. C., & Tohme, J. M. (2012). Unraveling whitefly resistance in cassava (*Manihot esculenta*). [Online] CIAT, Cali, Colombia. Available from: http://ciat-library.ciat.cgiar.org/Articulos_Ciat/biblioteca/Unraveling%20whitefly%20resistance%20in%20cassava.pdf.
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114-2120.
- Bredeson, J. V., Lyons, J. B., Prochnik, S. E., Wu, G. A., Ha, C. M., Edsinger-Gonzales, E., Grimwood, J., Schmutz, J., Rabbi, I. Y., & Egesi, C. (2016). Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. *Nature Biotechnology*, **34**, 562-570.
- Bricchi, I., Berte, C. M., Occhipinti, A., Paponov, I. A., & Maffei, M. E. (2012). Dynamics of membrane potential variation and gene expression induced by *Spodoptera littoralis*, *Myzus persicae*, and *Pseudomonas syringae* in *Arabidopsis*. *PloS One*, **7**, 46673-46692.
- Broekgaarden, C., Poelman, E. H., Steenhuis, G., Voorrips, R. E., Dicke, M., & Vosman, B. (2007). Genotypic variation in genome-wide transcription profiles induced by insect feeding: *Brassica oleracea*-*Pieris rapae* interactions. *BMC Genomics*, **8**, 239-251.
- Broekgaarden, C., Poelman, E. H., Steenhuis, G., Voorrips, R. E., Dicke, M., & Vosman, B. (2008). Responses of *Brassica oleracea* cultivars to infestation by the aphid *Brevicoryne brassicae*: an ecological and molecular approach. *Plant, Cell & Environment*, **31**, 1592-1605.
- Broekgaarden, C., Snoeren, T. A., Dicke, M., & Vosman, B. (2011). Exploiting natural variation to identify insect- resistance genes. *Plant Biotechnology Journal*, **9**, 819-825.
- Buermans, H., & Den Dunnen, J. (2014). Next generation sequencing technology: advances and applications. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*, **1842**, 1932-1941.
- Bull, S., Owiti, J., Niklaus, M., Beeching, J. R., Grissem, W., & Vanderschuren, H. (2009). Agrobacterium-mediated transformation of friable embryogenic calli and regeneration of transgenic cassava. *Nature Protocols*, **4**, 1845-1854.
- Caarls, L., Pieterse, C. M., & Van Wees, S. C. (2015). How salicylic acid takes transcriptional control over jasmonic acid signaling. *Frontiers in Plant Science*, **6**, 1-11.
- Cahais, V., Gayral, P., Tsagkogeorga, G., Melo- Ferreira, J., Ballenghien, M., Weinert, L., Chiari, Y., Belkhir, K., Ranwez, V., & Galtier, N. (2012). Reference- free transcriptome assembly in non-model animals from next- generation sequencing data. *Molecular Ecology Resources*, **12**, 834-845.
- Calatayud, P. A., & Le Rü, B. P. (2006). Cassava-mealybug interactions: In: Institut De Recherche (eds), CIAT, Cali, Colombia pp 11-21.
- Calatayud, P. A., Rahbé, Y., Tjallingii, W., Tertuliano, M., & Rü, B. (1994). Electrically recorded feeding behaviour of cassava mealybug on host and non- host plants. *Entomologia Experimentalis et Applicata*, **72**, 219-232.
- Calvert, L., & Thresh, J. M. (2002). The viruses and virus diseases of cassava. In: Hillocks, R.J., Thresh, J.M. and Bellotti, A.C. (eds). Cassava: Biology, Production and Utilization, CIAT, Cali, Colombia and University of Greenwich, Kent, UK, pp 237-260.
- Carabalí, A., Bellotti, A., Montoya-Lerma, J., & Fregene, M. (2010a). Resistance to the whitefly, *Aleurotrachelus socialis*, in wild populations of cassava, *Manihot tristis*. *Journal of Insect Science*, **10**, 170-180.

- Carabalí, A., Bellotti, A. C., Montoya-Lerma, J., & Fregene, M. (2010b). *Manihot flabellifolia* Pohl, wild source of resistance to the whitefly *Aleurotrachelus socialis* Bondar (Hemiptera: Aleyrodidae). *Crop Protection*, **29**, 34-38.
- Cavaliere, V., Mazzeo, G., Garzia, G. T., Buonocore, E., & Russo, A. (2008). Identification of *Planococcus ficus* and *Planococcus citri* (Hemiptera: Pseudococcidae) by PCR-RFLP of coi gene. *Zootaxa*, **1816**, 65-68.
- Ceballos, H., Iglesias, C. A., Pérez, J. C., & Dixon, A. G. (2004). Cassava breeding: opportunities and challenges. *Plant Molecular Biology*, **56**, 503-516.
- Chang, K. N., Zhong, S., Weirauch, M. T., Hon, G., Pelizzola, M., Li, H., Huang, S.-s. C., Schmitz, R. J., Urich, M. A., & Kuo, D. (2013). Temporal transcriptional response to ethylene gas drives growth hormone cross-regulation in *Arabidopsis*. *Elife*, **2**, 1-20.
- Chauhan, R. D., Beyene, G., Kalyaeva, M., Fauquet, C. M., & Taylor, N. (2015). Improvements in Agrobacterium-mediated transformation of cassava (*Manihot esculenta* Crantz) for large-scale production of transgenic plants. *Plant Cell, Tissue and Organ Culture (PCTOC)*, **121**, 591-603.
- Chauynarong, N., Elangovan, A., & Iji, P. (2009). The potential of cassava products in diets for poultry. *World's Poultry Science Journal*, **65**, 23-36.
- Chavarriaga-Aguirre, P., & Halsey, M. (2005). Cassava (*Manihot esculenta* Crantz): Reproductive biology and practices for confinement of experimental field trials. In: *Report prepared for the Program for Biosafety Systems*. Washington, DC, USA, pp 20.
- Chen, J., Dai, Z., Cao, C., Zhang, Q., Liu, H., & Sun, X. (2012). Next-generation sequencing data processing: Analysis of unmapped reads and extremely high mapped peaks. In: *5th International Conference on Biomedical Engineering and Informatics (BMEI)*, Chongqing, China, pp 893-897.
- Chen, M. S. (2008). Inducible direct plant defense against insect herbivores: a review. *Insect Science*, **15**, 101-114.
- Chen, S., Yang, P., Jiang, F., Wei, Y., Ma, Z., & Kang, L. (2010). *De novo* analysis of transcriptome dynamics in the migratory locust during the development of phase traits. *PloS One*, **5**, 1-15.
- Chung, H. S., Niu, Y., Browse, J., & Howe, G. A. (2009). Top hits in contemporary JAZ: an update on jasmonate signaling. *Phytochemistry*, **70**, 1547-1559.
- Cipollini, D., Purrington, C. B., & Bergelson, J. (2003). Costs of induced responses in plants. *Basic and Applied Ecology*, **4**, 79-89.
- Cock, J. H. (1985). Cassava: New potential for a neglected crop: Steven A. Breth (ed), Westview Press, Colorado, USA, pp 12-22.
- Cook, L., Edwards, R., Crisp, M., & Hardy, N. (2010). Need morphology always be required for new species descriptions? *Invertebrate Systematics*, **24**, 322-326.
- Cooper, W., & Goggin, F. (2005). Effects of jasmonate- induced defenses in tomato on the potato aphid, *Macrosiphum euphorbiae*. *Entomologia Experimentalis et Applicata*, **115**, 107-115.
- Coppola, M., Cascone, P., Chiusano, M. L., Colantuono, C., Lorito, M., Pennacchio, F., Rao, R., Woo, S. L., Guerrieri, E., & Digilio, M. C. (2017). *Trichoderma harzianum* enhances tomato indirect defense against aphids. *Insect Science*. doi:10.1111/1744-7917.12475
- Coppola, V., Coppola, M., Rocco, M., Digilio, M. C., D'Ambrosio, C., Renzone, G., Martinelli, R., Scaloni, A., Pennacchio, F., & Rao, R. (2013). Transcriptomic and proteomic analysis of a compatible tomato-aphid interaction reveals a predominant salicylic acid-dependent plant response. *BMC Genomics*, **14**, 515-532.
- Couldridge, C., Newbury, H., Ford-Lloyd, B., Bale, J., & Pritchard, J. (2007). Exploring plant responses to aphid feeding using a full *Arabidopsis* microarray reveals a small number of genes with significantly altered expression. *Bulletin of Entomological Research*, **97**, 523-532.
- da Silva, V. C. P., Bertin, A., Blin, A., Germain, J.-F., Bernardi, D., Rignol, G., Botton, M., & Malausa, T. (2014). Molecular and morphological identification of mealybug species (Hemiptera: Pseudococcidae) in Brazilian vineyards. *PLoS One*, **9**, 1-13.
- Daane, K. M., Almeida, R. P., Bell, V. A., Walker, J. T., Botton, M., Fallahzadeh, M., Mani, M., Miano, J. L., Sforza, R., & Walton, V. M. (2012). Biology and management of mealybugs in vineyards. In: Bostanian N.J., Isaacs R. and Vincent V. (eds) *Arthropod Management in Vineyards*. pp 271-307
- Dangl, J. L., & Jones, J. D. (2001). Plant pathogens and integrated defence responses to infection. *Nature*, **411**, 826-833.

- Danielson, P. (2002). The cytochrome *P450* superfamily: biochemistry, evolution and drug metabolism in humans. *Current Drug Metabolism*, **3**, 561-597.
- Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, **12**, 499-510.
- Dawit, B. (2010). Micropropagation of selected cassava varieties (*Manihot esculenta* Crantz) from meristem culture. A Thesis submitted for a Masters of Science in Biological Science, Addis Ababa University, Addis Ababa, Ethiopia.
- De Cosa, B., Moar, W., Lee, S.-B., Miller, M., & Daniell, H. (2001). Overexpression of the *Bt Cry2Aa2* operon in chloroplasts leads to formation of insecticidal crystals. *Nature Biotechnology*, **19**, 71-74.
- de Ilarduya, O. M., Xie, Q., & Kaloshian, I. (2003). Aphid-induced defense responses in *Mi-1*-mediated compatible and incompatible tomato interactions. *Molecular Plant-Microbe Interactions*, **16**, 699-708.
- Demontis, M. A., Ortu, S., Cocco, A., Lentini, A., & Migheli, Q. (2007). Diagnostic markers for *Planococcus ficus* (Signoret) and *Planococcus citri* (Risso) by random amplification of polymorphic DNA-polymerase chain reaction and species-specific mitochondrial DNA primers. *Journal of Applied Entomology*, **131**, 59-64.
- De Vos, M., Van Oosten, V. R., Van Poecke, R. M., Van Pelt, J. A., Pozo, M. J., Mueller, M. J., Buchala, A. J., Métraux, J.P., Van Loon, L., & Dicke, M. (2005). Signal signature and transcriptome changes of *Arabidopsis* during pathogen and insect attack. *Molecular Plant-Microbe Interactions*, **18**, 923-937.
- De Vos, M., Van Zaanen, W., Koornneef, A., Korzelius, J. P., Dicke, M., Van Loon, L., & Pieterse, C. M. (2006). Herbivore-induced resistance against microbial pathogens in *Arabidopsis*. *Plant Physiology*, **142**, 352-363.
- Dogimont, C., Bendahmane, A., Chovelon, V., & Boissot, N. (2010). Host plant resistance to aphids in cultivated crops: genetic and molecular bases, and interactions with aphid populations. *Comptes Rendus Biologies*, **333**, 566-573.
- Dong, X. (1998). SA, JA, ethylene, and disease resistance in plants. *Current Opinion in Plant Biology*, **1**, 316-323.
- Douglas, N., & Krüger, K. (2008). Transmission efficiency of grapevine leafroll-associated virus (*glrav-3*) by the mealybugs *Planococcus ficus* and *Pseudococcus longispinus* (Hemiptera: Pseudococcidae). *European Journal of Plant Pathology*, **122**, 207-212.
- Dowton, M., & Austin, A. (1998). Phylogenetic relationships among the microgastroid wasps (Hymenoptera: Braconidae): combined analysis of 16S and 28S rDNA genes and morphological data. *Molecular Phylogenetics and Evolution*, **10**, 354-366.
- Du, B., Zhang, W., Liu, B., Hu, J., Wei, Z., Shi, Z., He, R., Zhu, L., Chen, R., & Han, B. (2009). Identification and characterization of *Bph14*, a gene conferring resistance to brown planthopper in rice. *Proceedings of the National Academy of Sciences*, **106**, 22163-22168.
- Duan, H., & Schuler, M.A. (2005). Differential expression and evolution of the *Arabidopsis* CYP86A subfamily. *Plant Physiology*, **137**, 1067-1081.
- Dubey, N. K., Goel, R., Ranjan, A., Idris, A., Singh, S. K., Bag, S. K., Chandrashekar, K., Pandey, K. D., Singh, P. K., & Sawant, S. V. (2013). Comparative transcriptome analysis of *Gossypium hirsutum* L. in response to sap sucking insects: aphid and whitefly. *BMC Genomics*, **14**, 241-261.
- Dudareva, N., Klempien, A., Muhlemann, J. K., & Kaplan, I. (2013). Biosynthesis, function and metabolic engineering of plant volatile organic compounds. *New Phytologist*, **198**, 16-32.
- Duffey, S. S., & Stout, M. J. (1996). Antinutritive and toxic components of plant defense against insects. *Archives of Insect Biochemistry and Physiology*, **32**, 3-37.
- Egan, A. N., Schlueter, J., & Spooner, D. M. (2012). Applications of next-generation sequencing in plant biology. *American Journal of Botany*, **99**, 175-185.
- Ehrling, J., Chowrira, S. G., Mattheus, N., Aeschliman, D. S., Arimura, G.-I., & Bohlmann, J. (2008). Comparative transcriptome analysis of *Arabidopsis thaliana* infested by diamond back moth (*Plutella xylostella*) larvae reveals signatures of stress response, secondary metabolism, and signalling. *BMC Genomics*, **9**, 154-173.

- Falade, K.O., & Akingbala, J.O. (2010). Utilization of cassava for food. *Food Reviews International*, **27**, 51-83.
- FAOSTAT. (2016). Global cassava production trends, 2005-2014. Available: www.faostat3.fao.org/browse/rankings/commodities_by_regions/E.
- Ferreira-Leitao, V., Gottschalk, L. M. F., Ferrara, M. A., Nepomuceno, A. L., Molinari, H. B. C., & Bon, E. P. (2010). Biomass residues in Brazil: availability and potential uses. *Waste and Biomass Valorization*, **1**, 65-76.
- Fokunang, C., Iktun, T., Dixon, A., & Akem, C. (2000). Field reaction of cassava genotypes to anthracnose, bacterial blight, cassava mosaic disease and their effects on yield. *African Crop Science Journal*, **8**, 179-186.
- Fregene, M., Angel, F., Gómez, R., Rodríguez, F., Chavarriaga, P., Roca, W., Tohme, J., & Bonierbale, M. (1997). A molecular genetic map of cassava (*Manihot esculenta* Crantz). *Theoretical and Applied Genetics*, **95**, 431-441.
- Fregene, M., Matsumura, H., Akano, A., Dixon, A., & Terauchi, R. (2004). Serial analysis of gene expression (SAGE) of host-plant resistance to the cassava mosaic disease (CMD). *Plant Molecular Biology*, **56**, 563-571.
- Fregene, M., & Puonti-Kaerlas, J. (2002). Cassava biotechnology. In: Hillocks, R.J., Thresh, J.M. and Bellotti, A.C. (eds). *Cassava: Biology, Production and Utilization*, CABI Publishing, CIAT, Cali, Colombia and University of Greenwich, Kent, UK, pp 179-207.
- Fu, Z. Q., & Dong, X. (2013). Systemic acquired resistance: turning local infection into global defense. *Annual Review of Plant Biology*, **64**, 839-863.
- Gao, L., Wang, Y., Li, Z., Zhang, H., Ye, J., & Li, G. (2016). Gene Expression Changes during the gummosis development of peach shoots in response to *Lasiodiplodia theobromae* infection using RNA-Seq. *Frontiers in Physiology*, **7**, 1-12.
- Gao, L. L., Kamphuis, L. G., Kakar, K., Edwards, O. R., Udvardi, M. K., & Singh, K. B. (2010). Identification of potential early regulators of aphid resistance in *Medicago truncatula* via transcription factor expression profiling. *New Phytologist*, **186**, 980-994.
- Gatehouse, J. A. (2002). Plant resistance towards insect herbivores: a dynamic interaction. *New Phytologist*, **156**, 145-169.
- Glazebrook, J. (2005). Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annual Review of Phytopathology*, **43**, 205-227.
- Gleadow, R. M., Evans, J. R., McCaffery, S., & Cavagnaro, T. R. (2009). Growth and nutritive value of cassava (*Manihot esculenta* Cranz.) are reduced when grown in elevated CO₂. *Plant Biology*, **11**, 76-82.
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., & Zeng, Q. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, **29**, 644-652.
- Grant, M., & Lamb, C. (2006). Systemic immunity. *Current Opinion in Plant Biology*, **9**, 414-420.
- Greninger, A.L., Naccache, S. N., Federman, S., Yu, G., Mbala, P., Bres, V., Stryke, D., Bouquet, J., Soma sekhar, S., Linnen, J. M., Dodd, R., Mulembakani, P., Schneider, B. S., Muyembe-Tamfum, J. J., Stramer, S. L., & Chiu, C. Y. (2015). Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis. *Genome Medicine*, **7**, 1-13.
- Guo, H., & Ecker, J. R. (2004). The ethylene signaling pathway: new insights. *Current Opinion in Plant Biology*, **7**, 40-49.
- Gururani, M. A., Venkatesh, J., Upadhyaya, C. P., Nookaraju, A., Pandey, S. K., & Park, S. W. (2012). Plant disease resistance genes: current status and future directions. *Physiological and Molecular Plant Pathology*, **78**, 51-65.
- Gutierrez, A., Yaninek, J., Wermelinger, B., Herren, H., & Ellis, C. (1988). Analysis of biological control of cassava pests in Africa. III. Cassava green mite *Mononychellus tanajoa*. *Journal of Applied Ecology*, **25**, 941-950.
- Gutsche, A., Heng-Moss, T., Sarath, G., Twigg, P., Xia, Y., Lu, G., & Mornhinweg, D. (2009). Gene expression profiling of tolerant barley in response to *Diuraphis noxia* (Hemiptera: Aphididae) feeding. *Bulletin of Entomological Research*, **99**, 163-173.

- Gutzat, R., & Scheid, O. M. (2012). Epigenetic responses to stress: triple defense? *Current Opinion in Plant Biology*, **15**, 568-573.
- Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., & Lieber, M. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, **8**, 1494-1512.
- Halitschke, R., & Baldwin, I. T. (2004). Jasmonates and related compounds in plant-insect interactions. *Journal of Plant Growth Regulation*, **23**, 238-245.
- Halitschke, R., Schittko, U., Pohnert, G., Boland, W., & Baldwin, I. T. (2001). Molecular interactions between the specialist herbivore *Manduca sexta* (Lepidoptera, Sphingidae) and its natural host *Nicotiana attenuata*. III. Fatty acid-amino acid conjugates in herbivore oral secretions are necessary and sufficient for herbivore-specific plant responses. *Plant Physiology*, **125**, 711-717.
- Hall, N. (2007). Advanced sequencing technologies and their wider impact in microbiology. *Journal of Experimental Biology*, **210**, 1518-1525.
- Halsey, M. E., Olsen, K. M., Taylor, N. J., & Chavarriaga-Aguirre, P. (2008). Reproductive Biology of Cassava (Crantz) and Isolation of Experimental Field Trials. *Crop Science*, **48**, 49-58.
- Halter, T., Imkampe, J., Mazzotta, S., Wierzba, M., Postel, S., Bücherl, C., Kiefer, C., Stahl, M., Chinchilla, D., & Wang, X. (2014). The leucine-rich repeat receptor kinase *BIR2* is a negative regulator of *BAK1* in plant immunity. *Current Biology*, **24**, 134-143.
- Hammond-Kosack, K. E., & Jones, J. D. (1997). Plant disease resistance genes. *Annual Review of Plant Biology*, **48**, 575-607.
- Handley, R., Ekbom, B., & Agren, J. (2005). Variation in trichome density and resistance against a specialist insect herbivore in natural populations of *Arabidopsis thaliana*. *Ecological Entomology*, **30**, 284-292.
- Hanif, M., Mahlia, T., Aditiya, H., & Chong, W. (2016). Techno-economic and environmental assessment of bioethanol production from high starch and root yield Sri Kanji 1 cassava in Malaysia. *Energy Reports*, **2**, 246-253.
- Hardy, N. B., Gullan, P. J., & Hodgson, C. J. (2008). A subfamily- level classification of mealybugs (Hemiptera: Pseudococcidae) based on integrated molecular and morphological data. *Systematic Entomology*, **33**, 51-71.
- Hegde, M., Oliveira, J. N., Da Costa, J. G., Bleicher, E., Santana, A. E., Bruce, T. J., Caulfield, J., Dewhurst, S. Y., Woodcock, C. M., & Pickett, J. A. (2011). Identification of semiochemicals released by cotton, *Gossypium hirsutum*, upon infestation by the cotton aphid, *Aphis gossypii*. *Journal of Chemical Ecology*, **37**, 741-750.
- Henderson, L. (2015). Cassava: an economically viable animal feed?: feeds & grazing. *Stockfarm*, **5**, 32-33.
- Henry, G., & Hershey, C. (2002). Cassava in South America and the Caribbean. In: Hillocks, R.J., Thresh, J.M. and Bellotti, A.C. (eds). Cassava: Biology, Production and Utilization, CABI Publishing, CIAT, Cali, Colombia and University of Greenwich, Kent, UK, pp 17-40.
- Herren, H. R., & Neuenschwander, P. (1991). Biological control of cassava pests in Africa. *Annual Review of Entomology*, **36**, 257-283.
- Hershey, C. H., Álvarez, E., Aye, T. M., Becerra, L. A., Bellotti, A., Ceballos, H., Fahrney, K., Howeler, R., Lefroy, R., & Ospina, B. (2012). Eco-efficient interventions to support cassava's multiple roles in improving the lives of smallholders. In: Hershey, C.H. and Neate, P. (eds), Eco-Efficiency: From Vision to Reality, CIAT, Cali, Colombia, pp 135-160.
- Hilfiker, O., Groux, R., Bruessow, F., Kiefer, K., Zeier, J., & Reymond, P. (2014). Insect eggs induce a systemic acquired resistance in *Arabidopsis*. *Plant Journal*, **80**, 1085-1094.
- Hillocks, R., & Jennings, D. (2003). Cassava brown streak disease: a review of present knowledge and research needs. *International Journal of Pest Management*, **49**, 225-234.
- Hillocks, R. J. (2002). Cassava in Africa. In: Hillocks, R.J., Thresh, J.M. and Bellotti, A.C. (eds). Cassava: Biology, Production and Utilization, CABI Publishing, CIAT, Cali, Colombia and University of Greenwich, Kent, UK, pp 41-54.
- Hillocks, R. J., & Wydra, K. (2002). Bacterial, fungal and nematode diseases. In: Hillocks, R.J., Thresh, J.M. and Bellotti, A.C. Cassava: Biology, Production and Utilization, CABI Publishing, CIAT, Cali, Colombia and University of Greenwich, Kent, UK, pp 261-280.

- Hiremath, P. J., Farmer, A., Cannon, S. B., Woodward, J., Kudapa, H., Tuteja, R., Kumar, A., BhanuPrakash, A., Mulaosmanovic, B., & Gujaria, N. (2011). Large- scale transcriptome analysis in chickpea (*Cicer arietinum* L.), an orphan legume crop of the semi- arid tropics of Asia and Africa. *Plant Biotechnology Journal*, **9**, 922-931.
- Hirsch, C., Hirsch, C. D., Brohammer, A. B., Bowman, M. J., Soifer, I., Barad, O., Shem-Tov, D., Baruch, K., Lu, F., Hernandez, A. G., Fields, C. J., Wright, C. L., Koehler, K., Springer, N. M., Buckler, E. S., Buell, C. R., de Leon, N., Kaeppler, S. M., Childs, K., & Mikel, M. A. (2016). Draft assembly of elite inbred PH207 provides insights into genomic and transcriptome diversity in maize. *Plant Cell*, **28**, 2700-2714.
- Hogenhout, S. A., & Bos, J. I. (2011). Effector proteins that modulate plant–insect interactions. *Current Opinion in Plant Biology*, **14**, 422-428.
- Hou, R., Bao, Z., Wang, S., Su, H., Li, Y., Du, H., Hu, J., Wang, S., & Hu, X. (2011). Transcriptome sequencing and *de novo* analysis for Yesso scallop (*Patinopecten yessoensis*) using 454 GS FLX. *PloS One*, **6**, 21560-21566.
- Howe, G. A., & Jander, G. (2008). Plant immunity to insect herbivores. *Annual Review of Plant Biology*, **59**, 41-66.
- Howeler, R. H. (2006). Cassava in Asia. Workshop on Cassava productuion and technology dissemination in Indonesia and East Timor, CIAT Asia office- East Timor, Department of Agriculture, Dilli, East Timor. Available: http://www.ciat.cgiar.org/asia_cassava/proceedings_workshop.
- Howeler, R.H. (2012). Recent trends in production and utilization of cassava in Asia. In: Howeler, R.H. (eds). The Cassava Handbook, Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia, pp. 1-22.
- Hu, Y. X., Tao, Y. B., & Xu, Z. F. (2017). Overexpression of *Jatropha Gibberellin 2-oxidase6* (*JcGA2ox6*) Induces Dwarfism and Smaller Leaves, Flowers and Fruits in *Arabidopsis* and *Jatropha*. *Frontiers in Plant Science*, **8**, 1-12.
- Huang, C.-G., Hsu, J.-C., Haymer, D. S., Lin, G.-C., & Wu, W.-J. (2009a). Rapid identification of the Mediterranean fruit fly (Diptera: Tephritidae) by loop-mediated isothermal amplification. *Journal of Economic Entomology*, **102**, 1239-1246.
- Huang, S., Li, R., Zhang, Z., Li, L., Gu, X., Fan, W., Lucas, W. J., Wang, X., Xie, B., & Ni, P. (2009b). The genome of the cucumber, *Cucumis sativus* L. *Nature Genetics*, **41**, 1275-1281.
- Huang, X. Z., Chen, J. Y., Xiao, H. J., Xiao, Y. T., Wu, J., Wu, J. X., Zhou, J.-J., Zhang, Y.-J., & Guo, Y. Y. (2015). Dynamic transcriptome analysis and volatile profiling of *Gossypium hirsutum* in response to the cotton bollworm *Helicoverpa armigera*. *Scientific Reports*, **5**, 1-14.
- Huttly, A. K., & Phillips, A. L. (1995). Gibberellin-regulated expression in oat aleurone cells of two kinases that show homology to MAP kinase and a ribosomal protein kinase. *Plant Molecular Biology*, **27**, 1043-1052.
- Jackson, B., Oladipo, N., & Agaja, M. (2014). Cassava: a potential crop for industrial raw material in Nigeria. *International Journal of Life Science*, **3**, 105-112.
- Janssens, M. (2001). Cassava. Crop Production in Tropical Africa. Directorate General for International Co-operation (DGIC), Brussels, pp 165-187.
- Jansson, C., Westerbergh, A., Zhang, J., Hu, X., & Sun, C. (2009). Cassava, a potential biofuel crop in (the) People's Republic of China. *Applied Energy*, **86**, S95-S99.
- Jennings, D. (1963). Variation in pollen and ovule fertility in varieties of cassava, and the effect of interspecific crossing on fertility. *Euphytica*, **12**, 69-76.
- Jennings, D., & Iglesias, C. (2002). Breeding for crop improvement. In: Hillocks, R.J., Thresh, J.M. and Bellotti, A.C. (eds). Cassava: Biology, Production and Utilization, Molecular taxonomy using single- strand conformation polymorphism (SSCP) analysis of mitochondrial ribosomal DNA genes, pp 149-166.
- Jensen, L. J., Kuhn, M., Stark, M., Chaffron, S., Creevey, C., Muller, J., Doerks, T., Julien, P., Roth, A., & Simonovic, M. (2009). STRING 8—a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Research*, **37**, 412-416.
- Johnson, S. M., Lim, F.-L., Finkler, A., Fromm, H., Slabas, A. R., & Knight, M. R. (2014). Transcriptomic analysis of *sorghum bicolor* responding to combined heat and drought stress. *BMC Genomics*, **15**, 456-474.

- Jones, J. D., & Dangl, J. L. (2006). The plant immune system. *Nature*, **444**, 323-329.
- Jones, W. O. (1959). Manioc in Africa. In: P. Stanley King (eds), Stanford University Press, Stanford, California, USA.
- Jorge, V., Fregene, M. A., Duque, M., Bonierbale, M. W., Tohme, J., & Verdier, V. (2000). Genetic mapping of resistance to bacterial blight disease in cassava (*Manihot esculenta* Crantz). *Theoretical and Applied Genetics*, **101**, 865-872.
- Jørgensen, K., Morant, A. V., Morant, M., Jensen, N. B., Olsen, C. E., Kannangara, R., Motawia, M. S., Møller, B. L., & Bak, S. (2011). Biosynthesis of the cyanogenic glucosides linamarin and lotaustralin in cassava: isolation, biochemical characterization, and expression pattern of *CYP71E7*, the oxime-metabolizing cytochrome *P450* enzyme. *Plant Physiology*, **155**, 282-292.
- Kamphuis, L., Gao, L., & Singh, K. (2012). Identification and characterization of resistance of cowpea aphid (*Aphis craccivora* Koch) in *Medicago truncatula*. *BMC Plant Biology*, **12**, 1-12.
- Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y., & Hattori, M. (2004). The KEGG resource for deciphering the genome. *Nucleic Acids Research*, **32**, 277-280.
- Kang, J., Park, J., Choi, H., Burla, B., Kretschmar, T., Lee, Y., & Martinoia, E. (2011). Plant ABC transporters, *American Society for Plant Biologist*, **153**, 1-25.
- Kawai, Y., Ono, E., & Mizutani, M. (2014). Evolution and diversity of the 2-oxoglutarate-dependent dioxygenase superfamily in plants. *The Plant Journal*, **78**, 328-343.
- Kazan, K., & Manners, J. M. (2013). MYC2: the master in action. *Molecular Plant*, **6**, 686-703.
- Kempema, L. A., Cui, X., Holzer, F. M., & Walling, L. L. (2007). *Arabidopsis* transcriptome changes in response to phloem-feeding silverleaf whitefly nymphs. Similarities and distinctions in responses to aphids. *Plant Physiology*, **143**, 849-865.
- Kenyon, L., Anandajayasekeram, P., & Ochieng, C. (2006). A synthesis/lesson-learning study of the research carried out on root and tuber crops commissioned through the DFID RNRRS research programmes between 1995 and 2005. A report submitted to the Crop Protection Programme (CPP) of the UK Department For International Development (DFID), January, 2006.
- Kessler, A., & Baldwin, I. T. (2001). Defensive function of herbivore-induced plant volatile emissions in nature. *Science*, **291**, 2141-2144.
- Kessler, A., & Baldwin, I. T. (2002). Plant responses to insect herbivory: the emerging molecular analysis. *Annual Review of Plant Biology*, **53**, 299-328.
- Kessler, A., Halitschke, R., & Baldwin, I. T. (2004). Silencing the jasmonate cascade: induced plant defenses and insect populations. *Science*, **305**, 665-668.
- Kim, D., Perte, G., Trapnell, C., Pimentel, H., Kelley, R., & Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology*, **14**, 36-48.
- Kim, K.-S., Bellendir, S., Hudson, K. A., Hill, C. B., Hartman, G. L., Hyten, D. L., Hudson, M. E., & Diers, B. W. (2010a). Fine mapping the soybean aphid resistance gene *Rag1* in soybean. *Theoretical and Applied Genetics*, **120**, 1063-1071.
- Kim, K.-S., Hill, C. B., Hartman, G. L., Hyten, D. L., Hudson, M. E., & Diers, B. W. (2010b). Fine mapping of the soybean aphid-resistance gene *Rag2* in soybean PI 200538. *Theoretical and Applied Genetics*, **121**, 599-610.
- Kitaoka, N., Matsubara, T., Sato, M., Takahashi, K., Wakuta, S., Kawaide, H., Matsui, H., Nabeta, & K., Matsuura, H. (2011). *Arabidopsis* *CYP94B3* encodes jasmonyl-L-isoleucine 12-hydroxylase, a key enzyme in the oxidative catabolism of jasmonate. *Plant Cell Physiology*, **52**, 1757-1765.
- Klauser, D., Desurmont, G. A., Glauser, G., Vallat, A., Flury, P., Boller, T., Turlings, T. C., & Bartels, S. (2015). The *Arabidopsis* Pep-PEPR system is induced by herbivore feeding and contributes to JA-mediated plant defence against herbivory. *Journal of Experimental Botany*, **66**, 5327-5336.
- Klingler, J., Creasy, R., Gao, L., Nair, R. M., Calix, A. S., Jacob, H. S., Edwards, O. R., & Singh, K. B. (2005). Aphid resistance in *Medicago truncatula* involves antixenosis and phloem-specific, inducible antibiosis, and maps to a single locus flanked by NBS-LRR resistance gene analogs. *Plant Physiology*, **137**, 1445-1455.
- Klingler, J., Kovalski, I., Silberstein, L., Thompson, G. A., & Perl-Treves, R. (2001). Mapping of cotton-melon aphid resistance in melon. *Journal of the American Society for Horticultural Science*, **126**, 56-63.

- Klingler, J. P., Nair, R. M., Edwards, O. R., Singh, K. B. (2009). A single gene, *AIN*, in *Medicago truncatula* mediates a hypersensitive response to both bluegreen aphid and pea aphid, but confers resistance only to bluegreen aphid. *Journal Experimental Botany*, **60**, 4115-4127.
- Koo A.J., Cooke T.F., & Howe G.A. (2011). Cytochrome P450 CYP94B3 mediates catabolism and inactivation of the plant hormone jasmonoyl-L-isoleucine. *Proceedings of the National Academy of Sciences*, **108**, 9298-9303.
- Korth, K. L. (2003). Profiling the response of plants to herbivorous insects. *Genome Biology*, **4**, 211-212.
- Kozziel, M. G., Beland, G. L., Bowman, C., Carozzi, N. B., Crenshaw, R., Crossland, L., Dawson, J., Desai, N., Hill, M., & Kadwell, S. (1993). Field performance of elite transgenic maize plants expressing an insecticidal protein derived from *Bacillus thuringiensis*. *Nature Biotechnology*, **11**, 194-200.
- Kress, W. J., & Erickson, D. L. (2008). DNA barcodes: genes, genomics, and bioinformatics. *Proceedings of the National Academy of Sciences*, **105**, 2761-2762.
- Krimmel, B. A., & Pearse, I. (2013). Sticky plant traps insects to enhance indirect defence. *Ecology letters*, **16**, 219-224.
- Kumar, S., Kanakachari, M., Gurusamy, D., Kumar, K., Narayanasamy, P., Kethireddy Venkata, P., Solanke, A., Gamanagatti, S., Hiremath, V., & Katageri, I. S. (2015). Genome- wide transcriptomic and proteomic analyses of bollworm- infested developing cotton bolls revealed the genes and pathways involved in the insect pest defence mechanism. *Plant Biotechnology Journal*, **14**, 1438-1455.
- Kuniyuki, H., Gioria, R., Rezende, J. A. M., Willink, C. G. d., Novo, J. P. S., & Yuki, V. A. (2006). Transmission of the Grapevine virus B by the mealybug *Pseudococcus longispinus* Targioni-Tozzetti (Hemiptera: Pseudococcidae) in Brazil. *Summa Phytopathologica*, **32**, 151-155.
- Kuśnierczyk, A., Winge, P. E. R., Jørstad, T. S., Troczyńska, J., Rossiter, J. T., & Bones, A. M. (2008). Towards global understanding of plant defence against aphids: timing and dynamics of early *Arabidopsis* defence responses to cabbage aphid (*Brevicoryne brassicae*) attack. *Plant, Cell and Environment*, **31**, 1097-1115.
- Lai, H. G., Chen, X., Chen, Z., Zhou, Y. Q., Ou, W. J., Li, K. M., Ye, J. Q., & Chen, S. B. (2014). In Situ Pollen Germination and Artificial Pollination Compatibility in Cassava (*Manihot esculenta* Crantz). In: Conference of Applied Mechanics and Materials, **654-653**, 245-251.
- Lamesch, P., Berardini, T. Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D. L., Garcia-Hernandez, M., Karthikeyan, A. S., Lee, C. H., Nelson, W. D., Ploetz, L., Singh, S., Wensel, A., & Huala, E. (2012). The *Arabidopsis* information resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Research*, **40**, 1202-1210.
- Lapidot, M., & Friedmann, M. (2002). Breeding for resistance to whitefly- transmitted geminiviruses. *Annals of Applied Biology*, **140**, 109-127.
- Le Ru, B., & Calatayud, P. A. (1994). Interactions between cassava and arthropod pests. *African Crop Science Journal*, **2**, 385-390.
- Legg, J. (1999). Emergence, spread and strategies for controlling the pandemic of cassava mosaic virus disease in east and central Africa. *Crop Protection*, **18**, 627-637.
- Legg, J., Owor, B., Sseruwagi, P., & Ndunguru, J. (2006). Cassava mosaic virus disease in east and central Africa: epidemiology and management of a regional pandemic. *Advances in Virus Research*, **67**, 355-418.
- Leite, G. L. D., Picanço, M., Jham, G. N., & Gusmão, M. R. (2002). Effects of leaf compounds, climate and natural enemies on the incidence of thrips in cassava. *Pesquisa Agropecuária Brasileira*, **37**, 1657-1662.
- Lenfant, N., Hotelier, T., Bourne, Y., Marchot, P., & Chatonnet, A. (2013). Proteins with an alpha/beta hydrolase fold: relationships between subfamilies in an ever-growing superfamily. *Chemico-biological interactions*, **203**, 266-268.
- Li, H., Mahmood, T., Antony, G., Lu, N., Pumphreys, M., Gill, B., Kang, Z., White, F. F., & Bai, J. (2016). The non-host pathogen *Puccinia tritricina* elicits an active transcriptional response in rice. *European Journal of Plant Pathology*, **147**, 553-569.
- Li, J., Zhu, L., Hull, J. J., Liang, S., Daniell, H., Jin, S., & Zhang, X. (2016). Transcriptome analysis reveals a comprehensive insect resistance response mechanism in cotton to infestation by the phloem feeding insect *Bemisia tabaci* (whitefly). *Plant Biotechnology Journal*, **14**, 1956-1975.

- Li, Y., Zou, J., Li, M., Bilgin, D. D., Vodkin, L. O., Hartman, G. L., & Clough, S. J. (2008). Soybean defense responses to the soybean aphid. *New Phytologist*, **179**, 185-195.
- Liebrand, T. W., van den Berg, G. C., Zhang, Z., Smit, P., Cordewener, J. H., America, A. H., Sklenar, J., Jones, A. M., Tameling, W. I., & Robatzek, S. (2013). Receptor-like kinase *SOBIR1/EVR* interacts with receptor-like proteins in plant immunity against fungal infection. *Proceedings of the National Academy of Sciences*, **110**, 10010-10015.
- Liu, J., Zheng, Q., Ma, Q., Gadidasu, K. K., & Zhang, P. (2011). Cassava Genetic Transformation and its Application in Breeding. *Journal of Integrative Plant Biology*, **53**, 552-569.
- Liu, F., Jiang, H., Ye, S., Chen, W. P., Liang, W., Xu, Y., Sun, B., Sun, J., Wang, Q., & Li, C. (2010). The *Arabidopsis* P450 protein CYP82C2 modulates jasmonate-induced root growth inhibition, defense gene expression and indole glucosinolate biosynthesis. *Cell Research*, **20**, 539-552.
- Liu, X. M., Smith, C. M. & Gill, B. S. (2005). Allelic relationships among Russian wheat aphid resistance genes. *Crop Science*, **45**, 2273-2280.
- Livak, K. J., & Schmittgen, T. D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻(DeltaDeltaCT) method. *Methods*, **25**, 402-408.
- Lo, S. F., Ho, T. H. D., Liu, Y. L., Jiang, M. J., Hsieh, K. T., Chen, K. T., Yu, L. C., Lee, M. H., Chen, C. y., & Huang, T. P. (2016). Ectopic expression of specific GA-2 oxidase mutants promotes yield and stress tolerance in rice. *Plant Biotechnology Journal*, **15**, 850-864.
- Lopes-Caitar, V. S., de Carvalho, M. C., Darben, L. M., Kuwahara, M. K., Nepomuceno, A. L., Dias, W. P., Abdelnoor, R. V., & Marcelino-Guimaraes, F. C. (2013). Genome-wide analysis of the Hsp20 gene family in soybean: comprehensive sequence, genomic organization and expression profile analysis under abiotic and biotic stresses. *BMC Genomics*, **14**, 1-17.
- Lopez, C., Piegu, B., Cooke, R., Delseny, M., Tohme, J., & Verdier, V. (2005a). Using cDNA and genomic sequences as tools to develop SNP strategies in cassava (*Manihot esculenta* Crantz). *Theoretical and Applied Genetics*, **110**, 425-431.
- Lopez, C., Soto, M., Restrepo, S., Piégu, B., Cooke, R., Delseny, M., Tohme, J., & Verdier, V. (2005b). Gene expression profile in response to *Xanthomonas axonopodis* pv. *manihotis* infection in cassava using a cDNA microarray. *Plant Molecular Biology*, **57**, 393-410.
- López, C. E., & Bernal, A. J. (2012). Cassava bacterial blight: using genomics for the elucidation and management of an old problem. *Tropical Plant Biology*, **5**, 117-126.
- Lorenzo, O., Chico, J. M., Sánchez-Serrano, J. J., & Solano, R. (2004). *Jasmonate-Insensitive1* encodes a MYC transcription factor essential to discriminate between different jasmonate-regulated defense responses in *Arabidopsis*. *The Plant Cell*, **16**, 1938-1950.
- Lorenzo, O., & Solano, R. (2005). Molecular players regulating the jasmonate signalling network. *Current Opinion in Plant Biology*, **8**, 532-540.
- Louis, B., & Rey, C. (2015). Resistance gene analogs involved in tolerant cassava–geminivirus interaction that shows a recovery phenotype. *Virus Genes*, **51**, 393-407.
- Lu, Y., Wu, K., Jiang, Y., Guo, Y., & Desneux, N. (2012). Widespread adoption of *Bt* cotton and insecticide decrease promotes biocontrol services. *Nature*, **487**, 362-365.
- Luna, E., Bruce, T. J. A., Roberts, M. R., Flors, V., & Ton, J. (2012). Next-generation systemic acquired resistance. *Plant Physiology*, **158**, 844-853.
- Madeley, J. (2011). Cassava beer. *Appropriate Technology*, **38**, 16-17.
- Maffei, M. E., Mithofer, A., & Boland, W. (2007). Before gene expression: early events in plant- insect interaction. *Trends in Plant Science*, **12**, 310-316.
- Mahungu, N., Dixon, A. G., & Kumbira, J. (2009). Breeding cassava for multiple pest resistance in Africa. *African Crop Science Journal*, **2**, 539-552.
- Makwarela, M., & Rey, C. (2006). Cassava biotechnology, a southern African perspective. *Biotechnology and Molecular Biology Reviews*, **1**, 2-11.
- Malaus, T., Fenis, A., Warot, S., Germain, J. F., Ris, N., Prado, E., Botton, M., Vanlerberghe- Masutti, F., Sforza, R., & Cruaud, C. (2011). DNA markers to disentangle complexes of cryptic taxa in mealybugs (Hemiptera: Pseudococcidae). *Journal of Applied Entomology*, **135**, 142-155.
- Mardis, E. R. (2008a). The impact of next-generation sequencing technology on genetics. *Trends in Genetics*, **24**, 133-141.

