Molecular responses of maize to its foliar pathogen, *Cercospora zeina*

By

Jeanne Nicola Korsman

Submitted in partial fulfilment of the requirements for the degree

PHILOSOPHIAE DOCTOR

In the Department of Plant Science

Faculty of Natural & Agricultural Sciences

University of Pretoria

Pretoria

July 2015

Supervisor: Prof D. K. Berger

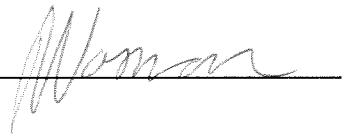
Co-supervisor: Dr. B. G. Crampton

in memory of Darko Tonžetić
19.06.1984 – 05.09.2008

You told me I should… and I did it…

Declaration

I, Jeanne Nicola Korsman declare that the thesis, which I hereby submit for the degree *Philosophiae Doctor* at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

SIGNATURE: 

DATE: 9/7/2015

Table of contents

Declaration ii
Acknowledgements vii
Preface viii
Abstract xii
List of Figures xiii
List of Tables xvi
List of abbreviations xviii

Chapter 1:

Literature Review: The maize - *Cercospora zeina* conflict 1
1.1 The value of maize 1
1.2 Maize pathogens 2
1.3 Grey Leaf Spot 3
1.4 Quantitative disease resistance 5
1.5 Plant-pathogen defence 6
1.5.1 Structural and preformed immunity 7
1.5.2 Innate immunity 7
1.5.2.1 PAMP-triggered immunity 7
1.5.2.2 Effector triggered immunity 8
1.5.2.3 Signalling in defence response 9
1.5.2.3.1 Plant hormones in defence signalling 10
1.5.3 Induced plant defence responses 11
1.5.3.1 Pathogenesis related proteins 12
1.5.3.2 Systemic acquired resistance 13
1.5.3.3 Reactive oxygen species 13
1.5.3.4 Hypersensitive response 14
1.5.3.5 Secondary antimicrobial compounds 15
1.5.3.6 Papillae 15
1.6 Conclusions 16
1.7 References 17

Chapter 2:

Quantitative phenotyping of grey leaf spot disease in maize using real-time PCR 27

Abstract 28

2.1 Introduction 29

2.2 Materials and Methods 31

2.2.1 Biological material and host plant infection 31

2.2.2 Digital image analysis 32

2.2.3 DNA extraction from fungi and plants 32

2.2.4 Conventional and quantitative PCR amplification 32

2.3 Results 35

2.3.1 Design of a specific PCR assay for *Cercospora* spp. in infected maize leaves 35

2.3.2 qPCR method to quantify *Cercospora* spp. in infected maize tissue 37

2.3.3 qPCR quantification of *C. zeina* DNA in glasshouse inoculated maize leaves before and after lesion development 39

2.3.4 *C. zeina* fungal content in GLS resistant and susceptible maize lines 40

2.3.5 Implementation of qPCR assay to differentiate between *C. zeae-maydis* and *C. zeina* 41

2.4 Discussion 44

2.5 References 47

Chapter 3:

Comparison of QTL for pathogen content in leaves measured by qPCR assay with QTL for whole plant disease phenotype in the GLS maize pathosystem 49

Abstract 50

3.1 Introduction 51

3.2 Materials and Methods 53

3.2.1 Field trials and maize germplasm 53

3.2.2 GLS disease scores 53

3.2.3 Determination of pathogen content in leaves 53

3.2.3.1 Molecular quantification of *C. zeina* by qPCR assay 53

3.2.3.2 Digital image analysis 54

3.2.4 Genetic map 55

3.2.5 QTL Analysis 55

3.2.6 Statistical analysis 56

3.3 Results 57

3.3.1 GLS disease assessment and quantitative phenotyping by qPCR assay 57
3.3.2 Frequency distributions 58
3.3.3 Sources of variation in disease quantification methods 58
3.3.4 Correlation of GLS disease assessment and *C. zeina* quantification values 60
3.3.5 Genetic map 61
3.3.6 Identification of QTL for qPCR assay and lesion area 62
3.4 Discussion 68
3.5 References 74

Chapter 4:

Candidate gene discovery using a bulked segregant approach to expression profiling 78
Abstract 79
4.1 Introduction 80
4.2 Materials and methods 82
4.2.1 Biological material 82
4.2.2 Choice of RILs for expression profiling 82
4.2.3 RNA extractions 83
4.2.4 Microarray 85
4.2.4.1 Experimental design 85
4.2.4.2 Sample preparation 85
4.2.4.3 Hybridisation and data capture 86
4.2.5 Microarray data analysis 87
4.2.6 RNAseq expression profiling 88
4.2.7 Annotation 88
4.2.8 Gene ontologies 89
4.2.9 Association of genes with QTL 89
4.3 Results 91
4.3.1 Expression profiling of resistant and susceptible bulked RILs 91
4.3.1.1 Microarray analysis 92
4.3.1.2 RNA-seq analysis 94
4.3.2 Over representation analysis 97
4.3.3 Candidate genes for resistance model 97
4.3.4 Candidate genes for QTL 102
4.3.5 Microarray and RNA-seq comparison 104

4.4 Discussion 108
4.5 References 118

Chapter 5:
Concluding remarks 126
References 131

Appendix A: Quality control for microarray and RNA-seq 133
Appendix B: GO terms and gene lists from Microarray and RNA-seq 140

Acknowledgements

I wish to thank my supervisor, Prof Dave Berger, for his intellectual input, guidance, and supervision throughout my PhD.

My thanks also go to my co-supervisor, Dr Bridget Crampton, for her intellectual input, guidance, friendship, and all the cappuccinos.

The input from the members of the "Maize eQTL" collaborators is also greatly appreciated. Special thanks go to Dr Barbara Meisel, Dr Shane Murray, Anelda van der Walt, Dr Boney Kuriakose, Dr Maryke Carstens, Dr Sonia Phillips, Adrie Veale, Grieta Mahlangu, Dr Nannette Christie, Thomas Schmidt, Monique Heystek, Jacqueline Meyer, Felix Middleton, Dr Rikkus Kloppers, Prof P Tongoona, and Prof A Myburg.

To my lab mates, present and past, for their help, support, encouragement and friendship, and for putting up with my antics, a big thank you! Thanks to Miekie Haasbroek, Velushka Birkenbach and Dr Maryke Carstens for proof reading parts of this manuscript. And thanks to Dr Barbara Meisel, Dr Sonia Phillips, Dr Maryke Carstens, Erika Viljoen and Mischa Muller for emotional support. Also, thank you to all other FABIans for their help and friendship, especially to Dr Shuai-Fei Chen.

My thanks also go to Nicky Olivier for help with microarray analysis, as well as keeping the lab running, and to René Swart and all the "behind the scenes" people in the Plant Science department for keeping the department running and making it such a pleasant one.

I am very grateful to my parents, Shirley and Nico Korsman, for their love, support and understanding, and my brother, Stephen, for a place to write without distractions. To Yvette Tonžetić for her support and encouragement. To my friends for their support, encouragement, and for keeping me sane, there are too many of you to mention, I love you all. My other "family", Janet, Jac, Sebastian and Benjamin Steenekamp, for wake up calls and making sure I laugh and eat. And to God for infinite blessings.

I wish to thank the Technology Innovation Agency (TIA), the Department of Science and Technology, South Africa, and PANNAR SEED (Pty) Ltd for funding the project. And TIA, for personal financial support.

The financial assistance of the Technology Innovation Agency (TIA) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to TIA.

Preface

Maize is an important food crop in South Africa. It is vulnerable to Grey Leaf Spot (GLS), a fungal leaf disease caused by two related species, namely *Cercospora zeae-maydis* and *Cercospora zeina* (Crous et al. 2006). *C. zeina* is the causative agent of GLS in South Africa (Crous et al. 2006; Meisel et al. 2009). GLS can have disastrous effects on maize crop yield. Conventional measures used to protect maize crops against GLS are fungicide sprays, conventional tillage and breeding of resistant hybrids (Ward and Nowell 1998). Quantitative resistance is important for breeding maize with durable resistance to *C. zeina*, as quantitative resistance is more robust and durable than single genes conferring resistance (Lindhout 2002; Parlevliet 2002). Thus we set out to study molecular aspects involved in quantitative resistance to GLS. First, a fungal quantification assay was developed. The assay was then used to obtain data for mapping QTL controlling *in planta* fungal quantity from a maize field trial of a RIL population with a range of GLS resistance phenotypes. Expression profiling was carried out on the *C. zeina* infected maize plants from that trial to identify genes involved in the resistance response to *C. zeina*.

The work carried out for this thesis constituted part of a greater collaborative project, Genomics of Quantitative Disease Resistance in African Maize Varieties, between the University of Pretoria, the Centre for Proteomic and Genomic Research (CPGR), the University of KwaZulu-Natal, and PANNAR SEED (Pty) Ltd. The project was funded by the Technology Innovation Agency (TIA), the Department of Science and Technology, South Africa, and PANNAR SEED (Pty) Ltd.

Each research chapter is presented as a publishable section written in the format of the European Journal of Plant Pathology. Content of the chapters is described below.

Chapter 1: The maize - *Cercospora zeina* conflict
This is a literature review focussing on maize, its interaction with *C. zeina*, and pathogen defence mechanisms in maize and other monocots.

Chapter 2: Quantitative phenotyping of grey leaf spot disease in maize using real-time PCR
The objective of Chapter 2 was to develop a qPCR assay to measure the amount of *Cercospora* sp. DNA in infected maize leaves. A fungal gene fragment was amplified using real-time PCR

to detect the fungal biomass; this was normalized against a real-time PCR amplified maize *gst* gene fragment. The assay proved to be sensitive, with a detection limit of 5 pg *C. zeina* or *C. zeae-maydis* DNA. The assay also proved to be specific; it did not amplify fragments from a range of other maize foliar pathogens. This chapter was published in the European Journal of Plant Pathology in 2012.

Chapter 3: Comparative analysis of QTL derived from fungal biomass and whole plant disease phenotype in the GLS maize pathosystem

In this Chapter we aimed to find QTL in the maize genome involved in limiting the growth of *C. zeina*, and to determine whether such QTL overlap with GLS resistance QTL in the same population. The CML444 × SC Malawi maize RIL population exhibits a range of GLS resistance phenotypes including more extreme phenotypes than the resistant parent, CML444 and the susceptible parent SC Malawi. The qPCR assay developed in Chapter 2 was used to determine the *in planta* content of *C. zeina* in a subset of the RILs. QTL controlling the growth of *C. zeina* within leaves were mapped for the RIL population and one of these QTL overlapped with a GLS disease resistance QTL from the same RIL population and environment. This demonstrates that GLS resistance QTL can be mapped using qPCR data. The chapter will be submitted to BMC Genetics for publication.

Chapter 4: Candidate gene discovery using a bulked segregant approach to expression profiling

In Chapter 4 we aimed to identify candidate genes that are differentially expressed between resistant and susceptible RILs during the later stages of *C. zeina* infection of maize. We also wanted to determine if any of these genes are located within the boundaries of GLS resistance QTL, which could indicate that an allelic difference in the gene or promoter may contribute to the QTL effect. Expression profiling of a bulk of resitant RILs and a bulk of susceptible RILs, chosen for their opposing alleles at GLS resistance QTL, was used to identify candidate genes involved in the resistance phenotype. A number of candidate genes with possible roles in the resistance phenotype coincided with GLS resistance QTL. Further study of the genes identified here will need to be carried out in future projects to confirm their role in resistance and their functions.

Chapter 5: Concluding remarks

Here I discuss the results and conclusions of the research, the limitations of the study, and propose further steps to confirm and characterize the candidate genes identified from the study.

Parts of the thesis that have appeared in publications:

Chapter 2 has been published:

- Korsman, J., Meisel, B., Kloppers, F., Crampton, B., & Berger, D. (2012). Quantitative phenotyping of grey leaf spot disease in maize using real-time PCR. *European Journal of Plant Pathology*, *133*, 461–471.

The modified genetic map from Chapter 3 was used in

- Berger, D. K., Carstens, M., Korsman, J. N., Middleton, F., Kloppers, F. J., Tongoona, P., & Myburg, A. A. (2014). Mapping QTL conferring resistance in maize to gray leaf spot disease caused by *Cercospora zeina*. *BMC Genetics*, *15*, 60.

This paper was published before the beginning of the PhD study, but laid the groundwork for the qPCR assay and provided knowledge of the pathogen causing GLS in the field trial used in Chapters 3 and 4. I contributed to the maintainence of *C. zeina* cultures, the inoculation of maize plants in the glasshouse, and the DNA extractions, PCR and sequencing.

- Meisel, B., Korsman, J., Kloppers, F. J., & Berger, D. K. (2009). Cercospora zeina is the causal agent of grey leaf spot disease of maize in southern Africa. *European Journal of Plant Pathology*, *124*, 577–583.

International congress poster (Presenting author: DK Berger):

- Korsman, J. N., Meisel, B., Crampton, B. G., Kloppers, F. J., Schmidt, T. G., & Berger, D. K. (2012). Precision phenotyping gray leaf spot disease using real-time PCR and digital image analysis, 52nd Maize Genetics Conference, Portland, Oregon, USA

National congress posters:

- Korsman, J., Meisel, B., Kloppers, F. J., & Berger, D. K. (2011). Quantification and species differentiation of *Cercospora zeina* and *Cercospora zeae-maydis in planta* using quantitative PCR, 47th Congress of the Southern African Society For Plant Pathology, Berg-en-Dal, Kruger National Park, South Africa

- Korsman, J. N., Schmidt, T. G., Meisel, B., Kloppers, F. J., Crampton, B.G. & Berger D. K. (2012). Comparison between grey leaf spot lesion area and *Cercospora zeina* DNA content

within maize leaves, South African Association of Botanists (SAAB), 38th Annual Conference, University of Pretoria

Regional congress posters:

- Korsman, J., Meisel, B., Kloppers, F. J., & Berger, D. K. (2010). Quantification and species differentiation of *Cercospora zeina* and *Cercospora zeae-maydis in planta* using quantitative PCR, ACGT Regional Plant Biotechnology Forum, FABI, University of Pretoria

Oral congress presentation (Presenting author: DK Berger):

- Berger, D. K., Carstens, M., Korsman, J. N., Middleton, F., Kloppers, F. J., Tongoona, P., & Myburg, A. A. (2014). QTL for resistance to grey leaf spot disease in a sub-tropical population of maize from multiple seasons and field sites, SASBi-SAGS JOINT 2014 CONGRESS, Kwalata Game Lodge, Pretoria

References:

Crous, P. W., Groenewald, J. Z., Groenewald, M., Caldwell, P., Braun, U., & Harrington, T. C. (2006). Species of *Cercospora* associated with grey leaf spot of maize. *Studies in Mycology*, *55*, 189–197.

Lindhout, P. (2002). The perspectives of polygenic resistance in breeding for durable disease. *Euphytica*, *124*, 217–226.

Meisel, B., Korsman, J., Kloppers, F. J., & Berger, D. K. (2009). *Cercospora zeina* is the causal agent of grey leaf spot disease of maize in southern Africa. *European Journal of Plant Pathology*, *124*, 577–583.

Parlevliet, J. E. (2002). Durability of resistance against fungal, bacterial and viral pathogens; present situation. *Euphytica*, *124*, 147–156.

Ward, J. M. J., & Nowell, D. C. (1998). Integrated management practices for the control of maize grey leaf spot. *Integrated Pest Management Reviews*, *3*, 177–188.

Abstract

The defence response of maize (*Zea mays*) to its foliar pathogen, *Cercospora zeina*, is not well characterized at the molecular and genetic level. *C. zeina* causes grey leaf spot (GLS), and high infection levels result in reduced crop yield. A molecular genetic study of the interaction between maize and *C. zeina* required a method to quantify the fungus within infected leaves. A quantitative PCR assay was developed for this purpose. It was based on amplification of a fungal cytochrome P450 reductase (*cpr1*) gene fragment, which was normalized using an amplified maize glutathione S-transferase III (*gst*) gene fragment. The assay was specific to *C. zeina* and a related species, *Cercospora zeae-maydis*, not yet found in Africa. In addition, a melt curve analysis enabled discrimination between these species. There was no amplification from a range of other maize foliar pathogens. The fungal quantification assay was successfully tested on glasshouse grown maize inoculated with *C. zeina*, and infected field grown maize. This assay can be implemented to quantify *C. zeina* at early stages of infection due to its sensitivity, and in the field due to its specificity. The fungal quantification assay was used to map GLS resistance QTL from a field trial of a sub-tropical maize RIL population in KwaZulu-Natal. We hypothesized that QTL involved in limiting *C. zeina* growth maize in leaves could be detected using *in planta* fungal quantity and lesion area data from digital image analysis, and that they would correspond with GLS resistance QTL previously mapped in the same population. Three QTL were mapped for *in planta* fungal quantity and one for lesion area. The strongest effect QTL was located on chromosome six. It was detected using both the fungal quantification assay and the lesion area data, and overlapped with a GLS resistance QTL from the same population in the same environment. Thus QTL can be successfully mapped from *in planta* fungal quantification and lesion area data and these QTL correspond to GLS resistance QTL. The molecular response of maize to *C. zeina* was studied using expression profiling of a pooled bulk of resistant RILs versus a pooled bulk of susceptible RILs from the same population used for GLS QTL mapping. We aimed to find genes with differential expression between the resistant and susceptible bulks during GLS field infection. Additionally, we hypothesized that genes contributing to the QTL effect could be identified within QTL. Candidate genes for the resistance response included an *RPP13-like* gene and an *mlo* gene. The genome positions of the differentially expressed genes were compared to the genome positions of the QTL. Candidate genes coinciding with QTL included a malectin containing RLK, an *EDR1-like* gene, a *GTPase* gene, a *cytochrome b561* gene, and a *chorismate synthase* gene. Comparison of gene ontologies from all the genes differentially expressed between the resistant and susceptible bulks indicated that cell death was a likely strategy for resistance. The resistant response is probably an early response, with a later and continuing response to biotic stress in the susceptible maize plants.