- Mardis, E. R. (2008b). Next-generation DNA sequencing methods. *Annual Review of Genomics and Human Genetics*, **9**, 387-402.
- Marone, D., Russo, M. A., Laidò, G., De Leonardis, A. M., & Mastrangelo, A. M. (2013). Plant nucleotide binding site-leucine-rich repeat (NBS-LRR) genes: active guardians in host defense responses. *International Journal of Molecular Sciences*, **14**, 7302-7326.
- Maruthi, M., Hillocks, R., Mtunda, K., Raya, M., Muhanna, M., Kiozia, H., Rekha, A., Colvin, J., & Thresh, J. (2005). Transmission of Cassava brown streak virus by *Bemisia tabaci* (Gennadius). *Journal of Phytopathology*, **153**(5), 307-312.
- Marx, S., & Nquma, T. Y. (2013). Cassava as feedstock for ethanol production in South Africa. *African Journal of Biotechnology*, **12**, 4975-4983.
- Mbanzibwa, D., Tian, Y., Tugume, A., Mukasa, S., Tairo, F., Kyamanywa, S., Kullaya, A., & Valkonen, J. P. (2009). Genetically distinct strains of Cassava brown streak virus in the Lake Victoria basin and the Indian Ocean coastal area of east Africa. *Archives of Virology*, **154**, 353-359.
- Mbanzibwa, D., Tian, Y., Tugume, A., Patil, B., Yadav, J., Bagewadi, B., Abarshi, M., Alicai, T., Changadeya, W., & Mkumbira, J. (2011). Evolution of cassava brown streak disease-associated viruses. *Journal of General Virology*, **92**, 974-987.
- McDowell, J. M., & Woffenden, B. J. (2003). Plant disease resistance genes: recent insights and potential applications. *Trends in Biotechnology*, **21**, 178-183.
- McManus, M. T. (2012). The plant hormone ethylene. *Annual Plant Reviews*, **44**, 1-17.
- Mello, M. O., & Silva-Filho, M. C. (2002). Plant-insect interactions: an evolutionary arms race between two distinct defense mechanisms. *Brazilian Journal of Plant Physiology*, **14**, 71-81.
- Metzker, M. L. (2010). Sequencing technologies-the next generation. *Nature Reviews Genetics*, **11**, 31-46.
- Meyer, J., Kasdorf, G., Nel, L., & Pietersen, G. (2008). Transmission of activated-episomal banana streak virus (bads) to cv. Williams banana (*Musa* sp.) by three mealybug species. *Plant Disease*, **92**, 1158-1163.
- Meyers, B. C., Galbraith, D. W., Nelson, T., & Agrawal, V. (2004). Methods for transcriptional profiling in plants. Be fruitful and replicate. *Plant Physiology*, **135**, 637-652.
- Mian, M., Hammond, R. B., & St Martin, S. K. (2008). New plant introductions with resistance to the soybean aphid. *Crop Science*, **48**, 1055-1061.
- Michelmore, R. W., & Meyers, B. C. (1998). Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Research*, **8**, 1113-1130.
- Millar, I. (2002). Mealybug genera (Hemiptera: Pseudococcidae) of South Africa: identification and review. *African Entomology*, **10**, 185-233.
- Millar, J.G., Daane, K.M., McElfresh, J.S., Moreira, J.A., Malakar-Kuenen, R., Guille, M., & Bently, W.J. (2002). Development and optimization methods for using sex pheromone for monitoring the mealybug *Planococcus ficus* (Homoptera: Pseudococcidae) in California vineyards. *Journal of Economic Entomology*, **4**, 706-714.
- Miller, D. R., Miller, G. L., Hodges, G. S., & Davidson, J. A. (2005). Introduced scale insects (Hemiptera: Coccoidea) of the United States and their impact on US agriculture. *Proceedings of the Entomological Society of Washington*, **107**, 123-158.
- Miller, D. R., Miller, G. L., & Watson, G. W. (2002). Invasive species of mealybugs (Hemiptera: Pseudococcidae) and their threat to US agriculture. *Proceedings-Entomological Society of Washington*, **104**, 825-836.
- Miller, D. R., & Rossman, A. Y. (1995). Systematics, biodiversity, and agriculture. *BioScience*, **45**, 680-686.
- Mizutani, M., & Ohta, D. (2010). Diversification of P450 genes during land plant evolution. *Annual Review of Plant Biology*, **61**, 291-315.
- Morales, F. J. (2001). Conventional breeding for resistance to *Bemisia tabaci*-transmitted geminiviruses. *Crop Protection*, **20**, 825-834.
- Moran, P. J., & Thompson, G. A. (2001). Molecular responses to aphid feeding in *Arabidopsis* in relation to plant defense pathways. *Plant Physiology*, **125**, 1074-1085.
- Morante, N., Marín, J. A., Gutiérrez A, J. P., Sánchez, T., Ospina, C., Bellotti, A. C., & Fregene, M. A. (2004). Mining the primary gene pool of cassava: Introgression of resistance to the cassava green mite and high root protein from accessions of *Manihot esculenta* sub spp *Fabellifolia* and *Manihot*

- tristis* into cassava. [Online], CIAT, Cali, Colombia, Available: http://ciat-library.ciat.cgiar.org/Articulos_ciat/cbn/Posters-PDF/PS-2/N_Morante.pdf.
- Moreno, I., Gruissem, W., & Vanderschuren, H. (2011). Reference genes for reliable potyvirus quantitation in cassava and analysis of Cassava brown streak virus load in host varieties. *Journal of Virological Methods*, **177**, 49-54.
- Morgan, N. K., & Choct, M. (2016). Cassava: Nutrient composition and nutritive value in poultry diets. *Animal Nutrition*, **2**, 253-261.
- Moritz, C., & Cicero, C. (2004). DNA barcoding: promise and pitfalls. *PLoS Biology*, **2**, 1529-1531.
- Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., & Kanehisa, M. (2007). KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research*, **35**, 182-185.
- Morozova, O., Hirst, M., & Marra, M. A. (2009). Applications of new sequencing technologies for transcriptome analysis. *Annual Review of Genomics and Human Genetics*, **10**, 135-151.
- Morozova, O., & Marra, M. A. (2008). Applications of next-generation sequencing technologies in functional genomics. *Genomics*, **92**, 255-264.
- Müller, M., & Munné-Bosch, S. (2015). Ethylene response factors: a key regulatory hub in hormone and stress signaling. *Plant Physiology*, **169**, 32-41.
- Murashige, T., & Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiologia Plantarum*, **15**, 473-497.
- Mutisya, D. L., El-Banhawy, E., Khamala, C., & Kariuki, C. W. (2015a). Management of cassava green mite *Mononychellus progresivus* (Acari: Tetranychidae) in different agro-ecological zones of Kenya. *Systematic and Applied Acarology*, **20**, 39-50.
- Mutisya, D. L., Wambua, J. M., Miano, D. W., & Kariuki, C. W. (2015b). Farmer perceptions of cassava green mite pest impact in eastern Kenya. *Journal of Entomology and Zoology Studies*, **3**:354-358
- Nadel, H., & Alphen, J. v. (1987). The role of host and host- plant odours in the attraction of a parasitoid, *Epidinocarsis lopezi*, to the habitat of its host, the cassava mealybug, *Phenacoccus manihoti*. *Entomologia Experimentalis et Applicata*, **45**, 181-186.
- Nagrare, V., Kranthi, S., Biradar, V., Zade, N., Sangode, V., Kakde, G., Shukla, R., Shivare, D., Khadi, B., & Kranthi, K. (2009). Widespread infestation of the exotic mealybug species, *Phenacoccus solenopsis* (Tinsley)(Hemiptera: Pseudococcidae), on cotton in India. *Bulletin of Entomological Research*, **99**, 537-541.
- Nakaune, R., Toda, S., Mochizuki, M., & Nakano, M. (2008). Identification and characterization of a new vitivirus from grapevine. *Archives of Virology*, **153**, 1827-1832.
- Narina, S. S., & Odeny, D. (2012). Cassava. In: Kole, C., Joshi, C.P., and Shonnard, D.R. (eds). Handbook of Bioenergy Crop Plants, CRC Press (Taylor and Francis Group), New York, USA, pp 303-314.
- Nassar, N. (2000). Wild cassava, *Manihot* spp.: biology and potentialities for genetic improvement. *Genetics and Molecular Biology*, **23**, 201-212.
- Nassar, N., & Ortiz, R. (2007). Cassava improvement: challenges and impacts. *The Journal of Agricultural Science*, **145**, 163-171.
- Nassar, N., & Ortiz, R. (2010). Breeding cassava to feed the poor. *Scientific American*, **302**, 78-84.
- Nassar, N. M. (2003). Cassava, *Manihot esculenta* Crantz genetic resources: VI. Anatomy of a diversity center. *Genetics and Molecular Research*, **2**, 214-222.
- Nassar, N. M., Hashimoto, D., & Fernandes, S. (2008). Wild *Manihot* species: botanical aspects, geographic distribution and economic value. *Genetics and Molecular Research*, **7**, 16-28.
- Neuenschwander, P. (2001). Biological control of the cassava mealybug in Africa: a review. *Biological Control*, **21**, 214-229.
- Nguyen, T. L. T., Gheewala, S. H., & Garivait, S. (2007). Full chain energy analysis of fuel ethanol from cassava in Thailand. *Environmental Science & Technology*, **41**, 4135-4142.
- Ni, X., Quisenberry, S.S., Heng-Moss, S., Markwell, J.P., & Sarath, G. (2001). Oxidative responses of resistant and susceptible cereal leaves to symptomatic and nonsymptomatic cereal aphid (Hemiptera: Aphididae) feeding. *Journal of Economic Entomology*, **94**, 743-751
- Nichols, R. (1947). Breeding cassava for virus resistance. *East African Agricultural Journal*, **12**, 184-194.
- Nkaa, F., Ene-Obong, E., Taylor, N., Fauquet, C., & Mbanaso, E. (2013). Elimination of African cassava mosaic virus (ACMV) and East African cassava mosaic virus (EACMV) from cassava (*Manihot*

- esculenta Crantz) cv. 'Nwugo' via somatic embryogenesis. *American Journal of Biomedical Sciences*, **3**, 33-40.
- Nkaa, F., Ene-Obong, E., Taylor, N., Fauquet, C., & Mbanaso, E. (2014). Elimination of East African Cassava Mosaic Virus (EACMV) from cassava (*Manihot esculenta* Crantz) through somatic embryogenesis. *Journal of Sustainable Agriculture and the Environment*, **15**, 58-68.
- Nweke, F., Otim-Nape, G., Dixon, A., Asadu, C., Bua, A., Ajobo, O., Ugwu, B., Baguma, Y., & Masembe-Kajubi, H. (1999). Production prospects for cassava in Uganda. Collaborative Study of Cassava in Africa (COSCA): Working paper No. 17, International Institute of Tropical Agriculture, Ibadan, Nigeria.
- Nweke, F., Spencer, D., & Lynam, J. (2002). The cassava transformation: Methods of COSCA study. Africa's best kept secret. Michigan State University Press, East Lansing, Michigan, pp 203-214.
- Nyaboga, E. N., Njiru, J. M., & Tripathi, L. (2015). Factors influencing somatic embryogenesis, regeneration, and *Agrobacterium*-mediated transformation of cassava (*Manihot esculenta* Crantz) cultivar TME14. *Frontiers in Plant Science*, **6**, 1-13.
- Nyaboga, E., Njiru, J., Nguu, E., Gruissem, W., Vanderschuren, H. & Tripathi, L. (2013). Unlocking potential of tropical root crop biotechnology in east Africa by establishing a genetic transformation platform for local farmer-preferred cassava cultivars. *Frontiers in Plant Science*, **4**, 1-11.
- Olsen, K. M. (2004). SNPs, SSRs and inferences on cassava's origin. *Plant Molecular Biology*, **56**, 517-526.
- Olsen, K. M., & Schaal, B. A. (1999). Evidence on the origin of cassava: phylogeography of *Manihot esculenta*. *Proceedings of the National Academy of Sciences*, **96**, 5586-5591.
- Onwueme, I. (2002). Cassava in Asia and the Pacific. In: Hillocks, R.J., Thresh, J.M. and Bellotti, A.C. (eds). Cassava: Biology, Production and Utilization, Oxon, UK and New York, USA pp. 55-65
- Otim-Nape, G. W., Bua, A., & Baguma, Y. (1994). Accelerating the transfer of improved production technologies: Controlling African cassava mosaic virus disease epidemics in Uganda. *African Crop Science Journal*, **2**, 479-495.
- Ouaked, F., Rozhon, W., Lecourieux, D., & Hirt, H. (2003). A MAPK pathway mediates ethylene signaling in plants. *The EMBO Journal*, **22**, 1282-1288.
- Owiti, J., Grossmann, J., Gehrig, P., Dessimoz, C., Laloi, C., Hansen, M. B., Gruissem, W., & Vanderschuren, H. (2011). iTRAQ- based analysis of changes in the cassava root proteome reveals pathways associated with post- harvest physiological deterioration. *The Plant Journal*, **67**, 145-156.
- Pan, Q., Wendel, J., & Fluhr, R. (2000). Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes. *Journal of Molecular Evolution*, **50**, 203-213.
- Park, D.-S., Leem, Y. J., Hahn, K.-W., Suh, S.-J., Hong, K.-J., & Oh, H.-W. (2010a). Molecular identification of mealybugs (Hemiptera: Pseudococcidae) found on Korean pears. *Journal of Economic Entomology*, **103**, 25-33.
- Park, D.-S., Suh, S.-J., Oh, H.-W., & Hebert, P. D. (2010b). Recovery of the mitochondrial COI barcode region in diverse *Hexapoda* through tRNA-based primers. *BMC Genomics*, **11**, 1-7.
- Park, S.-J., Huang, Y., & Ayoubi, P. (2006). Identification of expression profiles of sorghum genes in response to greenbug phloem-feeding using cDNA subtraction and microarray analysis. *Planta*, **223**, 932-947.
- Park, Y.-S., & Ryu, C.-M. (2014). Understanding cross-communication between aboveground and belowground tissues via transcriptome analysis of a sucking insect whitefly-infested pepper plants. *Biochemical and Biophysical Research Communications*, **443**, 272-277.
- Parkes, E. Y. (2011). Assessment of genetic diversity, combining ability, stability and farmer preference of cassava germplasm in Ghana. A thesis submitted for a degree of Doctor of Philosophy in Plant Breeding, University of Free State, Bloemfontein, South Africa.
- Parsa, S., Kondo, T., & Winotai, A. (2012). The cassava mealybug (*Phenacoccus manihoti*) in Asia: First records, potential distribution, and an identification key. *PLoS One*, **7**, 47675-47685.
- Patil, B. L., Ogowok, E., Wagaba, H., Mohammed, I. U., Yadav, J. S., Bagewadi, B., Taylor, N. J., Kreuze, J. F., Maruthi, M., & Alicai, T. (2011). RNAi- mediated resistance to diverse isolates belonging to two virus species involved in Cassava brown streak disease. *Molecular Plant Pathology*, **12**, 31-41.

- Petersen M, Brodersen P, Naested H, Andreasson E, Lindhart U, Johansen, B., Nielsen, H. B., Lacy, M., Austin, J., Parker, J. E., Sharma, S. B., Klessig, F., Martienssen, R., Mattsson, O., Jensen, A. B., & Mund, J. (2000). *Arabidopsis* map kinase 4 negatively regulates systemic acquired resistance. *Cell*, **103**, 1111-1120.
- Petit-Houdenot, Y., & Fudal, I. (2017). Complex interactions between fungal avirulence genes and their corresponding plant resistance genes and consequences for disease resistance management. *Frontiers in Plant Science*, **8**, 1-8.
- Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, **29**, 2003-2007.
- Pfaffl, M. W., Tichopad, A., Prgomet, C., & Neuvians, T. P. (2004). Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper–Excel-based tool using pair-wise correlations. *Biotechnology Letters*, **26**, 509-515.
- Phu, T. V., Thao, N. T., Son, C. K., Nga, L. H., Duc, N. T. H., Ben, B., Diego, N., Keith, T., & Anh, T. K. (2016). Potential use of by-products for animal feed from cassava in Vietnam. *Journal of Science*, **14**, 93-100.
- Pieterse, C. M., Leon Reyes, H., Does, D., Verhage, A., Koornneef, A., van Pelt, J. A., & van Wees, S. C. (2012). Networking by small-molecule hormones in plant immunity. *IOBC-WPRS Bulletin*, **83**, 77-80.
- Pieterse, C. M., Leon-Reyes, A., Van der Ent, S., & Van Wees, S. C. (2009). Networking by small-molecule hormones in plant immunity. *Nature Chemical Biology*, **5**, 308-316.
- Pingmuanglek, P., Jakrawatana, N., & Gheewala, S. H. (2017). Supply chain analysis for cassava starch production: Cleaner production opportunities and benefits. *Journal of Cleaner Production*, **162**, 1075-1084.
- Pitzschke, A., Schikora, A., & Hirt, H. (2009). MAPK cascade signalling networks in plant defence. *Current Opinion in Plant Biology*, **12**, 421-426.
- Poonsrisawat, A., Paemanee, A., Wanlapatit, S., Piyachomkwan, K., Eurwilaichitr, L., & Champreda, V. (2017). Simultaneous saccharification and viscosity reduction of cassava pulp using a multi-component starch-and cell-wall degrading enzyme for bioethanol production. *Biotech*, **7**, 290. <https://doi.org/10.1007/s13205-017-0924-1>.
- Prochaska, T. J., Donze-Reiner, T., Marchi-Werle, L., Palmer, N. A., Hunt, T. E., Sarath, G., & Heng-Moss, T. Transcriptional responses of tolerant and susceptible soybeans to soybean aphid (*Aphis glycines* Matsumura) herbivory. *Arthropod Plant Interactions Journal*, **9**, 347-359.
- Prochnik, S., Marri, P. R., Desany, B., Rabinowicz, P. D., Kodira, C., Mohiuddin, M., Rodriguez, F., Fauquet, C., Tohme, J., & Harkins, T. (2012). The cassava genome: current progress, future directions. *Tropical Plant Biology*, **5**, 88-94.
- Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., & Yamada, T. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*, **464**, 59-65.
- Rabbi, I., Hamblin, M., Gedil, M., Kulakow, P., Ferguson, M., Ikpan, A. S., Ly, D., & Jannink, J. (2014). Genetic mapping using genotyping-by-sequencing in the clonally-propagated cassava. *Crop Science*, **54**, 1384-1396.
- Rach, J., DeSalle, R., Sarkar, I. N., Schierwater, B., & Hadrys, H. (2008). Character-based DNA barcoding allows discrimination of genera, species and populations in Odonata. *Proceedings of the Royal Society of London B: Biological Sciences*, **275**, 237-247.
- Raji, A., Fawole, I., Gedil, M., & Dixon, A. (2009). Genetic differentiation analysis of African cassava (*Manihot esculenta*) landraces and elite germplasm using amplified fragment length polymorphism and simple sequence repeat markers. *Annals of Applied Biology*, **155**, 187-199.
- Rajniak, J., Barco, B., Clay, N. K., & Sattely, E. S. (2015). A new cyanogenic metabolite in *Arabidopsis* required for inducible pathogen defence. *Nature*, **525**, 376-379.
- Ratnasingham, S., & Hebert, P. D. (2007). BOLD: The Barcode of Life Data System (<http://www.barcodinglife.org>). *Molecular Ecology Notes*, **7**, 355-364.
- Ravindran, V. (1993). Cassava leaves as animal feed: potential and limitations. *Journal of the Science of Food and Agriculture*, **61**(2), 141-150.

- Rawat, N., Naga, N. C., Meenakshi, S. R., Nair, S., & Bentur, J. S. (2012). A novel mechanism of gall midge resistance in the rice variety Kavya revealed by microarray analysis. *Functional & Integrative Genomics*, **12**, 249-264.
- Ren, J. (2000). High-throughput single-strand conformation polymorphism analysis by capillary electrophoresis. *Journal of Chromatography B: Biomedical Sciences and Applications*, **741**, 115-128.
- Rensink, W. A., & Buell, C. R. (2005). Microarray expression profiling resources for plant genomics. *Trends in Plant Science*, **10**, 603-609.
- Reuter, J. A., Spacek, D. V., & Snyder, M. P. (2015). High-throughput sequencing technologies. *Molecular Cell*, **58**, 586-597.
- Reymond, P., Weber, H., Damond, M., & Farmer, E. E. (2000). Differential gene expression in response to mechanical wounding and insect feeding in *Arabidopsis*. *The Plant Cell*, **12**, 707-719.
- Riis, L., Bellotti, A. C., Bonierbale, M., & O'Brien, G. M. (2003). Cyanogenic potential in cassava and its influence on a generalist insect herbivore *Cyrtomenus bergi* (Hemiptera: Cydnidae). *Journal of Economic Entomology*, **96**, 1905-1914.
- Robertson, G., Schein, J., Chiu, R., Corbett, R., Field, M., Jackman, S. D., Mungall, K., Lee, S., Okada, H. M., & Qian, J. Q. (2010). *De novo* assembly and analysis of RNA-seq data. *Nature Methods*, **7**, 909-912.
- Rogers, D. J., & Appan, S. G. (1973). *Manihot* and *Manihotoides* (Euphorbiaceae): a computer assisted study. Flora Neotropica, Monograph 13, Hafner Press, New York, USA.
- Rosas-García, N. M., Sarmiento-Benavides, S. L., Villegas-Mendoza, J. M., Hernández-Delgado, S., & Mayek-Pérez, N. (2010). Genetic differentiation among *Maconellicoccus hirsutus* (Hemiptera: Pseudococcidae) populations living on different host plants. *Environmental Entomology*, **39**, 1043-1050.
- Rossi, M., Goggin, F. L., Milligan, S. B., Kaloshian, I., Ullman, D. E., & Williamson, V. M. (1998). The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. *Proceedings of the National Academy of Sciences*, **95**, 9750-9754.
- Rothberg, J. M., & Leamon, J. H. (2008). The development and impact of 454 sequencing. *Nature Biotechnology*, **26**, 1117-1124.
- Roy, S. (2016). Function of myb domain transcription factors in abiotic stress and epigenetic control of stress response in plant genome. *Plant signaling & Behavior*, **11**, 1-7.
- Rung, A., Scheffer, S., Evans, G., & Miller, D. (2008). Molecular identification of two closely related species of mealybugs of the genus *Planococcus* (Homoptera: Pseudococcidae). *Annals of the Entomological Society of America*, **101**, 525-532.
- Sánchez, A. S., Silva, Y. L., Kalid, R. A., Cohim, E., & Torres, E. A. (2017). Waste bio-refineries for the cassava starch industry: New trends and review of alternatives. *Renewable and Sustainable Energy Reviews*, **73**, 1265-1275.
- Sanchez, O. J., & Cardona, C. A. (2008). Trends in biotechnological production of fuel ethanol from different feedstocks. *Bioresource Technology*, **99**, 5270-5295.
- Santana, Q. C., Coetzee, M. P., Steenkamp, E. T., Mlonyeni, O. X., Hammond, G. N., Wingfield, M. J., & Wingfield, B. D. (2009). Microsatellite discovery by deep sequencing of enriched genomic libraries. *Biotechniques*, **46**, 217-223.
- Sayre, R., Beeching, J. R., Cahoon, E. B., Egesi, C., Fauquet, C., Fellman, J., Fregene, M., Grissem, W., Mallowa, S., & Manary, M. (2011). The BioCassava plus program: biofortification of cassava for sub-Saharan Africa. *Annual Review of Plant Biology*, **62**, 251-272.
- Schadt, E. E., Turner, S. & Kasarskis, A. (2010). A window into third-generation sequencing. *Human Molecular Genetics*, **19**, 227-240.
- Schena, M., Shalon, D., Davis, R. W., & Brown, P. O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, **270**, 467-470.
- Schneider, G.F. & Dekker, C. (2012). DNA sequencing with nanopores. *News and Views*, **30**, 326-328.
- Schweizer, F., Fernández-Calvo, P., Zander, M., Diez-Diaz, M., Fonseca, S., Glauser, G., Lewsey, M. G., Ecker, J. R., Solano, R., & Reymond, P. (2013). *Arabidopsis* basic helix-loop-helix transcription factors MYC2, MYC3, and MYC4 regulate glucosinolate biosynthesis, insect performance, and feeding behavior. *The Plant Cell*, **25**, 3117-3132.

- Sethusa, M., Millar, I., Yessoufou, K., Jacobs, A., Van der Bank, M., & Van der Bank, H. (2014). DNA barcode efficacy for the identification of economically important scale insects (Hemiptera: Coccoidea) in South Africa. *African Entomology*, **22**, 257-266.
- Seo, S., Sano, H., & Ohashi, Y. (1999). Jasmonate-based wound signal transduction requires activation of WIPK, a tobacco mitogen-activated protein kinase. *Plant Cell*, **11**, 289-298.
- Sforza, R., Boudon-Padieu, E., & Greif, C. (2003). New mealybug species vectoring Grapevine leafroll-associated viruses-1 and-3 (GLRaV-1 and-3). *European Journal of Plant Pathology*, **109**, 975-981.
- Sharma, K. K., & Ortiz, R. (2000). Program for the application of genetic transformation for crop improvement in the semi-arid tropics. *In Vitro Cellular & Developmental Biology-Plant*, **36**, 83-92.
- Sharpe, R. G., Hims, M. M., Harbach, R. E., & Butlin, R. K., 1999. PCR-based methods for identification of species of the *Anopheles minimus* group: allele-specific amplification and single-strand conformation polymorphism. *Medical and Veterinary Entomology*, **13**, 265-273.
- Shearer, A. E., Hildebrand, M. S., Sloan, C. M., & Smith, R. J. (2011). Deafness in the genomics era. *Hearing research*, **282**, 1-9.
- Silalertruksa, T., & Gheewala, S. H. (2010). Security of feedstocks supply for future bio-ethanol production in Thailand. *Energy Policy*, **38**, 7476-7486.
- Song, S., Qi, T., Wasternack, C., & Xie, D. (2014). Jasmonate signaling and crosstalk with gibberellin and ethylene. *Current Opinion in Plant Biology*, **21**, 112-119.
- Soto, J. C., Ortiz, J. F., Perlaza-Jiménez, L., Vásquez, A. X., Lopez-Lavalle, L. A. B., Mathew, B., León, J., Bernal, A. J., Ballvora, A., & López, C.E (2015). A genetic map of cassava (*Manihot esculenta* Crantz) with integrated physical mapping of immunity-related genes. *BMC Genomics*, **16**, 1-16.
- Sperling, F., Raske, A., & Otvos, I. (1999). Mitochondrial DNA sequence variation among populations and host races of *Lambdina fiscellaria* (Gn.) (Lepidoptera: Geometridae). *Insect Molecular Biology*, **8**, 97-106.
- Spoel, S. H., Johnson, J. S., & Dong, X. (2007). Regulation of tradeoffs between plant defenses against pathogens with different lifestyles. *Proceedings of the National Academy of Sciences*, **104**, 18842-18847.
- Staden, R., Judge, D. P., & Bonfield, J. K. (1998). Sequence assembly and finishing methods. In: Andreas D. Baxevanis, and B.F. Francis Ouellette (eds). *Bioinformatics: A practical guide to the analysis of genes and proteins*, New York, USA, pp 305-322
- Stam, J. M., Kroes, A., Li, Y., Gols, R., van Loon, J. J., Poelman, E. H., & Dicke, M. (2014). Plant interactions with multiple insect herbivores: from community to genes. *Plant Biology*, **65**, 689-713.
- Staswick, P. E., & Tiryaki, I. (2004). The oxylipin signal jasmonic acid is activated by an enzyme that conjugates it to isoleucine in *Arabidopsis*. *The Plant Cell*, **16**, 2117-2127.
- Steppuhn, A., & Baldwin, I. T. (2007). Resistance management in a native plant: nicotine prevents herbivores from compensating for plant protease inhibitors. *Ecology Letters*, **10**, 499-511.
- Stewart Jr, C. N., Adang, M. J., All, J. N., Boerma, H. R., Cardineau, G., Tucker, D., & Parrott, W. A. (1996). Genetic transformation, recovery, and characterization of fertile soybean transgenic for a synthetic *Bacillus thuringiensis cryIac* gene. *Plant Physiology*, **112**, 121-129.
- Studham, M. E., & MacIntosh, G. C. (2013). Multiple phytohormone signals control the transcriptional response to soybean aphid infestation in susceptible and resistant soybean plants. *Molecular Plant-Microbe Interactions*, **26**, 116-129.
- Swadener, C. (1994). *Bacillus thuringiensis* (BT). *Journal of Pesticide Reform*, **14**, 13-20.
- Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguez, P., Doerks, T., Stark, M., Muller, J., & Bork, P. (2011). The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Research*, **39**, 561-568.
- Taiwo, K. A. (2006). Utilization potentials of cassava in Nigeria: the domestic and industrial products. *Food Reviews International*, **22**, 29-42.
- Tang, H., Krishnakumar, V., Bidwell, S., Rosen, B., Chan, A., Zhou, S., Gentzbittel, L., Childs, K. L., Yandell, M., Gundlach, H., Mayer, K. F., Schwartz, D. C., & Town, C. D. (2014). An improved genome release (version MT4.0) for the model legume *Medicago Truncatula*. *BMC Genomics*, **15**, 1-14

- Tao, Y., Xie, Z., Chen, W., Glazebrook, J., Chang, H.-S., Han, B., Zhu, T., Zou, G., & Katagiri, F. (2003). Quantitative nature of *Arabidopsis* responses during compatible and incompatible interactions with the bacterial pathogen *Pseudomonas syringae*. *The Plant Cell*, **15**, 317-330.
- Tautz, D., Hancock, J. M., Webb, D. A., Tautz, C., & Dover, G. A. (1988). Complete sequences of the rRNA genes of *Drosophila melanogaster*. *Molecular Biology and Evolution*, **5**, 366-376.
- Taylor, N., Chavarriaga, P., Raemakers, K., Siritunga, D., & Zhang, P. (2004). Development and application of transgenic technologies in cassava. *Plant Molecular Biology*, **56**, 671-688.
- Tertaliano, M., S., Dossou-Gbete, S., & Le Rü, B. (1993). Antixenosis and antibiotic components of resistance to cassava mealybug *Phenacoccus manihoti* (Hom. Pseudococcidae) in various plants. *Insect Science and Its Application*. **5-6**, 657-665.
- The Tomato Genome Sequencing Consortium. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*, **485**, 635-641.
- Thresh, J., & Cooter, R. (2005). Strategies for controlling cassava mosaic virus disease in Africa. *Plant Pathology*, **54**, 587-614.
- Timm, A. E., & Reineke, A. (2014). First insights into grapevine transcriptional responses as a result of vine mealybug *Planococcus ficus* feeding. *Arthropod-Plant Interactions*, **8**, 495-505.
- Tomkins, J., Fregene, M., Main, D., Kim, H., Wing, R., & Tohme, J. (2004). Bacterial artificial chromosome (BAC) library resource for positional cloning of pest and disease resistance genes in cassava (*Manihot esculenta* Crantz). *Plant Molecular Biology*, **56**, 555-561.
- Tonukari, N. J. (2004). Cassava and the future of starch. *Electronic Journal of Biotechnology*, **7**, 5-8.
- Torii, K. U. (2004). Leucine-rich repeat receptor kinases in plants: structure, function, and signal transduction pathways. *International Review of Cytology*, **234**, 1-46.
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., Pimentel, H., Salzberg, S. L., Rinn, J. L., & Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols*, **7**, 562-578.
- Trung, N. T., Berg, J., Cuong, V. C., & Kjos, N. P. (2014). Influence of varying levels of supplemental cassava root meal without or with groundnut cake on performance of growing Laisind cattle. *Tropical Animal Health and Production*, **46**, 925-930.
- Tuskan, G. A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., & Salamov, A. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*, **313**, 1596-1604.
- Upadhyay, P., Rai, A., Kumar, R., Singh, M., & Sinha, B. (2014). Differential expression of pathogenesis related protein genes in tomato during inoculation with *A. solani*. *Journal of Plant Pathology & Microbiology*, **5**, 217-223.
- Utsumi, Y., Tanaka, M., Kurotani, A., Yoshida, T., Mochida, K., Matsui, A., Ishitani, M., Sraphet, S., Whankaew, S., & Asvarak, T. (2016). Cassava (*Manihot esculenta*) transcriptome analysis in response to infection by the fungus *Colletotrichum gloeosporioides* using an oligonucleotide-DNA microarray. *Journal of Plant Research*, **129**, 711-726.
- Utsumi, Y., Tanaka, M., Morosawa, T., Kurotani, A., Yoshida, T., Mochida, K., Matsui, A., Umemura, Y., Ishitani, M., & Shinozaki, K. (2012). Transcriptome analysis using a high-density oligomicroarray under drought stress in various genotypes of cassava: an important tropical crop. *DNA Research*, **19**, 335-345.
- Valentini, A., Pompanon, F., & Taberlet, P. (2009). DNA barcoding for ecologists. *Trends in Ecology & Evolution*, **24**, 110-117.
- Van Dam, N. M., Hadwich, K., & Baldwin, I. T. (2000). Induced responses in *Nicotiana attenuata* affect behavior and growth of the specialist herbivore *Manduca sexta*. *Oecologia*, **122**, 371-379.
- van Dam, S., Vösa, U., van der Graaf, A., Franke, L., & de Magalhães, J. P. (2017). Gene co-expression analysis for functional classification and gene-disease predictions. *Briefings in Bioinformatics*, **139**, 1-18.
- Van Poecke, R. M. (2007). *Arabidopsis*-insect interactions. *The Arabidopsis Book/American Society of Plant Biologists*, **5**, 107-140.
- Varshney, R. K., Chen, W., Li, Y., Bharti, A. K., Saxena, R. K., Schlueter, J. A., Donoghue, M. T., Azam, S., Fan, G., & Whaley, A. M. (2012). Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. *Nature Biotechnology*, **30**, 83-89.

- Varshney, R. K., Nayak, S. N., May, G. D., & Jackson, S. A. (2009). Next-generation sequencing technologies and their implications for crop genetics and breeding. *Trends in Biotechnology*, **27**, 522-530.
- Venu, R., Madhav, M. S., Sreerekha, M., Nobuta, K., Zhang, Y., Carswell, P., Boehm, M. J., Meyers, B. C., Korth, K. L., & Wang, G.-L. (2010). Deep and comparative transcriptome analysis of rice plants infested by the beet armyworm (*Spodoptera exigua*) and water weevil (*Lissorhoptrus oryzophilus*). *Rice*, **3**, 22-35.
- Verhage, A., van Wees, S. C., & Pieterse, C. M. (2010). Plant immunity: it's the hormones talking, but what do they say? *Plant Physiology*, **154**, 536-540.
- Vieira, L. d. J., Soares, T. L., Rossi, M. L., Alves, A. A. C., Santos, F. d. A. R. d., & Souza, F. V. D. (2012). Viability, production and morphology of pollen grains for different species in the genus *Manihot* (Euphorbiaceae). *Acta Botanica Brasilica*, **26**, 350-356.
- Von Dahl, C. C., & Baldwin, I. T. (2007). Deciphering the role of ethylene in plant-herbivore interactions. *Journal of Plant Growth Regulation*, **26**, 201-209.
- von Dohlen, C. D., & Moran, N. A. (1995). Molecular phylogeny of the Homoptera: a paraphyletic taxon. *Journal of Molecular Evolution*, **41**, 211-223.
- Wall, P. K., Leebens-Mack, J., Chanderbali, A. S., Barakat, A., Wolcott, E., Liang, H., Landherr, L., Tomsho, L. P., Hu, Y., & Carlson, J. E. (2009). Comparison of next generation sequencing technologies for transcriptome characterization. *BMC Genomics*, **10**, 347-365.
- Walton, V. M., Kruger, K., Saccaggi, D. L., & Millar, I. M. (2009). A survey of scale insects (Sternorrhyncha: Coccoidea) occurring on table grapes in South Africa. *Journal of Insect Science*, **9**, 1-6.
- Wang, W., Feng, B., Xiao, J., Xia, Z., Zhou, X., Li, P., Zhang, W., Wang, Y., Møller, B. L., & Zhang, P. (2014). Cassava genome from a wild ancestor to cultivated varieties. *Nature Communications*, **5**, 1-9.
- Wang, Y., & Wu, W.-H. (2013). Potassium transport and signaling in higher plants. *Annual Review of Plant Biology*, **64**, 451-476.
- Wang, Y., Zhou, L., & Yu, X. (2016). Transcriptome profiling of Huanglongbing (HLB) tolerant and susceptible citrus plants reveals the role of basal resistance in HLB tolerance. *Frontiers in Plant Science*, **7**, 1-13.
- War, A. R., Paulraj, M. G., Ahmad, T., Buhroo, A. A., Hussain, B., Ignacimuthu, S., & Sharma, H. C. (2012). Mechanisms of plant defense against insect herbivores. *Plant Signaling & Behavior*, **7**, 1306-1320.
- Wasswa, P., Alicai, A., & Mukasa, S. (2010). Optimisation of *in vitro* techniques for Cassava brown streak virus elimination from infected cassava clones. *African Crop Science Journal*, **18**, 235-241.
- Watson, G., & Kubiriba, J. (2005). Identification of mealybugs (Hemiptera: Pseudococcidae) on banana and plantain in Africa. *African Entomology*, **13**, 35-47.
- Weeto, M., Koekemoer, L., Kamau, L., Hunt, R., & Coetzee, M. (2004). Evaluation of a species-specific PCR assay for the *Anopheles funestus* group from eleven African countries and Madagascar. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **98**, 142-147.
- Wei, J., Wang, L., Zhu, J., Zhang, S., Nandi, O. I., & Kang, L. (2007). Plants attract parasitic wasps to defend themselves against insect pests by releasing hexenol. *PLoS One*, **2**, 852-852.
- Wei, X., Zhang, X., Shen, D., Wang, H., Wu, Q., Lu, P., Qiu, Y., Song, J., Zhang, Y., & Li, X. (2013). Transcriptome analysis of *Barbarea vulgaris* infested with diamondback moth (*Plutella xylostella*) larvae. *PLoS One*, **8**, 64481-64499.
- Wetten, A., Campbell, C., & Allainguillaume, J. (2016). High-resolution melt and morphological analyses of mealybugs (Hemiptera: Pseudococcidae) from cacao: tools for the control of Cacao swollen shoot virus spread. *Pest Management Science*, **72**, 527-533.
- Whiting, M. F., Carpenter, J. C., Wheeler, Q. D., & Wheeler, W. C. (1997). The Strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. *Systematic Biology*, **46**, 1-68.
- Wicker, T., Taudien, S., Houben, A., Keller, B., Graner, A., Platzer, M., & Stein, N. (2009). A whole-genome snapshot of 454 sequences exposes the composition of the barley genome and provides evidence for parallel evolution of genome size in wheat and barley. *The Plant Journal*, **59**, 712-722.

- Williams, D. J., & de Willink, M. C. G. (1992). *Mealybugs of central and South America*: CAB International, Wallingford, UK. pp 635.
- Williamson, C. E. (1950). Ethylene, a metabolic product of diseased and injured plants. *Phytopathology*, **40**, 205-208.
- Wilson, W. (2002). Soils utilized for gardens by Tukanoans in northwestern Amazonia and their impact on cassava (*Manihot esculenta* Crantz) cultivar selection. *Culture & Agriculture*, **24**, 20-30.
- Winz, R. A., & Baldwin, I. T. (2001). Molecular interactions between the specialist herbivore *Manduca sexta* (Lepidoptera, Sphingidae) and its natural host *Nicotiana attenuata*. IV. Insect-induced ethylene reduces jasmonate-induced nicotine accumulation by regulating putrescine N-methyltransferase transcripts. *Plant Physiology*, **125**, 2189-2202.
- Wise, R. P., Moscou, M. J., Bogdanove, A. J., & Whitham, S. A. (2007). Transcript Profiling in Host-Pathogen Interactions. *Annual Reviews of Phytopathology*, **45**, 329-369.
- Worthey, E. A., Mayer, A. N., Syverson, G. D., Helbling, D., Bonacci, B. B., Decker, B., Serpe, J. M., Dasu, T., Tschannen, M. R., & Veith, R. L. (2011). Making a definitive diagnosis: successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. *Genetics in Medicine*, **13**, 255-262.
- Wu, J., & Baldwin, I. T. (2010). New insights into plant responses to the attack from insect herbivores. *Annual Review of Genetics*, **44**, 1-24.
- Wu, Q., Xu, L., Guo, J., Su, Y., & Que, Y. (2013). Transcriptome profile analysis of sugarcane responses to *Sporisorium scitaminea* infection using Solexa sequencing technology. *BioMed Research International*, **2013**, 1-9.
- Wydra, K., & Verdier, V. (2002). Occurrence of cassava diseases in relation to environmental, agronomic and plant characteristics. *Agriculture, Ecosystems & Environment*, **93**, 211-226.
- Xie, D.-X., Feys, B. F., James, S., Nieto-Rostro, M., & Turner, J. G. (1998). *COI1*: an *Arabidopsis* gene required for jasmonate-regulated defense and fertility. *Science*, **280**, 1091-1094.
- Xu, L., Liu, F., Lechner, E., Genschik, P., Crosby, W. L., Ma, H., Peng, W., Huang, D., & Xie, D. (2002). The *SCFCOI1* ubiquitin-ligase complexes are required for jasmonate response in *Arabidopsis*. *The Plant Cell*, **14**, 1919-1935.
- Yan, J., Zhang, C., Gu, M., Bai, Z., Zhang, W., Qi, T., Cheng, Z., Peng, W., Luo, H., & Nan, F. (2009). The *Arabidopsis* *CORONATINE INSENSITIVE1* protein is a jasmonate receptor. *The Plant Cell*, **21**, 2220-2236.
- Yang, F., Zhang, Y., Huang, Q., Yin, G., Pennerman, K. K., Yu, J., Liu, Z., Li, D., & Guo, A. (2015). Analysis of key genes of jasmonic acid mediated signal pathway for defense against insect damages by comparative transcriptome sequencing. *Scientific Reports*, **5**, 16500; doi:10.1038/srep16500
- Yang, T., Bao, S.-y., Ford, R., Jia, T.-j., Guan, J.-p., He, Y.-h., Sun, X.-l., Jiang, J.-y., Hao, J.-j., & Zhang, X.-y. (2012). High-throughput novel microsatellite marker of faba bean via next generation sequencing. *BMC Genomics*, **13**, 602-612.
- Yaninek, S., & Hanna, R. (2003). Cassava Green Mite in Africa—A unique example of successful classical biological control of a mite pest on a continental scale. In: Neuenschwander, C.B. and Langewald, J. (eds). *Biological control in IPM systems in Africa*. Oxon, UK, pp 61-76.
- Ye, J., Fang, L., Zheng, H., Zhang, Y., Chen, J., Zhang, Z., Wang, J., Li, S., Li, R., & Bolund, L. (2006). WEGO: a web tool for plotting GO annotations. *Nucleic Acids Research*, **34**, 293-297.
- Zagrobelny, M., Scheibye-Alsing, K., Jensen, N. B., Møller, B. L., Gorodkin, J., & Bak, S. (2009). 454 pyrosequencing based transcriptome analysis of *Zygaena filipendulae* with focus on genes involved in biosynthesis of cyanogenic glucosides. *BMC Genomics*, **10**, 574.
- Zhang, G., Gu, C., & Wang, D. (2009). Molecular mapping of soybean aphid resistance genes in PI 567541B. *Theoretical and Applied Genetics*, **118**, 473-482.
- Zhang, J., Chiodini, R., Badr, A., & Zhang, G. (2011a). The impact of next-generation sequencing on genomics. *Journal of Genetics and Genomics*, **38**, 95-109.
- Zhang, M., Xie, L., Yin, Z., Khanal, S. K., & Zhou, Q. (2016). Biorefinery approach for cassava-based industrial wastes: current status and opportunities. *Bioresource Technology*, **215**, 50-62.
- Zhang, P., Vanderschuren, H., Fütterer, J., & Grisse, W. (2005). Resistance to cassava mosaic disease in transgenic cassava expressing antisense RNAs targeting virus replication genes. *Plant Biotechnology Journal*, **3**, 385-397.

- Zhang, P., Zhu, X., Huang, F., Liu, Y., Zhang, J., Lu, Y., & Ruan, Y. (2011b). Suppression of jasmonic acid-dependent defense in cotton plant by the mealybug *Phenacoccus solenopsis*. *PLoS One*, **6**, e22378.
- Zhang, P. J., Huang, F., Zhang, J. M., Wei, J. N., & Lu, Y. B. (2015a). The mealybug *Phenacoccus solenopsis* suppresses plant defense responses by manipulating JA-SA crosstalk. *Scientific Reports*, **5**, 9354-9360.
- Zhang, P. J., Broekgaarden, C., Zheng, S. J., Snoeren, T. A., Loon, J. J., Gols, R., & Dicke, M. (2013). Jasmonate and ethylene signaling mediate whitefly-induced interference with indirect plant defense in *Arabidopsis thaliana*. *New Phytologist*, **197**, 1291-1299.
- Zhang, X., Liu, T., Wei, X., Qiu, Y., Song, J., Wang, H., Shen, D., Agerbirk, N., & Li, X. (2015b). Expression patterns, molecular markers and genetic diversity of insect-susceptible and resistant *Barbarea* genotypes by comparative transcriptome analysis. *BMC Genomics*, **16**, 1-16.
- Zhang, P. J., Xu, C. X., Zhang, J. M., Lu, Y. B., Wei, J. N., Liu, Y. Q., David, A., Boland, W., & Turlings, T.C. (2013). Phloem-feeding whiteflies can fool their host plants, but not their parasitoids. *Functional Ecology*, **27**, 1304-1312.
- Zhou, B., Lin, J. Z., Peng, W. S., Peng, D., Zuo, Y. H., Zhu, D. F., Huang, X. Q., Tang, D. Y., Guo, M., He, R. Q., Zhang, J. H., Li, X. S., Zhao, X. Y., Liu, X. M. (2012). Dwarfism in *Brassica Napus* L. induced by the over-expression of a gibberellin 2-oxidase gene from *Arabidopsis thaliana*. *Molecular Breeding*, **29**, 115-127.
- Zhou, G., Wang, X., Yan, F., Li, R., Cheng, J., & Lou, Y. (2011). Genome- wide transcriptional changes and defence- related chemical profiling of rice in response to infestation by the rice striped stem borer *Chilo suppressalis*. *Physiologia Plantarum*, **143**, 21-40.
- Zhu, W.B., Lestander, T.A., Örberg, H., Wei, M., Hedman, B., Ren, J., Xie, G., & Xiong, S. (2015). Cassava stems: a new resource to increase food and fuel production. *Global Change Biology Bioenergy*, **7**, 72-83

APPENDIX

Appendix A1: Reagents, solutions and kits

Acetic acid - Sigma Aldrich, Switzerland
Agarose - Merck, Germany
Boric acid - Merck, Germany
 β -mercaptoethanol - Merck, Germany
Diethyl pyrocarbonate (DEPC) - Sigma Aldrich, Switzerland
Ethylene diaminetetraacetic acid (EDTA) - Sigma Aldrich, Switzerland
Ethanol - Merck, Germany
Ethidium bromide - Sigma Aldrich, Switzerland
Hydrochloric acid (HCl) - Sigma Aldrich, Switzerland
Murashige and Skoog medium powder - Sigma Aldrich, Switzerland
Plant agar - Sigma Aldrich, Switzerland
RiboRuler RNA ladder - Inqaba Biotec, SA
RNase AWAY – Whitehead Scientific, SA
Sodium hydroxide (NaOH) - Merck, Germany
Sugar (white) - Makro Supermarket, SA
Tris base - Sigma Aldrich, Switzerland
Qubit reagents - Life technologies Inc., USA
1kb DNA ladder - Inqaba Biotec, SA
6x loading dye - Inqaba Biotec, SA

Appendix A2: Solutions

Acid alcohol (95% glacial acetic acid and 50% alcohol)
Acid fuchsin (1.2% hydrochloric acid (v/v), 0.2% acid fuchsin (g/v) in 300 ml distilled water)
Carbol xylene (carbol xylene is made by dissolving one-part phenol crystals in three parts xylene)
Murashige and Skoog medium (4.4g MS media, 30g sugar, 7g plant agar)
TAE buffer (40mM Tris, 20mM acetic acid, 100mM EDTA)
TBE buffer (10.8g Tris base, 5.5 Boric acid, 0.5M Na₂EDTA)
TE buffer (10mM Tris-Cl, 1mM EDTA)
RNase free water (0.1% DEPC in deionized water)

Appendix A3: Extraction kits

DNeasy DNA extraction kit - Qiagen, Valencia, CA

Illumina TruSeq stranded mRNA sample preparation kit - Whitehead Scientific, SA
KAPA SYBR FAST Universal qPCR kit - KAPA Biosystems, SA
RNeasy Plant Mini kit - Qiagen, Valencia, CA
RNase-Free DNase set - Qiagen, Valencia, CA
TaKaRa PCR amplification kit (Clontech laboratories Inc, USA)
QIAquick PCR purification kit - Qiagen, Valencia, CA
QubitR dsRNA BR Assay kit (Invitrogen, Life Technologies, Oregon, USA)
QubitR dsRNA HS Assay kit (Invitrogen, Life Technologies, Oregon, USA)
QubitR ssDNA Assay kit (Invitrogen, Life Technologies, Oregon, USA)

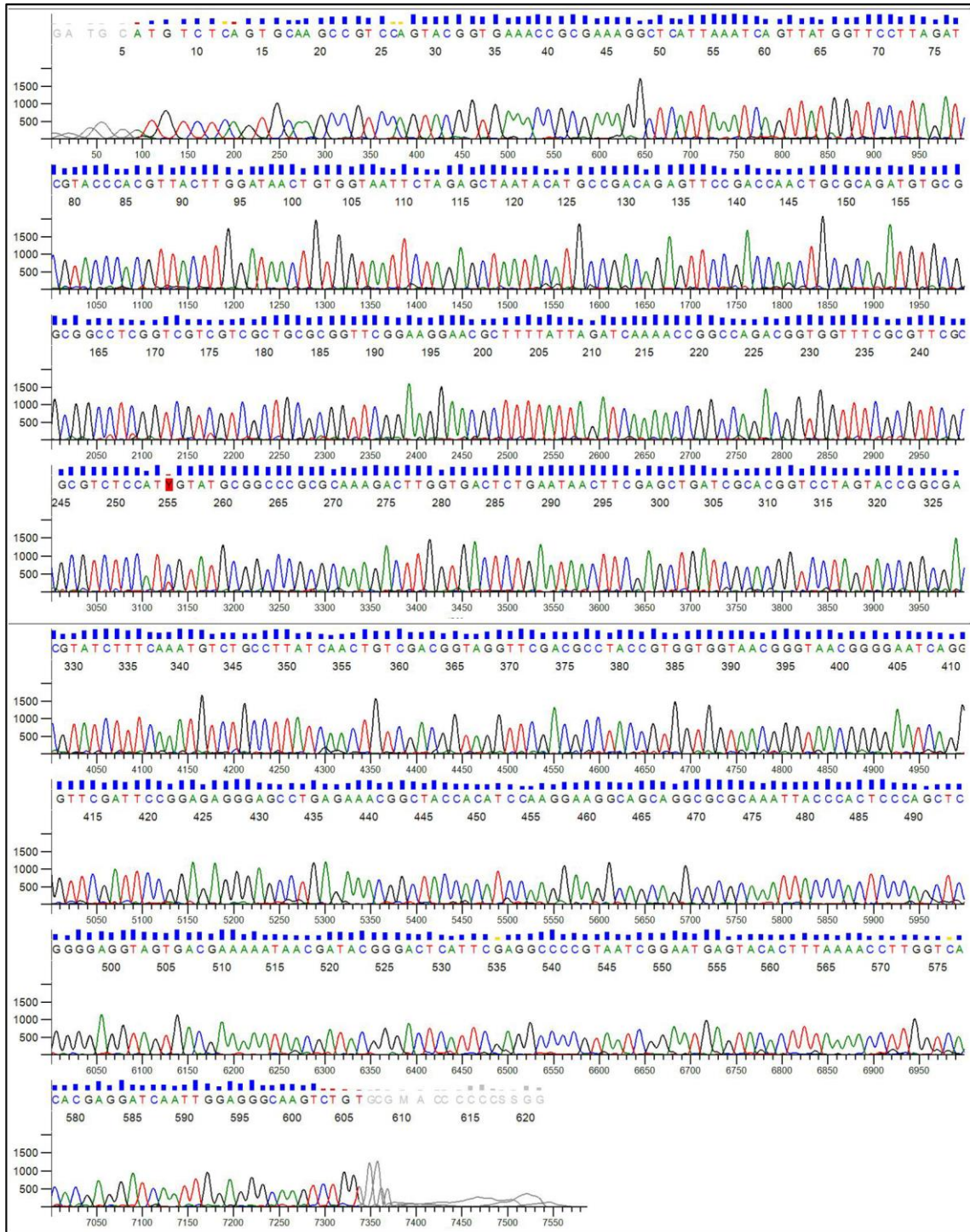


Figure A1: Example of raw Sanger CMB_1.1 18S forward rRNA sequences of mealybugs (*P. manihoti*) before alignments and analysis.

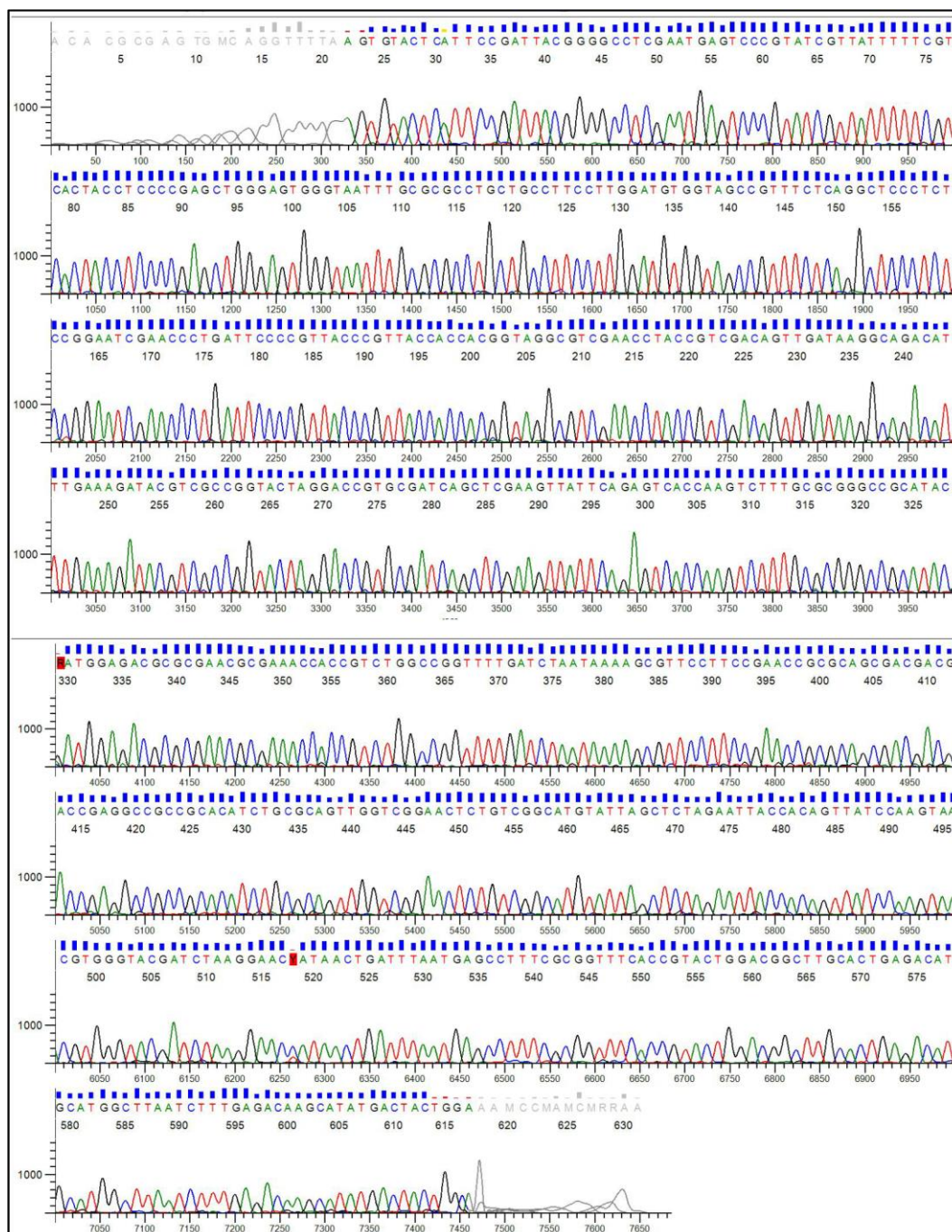
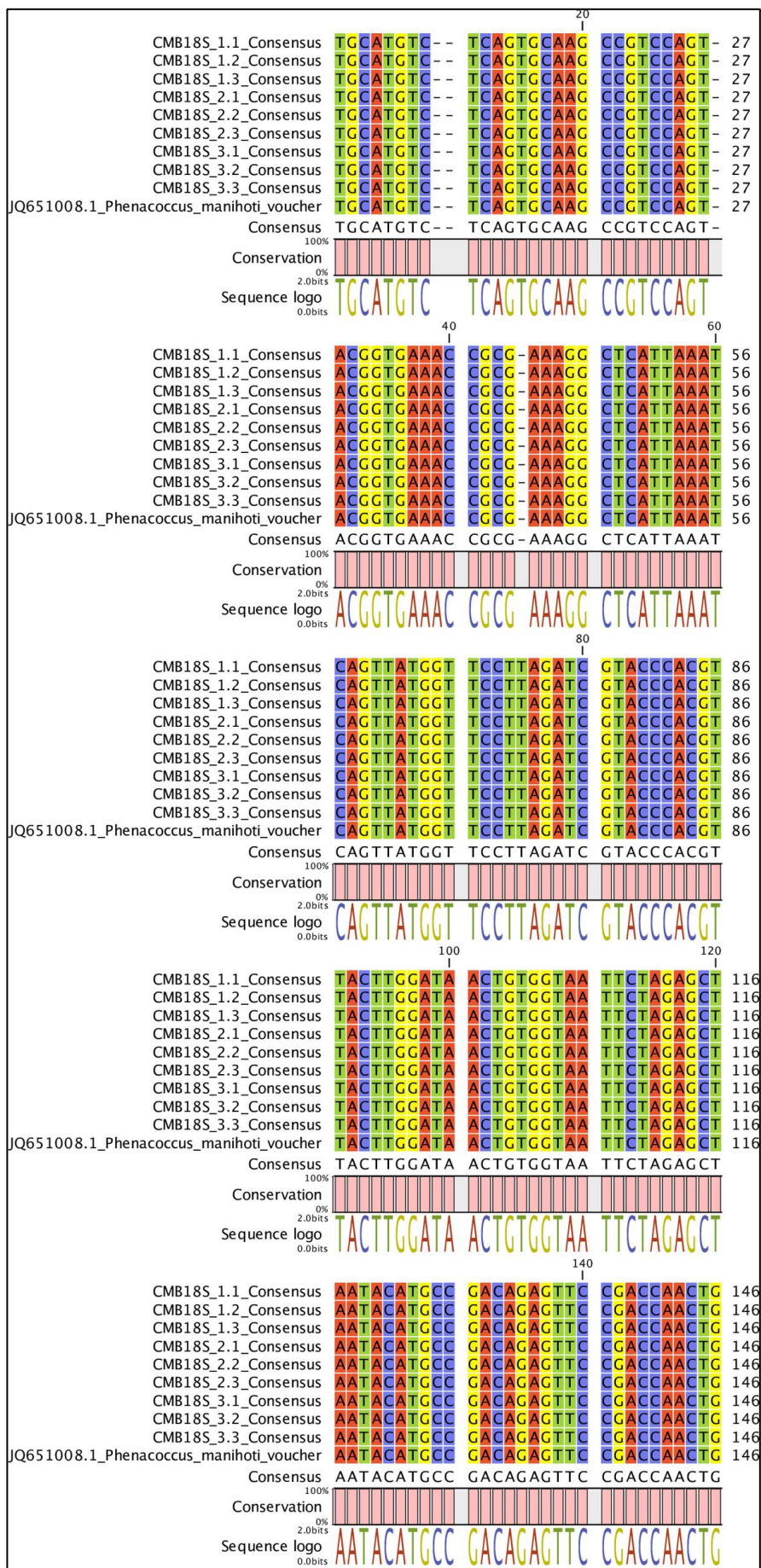
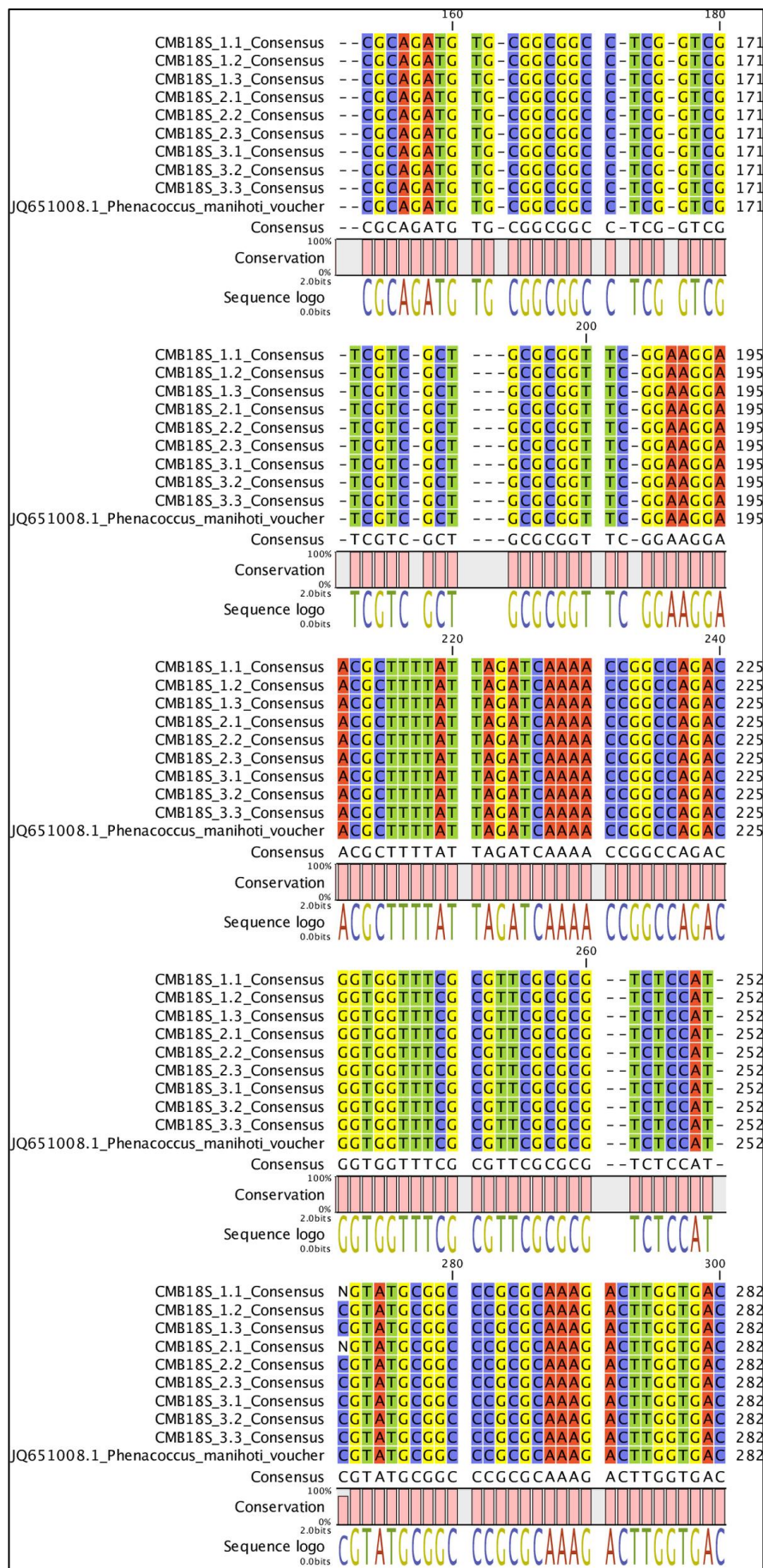
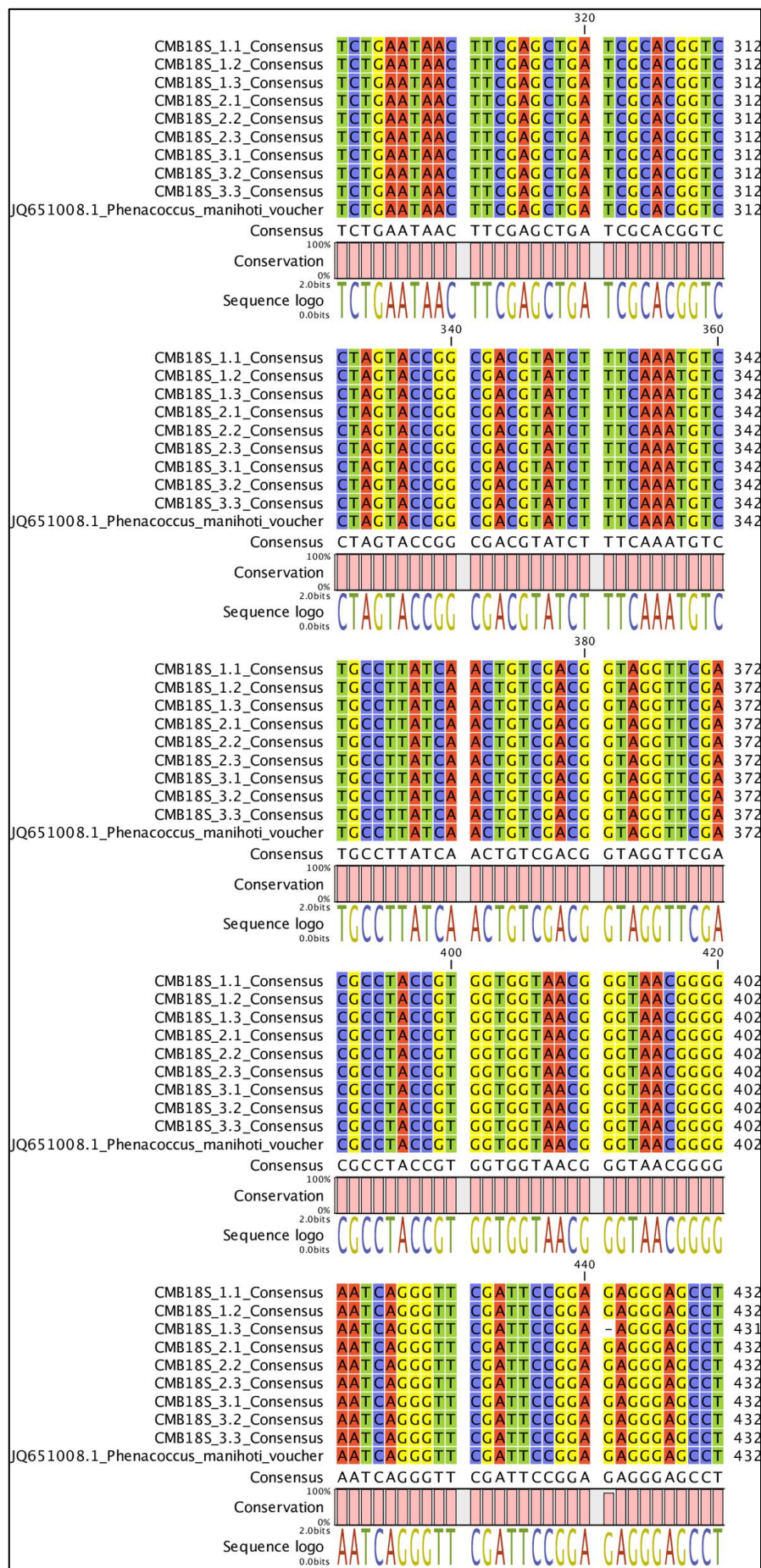


Figure A2: Example of raw Sanger CMB_1.1 18S reverse rRNA sequences of mealybugs (*P. manihoti*) before alignments and analysis.







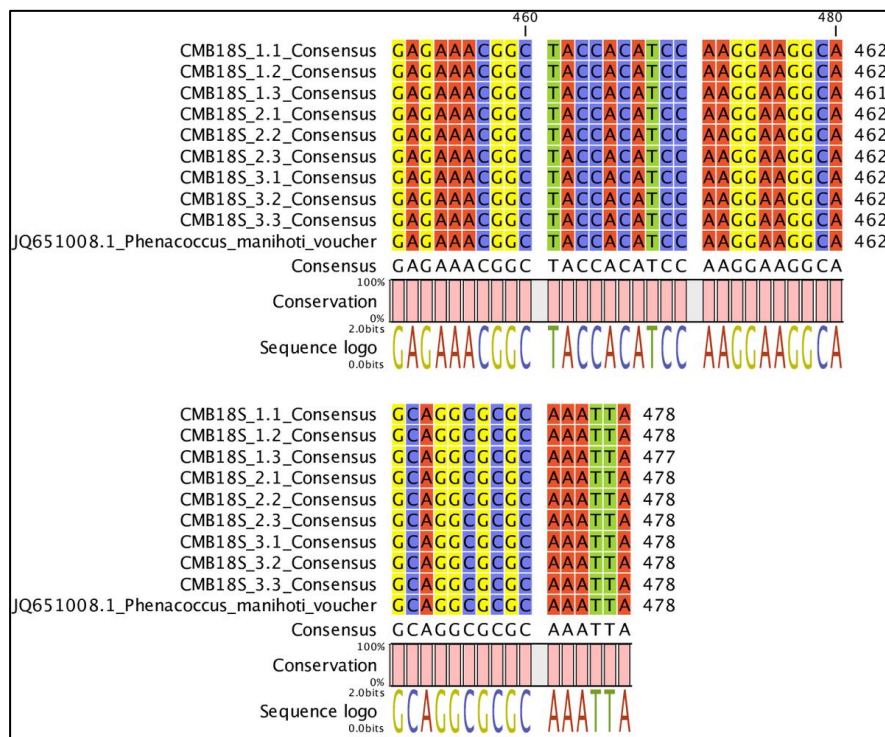
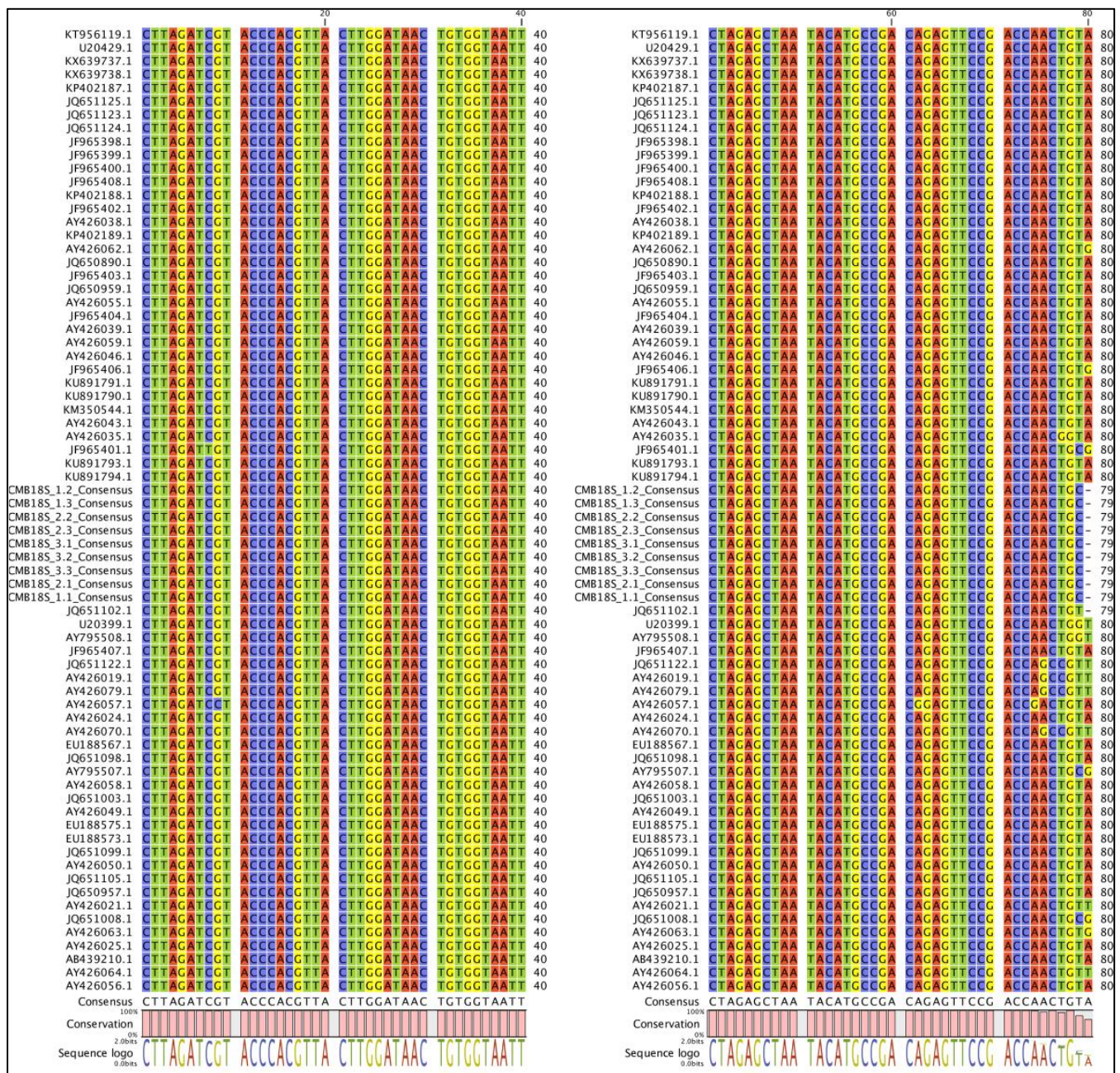
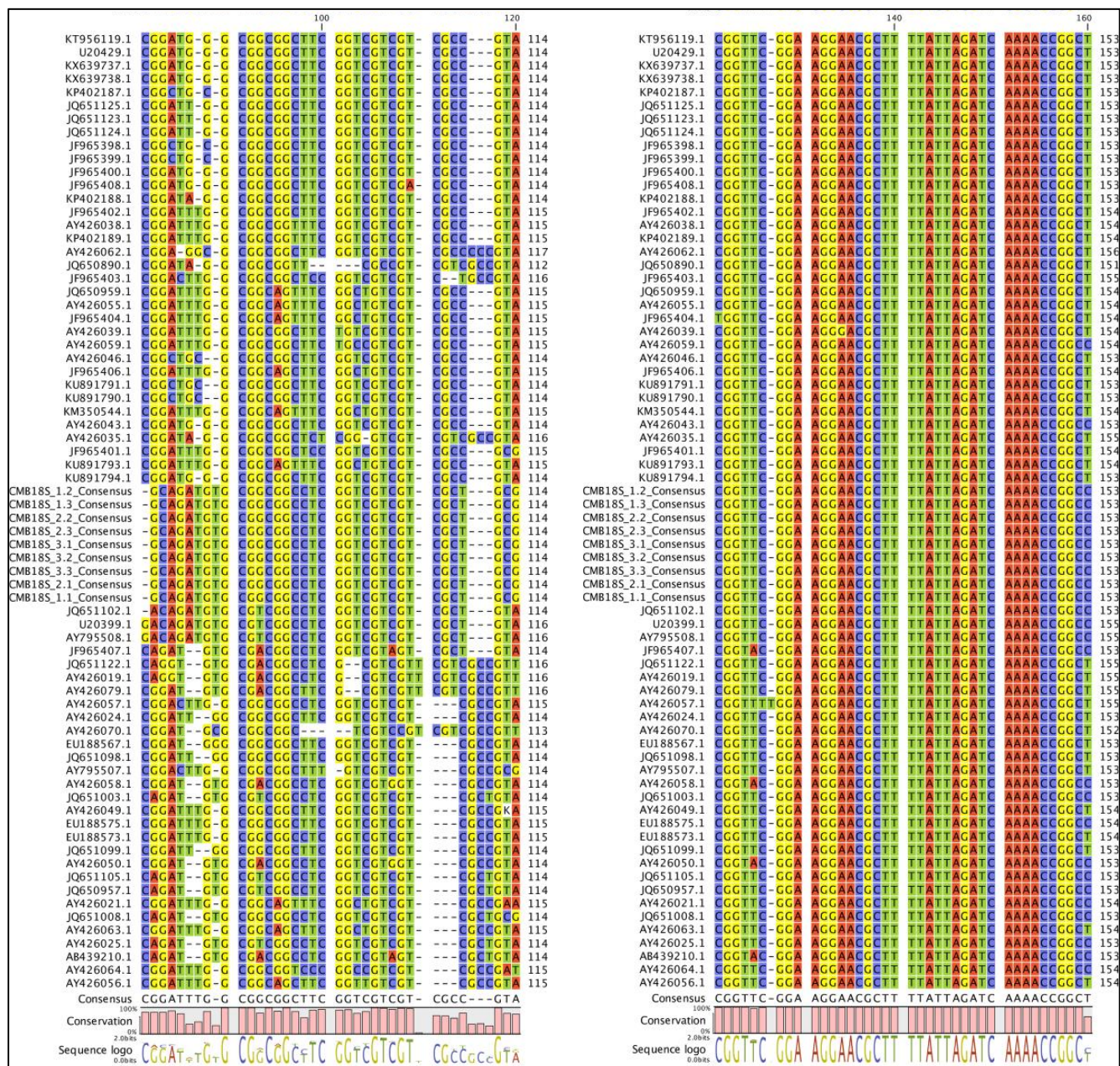


Figure A3: Sequence alignments of nine CMB 18S regions of species of mealybug from KwaZulu Natal and Mpumalanga in comparison with GenBank *P. manihoti* voucher species. The alignments were created using CLC Bio Genomics Workbench v9.0.





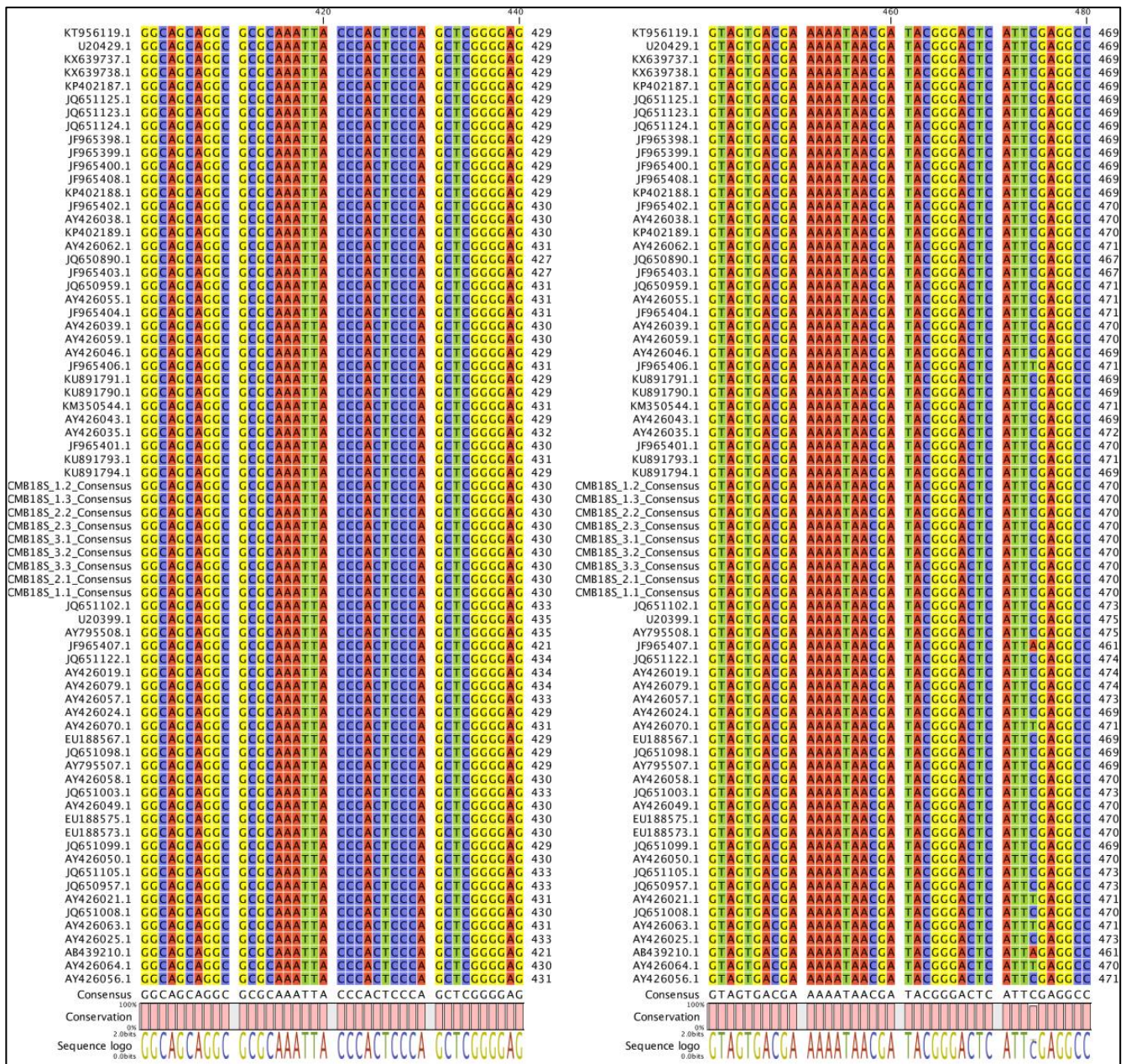


Figure A4: Multiple sequence alignments of 18S region of species of mealybug (*P. manihoti*) from KwaZulu Natal and Mpumalanga with comparison to the GenBank databases. The alignments were created using CLC Bio Genomics Workbench v9.0.

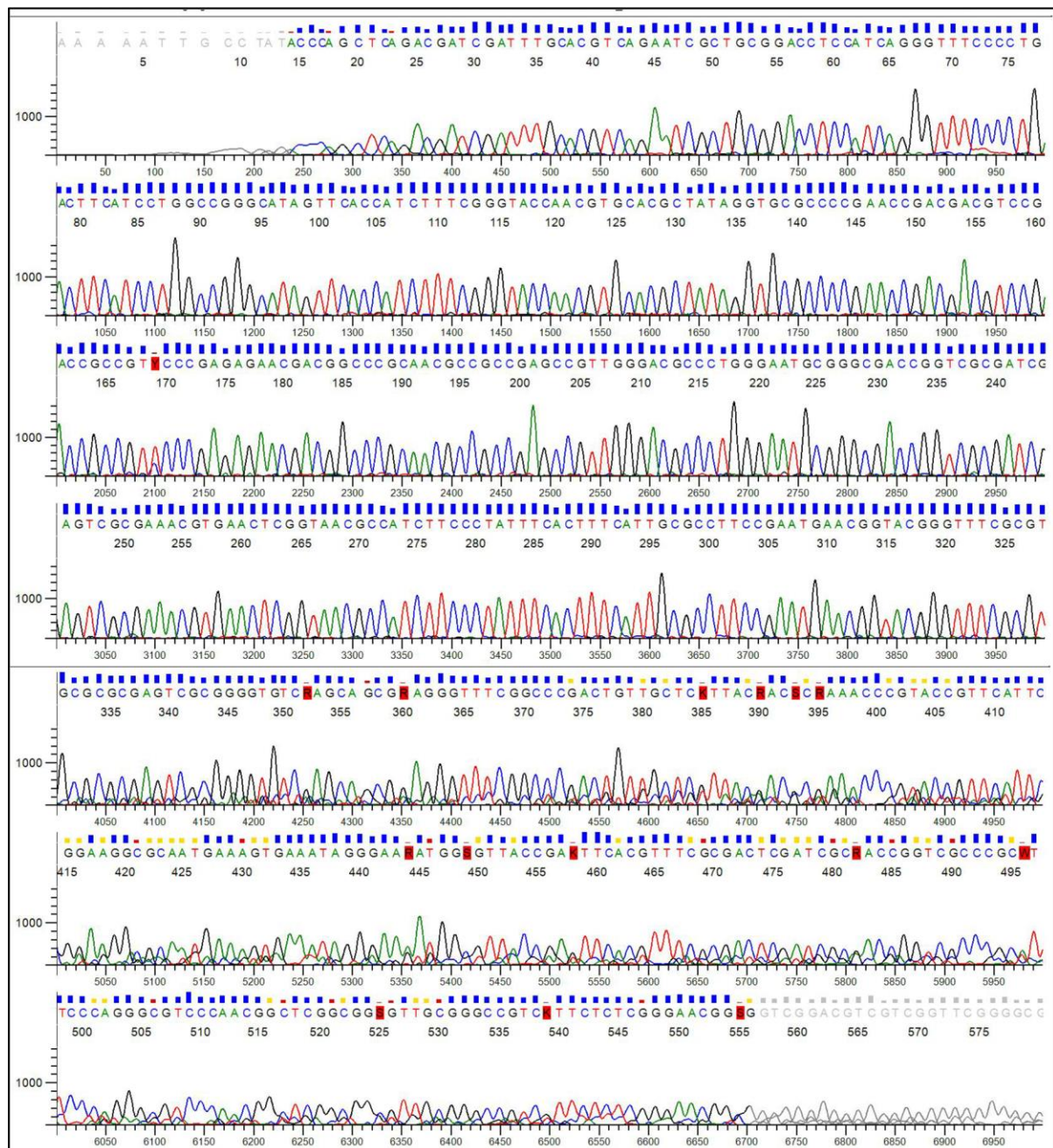


Figure A5: Example of raw Sanger CMB_1.1 28S forward rRNA sequences of mealybugs (*P. manihoti*) before alignments and analysis.