List of Figures

Figure 1.1: Comparison of maize production and its uses between the world and Africa. 1
Figure 1.2: A section of a maize leaf with GLS. 4

Figure 2.1: Alignment of a 164 bp fragment of *cpr1* from *C. zeae-maydis* and *C. zeina*. 35
Figure 2.2: PCR amplification of *cpr1* (a), *gst 3* (b) and ITS1/4 (c) gene fragments from *C. zeae-maydis* and *C. zeina* cultures. 36
Figure 2.3: Alignment of a 164 bp fragment of *cpr1* amplified with the CPR1_2 primer pair from *C. zeae-maydis, C. zeina* and *Cercospora* sp. 37
Figure 2.4: Standard curves used for qPCR assay to quantify *C. zeina* and *C. zeae-maydis in planta* .. 38
Figure 2.5: Photographs of leaves with symptoms of GLS from glasshouse and field grown maize that were used for GLS disease quantification. 39
Figure 2.6: Comparison of *C. zeina* infection in resistant and susceptible maize lines using (a) digital image analysis and (b) qPCR. 40
Figure 2.7: Melting temperature (Tm) graphs of 164 bp qPCR amplification products using CPR1_2 primers. 43
Figure 2.8: PCR-RFLP analysis of the 164 bp *cpr1* amplicon from *C. zeae-maydis* and *C. zeina*. 43

Figure 3.1: The range of GLS severity in the field. 57
Figure 3.2: Frequency distributions of GLS ratings, the qPCR assay scores and the lesion area scores from the Baynesfield 2009 field trial. 59
Figure 3.3: SSR markers used to fill regions of low marker density on the original genetic linkage map for the RIL population. 62
Figure 3.4: QMap 2.0, the genetic map based on the CML444 × SC Malawi RIL population. 63
Figure 3.5: Partial genetic linkage map of the CML444 × SC Malawi RIL population indicating QTL. 67

Figure 4.1: Experimental design of the microarray experiment. 85
Figure 4.2: Leaf samples of the six resistant RILs and six susceptible RILs used in the bulks. 91
Figure 4.3: Over representation analysis of GO terms from genes with relatively higher expression in the resistant bulk. 98
Figure 4.4: Over representation analysis of GO terms from genes with relatively higher expression in the susceptible bulk. 99
Figure 4.5: MapMan representation of the RNA-seq-identified genes involved in biotic stress. 101
Figure 4.6: Scatter plots of the relative expressions measured in $\log_2$ fold change of the R bulk/S bulk compared between the RNA-seq and the microarray analysis. 105

Figure 4.7: Venn diagram showing overlap of GO terms associated with the genes with higher expression in (a) the resistant bulk and (b) the susceptible bulk from the RNA-seq and microarray analyses. 106
Figure 4.8: Venn diagram showing the numbers of genes observed with higher expression in the resistant bulk for the different technologies and their overlap. 106
Figure 4.9: Venn diagram showing the numbers of genes observed with higher expression in the susceptible bulk for the different technologies and their overlap. 106
Figure 4.10: The proposed mechanism(s) of resistance based on the genes found to have higher expression in the resistant bulk compared to the susceptible bulk, based on RNA-seq analysis. 111

Figure A1: Denaturing agarose gel showing total RNA extracted from maize leaves from a plant grown under controlled conditions in the glasshouse and three plants from the Baynesfield 2008/2009 field trial. 133
Figure A2: MA plots of the four arrays of the microarray experiment before and after within slide global lowess normalisation. 134
Figure A3: Sequence quality score plot for the bulked RNA-seq showing the quality scores per base of the read length. 135
Figure A4: Volcano plots of the relative expression of genes in the R bulk compared to the S bulk from the microarray experiment (a) and the RNA-seq experiment (b). 136

Figure B1: Over representation analysis of biological process GO terms from genes with relatively higher expression in the resistant bulk from the RNA-seq analysis. 140
Figure B2: Over representation analysis of molecular function GO terms from genes with relatively higher expression in the resistant bulk from the RNA-seq analysis. 141
Figure B3: Over representation analysis of biological process GO terms from genes with relatively higher expression in the susceptible bulk from the RNA-seq analysis. 142
Figure B4: Over representation analysis of cellular compartment GO terms from genes with relatively higher expression in the susceptible bulk from the RNA-seq analysis. 143
Figure B5: Over representation analysis of molecular function GO terms from genes with relatively higher expression in the susceptible bulk from the RNA-seq analysis. 144
Figure B6: Over representation analysis of molecular function GO terms from genes with relatively higher expression in the susceptible bulk from the microarray analysis. 145
Figure B7: Over representation analysis of molecular function GO terms from genes with relatively higher expression in the resistant bulk from the subset of RNA-seq genes which fall within the borders of the QTL. 146

Figure B8: Over representation analysis of molecular function GO terms from genes with relatively higher expression in the susceptible bulk from the subset of RNA-seq genes which fall within the borders of the QTL. 147
Figure B9: Over representation analysis of biological process GO terms from genes with relatively higher expression in the susceptible bulk from the subset of RNA-seq genes which fall within the borders of the QTL. 148

List of Tables

Table 1.1: Diseases of maize is South Africa. 3
Table 1.2: List of selected, known R-genes in monocots. 9
Table 1.3: Types of PR proteins. 12

Table 2.1: Melting temperatures of 164 bp qPCR products obtained with CPR1_2 primers from *Cercospora* spp. isolated from maize. 42

Table 3.1: Transgressive segregation of RIL population. 58
Table 3.2: Analysis of variance of *C. zeina* quantification by the qPCR assay. 60
Table 3.3: Comparison of GLS disease data as separate replicates (blocks) for 100 RILs sampled. 60
Table 3.4: Comparison of time of disease assessment between different GLS disease data. 61
Table 3.5: QTL identified from *C. zeina* qPCR assay and lesion area phenotypic data 65

Table 4.1: A table showing the marker data for the QTL of the six resistant and six susceptible bulked RILs. 83
Table 4.2: Top ten genes with the highest expression in the resistant bulk compared to the susceptible bulk (R/S) as determined by microarray analysis. 93
Table 4.3: Top ten genes with the highest expression in the susceptible bulk compared to the resistant bulk (S/R) as determined by microarray analysis. 94
Table 4.4: Top ten genes with the highest expression in the resistant bulk compared to the susceptible bulk (R/S) as determined by RNA-seq analysis. 95
Table 4.5: Top ten genes with the highest expression in the susceptible bulk compared to the resistant bulk (S/R) as determined by RNA-seq analysis. 96
Table 4.6: Genes only expressed in the Susceptible bulk, as determined by RNA-seq analysis. 96
Table 4.7: Differentially expressed genes observed in both the microarray and RNA-seq analysis that coincide with QTL. 103

Table A1: QC statistics for bulked microarray analysis. 137
Table A2: RNA-seq reads obtained for samples and mapping statistics. 138
Table A3: Summarizing statistics for FPKMs calculated using Cufflinks. 139

Table B1: Top 100 genes with higher expression in the resistant bulk compared to the susceptible bulk as determined by microarray analysis. 149
Table B2: Top 100 genes with higher expression in the susceptible bulk compared to the resistant bulk as determined by microarray analysis. 153

Table B3: Top 100 genes with higher expression in the resistant bulk compared to the susceptible bulk as determined by RNA-seq analysis. 156
Table B4: Top 100 genes with higher expression in the susceptible bulk compared to the resistant bulk as determined by RNA-seq analysis 159
Table B5a: Differentially expressed genes with higher expression in R bulk observed in microarray analysis that coincide with QTL. 162
Table B5b: Differentially expressed genes with higher expression in S bulk observed in microarray analysis that coincide with QTL. 163
Table B6a: Differentially expressed genes with higher expression in R bulk observed in RNA-seq analysis that coincide with QTL. 165
Table B6b: Differentially expressed genes with higher expression in S bulk observed in RNA-seq analysis that coincide with QTL. 167

List of abbreviations

ABA	abscisic acid
aRNA	amplified RNA
Avr	avirulence
CC-NB-LRR	coil-coiled nucleotide binding leucine rich repeat
cM	centi Morgan
cpr1	cytochrome P450 reductase 1 gene
Ct	cycle threshold
DAMP	damage-associated molecular patterns
dap	days after planting
DMSO	dimethylsulphoxide
dpi	days post inoculation
EDR1	enhanced disease resistance 1
ETI	effector-triggered immunity
FDR	false discovery rate
FPKM	kilobase of transcript per million fragments mapped
GABA	gamma-aminobutyric acid
gDNA	genomic DNA
GLS	grey leaf spot
GO	gene ontology
gst3	glutathione *S*-transferase III gene
H_2O_2	hydrogen peroxide
ITS	internal transcribed spacer
JA	jasmonic acid
KZN	KwaZulu-Natal
LOD	logarithm of odds
LRR	leucine rich repeat
lsm	least square means
MAMP	microbe-associated molecular pattern
MAPK	mitogen-activated protein kinase
NO	nitric oxide
PAMP	pathogen-associated molecular pattern
PR protein	pathogenesis related protein
PTI	PAMP-triggered immunity
qPCR	quantitative PCR
QTL	quantitative trait locus/loci
R bulk	pooled bulk of resistant RILs
R genes	resistance genes
rDNA	ribosomal DNA
RIL	recombinant inbred line
RLK	receptor like kinase
RNase	Ribonuclease
ROS	reactive oxygen species
RPP13	Resitance to *Peronospora parasitica* 13
S bulk	pooled bulk of susceptible RILs
SA	salicylic acid
SSR	simple sequence repeat

Chapter 1

Literature review: The maize - *Cercospora zeina* conflict

1.1 The value of maize

The decimation of the classic Maya civilization towards the end of the first millennium AD has been proposed to be due to repeated maize (*Zea mays*) crop failures caused by an epidemic of the maize mosaic virus (Brewbaker 1979). In order to avoid starvation, the Mayans migrated from what is now the northern region of Guatemala, where the maize yield losses were encountered, to surrounding regions where the virus and insect vector could not thrive.

Today maize, also known as corn, is still a staple food crop for many and is important for food security. It is the grain crop with the third highest production for human consumption in the world, with the total available calories from maize being 5% of the world's available calories. Rice and wheat each provide 19% of the available calories. However, in Africa, which produces approximately 7.2% of the world's maize, a far greater percentage is consumed as food rather than being used feed animal feed or in other industries (http://faostat3.fao.org/). In Figure 1.1, a comparison of maize production in Africa and the world is shown. Maize is still a staple food source and important for food security in Africa where it is grown both on a commercial scale and as a subsistence crop.

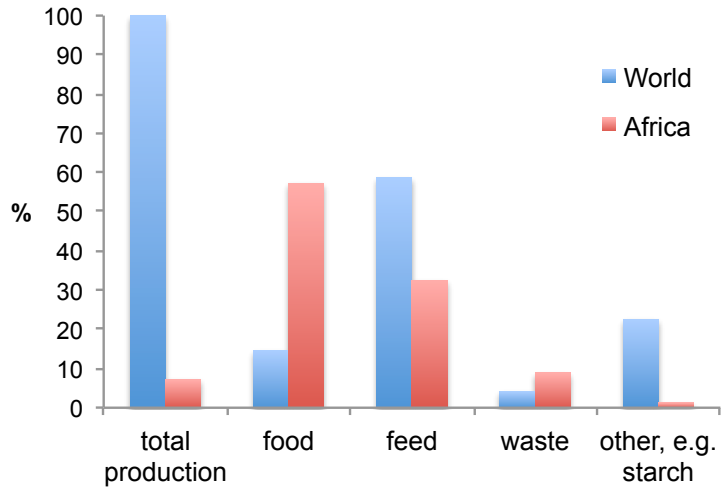


Figure 1.1: Comparison of maize production and its uses between the world and Africa. Graph compiled from data retrieved from http://faostat3.fao.org/.

The average yield of maize that is produced in Africa is far lower than the average worldwide yield. It is approximately 40% of the average worldwide yield, leaving a theoretical 60% for improvement. Many factors contribute to this lower yield. Subsistence farming produces relatively low yield due to the impact of environmental factors such as rainfall and soil quality. Weeds are not as effectively controlled on subsistence farms as on commercial farms and competition for nutrients may lead to decreased yields. Maize pests and pathogens, if not propoerly controlled can cause major declines in yield. In order to maximise maize yield in Africa, it is important to keep the pests and pathogens, as well as weeds, affecting maize yield under control (Sibiya et al. 2013).

One of the most common and effective ways of controlling weeds and maize pests and pathogens are chemical herbicide, insecticide and fungicide sprays. Plant based resistance to pests and pathogens is also greatly beneficial. Conventional breeding is the typical way of introducing resistance traits into crops. Genetically modified (GM) maize has been produced with tolerance to the herbicide Glyphosate for weed management (Roundup Ready®) and with proteins toxic to certain insects (Cry1Ab protein or *Bt*) for insect control, and most recently with drought tolerance (DroughtGard™). To date there is no commercial GM event for resistance to maize pathogens. However there have been a number of field trials for potential bacterial, fungal and viral resistant crops in the United States (Collinge et al. 2010). GM papaya with resistance to papaya ringspot virus has been grown in Hawaii since the field trials in 1995 (Ferreira et al. 2002).

1.2 Maize pathogens

There are a number of fungal, bacterial and viral pathogens that infect maize and cause yield loss in South Africa. Those that are of most concern to plant pathologists and maize breeders are listed in Table 1.1.

Out of the “top 10” fungal plant pathogens listed in the journal Molecular Plant Pathology, which were either chosen for their detrimental effects or their value as model organisms, six cause diseases of grain crops (Dean et al. 2012). *Magnaporthe oryzae* is a rice pathogen; *Puccinia graminis*, *Puccinia striiformis* and *Puccinia triticina* group together as major pathogens causing rust of wheat; *Fusarium graminearum* (teleomorph *Gibberella zeae*) causes diseases of all cereals and also produces mycotoxins; *Blumeria graminis* causes powdery mildew on wheat (*Triticum* spp.) and barley (*Hordeum vulgare*); *Mycosphaerella graminicola*

(anamorph *Septoria tritici*) a wheat pathogen; and *Ustilago maydis,* a maize pathogen, is known as a model organism of biotrophic plant pathogens.

Table 1.1 Diseases of maize is South Africa.

Disease	Pathogen
Foliar diseases	
Grey Leaf Spot	*Cercospora zeina*
Northern Corn Leaf Blight	*Exserohilum turcicum* (*Setosphaeria turcica*)
Diplodia leaf streak	*Stenocarpella macrospora*
Phaeosphaeria leaf spot	*Phaeosphaeria maydis*
Common rust	*Puccinia sorghi*
Polysora rust	*Puccinia polysora*
Eyespot	*Aureobasidium zeae*
Bacterial leaf streak	*Xanthamonas campestris* pv. *zeae* (bacteria)
Maize streak disease	Maize streak virus (transmitted by leafhoppers)
Maize stalk rots	
Stalk rot	*Fusarium verticillioides*
Gibberella stalk rot	*Fusarium graminearum* (*Gibberella zeae*)
Diplodia stalk rot	*Stenocarpella maydis*
Charcoal rot	*Macrophomina phaseolina*
Erwinia stalk rot	*Erwinia chrysanthemi* pv. *zeae* (bacteria)
Cob diseases	
Diplodia ear rot	*Stenocarpella maydis*
Fusarium ear rot	*Fusarium verticillioides*
Gibberella ear rot	*Fusarium graminearum* (*Gibberella zeae*)
Boil smut	*Ustilago maydis*
Cob and tassel smut	*Sphacelotheca reiliana*
Crazy top	*Sclerophthora macrospora*

Information obtained from http://www.pannar.com/diseases

Possible sources of new pathogen resistance for conventional breeding of maize lie in the wide variety of maize lines with an array of traits, including maize land races and teosinte. Teosinte, specifically *Zea mays* ssp. *parviglumis*, is widely considered to be the wild ancestor of maize, *Zea mays* ssp. *mays* (Doebley 2004). Some teosintes can hybridize with maize producing fertile progeny.

1.3 Grey Leaf Spot

Although Grey Leaf Spot (GLS) was not included in the top ten list, it can cause devastating yield losses of up to 60% in South Africa (Ward et al. 1997) in seasons with conditions favourable for disease development. GLS was first described in the USA and *Cercospora zeae-maydis* was identified as the causative agent (Tehon and Daniels 1925). It was more recently established that two closely related species, *Cercospora zeina* and *C. zeae-maydis,* are both responsible for GLS. While both species occur in the USA and Brazil (Wang et al. 1998;

Dunkle and Levy 2000; Goodwin et al. 2001; Brunelli et al. 2008), only *C. zeina* has been identified in South Africa (Crous et al. 2006; Meisel et al. 2009), and other African countries and China (Liu and Xu 2013; Okori et al. 2003).

GLS is a foliar disease of maize. The *C. zeina* conidia are wind dispersed either from old crop residue or nearby infected plants and only germinate under highly humid conditions (Ward et al. 1999) before entering the leaves via the stomata (Beckman and Payne 1982). The hyphae grow intercellularly (Kim et al 2011) and eventually kill off the host tissue, resulting in necrotic lesion formation. The symptoms are brownish-tan to grey lesions elongating within the borders of the major leaf veins (this can be visualised in Figure 1.2). The lesions become grey when conidiation occurs (Latterell and Rossi 1983) and conidiophores emerge through the stomata (Caldwell and Laing 2005). It can take from 14 to 28 days from infection to sporulation, with sporulation being closer to 28 days after infection in resistant plants (Beckman and Payne 1982; Ward et al. 1999). In severe infections the lesions coalesce leaving little living tissue available for photosynthesis, leading to limited grain filling, thus resulting in severe yield loss.

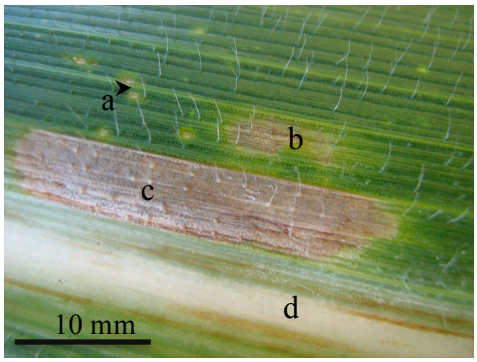


Figure 1.2: A section of a maize leaf with GLS lesions. Chlorotic spots can be seen where a lesion is beginning to develop (a), as well as a developing lesion (b), and a mature GLS lesion within the boundaries of the leaf veins (c). The leaf midrib is indicated by (d).

The occurrence of GLS in South Africa has increased in KwaZulu-Natal and since it was first observed in South Africa the pathogens has spread into neighbouring provinces . GLS has been has identified in other African countries, e.g. Zambia and Zimbabwe (Ward et al. 1999; Meisel et al. 2009). The most effective disease management practices for reducing GLS occurrence and severity are fungicide sprays, conventional tillage, crop rotation and utilization of resistant

hybrids (Ward and Nowell 1998; Ward et al. 1999). The most practical and cost effective of these for subsistence farmers in Africa would be resistant hybrids.

Multigene or quantitative resistance within resistant hybrids is more robust and durable and therefore more desirable than single genes conferring resistance (Lindhout 2002; Parlevliet 2002). Single gene resistance is more easily overcome by the pathogen (Lindhout, 2002; Parlevliet, 2002). GLS disease resistance is usually quantitative and several quantitative trait loci (QTL) for GLS resistance have been identified (Saghai Maroof et al. 1996; Clements et al. 2000; Lehmensiek et al. 2001; Gordon et al. 2004; Balint-Kurti et al. 2008; Juliatti et al. 2009; Pozar et al. 2009; Zhang et al. 2012; Berger et al. 2014; Benson et al. 2015). There are genes underlying these QTL, such as the maize wall associated kinase identified by Zuo et al. (2015) as being involved in a major maize head smut resistance QTL, and a glutathione *S*-transferase gene identified by Wisser et al. (2011) as being involved in resistance to GLS, northern corn leaf blight and southern leaf blight.

1.4 Quantitative disease resistance

Quantitative disease resistance is controlled by numerous genes (Churchill and Doerge 1994; Doerge et al. 1997) with the possibility of epistatic interactions between the genes (Tanksley 1993). Wisser et al. (2006) have summarised the mapping of quantitative disease resistance in maize looking at a total of twenty diseases, including GLS. They present eight clusters of disease resistance QTL where twice the number of QTL expected from gene density in the QTL regions were identified. These may represent multiple disease resistance loci.

There are numerous genomic regions populated with a multitude of genes involved in the response of maize to GLS. A number of GLS disease resistance QTL have been identified over a number of maize populations in different environments across continents. A number of meta-analyses of these have been conducted, the first on five populations (Wisser et al. 2006), the second on seven populations (Shi et al. 2007) and a third on thirteen populations (Berger et al. 2014). The most recent study by Benson et al. (2015) identified sixteen resistance QTL for GLS in a nested association mapping study, including seven novel QTL. They also identified a candidate detoxification-related gene, which could play a role in resistance, within one QTL region. As mentioned above, another detoxification-related gene, encoding a glutathione *S*-transferase, was found to be associated with resistance to GLS in a genome wide association study (Wisser et al. 2011).

Olukolu et al. (2014) reported a set of genes involved in the hypersensitive response of maize that are among the candidate genes for QTL hypersensitive response-related phenotypes identified in a genome wide association study of maize. The pathways implicated in the hypersensitive response as deduced by the observed genes included redox homeostasis, lignin biosynthesis, calcium signalling, programmed cell death, autophagy, ubiquitination and protein degradation, and interaction with R genes.

It would be interesting to elucidate other defence mechanisms underlying such resistance QTL. We should find the answer within the complex molecular mechanisms of the plant's immune system. Zuo et al. (2015) has achieved this for maize head smut by identifying a *ZmWAK* gene co-locating with a disease resistance QTL and conferring resistance to the fungus *Sporisorium reilianum*. These authors fine mapped a major resistance QTL to identify candidate genes and inserted the candidate *ZmWAK* into Hi-II maize, which resulted in increased resistance to *S. reilianum*.

1.5 Plant-pathogen defence

Most knowledge of the plant's defence system against pathogens has been gained from studies on dicots, with *Arabidopsis thaliana* being the favoured model system (Glazebrook et al. 1997). There has been some focus on monocots, especially the grain crops, with rice (*Oryza sativa*) having emerged as the model system (Chen and Ronald 2011). In this review we focus on monocot defence systems with particular reference to maize and the grain crops and as our interest lies in maize defence against *C. zeina*.

Non-host immunity is most likely based on structural or preformed immunity. For example, if *C. zeina* conidia land on a neighbouring field of soybean plants, the pathogen starts the infection process, but does not cause disease, probably because it is blocked before entering a cell. This is possibly due to the preformed immunity, or obstructions such as the formation of papillae, or an extremely efficient immune response (Lipka et al. 2008). Host immunity is based on an induced response after the pathogen enteris the cell or apoplast. If the plant succeeds in preventing disease, it is known as an incompatible interaction. If disease ensues, it is referred to as a compatible interaction (Hammond-Kosack and Parker 2003).

1.5.1 Structural and preformed immunity

Plant cuticles and cell walls form a physical barrier to pathogens as well as non-pathogens (Łaźniewska et al. 2012; Bigeard et al. 2015). Plants also produce preformed antimicrobial compounds, or phytoanticipins, including saponins, cyanogenic glycosides, and glucosinolates (Osbourn 1996). Such compounds are often toxic and are isolated in vacuoles or stored in an inactive form so as not to harm the plant itself. If the pathogen avoids contact with the active form it may bypass this defence mechanism and further defence mechanisms are required.

The benzoxazinoid hydroxamic acids are phytoanticipins often found in cereals (Niemeyer 2009). A number of studies in cereals have correlated benzoxazinoid hydroxamic acid levels with resistance to fungal pathogens, mostly in younger plants. The germination of spores of *Exserohilum turcicum* (teleomorph *Setosphaeria turcica*) is inhibited by a cyclic hydroxamate, 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOA), and maize seedlings without DIMBOA are more susceptible to infection by *E. turcicum* (Couture et al. 1971). However, many studies have found no effect of benzoxazinoid hydroxamic acids on resistance (Niemeyer 2009). The effect of such compounds on resistance will depend on the genotype and age of the plant as well as the pathogen involved.

1.5.2 Innate immunity

Plant defence in general has been extensively reviewed (Hammond-Kosack and Jones 1997; Hammond-Kosack and Parker 2003; Jones and Dangl 2006; Tena et al. 2011). The majority of what is known about pathogen defence in monocots has also been reviewed (Ayliffe and Lagudah 2004; Balmer et al. 2013a; Balmer et al. 2013b). Innate immunity starts with the recognition of the presence of a pathogen by pathogen-associated molecular pattern (PAMP)-triggered immunity **(**PTI) (Jones and Dangl 2006), also known as pattern-triggered immunity (Balmer et al. 2013b; Bigeard et al. 2015). Thereafter the pathogen responds by producing effectors, to which the host plants mount effector-triggered immunity (ETI) (Jones and Dangl 2006).

1.5.2.1 PAMP-triggered immunity

The plant needs to be able to perceive that a pathogen is present in order to initiate a defence response against it. This pathogen recognition begins with PAMPs or elicitors by plant receptor molecules. PAMPs are often referred to as microbe-associated molecular patterns (MAMPs) as the molecules that are perceived are present in non-pathogenic microbes as well. The classic

examples are flagellin from bacteria (Felix et al. 1999) and chitin from fungi (Felix et al. 1993). Another form of molecule perceived by plants are damage-associated molecular patterns (DAMPs), which comprise molecules from the host plant that result from the damage caused by the pathogen. Recognition of any of these compounds triggers PTI (Bigeard et al. 2015; Jones and Dangl 2006). PTI is fairly conserved between dicots and monocots (Balmer et al. 2013a).

A number of proteins, mostly receptor-like kinases (RLKs), have been identified in monocots that recognise PAMPs and induce the PTI defence process. A rice gene, *OsFLS2*, an orthologue of *AtFLS2* from Arabidopsis, encodes a receptor of flg22 (flagellin derived PAMP), which initiates a signalling pathway (Takai et al. 2008). A gene encoding ZmPep1, an orthologue of an Arabidopsis elicitor peptide 1 (AtPep1), is activated by *Cochliobolis heterostrophus*, jasmonic acid (JA), or ZmPep1 application (Huffaker et al. 2011a). Application of a synthesized ZmPep1 peptide induced JA and ethylene production, endochitinase, PR4, PRms (a homolog of tobacco PR1), and increased resistance of maize to southern leaf blight (*C. heterostrophus*) and anthracnose stalk rot (*Colletotrichum graminicola*). ZmPep1 seems to be endogenous elicitor, possibly a DAMP (Balmer et al. 2013a; Huffaker et al. 2011a).

1.5.2.2 Effector triggered immunity

Pathogens often produce molecules that assist in their virulence, by suppressing host defences. These virulence factors, called effectors, can be recognised by the plant which subsequently initiates ETI. In cases where effectors trigger an effective host immune response, they are termed avirulence factors. (Jones and Dangl 2006; Bigeard et al. 2015). There is what seems to be a "gene-for-gene" interaction, where an effector, or *Avr* gene product encoded by the pathogen and a resistance (R) gene encoded protein from the host plant interact to give rise to host resistance (Hammond-Kosack and Jones 1997). Cytoplasmic nucleotide-binding leucine-rich-repeat (NB-LRR) proteins are the most abundant class of R genes and are often involved in pathogen recognition (Van Der Biezen and Jones 1998; Jones and Dangl 2006). ETI often results in a downstream hypersensitive response (Greenberg and Yao 2004; Jones and Dangl 2006).

Gene-for-gene interactions have been have been characterised in various monocot-pathogen pathosystems. The interaction between maize and *Puccinia sorghi* is a good example of gene-

for-gene interaction, with the *rp1* locus conferring race-specific resistance to *P. sorghi*. Mutants at this locus can cause lesion mimics which have similar characteristics to the hypersensitive response (Hu et al. 1996). A gene encoding a CC-NB-LRR, with orthologues in maize and sorghum, has been recognized as conferring resistance to *E. turcicum* (Martin et al. 2011). The *mlo* gene, identified in barley, is a seven transmembrane protein which acts as a G-protein coupled receptor and confers resistance to powdery mildew (Büschges et al. 1997; Elliott et al. 2005). The wildtype *Mlo* gene is associated with susceptibility due to inhibition of cell death, but mutant *mlo* does not inhibit cell death, thus allowing resistance, possibly due to programmed cell death (Büschges et al. 1997).

There are a number of different types of R genes, some that have been identified in monocotyledonous species are listed in Table 1.2.

Table 1.2: List of selected, known resistance genes in monocots (adapted from Ayliffe and Lagudah (2004); Hammond-Kosack and Jones (1997); Hammond-Kosack and Parker (2003))

R-protein type	Gene	Plant	Pathogen	Reference
Detoxifying enzyme (HC toxin reductase)	*Hm1*	Maize	*Cochliobolus carbonum*	Johal and Briggs (1992)
Intracellular serine/threonine protein kinase	*Rpg1*	Barley	*Puccinia graminis*	Brueggeman et al. (2002)
CC-NB-LRR Intracellular protein	*Mla1/ Mla6*	Barley	*Blumeria graminis*	Zhou et al. (2001) Halterman et al. (2001)
	Rp1 *Rp1-D*	Maize	*Puccinia sorghi*	Hu et al. (1996) Collins et al. (1999)
	Rp3	Maize	*Puccinia sorghi*	Webb et al. (2002)
NB-LRD (nucleotide binding site-leucine rich domain)	*Pi-ta*	Rice	*Magnaporthe grisea*	Bryan et al. (2000)
Extracellular protein with single trans-membrane region and cytoplasmic kinase domain	*Xa21*	Rice	*Xanthomonas oryzae*	Song et al. (1995)
G protein coupled receptor with seven transmembrane regions	*mlo*	Barley	*Blumeria graminis*	Büschges et al. (1997)

1.5.2.3 Signalling in defence response

Pathogen perception via PTI or ETI results in activation of a multitude of signalling pathways. The signalling pathways/networks result in various defence mechanisms and are usually initiated by recognition of a pathogen by a RLK, which can be cell wall associated or cytoplasmic. The ensuing signalling kinase cascades result in transcriptional activation of defence leading to resistance or failure of resistance (Hammond-Kosack and Parker 2003).

An early event of PTI is the influx of calcium ions into the cell, possibly activated by MAPK signalling (Tena et al. 2011; Bigeard et al. 2015). The Ca^{2+} influx itself induces further mechanisms. The Ca^{2+} influx can begin a cascade leading to cell death, but can also be found downstream of oxidative burst (Levine et al. 1996). Calcium ions directly activate calcium-dependent protein kinases (CDPKs), and can act upon downstream gene expression or further signalling pathways (Harmon et al. 2000). A CDPK gene identified in maize, *ZmCPK10*, is transcriptionally activated by inoculation with spores from *Fusarium moniliforme* (also known as *Gibberella fujikuroi*) as well as fungal elicitor application (Murillo et al. 2001). Mitogen activated protein kinases (MAPKs) are often involved in altering transcription factor activity, often activating WRKY transcription factors (Bigeard et al. 2015).

A number of rice MAPKs are transcriptionally upregulated by various of plant hormones, discussed below, and the reactive oxygen species (ROS) H_2O_2 (Agrawal et al. 2003). Expression of the rice MAPK gene, *OsBWMK1*, was shown to be induced by *Magnaporthe grisea* infection of rice as well as by wounding (He et al. 1999). The expression of *OsBIMK1,* another MAPK gene, is activated in response to *M. grisea* in an incompatible reaction but not in a compatible reaction, indicating a role in resistance (Song and Goodman 2002).

Small GTPases belong to a diverse protein family functioning in a wide variety of pathways in the cell (Takai et al. 2001; Yang 2002). Some of them, from the Ras and Rho classes, are involved in defence signalling. Through signalling they may regulate H_2O_2 or other ROS production, or they may regulate abscisic acid (ABA) and other hormones (Yang 2002). GTPases have been shown to be important in callose deposition and papillae formation, which prevents pathogen penetration into cells (Böhlenius et al. 2010; Ellinger et al. 2014). Specifically, Rho-like GTP-binding proteins have been implicated in inducing callose deposition (Hong et al. 2001).

1.5.2.3.1 Plant hormones in defence signalling

The plant hormones involved in plant defence against pathogens are salicylic acid (SA), JA and ethylene, and ABA. JA and ethylene are usually associated with defence against necrotrophs and SA with defence against biotrophs. However defence responses are often far more complicated and both hormones may be involved with synergistic and conflicting interactions between the pathways (Beckers and Spoel 2006; Pieterse et al. 2009).

The role of SA and JA in plant defence is reviewed by Halim et al. (2006) and Beckers and Spoel (2006). The SA defence pathway seems to be involved in resistance of pearl millet to rust, caused by the biotroph *Puccinia substriata*, as resistance was conferred to pearl millet by SA application and not methyl jasmonate (MeJA) application (Crampton et al. 2009). However, an overlap in the transcript responses to SA and MeJA application was observed.

SA levels in uninfected maize plants are fairly low, and are induced strongly in plant tissues infected with *C. graminicola* causing anthracnose leaf blight, and *Bipolaris maydis* causing southern corn leaf blight (Morris et al. 1998). Surprisingly, SA is not induced in rice following infection with various pathogens, but rice has high constitutive SA levels (Silverman et al. 1995). The levels of SA in rice did however correlate with blast resistance.

A rice Myb transcription factor gene (*JAmyb*) is induced by JA as well as by the blast fungus (*M. grisea*) (Lee et al. 2001). A gene encoding a proteinase inhibitor (*MPI*) is induced by *F. moniliforme* infection as well as by ABA or MeJA application (Cordero et al. 1994). There is also both a local and systemic induction of the gene in response to wounding. Bravo et al. (2003) also found that wounding or ABA or MeJA application to maize plants increased *ZmPR4* transcript accumulation, which they also found in the response to *F. moniliforme* infection. Application of gibberellic acid or SA had no effect on *ZmPR4* transcript accumulation.

The role of ABA in pathogen resistance appears to be in initiating callose production for papillae formation (Ton and Mauch-Mani 2004). It can, however, have antagonistic interactions with the other hormone signalling pathways with negative effects on resistance (Mauch-Mani and Mauch 2005). Chen et al. (2013) have shown that application of ABA to barley increases resistance to *B. graminis*, which is associated with papillae formation to prevent penetration.

1.5.3 Induced plant defence responses

The recognition of a pathogen results in complex signalling networks (Hammond-Kosack and Parker 2003; Tena et al. 2011) that activate the production of pathogenesis related (PR) proteins (van Loon et al. 2006), some of which play a role in systemic acquired resistance (SAR) (Gaffney et al. 1993; Ryals et al. 1996). The hypersensitive response (Lam et al. 2001; Morel and Dangl 1997) or the production of secondary antimicrobial compounds (Großkinsky

et al. 2012) and paplillae formation (Underwood 2012) may follow to prevent further pathogen invasion of host cells. These are discussed in more detail below.

1.5.3.1 Pathogenesis related proteins

The PR proteins are transcribed in response to pathogen attack. There are a growing number of recognised PR protein families with various functions (Table 1.3). Although most PR proteins were first identified in dicots, monocots have been shown to produce orthologues for many of the PR proteins.

PR1 is up regulated in response to MeJA application in pearl millet (Crampton et al. 2009). Increased expression levels of a β-1,3-glucanase (PR2) in pearl millet showed lower susceptibility to the downy mildew causing oomycete, *Sclerospora graminicola* (O'Kennedy et al. 2011).

Table 1.3: Types of PR proteins (adapted from van Loon et al. (2006))

PR protein family	Protein class/properties	Type member
PR1	unknown	Tobacco PR-1a
PR2	β-1,3-glucanase	Tobacco PR-2
PR3	Chitinase type I, II, IV, V, VI, VII	Tobacco P, Q
PR4	Chitinase type I, II	Tobacco 'R'
PR5	Thaumatin-like	Tobacco S
PR6	Proteinase inhibitor	Tomato Inhibitor I
PR7	Endoproteinase	Tomato P69
PR8	Chitinase type III	Cucumber chitinase
PR9	Peroxidase	Tobacco "lignin-forming peroxidase"
PR10	Ribonuclease-like	Parsley "PR1"
PR11	Chitinase type I	Tobacco "class V" chitinase
PR12	Defensin	Radish Rs-AFP3
PR13	Thionin	Arabidopsis THI2.1
PR14	Lipid transfer protein	Barley LTP4
PR15	Oxalate oxidase	Barley OxOa (germin)
PR16	Oxalate-oxidase-like	Barley OxOLP
PR17	Unknown	Tobacco PRp27

The maize *ZmPR4* gene encodes a class II chitinase, which is a PR4 protein, and is induced upon *F. moniliforme* infection, wounding, and treatment with fungal elicitors or MeJA (Bravo et al. 2003). Rice PR4 (*OsPR4*) is expressed in leaves infected with *M. grisea*, but not induced by wounding (Agrawal et al. 2003). Again, this PR4 protein was up regulated by JA and ABA, but not SA.