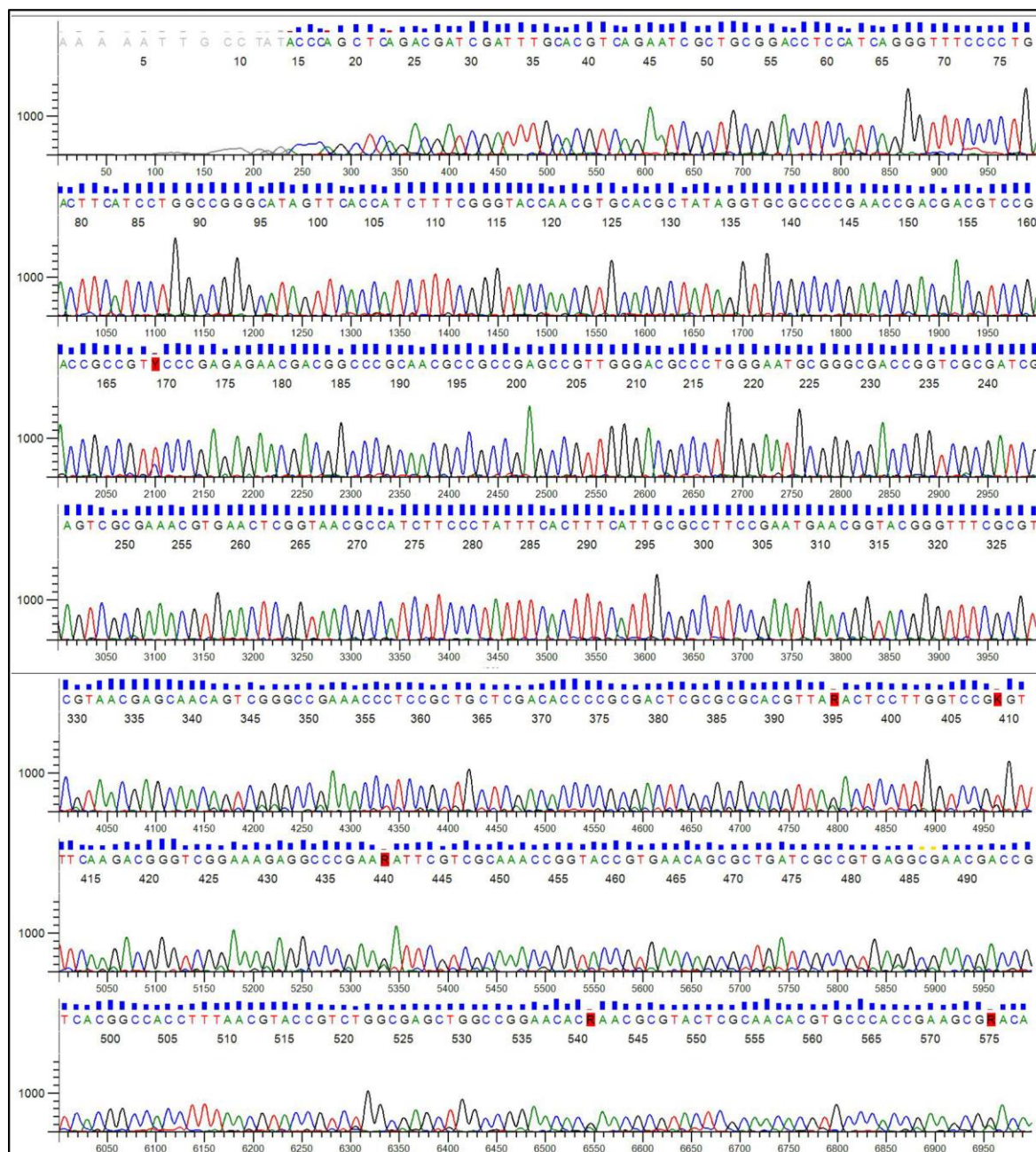
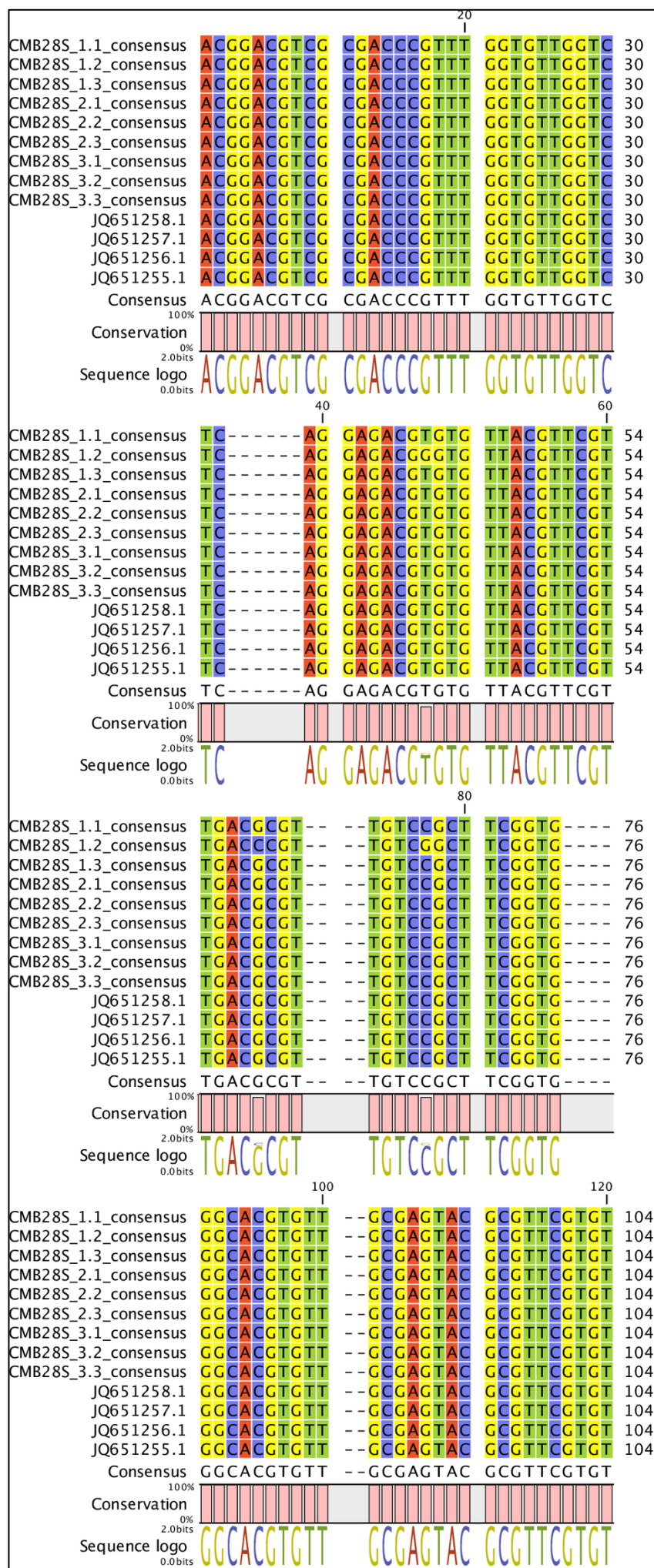
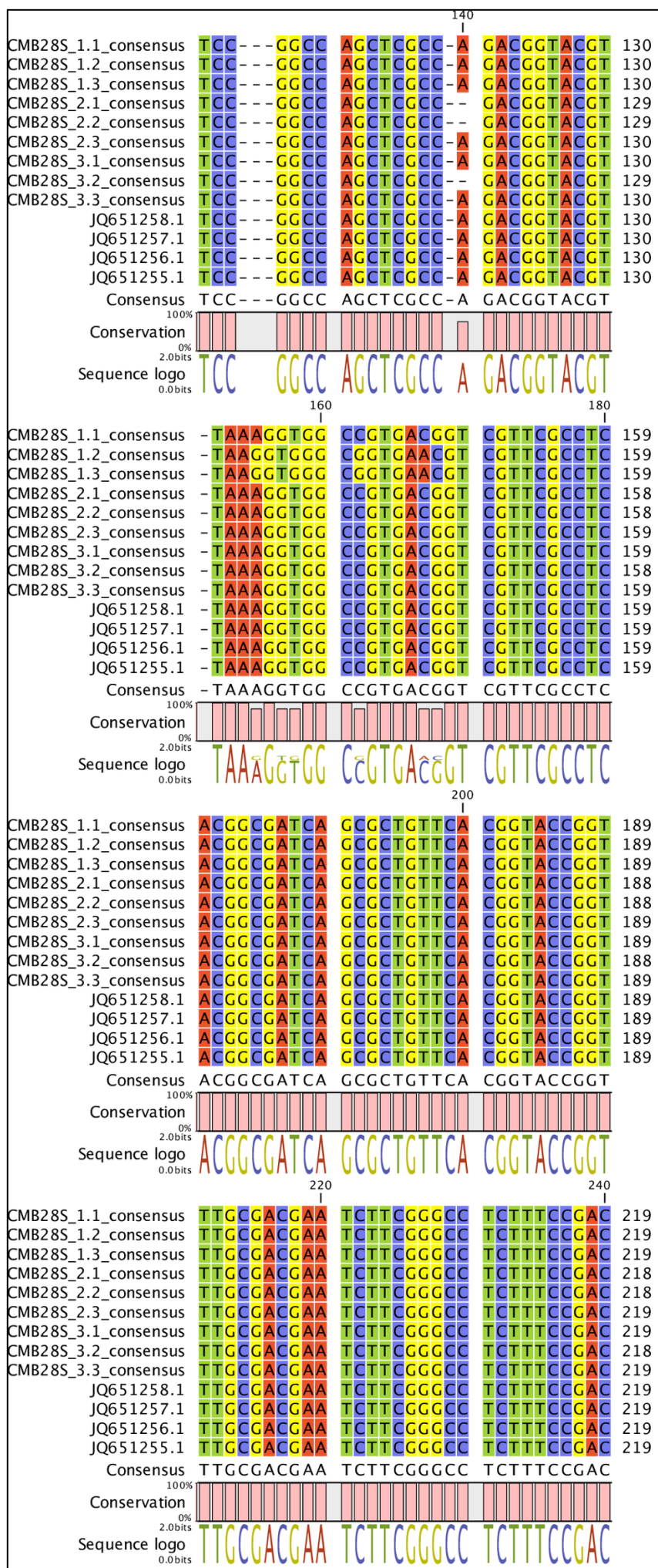
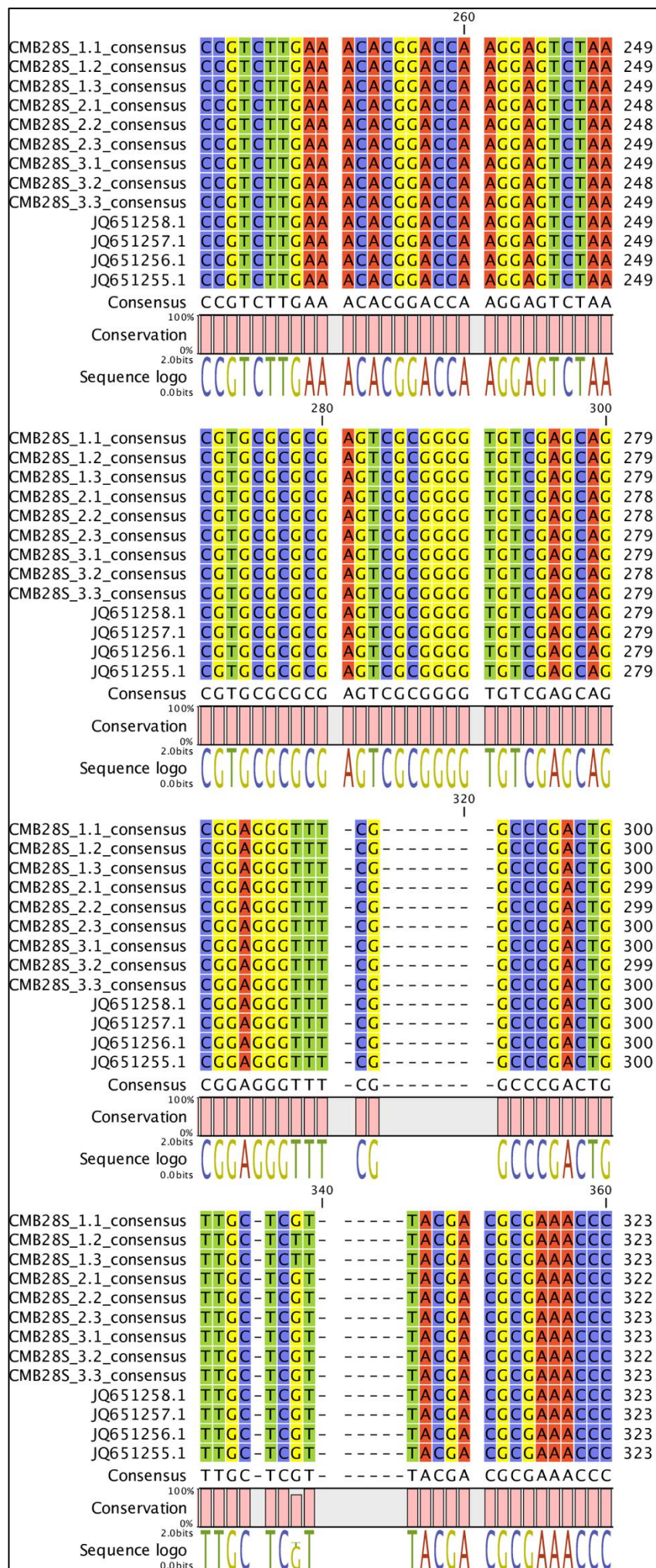
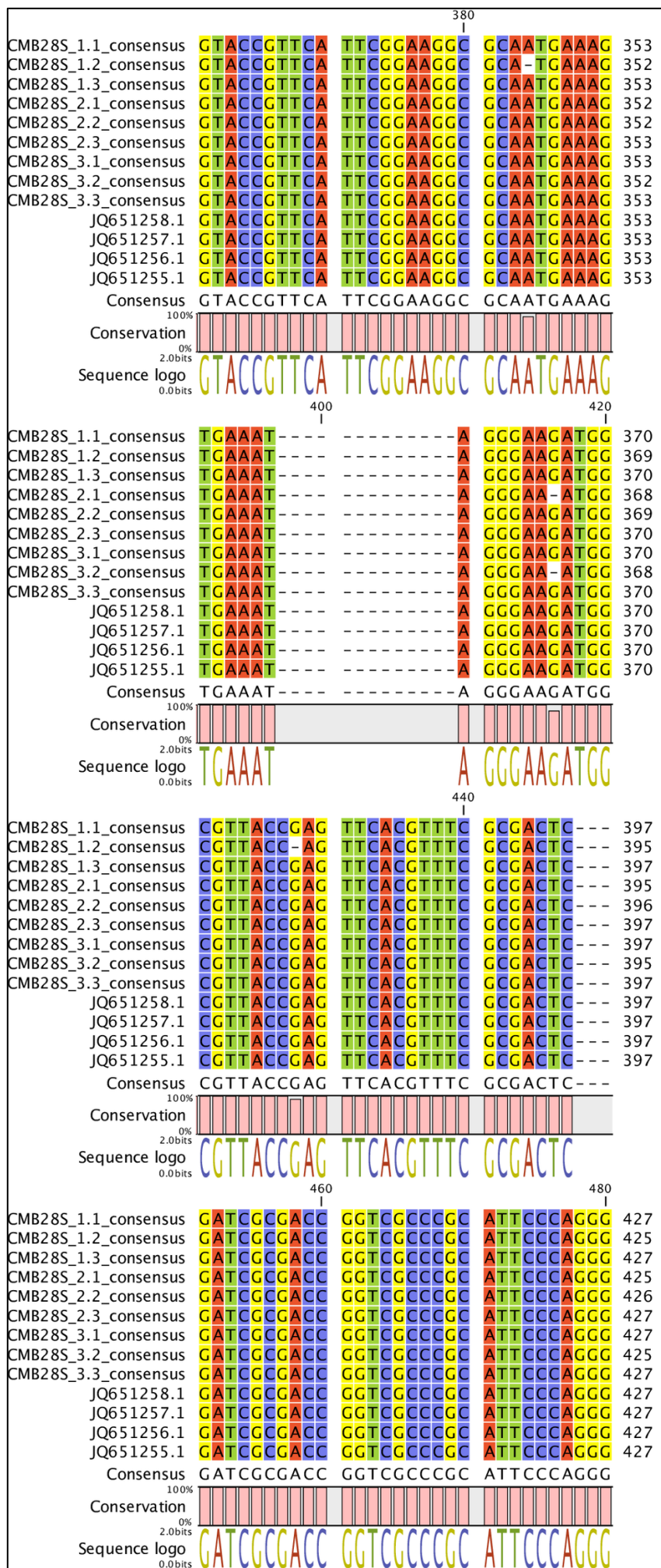


Figure A6: Example of raw Sanger CMB_1.1 28S reverse rRNA sequences of mealybugs (*P. manihoti*) before alignments and analysis.









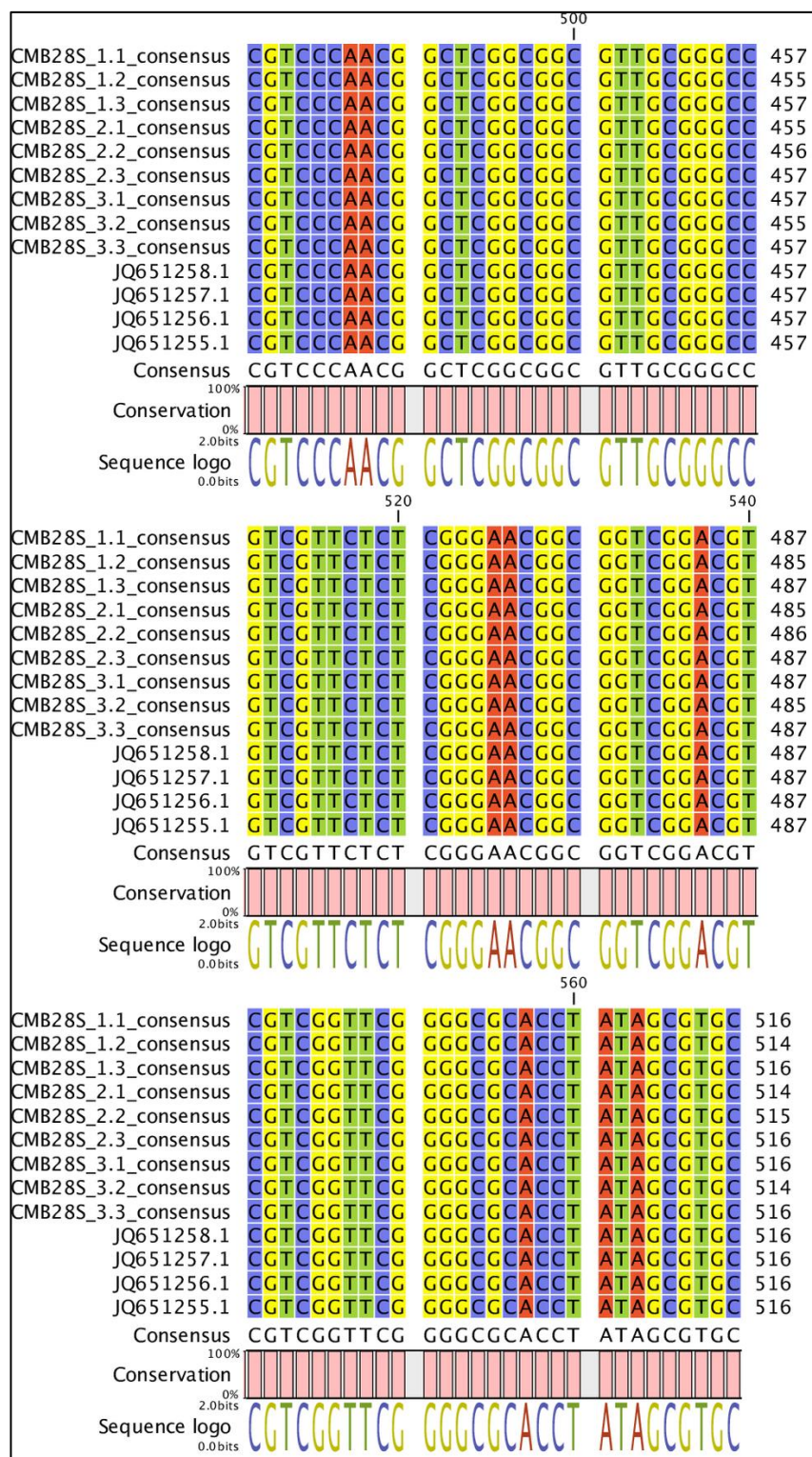
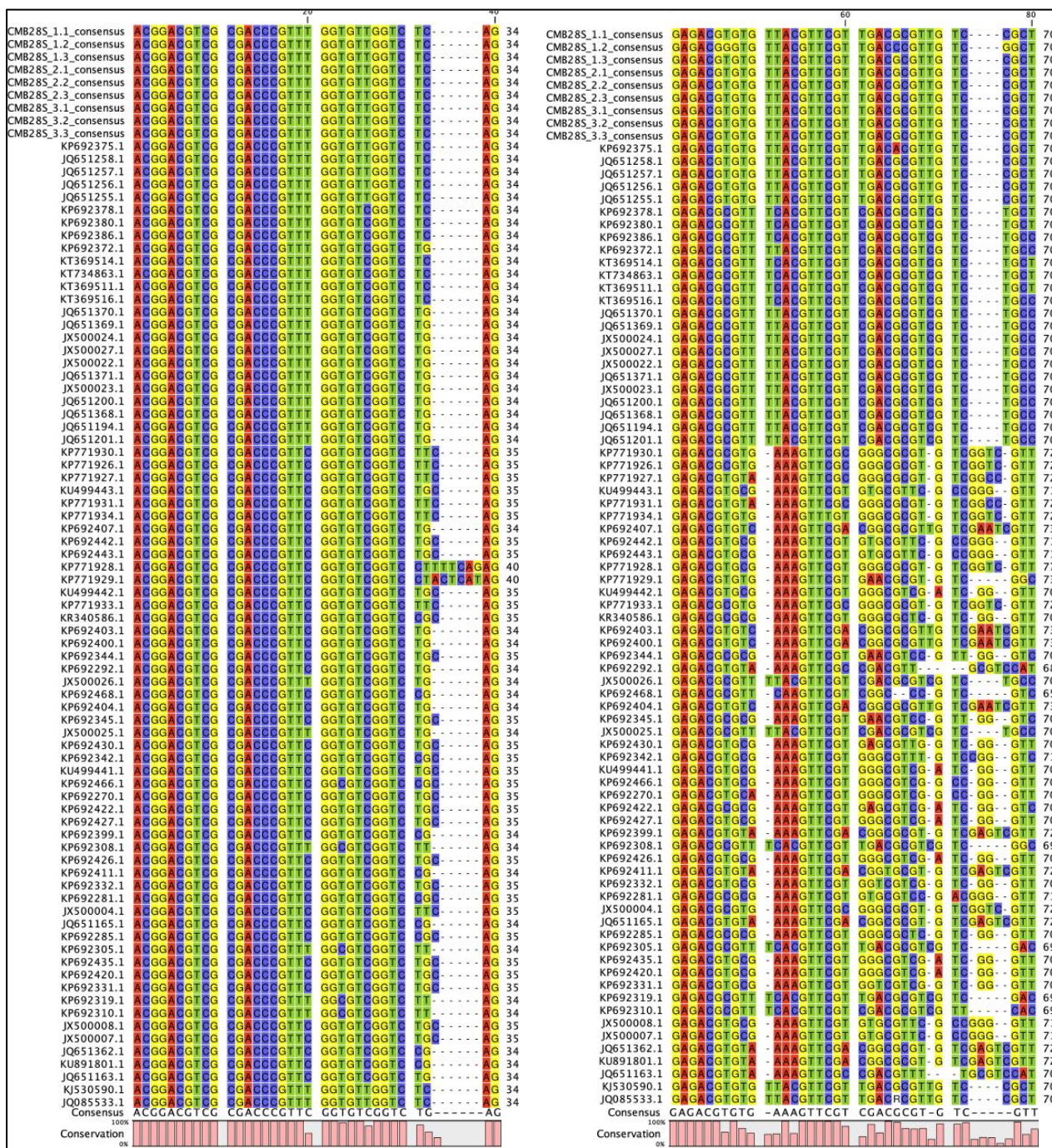
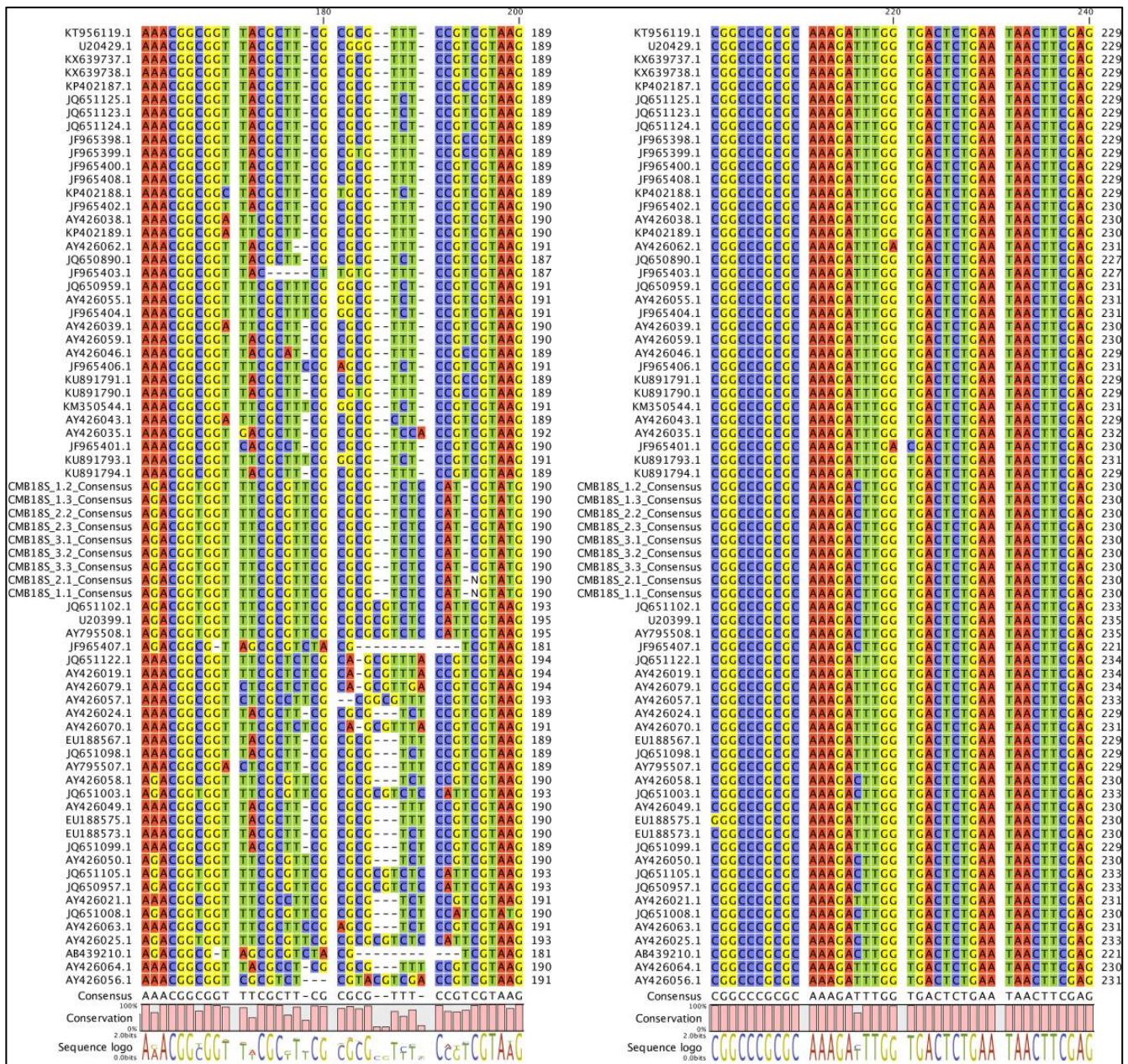
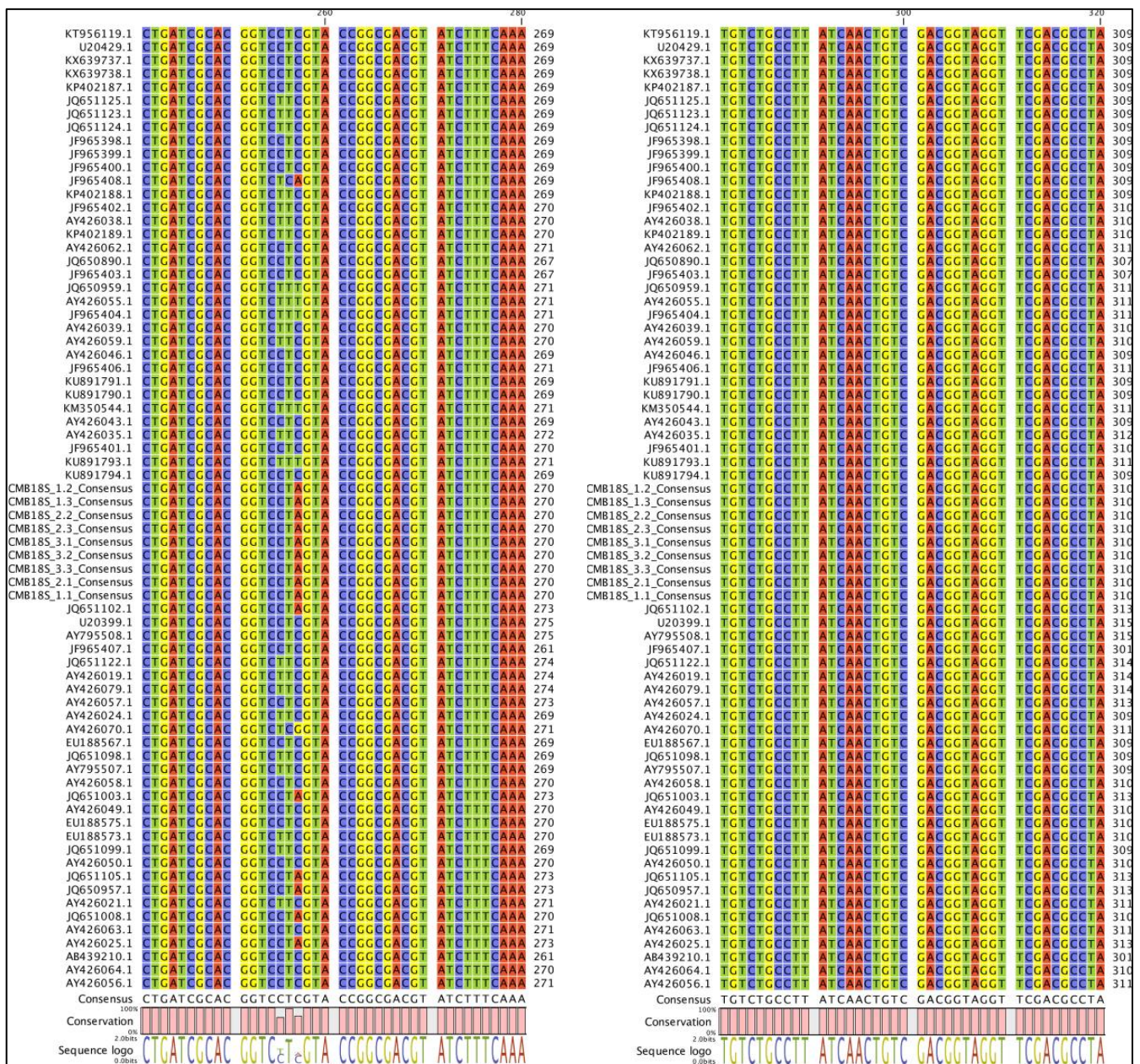


Figure A7: Sequence alignments of nine CMB 28S regions of species of mealybug from KwaZulu Natal and Mpumalanga in comparison with GenBank *P. manihoti* voucher species. The alignments were created using CLC Bio Genomics Workbench v9.0.



[illegible]





124

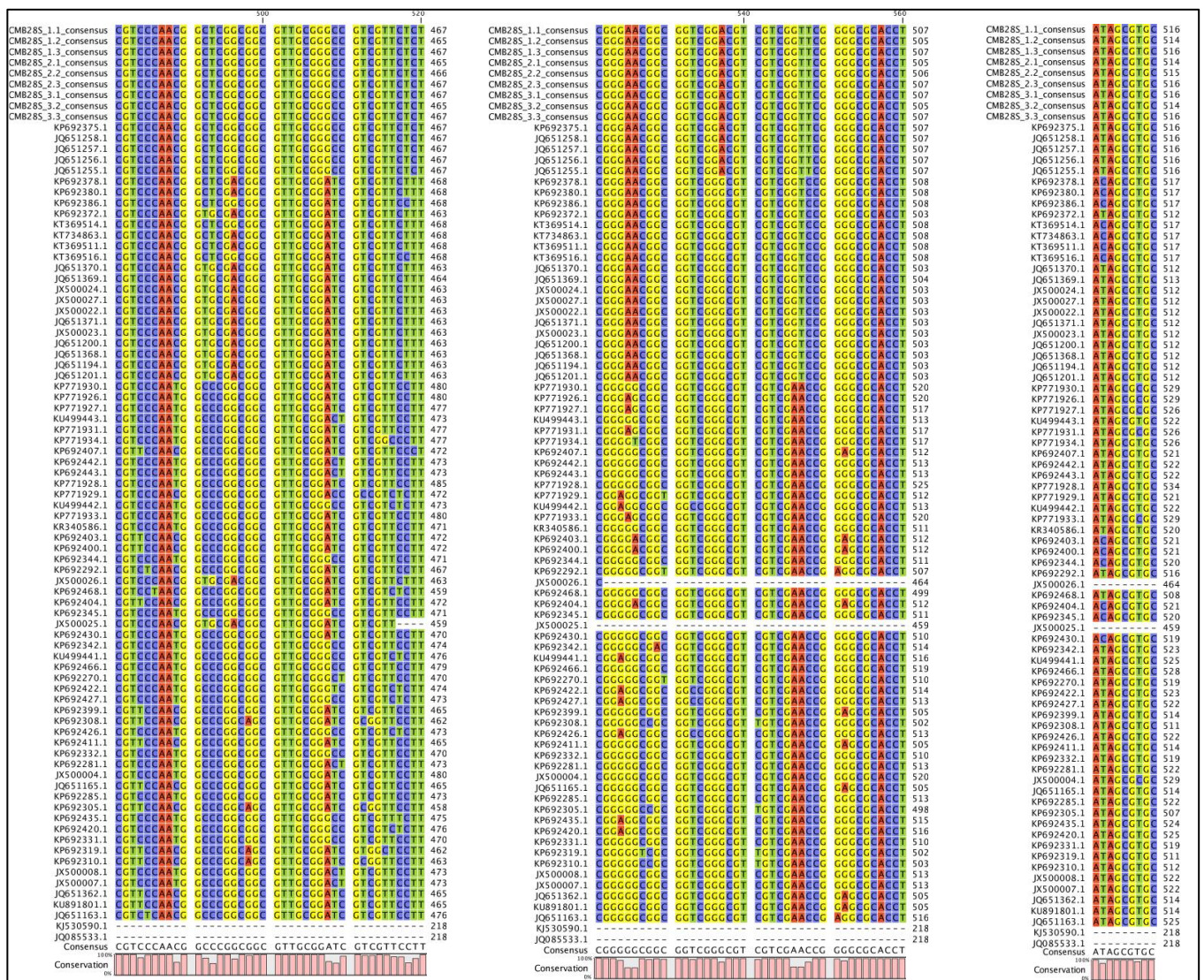


Figure A8: Multiple sequence alignments of 28S region of species of mealybug (*P. manihoti*) from KwaZulu Natal and Mpumalanga with comparison to the GenBank databases. The alignments were created using CLC Bio Genomics Workbench v9.0.

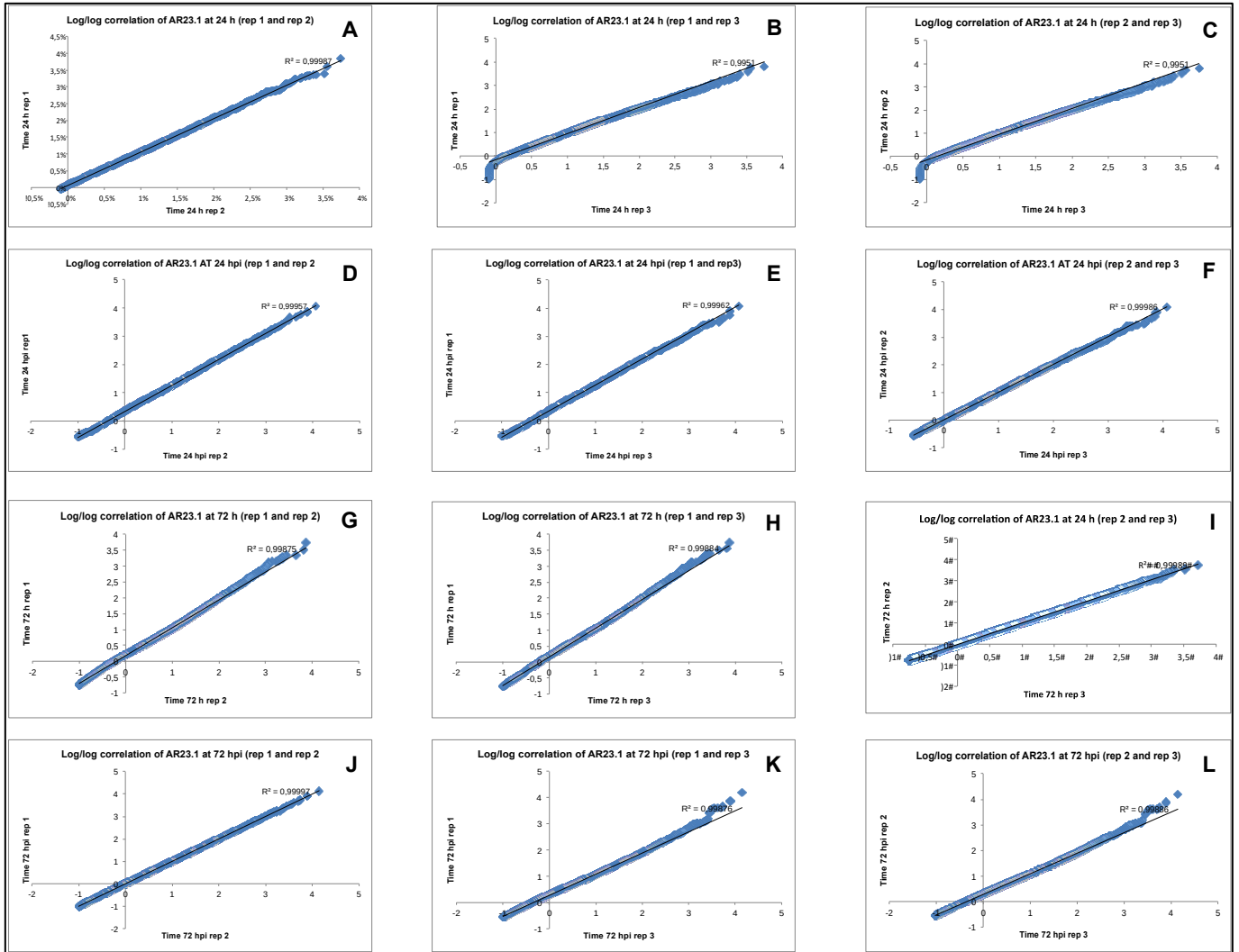


Figure A9: Log/log correlations of AR23.1 leaf samples infested with mealybugs (24 and 72 hours post-infestation) compared to non-infested controls. Images A-C show correlations of 3 biological replicates 24 h (non-infested control), D-F show correlations of 3 biological replicates at 24 hpi (infested at 24 h), G-I show correlations of 3 biological replicates at 72 h (non-infested control), while J-L show correlations of 3 biological replicates at 72 hpi (infested at 72 h).

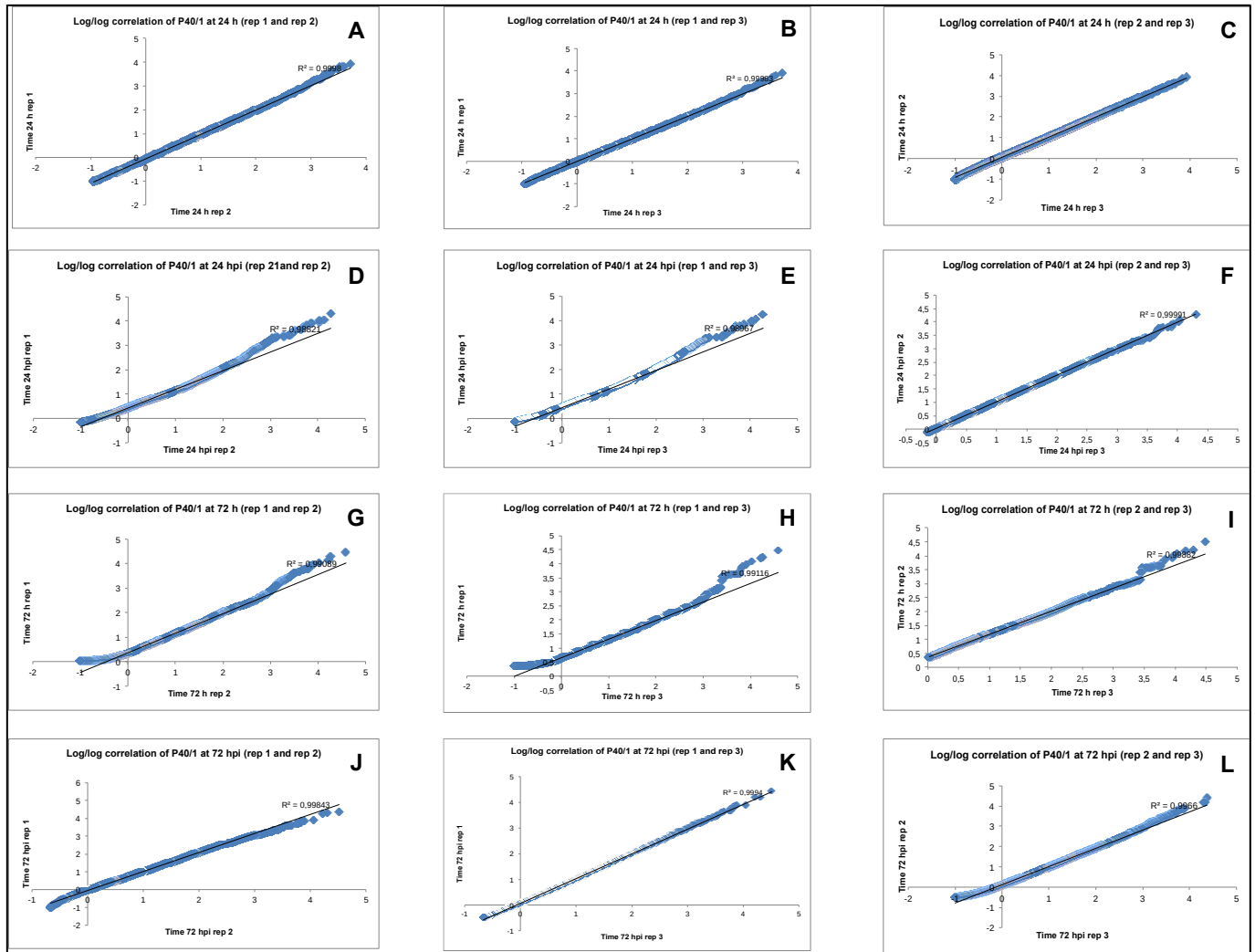


Figure A10: Log/log correlations of P40/1 leaf samples infested with mealybugs (24 and 72 hours post-infestation) compared to non-infested controls. Images A-C show correlations of 3 biological replicates 24 h (non-infested control), D-F show correlations of 3 biological replicates at 24 hpi (infested at 24 h), G-I show correlations of 3 biological replicates at 72 h (non-infested control), while J-L show correlations of 3 biological replicates at 72 hpi (infested at 72 h).

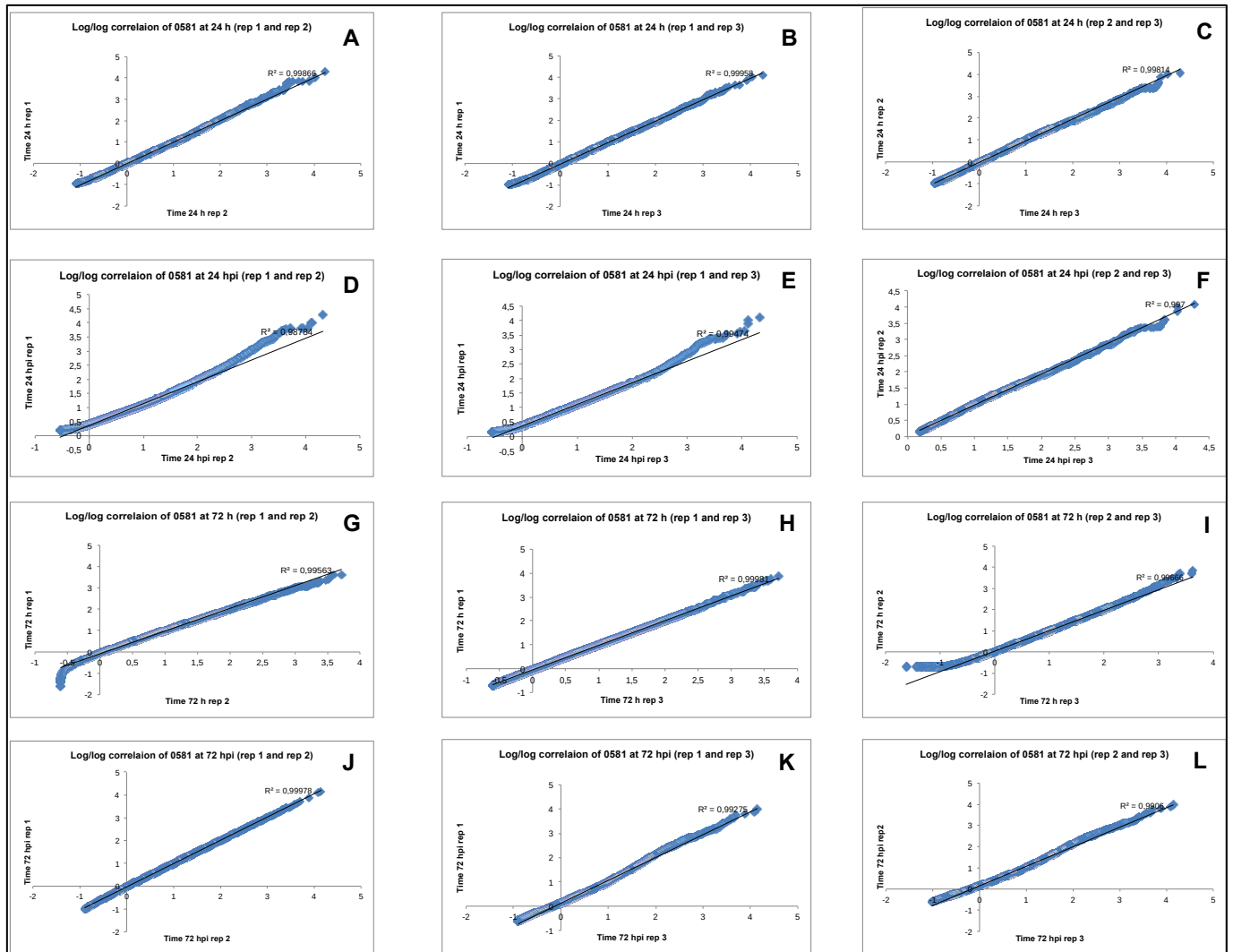


Figure A11: Log/log correlations of 98/0581 leaf samples infested with mealybugs (24 and 72 hours post-infestation) compared to non-infested controls. Images A-C show correlations of 3 biological replicates 24 h (non-infested control), D-F show correlations of 3 biological replicates at 24 hpi (infested at 24 h), G-I show correlations of 3 biological replicates at 72 h (non-infested control), while J-L show correlations of 3 biological replicates at 72 hpi (infested at 72 h).