Maize *PR1* and *PR5* genes are induced by common rust (*P. sorghi*) and Southern corn leaf blight (*Cochliobolus heterostrophus*) infection (Morris et al. 1998). Up regulation of both genes was observed at early time points in incompatible interactions and induction at lower levels was seen for compatible interactions at later time points.

1.5.3.2 Systemic acquired resistance

Systemic acquired resistance is a broad spectrum resistance that spreads through the plant, protecting it from future infections by a variety of pathogens (Ryals et al. 1994; Ryals et al. 1996). The SA signalling pathway induces *NPR1 (Non expressor of PR1)*, which activates SAR (Vlot et al. 2008). SAR is activated along with expression of genes associated with SAR, including PR proteins, in infected tissues as well as non-infected tissues (Ward et al. 1991). SA has been shown to be a necessary signalling molecule for SAR induction (Gaffney et al. 1993), while more recently, Park et al. (2007) have demonstrated that methyl salicylate (MeSA) is the mobile signal required for SAR. SA induces the transcription of the genes associated with SAR (Ward et al. 1991). In maize, *C. graminicola* leaf infection resulted in SAR against inoculation with the same fungus in neighbouring leaves (Balmer et al. 2013c) and it was associated with elevated SA and ABA levels.

1.5.3.3 Reactive oxygen species

Reactive oxygen species as well as nitric oxide (NO), are produced in a process of oxidative burst and play a role in cell death in response to pathogen invasion (Torres et al. 2006). NADPH oxidase is the main enzyme responsible for ROS production but other enzymes, such as cell wall peroxidases, may also be involved (Torres et al. 2006). Superoxide (O_2^-) and hydrogen peroxide (H_2O_2) are involved in signalling (Torres et al. 2006). The signal network involved in ROS production is complex, interacts with other pathways, and can produce different effects with different pathogens. H_2O_2 induces PR protein production, but to a lesser extent than SA (Lamb and Dixon 1997). In maize, H_2O_2 has been shown to induce Ca^{2+} influx into the cytosol and to induce calmodulin (Hu et al. 2007) and NO production (Zhang et al. 2007). Production of ROS in maize has been shown to occur early in the defence response to *C. graminicola*, and production of H_2O_2 containing vesicles were produced in the plant cells to target the fungal hyphae (Vargas et al. 2012). Oxidative burst plays an important role in SAR (Lamb and Dixon 1997; Torres et al. 2006). Accumulation of free radicals, such as H_2O_2 can result in the hypersensitive response (Apel and Hirt 2004).

1.5.3.4 Hypersensitive response

Non-compatible host-pathogen interactions appear to have a low-scale buildup of H_2O_2, followed by a second, larger accumulation of H_2O_2, while compatible host-pathogen interactions produce a smaller low scale accumulation, which is not followed by the second phase (Lamb and Dixon 1997). When H_2O_2 reaches a critical threshold the hypersensitive response ensues (Lamb and Dixon 1997). Trujillo et al. (2004) showed that H_2O_2 accumulation is associated with the hypersensitive response in the non-host resistance interaction of wheat and *B. graminis* f. sp. *hordei* (barley powdery mildew). Similar results were obtained for the non-host resistance interaction of barley and *B. graminis* f. sp. *tritici* (wheat powdery mildew) (Hückelhoven et al. 2001).

Following the initial hypersensitive response, ROS from different pathways can then inhibit the spread of cell death (Torres et al. 2006). Elicitors and effectors can trigger ROS accumulation resulting in increases in glutathione, ethylene, lipoxygenase production, and cell death (Lamb and Dixon 1997). Ca^{2+} influx is vital for oxidative burst and the hypersensitive response (Levine et al. 1996; Lamb and Dixon 1997; Torres et al. 2006).

The hypersensitive response is more effective against biotrophs, as biotrophs require living tissue. For example, the non-host resistance of wheat to the biotroph, *B. graminis* f. sp. *hordei*, which includes an effective hypersensitive response (Trujillo et al. 2004). In host resistance of barley to *B. graminis*, there is an accumulation of NO preceding the hypersensitive response (Prats et al. 2005). However, hemibiotrophs have an initial biotrophic phase before switching to a necrotrophic phase, e.g. *C. graminicola* (Vargas et al. 2012). The initial penetration and growth phase of necrotrophs, such as *C. zeae-maydis*, which grows intercellularly after entering the leaf via the stomata before changing to its necrotrophic stage (Kim et al. 2011), may be halted by a fast enough hypersensitive response.

QTL have been mapped for hypersensitive response-related phenotypes, related to the R gene lesion mimic mutant (*Rp1-D21*) described by Collins et al. (1999), in the maize nested association mapping population (Olukolu et al. 2014). Candidate genes that were identified were differentially expressed between maize isogenic lines polymorphic at the *Rp1-D21* locus. Certain of these candidate genes function in programmed cell death pathways.

1.5.3.5 Secondary antimicrobial compounds

Phytoalexins are secondary metabolites involved in induced plant defence and fall into three groups: alkaloids; terpenoids and shikimates (Balmer et al. 2013a). Research on phytoalexins has correlated the accumulation of these low molecular weight antimicrobial compounds to pathogen resistance (Hammerschmidt 1999). Schmelz et al. (2011) and Huffaker et al. (2011b) isolated and characterised two new maize terpenoid phytoalexins, kauralexin and zealexin. Kauralexins are produced in maize in response to European Corn Borer feeding and infection with *Rhizopus microsporus* and *C. graminicola* (Schmelz et al. 2011). They also inhibited the growth of both fungal species. Sesquiterpenoid phytoalexins in maize, or zealexins, are produced in response to *Fusarium graminearum* infection of maize, as well as other fungi and damage due to herbivory. Two terpene synthase genes (*Tsp6* and *Tsp11*) had the highest up regulation in the immune response (Huffaker et al. 2011b). Furthermore, in an expression profiling study of the maize response to *Fusarium verticillioides*, an increase in transcription of genes involved in secondary metabolism, specifically the biosynthesis of shikimate, lignin, flavonoids and terpenoids was detected in a resistant maize line compared to a susceptible maize line (Lanubile et al. 2014).

1.5.3.6 Papillae

Papillae are formed at the cell wall by deposition of callose, a linear β-1,3-glucan made from D-glucose molecules bound together by beta-1,3-glycosidic bonds. It is formed by callose synthases and deposited between the plant cell wall and plasma membrane in response to biotic stresses, such as pathogen attack (Humphrey et al. 2007; Chen and Kim 2009; Piršelová and Matušíková 2013). The formation of papillae, or thickening of the cell wall beneath the epidermis at the point of hyphal entry creates a physical barrier. The main component is callose with other polysaccharides, phenolic compounds and proteins present (Flors et al. 2005).

In the production of callose, the callose synthases work in association with other proteins in membrane associated complexes (Hong et al. 2001; Verma and Hong 2001). The callose synthase complexes seem to be activated directly or indirectly by a protease (Nakashima et al. 2003). Abscisic acid and beta-amino-butyric acid (BABA) signalling (Ton and Mauch-Mani 2004; Flors et al. 2005) as well as Rho-like GTP-binding proteins (Hong et al. 2001) have been implicated in inducing callose deposition. Although callose is an important constituent of papillae for defence, other compounds, such as phenylpropanoinds, also play vital roles in papillae formation (Aist et al. 1988).

Investigation of callose deposition in monocots during defence response has also been carried out. A functional *AtGSL5* (Ellinger et al. 2013) was transiently expressed in barley where it reduced penetration success of the powdery mildew, *Blumeria graminis*, on the *AtGSL5*-infiltrated leaves (Blümke et al. 2013). An example of papillae formation with evidence of callose deposition in a susceptible response of maize to *C. graminicola* is given by Mims and Vaillancourt (2002). Callose deposition in rice has also been shown to be induced by the feeding of a herbivorous insect (Hao et al. 2008). Early papillae formation and deposition of callose was found to be associated with low penetration efficiency of powdery mildew in barley while delayed callose formation led to susceptibility (Gold et al. 1986; Bayles et al. 1990; Inouz et al. 1994).

Considering fungal pathogens, an exo-β-1,3-glucanase has been found to be up regulated in *C. zeae-maydis* during sporulation as compared to vegetative growth in culture (Bluhm et al. 2008). This may allude to a struggle between β-1,3-glucanases and callose synthases with the fungus secreting cell wall degrading enzymes, and maize leaves reinforcing and replenishing cell wall defences.

1.6 Conclusions

As pathogens are a major threat to yield, the defence responses of crops to their pathogens are of interest to crop breeders and farmers, as well as for scientific knowledge. The most practical grey leaf spot control strategy in Africa is resistant hybrids with durable, quantitative resistance. The underlying mechanisms of quantitative resistance can be found in the complex molecular mechanisms of the plant's immune system summarized here. To aid maize farmers in protecting their yield from *C. zeina* infection, the mechanisms behind the plant-pathogen interaction must be understood so that these mechanisms can be used to produce better, more resistant hybrids.

We thus wanted to identify candidate genes involved in quantitative resistance underlying the resistance phenotype of maize to GLS. To do this we sought to quantify *C. zeina in planta* using a molecular quantitative PCR approach. This allowed us to undertake the mapping of QTL for limiting *C. zeina* growth in infected plants' leaves and compare them to GLS resistance QTL identified from whole plant disease scoring. We then aimed to find differentially expressed genes involved in the resistance response between resistant and susceptible maize RILs with known resistance or susceptible QTL regions.

1.7 References:

Agrawal, G. K., Iwahashi, H., & Rakwal, R. (2003). Rice MAPKs. *Biochemical and Biophysical Research Communications*, *302*, 171–180.

Agrawal, G. K., Jwa, N. S., Han, K. S., Agrawal, V. P., & Rakwal, R. (2003). Isolation of a novel rice PR4 type gene whose mRNA expression is modulated by blast pathogen attack and signaling components. *Plant Physiology and Biochemistry*, *41*, 81–90.

Aist, J. R., Gold, R. E., Bayles, C. J., Morrison, G. H., Chandra, S., & Israel, H. W. (1988). Evidence that molecular components of papillae may be involved in mlo resistance to barley powdery mildew. *Physiological and Molecular Plant Pathology*, *33*, 17–32.

Apel, K., & Hirt, H. (2004). Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annual Review of Plant Biology*, *55*, 373–399.

Ayliffe, M. A., & Lagudah, E. S. (2004). Molecular genetics of disease resistance in cereals. *Annals of Botany*, *94*, 765–773.

Balint-Kurti, P. J., Wisser, R., & Zwonitzer, J. C. (2008). Use of an advanced intercross line population for precise mapping of quantitative trait loci for gray leaf spot resistance in maize. *Crop Science*, *48*, 1696–1704.

Balmer, D., Flors, V., Glauser, G., & Mauch-mani, B. (2013a). Metabolomics of cereals under biotic stress: current knowledge and techniques. *Frontiers in Plant Science*, *4*, 1–12.

Balmer, D., Planchamp, C., & Mauch-Mani, B. (2013b). On the move: induced resistance in monocots. *Journal of Experimental Botany*, *64*, 1249–1261.

Balmer, D., De Papajewski, D. V., Planchamp, C., Glauser, G., & Mauch-Mani, B. (2013c). Induced resistance in maize is based on organ-specific defence responses. *The Plant Journal*, *74*, 213–225.

Bayles, C. J., Ghemawat, M. S., & Aist, J. R. (1990). Inhibition by 2-deoxy-D-glucose of callose formation, papilla deposition, and resistance to powdery mildew in an ml-o barley mutant. *Physiological and Molecular Plant Pathology*, *36*, 63–72.

Beckers, G. J. M., & Spoel, S. H. (2006). Fine-tuning plant defence signalling: salicylate versus jasmonate. *Plant Biology*, *8*, 1–10.

Beckman, P. M., & Payne, G. A. (1982). External growth, penetration, and development of *Cercospora zeae-maydis* in corn leaves. *Phytopathology*, *72*, 810–815.

Benson, J. M., Poland, J. A., Benson, B. M., Stromberg, E. L., & Nelson, R. J. (2015). Resistance to gray leaf spot of maize: genetic architecture and mechanisms elucidated through nested association mapping and near-isogenic line analysis. *PLOS Genetics*, *11*, e1005045.

Berger, D. K., Carstens, M., Korsman, J. N., Middleton, F., Kloppers, F. J., Tongoona, P., & Myburg, A. A. (2014). Mapping QTL conferring resistance in maize to gray leaf spot disease caused by *Cercospora zeina*. *BMC Genetics*, *15*, 60.

Bigeard, J., Colcombet, J., & Hirt, H. (2015). Signaling mechanisms in pattern-triggered immunity (PTI). *Molecular Plant*, *8*, 657–674.

Bluhm, B. H., Dhillon, B., Lindquist, E. A., Kema, G. H., Goodwin, S. B., & Dunkle, L. D. (2008). Analyses of expressed sequence tags from the maize foliar pathogen *Cercospora zeae-maydis* identify novel genes expressed during vegetative, infectious, and reproductive growth. *BMC Genomics*, *9*, 523.

Blümke, A., Somerville, S. C., & Voigt, C. A. (2013). Transient expression of the *Arabidopsis thaliana* callose synthase PMR4 increases penetration resistance to powdery mildew in barley. *Advances in Biosciences and Biotechnology*, *4*, 810–813.

Böhlenius, H., Mørch, S. M., Godfrey, D., Nielsen, M. E., & Thordal-Christensen, H. (2010). The multivesicular body-localized GTPase ARFA1b/1c is important for callose deposition and ROR2 syntaxin-dependent preinvasive basal defense in barley. *The Plant Cell*, *22*, 3831–3844.

Bravo, J. M., Campo, S., Murillo, I., Coca, M., & San Segundo, B. (2003). Fungus- and wound-induced accumulation of mRNA containing a class II chitinase of the pathogenesis-related protein 4 (PR-4) family of maize. *Plant Molecular Biology*, *52*, 745–759.

Brewbaker, J. L. (1979). Diseases of maize in the wet lowland tropics and the collapse of the classic Maya civilization. *Economic Botany*, *33*, 101–118.

Brueggeman, R., Rostoks, N., Kudrna, D., Kilian, A, Han, F., Chen, J., Druka, A., Steffenson, B., & Kleinhofs, A. (2002). The barley stem rust-resistance gene Rpg1 is a novel disease-resistance gene with homology to receptor kinases. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 9328–9333.

Brunelli, K. R., Dunkle, L. D., Sobrinho, C. A., Fazza, A. C., & Camargo, L. E. A. (2008). Molecular variability in the maize grey leaf spot pathogens in Brazil. *Genetics and Molecular Biology*, *942*, 938–942.

Bryan, G. T., Wu, K. S., Farrall, L., Jia, Y., Hershey, H. P., McAdams, S. A., Faulk, K. N., Donaldson, G. K., Tarchini, R., & Valent, B. (2000). A single amino acid difference distinguishes resistant and susceptible alleles of the rice blast resistance gene Pi-ta. *The Plant Cell, 12*, 2033–2046.

Büschges, R., Hollricher, K., Panstruga, R., Simons, G., Wolter, M., Frijters, A., van Daelen, R., van der Lee, T., Diergaarde, P., Groenendijk, J., Töpsch, S., Vos, P., Salamini, F., & Schulze-Lefert, P. (1997). The barley Mlo gene: a novel control element of plant pathogen resistance. *Cell*, *88*, 695–705.

Caldwell, P. M., & Laing, M. D. (2005). Light, scanning and transmission electron microscopy studies on the conidiogenesis of *Cercospora zeae-maydis* on maize. *South African Journal of Plant and Soil*, *22*, 207–213.

Chen, X., & Ronald, P. C. (2011). Innate immunity in rice. *Trends in Plant Science*, *16*, 451–459.

Chen, X.-Y., & Kim, J.-Y. (2009). Callose synthesis in higher plants. *Plant Signaling and Behavior*, *4*, 489–92.

Chen, Y. J., Perera, V., Christiansen, M. W., Holme, I. B., Gregersen, P. L., Grant, M. R., Collinge, D. B., & Lyngkjær, M. F. (2013). The barley *HvNAC6* transcription factor affects ABA accumulation and promotes basal resistance against powdery mildew. *Plant Molecular Biology*, *83*, 577–590.

Churchill, G. A., & Doerge, R. W. (1994). Empirical threshold values for quantitative trait mapping. *Genetics*, *138*, 963–971.

Clements, M. J., Dudley, J. W., & White, D. G. (2000). Quantitative trait loci associated with resistance to gray leaf spot of corn. *Phytopathology*, *90*, 1018–25.

Collinge, D. B., Jørgensen, H. J. L., Lund, O. S., & Lyngkjaer, M. F. (2010). Engineering pathogen resistance in crop plants: current trends and future prospects. *Annual Review of Phytopathology*, *48*, 269–291.

Collins, N., Drake, J., Ayliffe, M., Sun, Q., Ellis, J., Hulbert, S., & Pryor, T. (1999). Molecular characterization of the maize Rp1-D rust resistance haplotype and its mutants. *The Plant Cell*, *11*, 1365–1376.

Cordero, M. J., Raventós, D., & San Segundo, B. (1994). Expression of a maize proteinase inhibitor gene is induced in response to wounding and fungal infection: systemic wound-response of a monocot gene. *The Plant Journal*, *6*, 141–150.

Couture, R., Routley, D., & Dunn, G. (1971). Role of cyclic hydroxamic acids in monogenic resistance of maize to *Helminthosporium turcicum*. *Physiologial Plant Pathology*, *1*, 515–521.

Crampton, B. G., Hein, I., & Berger, D. K. (2009). Salicylic acid confers resistance to a biotrophic rust pathogen, *Puccinia substriata*, in pearl millet (*Pennisetum glaucum*). *Molecular Plant Pathology*, *10*, 291–304.

Crous, P. W., Groenewald, J. Z., Groenewald, M., Caldwell, P., Braun, U., & Harrington, T. C. (2006). Species of *Cercospora* associated with grey leaf spot of maize. *Studies in Mycology*, *55*, 189–197.

Dean, R., Van Kan, J. A. L., Pretorius, Z. A., Hammond-Kosack, K. E., Di Pietro, A., Spanu, P. D., Rudd, J. J., Dickman, M., Kahmann, R., Ellis, J., & Foster, G. D. (2012). The top 10 fungal pathogens in molecular plant pathology. *Molecular Plant Pathology*, *13*, 414–430.