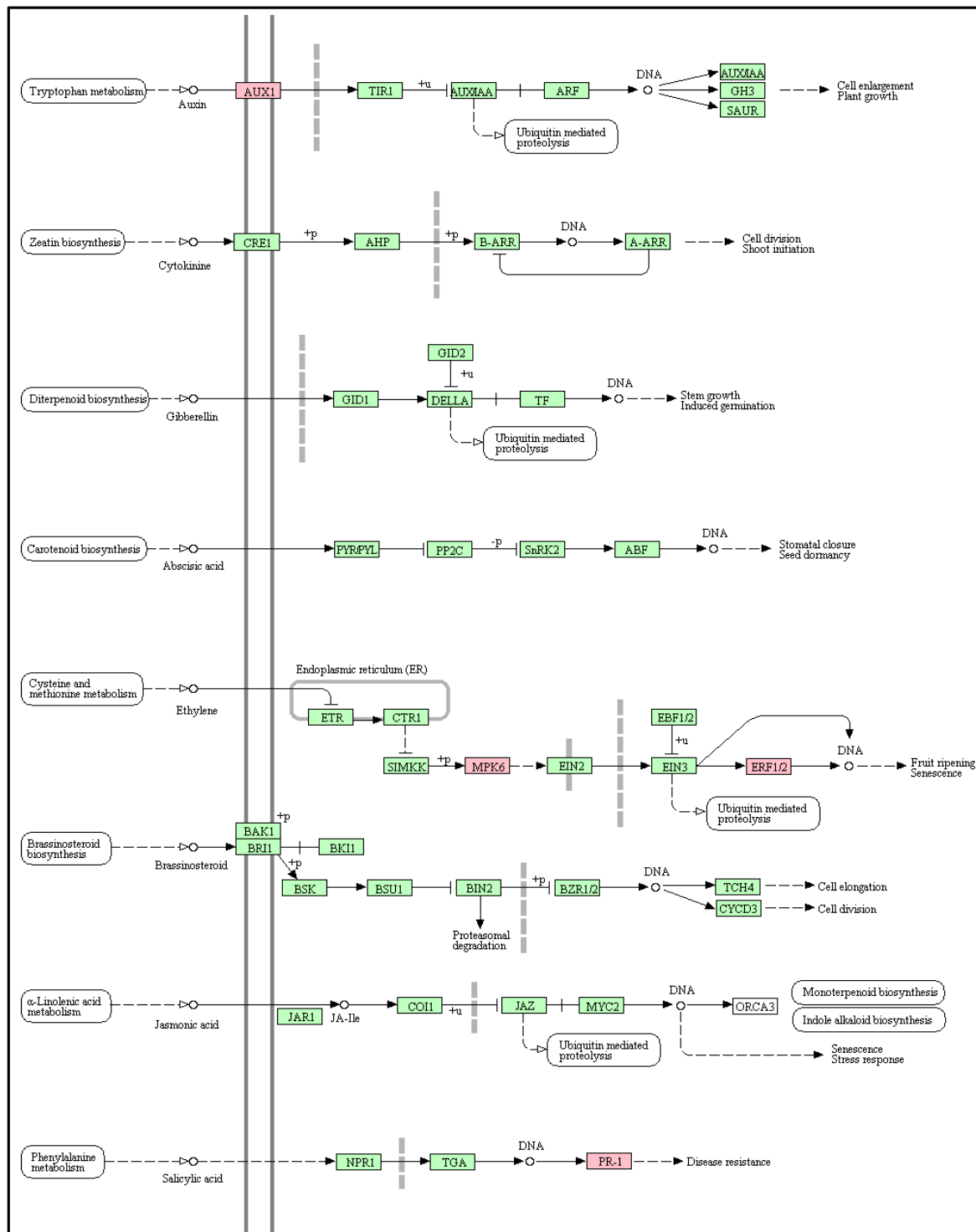


Figure A12: Localization of up-regulated plant hormone signal transduction pathway enzymes on the leaf tissues of AR23.1 post-infestation with mealybugs. Boxes highlighted in green represents genes that are present in the castor bean (*Ricinus communis*) genome (organism used for KEGG maps), while boxes highlighted in pink are the up-regulated DEGs in the leaf tissues post mealybug infestation. Boxes that are not highlighted are the genes that are not present in the castor bean genome.

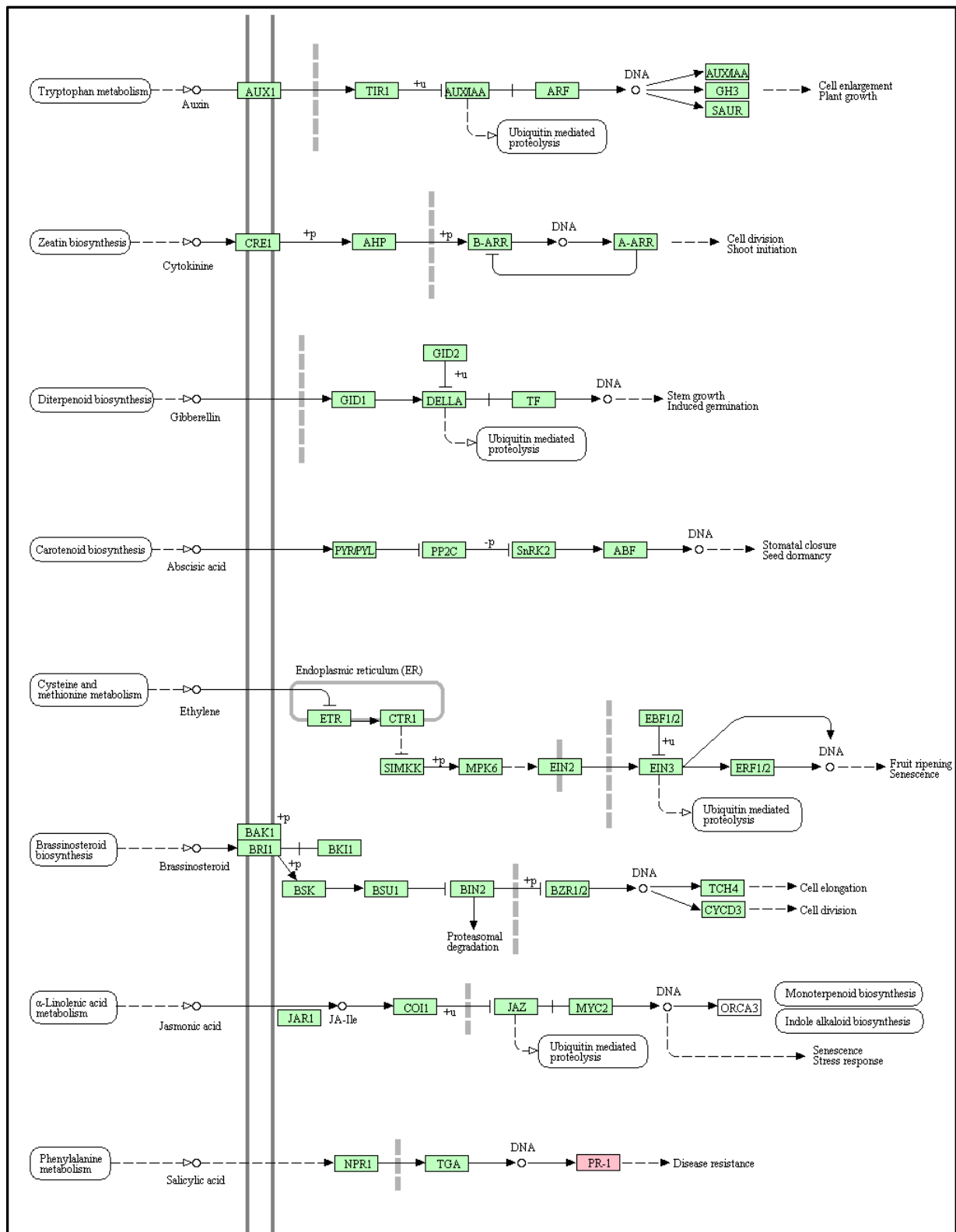


Figure A13: Localization of down-regulated plant hormone signal transduction pathway enzymes on the leaf tissues of AR23.1 post-infestation with mealybugs. Boxes highlighted in green represents genes that are present in the castor bean (*Ricinus communis*) genome (organism used for KEGG maps), while boxes highlighted in pink are the down-regulated DEGs in the leaf tissues post mealybug infestation. Boxes that are not highlighted are the genes that are not present in the castor bean genome.

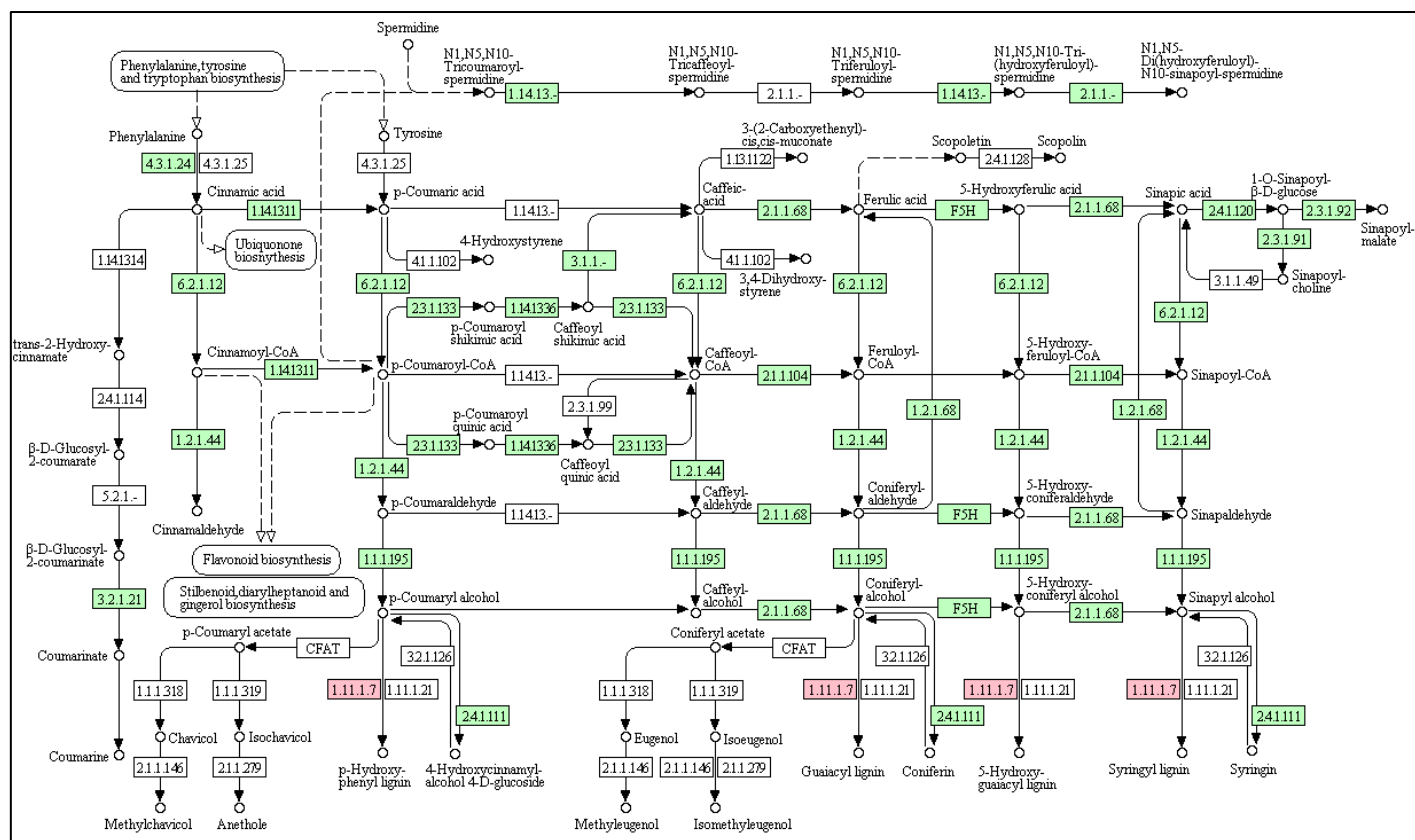


Figure A14: Localization of up-regulated phenylpropanoid biosynthesis pathway enzymes on the leaf tissues of AR23.1 post-infestation with mealybugs. Boxes highlighted in green represents genes that are present in the castor bean (*Ricinus communis*) genome (organism used for KEGG maps), while boxes highlighted in pink are the up-regulated DEGs in the leaf tissues post mealybug infestation. Boxes that are not highlighted are the genes that are not present in the castor bean genome.

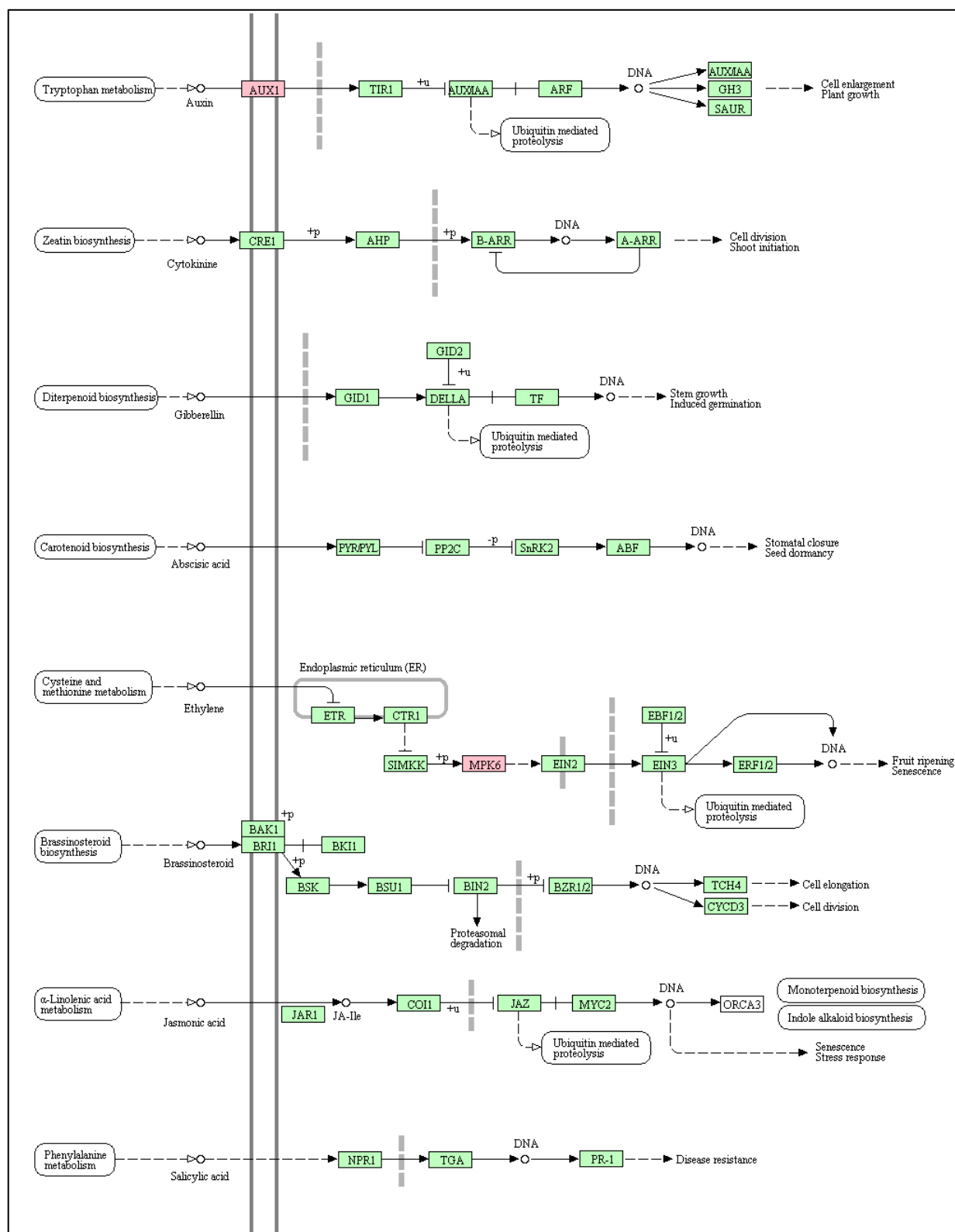


Figure A18: Localization of up-regulated plant hormone signal transduction pathway enzymes on the leaf tissues of P40/1 post-infestation with mealybugs. Boxes highlighted in green represents genes that are present in the castor bean (*Ricinus communis*) genome (organism used for KEGG maps), while boxes highlighted in pink are the up-regulated DEGs in the leaf tissues post mealybug infestation. Boxes that are not highlighted are the genes that are not present in the castor bean genome.

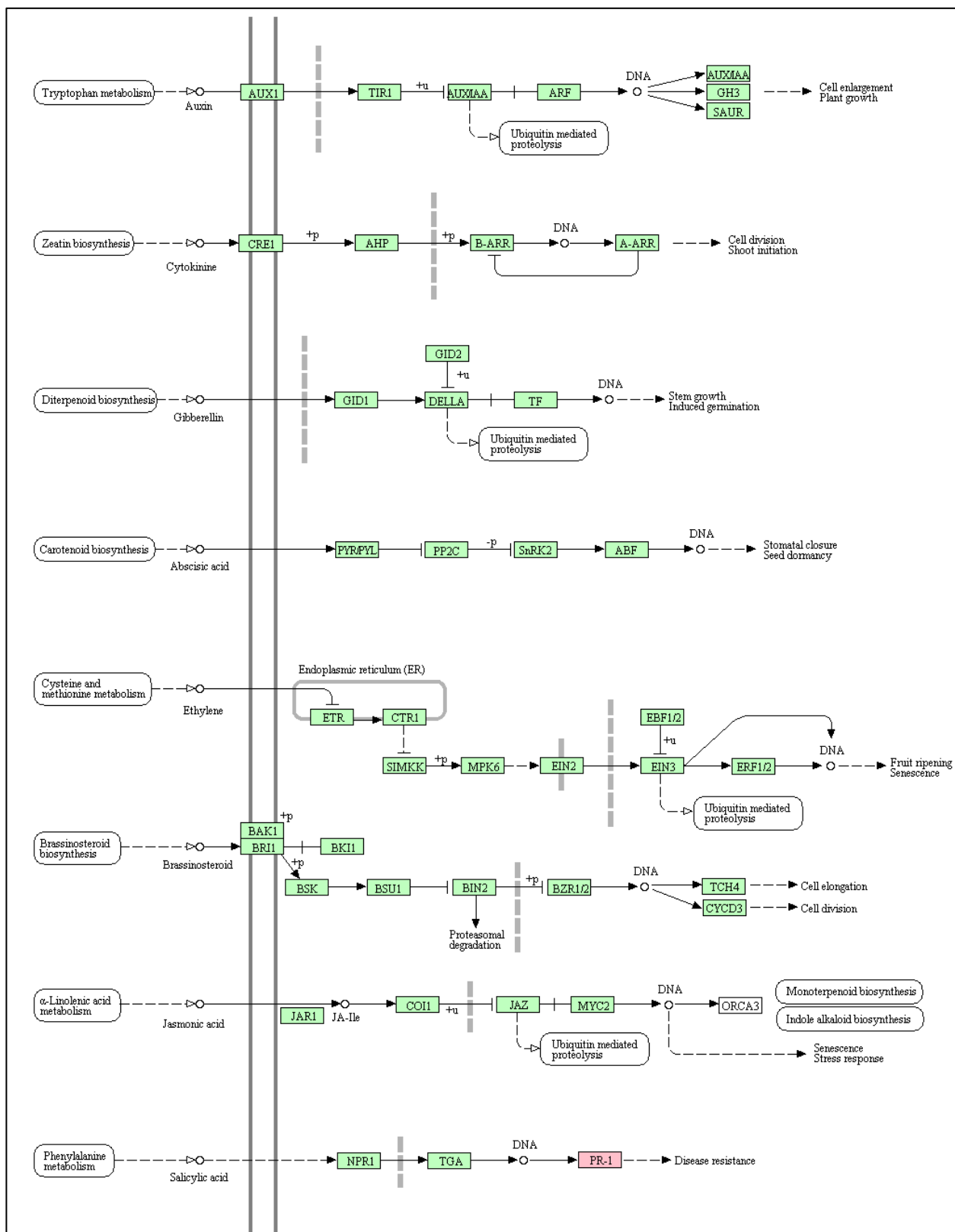


Figure A19: Localization of down-regulated plant hormone signal transduction pathway enzymes on the leaf tissues of P40/1 post-infestation with mealybugs. Boxes highlighted in green represents genes that are present in the castor bean (*Ricinus communis*) genome (organism used for KEGG maps), while boxes highlighted in pink are the down-regulated DEGs in the leaf tissues post mealybug infestation. Boxes that are not highlighted are the genes that are not present in the castor bean genome.

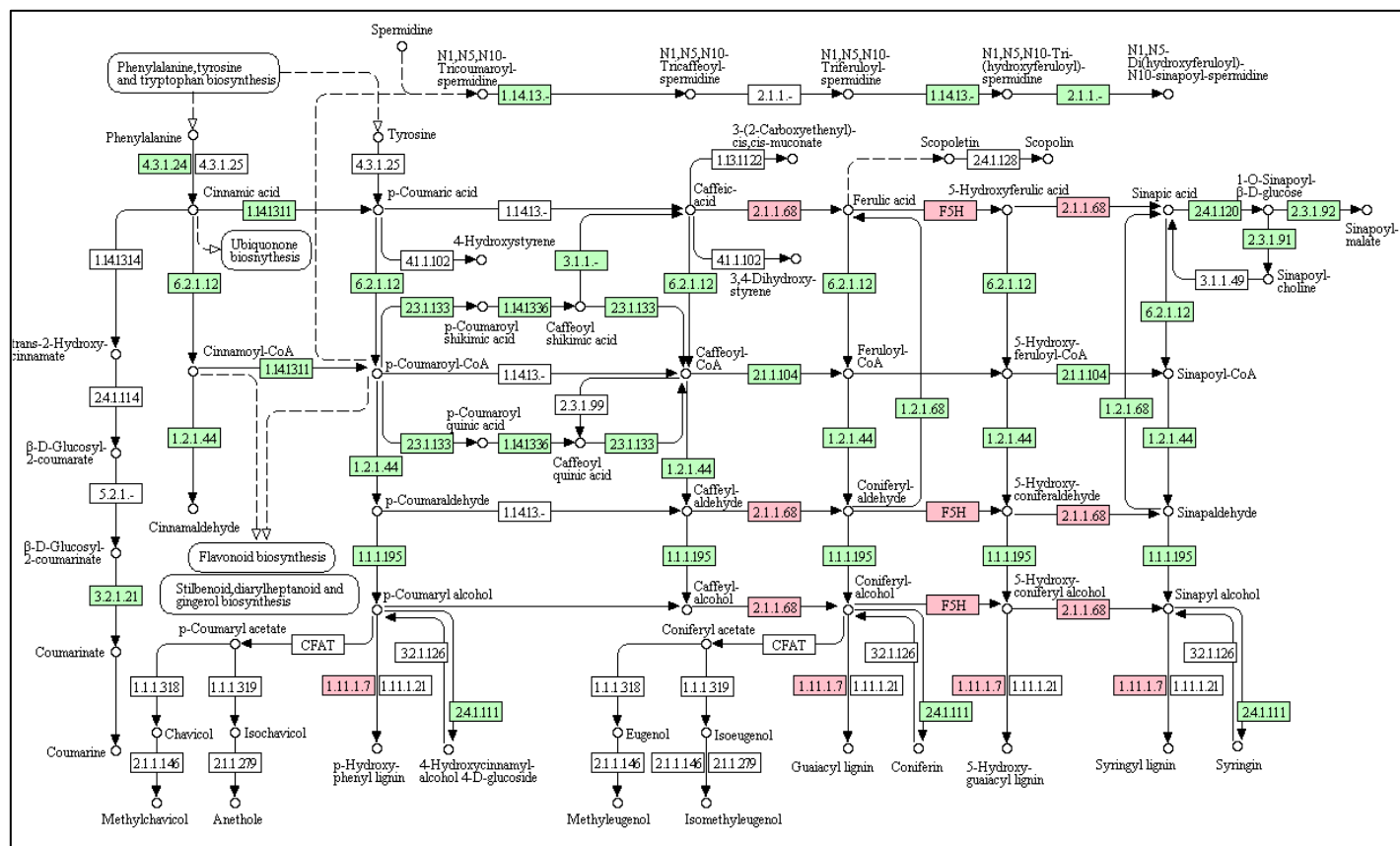


Figure A21: Localization of down-regulated phenylpropanoid biosynthesis pathway enzymes on the leaf tissues of P40/1 post-infestation with mealybugs. Boxes highlighted in green represents genes that are present in the castor bean (*Ricinus communis*) genome (organism used for KEGG maps), while boxes highlighted in pink are the down-regulated DEGs in the leaf tissues post mealybug infestation. Boxes that are not highlighted are the genes that are not present in the castor bean genome.

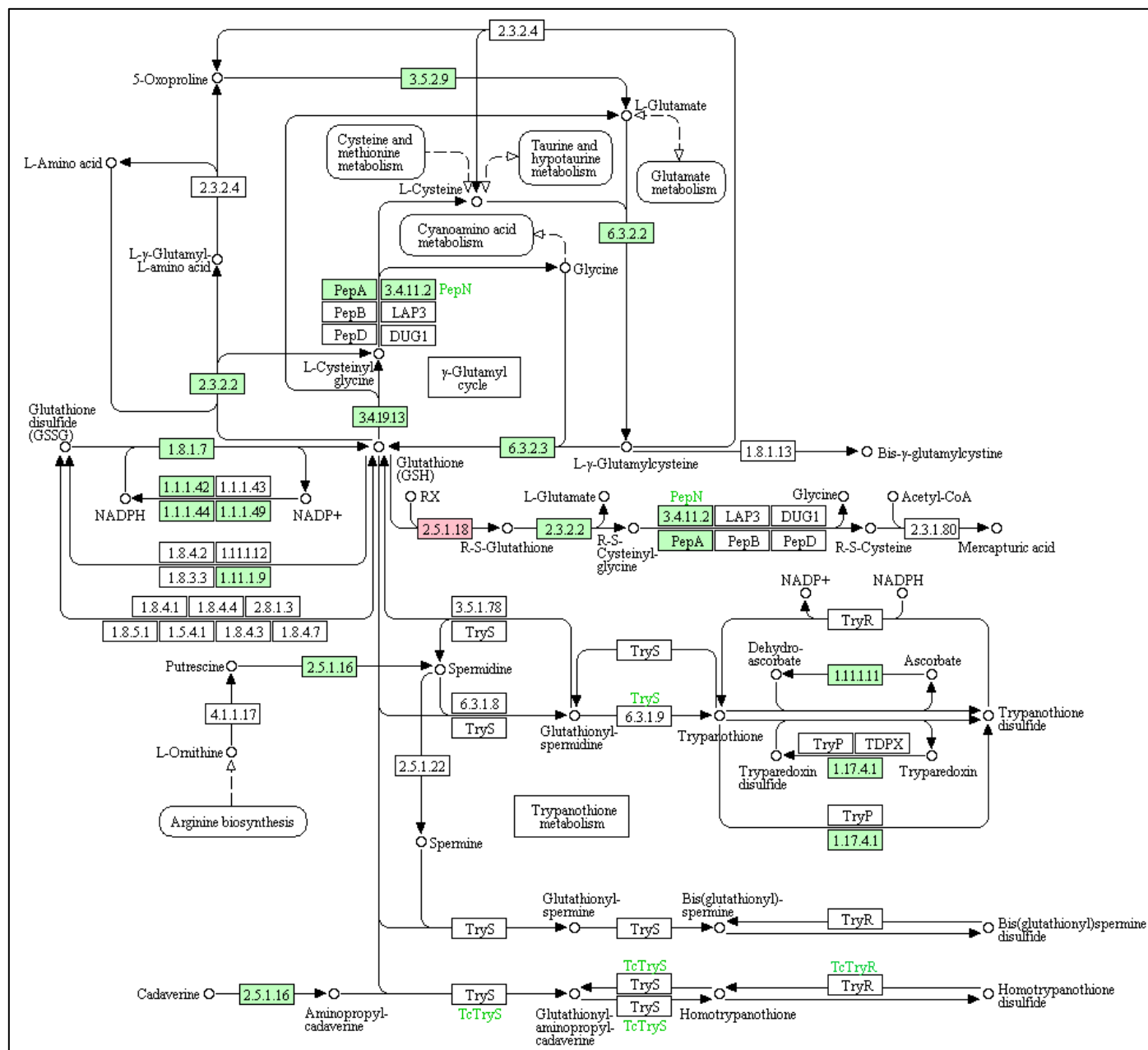


Figure A22: Localization of up-regulated glutathione metabolism pathway enzymes on the leaf tissues of P40/1 post-infestation with mealybugs. Boxes highlighted in green represents genes that are present in the castor bean (*Ricinus communis*) genome (organism used for KEGG maps), while boxes highlighted in pink are the up-regulated DEGs in the leaf tissues post mealybug infestation. Boxes that are not highlighted are the genes that are not present in the castor bean genome.

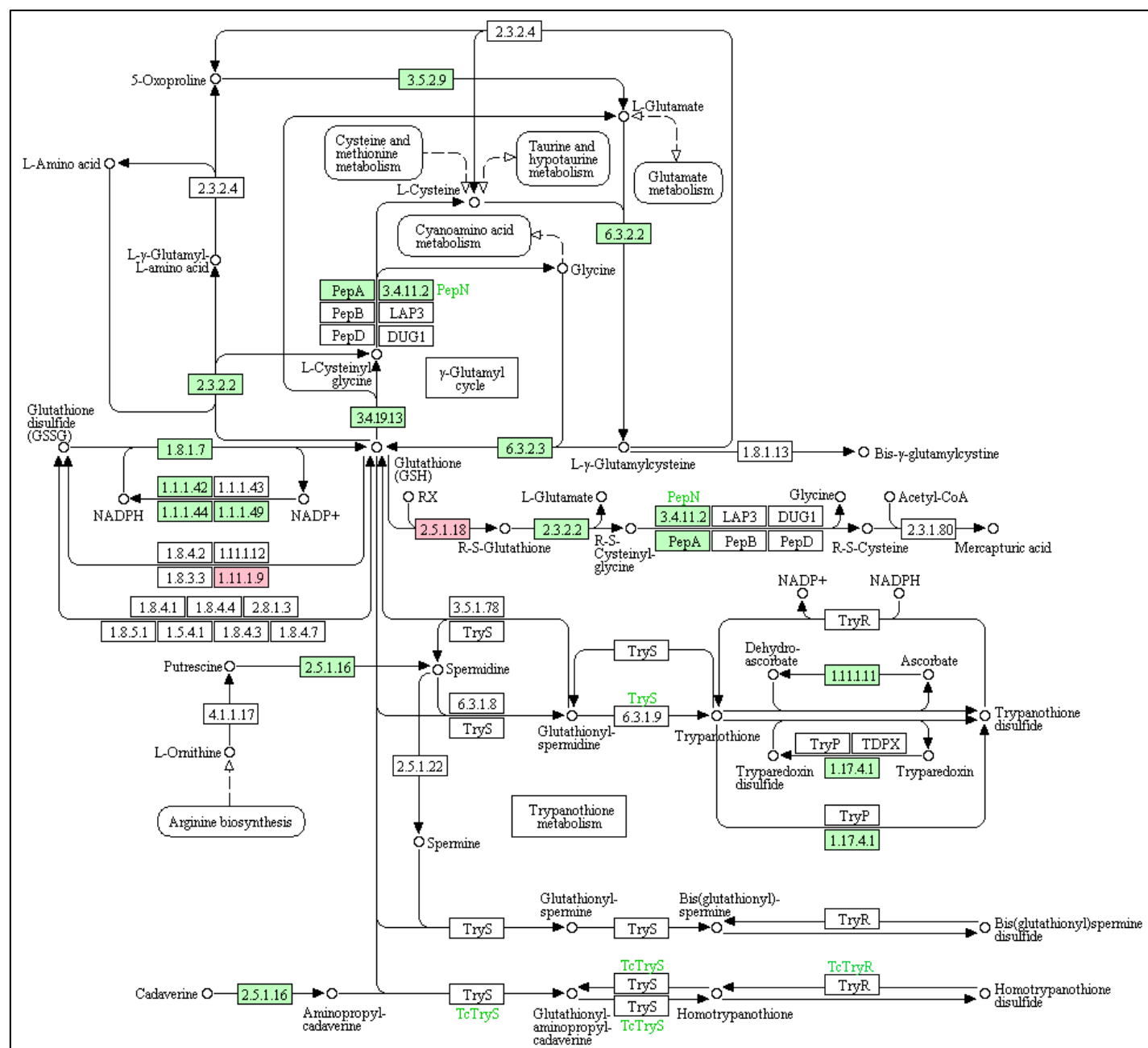


Figure A23: Localization of down-regulated glutathione metabolism pathway enzymes on the leaf tissues of P40/1 post-infestation with mealybugs. Boxes highlighted in green represents genes that are present in the castor bean (*Ricinus communis*) genome (organism used for KEGG maps), while boxes highlighted in pink are the down-regulated DEGs in the leaf tissues post mealybug infestation. Boxes that are not highlighted are the genes that are not present in the castor bean genome.

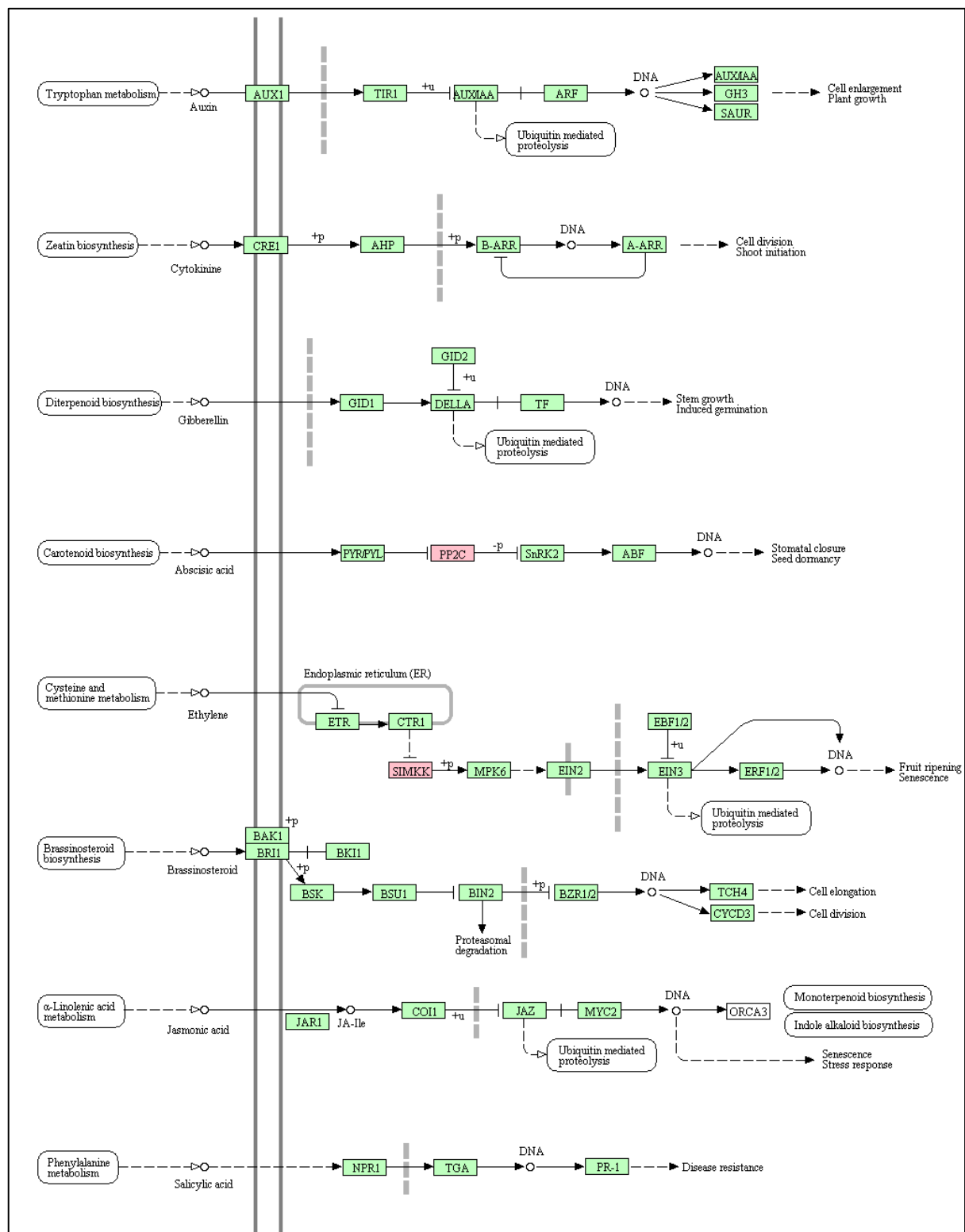


Figure A24: Localization of up-regulated plant hormone signal transduction pathway enzymes on the leaf tissues of 98/0581 post-infestation with mealybugs. Boxes highlighted in green represents genes that are present in the castor bean (*Ricinus communis*) genome (organism used for KEGG maps), while boxes highlighted in pink are the up-regulated DEGs in the leaf tissues post mealybug infestation. Boxes that are not highlighted are the genes that are not present in the castor bean genome.

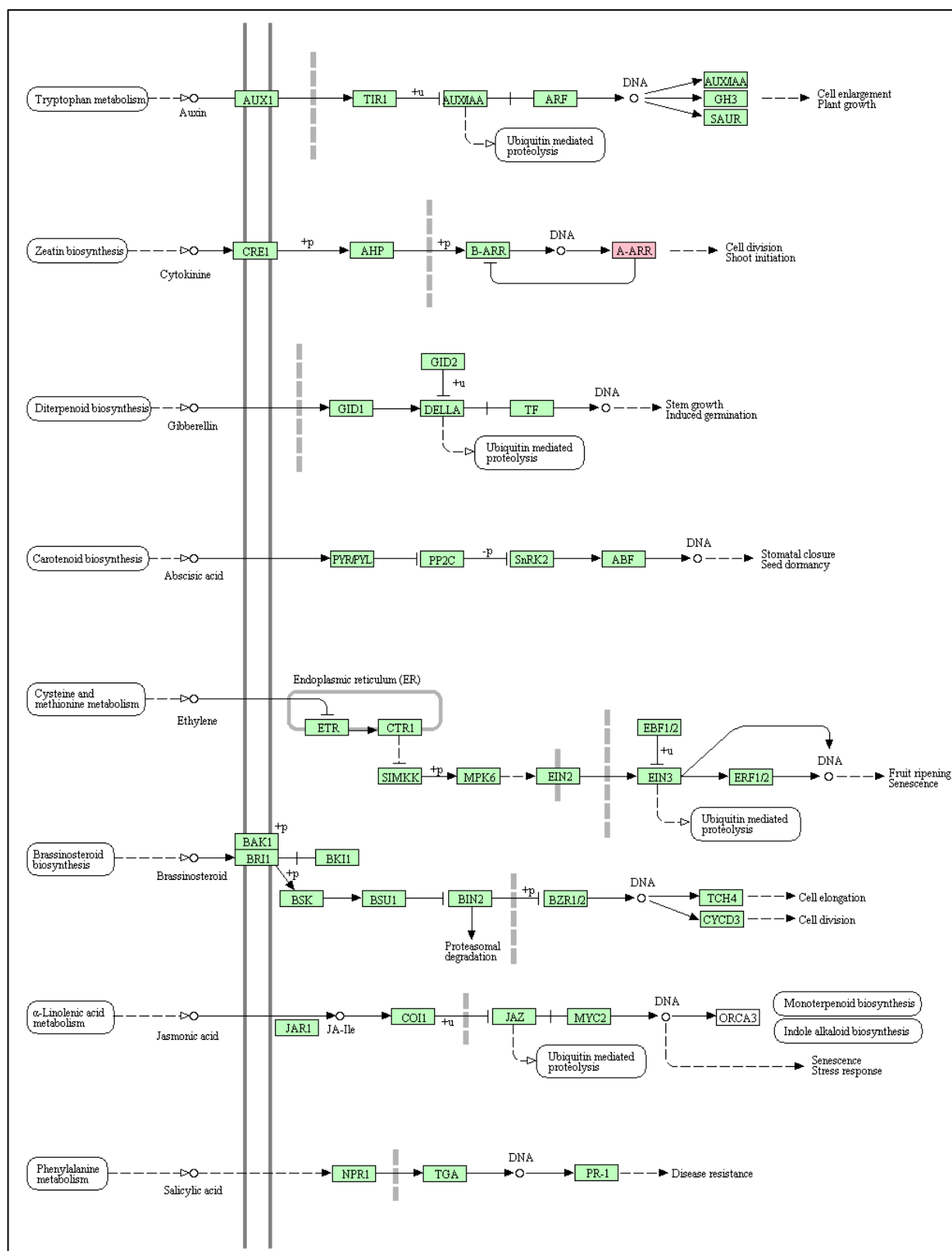


Figure A25: Localization of down-regulated plant hormone signal transduction pathway enzymes on the leaf tissues of 98/0581 post-infestation with mealybugs. Boxes highlighted in green represents genes that are present in the castor bean (*Ricinus communis*) genome (organism used for KEGG maps), while boxes highlighted in pink are the down-regulated DEGs in the leaf tissues post mealybug infestation. Boxes that are not highlighted are the genes that are not present in the castor bean genome.

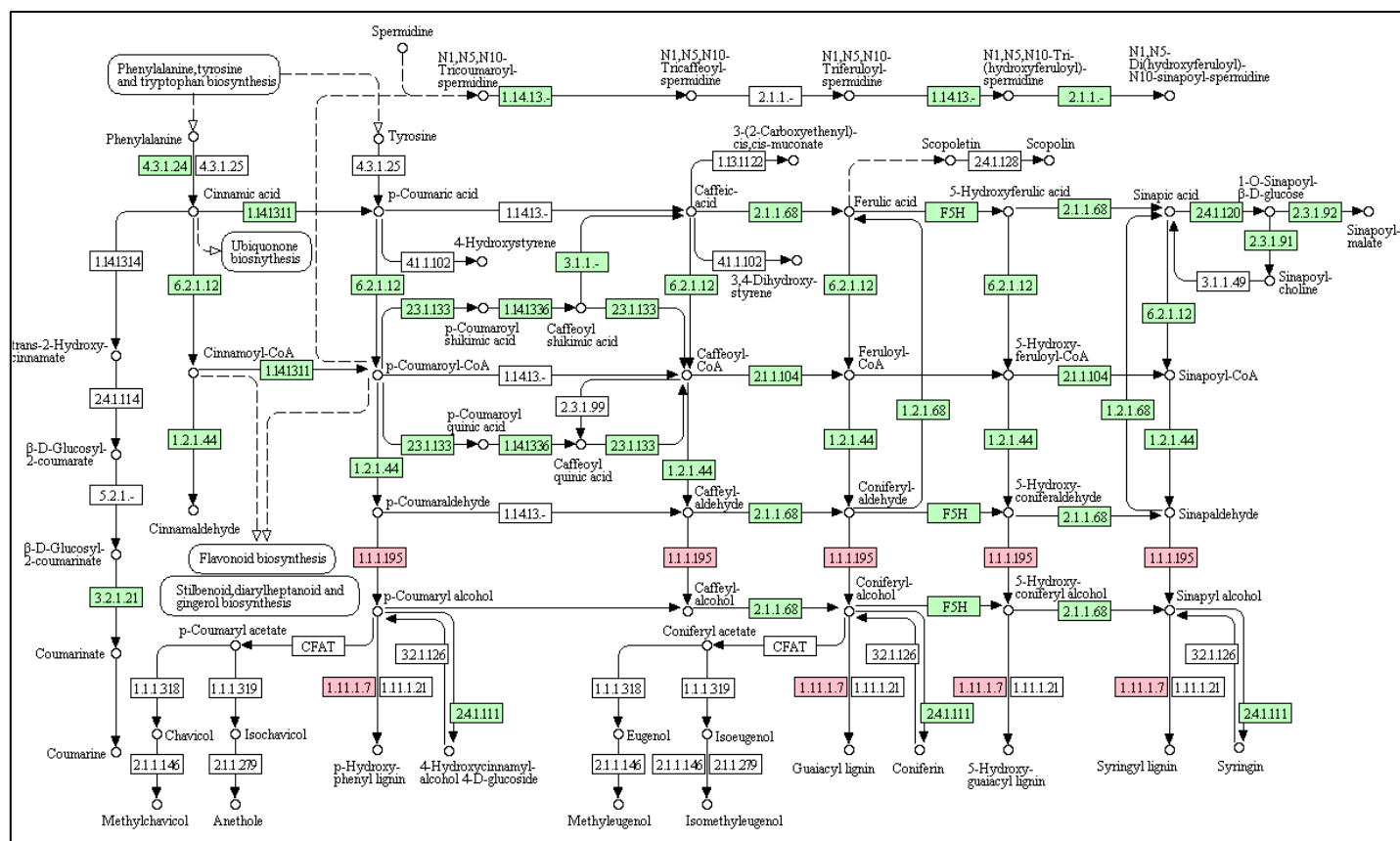


Figure A26: Localization of up-regulated phenylpropanoid biosynthesis pathway enzymes on the leaf tissues of 98/0581 post-infestation with mealybugs. Boxes highlighted in green represents genes that are present in the castor bean (*Ricinus communis*) genome (organism used for KEGG maps), while boxes highlighted in pink are the up-regulated DEGs in the leaf tissues post mealybug infestation. Boxes that are not highlighted are the genes that are not present in the castor bean genome.

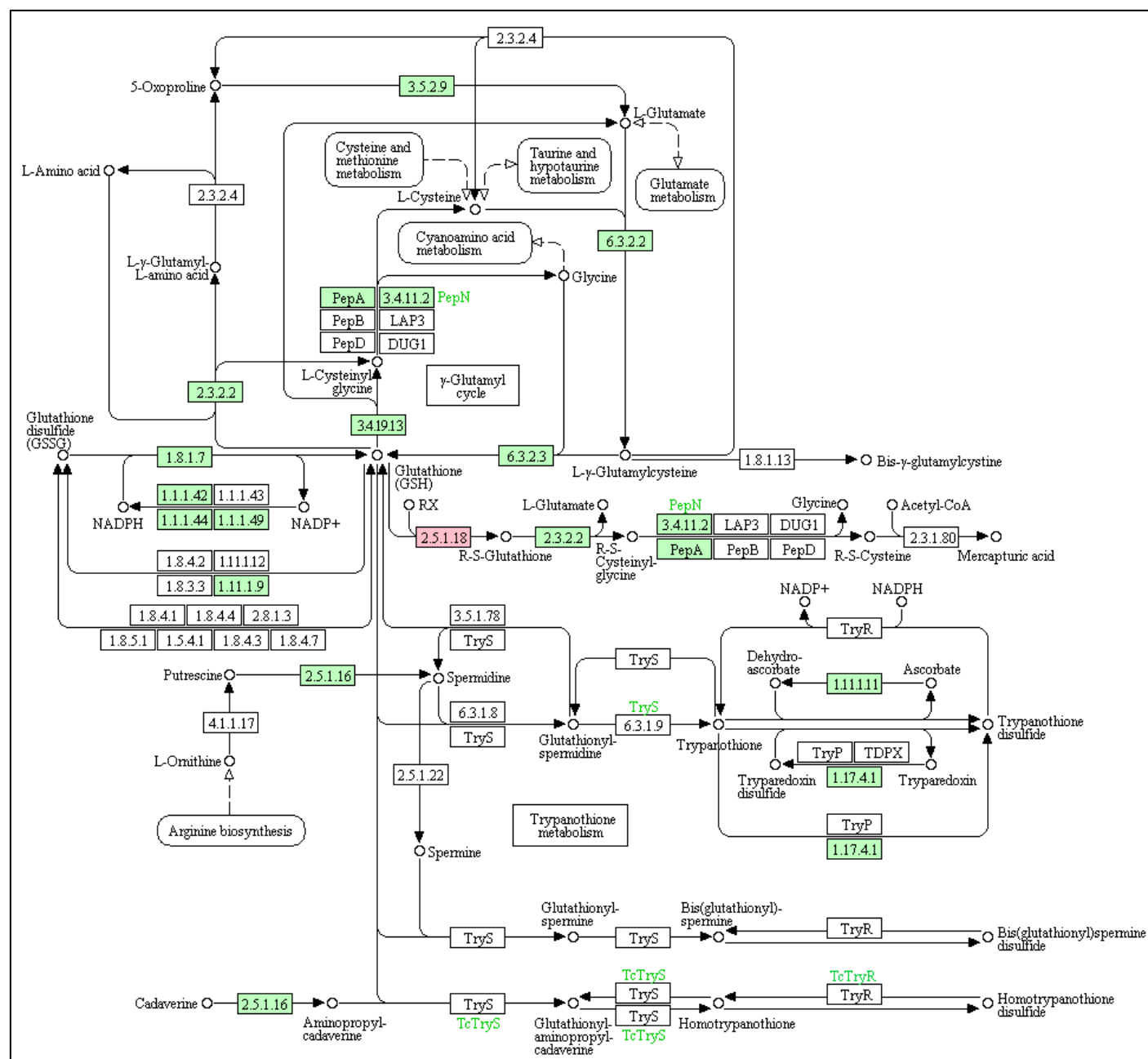


Figure A28: Localization of up-regulated glutathione metabolism pathway enzymes on the leaf tissues of 98/0581 post-infestation with mealybugs. Boxes highlighted in green represents genes that are present in the castor bean (*Ricinus communis*) genome (organism used for KEGG maps), while boxes highlighted in pink are the up-regulated DEGs in the leaf tissues post mealybug infestation. Boxes that are not highlighted are the genes that are not present in the castor bean genome.

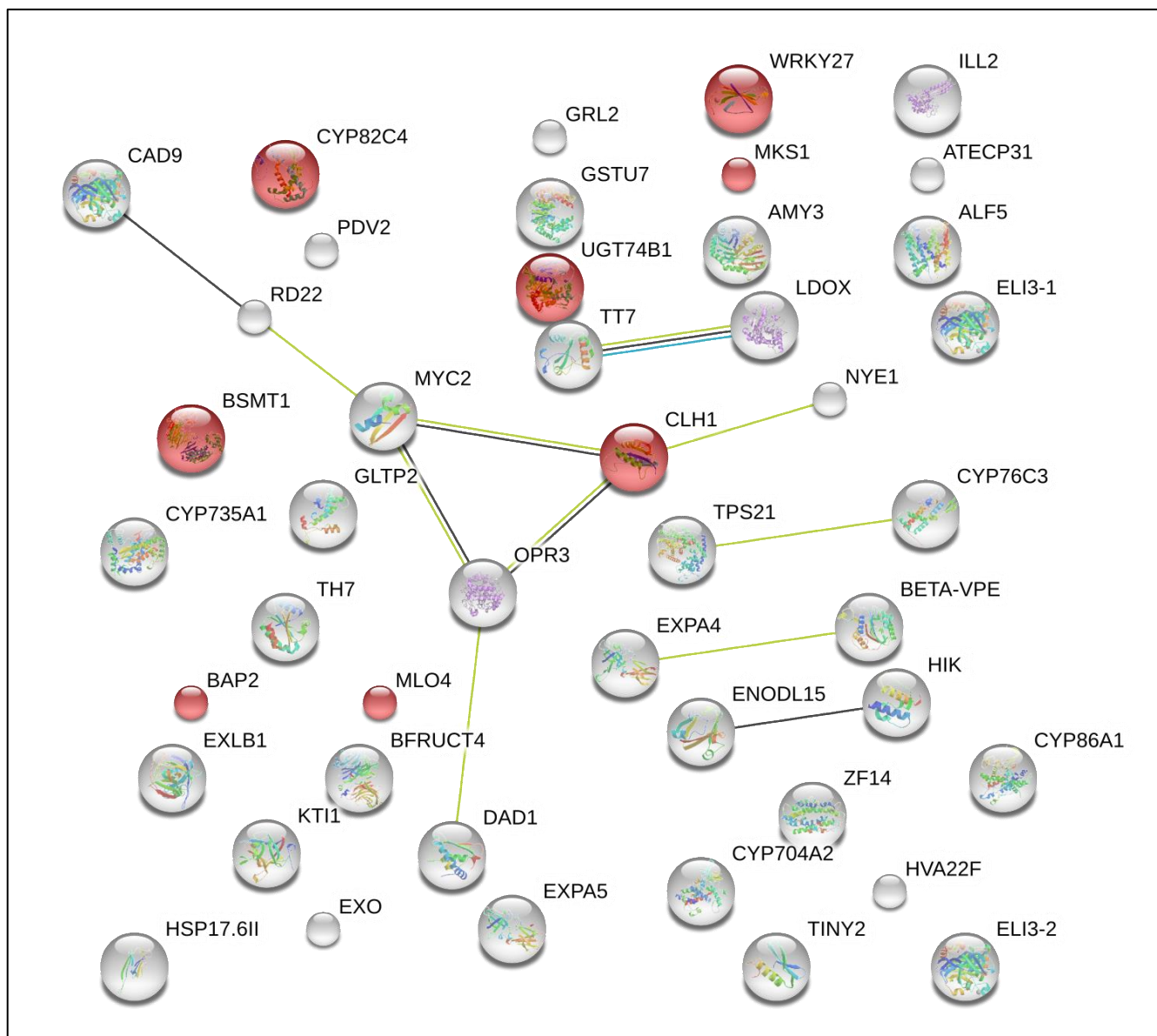


Figure A30: Network representation of the interactions of the up-regulated differentially expressed genes (DEGs) in P40/1 post-infestation with mealybugs. *Arabidopsis* genome was used to infer interactions in cassava. Different lines and colours identify types of evidence of interactions. Green line represents gene neighbourhood, red: gene fusion, light blue: protein homology, dark blue: gene occurrence, black: co-expression, and pink is the experimentally determined genes. DEGs, in the leaf tissues post mealybug infestation (highlighted in red) represent defence response genes.

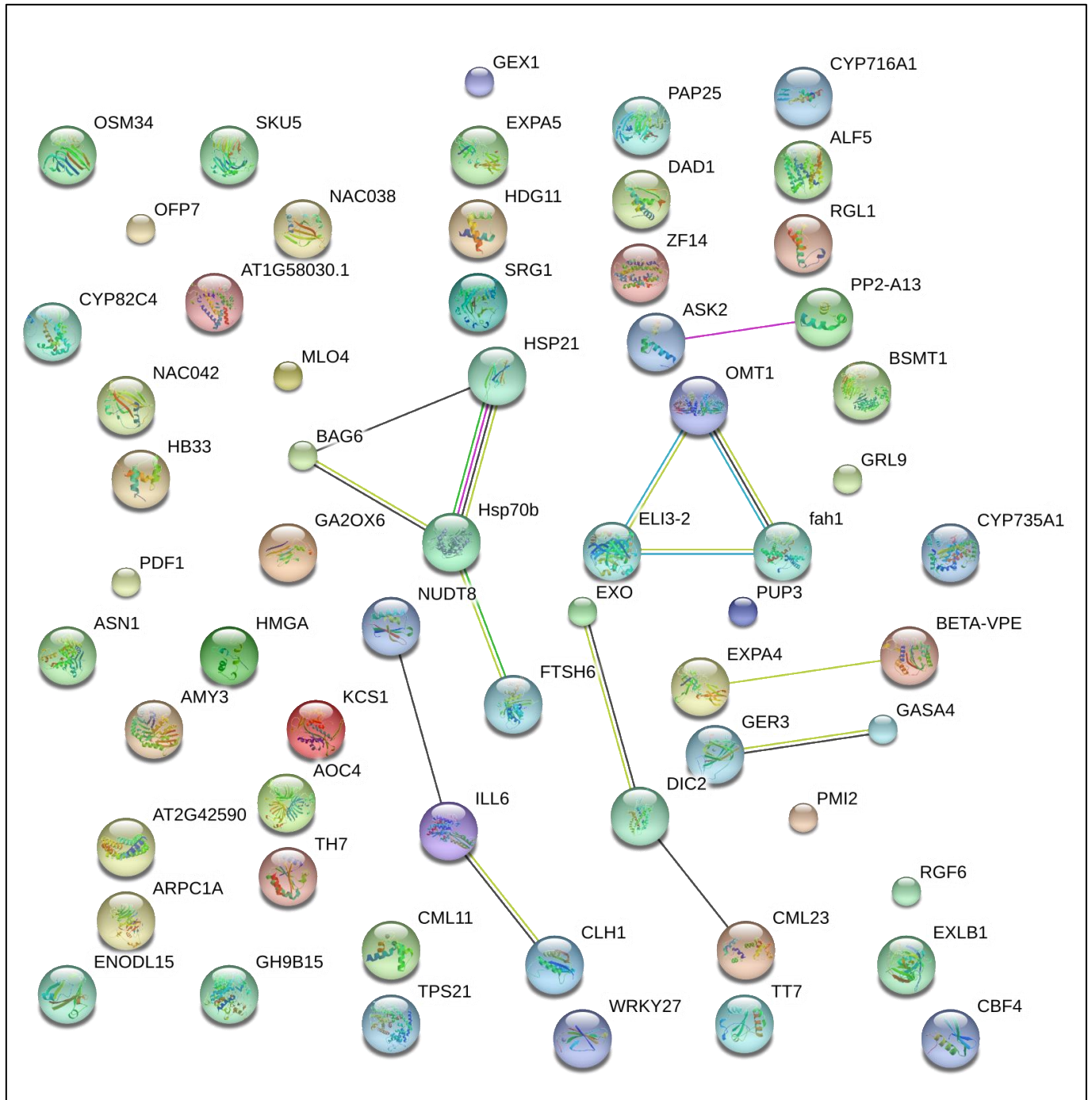


Figure A31: Network representation of the interactions of the down-regulated differentially expressed genes (DEGs) in P40/1 post-infestation with mealybugs. *Arabidopsis* genome was used to infer interactions in cassava. Different lines and colours identify types of evidence of interactions. Green line represents gene neighbourhood, red: gene fusion, light blue: protein homology, dark blue: gene occurrence, black: co-expression, and pink is the experimentally determined genes. DEGs, in the leaf tissues post mealybug infestation (highlighted in red) represent defence response genes.

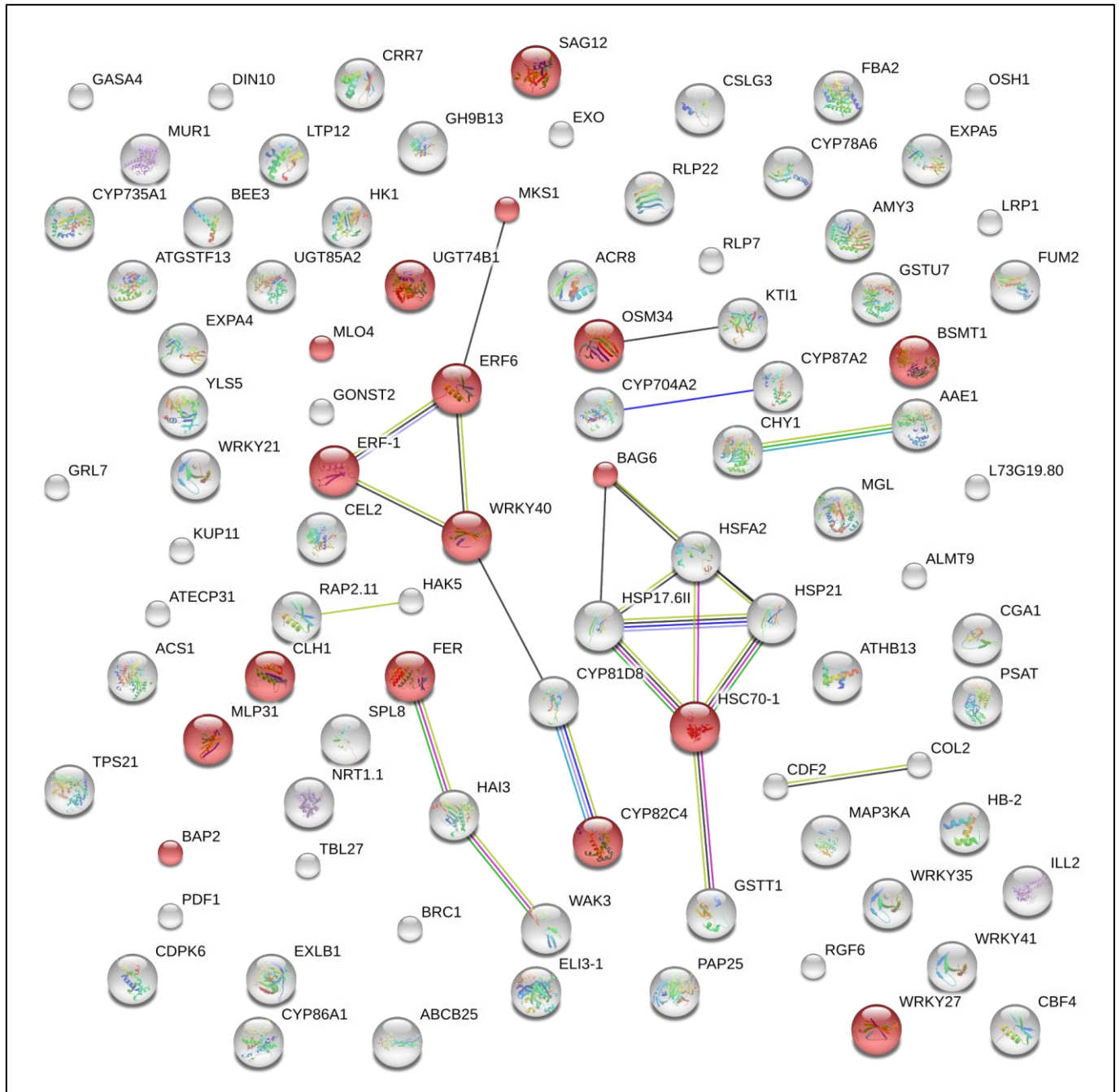


Figure A32: Network representation of the interactions of the up-regulated differentially expressed genes (DEGs) in 98/0581 post-infestation with mealybugs. *Arabidopsis* genome was used to infer interactions in cassava. Different lines and colours identify types of evidence of interactions. Green line represents gene neighbourhood, red: gene fusion, light blue: protein homology, dark blue: gene occurrence, black: co-expression, and pink is the experimentally determined genes. DEGs, in the leaf tissues post mealybug infestation (highlighted in red) represent defence response genes.

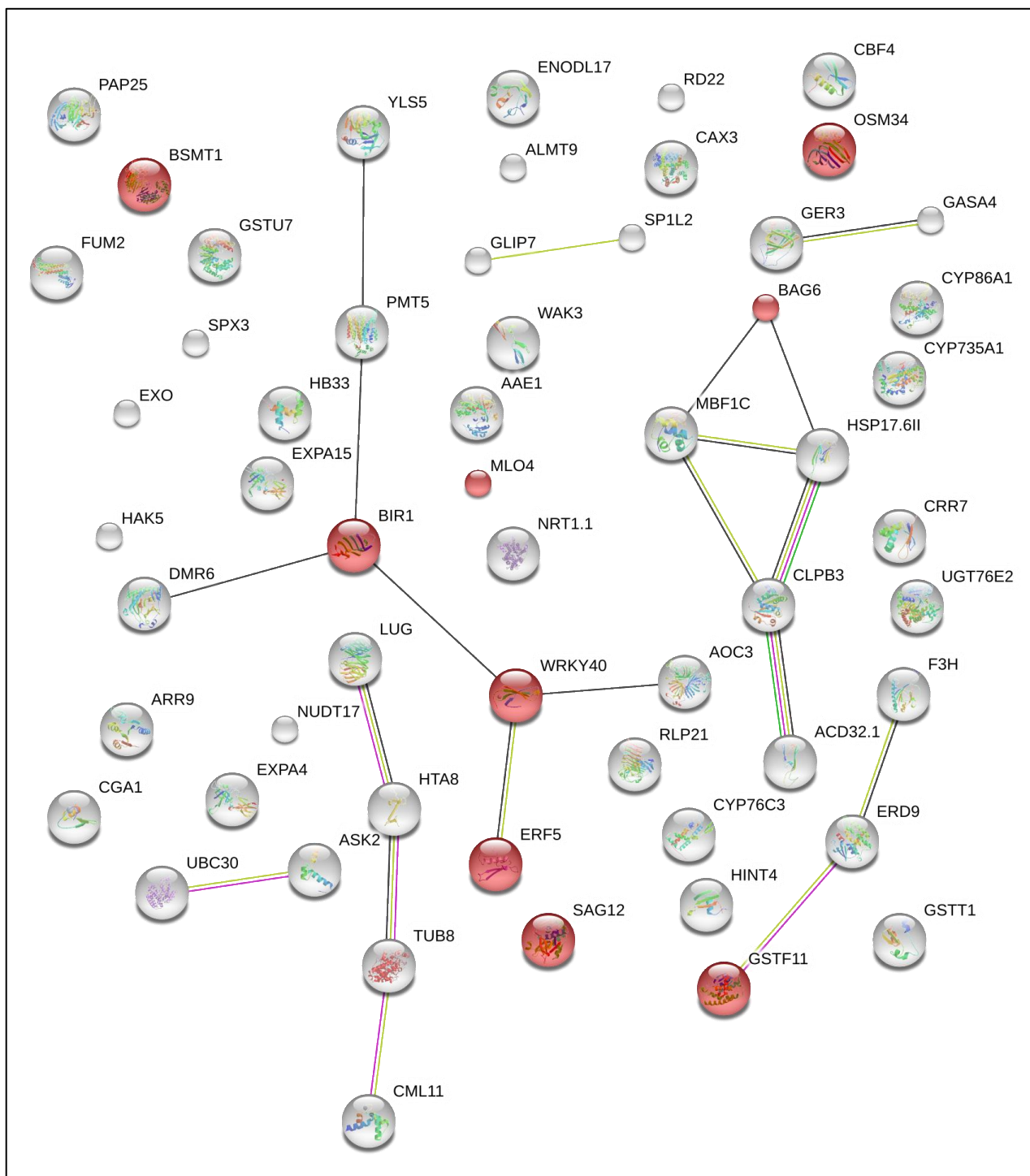


Figure A33: Network representation of the interactions of the down-regulated differentially expressed genes (DEGs) in 98/0581 post-infestation with mealybugs. *Arabidopsis* genome was used to infer interactions in cassava. Different lines and colours identify types of evidence of interactions. Green line represents gene neighbourhood, red: gene fusion, light blue: protein homology, dark blue: gene occurrence, black: co-expression, and pink is the experimentally determined genes. DEGs in the leaf tissues post mealybug infestation DEGs highlighted in red represents defence response genes.

CP data of housekeeping Genes:										
	UBQ10	GTPb	EF1	ACTIN						
n	12	12	12	12	0	0	0	0	0	0
geo Mean [CP]	31,77	24,73	30,14	35,89						
ar Mean [CP]	31,77	24,73	30,34	35,90						
min [CP]	30,60	24,25	25,12	34,82						
max [CP]	32,18	25,41	35,87	36,98						
std dev [± CP]	0,40	0,29	2,68	0,42						
CV [% CP]	1,25	1,16	8,83	1,16						
min [x-fold]	-2,25	-1,39	-32,50	-2,11						
max [x-fold]	1,33	1,60	52,99	2,12						
std dev [± x-fold]	1,32	1,22	6,49	1,33						

Pearson correlation coefficient (r)										
vs.	HKG 1	HKG 2	HKG 3	HKG 4	HKG 5	HKG 6	HKG 7	HKG 8	HKG 9	HKG 10
HKG 2	0,570	-	-	-	-	-	-	-	-	-
p-value	0,053	-	-	-	-	-	-	-	-	-
HKG 3	0,681	0,790	-	-	-	-	-	-	-	-
p-value	0,015	0,002	-	-	-	-	-	-	-	-
HKG 4	-0,460	0,260	-0,853	-	-	-	-	-	-	-
p-value	0,132	0,415	0,868	-	-	-	-	-	-	-
HKG 5					-	-	-	-	-	-
p-value					-	-	-	-	-	-
HKG 6						-	-	-	-	-
p-value						-	-	-	-	-
HKG 7							-	-	-	-
p-value							-	-	-	-
HKG 8								-	-	-
p-value								-	-	-
HKG 9									-	-
p-value									-	-
HKG 10										-
p-value										-
BestKeeper vs.	HKG 1	HKG 2	HKG 3	HKG 4	HKG 5	HKG 6	HKG 7	HKG 8	HKG 9	HKG 10
coeff. of corr. [r]	0,708	0,858	0,990	0,021						
p-value	0,019	0,001	0,001	0,946						

Regression Analysis: HKG vs. BestKeeper										
	UBQ10	GTPb	EF1	ACTIN						
vs.	HKG 1	HKG 2	HKG 3	HKG 4	HKG 5	HKG 6	HKG 7	HKG 8	HKG 9	HKG 10
vs. BK	vs. BK	vs. BK	vs. BK	vs. BK	vs. BK	vs. BK	vs. BK	vs. BK	vs. BK	vs. BK
coeff. of corr. [r]	0,71	0,86	0,99	0,02						
coeff. of det. [r ²]	0,50	0,74	0,98	0,00						
intercept [CP]	21,68	16,10	-70,23	35,55						
slope [CP]	0,33	0,28	3,31	0,01						
SE [CP]	±0,375	±0,193	±0,536	±0,598						
p-value	0,000	0,001	0,001	0,946						
Power of HKG [x-fold]	1,26	1,22	9,92	1,01						

Figure A34: Descriptive statistics of four candidate housekeeping genes (HKG) based on their crossing point (CP) values. BestKeeper index is computed together with the same descriptive parameters for four genes (UBQ10, GTPb, EF1 and Actin). Abbreviations: n: number of samples; geo Mean [CP]: the geometric mean of CP; ar Mean [CP]: the arithmetic mean of CP; Min [CP] and Max [CP]: the extreme values of CP; std dev [±CP]: the standard deviation of the CP; CV [%CP]: the coefficient of variance expressed as a percentage on the CP level; Min [x-fold] and Max [x-fold]: the extreme values of expression levels expressed as an absolute x-fold over- or under-regulation coefficient; std dev [± x-fold]: standard deviation of the absolute regulation coefficients.



Figure A35: Expression stability of the four candidate housekeeping genes (UBQ10, GTPb, EF1 and Actin) in cassava leaves. The figure illustrates the crossing points of each sample replicate for each gene.

Table A1: A list of of raw and trimmed reads generated per time point for each biological replicate and subsequently mapping to the cassava reference genome (Mesculenta_305_v6.1).

Genotype	Time	Infestation	No. of raw reads	Trimmed reads	% Trimmed reads	Mapped reads	Mapped reads/million sequence reads
AR23,1 Biological replicate 3	24 h	Non-infested	21,444,473	19,110,834	89.1	16,921,128	885 420
	72 h	Non-infested	29,388,325	25,197,198	85.7	22,067,654	875 797
	24 hpi	Infested	42,212,956	41,500,026	98.3	30,032,612	723 538
	72 hpi	Infested	45,953,442	41,507,964	90.3	36,210,712	872 379
AR23,1 Biological replicate 2	24 h	Non-infested	29,065,379	25,218,186	86.7	23,190,120	919 579
	72 h	Non-infested	25,901,916	22,107,187	85.3	19,917,821	900 965
	24 hpi	Infested	40,000,000	36,490,056	91.2	32,054,329	878 440
	72 hpi	Infested	39,115,704	37,114,004	94.9	33,790,451	910 450
AR23,1 Biological replicate 3	24 h	Non-infested	37,694,984	33,107,823	87.7	29,108,912	879 215
	72 h	Non-infested	29,614,889	25,108,197	85.0	21,109,432	840 738
	24 hpi	Infested	38,453,272	34,989,258	90.9	32,157,979	919 081
	72 hpi	Infested	31,073,358	28,957,014	93.1	23,822,215	822 675
P40/1 Biological replicate 1	24 h	Non-infested	32,536,388	28,927,448	89.0	24,978,358	863 482
	72 h	Non-infested	29,514,034	25,216,917	85.4	22,421,686	889 152
	24 hpi	Infested	29,378,808	27,764,750	93.3	22,816,550	821 781
	72 hpi	Infested	23,424,700	20,774,916	88.7	19,526,920	843 657
P40/1 Biological replicate 2	24 h	Non-infested	26,937,496	23,928,218	89.0	20,684,329	864 432
	72 h	Non-infested	26,726,775	24,210,721	90.5	20,463,976	845 244
	24 hpi	Infested	35,954,764	34, 860,410	96.9	31,986,993	917 573
	72 hpi	Infested	33,394,144	32,224,992	95.0	25,158,790	780 722
P40/1 Biological replicate 3	24 h	Non-infested	22,016,380	20,219,012	91.0	17,989,129	889 713
	72 h	Control	29,663,515	26,018,219	88.0	22,129,719	850 547
	24 hpi	Infested	33,346,560	28,740,636	86.1	24,109,731	838 872
	72 hpi	Infested	38,371,700	37,807,994	98.5	30,223,924	799 405
98/0581 Biological replicate 1	24 h	Non-infested	16,716,348	15,969,108	95.5	13,108,918	820 892
	72 h	Non-infested	17,019,479	15,102,791	88.7	13,977,198	925 471
	24 hpi	Infested	26,000,292	22,880,320	91.3	19,751,496	863 252
	72 hpi	Infested	25,796,692	20,000,141	87.3	17,854,053	892 696
98/0581 Biological replicate 2	24 h	Non-infested	17,432,730	16,107,493	92.3	14,917,893	926 146
	72 h	Non-infested	14,201,969	12,012,718	85.0	10,213,642	850 235
	24 hpi	Infested	25,498,497	24,234,738	95.8	22,406,273	937 455
	72 hpi	Infested	21,544,931	20,526,819	95.2	18,693,128	910 668
98/0581 Biological replicate 3	24 h	Non-infested	15,850,922	14,916,486	94.1	11,917,836	798 970
	72 h	Non-infested	16,729,996	14,194,653	85.0	12,917,826	910 048
	24 hpi	Infested	17,617,272	16,200,003	91.9	14,234,900	878 697
	72 hpi	Infested	21,332,252	19,141,698	89.7	16,242,109	848 519

Table A2: List of AR23.1 differentially expressed genes in cassava leaves in response to mealybug infestation at 24 and 72 hours post-infestation normalized against mock (non-infested).

Gene ID	GO ID	Gene annotation	Log2(fold_change) 24 hpi	P-value
Manes.01G059800	GO:0055114,GO:0020037,GO:016705,GO:0005506	Cytochrome P450, family 86, subfamily A, polypeptide 8	6.07	0.00005
Manes.02G151200	-	Phosphate-responsive 1 family protein	6.01	0.00005
Manes.01G043500	GO:0055114,GO:0016491	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	5.45	0.00005
Manes.17G058800	-	Ucharacterized protein	5.33	0.00005
Manes.17G070900	GO:0003824	Camelliol C synthase 1	5.25	0.00005
Manes.02G189200	GO:0016788	GDSL-like Lipase/Acylhydrolase superfamily protein	4.72	0.00005
Manes.09G175800	-	Hypothetical protein	4.54	0.0002
Manes.01G200000	GO:0009611,GO:0004867,GO:0055114,GO:0020037,GO:0016705,GO:0005506	Cytochrome P450, family 94, subfamily B, polypeptide 3	4.39	0.00005
Manes.04G050300	-	RHO GTPASE Activating protein	4.2.0	0.00005
Manes.01G020600	GO:0030983,GO:0006298,GO:0005524	MUTS homolog 2	4.18	0.00015
Manes.02G086300	GO:0016829,GO:0010333,GO:0008152,GO:0000287	Terpene synthase 21	4.17	0.00005
Manes.01G243800	GO:0015743	Aluminium activated malate transporter family protein	4.07	0.0001
Manes.15G094100	GO:0048544	D-mannose binding lectin protein with Apple-like carbohydrate-binding domain	3.93	0.00015
Manes.08G103300,Manes.08G103400	-	Hypothetical protein	3.82	0.0001
Manes.01G147500	-	Myb domain protein 4	3.79	0.00015
Manes.07G061600	GO:0009958	Phototropic-responsive NPH3 family protein	3.69	0.00015
Manes.02G124600	-	HSP20-like chaperones superfamily protein	3.69	0.00005
Manes.01G158000	GO:0055085,GO:0016021	Auxin efflux carrier family protein	3.54	0.0001
Manes.01G067000	GO:0055085,GO:0016020,GO:0015297,GO:0015238,GO:0006855	MATE efflux family protein	3.37	0.00005
Manes.17G078600	-	CLAVATA3/ESR-RELATED 16	3.35	0.0001
Manes.01G076500	GO:0004575,GO:0004564	Glycosyl hydrolases family 32 protein	3.21	0.00005
Manes.01G231100	GO:0055114,GO:0016491	Gibberellin 2-oxidase 6	3.06	0.00005
Manes.02G051100	GO:0055085,GO:0016020,GO:0006811,GO:0005216	Cyclic nucleotide-regulated ion channel family protein	2.88	0.00005
Manes.06G136100	GO:0004857	Plant invertase/pectin methylesterase inhibitor superfamily protein	2.59	0.00005
Manes.06G131100	GO:0006355,GO:0003700	Ethylene responsive element binding factor 1	2.41	0.0002
Manes.09G040500	-	Alpha/beta-Hydrolases superfamily protein	2.37	0.0001
Manes.01G274500	GO:0006629	Alpha/beta-Hydrolases superfamily protein	2.17	0.00005
Manes.11G057100	GO:0005515	Leucine-rich repeat (LRR) family protein	2.10	0.00005
Manes.14G096200	GO:0005515	RNI-like superfamily protein	2.00	0.00005
Manes.05G143800	-	Expansin 11	1.89	0.00005
Manes.03G048300	GO:0005515,GO:0006468,GO:0004672	Leucine-rich repeat transmembrane protein kinase family protein	1.68	0.00005
Manes.15G094300	GO:0048544	D-mannose binding lectin protein with Apple-like carbohydrate-binding domain	1.58	0.00005
Manes.01G142200	GO:0005515	Leucine-rich repeat receptor-like protein kinase family protein	1.58	0.00005

Manes.01G273700	GO:0043565,GO:0006355,GO:003700	WRKY family transcription factor	1.58	0.00005
Manes.05G132500	GO:0008017,GO:0007018,GO:005524,GO:0003777,GO:0005871	ATP binding microtubule motor family protein	1.37	0.00005
Manes.01G235800	-	Myb domain protein 55	1.37	0.00005
Manes.07G059300	GO:0005975,GO:0004553	Glycosyl hydrolase superfamily protein	1.35	0.0001
Manes.10G105100	GO:0016887,GO:0005524,GO:0016020	White-brown complex homolog protein 11	1.16	0.00015
Manes.03G079300	GO:0008270,GO:0005515	RING/U-box superfamily protein	0.71	0.00005
Manes.13G137200	GO:0045735	RmlC-like cupins superfamily protein	0.24	0.00005
Manes.02G026300	GO:0055114,GO:0020037,GO:006979,GO:0004601	Peroxidase superfamily protein	-0.12	0.00005
Manes.16G045900	GO:0006468,GO:0005524,GO:0004672	MAP kinase kinase 5	-1.91	0.00005
Manes.11G052600	-	Late embryogenesis abundant protein (LEA) family protein	-2.61	0.0001
Manes.01G129800	-	Alpha-crystallin domain 32.1	-2.74	0.0001
Manes.S110500	-	Hypothetical protein	-3.11	0.00005
Manes.14G051800	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Cytochrome P450, family 94, subfamily B, polypeptide 3	-3.25	0.00005
Manes.02G081800	GO:0005515	Glutathione S-transferase F11	-3.34	0.00005
Manes.02G023600	-	Protein of unknown function (DUF567)	-3.46	0.00005
Manes.02G025100	-	Osmotin 34	-3.53	0.00005
Manes.14G143500	-	Heat shock protein 70B	-3.53	0.0001
Manes.01G252200	GO:0043565,GO:0006355,GO:003700	WRKY DNA-binding protein 56	-3.57	0.00005
Manes.01G114100	-	Gamma interferon responsive lysosomal thiol (GILT) reductase family protein	-3.76	0.00005
Manes.02G005800	GO:0016810,GO:0006807	Carbon-nitrogen hydrolase family protein	-4.03	0.00005
Manes.01G187700	GO:0046983	Basic helix-loop-helix (bHLH)	-4.10	0.00015
Manes.02G008500	-	DNA-binding family protein	-4.15	0.00005
Manes.02G015400	GO:0009378,GO:0006310,GO:006281	Galactose-binding protein	-4.17	0.00005
Manes.01G110800	GO:0008270,GO:0005975,GO:004476,GO:0009298	P-loop containing nucleoside triphosphate hydrolases superfamily protein	-4.17	0.00005
Manes.01G178000	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Mannose-6-phosphate isomerase, type I	-4.19	0.00015
Manes.02G040000	GO:0016021,GO:0016020,GO:0022857	Cytochrome P450 superfamily protein	-4.19	0.00005
Manes.01G268700	-	Nodulin MtN21 /EamA-like transporter family protein	-4.46	0.00005
Manes.11G034200	GO:0016887,GO:0005524,GO:0016020	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	-4.47	0.00005
Manes.03G043600	GO:0004857	ABC-2 and Plant PDR ABC-type transporter family protein	-4.53	0.0001
Manes.01G020700	-	Plant invertase/pectin methylesterase inhibitor superfamily protein	-4.80	0.00005
Manes.01G250800	GO:0051087	TCP family transcription factor	-4.82	0.0003
Manes.02G174100	-	BCL-2-associated athanogene 6	-4.82	0.00005
Manes.03G061200	-	Nodulin MtN3 family protein	-4.90	0.00005
Manes.01G058100	-	Tetraspanin family protein	-4.95	0.00005
Manes.01G082400	GO:0030983,GO:0006298,GO:005524	Pectin lyase-like superfamily protein	-5.11	0.00005
Manes.01G013000	GO:0006468,GO:0005524,GO:004672	MUTS-like protein 4	-5.34	0.00005
Manes.01G138100	GO:0005975,GO:0004553	Wall-associated kinase family protein	-5.53	0.00005
Manes.01G161800	GO:0055085,GO:0016021	Glycosyl hydrolase superfamily protein	-5.63	0.00005
Manes.01G235200	-	Cation exchanger 3	-5.63	0.00005
Manes.03G018600	GO:0016887,GO:0005524,GO:0016020	Pectin lyase-like superfamily protein	-5.65	0.00015

Manes.03G062400	GO:0005515	Transducin/WD40 repeat-like superfamily protein	-5.65	0.00025
Manes.15G189500	-	HSP20-like chaperones superfamily protein	-5.73	0.00015
Manes.02G084300	-	Phy rapidly regulated 1	-5.74	0.00005
Manes.01G087800	-	Expansin-like B1	-5.74	0.00005
Manes.02G109900	-	Protein of unknown function (DUF630 and DUF632)	-5.86	0.00015
Manes.01G031800	GO:0008270	DHHC-type zinc finger family protein	-5.87	0.0001
Manes.02G001500	GO:0055114,GO:0016491,GO:0008270	GroES-like zinc-binding dehydrogenase family protein	-5.91	0.00005
Manes.01G070800	GO:0006355,GO:0005634,GO:005524,GO:0032502,GO:0006351	Growth-regulating factor 2	-5.95	0.00005
Manes.01G254100	GO:0016757,GO:0016020	Protein of unknown function (DUF604)	-5.97	0.00025
Manes.02G106500	GO:0016021,GO:0006952	Seven transmembrane MLO family protein	-6.03	0.00005
Manes.03G120700	GO:0045454,GO:0005681,GO:000398	mRNA splicing factor, thioredoxin-like U5 snRNP	-6.11	0.0001
Manes.01G070200	GO:0055114,GO:0016491	Leucoanthocyanidin dioxygenase	-6.11	0.00005
Manes.01G173300	GO:0006468,GO:0004672	Protein kinase superfamily protein	-6.13	0.00005
Manes.02G104600	-	GAST1 protein homolog 1	-6.34	0.0002
Manes.01G013100	-	PLAC8 family protein	-6.57	0.00005
Manes.03G066400	GO:0016021,GO:0015116,GO:0008272,GO:0055085,GO:0016020,GO:0008271,GO:1902358	Sulfate transporter 3;5	-6.58	0.0001
Manes.02G099700	GO:0016787	purple acid phosphatase 25	-6.64	0.00005
Manes.15G154400	-	HSP20-like chaperones superfamily protein	-6.80	0.00005
Manes.S053400	-	HSP20-like chaperones superfamily protein	-7.27	0.00005
Gene ID	GO ID	Gene annotation	Log2(fold_change) 72 hpi	P-value
Manes.01G200000	GO:0009611,GO:0004867,GO:0055114,GO:0020037,GO:0016705,GO:0005506	cytochrome P450, family 94, subfamily B, polypeptide 3	6.94	0.00005
Manes.03G040000	-	Mob1/phocein family protein	6.83	0.00005
Manes.02G115200	GO:0055085,GO:0016021	Auxin efflux carrier family protein	5.93	0.00005
Manes.01G043500	GO:0055114,GO:0016491	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	5.88	0.00005
Manes.02G126100	-	Glutathione S-transferase tau 7	5.69	0.00005
Manes.01G264600	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Cytochrome P450, family 704, subfamily A, polypeptide 2	5.68	0.00005
Manes.13G043400	-	Glycine-rich protein 5-like	5.51	0.00005
Manes.02G208500	GO:0006355,GO:0003700	Integrase-type DNA-binding superfamily protein	5.47	0.00005
Manes.01G070800	GO:0006355,GO:0005634,GO:005524,GO:0032502,GO:0006351	Growth-regulating factor 2	5.08	0.00005
Manes.02G032300	GO:0055085,GO:0016020,GO:0015297,GO:0015238,GO:0006855	MATE efflux family protein	5.03	0.00005
Manes.06G136300	GO:0004857	Plant invertase/pectin methylesterase inhibitor superfamily protein	5.02	0.00005
Manes.01G189000	GO:0043565,GO:0006355,GO:0003700	WRKY DNA-binding protein 72	5.00	0.00005
Manes.03G056100	GO:0006355,GO:0003700	Integrase-type DNA-binding superfamily protein	4.97	0.00005
Manes.02G105300	-	Pathogenesis-related thaumatin superfamily protein	4.80	0.00005
Manes.02G002200	-	Alpha/beta-Hydrolases superfamily protein	4.56	0.00005
Manes.17G058800	-	Uncharacterized protein	4.55	0.0001
Manes.02G145900	GO:0006508,GO:0004190	Eukaryotic aspartyl protease family protein	4.54	0.00005
Manes.10G146700	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Cytochrome P450, family 705, subfamily A, polypeptide 27	4.37	0.0001

Manes.02G086300	GO:0016829,GO:0010333,GO:008152,GO:0000287	Terpene synthase 21	4.19	0.00005
Manes.01G231100	GO:0055114,GO:0016491	Gibberellin 2-oxidase 6	3.81	0.00005
Manes.01G076500	GO:0004575,GO:0004564	Glycosyl hydrolases family 32 protein	3.79	0.00005
Manes.02G086200	GO:0016829,GO:0010333,GO:008152,GO:0000287	Terpenoid cyclases/Protein prenyltransferases superfamily protein	3.77	0.00005
Manes.02G150900	-	Phosphate-responsive 1 family protein	3.62	0.00005
Manes.08G103300,Manes.08G103400	-	Hypothetical protein	3.53	0.00005
Manes.01G129500	GO:0005515,GO:0006468,GO:005524,GO:0004672	Leucine-rich repeat protein kinase family protein	3.32	0.00005
Manes.01G231200	GO:0042545,GO:0030599,GO:005618	Pectin lyase-like superfamily protein	3.30	0.00005
Manes.S053400	-	HSP20-like chaperones superfamily protein	3.30	0.00005
Manes.02G189200	GO:0016788	GDSL-like Lipase/Acylhydrolase superfamily protein	3.01	0.00005
Manes.06G154000	GO:0016887,GO:0005524,GO:0016020	ABC-2 type transporter family protein	2.95	0.00005
Manes.02G124600	-	HSP20-like chaperones superfamily protein	2.83	0.00005
Manes.01G274500	GO:0006629	Alpha/beta-Hydrolases superfamily protein	2.79	0.00015
Manes.15G130200	GO:0009607,GO:0006952	Polyketide cyclase/dehydrase and lipid transport superfamily protein	2.71	0.0001
Manes.15G015600	GO:0045454,GO:0015035,GO:009055	Glutaredoxin family protein	2.58	0.00005
Manes.01G147500	-	Myb domain protein 4	2.44	0.00005
Manes.11G080400	-	dsRNA-binding protein 5	2.31	0.0001
Manes.01G126300	GO:0004866	Kunitz trypsin inhibitor 1	2.15	0.00005
Manes.15G130300	GO:0009607,GO:0006952	MLP-like protein 43	1.87	0.0002
Manes.04G120100	GO:0055114,GO:0016491	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	1.81	0.00005
Manes.01G188800	-	Hypothetical protein	1.39	0.00005
Manes.15G015500	GO:0045454,GO:0015035,GO:009055	Thioredoxin superfamily protein	0.86	0.00005
Manes.03G012600	GO:0016798,GO:0016020	glucuronidase 3	0.41	0.00005
Manes.02G040700	-	Protein of unknown function, DUF547	0.39	0.00005
Manes.10G105100	GO:0016887,GO:0005524,GO:0016020	White-brown complex homolog protein 11	-0.30	0.00005
Manes.07G059300	GO:0005975,GO:0004553	Glycosyl hydrolase superfamily protein	-0.61	0.0001
Manes.10G106000	-	homeobox protein 31	-0.84	0.00025
Manes.01G041100	GO:0005515	Disease resistance family protein / LRR family protein	-0.85	0.00005
Manes.14G096200	GO:0005515	RNI-like superfamily protein	-1.22	0.00005
Manes.05G143800	-	expansin 11	-1.30	0.0001
Manes.03G138800	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Cytochrome P450, family 71, subfamily B, polypeptide 34	-1.36	0.0001
Manes.02G124400	-	Protein of unknown function (DUF3527)	-1.54	0.0001
Manes.14G051800	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Cytochrome P450, family 94, subfamily B, polypeptide 3	-1.60	0.00005
Manes.07G061600	GO:0009958	Phototropic-responsive NPH3 family protein	-3.01	0.00005
Manes.01G108600	GO:0016020	Uncharacterized protein family (UPF0016)	-3.23	0.00015
Manes.15G148400	-	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	-3.27	0.0001
Manes.02G040000	GO:0016021,GO:0016020,GO:0022857	Nodulin MtN21 /EamA-like transporter family protein	-3.31	0.00015
Manes.02G076300	GO:0006412,GO:0005840,GO:005622,GO:0003735	Nucleic acid-binding, OB-fold-like protein	-3.35	0.00005
Manes.17G113800	-	RmlC-like cupins superfamily protein	-3.46	0.00005
Manes.01G086900	GO:0055085,GO:0016021	Major facilitator superfamily protein	-3.51	0.00005