Doebley, J. (2004). The genetics of maize evolution. *Annual Review of Genetics*, *38*, 37–59.

Doerge, R. W., Zeng, Z., & Weir, B. S. (1997). Statistical issues in the search for genes affecting quantitative traits in experimental populations. *Statistical Science*, *12*, 195–219.

Dunkle, L. D., & Levy, M. (2000). Genetic relatedness of African and United States populations of *Cercospora zeae-maydis*. *Phytopathology*, *90*, 486–490.

Ellinger, D., Glockner, A., Koch, J., Naumann, M., Sturtz, V., Schutt, K., Manisseri, C., Somerville, S. C., & Voigt, C. A. (2014). Interaction of the Arabidopsis GTPase RabA4c with its effector PMR4 results in complete penetration resistance to Powdery Mildew. *The Plant Cell*, *26*, 3185–3200.

Ellinger, D., Naumann, M., Falter, C., Zwikowics, C., Jamrow, T., Manisseri, C., Somerville, S. C., & Voigt, C. A. (2013). Elevated early callose deposition results in complete penetration resistance to powdery mildew in Arabidopsis. *Plant Physiology*, *161*, 1433–44.

Elliott, C., Müller, J., Miklis, M., Bhat, R. A., Schulze-Lefert, P., & Panstruga, R. (2005). Conserved extracellular cysteine residues and cytoplasmic loop-loop interplay are required for functionality of the heptahelical MLO protein. *The Biochemical Journal*, *385*, 243–254.

Felix, G., Duran, J. D., Volko, S., & Boller, T. (1999). Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *The Plant Journal*, *18*, 265–276.

Felix, G., Regenass, M., & Boller, T. (1993). Specific perception of subnanomolar concentrations of chitin fragments by tomato cells: induction of extracellular alkalinization, changes in protein phosphorylation, and establishment of a refractory state. *The Plant Journal*, *4*, 307–316.

Ferreira, S. A., Pitz, K. Y., Manshardt, R., Zee, F., Fitch, M., & Gonsalves, D. (2002). Virus coat protein transgenic papaya provides practical control of papaya ringspot virus in Hawaii. *Plant Disease*, *86*, 101–105.

Flors, V., Ton, J., Jakab, G., & Mauch-Mani, B. (2005). Abscisic acid and callose: team players in defence against pathogens? *Journal of Phytopathology*, *153*, 377–383.

Gaffney, T., Friedrich, L., Vernooij, B., Negrotto, D., Nye, G., Uknes, S., Ward, E., Kessmann, H., & Ryals, J. (1993). Requirement of salicylic acid for the induction of systemic acquired resistance. *Science*, *261*, 754–756.

Glazebrook, J., Rogers, E. E., & Ausubel, F. M. (1997). Use of Arabidopsis for genetic dissection of plant defense responses. *Annual Review of Genetics*, *31*, 547–569.

Gold, R. E., Aist, J. R., Hazen, B. E., Stolzenburg, M. C., Marshall, M. R., & Israel, H. W. (1986). Effects of calcium nitrate and chlortetracycline on papilla formation, ml-o resistance and susceptibility of barley to powdery mildew. *Physiological and Molecular Plant Pathology*, *29*, 115–129.

Goodwin, S. B., Dunkle, L. D., & Zismann, V. L. (2001). Phylogenetic analysis of *Cercospora* and *Mycosphaerella* based on the internal transcribed spacer region of ribosomal DNA. *Phytopathology*, *91*, 648–658.

Gordon, S. G., Bartsch, M., Matthies, I., Gevers, H. O., Lipps, P. E., & Pratt, R. C. (2004). Linkage of molecular markers to *Cercospora zeae-maydis* resistance in maize. *Crop Science*, *44*, 628–636.

Greenberg, J. T., & Yao, N. (2004). The role of regulation of programmed cell death in plant-pathogen interactions. *Cellular Microbiology*, *6*, 201–211.

Großkinsky, D. K., van der Graaff, E., & Roitsch, T. (2012). Phytoalexin transgenics in crop protection-fairy tale with a happy end? *Plant Science*, *195*, 54–70.

Halim, V. A., Vess, A., Scheel, D., & Rosahl, S. (2006). The role of salicylic acid and jasmonic acid in pathogen defence. *Plant Biology*, *8*, 307–313.

Halterman, D., Zhou, F., Wei, F., Wise, R. P., & Schulze-Lefert, P. (2001). The MLA6 coiled-coil, NBS-LRR protein confers AvrMla6-dependent resistance specificity to *Blumeria graminis* f. sp. *hordei* in barley and wheat. *The Plant Journal*, *25*, 335–348.

Hammerschmidt, R. (1999). Phytoalexins: what have we learned after 60 years? *Annual Review of Phytopathology*, *37*, 285–306.

Hammond-Kosack, K. E., & Jones, J. D. G. (1997). Plant disease resistance genes. *Annual Review of Plant Physiology and Plant Molecular Biology*, *48*, 575–607.

Hammond-Kosack, K. E., & Parker, J. E. (2003). Deciphering plant-pathogen communication: fresh perspectives for molecular resistance breeding. *Current Opinion in Biotechnology*, *14*, 177–193.

Hao, P., Liu, C., Wang, Y., Chen, R., Tang, M., Du, B., Zhu, Lili., & He, G. (2008). Herbivore-induced callose deposition on the sieve plates of rice: an important mechanism for host resistance. *Plant Physiology*, *146*, 1810–20.

Harmon, A. C., Gribskov, M., & Harper, J. F. (2000). CDPKs – a kinase for every Ca^{2+} signal? *Trends in Plant Science*, *5*, 154–159.

He, C., Fong, S. H., Yang, D., & Wang, G. L. (1999). BWMK1, a novel MAP kinase induced by fungal infection and mechanical wounding in rice. *Molecular Plant-Microbe Interactions. 12*, 1064–1073.

Hong, Z., Zhang, Z., Olson, J. M., & Verma, D. P. (2001). A novel UDP-glucose transferase is part of the callose synthase complex and interacts with phragmoplastin at the forming cell plate. *The Plant Cell*, *13*, 769–79.

Hu, G., Richter, T., Hulbert, S., & Pryor, T. (1996). Disease lesion mimicry caused by mutations in the rust resistance gene *rp1*. *The Plant Cell*, *8*, 1367–1376.

Hu, X., Jiang, M., Zhang, J., Zhang, A., Lin, F., & Tan, M. (2007). Calcium–calmodulin is required for abscisic acid-induced antioxidant defense and functions both upstream and downstream of H2O2 production in leaves of maize (*Zea mays*) plants. *New Phytologist*, *173*, 27–38.

Hückelhoven, R., Dechert, C., & Kogel, K. H. (2001). Non-host resistance of barley is associated with a hydrogen peroxide burst at sites of attempted penetration by wheat powdery mildew fungus. *Molecular Plant Pathology*, *2*, 199–205.

Huffaker, A., Dafoe, N. J., & Schmelz, E. A. (2011a). ZmPep1, an ortholog of Arabidopsis elicitor peptide 1, regulates maize innate immunity and enhances disease resistance. *Plant Physiology*, *155*, 1325–1338.

Huffaker, A., Kaplan, F., Vaughan, M. M., Dafoe, N. J., Ni, X., Rocca, J. R., Alborn, H. T., Teal, P. E. A., & Schmelz, E. A. (2011b). Novel acidic sesquiterpenoids constitute a dominant class of pathogen-induced phytoalexins in maize. *Plant Physiology*, *156*, 2082–2097.

Humphrey, T. V, Bonetta, D. T., & Goring, D. R. (2007). Sentinels at the wall: cell wall receptors and sensors. *New Phytologist*, *176*, 7–21.

Inouz, S., Aist, J. R., & Macko, V. (1994). Earlier papilla formation and resistance to barley powdery mildew induced by a papilla-regulating extract. *Physiological and Molecular Plant Pathology*, *44*, 433–440.

Johal, G. S., & Briggs, S. P. (1992). Reductase activity encoded by the *HM1* disease resistance gene in maize. *Science* , *258*, 985–987.

Jones, J. D. G., & Dangl, J. L. (2006). The plant immune system. *Nature*, *444*, 323–329.

Juliatti, F. C., Pedrosa, M. G., Silva, H. D., & Silva, J. V. C. (2009). Genetic mapping for resistance to gray leaf spot in maize. *Euphytica*, *169*, 227–238.

Kim, H., Ridenour, J. B., Dunkle, L. D., & Bluhm, B. H. (2011). Regulation of pathogenesis by light in *Cercospora zeae-maydis*: an updated perspective. *The Plant Pathology Journal*, *27*, 103–109.

Lam, E., Kato, N., & Lawton, M. (2001). Programmed cell death, mitochondria and the plant hypersensitive response. *Nature*, *411*, 848–853.

Lamb, C., & Dixon, R. A. (1997). The oxidative burst in plant disease resistance. *Annual Review of Plant Physiology and Plant Molecular Biology*, *48*, 251–275.

Lanubile, A., Ferrarini, A., Maschietto, V., Delledonne, M., Marocco, A., & Bellin, D. (2014). Functional genomic analysis of constitutive and inducible defense responses to *Fusarium verticillioides* infection in maize genotypes with contrasting ear rot resistance. *BMC Genomics*, *15*, 710.

Latterell, F., & Rossi, A. (1983). Gray leaf spot of corn: a disease on the move. *Plant Disease*, *67*, 842–847.

Łaźniewska, J., Macioszek, V. K., & Kononowicz, A. K. (2012). Plant-fungus interface: the role of surface structures in plant resistance and susceptibility to pathogenic fungi. *Physiological and Molecular Plant Pathology*, *78*, 24–30.

Lee, M. W., Qi, M., & Yang, Y. (2001). A novel jasmonic acid-inducible rice myb gene associates with fungal infection and host cell death. *Molecular Plant-Microbe Interaction*, *14*, 527–535.

Lehmensiek, A., Esterhuizen, A., van Staden, D., Nelson, S., & Retief, A. (2001). Genetic mapping of gray leaf spot (GLS) resistance genes in maize. *Theoretical and Applied Genetics*, *103*, 797–803.

Levine, A., Pennell, R. I., Alvarez, M. E., Palmer, R., & Lamb, C. (1996). Calcium-mediated apoptosis in a plant hypersensitive disease resistance response. *Current Biology*, *6*, 427–437.

Lindhout, P. (2002). The perspectives of polygenic resistance in breeding for durable disease. *Euphytica*, *124*, 217–226.

Lipka, U., Fuchs, R., & Lipka, V. (2008). Arabidopsis non-host resistance to powdery mildews. *Current Opinion in Plant Biology*, *11*, 404–411.

Liu, K.-J., & Xu, X.-D. (2013). First report of gray leaf spot of maize caused by *Cercospora zeina* in China. *Plant Disease*, *97*, 1656.

Martin, T., Biruma, M., Fridborg, I., Okori, P., & Dixelius, C. (2011). A highly conserved NB-LRR encoding gene cluster effective against *Setosphaeria turcica* in sorghum. *BMC Plant Biology*, *11*, 151.

Mauch-Mani, B., & Mauch, F. (2005). The role of abscisic acid in plant-pathogen interactions. *Current Opinion in Plant Biology*, *8*, 409–414.

Meisel, B., Korsman, J., Kloppers, F. J., & Berger, D. K. (2009). *Cercospora zeina* is the causal agent of grey leaf spot disease of maize in southern Africa. *European Journal of Plant Pathology*, *124*, 577–583.

Mims, C. W., & Vaillancourt, L. J. (2002). Ultrastructural characterization of infection and colonization of maize leaves by *Colletotrichum graminicola*, and by a *C. graminicola* pathogenicity mutant. *Phytopathology*, *92*(7), 803–812.

Morel, J. B., & Dangl, J. L. (1997). The hypersensitive response and the induction of cell death in plants. *Cell Death and Differentiation*, *4*, 671–683.

Morris, S. W., Vernooij, B., Titatarn, S., Starrett, M., Thomas, S., Wiltse, C. C., Frederiksen, R. A., Bhandhufalck, A., Hulbert, S., & Uknes, S. (1998). Induced resistance responses in maize. *Molecular Plant-Microbe Interactions*, *11*, 643–658.

Murillo, I., Jaeck, E., Cordero, M., & Segundo, B. (2001). Transcriptional activation of a maize calcium-dependent protein kinase gene in response to fungal elicitors and infection. *Plant Molecular Biology*, *45*, 145–158.

Nakashima, J., Laosinchai, W., Cui, X., & Brown, R. M. (2003). New insight into the mechanism of cellulose and callose biosynthesis: proteases may regulate callose biosynthesis upon wounding. *Cellulose*, *10*, 369–389.

Niemeyer, H. (2009). Hydroxamic acids derived from 2-Hydroxy-2H-1,4-Benzoxazin-3(4H)-one: key defense chemicals of cereals. *Journal of Agricultural and Food Chemistry*, *3*, 1677–1696.

O'Kennedy, M. M., Crampton, B. G., Lorito, M., Chakauya, E., Breese, W. A., Burger, J. T., & Botha, F. C. (2011). Expression of a β-1,3-glucanase from a biocontrol fungus in transgenic pearl millet. *South African Journal of Botany*, *77*, 335–345.

Okori, P., Fahleson, J., Rubaihayo, P., Adipala, E., & Dixelius, C. (2003). Assessment of genetic variation among East African *Cercospora zeae-maydis*. *African Crop Science Journal*, *11*, 75–85.

Olukolu, B. A., Wang, G.-F., Vontimitta, V., Venkata, B. P., Marla, S., Ji, J., Gachomo, E., Chu, K., Negeri, A., Benson, J., Nelson, R., Bradbury, P., Nielsen, D., Holland, J. B., Balint-Kurti, P. J., & Johal, G. (2014). A genome-wide association study of the maize hypersensitive defense response identifies genes that cluster in related pathways. *PLoS Genetics*, *10*, e1004562.

Osbourn, A. (1996). Preformed antimicrobial compounds and plant defense against fungal attack. *The Plant Cell*, *8*, 1821–1831.

Park, S., Kaimoyo, E., Kumar, D., Mosher, S., & Klessig, D. F. (2007). Methyl salicylate is a critical mobile signal for plant systemic acquired resistance. *Science*, *318*, 113–116.

Parlevliet, J. E. (2002). Durability of resistance against fungal, bacterial and viral pathogens: present situation. *Euphytica*, *124*, 147–156.

Pieterse, C. M. J., Leon-Reyes, A., Van der Ent, S., & Van Wees, S. C. M. (2009). Networking by small-molecule hormones in plant immunity. *Nature Chemical Biology*, *5*, 308–316.

Piršelová, B., & Matušíková, I. (2013). Callose: the plant cell wall polysaccharide with multiple biological functions. *Acta Physiologiae Plantarum*, *35*, 635–644.

Pozar, G., Butruille, D., Silva, H. D., McCuddin, Z. P., & Penna, J. C. V. (2009). Mapping and validation of quantitative trait loci for resistance to *Cercospora zeae-maydis* infection in tropical maize (*Zea mays* L.). *Theoretical and Applied Genetics*, *118*, 553–564.

Prats, E., Mur, L. A. J., Sanderson, R., & Carver, T. L. W. (2005). Nitric oxide contributes both to papilla-based resistance and the hypersensitive response in barley attacked by *Blumeria graminis* f. sp. *hordei*. *Molecular Plant Pathology*, *6*, 65–78.

Ryals, J. A., Neuenschwander, U. H., Willits, M. G., Molina, A., Steiner, H. Y., & Hunt, M. D. (1996). Systemic acquired resistance. *The Plant Cell*, *8*, 1809–1819.

Ryals, J., Uknes, S., & Ward, E. (1994). Systemic acquired resistance. *Plant Physiology*, *104*, 1109–1112.

Saghai Maroof, M. A., Yue, Y. G., Xiang, Z. X., Stromberg, E. L., & Rufener, G. K. (1996). Identification of quantitative trait loci controlling resistance to gray leaf spot disease in maize. *Theoretical and Applied Genetics. 93*, 539–46.

Schmelz, E. A., Kaplan, F., Huffaker, A., Dafoe, N. J., Vaughan, M. M., Ni, X., Rocca, J. R., Alborn, H. T., & Teal, P. E. (2011). Identity, regulation, and activity of inducible diterpenoid phytoalexins in maize. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 5455–5460.

Shi, L., Li, X., Hao, Z., Xie, C., Ji, H., Lu, X., Zhang, S., & Pan, G. (2007). Comparative QTL mapping of resistance to gray leaf spot in maize based on bioinformatics. *Agricultural Sciences in China*, *6*, 1411–1419.

Sibiya, J., Tongoona, P., Derera, J., & Makanda, I. (2013). Smallholder farmers' perceptions of maize diseases, pests, and other production constraints, their implications for maize breeding and evaluation of local maize cultivars in KwaZulu-Natal, South Africa. *African Journal of Agricultural Research*, *8*, 1790–1798.

Silverman, P., Seskar, M., Kanter, D., Schweizer, P., Métraux, J.-P., & Raskin, L. (1995). Salicylic acid in rice: biosynthesis, conjugation, and possible role. *Plant Physiology*, *108*, 633–639.

Song, F., & Goodman, R. M. (2002). *OsBIMK1*, a rice MAP kinase gene involved in disease resistance responses. *Planta*, *215*, 997–1005.

Song, W. Y., Wang, G. L., Chen, L. L., Kim, H. S., Pi, L. Y., Holsten, T., Gardner, J., Wang, B., Zhai, W. X., Zhu, L. H., Fauquet, C., & Ronald, P. (1995). A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. *Science, 270*, 1804–1806.

Takai, R., Isogai, A., Takayama, S., & Che, F.-S. (2008). Analysis of flagellin perception mediated by flg22 receptor OsFLS2 in rice. *Molecular Plant-Microbe Interactions*, *21*, 1635–1642.

Takai, Y., Sasaki, T., & Matozaki, T. (2001). Small GTP-binding proteins. *Physiological Reviews*, *81*, 153–209.

Tanksley, S. (1993). Mapping polygenes. *Annual Review of Genetics*, *27*, 205–233.

Tehon, L., & Daniels, E. (1925). Notes on the parasitic fungi of Illinois: II. *Mycologia*, *17*, 240–249.