Manes.02G205700	GO:0007165,GO:0005515,GO:0043531	Disease resistance protein (TIR-NBS-LRR class), putative	-3.57	0.00005
Manes.02G039500	GO:0042545,GO:0030599,GO:0005618	Pectin lyase-like superfamily protein	-3.60	0.0001
Manes.01G040000	GO:0005975,GO:0004553	O-Glycosyl hydrolases family 17 protein	-3.82	0.0001
Manes.01G202400	-	Class I glutamine amidotransferase-like superfamily protein	-4.12	0.00005
Manes.02G017000	GO:0005975,GO:0004553	Glycosyl hydrolase 9B15	-4.30	0.00005
Manes.03G021100	GO:0055085,GO:0022857,GO:0016021	Polyol/monosaccharide transporter 5	-4.30	0.00005
Manes.01G142200	GO:0005515	Leucine-rich repeat receptor-like protein kinase family protein	-4.37	0.00005
Manes.02G170200	GO:0005509,GO:0006468,GO:0005524,GO:0004672	Wall associated kinase 3	-4.51	0.00005
Manes.02G054300	GO:0016758,GO:0008152	UDP-Glycosyltransferase superfamily protein	-4.52	0.00005
Manes.02G169700	GO:0016021	Nodulin MtN3 family protein	-4.57	0.00045
Manes.01G065500	GO:0055114,GO:0020037,GO:0006979,GO:0004601	Peroxidase superfamily protein	-4.68	0.0004
Manes.16G016100,Manes.16G016200,Manes.16G016300	GO:0016829,GO:0010333,GO:0008152,GO:0000287	Terpene synthase 14	-4.83	0.00005
Manes.03G027900	-	Oxidative stress 3	-4.85	0.0001
Manes.03G097700	GO:0006468,GO:0005524,GO:0004672	MAP kinase kinase 5	-4.86	0.00005
Manes.02G044700	-	Expansin-like B1	-4.9	0.00005
Manes.02G108300	GO:0008234,GO:0006508	Senescence-associated gene 12	-4.92	0.00005
Manes.03G061200	-	Tetraspanin family protein	-4.94	0.00025
Manes.02G025800	GO:0055114,GO:0020037,GO:0006979,GO:0004601	Peroxidase superfamily protein	-4.94	0.00005
Manes.01G002800	GO:0005975,GO:0004553	Cellulase 2	-4.96	0.0002
Manes.03G032500	GO:0055114,GO:0016491	ACC oxidase 1	-5	0.00005
Manes.01G203500	-	Calmodulin-like 11	-5.11	0.00005
Manes.01G047000	-	Alpha/beta-Hydrolases superfamily protein	-5.17	0.00005
Manes.02G071200	GO:0008152,GO:0003824	Cytochrome P450, family 77, subfamily A, polypeptide 5	-5.17	0.00005
Manes.18G024300	-	pseudogene	-5.41	0.00005
Manes.03G062400	GO:0005515	RAD-like 6	-5.46	0.00005
Manes.02G067700	-	Transducin/WD40 repeat-like superfamily protein	-5.47	0.00005
Manes.01G067400	GO:0006355,GO:0003700	Regulator of Vps4 activity in the MVB pathway protein	-5.89	0.00005
Manes.01G002600	-	C-repeat-binding factor 4	-5.92	0.00005
Manes.01G078200	-	Alpha/beta-Hydrolases superfamily protein	-6.41	0.0001
Manes.02G151300	-	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	-6.70	0.00025
Manes.03G063600	GO:0006355,GO:0005634,GO:0003677	Phosphate-responsive 1 family protein	-6.76	0.00025
Manes.02G089000	GO:0006355,GO:0003677	S1FA-like DNA-binding protein	-6.79	0.0001
Manes.01G070200	GO:0055114,GO:0016491	NAC domain containing protein 42	-6.82	0.00025
Manes.02G025400	-	Leucoanthocyanidin dioxygenase	-6.92	0.00005
Manes.01G087800	-	Uncharacterised protein family (UPF0497)	-6.95	0.00005
Manes.02G099700	GO:0016787	Expansin-like B1	-6.96	0.0002
Manes.01G067500	GO:0006355,GO:0003700	purple acid phosphatase 25	-6.96	0.00005
Manes.01G178000	GO:0055114,GO:0020037,GO:0016705,GO:0005506	C-repeat-binding factor 4	-6.99	0.00005
Manes.02G148300	GO:0016021,GO:0005215	Cytochrome P450 superfamily protein	-7.09	0.0001
Manes.02G025100	-	purine permease 3	-7.51	0.00005
Manes.02G015400	GO:0009378,GO:0006310,GO:0006281	osmotin 34	-7.94	0.0001
		P-loop containing nucleoside triphosphate hydrolases superfamily protein		
Gene ID	GO ID	Gene annotation	Log2(fold_change) 24 hpi vs 72 hpi	
Manes.01G178000	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Cytochrome P450 superfamily protein	6.81	0.00005

Manes.02G099700	GO:0016787	Purple acid phosphatase 25	6.69	0.00005
Manes.02G151000	-	Phosphate-responsive 1 family protein	6.54	0.00005
Manes.01G236500	-	Germin-like protein 10	5.79	0.00005
Manes.02G150900	-	Phosphate-responsive 1 family protein	5.75	0.00005
Manes.02G170600	GO:0006468,GO:0004672	Wall associated kinase 3	5.75	0.00005
Manes.01G121300	GO:0016567,GO:0004842	ARM repeat superfamily protein	5.70	0.00005
Manes.01G235200	-	Pectin lyase-like superfamily protein	5.60	0.0001
Manes.03G044100	GO:0006355,GO:0003700	Integrase-type DNA-binding superfamily protein	5.26	0.00005
Manes.01G047200	GO:0043565,GO:0006355,GO:003700	WRKY family transcription factor	4.86	0.00005
Manes.02G075400	-	Calmodulin like 23	4.64	0.00005
Manes.02G150200	GO:0006468,GO:0005524,GO:004672,GO:0005515	Leucine-rich repeat receptor-like protein kinase family protein	4.57	0.00005
Manes.02G130000	GO:0008270,GO:0005515	Zinc finger (C3HC4-type RING finger) family protein	4.51	0.00005
Manes.01G085400	GO:0006355,GO:0003700	Ethylene responsive element binding factor 1	4.14	0.00005
Manes.01G260700	GO:0005515	BON association protein 2	4.08	0.0001
Manes.S053400	GO:0006529,GO:0004066	HSP20-like chaperones superfamily protein	4.01	0.0001
Manes.02G146000	-	Dicarboxylate carrier 2	3.92	0.0001
Manes.12G152600	-	MAP kinase substrate 1	3.67	0.00005
Manes.14G006300	GO:0005515	CCT motif family protein	3.22	0.00005
Manes.16G132600	GO:0006508,GO:0004190	Eukaryotic aspartyl protease family protein	2.95	0.0001
Manes.15G121600	GO:0055114,GO:0016491	Gibberellin 2-oxidase	2.95	0.00005
Manes.02G110100	GO:0008270,GO:0005515	RING/U-box superfamily protein	2.87	0.00005
Manes.15G122400	-	Nine-cis-epoxycarotenoid dioxygenase 5	2.81	0.00005
Manes.15G035200	GO:0007165	CBL-interacting protein kinase 7	2.72	0.00005
Manes.13G112500	GO:0008270,GO:0006355,GO:005634,GO:0004402,GO:0003712,GO:0005515	BTB and TAZ domain protein 4	2.48	0.00005
Manes.01G192100	GO:0045454,GO:0015035,GO:006662	Thioredoxin H-type 7	2.47	0.00005
Manes.01G267100	-	Vegetative cell wall protein gp1-like	2.41	0.00005
Manes.02G039000	GO:0006508,GO:0004190	Eukaryotic aspartyl protease family protein	2.38	0.00005
Manes.05G061300	GO:0006950	Late embryogenesis abundant 3 (LEA3) family protein	2.28	0.00005
Manes.01G138500,Manes.01G138600	GO:0008168	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein	2.18	0.00005
Manes.02G092000	GO:0008270,GO:0005634	DWNN domain, a CCHC-type zinc finger	1.77	0.00005
Manes.01G123500	GO:0046872	Regulator of chromosome condensation (RCC1) family with FYVE zinc finger domain	1.23	0.0001
Manes.01G241000	GO:0016747,GO:0008610,GO:016020,GO:0006633	3-ketoacyl-CoA synthase 1	-1.61	0.0001
Manes.04G140500	-	Glycosyl hydrolases family 32 protein	0.23	0.00005
Manes.06G004300	GO:0016887,GO:0005524,GO:016020	ABC-2 type transporter family protein	-0.41	0.00005
Manes.15G149600	-	Myb domain protein 24	-1.15	0.00005
Manes.15G060200	-	Putative uncharacterized protein DDB_G0290521	-1.36	0.0001
Manes.01G089900	-	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	-1.38	0.00005
Manes.01G189000	GO:0043565,GO:0006355,GO:003700	WRKY DNA-binding protein 72	-1.91	0.0001
Manes.15G015600	GO:0045454,GO:0015035,GO:009055	Glutaredoxin family protein	-1.98	0.0001
Manes.04G121600	-	Galactose oxidase/kelch repeat superfamily protein	-2.10	0.0001
Manes.09G062300	GO:0016020,GO:0006810,GO:005215	Gamma tonoplast intrinsic protein	-2.10	0.00005

Manes.06G007600	-	Calcium-binding EF-hand family protein	-2.23	0.00005
Manes.04G013300	GO:0008270,GO:0005515	CHY-type/CTCHY-type/RING-type Zinc finger protein	-2.29	0.00005
Manes.01G233000	GO:0006508	tolB protein-related	-2.29	0.0001
Manes.05G096400	-	PAS/LOV protein B	-2.33	0.00005
Manes.10G071700	GO:0055114,GO:0016491,GO:0051536,GO:0020037	Nitrite reductase 1	-2.33	0.00005
Manes.18G036200	GO:0005509	Calcium-binding EF-hand family protein	-2.35	0.0001
Manes.04G094700	GO:0071805,GO:0016020,GO:0015079	K ⁺ uptake transporter 3	-2.37	0.00005
Manes.15G085500	-	Aluminium induced protein with YGL and LRDR motifs	-2.40	0.00005
Manes.08G067800	-	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	-2.47	0.0001
Manes.11G067500	GO:0003824,GO:0006470,GO:0004722	Heat shock protein 70B	-2.53	0.00005
Manes.02G128000	GO:0045454,GO:0015035,GO:0009055	Highly ABA-induced PP2C gene 3	-2.54	0.00005
Manes.15G015500	GO:0006529,GO:0004066	Thioredoxin superfamily protein	-2.61	0.0001
Manes.02G212800	GO:0008081,GO:0006629	Glutamine-dependent asparagine synthase 1	-2.61	0.00005
Manes.04G089800	-	Senescence-related gene 3	-2.62	0.0001
Manes.11G067600	-	Heat shock protein 70B	-2.63	0.00005
Manes.11G068500	-	Chaperone DnaJ-domain superfamily protein	-2.67	0.00005
Manes.01G119000	-	HSP20-like chaperones superfamily protein	-2.67	0.00005
Manes.09G112800	-	RNI-like superfamily protein	-2.71	0.00005
Manes.17G076700	GO:0016829,GO:0010333,GO:0008152,GO:0000287	Chaperone DnaJ-domain superfamily protein	-2.85	0.0001
Manes.04G052400	GO:0004866	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family	-2.94	0.00005
Manes.S099100	-	Terpene synthase-like sequence-1,8-cineole	-2.99	0.00005
Manes.01G126300	-	Kunitz trypsin inhibitor 1	-3.04	0.00005
Manes.07G123200	GO:0020037,GO:0017004	HSP20-like chaperones superfamily protein	-3.11	0.00005
Manes.02G051000	-	Cyclic nucleotide-regulated ion channel family protein	-3.15	0.00005
Manes.07G086000	-	Cytochrome C assembly protein	-3.16	0.0001
Manes.10G105900	GO:0008080	HSP20-like chaperones superfamily protein	-3.28	0.00005
Manes.08G103000	-	GRAM domain family protein	-3.30	0.00005
Manes.03G040200	GO:0004869	Acyl-CoA N-acyltransferases (NAT) superfamily protein	-3.30	0.0002
Manes.15G028300	-	heat-shock protein 70T-2	-3.32	0.00015
Manes.09G039400	GO:0043565,GO:0006355,GO:0003700	Cystatin/monellin superfamily protein	-3.42	0.00005
Manes.16G048700	GO:0055114,GO:0020037,GO:0006979,GO:0004601	HSP20-like chaperones superfamily protein	-3.46	0.0001
Manes.02G061900	GO:0070569,GO:0008152	Basic leucine-zipper 42	-3.48	0.00005
Manes.08G002700	-	Ascorbate peroxidase 2	-3.52	0.0001
Manes.02G169300,Manes.02G169400	GO:0006468,GO:0005524,GO:0004672	UDP-sugar pyrophosphorylase	-3.61	0.0001
Manes.13G028800	-	enoyl-CoA hydratase/isomerase D	-3.62	0.00005
Manes.03G076800	-	Protein serine/threonine kinases;protein kinases;ATP binding;sugar binding;kinases;carbohydrate binding	-3.8	0.00005
Manes.02G087900	-	Predicted AT-hook DNA-binding family protein	-4.00	0.00005
Manes.08G019800	-	Programmed cell death 4	-4.07	0.00005
Manes.02G212800	GO:0043565,GO:0006355,GO:0003700,GO:0003677	Glutamine-dependent asparagine synthase 1	-4.2	0.0001
Manes.13G124700	-	Uncharacterized protein	-4.23	0.00005
Manes.01G242300	-	Homeobox 7	-4.25	0.0001

Manes.07G031300	GO:0055114,GO:0020037,GO:016705,GO:0005506	HSP20-like chaperones superfamily protein	-4.28	0.0001
Manes.01G264600	GO:0008270,GO:0004089	Cytochrome P450, family 704, subfamily A, polypeptide 2	-4.47	0.00005
Manes.02G071200	GO:0045156,GO:0019684,GO:009772	Cytochrome P450, family 77, subfamily A, polypeptide 5 pseudogene	-4.64	0.00005
Manes.17G116800	GO:0005524,GO:0006508,GO:004222	Carbonic anhydrase 1	-4.65	0.00015
Manes.S113100	GO:0016798,GO:0016020	photosystem II reaction center protein A	-4.73	0.00005
Manes.02G046700	GO:0016984,GO:0015977,GO:000287	FTSH protease 6	-4.83	0.00005
Manes.03G012600	GO:0016829,GO:0010333,GO:000287	Glucuronidase 3	-4.91	0.00005
Manes.S113700	GO:0016829,GO:0010333,GO:000287	Ribulose-bisphosphate carboxylases	-4.94	0.00005
Manes.02G086100	GO:0055114,GO:0020037,GO:016705,GO:0005506	Terpene synthase 21	-5.08	0.00005
Manes.08G138500,Manes.08G138600	GO:0016829,GO:0010333,GO:008152,GO:0000287	Terpene synthase 21	-5.18	0.00005
Manes.01G109800	GO:0006468,GO:0005524,GO:004672	Cytochrome P450, family 735, subfamily A, polypeptide 1	-5.20	0.00005
Manes.02G086300	-	Terpene synthase 21	-5.24	0.00005
Manes.02G172500	GO:0016829,GO:0010333,GO:008152,GO:0000287	Protein kinase protein with adenine nucleotide alpha hydrolases-like domain	-5.30	0.00005
Manes.02G025400	GO:0016829,GO:0010333,GO:008152,GO:0000287	Uncharacterised protein family (UPF0497)	-5.36	0.00005
Manes.12G076600	-	Terpene synthase 10	-5.38	0.00005
Manes.02G086200	GO:0055114,GO:0016491,GO:006633,GO:0005506	Terpenoid cyclases/Protein prenyltransferases superfamily protein	-5.39	0.00015
Manes.02G221500	-	Protein of unknown function (DUF1442)	-5.42	0.0001
Manes.01G274900	-	Fatty acid hydroxylase superfamily	-5.46	0.00005
Manes.01G002600	GO:0055114,GO:0020037,GO:006979,GO:0004601	Alpha/beta-Hydrolases superfamily protein	-5.52	0.00005
Manes.02G079400	-	Gamete expressed protein 1	-5.57	0.0001
Manes.02G026300	-	Peroxidase superfamily protein	-5.73	0.00005
Manes.01G095400	-	HXXXD-type acyl-transferase family protein	-5.81	0.00005
Manes.02G044700	GO:0033177,GO:0015991,GO:0015078	Expansin-like B1	-5.98	0.00005
Manes.02G067700	-	Regulator of Vps4 activity in the MVB pathway protein	-6.34	0.00005
Manes.S091000	-	ATP synthase subunit C family protein	-7.16	0.0001
Manes.03G027900	-	Oxidative stress 3	-7.27	0.00005
Manes.02G124700	-	HSP20-like chaperones superfamily protein	-7.45	0.00005

Table A3: List of P40/1 differentially expressed genes in cassava leaves in response to mealybug infestation at 24 and 72 hours post-infestation normalized against mock (non-infested)

Gene ID	GO ID	Gene annotation	Log2(fold_change) 24 hpi	P-value
Manes.01G138100	GO:0005975,GO:0004553	Glycosyl hydrolase superfamily protein	5.53	0.00005
Manes.02G131900	-	Expansin A4	5.31	0.00005
Manes.03G048300	GO:0005515,GO:0006468,GO:004672	Leucine-rich repeat transmembrane protein kinase family protein	4.92	0.00005
Manes.01G264600	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Cytochrome P450, family 704, subfamily A, polypeptide 2	4.82	0.0001
Manes.03G043600	GO:0004857	Plant invertase/pectin methylesterase inhibitor superfamily protein	4.53	0.00005
Manes.01G212600	GO:0007165	Rho GTPase activating protein with PAK-box/P21-Rho-binding domain	4.38	0.00005

Manes.01G014700	GO:0008017,GO:0007018,GO:005524,GO:0003777,GO:0005871	ATP binding microtubule motor family protein	4.30	0.00005
Manes.03G059300	GO:0006468,GO:0005524,GO:004672,GO:0005515	Leucine-rich receptor-like protein kinase family protein	4.26	0.00005
Manes.05G012900	GO:0016746,GO:0008152	Glycerol-3-phosphate acyltransferase 3	3.88	0.00005
Manes.15G148400	-	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	3.77	0.00005
Manes.01G189100	-	Cystathionine beta-synthase (CBS) protein	3.61	0.00015
Manes.15G130000	GO:0009607,GO:0006952	MLP-like protein 28	3.29	0.00005
Manes.02G174100	-	Nodulin MtN3 family protein	3.21	0.00005
Manes.02G104600	-	GAST1 protein homolog 1	3.11	0.0001
Manes.01G241000	GO:0016747,GO:0008610,GO:016020,GO:0006633	3-ketoacyl-CoA synthase 1	2.85	0.00005
Manes.02G145900	GO:0006508,GO:0004190	Eukaryotic aspartyl protease family protein	2.80	0.00005
Manes.17G070900	GO:0003824	Camelliol C synthase 1	2.70	0.00015
Manes.05G143800	-	Expansin 11	2.17	0.00005
Manes.13G143700	GO:0046872	Zinc finger C-x8-C-x5-C-x3-H type family protein	2.16	0.00005
Manes.04G155800	GO:0008270,GO:0005515	RING/U-box superfamily protein	2.14	0.00005
Manes.07G059300	GO:0005975,GO:0004553	Glycosyl hydrolase superfamily protein	1.91	0.0002
Manes.12G031700	GO:0006355,GO:0003700	Ethylene response factor 1	1.62	0.00015
Manes.12G103400	GO:0046983	Basic helix-loop-helix (bHLH) DNA-binding superfamily protein	1.57	0.00005
Manes.10G105100	GO:0016887,GO:0005524,GO:016020	White-brown complex homolog protein 11	1.55	0.00005
Manes.06G121800	-	hydroxysteroid dehydrogenase 6	1.43	0.0001
Manes.12G137800	GO:0008654,GO:0006021,GO:004512	Myo-inositol-1-phosphate synthase 3	1.36	0.00005
Manes.02G017100	GO:0043565,GO:0006355,GO:0003700	WRKY DNA-binding protein 27	1.35	0.00005
Manes.02G189200	GO:0016788	GDSL-like Lipase/Acylhydrolase superfamily protein	1.30	0.00005
Manes.03G181100	-	probable membrane-associated kinase regulator 1	1.10	0.00005
Manes.06G121800	-	Hydroxysteroid dehydrogenase 6	1.07	0.0001
Manes.16G011800	-	Homeodomain-like superfamily protein	0.84	0.00005
Manes.01G018900	GO:0006468,GO:0004672	BAK1-interacting receptor-like kinase 1	-0.29	0.00005
Manes.05G167500	GO:0055114,GO:0020037,GO:016705,GO:0005506	Cytochrome P450, family 82, subfamily C, polypeptide 2	-0.66	0.0002
Manes.02G203800	-	Hypothetical protein	-1.02	0.00005
Manes.02G040000	GO:0016021,GO:0016020,GO:0022857	Nodulin MtN21 /EamA-like transporter family protein	-1.12	0.00005
Manes.06G149700	-	Uncharacterized protein At3g49055	-1.14	0.00005
Manes.01G196500	GO:0016758,GO:0008152	UDP-glucosyl transferase 74B1	-1.2	0.0001
Manes.12G109400	-	RING/U-box superfamily protein	-1.61	0.00015
Manes.17G079800	GO:0009607,GO:0006952	MLP-like protein 28	-1.72	0.00005
Manes.06G156700	GO:0055085,GO:0022857,GO:016021	Phosphate transporter 1;3	-1.88	0.00005
Manes.18G115900,		O-acyltransferase (WSD1-like) family protein	-2.19	0.00005
Manes.18G116000	GO:0045017,GO:0004144	protein	-2.31	0.0001
Manes.03G165000	-	Early nodulin-75-like	-2.32	0.00005
Manes.03G061200	-	Tetraspanin family protein	-2.68	0.00005
Manes.01G231100	GO:0055114,GO:0016491	Gibberellin 2-oxidase 6	-2.89	0.00005
Manes.08G154900	GO:0046983,GO:0008171	O-methyltransferase 1	-3.41	0.00005
Manes.10G105900	-	HSP20-like chaperones superfamily protein	-3.49	0.00005
Manes.02G099700	GO:0016787	Purple acid phosphatase 25	-3.63	0.00005
Manes.12G028900	GO:0003677	Basic helix-loop-helix (bHLH) DNA-binding superfamily protein	-3.93	0.00035
Manes.02G025100	-	Osmotin 34	-3.97	0.00015
Manes.07G040100	-	Hydroxyproline-rich glycoprotein family protein	-4.36	0.00005
Manes.01G212700	-	Uncharacterized protein		

Manes.11G105200	GO:0016788	GDSL-like Lipase/Acylhydrolase superfamily protein	-4.52	0.00005
Manes.02G166800	GO:0005634	Cyclin A1;1	-4.56	0.00005
Manes.18G052600	-	SPX domain gene 2	-4.78	0.00005
Manes.01G274500	GO:0006629	Alpha/beta-Hydrolases superfamily protein	-5.15	0.00005
Manes.14G173500	GO:0016787	Purple acid phosphatase 17	-5.16	0.00005
Manes.02G002300	GO:0046983	Basic helix-loop-helix (bHLH) DNA-binding family protein	-5.23	0.00005
Manes.01G147400	-	Tetratricopeptide repeat (TPR)-like superfamily protein	-5.58	0.00025
Manes.02G039500	GO:0042545,GO:0030599,GO:005618	Pectin lyase-like superfamily protein	-5.60	0.00015
Manes.02G164500	GO:0016787,GO:0008152	IAA-leucine resistant (ILR)-like gene 6	-5.65	0.00005
Manes.03G047200	GO:0016791,GO:0008152	Pyridoxal phosphate phosphatase-related protein	-5.66	0.00005
Manes.03G031600	-	Uncharacterized protein	-5.91	0.0001
Manes.02G189200	GO:0016788	GDSL-like Lipase/Acylhydrolase superfamily protein	-6.00	0.00005
Manes.S095100	-	Lipid transfer protein 3	-6.02	0.00005
Manes.01G148900	GO:0046872,GO:0030001	Heavy metal transport/detoxification superfamily protein	-6.16	0.0003
Manes.02G086200	GO:0016829,GO:0010333,GO:0008152,GO:0000287	Terpenoid cyclases/Protein prenyltransferases superfamily protein	-6.86	0.0004
Manes.S021300,Manes.S021400	GO:0031012,GO:0008270,GO:0006508,GO:0004222	Matrix metalloproteinase	-6.91	0.00005
Manes.02G124600	-	HSP20-like chaperones superfamily protein	-7.65	0.00005
Gene ID	GO ID	Gene annotation	Log2(fold_change) 72 hpi	
Manes.15G148400	-	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	3.89	0.00005
Manes.01G087800	-	Expansin-like B1	3.73	0.00005
Manes.05G063900	GO:0016021	TSPO(outer membrane tryptophan-rich sensory protein)-related	3.72	0.00005
Manes.15G130200	GO:0009607,GO:0006952	Polyketide cyclase/dehydrase and lipid transport superfamily protein	3.68	0.00005
Manes.01G129500	GO:0005515,GO:0006468,GO:0005524,GO:0004672	Leucine-rich repeat protein kinase family protein	3.32	0.00005
Manes.18G101700, Manes.18G101800, Manes.18G101900	-	Terpenoid cyclases/Protein prenyltransferases superfamily protein	3.09	0.0001
Manes.13G143700	GO:0046872	Zinc finger C-x8-C-x5-C-x3-H type family protein	2.71	0.00005
Manes.02G115200	GO:0055085,GO:0016021	Auxin efflux carrier family protein	2.61	0.00005
Manes.11G137300	GO:0005975,GO:0004553	Chitinase	2.53	0.00005
Manes.05G143800	-	Expansin 11	2.42	0.00005
Manes.18G032200	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Cytochrome P450, family 82, subfamily C, polypeptide 4	2.30	0.00005
Manes.12G103400	GO:0046983	Basic helix-loop-helix (bHLH) DNA-binding superfamily protein	2.14	0.00005
Manes.02G039500	GO:0042545,GO:0030599,GO:005618	Pectin lyase-like superfamily protein	2.09	0.00005
Manes.02G151200	-	Phosphate-responsive 1 family protein	2.05	0.00015
Manes.01G196200	GO:0016787,GO:0008152	IAA-leucine resistant (ILR)-like gene 2	1.88	0.00005
Manes.13G101600	GO:0008152,GO:0003824,GO:0016114,GO:0008661	Deoxyxylulose-5-phosphate synthase	1.85	0.00005
Manes.06G030200	-	Ribonuclease H-like superfamily protein	1.78	0.00005
Manes.13G056600	GO:0016758,GO:0008152	UDP-Glycosyltransferase superfamily protein	1.77	0.0001
Manes.13G141100	GO:0016021,GO:0006952	Seven transmembrane MLO family protein	1.70	0.00005
Manes.S043400,Manes.S043500	GO:0005975,GO:0004553	Beta glucosidase 40	1.47	0.00005
Manes.04G159500	GO:0055085,GO:0016021,GO:0016020,GO:0006810,GO:0005215	Major facilitator superfamily protein	1.28	0.00015
Manes.05G029900	GO:0030170,GO:0009058	ACC synthase 1	1.25	0.00005

Manes.01G189000	GO:0043565,GO:0006355,GO:003700	WRKY DNA-binding protein 72	1.03	0.00005
Manes.07G077900	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Cytochrome P450, family 707, subfamily A, polypeptide 4	0.92	0.00005
Manes.12G076900	GO:0016829,GO:0010333,GO:0008152,GO:0000287	Terpene synthase 02	0.59	0.00005
Manes.06G108300		Raffinose synthase family protein	0.09	0.0001
Manes.02G040000	GO:0016021,GO:0016020,GO:0022857	Nodulin MtN21 /EamA-like transporter family protein	-0.69	0.00005
Manes.17G017700	-	Tetratricopeptide repeat (TPR)-like superfamily protein	-0.76	0.00005
Manes.12G109400	-	RING/U-box superfamily protein	-1.22	0.0004
Manes.02G105300	-	Pathogenesis-related thaumatin superfamily protein	-1.52	0.00005
Manes.18G040300	GO:0003677	WUSCHEL related homeobox 4	-1.57	0.00005
Manes.06G026500	GO:0048544,GO:0006468,GO:0004672	S-locus lectin protein kinase family protein	-1.63	0.00005
Manes.06G066800	GO:0006308,GO:0004519,GO:0003676	Endonuclease 2	-1.71	0.00005
Manes.14G118700	GO:0055114,GO:0016491	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	-1.79	0.00015
Manes.01G220700	-	Protein of unknown function, DUF538	-1.81	0.00015
Manes.02G065400	GO:0055114,GO:0016616,GO:0006694,GO:0003854	NAD(P)-binding Rossmann-fold superfamily protein	-1.96	0.0001
Manes.01G196500	GO:0016758,GO:0008152	UDP-glucosyl transferase 74B1	-2.23	0.0002
Manes.03G061200		Tetraspanin family protein	-2.54	0.00015
Manes.08G154900	GO:0046983,GO:0008171	O-methyltransferase 1	-2.56	0.00005
Manes.01G041100	GO:0005515	Disease resistance family protein / LRR family protein	-2.85	0.0002
Manes.16G039100	-	Hypothetical protein	-2.85	0.00005
Manes.17G075600	-	Myb domain protein 4	-2.95	0.00015
Manes.01G204500	GO:0047746,GO:0015996	Chlorophyllase 1	-3.02	0.0001
Manes.07G064200	GO:0005525	ADP-ribosylation factor family protein	-3.12	0.00005
Manes.07G050700	-	CAP (Cysteine-rich secretory proteins, Antigen 5, and Pathogenesis-related 1 protein) superfamily protein	-3.26	0.0001
Manes.02G164500	GO:0016787,GO:0008152	IAA-leucine resistant (ILR)-like gene 6	-3.31	0.0001
Manes.13G127800	GO:0055085,GO:0016020,GO:0015297,GO:0015238,GO:0006855	MATE efflux family protein	-3.35	0.0002
Manes.02G099700	GO:0016787	Purple acid phosphatase 25	-3.4	0.0002
Manes.01G250800	GO:0051087	BCL-2-associated athanogene 6	-3.46	0.00015
Manes.02G205700	GO:0007165,GO:0005515,GO:00043531	Disease resistance protein (TIR-NBS-LRR class), putative	-3.57	0.00005
Manes.02G035800	GO:0055114,GO:0016491,GO:0005507	Cupredoxin superfamily protein	-3.88	0.0002
Manes.02G086100, Manes.02G086200, Manes.02G086300	GO:0016829,GO:0010333,GO:0000287	Terpene synthase 21	-3.98	0.00005
Manes.07G070800	GO:0016758,GO:0008152	UDP-Glycosyltransferase superfamily protein	-4.02	0.0001
Manes.01G241000	GO:0016747,GO:0008610,GO:0016020,GO:0006633	3-ketoacyl-CoA synthase 1	-4.07	0.00015
Manes.11G105200	GO:0016788	GDSSL-like Lipase/Acylhydrolase superfamily protein	-4.15	0.0001
Manes.12G028900	GO:0003677	Basic helix-loop-helix (bHLH) DNA-binding superfamily protein	-4.26	0.00005
Manes.02G025100	-	Osmotin 34	-4.31	0.00005
Manes.02G086300	GO:0016829,GO:0010333,GO:0008152,GO:0000287	Terpene synthase 21	-4.52	0.0001
Manes.02G028300	-	Osmotin 34	-4.86	0.00035
Manes.13G125100	GO:0006355,GO:0003700	Integrase-type DNA-binding superfamily protein	-5.10	0.00005
Manes.01G203500	-	Calmodulin-like 11	-5.11	0.00005
Manes.01G274500	GO:0006629	Alpha/beta-Hydrolases superfamily protein	-5.17	0.00005

Manes.S021200	GO:0031012,GO:0008270,GO:006508,GO:0004222	Matrixin family protein	-5.35	0.0001
Manes.01G076500	GO:0004575,GO:0004564	Glycosyl hydrolases family 32 protein	-5.58	0.00005
Manes.14G173500	GO:0016787	Purple acid phosphatase 17	-5.81	0.00005
Manes.02G189200	GO:0016788	GDSL-like Lipase/Acylhydrolase superfamily protein	-6.01	0.00005
Manes.01G142100	GO:0005515	Leucine-rich repeat (LRR) family protein	-6.16	0.0001
Manes.04G089800	GO:0008081,GO:0006629	Senescence-related gene 3	-6.17	0.00005
Manes.08G055600	-	Hypothetical protein	-6.17	0.00005
Manes.02G124700	-	HSP20-like chaperones superfamily protein	-7.07	0.0003
Gene ID	GO ID	Gene annotation	Log2(fold_change) 24 hpi vs 72 hpi	
Manes.02G151000	-	Phosphate-responsive 1 family protein	5.48	0.00005
Manes.04G029500	GO:0055114,GO:0016491	Exocyst subunit exo70 family protein	5.42	0.00005
Manes.05G195300	GO:0016021,GO:0015012,GO:006024	H7	5.38	0.0001
Manes.03G044100	GO:0043565,GO:0006355,GO:005634,GO:0003700	Transmembrane amino acid transporter family protein	5.31	0.00005
Manes.01G121300	GO:0016758,GO:0008152	Integrase-type DNA-binding superfamily protein	5.13	0.0004
Manes.07G072600	GO:0050832,GO:0042742	ARM repeat superfamily protein	5.04	0.00005
Manes.13G125100	-	Matrixin family protein	4.98	0.00055
Manes.S055900	-	Integrase-type DNA-binding superfamily protein	4.64	0.00005
Manes.S002300	GO:0016829,GO:0010333,GO:008152,GO:0000287	Xyloglucan endotransglycosylase 6	4.63	0.00005
Manes.02G124700,	-	C2H2 and C2HC zinc fingers superfamily protein	4.51	0.00005
Manes.02G124800	-	HSP20-like chaperones superfamily protein	4.19	0.00005
Manes.18G136400	-	Protein phosphatase 2C family protein	4.15	0.00005
Manes.05G113200	-	Calmodulin like 37	3.98	0.00035
Manes.02G146000	-	Dicarboxylate carrier 2	3.69	0.00005
Manes.02G130000	-	Zinc finger (C3HC4-type RING finger) family protein	3.58	0.0001
Manes.06G020900	GO:0016747,GO:0008610,GO:0016020,GO:0006633	Glycosyltransferase family protein 47	3.55	0.00005
Manes.16G116900	-	Arabinogalactan protein 22	3.52	0.0001
Manes.18G034400	-	Protein of unknown function (DUF1645)	3.35	0.00005
Manes.11G135900	GO:0043531	Exocyst subunit exo70 family protein	3.32	0.00005
Manes.02G017100	-	H7	3.27	0.00005
Manes.03G118300	GO:0055114,GO:0020037,GO:0016705,GO:0005506	WRKY DNA-binding protein 27	3.26	0.00005
Manes.09G046000	-	Calcium-binding EF-hand family protein	3.17	0.00005
Manes.02G125200	GO:0003824,GO:0006470,GO:004722	plant U-box 23	3.10	0.00015
Manes.17G071600	-	HSP20-like chaperones superfamily protein	2.85	0.00005
Manes.18G065800	GO:0006355,GO:0003700	Jasmonate-zim-domain protein 8	2.84	0.0003
Manes.02G110100	-	Phosphate-responsive 1 family protein	2.82	0.00005
Manes.09G112700	-	RING/U-box superfamily protein	2.47	0.00005
Manes.03G106300	-	WRKY DNA-binding protein 33	2.37	0.00005
Manes.05G030900	GO:0005515	Ethylene responsive element binding factor 4	2.36	0.00005
Manes.05G165900	-	WRKY family transcription factor	2.33	0.00015
Manes.07G063300	GO:0031012,GO:0008270,GO:006508,GO:0004222	Uncharacterised protein family (UPF0114)	1.78	0.00005
Manes.15G189500	-	Protein phosphatase 2C family protein	1.73	0.00005
Manes.05G038800	-	HSP20-like chaperones superfamily protein	1.49	0.00005
Manes.13G141100	-	BON association protein 2	1.45	0.00005
Manes.09G113500	GO:0006887,GO:0000145	Seven transmembrane MLO family protein	1.35	0.00005
Manes.18G118900	GO:0003824,GO:0006470,GO:004722	Hypothetical protein	0.82	0.0003
Manes.06G051100	-	Integrase-type DNA-binding superfamily protein	0.54	0.00005
Manes.04G005600	GO:0006887,GO:0000145	3-ketoacyl-CoA synthase 11	0.23	0.00025
Manes.13G152300	-	Cytochrome P450, family 94, subfamily C, polypeptide 1		
		Wound-responsive family protein		

Manes.S051300	GO:0048046,GO:0016762,GO:006073,GO:0005618,GO:0005975,GO:0004553	Terpenoid cyclases/Protein prenyltransferases superfamily protein	-0.32	0.00005
Manes.13G056600	GO:0043565,GO:0006355,GO:005634,GO:0003700	UDP-Glycosyltransferase superfamily protein	-0.59	0.00005
Manes.04G109700	-	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	-0.75	0.00035
Manes.14G143500	GO:0043531	Heat shock protein 70B	-0.96	0.00005
Manes.03G040200	GO:0006355,GO:0003700	Acyl-CoA N-acyltransferases (NAT) superfamily protein	-0.99	0.00005
Manes.S067700	-	Tetratricopeptide repeat (TPR)-like superfamily protein	-1.03	0.00005
Manes.08G135300	GO:0016567,GO:0004842	Pathogenesis-related 4	-1.09	0.00025
Manes.11G156500	-	Disease resistance protein (TIR-NBS-LRR class) family	-1.11	0.00005
Manes.12G024900	-	Alpha/beta-Hydrolases superfamily protein	-1.13	0.00005
Manes.09G051500	GO:0043565,GO:0006355,GO:003700	TRICHOME BIREFRINGENCE-LIKE 8	-1.29	0.00025
Manes.16G025700	-	Aluminium induced protein with YGL and LRDR motifs	-1.36	0.00005
Manes.02G044700	GO:0016829,GO:0010333,GO:000287	Expansin-like B1	-1.40	0.00025
Manes.18G023000	GO:0055114,GO:0020037,GO:0016705,GO:0005506	uncharacterized protein	-1.42	0.00005
Manes.14G165100	-	LOC110606054	-1.42	0.00005
Manes.13G139000	GO:0016021,GO:0006952	Disease resistance protein (TIR-NBS-LRR class) family	-1.45	0.00005
Manes.02G221500	GO:0008080	Late embryogenesis abundant domain-containing protein / LEA domain-containing protein	-1.47	0.00005
Manes.04G156300	GO:0043565,GO:0006355,GO:003700	Protein of unknown function (DUF1442)	-1.61	0.00005
Manes.18G032200	-	Protein LNK3-like	-1.68	0.00005
Manes.02G128000	GO:0008270,GO:0005515	Cytochrome P450, family 82, subfamily C, polypeptide 4	-2.04	0.00015
Manes.01G119000	GO:0016567,GO:0004842	Highly ABA-induced PP2C gene 3	-2.47	0.00005
Manes.01G117900	-	HSP20-like chaperones superfamily protein	-3.02	0.00005
Manes.02G086100, Manes.02G086200, Manes.02G086300	GO:0008270,GO:0005515	Senescence-associated family protein	-3.12	0.00015
Manes.13G124500	GO:0006355,GO:0003700	Terpene synthase 21	-3.61	0.0002
Manes.06G091300	GO:0051082,GO:0006950,GO:006457,GO:0005524	Heat shock factor 4	-3.73	0.00005
Manes.01G196500	GO:0043565,GO:0006355,GO:003700	Expansin A1	-3.89	0.00005
Manes.08G135200	GO:0008061,GO:0050832,GO:0042742	UDP-glucosyl transferase 74B1	-4.21	0.00015
Manes.02G124600	-	Pathogenesis-related 4	-4.39	0.0001
Manes.01G116700	-	HSP20-like chaperones superfamily protein	-4.71	0.00005
Manes.03G049500	GO:0006355,GO:0003700	Calmodulin like 23	-4.71	0.0001
Manes.06G152800	GO:0003824,GO:0006470,GO:004722	Heat shock transcription factor A2	-4.91	0.0002
Manes.13G013400	GO:0016758,GO:0008152	Heat shock protein 90.1	-5.21	0.00005
		Heat shock protein 21	-5.31	0.00005

Table A4: List of 98/0581 differentially expressed genes in cassava leaves in response to mealybug infestation at 24 and 72 hours post-infestation normalized against mock (non-infested).