Tena, G., Boudsocq, M., & Sheen, J. (2011). Protein kinase signaling networks in plant innate immunity. *Current Opinion in Plant Biology*, *14*, 519–529.

Ton, J., & Mauch-Mani, B. (2004). Beta-amino-butyric acid-induced resistance against necrotrophic pathogens is based on ABA-dependent priming for callose. *The Plant Journal*, *38*, 119–30.

Torres, M. A., Jones, J. D. G., & Dangl, J. L. (2006). Reactive oxygen species signaling in response to pathogens. *Plant Physiology*, *141*, 373–378.

Trujillo, M., Kogel, K.-H., & Hückelhoven, R. (2004). Superoxide and hydrogen peroxide play different roles in the nonhost interaction of barley and wheat with inappropriate *formae speciales* of *Blumeria graminis*. *Molecular Plant-Microbe Interactions*, *17*, 304–312.

Underwood, W. (2012). The plant cell wall: a dynamic barrier against pathogen invasion. *Frontiers in Plant Science*, *3*, 1–6.

Van Der Biezen, E. A., & Jones, J. D. G. (1998). Plant disease-resistance proteins and the gene-for-gene concept. *Trends in Biochemical Sciences*, *23*, 454–456.

Van Loon, L. C., Rep, M., & Pieterse, C. M. J. (2006). Significance of inducible defense-related proteins in infected plants. *Annual Review of Phytopathology*, *44*, 135–162.

Vargas, W. A., Sanz Martı´n, J. M., Rech, G. E., Rivera, L. P., Benito, E. P., Dı´az-Mı´nguez, J. M., Thon, M. R., & Sukno, S., A. (2012). Plant defense mechanisms are activated during biotrophic and necrotrophic development of *Colletotricum graminicola* in maize. *Plant Physiology*, *158*, 1342–1358.

Verma, D. P., & Hong, Z. (2001). Plant callose synthase complexes. *Plant Molecular Biology*, *47*, 693–701.

Vlot, A. C., Klessig, D. F., & Park, S. W. (2008). Systemic acquired resistance: the elusive signal(s). *Current Opinion in Plant Biology*, *11* 436–442.

Wang, J., Levy, M., & Dunkle, L. D. (1998). Sibling species of *Cercospora* associated with gray leaf spot of maize. *Phytopathology*, *88*, 1269–1275.

Ward, E., Uknes, S., Williams, S., Dincher, S., Wiederhold, D., Alexander, D., Ahl-Goy, P., Metraux, J. P., & Ryals, J. (1991). Coordinate gene activity in response to agents that induce systemic acquired resistance. *The Plant Cell*, *3*, 1085–1094.

Ward, J. M. J., Laing, M. D., & Rijkenberg, F. H. J. (1997). Frequency and timing of fungicide applications for the control of gray leaf spot in maize. *Plant Disease*, *81*, 41–48.

Ward, J. M. J., & Nowell, D. C. (1998). Integrated management practices for the control of maize grey leaf spot. *Integrated Pest Management Reviews*, *3*, 177–188.

Ward, J. M. J., Stromberg, E. L., Nowell, D. C., & W, N. (1999). Gray leaf spot: a disease of global importance in maize production. *Plant Disease*, *83*, 884–895.

Webb, C. A, Richter, T. E., Collins, N. C., Nicolas, M., Trick, H. N., Pryor, T., & Hulbert, S. H. (2002). Genetic and molecular characterization of the maize *rp3* rust resistance locus. *Genetics*, *162*, 381–394.

Wisser, R. J., Balint-Kurti, P. J., & Nelson, R. J. (2006). The genetic architecture of disease resistance in maize: a synthesis of published studies. *Phytopathology*, *96*, 120–129.

Wisser, R. J., Kolkman, J. M., Patzoldt, M. E., Holland, J. B., Yu, J., Krakowsky, M., Nelson, R. J., & Balint-Kurti, P. J. (2011). Multivariate analysis of maize disease resistances suggests a pleiotropic genetic basis and implicates a *GST* gene. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 7339–44.

Yang, Z. (2002). Small GTPases: versatile signaling switches in plants. *The Plant Cell*, S375–S389.

Zhang, A., Jiang, M., Zhang, J., Ding, H., Xu, S., Hu, X., & Tan, M. (2007). Nitric oxide induced by hydrogen peroxide mediates abscisic acid-induced activation of the mitogen-activated protein kinase cascade involved in antioxidant defense in maize leaves. *New Phytologist*, *175*, 36–50.

Zhang, Y., Xu, L., Fan, X., Tan, J., Chen, W., & Xu, M. (2012). QTL mapping of resistance to gray leaf spot in maize. *Theoretical and Applied Genetics*, *125*, 1797–808.

Zhou, F., Kurth, J., Wei, F., Elliott, C., Valè, G., Yahiaoui, N., Keller, B., Somerville, S., Wise, R., Schulze-Lefert, P. (2001). Cell-autonomous expression of barley *Mla1* confers race-specific resistance to the powdery mildew fungus via a *Rar1*-independent signaling pathway. *The Plant Cell*, *13*, 337–350.

Zuo, W., Chao, Q., Zhang, N., Ye, J., Tan, G., Li, B., Xing, Y., Zhang, B., Liu, H., Fengler, K. A., Zhao, J., Zhao, X., Chen, Y., Lai, J., Yan, J., & Xu, M. (2015). A maize wall-associated kinase confers quantitative resistance to head smut. *Nature Genetics*, *47*, 151–157

Chapter 2

Quantitative phenotyping of grey leaf spot disease in maize using real-time PCR

J Korsman[1], B Meisel[1,$], FJ Kloppers[2], BG Crampton[1] and DK Berger[1,*]

[1]Department of Plant Science, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa
[2] PANNAR SEED (Pty) Ltd, PO Box 19, Greytown, South Africa
[$]Current address: Monsanto, PO Box 69933, Bryanston, 2021, South Africa

This chapter has been published in the European Journal of Plant Pathology.

I carried out the molecular laboratory work, including maintaining fungal cultures, DNA extractions, PCR, qPCR and PCR-RFLP analysis, and all data analysis, unless otherwise stated. B Meisel designed the experiment, chose the genes to work with, and conducted the DNA isolations under quarantine conditions from diseased maize material collected outside South Africa. She also contributed photographs of infected leaves. I reanalysed the fungal gene sequence and redesigned primers for more specific amplification. FJ Kloppers collected and supplied diseased maize leaf material and identified the pathogens. He also organized the field trials from which the leaf material was collected. BG Crampton helped me compile the manuscript for submission and provided intellectual input. DK Berger provided intellectual input, funding, leaf photography, and edited the manuscript.

We thank W-B. Shim for DNA from *Cercospora* spp., and T.G. Schmidt for digital image analysis.

Abstract

Grey leaf spot is an important maize foliar disease caused by the fungal pathogens *Cercospora zeae-maydis* and *Cercospora zeina*. Although methods exist to detect these Cercospora species in maize, current techniques do not allow quantification of the fungi *in planta*. We developed a real-time SYBR® Green PCR assay for quantification of grey leaf spot disease in maize based on the amplification of a fragment of a cytochrome P450 reductase (*cpr1*) gene. *In planta* fungal DNA content was normalised to a maize glutathione S-transferase III gene (*gst3*) to yield values of ng Cercospora DNA/mg maize DNA. The assay was specific to the two *Cercospora* spp., and we observed no amplification of the *cpr1* fragment in non-target maize leaf pathogens or saprophytes. The assay was employed to quantify *C. zeina* in glasshouse inoculated maize plants and grey leaf spot infected field plants of resistant and susceptible maize lines. In both instances, *C. zeina* DNA content correlated with symptomatic leaf lesion area, and the susceptible maize line contained significantly more *C. zeina* DNA than the resistant line. Sequence differences between the *C. zeina* and *C. zeae-maydis cpr1* amplicons enabled us to perform melt curve analyses to identify the Cercospora species causing grey leaf spot at a particular location. This assay has application in the early detection and quantification of *Cercospora* spp., both of which are important tools in grey leaf spot disease management and maize breeding programmes.

Keywords:

Cercospora zeae-maydis, Cercospora zeina, Precision phenotyping, Quantitative PCR, Grey leaf spot, Gray leaf spot, Real-time PCR, qPCR

Abbreviations:

Ct: cycle threshold
cpr1: cytochrome P450 reductase 1 gene
dpi: days post inoculation
GLS: grey leaf spot / gray leaf spot
gst3: glutathione S-transferase III gene
ITS: internal transcribed spacer
rDNA: ribosomal DNA

2.1 Introduction

Grey leaf spot (GLS) caused by fungi of the genus Cercospora is a foliar disease of maize of great economic importance in many countries. GLS can reduce grain yield by 20 to 60% depending on the level of susceptibility of the hybrid (Latterell and Rossi 1983, Ward et al. 1994). Infected maize crop debris is the source of primary inoculum. After periods of high humidity spores are wind-dispersed or rain-splashed onto bottom leaves of the plant and the disease then progresses upwards. Some infection later in the season also takes place when spores are blown in from adjacent fields. Typically, lesions develop from small spots, which then become tan and rectangular. GLS has a long latent period with 14 to 28 days after infection before lesions produce conidia under humid conditions, which can be seen as a greyish cast. Lesions first run parallel with leaf veins but lesion expansion can result in blighting of entire leaves due to coalescing of lesions (Ward et al. 1999).

Depending on the geographic region, GLS disease of maize can be either caused by *Cercospora zeae-maydis* (formerly known as *C. zeae-maydis* Group I), which has been found throughout various areas of the U.S.A., Canada, Mexico and Brazil, or *Cercospora zeina* (formerly known as *C. zeae-maydis* Group II), which has been reported in the Eastern maize growing states of the U.S.A., Brazil and sub-Saharan Africa (Wang et al. 1998, Dunkle and Levy 2000, Goodwin et al. 2001, Zhu et al. 2002, Okori et al. 2003, Shim and Dunkle 2005, Crous et al. 2006, Brunelli et al. 2008, Meisel et al. 2009). Characteristics and dimensions of conidia and conidiophores produced on infected plants or nutrient media were found to be unreliable criteria for taxonomic differentiation of *C. zeae-maydis* and *C. zeina* (Wang et al. 1998). Crous et al. (2006) developed a PCR-based test to distinguish the two species based on species-specific primers designed from the histone H3 gene. Dunkle and Levy (2000) employed restriction analysis of internal transcribed spacer (ITS) and 5.8S ribosomal DNA (rDNA) regions to differentiate *C. zeae-maydis* and *C. zeina* isolates. Although these methods can distinguish the causal agents of GLS in maize, there are currently no available techniques to quantify *Cercospora* spp. levels in infected maize leaves.

Quantitative PCR (qPCR) offers a reliable and sensitive method for detecting and quantifying fungi *in planta*, and has particular application in pathogen detection prior to symptom development. Recently, De Coninck and coworkers (2012) developed a qPCR assay to quantify *Cercospora beticola* in sugar beet leaves, which will help breeders to discriminate minor differences in resistance in a segregating population. A number of recent studies have

employed qPCR to quantify pathogens on maize material. *Fusarium* spp. are a major source of mycotoxin contamination in maize, and quantification of fungal biomass is essential to understanding interactions between individual species in disease development. Waalwijk et al. (2008) developed a qPCR assay based on a mycotoxin biosynthesis gene to detect mycotoxin producing *Fusarium verticillioides* isolates in maize grain from subsistence farmers in South Africa, and correlated fungal DNA content with mycotoxin levels. Nicolaisen et al. (2009) established a SYBR Green qPCR assay based on the elongation factor 1 alpha gene to detect and quantify eleven *Fusarium* sp. in maize and wheat field material. These studies demonstrate the importance of qPCR for quantification of fungal pathogens in order to select resistant plants in breeding programmes.

The objective of this study was to develop a quantitative PCR assay (qPCR) to measure the amount of *Cercospora* sp. DNA in infected maize leaves. We hypothesised that a qPCR assay based on *cpr1*, a putative cytochrome P450 reductase (van den Brink et al. 1995), would accurately quantify *Cercospora* spp. DNA content in glasshouse inoculated and field infected maize leaves. Results obtained were significant as the qPCR assay was specific to *Cercospora* spp., and showed good correlation between *C. zeina* DNA content and leaf lesion area measured by digital imaging. In addition, we could distinguish between *C. zeae-maydis* and *C. zeina*, both causal agents of GLS in maize. Quantification of *Cercospora* spp. in maize leaves has important application in the breeding and selection of GLS resistant maize lines.

2.2 Materials and Methods

2.2.1 Biological material and host plant infection

The *C. zeae-maydis* ex-type culture CBS 117757 as well as other *C. zeae-maydis* cultures CBS 117755, CBS 117758, CBS 117759, CBS 117760, CBS 117761, CBS 117762, CBS 117763, the *C. zeina* ex-type culture CBS118820, and *Cercospora* sp. CPC 12062 were obtained from the culture collection of the Centraalbureau voor Schimmelcultures, Utrecht, The Netherlands. The *C. zeina* cultures CMW 25463, CMW 25467, CMW 25445, CMW 25442, CMW 25459, CMW 25465, CMW 25462, CMW 25454, CMW 25452 and CMW 25466 are described in Meisel et al. 2009. Fungal cultures were grown on V8 medium (800 ml of distilled water, 200 ml of V8 juice, 15 g of agar, and 2 g of $CaCO_3$) in the dark at 25°C. DNA from *C. zeae-maydis* cultures WB IL and WB KS, and from *C. zeina* cultures WB VA, WB UG and WB BZ were obtained from W-B Shim, Texas A & M University, College Station, Texas, USA.

Maize leaves showing typical symptoms of GLS (*C. zeina*), northern corn leaf blight (*Exserohilum turcicum)*, Phaeosphaeria leaf spot *(Phaeosphaeria maydis*), Hyalothyridium leaf spot (*Hyalothyridium maydis*), zonate leaf spot (*Gloeocercospora sorghi)* and tropical Diplodia (*Stenocarpella macrospora*) were collected from Greytown and other sites in South Africa. Maize leaves showing symptoms of GLS were collected from fields in Brookston, Indiana, USA; Argentina, Mpongwe and Gart farm, Zambia. DNA was isolated under quarantine conditions from diseased maize material collected outside South Africa.

Saprophytes were isolated from GLS lesions by scraping a needle over a lesion and streaking it out on V8 medium. Single spores of *Epicoccum* sp., *Cladosporium* sp., *Phoma* sp., *Tilletiopsis* sp., *Sporobolomyces* sp. and *Cryptococcus* sp. were placed on separate plates and grown for DNA isolation. Sequences of the ITS region amplified using primers ITS1/4 (White et al., 1990) were used for identification by BLASTN analysis against GenBank (Altschul et al. 1990).

Maize plants of a *C. zeina*-susceptible hybrid PAN 6724B were planted in a glasshouse at 28 +/- 4°C under a 16 h day length. Sporulation of *C. zeina* culture CMW 25463 grown on V8 medium was achieved by incubating the plates under equal periods of light and dark (12 h of fluorescent light, 12 h of dark). Conidia were dislodged with a brush and rinsed with 0.01% Tween 20. The inoculum was diluted to 3 x 10^5 conidia/ml and applied onto both surfaces of all three leaves (V3-stage) of the maize plants with a small brush. Control plants were treated

with 0.01% Tween 20. Plants were covered with transparent plastic bags for 5 days to increase humidity levels. Leaves developed typical lesions 13 to 19 dpi. Leaf pieces from different plants were photographed using a SLR digital camera for image analysis, and sampled at 19 days post inoculation (dpi) for DNA extraction.

Field grown maize material was collected from two inbred lines, one that is moderately resistant to grey leaf spot disease (CML444), and another that is susceptible (SC Malawi). Leaf samples were taken 15 weeks after planting from nine individual plants from each line that were positioned randomly within a field site situated in KwaZulu-Natal Province, South Africa. Grey leaf spot disease is prevalent in this region, and therefore the plants were exposed to a natural inoculum of *C. zeina* (Meisel et al, 2009). A 12 cm length of the leaf immediately above the lowest cob from each plant was marked with a permanent black marker pen. Photographs were taken prior to sampling of the leaf piece for subsequent DNA extraction.

2.2.2 Digital image analysis

The proportional area of maize leaf pieces with grey leaf spot lesions from the glasshouse and field experiments were calculated from the photographs of the 12 cm leaf pieces using ASSESS 2.7 software (L. Lamari, American Phytopathological Society, St. Paul, MN, USA), as described in De Coninck et al., 2012. The software carries out automatic measurements by selecting pixels that match certain colour criteria representing lesion or leaf. The lesion area was expressed as percent lesion area relative to the leaf area.

2.2.3 DNA extraction from fungi and plants

DNA from 50 to 100 mg of fungal cultures, as well as from 50 to 100 mg of control or diseased maize leaves was isolated according to Möller et al. (1992). Integrity of the DNA was checked on a 0.8% agarose gel and concentration was determined with a NanoDrop™ 1000 spectrophotometer (Thermo Scientific) at 260 nm.

2.2.4 Conventional and quantitative PCR amplification

Two cDNA sequences of a *C. zeae-maydis* putative cytochrome P450 reductase (GenBank accession numbers AF448828 and FG242129) were used to compile a composite sequence for primer design, resulting in primers CPR1_1F (5'-TCCACTCTCGCTCAATTCG-3') and CPR1_1R (5'-GCCTTCATCGCCATATGTTC-3'). These primers, designated CPR1_1, were used to amplify a 1.1 kb product from *C. zeina* genomic DNA, that was sequenced on an ABI

PRISM™ 3100 Automated DNA sequencer (Applied Biosystems) using the ABI Prism Big Dye Terminator Cycle sequencing reaction kit v3.1 (Applied Biosystems). The internal primers, designated CPR1_2, and named CPR1_2F (5'-TGAACTACGCGCTCAATG-3') and CPR1_2R (5'-TCTCTCTTGGACGAAACC-3') were designed in two short regions of 100% identity between the two species. The primers GST3F (5'-GGAGCCCTGAGTCGAATAAAAG-3') and GST3R (5'-AACACACACGAAAGGCAACAGT-3') were used to amplify a 106 bp-fragment of the glutathione S-transferase III gene *gst3* from maize (GenBank accession number X06755).