Gene ID	GO ID	Gene annotation	Log2(fold_change) 24 hpi	P-value
Manes.11G029700	GO:0046983	Basic helix-loop-helix (bHLH) DNA-binding superfamily protein	5.74	0.0001

Manes.06G046800	-	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	5.36	0.00005
Manes.S016300	GO:0005975,GO:0004559	Syntaxin of plants 51	5.32	0.00005
Manes.15G148400	-	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	4.43	0.0001
Manes.12G137800	GO:0008654,GO:0006021,GO:004512	Myo-inositol-1-phosphate synthase 3	4.31	0.0001
Manes.15G094200	GO:0048544	D-mannose binding lectin protein with Apple-like carbohydrate-binding domain	4.17	0.00005
Manes.05G155300	GO:0016773,GO:0016020,GO:0005524	Mitogen-activated protein kinase 1	4.01	0.00005
Manes.18G053300	-	Amino acid permease 6	3.43	0.0001
Manes.16G117300	GO:0055114,GO:0016491,GO:0008152	Aldehyde dehydrogenase 2C4	2.97	0.00005
Manes.15G130000	GO:0009607,GO:0006952	MLP-like protein 28	2.89	0.00005
Manes.01G004300	-	Arabinogalactan protein 14	2.85	0.00005
Manes.09G040500	-	Alpha/beta-Hydrolases superfamily protein	2.36	0.00005
Manes.10G050800	GO:0016747	HXXXD-type acyl-transferase family protein	2.25	0.00005
Manes.08G109900	-	Protein of unknown function, DUF538	2.09	0.0008
Manes.11G112600	GO:0005507,GO:0055114,GO:0016491	Plant L-ascorbate oxidase	2.03	0.0008
Manes.13G130400	GO:0016788	GDLS-like Lipase/Acylhydrolase superfamily protein	1.88	0.0006
Manes.14G039700	GO:0055114,GO:0016491	KAR-UP oxidoreductase 1	1.67	0.00005
Manes.15G130300	GO:0009607,GO:0006952	MLP-like protein 43	1.63	0.0003
Manes.02G174900	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Cytochrome P450, family 87, subfamily A, polypeptide 2	1.48	0.00005
Manes.05G020400	GO:0016567,GO:0004842	Plant U-box 23	1.31	0.00005
Manes.13G076400	-	Hypothetical protein MANES_13G076400	1.23	0.00005
Manes.01G059800	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Cytochrome P450, family 86, subfamily A, polypeptide 8	1.18	0.0001
Manes.10G129500	GO:0005515	Disease resistance family protein / LRR family protein	1.11	0.00005
Manes.17G079900	GO:0016868,GO:0005975	Phosphoglucosyltransferase, putative / glucose phosphomutase, putative	1.04	0.00045
Manes.04G140800	-	Myb domain protein 26	0.98	0.00005
Manes.05G153700	GO:0006355,GO:0003700	Integrase-type DNA-binding superfamily protein	0.89	0.0002
Manes.10G024100	GO:0005515	Receptor like protein 6	0.89	0.00015
Manes.12G142700	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Cytochrome P450, family 78, subfamily A, polypeptide 7	0.65	0.0001
Manes.S040700	GO:0005515	Receptor like protein 14	0.59	0.0001
Manes.02G124600	-	HSP20-like chaperones superfamily protein	0.58	0.0002
Manes.12G115900	GO:0016021	Heptahelical protein 4	0.35	0.00005
Manes.10G149400	GO:0005975,GO:0004553	Heta glucosidase 12	0.18	0.0003
Manes.01G163600	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Cytochrome P450, family 86, subfamily A, polypeptide 1	0.13	0.00015
Manes.09G112400	GO:0003824,GO:0006470,GO:0004722	Protein phosphatase 2C family protein	0.06	0.00005
Manes.07G048600	GO:0046983	Basic helix-loop-helix (bHLH) DNA-binding superfamily protein	-0.07	0.00005
Manes.08G000100	GO:0016887,GO:0005524,GO:0016020	ABC-2 type transporter family protein	-0.08	0.0002
Manes.07G035000	GO:0055114,GO:0016491	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	-0.70	0.00005
Manes.05G167500	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Cytochrome P450, family 82, subfamily C, polypeptide 2	-0.71	0.00005
Manes.03G165400	GO:0008270,GO:0005515	RING/U-box superfamily protein	-1.11	0.00035
Manes.06G177200	-	Uncharacterized protein	-1.25	0.00005
Manes.17G051400	GO:0007264,GO:0005525	LOC110617606	-1.29	0.00005
Manes.16G072800	-	RAC-like 2	-1.38	0.00045
Manes.13G074200	GO:0009055	Ankyrin repeat family protein	-1.48	0.00005
		Uclacyanin 1		

Manes.06G091800	GO:0016788	GDSL-like Lipase/Acylhydrolase superfamily protein	-1.5	0.00005
Manes.06G012500	GO:0005515,GO:0006468,GO:005524,GO:0004672	Leucine-rich repeat receptor-like protein kinase family protein	-1.59	0.00005
Manes.07G102200	-	LOB domain-containing protein 1	-1.62	0.0001
Manes.15G189500	-	HSP20-like chaperones superfamily protein	-1.74	0.00025
Manes.10G097000	GO:0016758,GO:0008152	UDP-glucosyl transferase 78D2	-1.81	0.00005
Manes.07G036100	-	Hydroxyproline-rich glycoprotein family protein	-2.67	0.0001
Manes.14G147100	-	Barwin-related endoglucanase	-2.74	0.0008
Manes.05G020800	-	Myb domain protein 3	-2.75	0.0008
Manes.07G048700	GO:0016021,GO:0015914,GO:005524,GO:0004012,GO:0000287	ATPase E1-E2 type family protein / haloacid dehalogenase-like hydrolase family protein	-2.79	0.00005
Manes.06G009100	GO:0034968,GO:0018024,GO:008270,GO:0005634,GO:0005515	SU(VAR)3-9 homolog 4	-2.88	0.0001
Manes.12G028900	GO:0003677	Basic helix-loop-helix (bHLH) DNA-binding superfamily protein	-2.96	0.00005
Manes.08G155100,	GO:0046983,GO:0008171	O-methyltransferase 1	-3.15	0.00005
Manes.08G155200	GO:0005525	ADP-ribosylation factor family protein	-3.26	0.00005
Manes.07G064200	GO:0051082,GO:0006950,GO:006457,GO:0005524	Heat shock protein 90.1	-3.59	0.00005
Manes.06G152800	-	Geranylgeranyl reductase	-3.62	0.00005
Manes.15G071000	-	Hydroxyproline-rich glycoprotein family protein	-3.65	0.00005
Manes.07G040700	-	Hydroxyproline-rich glycoprotein family protein	-3.69	0.00005
Manes.07G040100	-	Monogalactosyldiacylglycerol synthase 2	-3.76	0.0001
Manes.14G118100	GO:0016758	P-loop containing nucleoside triphosphate hydrolases superfamily protein	-3.83	0.00005
Manes.02G200000	GO:0019205,GO:0006139,GO:005524	Pectin methylesterase 3	-4.01	0.00005
Manes.16G089000	GO:0004857,GO:0042545,GO:030599,GO:0005618	Proteinase inhibitor I25, cystatin, conserved region	-4.06	0.00005
Manes.13G041600	-	WRKY DNA-binding protein 70	-4.08	0.0008
Manes.08G170600	GO:0043565,GO:0006355,GO:003700	SPX domain gene 2	-4.09	0.00005
Manes.18G052600	GO:0005525,GO:0003924	Dynamin related protein 4C	-4.13	0.0008
Manes.09G183900,	-	Glutathione S-transferase TAU 28	-4.14	0.00005
Manes.09G184000	GO:0031012,GO:0008270,GO:006508,GO:0004222	Matrix metalloproteinase	-4.22	0.00005
Manes.09G086300	GO:0055114,GO:0016491,GO:008270	Alkenal reductase	-4.30	0.0001
Manes.S021300,Manes.S021400	GO:0006355,GO:0003700	Integrase-type DNA-binding superfamily protein	-4.40	0.0008
Manes.05G121500,	GO:0050662,GO:0003824	NAD(P)-binding Rossmann-fold superfamily protein	-4.40	0.00025
Manes.05G121600	-	Plant natriuretic peptide A	-4.46	0.00005
Manes.06G058200	GO:0008081,GO:0006629	Senescence-related gene 3	-4.70	0.00005
Manes.16G016400	-	Hypothetical protein MANES_06G024100	-4.71	0.0004
Manes.14G147000	-	Alpha/beta-Hydrolases superfamily protein	-4.72	0.0004
Manes.04G089800	GO:0006457	rotamase FKBP 1	-4.76	0.00005
Manes.06G024100	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Cytochrome P450 superfamily protein	-4.94	0.00005
Manes.11G008300	GO:0016787	Purple acid phosphatase 17	-5.05	0.00005
Manes.05G198600	GO:0031012,GO:0008270,GO:006508,GO:0004222	Matrixin family protein	-5.21	0.00005
Manes.06G085500	GO:0043565	multiprotein bridging factor 1C	-5.21	0.00005
Manes.14G173500				
Manes.S021200				
Manes.03G019200				

Gene ID	GO ID	Gene annotation	Log2(fold_change) 72 hpi	
Manes.11G029700	GO:0046983	Basic helix-loop-helix (bHLH) DNA-binding superfamily protein	5.13	0.00005
Manes.02G174100		Nodulin MtN3 family protein	4.33	0.00005

Manes.05G066700	GO:0045454,GO:0015035,GO:009055	Thioredoxin superfamily protein PPPDE putative thiol peptidase	4.07	0.0001
Manes.13G045400	-	family protein	3.95	0.00005
Manes.02G119600	GO:0055085,GO:0016021	Cation exchanger 3	3.91	0.00005
Manes.07G081100	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Cytochrome P450, family 707, subfamily A, polypeptide 1	3.90	0.00005
Manes.18G022300	GO:0050662,GO:0003824	Dihydroflavonol 4-reductase	3.87	0.0001
Manes.18G056800	-	Cox19-like CHCH family protein	3.34	0.00005
Manes.11G139700	GO:0055114,GO:0020037,GO:006979,GO:0004601	Ascorbate peroxidase 1	3.13	0.00005
Manes.08G102900	-	Uncharacterized proline-rich protein- like isoform X3	3.00	0.00005
Manes.12G137800	GO:0008654,GO:0006021,GO:004512	Myo-inositol-1-phosphate synthase 3 HSP20-like chaperones superfamily protein	2.89	0.00005
Manes.01G167000	-	Transmembrane amino acid transporter family protein	2.82	0.00005
Manes.07G095400	-		2.70	0.00005
Manes.03G084300	GO:0055114,GO:0020037,GO:006979,GO:0004601	Peroxidase superfamily protein HSP20-like chaperones superfamily protein	2.64	0.00005
Manes.02G125200	-	Barwin-related endoglucanase	2.58	0.00005
Manes.14G147100	-		2.43	0.00005
Manes.16G043300	GO:0003824,GO:0042127,GO:0018343,GO:0005965	Prenyltransferase family protein ATPase E1-E2 type family protein / haloacid dehalogenase-like hydrolase family protein	2.37	0.00005
Manes.07G048700	GO:0016021,GO:0015914,GO:0005524,GO:0004012,GO:0000287		2.36	0.00005
Manes.05G078200	-	Seed maturation protein	2.35	0.00005
Manes.09G040200	GO:0003676	SC35-like splicing factor 30A	2.33	0.00005
Manes.09G020300	GO:0043565,GO:0006355,GO:0005634,GO:0003700	Heat shock transcription factor B3	2.33	0.00005
Manes.05G031200	GO:0006468,GO:0005524,GO:004672	MAP kinase 4	2.33	0.00005
Manes.05G020800	-	Myb domain protein 3 Hypothetical protein MANES_01G002000	2.09	0.00005
Manes.01G002000	-		1.76	0.00005
Manes.07G114100,	-	Heat shock cognate protein 70-1	1.64	0.00005
Manes.07G114200	-	CCT motif family protein	1.64	0.00015
Manes.11G101000	GO:0005515			
Manes.06G021700	GO:0005515,GO:0055114,GO:0046872,GO:0016702	Lipoxygenase 1 HSP20-like chaperones superfamily protein	1.27	0.00005
Manes.02G124600	-	HSP20-like chaperones superfamily protein	0.97	0.00005
Manes.10G020600	-		0.80	0.00005
Manes.14G148600,	-	HSP20-like chaperones superfamily protein	0.60	0.00005
Manes.14G148700	-	Disease resistance family protein / LRR family protein	0.50	0.00005
Manes.07G042900.1	GO:0005515	Cytochrome P450, family 87, subfamily A, polypeptide 2	0.39	0.00005
Manes.02G174900	GO:0055114,GO:0020037,GO:0016705,GO:0005506			
Manes.S104500	GO:0006412,GO:0005840,GO:0005622,GO:0003735	R-protein L3 B Cysteine-rich RLK (RECEPTOR-like protein kinase) 29	0.37	0.00005
Manes.16G018000	GO:0006468,GO:0004672	Protein of unknown function (DUF594)	0.34	0.00015
Manes.16G012100	-	Nodulin MtN21 /EamA-like transporter family protein	0.22	0.00015
Manes.02G040000	GO:0016021,GO:0016020,GO:0022857	Phytosulfokin receptor 1	0.20	0.00005
Manes.10G104200	GO:0005515	ENTH/ANTH/VHS superfamily protein	0.12	0.00005
Manes.14G158100	GO:0005543	UDP-Glycosyltransferase superfamily protein	-0.14	0.00005
Manes.06G168400	GO:0016758,GO:0008152		-0.24	0.0001
Manes.16G096700,	-	Myb domain protein 3	-0.53	0.0002
Manes.16G096800	-	Leucine-rich repeat receptor-like protein kinase family protein	-0.58	0.00005
Manes.11G118000	GO:0005515,GO:0006468,GO:0005524,GO:0004672	Nudix hydrolase homolog 13	-0.67	0.00005
Manes.09G173300	GO:0016787			

Manes.S092200	GO:0016020,GO:0006810,GO:005215	NRAMP metal ion transporter 6	-0.82	0.00005
Manes.06G086200	GO:0016758,GO:0008152	UDP-Glycosyltransferase superfamily protein	-0.89	0.00015
Manes.07G048600	GO:0046983	Basic helix-loop-helix (bHLH) DNA-binding superfamily protein	-0.95	0.0002
Manes.13G121400	GO:0016747	HXXXD-type acyl-transferase family protein	-1.05	0.00005
Manes.17G017700	-	Tetratricopeptide repeat (TPR)-like superfamily protein	-1.50	0.0008
Manes.02G124700, Manes.02G124800	-	HSP20-like chaperones superfamily protein	-1.51	0.0004
Manes.04G118800	-	Bromo-adjacent homology (BAH) domain-containing protein	-1.54	0.00015
Manes.09G108600	-	PEBP (phosphatidylethanolamine-binding protein) family protein	-1.55	0.00005
Manes.12G058100	-	Uncharacterized protein	-1.62	0.00005
Manes.11G140200	-	LOC110628434	-1.71	0.00005
Manes.10G097000	GO:0016758,GO:0008152	Cysteine proteinases superfamily protein	-1.84	0.00005
Manes.09G160400	GO:0055114,GO:0046872,GO:006801	UDP-glucosyl transferase 78D2	-2.41	0.00005
Manes.05G081000	-	Copper/zinc superoxide dismutase 1	-2.44	0.00005
Manes.13G084000	-	chlorophyllase 1	-2.49	0.00005
Manes.13G094200	GO:0055114,GO:0020037,GO:0016705,GO:0005506	DNAJ heat shock N-terminal domain-containing protein	-2.54	0.00015
Manes.17G080500	-	Cytochrome P450, family 79, subfamily B, polypeptide 2	-2.57	0.00005
Manes.12G045600	-	ARM repeat superfamily protein	-2.66	0.00005
Manes.03G116500	GO:0055114,GO:0017004,GO:0016020	Uncharacterised protein family SERF	-2.72	0.00015
Manes.08G094300	-	Cytochrome c biogenesis protein family	-2.74	0.00005
Manes.17G079900	GO:0016868,GO:0005975	DNAJ heat shock N-terminal domain-containing protein	-2.77	0.00005
Manes.12G067800	GO:0006508,GO:0004252	Phosphoglucomutase, putative / glucose phosphomutase, putative	-2.91	0.00005
Manes.01G192700	GO:0005515	Subtilisin-like serine endopeptidase family protein	-2.98	0.00005
Manes.06G052400	-	Glutathione S-transferase family protein	-3.24	0.00005
Manes.04G063000	GO:0051082,GO:0031072	Protein of unknown function (DUF3511)	-3.30	0.00005
Manes.15G154400	-	DNAJ homologue 2	-3.34	0.0001
Manes.12G132900	GO:0055114,GO:0020037,GO:0016705,GO:0005506	HSP20-like chaperones superfamily protein	-3.38	0.00005
Manes.08G133300, Manes.08G133400, Manes.08G133500, Manes.08G133700, Manes.08G133800	-	Cytochrome P450, family 71, subfamily B, polypeptide 10	-3.39	0.00005
Manes.11G067600	-	NAD(P)-linked oxidoreductase superfamily protein	-3.43	0.0002
Manes.08G082800	-	heat shock protein 70B	-3.64	0.00005
Manes.01G202400, Manes.01G202500, Manes.01G202600	-	HSP20-like chaperones superfamily protein	-3.69	0.00005
Manes.08G155100, Manes.08G155200	GO:0046983,GO:0008171	Class I glutamine amidotransferase-like superfamily protein	-3.87	0.0001
Manes.05G121500, Manes.05G121600	GO:0055114,GO:0016491,GO:0008270	O-methyltransferase 1	-4.00	0.00005
Manes.17G070300	-	Alkenal reductase	-4.07	0.00005
Manes.07G064200	GO:0005525	SGS domain-containing protein	-4.09	0.0002
Manes.06G085600	GO:0019538,GO:0016887,GO:0005524	ADP-ribosylation factor family protein	-4.32	0.00005
Manes.13G041600	-	Heat shock protein 101	-4.48	0.00005
Manes.02G028600	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Proteinase inhibitor I25, cystatin, conserved region	-4.49	0.00005
		Cytochrome P450, family 76, subfamily C, polypeptide 3		

Manes.03G049500	GO:0043565,GO:0006355,GO:005634,GO:0003700	Heat shock transcription factor A2	-4.49	0.00005
Manes.11G099500	GO:0051082,GO:0031072	DNAJ homologue 2	-4.61	0.00005
Manes.14G121300	GO:0005509	calcium-binding EF hand family protein	-4.7	0.00005
Manes.01G250800	GO:0051087	BCL-2-associated athanogene 6	-4.71	0.00006
Manes.15G189500		HSP20-like chaperones superfamily protein	-5.15	0.00025
Manes.14G022300	GO:0051082,GO:0006950,GO:006457,GO:0005524	Heat shock protein 90.1	-5.44	0.00005
Manes.08G080400, Manes.08G080500	-	Regulator of chromosome condensation (RCC1) family protein	-5.59	0.00008
Manes.14G036000	-	HSP20-like chaperones superfamily protein	-5.76	0.00005
Manes.03G051500	-	17.6 kDa class II heat shock protein	-5.77	0.00005
Manes.06G134700	-	Mitochondrion-localized small heat shock protein 23.6	-6.02	0.00005
Manes.06G152800	GO:0051082,GO:0006950,GO:006457,GO:0005524	Heat shock protein 90.1	-6.11	0.00005
Gene ID	GO ID	Gene annotation	Log2(fold_change) 24 hpi vs 72 hpi	
Manes.18G007900	GO:0008236,GO:0006508,GO:0070008,GO:0004252	Prolyl oligopeptidase family protein	6.26	0.00005
Manes.03G051500	-	17.6 kDa class II heat shock protein	5.65	0.00001
Manes.02G124500	-	HSP20-like chaperones superfamily protein	4.69	0.00015
Manes.08G054400	GO:0005975,GO:0004553	Beta-1,3-glucanase 3	4.35	0.00005
Manes.04G086000	-	Purple acid phosphatase 10	4.29	0.00005
Manes.14G118100	GO:0016758	Monogalactosyldiacylglycerol synthase 2	4.24	0.00015
Manes.07G050300	-	Basic pathogenesis-related protein 1	4.2	0.00025
Manes.09G146400	GO:0048544	Curculin-like (mannose-binding) lectin family protein	4.18	0.00005
Manes.09G131200	GO:0016747	HXXXD-type acyl-transferase family protein	4.15	0.00025
Manes.S111300	-	Alpha/beta-Hydrolases superfamily protein	4.07	0.00015
Manes.08G135900	GO:0005515	LisH dimerisation motif;WD40/YVTN repeat-like-containing domain	4.05	0.00005
Manes.08G171800	-	Lipid transfer protein 3	3.97	0.00005
Manes.17G101600	-	RAD-like 1	3.89	0.00025
Manes.18G020900	GO:0016773,GO:0016020,GO:005524	Mitogen-activated protein kinase 1	3.85	0.00001
Manes.13G007700	GO:0005515,GO:0051087	BCL-2-associated athanogene 5	3.82	0.00005
Manes.08G063100	GO:0016021,GO:0016020,GO:0022857	Nodulin MtN21 /EamA-like transporter family protein	3.81	0.00005
Manes.15G077200	GO:0045454,GO:0015035,GO:006662	C-terminal cysteine residue is changed to a serine 1	3.8	0.00005
Manes.02G107700	GO:0016788	SGNH hydrolase-type esterase superfamily protein	3.73	0.00005
Manes.06G046800	-	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	3.72	0.00005
Manes.06G091300	-	Expansin A1	3.61	0.00002
Manes.11G008300	-	Alpha/beta-Hydrolases superfamily protein	3.38	0.00005
Manes.02G107900	GO:0016788	GDSSL-like Lipase/Acylhydrolase superfamily protein	3.36	0.00005
Manes.S053400	-	HSP20-like chaperones superfamily protein	3.11	0.00015
Manes.08G130700	GO:0045735	Germin-like protein 6	2.99	0.00005
Manes.11G130700	GO:0016998,GO:0006032,GO:004568	Homolog of carrot EP3-3 chitinase	2.78	0.00005
Manes.08G130900	GO:0045735	RmlC-like cupins superfamily protein	1.99	0.00005
Manes.02G115700	GO:0055085,GO:0016021,GO:0022857	Tonoplast monosaccharide transporter3	1.77	0.00005
Manes.02G143800		SNARE-like superfamily protein	1.52	0.00005
Manes.14G039700	GO:0055114,GO:0016491	KAR-UP oxidoreductase 1	1.39	0.00005
Manes.06G108600	GO:0016021,GO:0012511	Oleosin family protein	1.12	0.00005
Manes.06G131000	GO:0016021,GO:0012511	Oleosin 1	0.96	0.00005

Manes.05G153700	GO:0006355,GO:0003700	Integrase-type DNA-binding superfamily protein	0.91	0.00005
Manes.05G020400	GO:0016567,GO:0004842	Plant U-box 23	0.79	0.00005
Manes.07G072600	GO:0031012,GO:0008270,GO:006508,GO:0004222	Matrixin family protein	-0.19	0.00005
Manes.18G104400	-	Late embryogenesis abundant domain-containing protein / LEA domain-containing protein	-0.48	0.00005
Manes.07G059300	GO:0005975,GO:0004553	Glycosyl hydrolase superfamily protein	-0.67	0.00015
Manes.12G123800	GO:0006468,GO:0005524,GO:004672,GO:0030246	Concanavalin A-like lectin protein kinase family protein	-1.03	0.00005
Manes.06G156700	GO:0055085,GO:0022857,GO:0016021	Phosphate transporter 1;3	-1.09	0.00005
Manes.02G054300	GO:0016758,GO:0008152	UDP-Glycosyltransferase superfamily protein	-1.10	0.00005
Manes.01G236100	-	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein	-1.38	0.00005
Manes.12G004900	GO:0006508,GO:0004185	Serine carboxypeptidase-like 20	-2.07	0.00005
Manes.S095100	-	Lipid transfer protein 3	-2.30	0.00005
Manes.05G156900	GO:0005975,GO:0004553	Glycosyl hydrolase superfamily protein	-2.36	0.00005
Manes.08G044800	GO:0055114,GO:0016717,GO:006629	Fatty acid desaturase 2	-2.39	0.00005
Manes.09G040500	-	Alpha/beta-Hydrolases superfamily protein	-2.39	0.00005
Manes.01G074200	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Cytochrome P450, family 706, subfamily A, polypeptide 6	-2.39	0.00005
Manes.08G170600	GO:0043565,GO:0006355,GO:0003700	WRKY DNA-binding protein 70	-2.48	0.00005
Manes.06G085500	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Cytochrome P450 superfamily protein	-2.49	0.00005
Manes.11G112600	GO:0005507,GO:0055114,GO:0016491	Plant L-ascorbate oxidase	-2.54	0.00005
Manes.09G172000	-	Nodulin MtN3 family protein	-2.56	0.00005
Manes.16G040500	GO:0008234,GO:0006508	Granulin repeat cysteine protease family protein	-2.57	0.00005
Manes.16G089000	GO:0004857,GO:0042545,GO:0030599,GO:0005618	Pectin methylesterase 3	-2.60	0.00005
Manes.04G011700	GO:0008236,GO:0006508	Serine carboxypeptidase S28 family protein	-2.60	0.00005
Manes.17G081100	GO:0004857,GO:0042545,GO:0030599,GO:0005618	Plant invertase/pectin methylesterase inhibitor superfamily	-2.68	0.00005
Manes.06G147200	-	CAP160 protein	-2.68	0.00006
Manes.04G094200	GO:0016747,GO:0008610	Chalcone and stilbene synthase family protein	-2.70	0.00006
Manes.08G155100, Manes.08G155200	GO:0046983,GO:0008171	O-methyltransferase 1	-2.73	0.00004
Manes.05G033900	GO:0016567,GO:0004842,GO:0006468,GO:0005524,GO:0004672	U-box domain-containing protein kinase family protein	-2.79	0.00005
Manes.16G036500	GO:0016788	GDSSL-like Lipase/Acylhydrolase superfamily protein	-2.83	0.00005
Manes.08G109900	-	Protein of unknown function, DUF538	-2.85	0.00015
Manes.15G052000	GO:0005975,GO:0004553	Beta glucosidase 46	-2.87	0.00035
Manes.09G103000	-	Sulfoquinovosyldiacylglycerol 2	-2.88	0.00005
Manes.09G148300	GO:0030170,GO:0009058	Threonine aldolase 1	-2.9	0.00005
Manes.02G099700	GO:0016787	Purple acid phosphatase 25	-2.96	0.00005
Manes.18G088500, Manes.18G088600	-	Pathogenesis-related family protein 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	-2.99	0.00005
Manes.01G043500	GO:0055114,GO:0016491	Hypothetical protein	-3.01	0.00025
Manes.10G101300	-	MANES_10G101300	-3.04	0.00005
Manes.14G132600	-	SPX domain gene 2	-3.09	0.00005
Manes.02G095500	GO:0016853,GO:0005975	Galactose mutarotase-like superfamily protein	-3.10	0.00005
Manes.06G028100	GO:0008152,GO:0003824,GO:0016114,GO:0008661	Deoxyxylulose-5-phosphate synthase	-3.16	0.00005

Manes.10G060000	GO:0006412,GO:0005840,GO:005622,GO:0003735	Nucleic acid-binding, OB-fold-like protein	-3.18	0.00005
Manes.18G039200	GO:0006508,GO:0004252	Subtilisin-like serine endopeptidase family protein	-3.21	0.00005
Manes.16G098200	-	Phenylalanine ammonia-lyase 4	-3.27	0.00005
Manes.15G171200	GO:0016161,GO:0000272	Chloroplast beta-amylase	-3.29	0.0002
Manes.S053300	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Cytochrome P450, family 94, subfamily B, polypeptide 3	-3.31	0.00005
Manes.04G089500	-	F-box family protein	-3.32	0.00005
Manes.14G125300	GO:0055114,GO:0050660,GO:0016491	FAD-binding Berberine family protein	-3.32	0.00005
Manes.13G125300	GO:0055114,GO:0020037,GO:0006979,GO:0004601	Peroxidase superfamily protein	-3.35	0.00005
Manes.16G102300,	GO:0030246,GO:0005975,GO:0004553	Beta galactosidase 1	-3.35	0.00005
Manes.16G102400	-	Expansin-like B1	-3.39	0.00005
Manes.02G044700	GO:0016021	Heptahelical protein 4	-3.41	0.00005
Manes.12G115900	GO:0047746,GO:0015996	Chlorophyllase 1	-3.51	0.00005
Manes.05G081100	-	Plant natriuretic peptide A	-3.59	0.00005
Manes.14G147000	-	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein	-4.51	0.00015
Manes.02G096900	GO:0008168	Cytochrome P450, family 96, subfamily A, polypeptide 1	-4.70	0.00005
Manes.07G018100	GO:0055114,GO:0020037,GO:0016705,GO:0005506	Chitinase A	-4.76	0.00005
Manes.06G155500	GO:0005975,GO:0004553	Protein of unknown function (DUF1313)	-5.26	0.0004
Manes.07G129600	-	Senescence-related gene 3	-5.48	0.0008
Manes.04G089800	GO:0008081,GO:0006629	phospholipase D P1	-5.53	0.0001
Manes.08G114700	GO:0003824	WUSCHEL related homeobox 1	-5.81	0.00005
Manes.15G129200,	GO:0003677	Terpene synthase 14	-6.01	0.00005
Manes.15G129300	GO:0016829,GO:0010333,GO:0008152,GO:0000287	Senescence-associated gene 12	-6.06	0.0008
Manes.16G016100,	GO:0008234,GO:0006508	Matrix metalloproteinase	-6.39	0.0008
Manes.16G016200,	GO:0031012,GO:0008270,GO:0006508,GO:0004222	Ribonuclease 1	-6.69	0.00005
Manes.16G016300	GO:0033897,GO:0003723	Plasma membrane intrinsic protein 3	-6.82	0.00035
Manes.16G038400	GO:0016020,GO:0006810,GO:0005215	Beta glucosidase 32	-7.04	0.00035
Manes.17G119800	GO:0005975,GO:0004553	Pyridoxal phosphate phosphatase-related protein	-7.65	0.0008
Manes.05G102500	GO:0016791,GO:0008152	MLP-like protein 423	-9.06	0.0008
Manes.05G144100	GO:0009607,GO:0006952			
Manes.10G149500				

Table A5: Gene Ontologies (GO) for AR23.1, P40/1 and 98/0581 cassava genotypes following mealybug infestation categorized in Biological process, Molecular function and Cellular component.

GO term	Gene ID	GO ID	Gene annotation
Biological process			UDP-Glycosyltransferase superfamily protein
1. Metabolic process	Manes.02G054300	GO:0016758,GO:0008152	Cytochrome P450, family 77, subfamily A, polypeptide 5 pseudogene
	Manes.02G071200	GO:0008152,GO:0003824	Terpenoid cyclases/Protein
	Manes.02G086200	GO:0016829,GO:0010333,GO:0008152,GO:0000287	prenyltransferases superfamily protein
	Manes.12G076600	GO:0016829,GO:0010333,GO:0008152,GO:0000287	Terpene synthase 10
	Manes.16G016100	GO:0016829,GO:0010333,GO:0008152,GO:0000287	Terpene synthase 14
	Manes.S099100	GO:0016829,GO:0010333,GO:0008152,GO:0000287	Terpene synthase-like sequence-1,8-cineole

	Manes.01G231100	GO:0055114,GO:0016491	Tibberellin 2-oxidase 6
		GO:0055114,GO:0016491,GO:0006	
	Manes.01G274900	633,GO:0005506	Fatty acid hydroxylase superfamily
	Manes.03G032500	GO:0055114,GO:0016491	ACC oxidase 1
			2-oxoglutarate (2OG) and Fe(II)-
	Manes.04G120100	GO:0055114,GO:0016491	dependent oxygenase superfamily protein
		GO:0055114,GO:0020037,GO:0006	
	Manes.08G002700	979,GO:0004601	Tscorbate peroxidase 2
		GO:0055114,GO:0016491,GO:0051	
	Manes.10G071700	536,GO:0020037	Titrite reductase 1
		GO:0055114,GO:0020037,GO:0016	Tytochrome P450, family 705, subfamily
	Manes.10G146700	705,GO:0005506	A, polypeptide 27
		GO:0055114,GO:0020037,GO:0016	Tytochrome P450, family 94, subfamily
	Manes.14G051800	705,GO:0005506	B, polypeptide 3
	Manes.15G121600	GO:0055114,GO:0016491	Tibberellin 2-oxidase
		GO:0045156,GO:0019684,GO:0009	
	Manes.S113100	772	Thotosystem II reaction center protein A
		GO:0008270,GO:0005975,GO:0004	
	Manes.01G110800	476,GO:0009298	Mannose-6-phosphate isomerase, type I
		GO:0016984,GO:0015977,GO:0000	
	Manes.S113700	287	Tibulose-bisphosphate carboxylases
		GO:0045454,GO:0015035,GO:0006	
	Manes.01G192100	662	Thioredoxin H-type 7
			Glutamine-dependent asparagine synthase
	Manes.02G212800	GO:0006529,GO:0004066	1
		GO:0043565,GO:0006355,GO:0003	
	Manes.01G242300	700,GO:0003677	Homeobox 7
			Integrase-type DNA-binding superfamily
	Manes.03G044100	GO:0006355,GO:0003700	protein
			Ethylene responsive element binding
	Manes.06G131100	GO:0006355,GO:0003700	factor 1
		GO:0008270,GO:0006355,GO:0005	
		634,GO:0004402,GO:0003712,GO:0	
	Manes.13G112500	005515	BTB and TAZ domain protein 4
		GO:0045454,GO:0005681,GO:0000	mRNA splicing factor, thioredoxin-like
	Manes.03G120700	398	U5 snRNP
2. Cellular process	Manes.01G067400	GO:0006355,GO:0003700	C-repeat-binding factor 4
			Ethylene responsive element binding
	Manes.01G085400	GO:0006355,GO:0003700	factor 1
	Manes.02G089000	GO:0006355,GO:0003677	NAC domain containing protein 42
			Integrase-type DNA-binding superfamily
	Manes.03G056100	GO:0006355,GO:0003700	protein
			Ethylene responsive element binding
	Manes.06G131100	GO:0006355,GO:0003700	factor 1
		GO:0008270,GO:0006355,GO:0005	
		634,GO:0004402,GO:0003712,GO:0	
	Manes.13G112500	005515	BTB and TAZ domain protein 4

	Manes.03G044100	GO:0006355,GO:0003700 GO:0045454,GO:0015035,GO:0009055	Integrase-type DNA-binding superfamily protein Glutaredoxin family protein
	Manes.15G015600		D-mannose binding lectin protein with Apple-like carbohydrate-binding domain
	Manes.15G094100	GO:0048544	Cytochrome C assembly protein
	Manes.07G086000	GO:0020037,GO:0017004 GO:0008017,GO:0007018,GO:0005524,GO:0003777,GO:0005871	ATP binding microtubule motor family protein
	Manes.05G132500		Late embryogenesis abundant 3 (LEA3) family protein
3. Response to stress	Manes.05G061300	GO:0006950 GO:0055114,GO:0020037,GO:0006979,GO:0004601	Peroxidase superfamily protein
	Manes.01G065500	GO:0055114,GO:0020037,GO:0006979,GO:0004601	Peroxidase superfamily protein
	Manes.02G025800	GO:0055114,GO:0020037,GO:0006979,GO:0004601	Peroxidase superfamily protein
	Manes.02G026300	GO:0055114,GO:0020037,GO:0006979,GO:0004601	Peroxidase superfamily protein
	Manes.08G002700	GO:0009378,GO:0006310,GO:0006281	Escorbate peroxidase 2
	Manes.02G015400	GO:0030983,GO:0006298,GO:0005524	P-loop containing nucleoside triphosphate hydrolases superfamily protein
	Manes.01G020600	GO:0030983,GO:0006298,GO:0005524	MUTS homolog 2
	Manes.01G082400		MUTS-like protein 4
	Manes.02G106500	GO:0016021,GO:0006952	Seven transmembrane MLO family protein
	Manes.15G130200	GO:0009607,GO:0006952	Polyketide cyclase/dehydrase and lipid transport superfamily protein
	Manes.15G130300	GO:0009607,GO:0006952	MLP-like protein 43
	Manes.13G141100	GO:0016021,GO:0006952	Seven transmembrane MLO family protein
	Manes.15G130000	GO:0009607,GO:0006952	MLP-like protein 28
	Manes.17G079800	GO:0009607,GO:0006952	MLP-like protein 28
	Manes.08G135200	GO:0050832,GO:0042742 GO:0008061,GO:0050832,GO:0042742	Pathogenesis-related 4
	Manes.08G135300	742	Pathogenesis-related 4
Molecular function			Pytochrome P450, family 77, subfamily A, polypeptide 5 pseudogene
1. Catalytic activity	Manes.02G071200	GO:0008152,GO:0003824 GO:0003824,GO:0006470,GO:0004722	Highly ABA-induced PP2C gene 3
	Manes.02G128000	722	Camelliol C synthase 1
	Manes.17G070900	GO:0003824 GO:0016829,GO:0010333,GO:0000287	Terpene synthase 21
	Manes.08G138500	287	Terpene synthase 10
	Manes.12G076600	152,GO:0000287 GO:0016829,GO:0010333,GO:0000287	Terpene synthase 14
	Manes.16G016100	152,GO:0000287	

		GO:0016984,GO:0015977,GO:0000	
	Manes.S113700	287	Ribulose-bisphosphate carboxylases
	Manes.17G116800	GO:0008270,GO:0004089	Carbonic anhydrase 1
		GO:0016829,GO:0010333,GO:0008	terpene synthase-like sequence-1,8-
	Manes.S099100	152,GO:0000287	cineole
		GO:0016887,GO:0005524,GO:0016	
	Manes.03G018600	020	Pleiotropic drug resistance 12
			Eukaryotic aspartyl protease family
	Manes.02G039000	GO:0006508,GO:0004190	protein
			Eukaryotic aspartyl protease family
	Manes.02G145900	GO:0006508,GO:0004190	protein
			Eukaryotic aspartyl protease family
	Manes.16G132600	GO:0006508,GO:0004190	protein
	Manes.07G059300	GO:0005975,GO:0004553	Glycosyl hydrolase superfamily protein
	Manes.01G076500	GO:0004575,GO:0004564	Glycosyl hydrolases family 32 protein
		GO:0042545,GO:0030599,GO:0005	
2. Binding	Manes.01G231200	618	Pectin lyase-like superfamily protein
		GO:0008270,GO:0005975,GO:0004	
	Manes.01G110800	476,GO:0009298	Mannose-6-phosphate isomerase, type I
		GO:0008270,GO:0006355,GO:0005	
		634,GO:0004402,GO:0003712,GO:0	
	Manes.13G112500	005515	BTB and TAZ domain protein 4
		GO:0006468,GO:0005524,GO:0004	
	Manes.01G013000	672	Wall-associated kinase family protein
		GO:0030983,GO:0006298,GO:0005	
	Manes.01G020600	524	MUTS homolog 2
		GO:0006355,GO:0005634,GO:0005	
	Manes.01G070800	524,GO:0032502,GO:0006351	Growth-regulating factor 2
		GO:0005509,GO:0006468,GO:0005	
	Manes.02G170200	524,GO:0004672	Wall associated kinase 3
		GO:0006468,GO:0005524,GO:0004	Protein kinase protein with adenine
	Manes.02G172500	672	nucleotide alpha hydrolases-like domain
		GO:0016887,GO:0005524,GO:0016	
	Manes.03G018600	020	Pleiotropic drug resistance 12
			protein serine/threonine kinases;protein
		GO:0006468,GO:0005524,GO:0004	Kinases;ATP binding;sugar
	Manes.03G076800	672	binding;kinases;carbohydrate binding
		GO:0008017,GO:0007018,GO:0005	ATP binding microtubule motor family
	Manes.05G132500	524,GO:0003777,GO:0005871	protein
		GO:0016887,GO:0005524,GO:0016	
	Manes.06G154000	020	ABC-2 type transporter family protein
		GO:0016887,GO:0005524,GO:0016	White-brown complex homolog protein
	Manes.10G105100	020	11
			Integrase-type DNA-binding superfamily
	Manes.03G056100	GO:0006355,GO:0003700	protein
			Ethylene responsive element binding
	Manes.06G131100	GO:0006355,GO:0003700	factor 1

		GO:0016829,GO:0010333,GO:0000	
	Manes.08G138500	287	Terpene synthase 21
		GO:0016829,GO:0010333,GO:0008	
	Manes.12G076600	152,GO:0000287	Terpene synthase 10
		GO:0016829,GO:0010333,GO:0008	
	Manes.16G016100	152,GO:0000287	Terpene synthase 14
		GO:0016829,GO:0010333,GO:0008	Terpene synthase-like sequence-1,8-
	Manes.S099100	152,GO:0000287	cineole
		GO:0016984,GO:0015977,GO:0000	
	Manes.S113700	287	Ribulose-bisphosphate carboxylases
		GO:0055114,GO:0020037,GO:0016	Cytochrome P450, family 705, subfamily
	Manes.10G146700	705,GO:0005506	A, polypeptide 27
		GO:0055114,GO:0020037,GO:0016	Cytochrome P450, family 94, subfamily
	Manes.14G051800	705,GO:0005506	B, polypeptide 3
		GO:0055114,GO:0020037,GO:0006	
	Manes.08G002700	979,GO:0004601	Ascorbate peroxidase 2
	Manes.17G098200	GO:0020037,GO:0019825	Hemoglobin 1
		GO:0008270,GO:0006355,GO:0005	
Cellular component		634,GO:0004402,GO:0003712,GO:0	
1. Cell	Manes.13G112500	005515	BTB and TAZ domain protein 4
	Manes.14G006300	GO:0005515	CCT motif family protein
	Manes.14G096200	GO:0005515	RNI-like superfamily protein
		GO:0016021,GO:0015116,GO:0008	
		272,GO:0055085,GO:0016020,GO:0	
	Manes.03G066400	008271,GO:1902358	Sulfate transporter 3;5
		GO:0071805,GO:0016020,GO:0015	
	Manes.04G094700	079	K ⁺ uptake transporter 3
		GO:0016887,GO:0005524,GO:0016	
	Manes.06G154000	020	ABC-2 type transporter family protein
		GO:0016020,GO:0006810,GO:0005	
	Manes.09G062300	215	Gamma tonoplast intrinsic protein
		GO:0016887,GO:0005524,GO:0016	
	Manes.10G105100	020	white-brown complex homolog protein 11
	Manes.01G086900	GO:0055085,GO:0016021	Major facilitator superfamily protein
	Manes.01G161800	GO:0055085,GO:0016021	Cation exchanger 3
		GO:0016021,GO:0016020,GO:0022	Nodulin MtN21 /EamA-like transporter
	Manes.02G040000	857	family protein
			Seven transmembrane MLO family
	Manes.02G106500	GO:0016021,GO:0006952	protein
		GO:0008270,GO:0006355,GO:0005	
		634,GO:0004402,GO:0003712,GO:0	
	Manes.13G112500	005515	BTB and TAZ domain protein 4
		GO:0008017,GO:0007018,GO:0005	ATP binding microtubule motor family
	Manes.05G132500	524,GO:0003777,GO:0005871	protein
		GO:0045454,GO:0005681,GO:0000	mRNA splicing factor, thioredoxin-like
	Manes.03G120700	398	U5 snRNP
		GO:0042545,GO:0030599,GO:0005	
	Manes.02G039500	618	Pectin lyase-like superfamily protein

	GO:0055085,GO:0016020,GO:0015	
Manes.01G067000	297,GO:0015238,GO:0006855	MATE efflux family protein Uncharacterized protein family
Manes.01G108600	GO:0016020	(UPF0016)
Manes.01G254100	GO:0016757,GO:0016020	Protein of unknown function (DUF604)
	GO:0055085,GO:0016020,GO:0015	
Manes.02G032300	297,GO:0015238,GO:0006855	MATE efflux family protein
	GO:0016021,GO:0016020,GO:0022	Nodulin MtN21 /EamA-like transporter
Manes.02G040000	857	family protein
	GO:0055085,GO:0016020,GO:0006	Cyclic nucleotide-regulated ion channel
Manes.02G051000	811,GO:0005216	family protein
	GO:0033177,GO:0015991,GO:0015	
Manes.S091000	078	ATP synthase subunit C family protein
	GO:0055085,GO:0016020,GO:0006	Cyclic nucleotide-regulated ion channel
Manes.02G051100	811,GO:0005216	family protein
	GO:0008270,GO:0006355,GO:0005	
	634,GO:0004402,GO:0003712,GO:0	
Manes.13G112500	005515	BTB and TAZ domain protein 4
	GO:0045454,GO:0005681,GO:0000	mRNA splicing factor, thioredoxin-like
Manes.03G120700	398	U5 snRNP
	GO:0008017,GO:0007018,GO:0005	ATP binding microtubule motor family
Manes.05G132500	524,GO:0003777,GO:0005871	protein
	GO:0042545,GO:0030599,GO:0005	
Manes.02G039500	618	Pectin lyase-like superfamily protein

Table A6: An example sample RT-qPCR data for analysis. In this data set, housekeeping genes (UBQ10 and GTPb) and genes of interest are represented by three replicates at 24 and 72 hpi compared to mock infested. The Ct values of the housekeeping genes UBQ10 and GTPb were consistent across the treatments.

Sample name	Cq value	Cq mean	Cq error	Slope
Control UBQ10	32.15	31.99	0.537	0.015
	32.13	31.99	0.537	0.1641
	31.69	31.99	0.537	0.0162
Control GTPB	24.46	24.53	0.070	0.8978
	24.54	24.53	0.070	0.9802
	24.60	24.53	0.070	0.5962
Control CYP	28.58	28.523	0.051	0.1193
	28.48	28.523	0.051	0.5187
	28.51	28.523	0.051	0.5726
Control MYB	33.10	33.466	0.318	0.6058
	33.67	33.466	0.318	0.3682
	33.63	33.466	0.318	0.5441
Control 2OG	24.85	24.846	0.185	0.5448
	24.66	24.846	0.185	0.8927
	25.03	24.846	0.185	0.9892
24HR UBQ10	31.82	28.523	6.125	0.172
	32.05	28.523	6.125	0.4189
	32.05	28.523	6.125	0.0923

24HR GTPB	24.12	24.84	0.687	0.7839
	25.49	24.84	0.687	0.6003
	24.91	24.84	0.687	0.5863
24HR CYP	28.56	28.91	0.325	0.6435
	28.98	28.91	0.325	0.527
	29.2	28.91	0.325	0.546
24HR MYB	34.64	34.39	0.390	0.2773
	34.59	34.39	0.390	0.3198
	33.94	34.39	0.390	0.3724
24HR 2OG	25.24	25.02	1.011	0.5642
	25.01	25.02	1.011	0.0093
	24.81	25.02	1.011	1.178
72HR UBQ10	31.13	31.00	4.558	0.2899
	30.60	31.00	4.558	0.2171
	31.29	31.00	4.558	0.192
72HR GTPB	24.59	24.36	0.268	0.8129
	24.25	24.36	0.268	0.8366
	24.26	24.36	0.268	0.8279
72HR CYP	33.43	33.67	11.77	0.6291
	33.4	33.67	11.77	0.5579
	34.22	33.67	11.77	0.5063
72HR MYB	33.6	33.99	1.320	0.5069
	34.2	33.99	1.320	0.006
	33.91	33.99	1.320	0.006
72HR 2OG	34.05	33.73	2.623	0.2736
	33.82	33.73	2.623	0.0057
	33.34	33.73	2.623	0.4296