For conventional PCRs, reaction volumes of 15 µl consisted of 1x NH_4 PCR reaction buffer, 2.7mM of $MgCl_2$, 0.1µM of each dNTP (Bioline), 0.2µM each of the primers, 1 U of BIOTAQ™ DNA Polymerase (Bioline), 11.15 µl of sterile distilled water and 10 ng of DNA template. Cycling conditions were 2 min at 94°C followed by 35x (15 s at 94°C, 15 s at 60°C, 20 s at 72°C). A final elongation step was performed at 72°C for 5 min. Quality of all PCR products was verified by gel electrophoresis.

The *cpr1*-and *gst3*-fragments amplified in conventional PCRs were ligated into the pJET1.2/blunt vector from the CloneJET™ PCR Cloning Kit (Fermentas) and subsequently transformed into chemically competent *Escherichia coli* using manufacturer's instructions. Transformed *E. coli* were grown at 37°C overnight in Luria Bertani broth supplemented with 100 mg/ml ampicillin. Purification of the plasmid DNA was performed using the Invisorb Spin Plasmid Mini Two Kit (Invitek). Plasmid DNA was used as template for sequencing reactions as described above. Sequences were checked by BLASTN (Altschul et al. 1990).

Quantitative PCR amplifications were performed in a total volume of 10 µl on a LightCycler® 480 instrument (Roche Diagnostics Corp.). Each reaction contained 1 µl of the DNA template or water in the non-template controls, 5 µl of the LightCycler® 480 SYBR Green I Master Mix (Roche Diagnostics Corp.), 0.2µM of each primer, 3.6 µl of sterile distilled water. Each PCR was run in triplicate. Cycling conditions consisted of 5 min at 95°C followed by 45x (10 s at 95°C, 10 s at 60°C, 5 s at 72°C). Fluorescence was detected at the end of each elongation step. The amplification specificity of the PCR was investigated by melting curve analysis after every run.

The specificity of the CPR1_2 primers was tested on 10 ng of DNA isolated from healthy maize leaves and maize leaves with typical symptoms of GLS (*C. zeae-maydis* and *C. zeina*), northern corn leaf blight (*E. turcicum*), Phaeosphaeria leaf spot *(P. maydis*), Hyalothyridium leaf spot (*H. maydis*), zonate leaf spot (*G. sorghi)*, tropical Diplodia (*S. macrospora*), and a number of fungi and yeast isolated from GLS lesions (*Epicoccum* sp.*, Cladosporium* sp.*, Phoma* sp.*, Tilletiopsis* sp.*, Sporobolomyces* sp.*, Cryptococcus* sp.). The specificity of the GST3 primers was tested on 10 ng of DNA isolated from healthy maize leaves*, C. zeina,* and *C. zeae-maydis* cultures and lesion samples. To exclude false negative results template DNA samples from lesions were tested for PCR amplification using primers ITS1/4 following the method of White et al. (1990).

In order to estimate the amount of fungal DNA in the infected leaf samples, DNA extracted from pure *C. zeina* and *C. zeae-maydis* cultures was diluted in maize DNA of concentration 10 ng μl^{-1}. The final fungal DNA concentrations were 5, 1, 5×10^{-1}, 2.5×10^{-1}, 1×10^{-1}, 5×10^{-2}, 2.5×10^{-2}, 1×10^{-2}, 5×10^{-3}, 2.5×10^{-3}, 1×10^{-3} ng μl^{-1}. These dilutions were used to determine the detection limits of the CPR1_2 primer pair in a simulated lesion sample. A serial dilution of DNA extracted from healthy leaves (2×10^{1}, 1.5×10^{1}, 1.25×10^{1}, 1×10^{1}, 5, 1, 5×10^{-1}, 1×10^{-1}, 5×10^{-2} ng μl^{-1}) was prepared to measure detection limits of the GST3 primers.

Standard curves were prepared by plotting $\log_{10}$ of the DNA concentration of known standards against the cycle number at which the fluorescent signal from the amplified PCR products surpassed the detection threshold. Regression curves were drawn and the qPCR efficiency was calculated as equal to $10^{(-1/slope)}$, the optimal efficiency thus equalling two. The amount of target DNA for unknown samples was extrapolated from the respective standard curves. To normalise gene quantification between different samples the amount of fungal *cpr1* was divided by the amount of plant *gst3* quantified in infected leaves.

2.3 Results

2.3.1 Design of a specific PCR assay for *Cercospora* spp. in infected maize leaves

CPR1_1 primers were designed to the sequence of the *C. zeae-maydis cpr1* gene, which encodes a cytochrome P450 reductase first described in *Aspergillus niger* (van den Brink et al. 1995). These primers were used to amplify 967 bp of this gene from *C. zeae-maydis* and 1,024 bp from *C. zeina*. Alignment of the *cpr1* gene sequences from *C. zeae-maydis* and *C. zeina* enabled us to design internal primers (designated CPR1_2; see Materials and Methods) to conserved areas of the gene sequence (Figure 2.1). Conventional and quantitative PCRs produced a 164bp fragment in *C. zeina* and *C. zeae-maydis* isolates, but no amplicon was obtained from uninfected maize leaves or leaves infected with non-target maize pathogens (*Exserohilum turcicum, Phaeosphaeria maydis, Hyalothyridum maydis, Gloeocercospora sorghi, Stenocarpella macrosporum*) (Figure 2.2a). In addition, the CPR1_2 primer set did not produce an amplification product from saprophytic organisms (*Epicoccum* sp., *Cladosporium* sp., *Phoma* sp., *Tilletiopsis* sp., *Sporobolomyces coprosmae, Cryptococcus flavescens*) isolated from maize leaves. As a check for DNA quality, all samples from infected leaves and saprophytic organisms were successfully subjected to conventional PCR analysis using the ITS1 and ITS4 primers (Figure 2.2c).

```
          5' TGAACTACGCGCTCAATG 3' CPR1_2F primer
             1                                                          60
Czm_cpr1     TGAACTACGCGCTCAATGGTCCACGAAACAAGTACGATGGCATCCACGTCCCGGTCCACA
Cz_cpr1      TGAACTACGCGCTCAATGGTCCACGAAACAAATACGATGGCATCCACTTCCCGGTTCACA
             ****************************** ************** ******* ****

             61                                                        120
Czm_cpr1     TTCGACACTCGAACTTCAAGCTCCCCTCGGATCCAAGCAAGCCGATCATCATGGTTGGCC
Cz_cpr1      TTCGACACTCGAACTTCAAGCTCCCCTCAGACCCAAGCAAGCCCATTATCATGGTTGGTC
             *************************** ** *********** ** *********** *

             121                                    164
Czm_cpr1     CTGGTACCGGCGTTGCTCCATTCCGCGGTTTCGTCCAAGAGAGA
Cz_cpr1      CTGGCACCGGCGTTGCTCCATTCCGCGGTTTCGTCCAAGAGAGA
             **** ****************************************
                     CPR1_2R primer  3'CCAAAGCAGGTTCTCTCT 5'
```

Figure 2.1: Alignment of a 164 bp fragment of *cpr1* from *C. zeae-maydis* CBS 117757 (Czm_cpr1) and *C. zeina* CBS 118820 (Cz_cpr1). The CPR1_2F and CPR1_2R primers used for the qPCR assay that are identical between the species are shown. *Nla*IV sites are underlined. Consensus nucleotides are indicated with an asterisk.

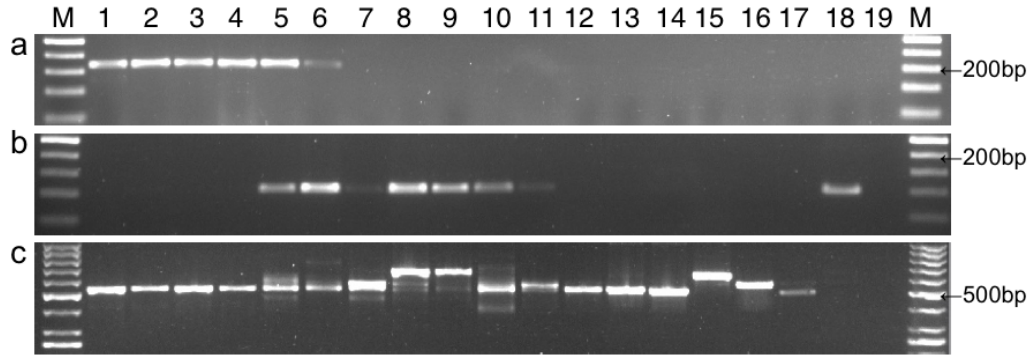


Figure 2.2: PCR amplification of *cpr1* (a), *gst 3* (b) and ITS1/4 (c) gene fragments from *C. zeae-maydis* and *C. zeina* cultures, DNA extracted from grey leaf spot lesions, other maize foliar pathogens, and saprophytic organisms isolated from grey leaf spot lesions. (a) CPR1_2 primers amplified fragments of 164 bp from *C. zeae-maydis* and *C. zeina*, but not from other fungal pathogens and saprophytes isolated from grey leaf spot lesions. (b) GST primers amplified the *gst3* fragment only from samples derived from lesions on maize leaves. (c) The ITS1/4 region was amplified from all samples containing fungi/yeast and indicates DNA quality is adequate for DNA amplification. Lanes M, 50 bp molecular marker (Fermentas). Templates used for each reaction were as follows: lanes 1-4, *C. zeae-maydis* CBS117755, *C. zeae-maydis* CBS117757, *C. zeina* CBS118820, *C. zeina* CMW25467 respectively; lanes 5-6, DNA isolated from lesions containing *C. zeae-maydis* from Argentina, and *C. zeina* from Greytown, RSA, respectively; lanes 7-11, DNA isolated from maize foliar diseases caused by *Exserohilum turcicum* (Northern corn leaf blight), *Phaeosphaeria maydis* (Phaeosphaeria leaf spot), *Hyalothyridum maydis* (Hyalothyridum leaf spot), *Gloeocercospora sorghi* (zonate leaf spot), *Stenocarpella macrosporum* (tropical diplodia), respectively; lanes 12-17, grey leaf spot saprophytes *Epicoccum* sp., *Cladosporium* sp., *Phoma* sp., *Tilletiopsis* sp., *Sporobolomyces coprosmae* and *Cryptococcus flavescens*, respectively; lane 18, non-diseased maize tissue; lane 19, no template PCR control.

Since *cpr1* appears to be conserved in many fungal species, the specificity of the primers was verified by comparison to corresponding regions of the *cpr1* orthologues from other fungi available from GenBank (Figure 2.3). The closest *cpr1* orthologue is from the non-plant pathogen *Podospora anserina*, and this shows five and four mismatches with the CPR1_2F and CPR12R primers, respectively (Figure 2.3). Other close orthologues from plant pathogens such as *Pyrenophora* spp. pathogenic on wheat, as well as *Alternaria brassicicola,* show increasing numbers of mismatches with the primers. It is therefore unlikely that amplification would occur from related pathogens on maize under the stringent conditions of the qPCR assay. This was borne out by the empirical evidence that shows no amplification from other fungi from maize (Figure 2.2a), whose *cpr1* sequences are not yet available on GenBank.

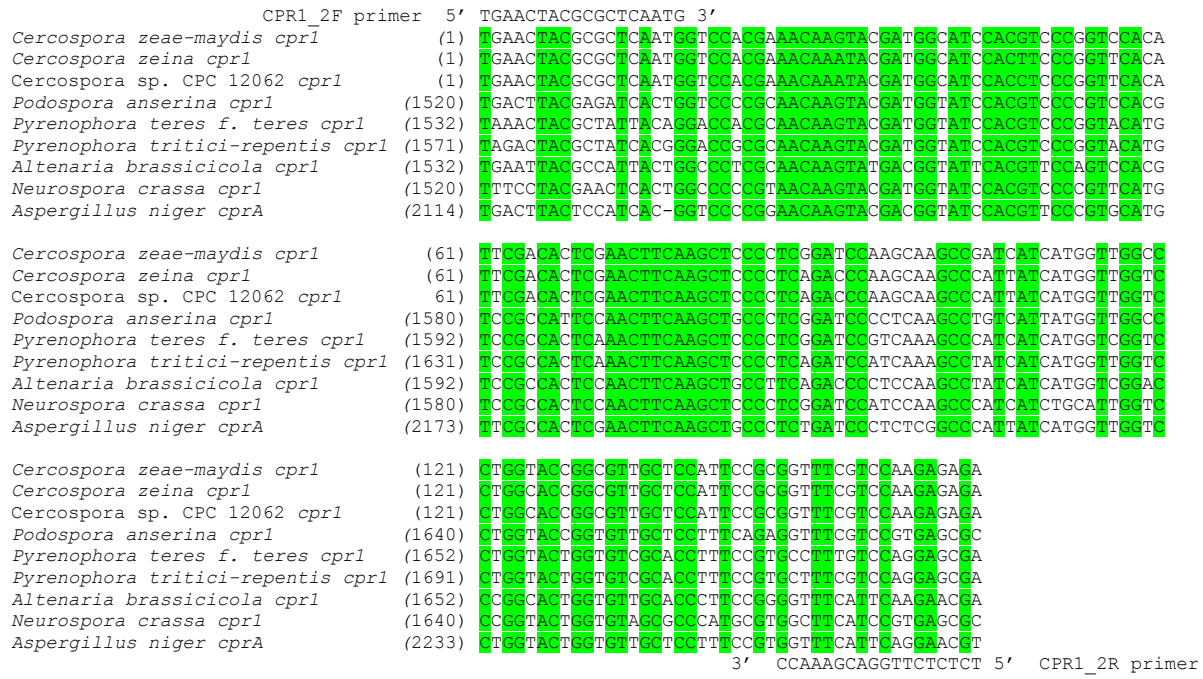


Figure 2.3: Alignment of a 164 bp fragment of *cpr1* amplified with the CPR1_2 primer pair from *C. zeae-maydis, C. zeina* and *Cercospora* sp. CPC12062 with putative close orthologues from other fungi. The sequences of the CPR1_2F and CPR1_2R primers are shown. The GenBank accession numbers of putative orthologues of *cpr1* obtained by BLASTN search of the *C. zeae-maydis cpr1* sequence against GenBank (nr database) were as follows: *Podospora anserina cpr1* (XM_001910139); *Pyrenophora teres f. teres cpr1* (XM_003297092); *Pyrenophora tritici-repentis cpr1* (XM_001939488); *Altenaria brassicicola cpr1* (AB506083); *Neurospora crassa cpr1* (XM_959350); and *Aspergillus niger cprA* (Z26938). Identical nucleotides are indicated in green.

Conventional PCR with GST3F/GST3R primers produced specific amplicons of the expected 106 bp when templates contained DNA from maize (Figure 2.2b). Sequencing of *gst3* amplicons showed a nucleotide sequence exhibiting 100% identity with the *gst3* gene from *Zea mays* (GenBank accession number X06755). No PCR products were amplified from pure fungal cultures of *C. zeae-maydis* or *C. zeina* using the GST3F/GST3R primers (Figure 2.2b).

2.3.2 qPCR method to quantify *Cercospora* spp. in infected maize tissue

To quantify the pathogen in biological samples, a dilution series of genomic DNA from *C. zeina* and *C. zeae-maydis* was added to 10 ng of DNA extracted from healthy maize creating progressively lower fungal DNA concentrations from 5 to 1×10^{-3}ng μl^{-1} to generate standard curves for *cpr1* (Figure 2.4a and b, respectively). Efficiency of amplification of the 164 bp *cpr1* fragment in maize carrier DNA was 1.84, with a minimum detection limit of 5 pg *C. zeina* or *C. zeae-maydis* DNA. Quantifications of the pathogens showed a linear relationship between

log_{10} values of the amount of DNA and cycle threshold values, indicating that the method is suitable as a quantitative assay over the concentration range tested. The qPCR efficiency of maize *gst3* was 1.68 (Figure 2.4c). Both melt curve analyses and agarose gel electrophoresis of qPCR reactions indicated that a single product was obtained for *C. zeina* and *C. zeae-maydis cpr1* and for maize *gst3*.

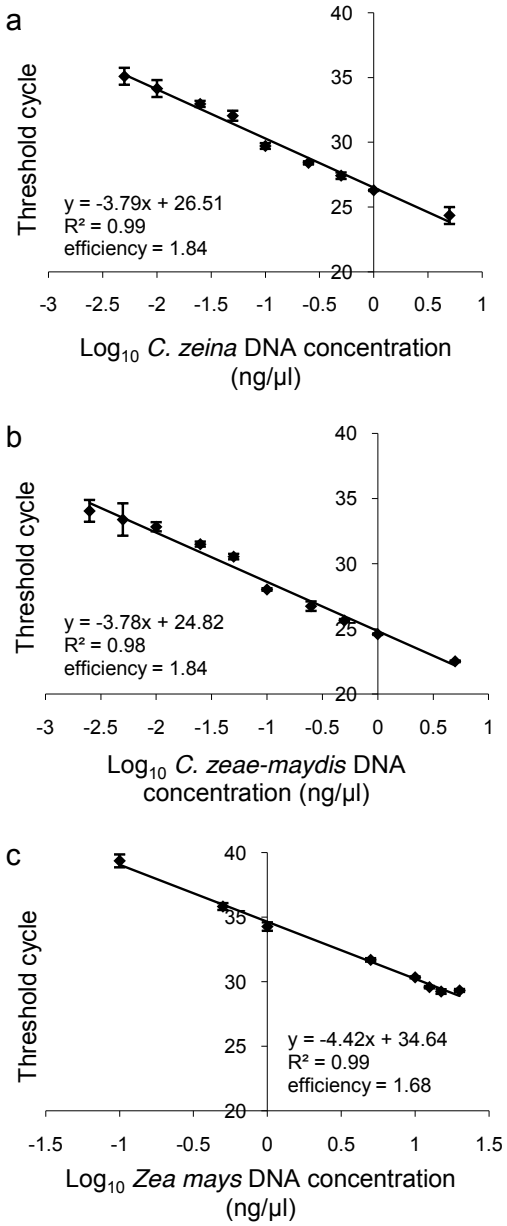


Figure 2.4: Standard curves used for qPCR assay to quantify *C. zeina* and *C. zeae-maydis in planta*. The graphs illustrate the efficiency and sensitivity of *cpr1* and *gst3* fragment amplification using the CPR1_2 primers and GST primers, respectively. Figures (a) and (b) represent standard curves of *C. zeina* and *C. zeae-maydis* DNA diluted in maize carrier DNA to a total of 10 ng DNA, respectively. Figure (c) indicates a standard curve for *gst3* amplification from serial dilutions of maize genomic DNA.

2.3.3 qPCR quantification of *C. zeina* DNA in glasshouse inoculated maize leaves before and after lesion development

Amplification of the *cpr1* gene by qPCR was undertaken in order to assess *C. zeina* content in maize leaves after inoculation with *C. zeina* in the glasshouse. DNA extractions were performed on three leaf segments (Figure 2.5 a, b and c) with approximately 2, 4 and 16% infected leaf area, respectively, and subjected to *cpr1* and *gst3* qPCR. The amount of *C. zeina* DNA in each leaf sample was determined from a *cpr1* standard curve (Figure 2.4a), and normalised to the amount of maize DNA as calculated from a *gst3* standard curve (Figure 2.4c). In this manner, leaf samples in Figure 2.5 a-c were estimated to contain 4.7 ± 0.3, 7.8 ± 1.1 and 46.0 ± 9.2 ng *C. zeina* DNA/mg maize DNA, which was proportional to 2, 4 and 16% infected leaf area, respectively. In addition, we were able to detect *C. zeina* DNA in inoculated maize leaves prior to the development of GLS symptoms (results not shown). These results verified that the qPCR assay using *cpr1* and *gst3* genes was an effective method for quantifying grey leaf spot disease in glasshouse experiments.

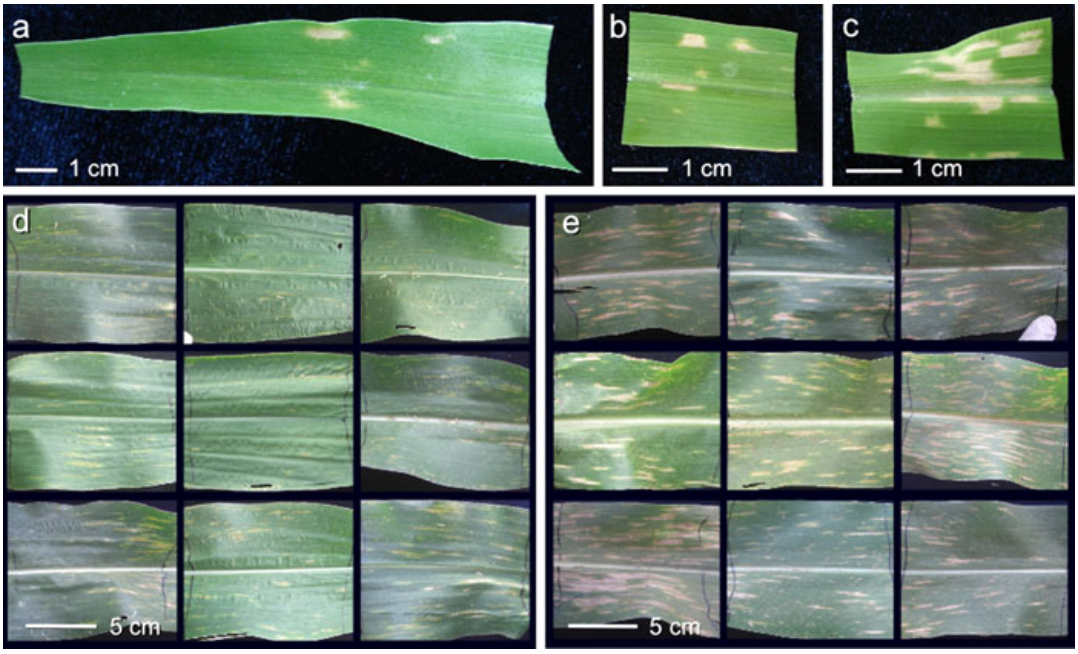


Figure 2.5: Photographs of leaves with symptoms of GLS from glasshouse and field grown maize that were used for GLS disease quantification shown in Figure 2.6. Figures (a), (b) and (c) indicate leaf pieces with different areas of GLS lesions from glasshouse grown maize plants (hybrid PAN 6724B) at 19 days post inoculation with *C. zeina*. Figures d and e show leaf pieces of nine biological replicates each from two field grown maize lines CML444 and SC Malawi that are resistant and susceptible to grey leaf spot, respectively.

2.3.4 *C. zeina* fungal content in GLS resistant and susceptible maize lines

The qPCR assay was applied to quantify *C. zeina* in field infected maize leaves. Lesion area was determined for nine biological replicates from each of two maize inbred lines CML444 and SC Malawi (resistant and susceptible, respectively) (Figure 2.5d and e) using digital image analysis. Differences in the number and area of lesions were observed between the two maize lines (Figure 2.5d and 4e). DNA was subsequently extracted from these samples and subjected to *cpr1* and *gst3* qPCR. The amount of *C. zeina* DNA was normalised relative to the amount of maize DNA (*gst3* qPCR), and a comparison was made to leaf lesion area (Figure 2.6). There was good correlation between amount of *C. zeina* DNA and leaf lesion area (Pearson correlation = 0.76), and the quantification of *C. zeina* in the resistant maize line was significantly less than the susceptible line as determined by ng *C. zeina*/mg maize DNA (Student's *T*-test; $p<0.005$) and scanned lesion area (Student's *T*-test; $p<5 \times 10^{-5}$).

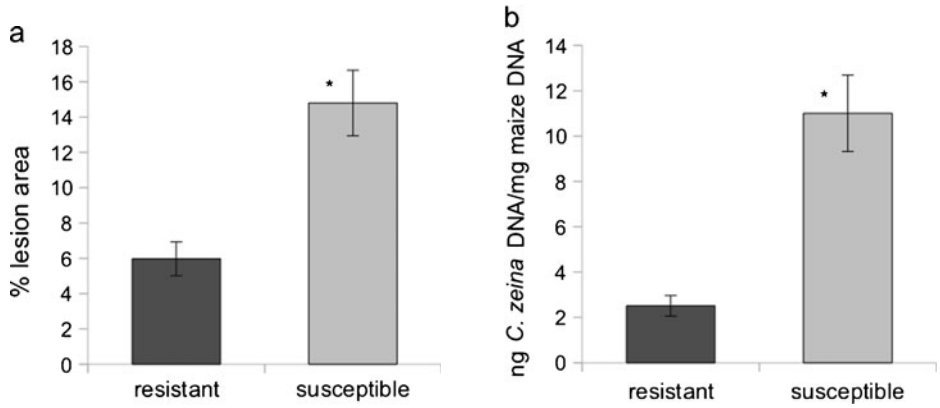


Figure 2.6: Comparison of *C. zeina* infection in resistant and susceptible maize lines using (a) digital image analysis and (b) qPCR. (a) Nine biological replicate leaf pieces from each resistant (CML444) and susceptible (SC Malawi) maize line were photographed and GLS lesion area relative to total leaf area was determined for the two lines. (b) Thereafter, DNA was extracted from these maize leaf pieces and subjected to the qPCR assay to determine ng Cercospora DNA per mg maize DNA. Bars on graphs indicate standard error. Asterisks indicate that the susceptible cultivar shows significantly different values to the resistant cultivar (Student's *T*-test; $p<0.005$).

2.3.5 Implementation of qPCR assay to differentiate between *C. zeae-maydis* and *C. zeina*

GLS in maize is caused by both *C. zeae-maydis* and *C. zeina*. Sequence differences between *C. zeina* and *C. zeae-maydis* in the 164 bp *cpr1* amplicon (Figure 2.1: nine nucleotides differ) allowed us to detect which of the two species was the causal agent of GLS on a symptomatic maize leaf. Melting curve analysis of known *C. zeae-maydis* and *C. zeina* isolates yielded significantly different single peaks of 85.28±0.07 and 84.53±0.07 respectively (two-tailed T-test, $p<0.05 \times 10^{-15}$) (Table 1; Figure 2.7). Importantly, this differentiation was clear for a wide range of isolates tested, namely ten isolates of *C. zeae-maydis* from different sites in the USA, and 14 isolates of *C. zeina* from diverse sites in the USA and Africa, including samples amplified directly from GLS lesions (Table 1).

The qPCR assay was also successful in amplification of the 164 bp *cpr1* product from a distinct Cercospora isolate from maize in South Africa, named *Cercospora* sp. CPC 12062 (Crous et al. 2006), which yielded a Tm = 84.59±0.14, which is not significantly different from *C. zeina* (Table 1; Figure 2.7). This was to be expected since the sequences corresponding to the CPR1_2 primers were identical, and there was only one base pair difference between the 164 bp *cpr1* amplicons of these two species (Figure 2.3).

An alternative differentiation between the two species known to be pathogens was evident from PCR-RFLP analysis of the 164 bp *cpr1* product, since the *C. zeae-maydis* amplicon has two *Nla*IV sites whereas *C. zeina* (as well as *Cercospora* sp. CPC 12062; Figure 2.3) has one *Nla*IV site. *Nla*IV digestion of the *C. zeae-maydis* amplicon yielded three fragments of 91, 39 and 34 bp, whereas the *C. zeina* amplicon yielded two fragments of 125 and 39 bp (Figure 2.8).

Table 2.1: Melting temperatures of 164 bp qPCR products obtained with CPR1_2 primers from *Cercospora* spp. isolated from maize.

Species	Isolate no./code[a]	Location[b]	Year[c]	Tm[d]	aveTm[e]
C. zeae-maydis	CBS 117755	Indiana, USA	2003	85.27±0.05	
	CBS 117757*	Wisconsin, USA	2000	85.31±0.03	
	CBS 117758	Iowa, USA	2004	85.22±0.05	
	CBS 117759	Tennessee, USA	1999	85.24±0.03	
	CBS 117760	Pennsylvania, USA	1999	85.33±0.02	**85.28±0.07**
	CBS 117761	Indiana, USA	1999	85.33±0.02	
	CBS 117762	Missouri, USA	2000	85.32±0.02	
	CBS 117763	Iowa, USA	1999	85.33±0.03	
	WB IL	Illinois, USA	-	85.12±0.05	
	WB KS	Kansas, USA	-	85.32±0.06	
C. zeina	CBS 118820*	KZN, South Africa	2005	84.48±0.05	
	CMW 25442	Art Farm, Zimbabwe	2007	84.56±0.04	
	CMW 25445	Gart Farm, Zambia	2007	84.52±0.02	
	CMW 25452	Karkloof, KZN, SA	2007	84.49±0.03	
	CMW 25454	Karkloof, KZN, SA	2007	84.45±0.02	
	CMW 25459	Winterton, KZN, SA	2007	84.53±0.00	
	CMW 25462	Cedara, KZN, SA	2007	84.44±0.01	
	CMW 25463	Greytown, KZN, SA	2007	84.48±0.02	
	CMW 25465	Cedara, KZN, SA	2007	84.49±0.06	**84.53±0.07**
	CMW 25466	Karkloof, KZN, SA	2007	84.56±0.03	
	CMW 25467	Mkushi, Zambia	2007	84.51±0.02	
	lesion DNA GF	Gart Farm, Zambia	2010	84.65±0.01	
	lesion DNA Mp	Mpongwe, Zambia	2010	84.64±0.01	
	lesion DNA Gtn	Greytown, KZN, SA	2010	84.68±0.06	
	WB VA	Virginia, USA	-	84.49±0.01	
	WB UG	Kamengo-Mpigi, Uganda	-	84.49±0.04	
	WB BZ	Goias, Brazil	-	84.57±0.07	
Cercospora sp.	CPC 12062	KZN, South Africa	-	84.59±0.14	

* ex-type cultures
[a] CBS = culture collection of the Centraalbureau voor Schimmelcultures (CBS), Utrecht, The Netherlands.
CMW = culture collection of the Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, South Africa.
CPC: Culture collection of Pedro Crous, housed at CBS.
WB = isolates from W-B. Shim.
[b] Location where the diseased maize material was collected.
[c] Year of GLS lesion collection.
[d] Melting temperature analyses were performed on the three technical replicates for each sample to obtain an average melting temperature and standard deviation.
[e] Melting temperatures for isolates from each species were averaged to obtain a species melting temperature.

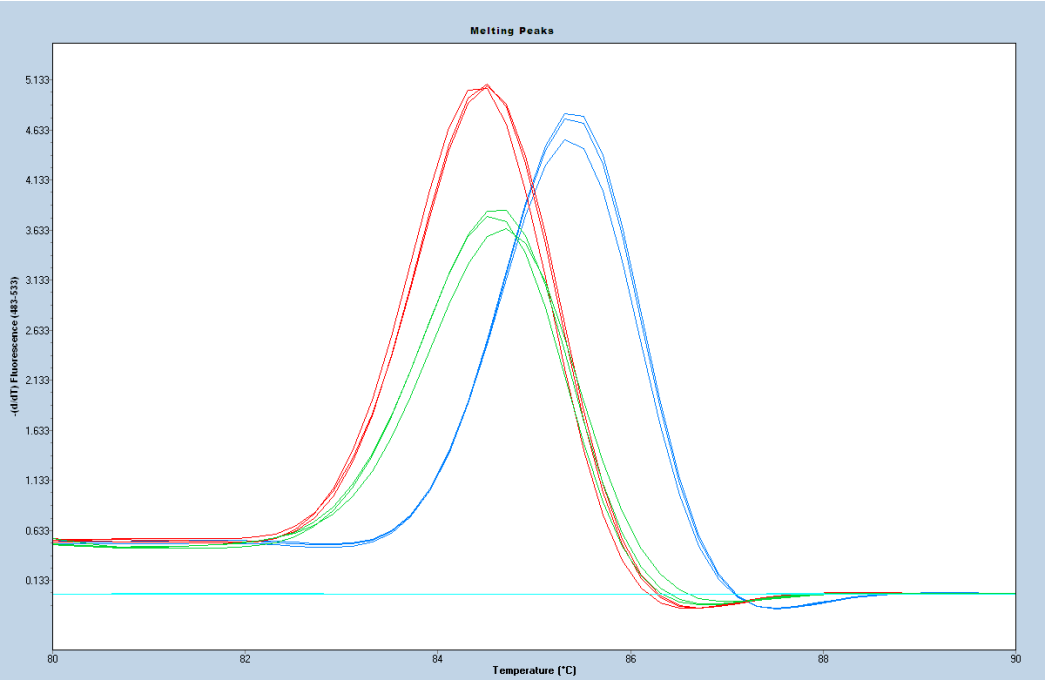


Figure 2.7: Melting temperature (Tm) graphs of 164 bp qPCR amplification products using CPR1_2 primers for *C. zeina* (CMW25463) (red), *C. zeae-maydis* (CBS117757) (blue) and *Cercospora* sp. (CPC12062) (green). Graphs for three technical replicates of each sample are shown. The peaks of the graphs indicate the temperature at which 50% of the amplicons have denatured (Tm). The samples were analysed on a LightCycler® 480 instrument with SYBR Green I labelling (Roche Diagnostics Corporation, Basel, Switzerland).

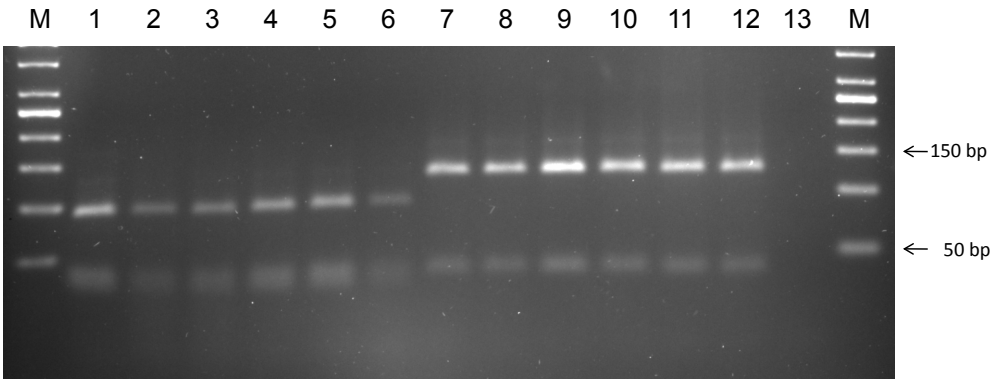


Figure 2.8: PCR-RFLP analysis of the 164 bp *cpr1* amplicon from *C. zeae-maydis* and *C. zeina*. *Nla*IV digestion of the *C. zeae-maydis* 164 bp *cpr1* amplicon yielded three fragments of 91, 39 and 34 bp (lanes 1-6), whereas the *C. zeina* amplicon yielded two fragments of 125 and 39 bp (lanes 7-12). Lane M, 50 bp molecular marker (Fermentas). Templates used for each reaction were as follows: lanes 1-6, *C. zeae-maydis* cultures CBS117757, CBS117755, CBS117759, CBS117760, WB KS, DNA extracted from a GLS lesion from Argentina, respectively; lanes 7-12, *C. zeina* cultures CBS118820, CMW25467, WB VA, WB UG, WB BZ, CMW25442; and lane 13, no template PCR control.

2.4 Discussion

In this study, we developed a qPCR method to detect and quantify *C. zeae-maydis* and *C. zeina* DNA in maize leaves. Primers were designed to a cytochrome P450 reductase gene (*cpr1*) (Figure 2.1), and the specificity was confirmed by testing against uninfected maize leaves and leaves infected with non-target maize pathogens (Figure 2.2). Primers designed to the glutathione S-transferase III gene (*gst*3) from maize proved to be plant specific by the absence of cross-reaction with fungal DNA. Normalisation with the maize *gst3* was performed to compensate for differences in the amount of maize DNA in each sample.

The functional role of the product of the *cpr1* gene in *Cercospora* spp. is not yet known, although sequence similarity indicates that it has cytochrome p450 reductase activity. One of the ESTs from which the CPR1_1 primers were designed (AF448828) was initially annotated as *ctb4*, one of the cercosporin biosynthesis genes (Shim and Dunkle 2002). However, this EST does not correspond to any of the three oxidoreductases of the well-characterized cercosporin biosynthetic cluster in *Cercospora nicotianae* (Chen et al 2007). This was supported by the fact that the second EST (FG242129) was cloned from *C. zeae-maydis* growing vegetatively, which are conditions when cercosporin biosynthesis is repressed (Bluhm et al 2008). CPR1 was first described in *A. niger* (GenBank accession number Z26938), and appears to be conserved in many fungal species (Figure 2.3). Importantly, the gene regions corresponding to the primers used in the qPCR assay for quantifying *Cercospora* spp. in this study are sufficiently variable in other fungal species (Figure 2.3), and were shown in our study to be specific to *Cercospora* spp. (Figure 2.2).

De Coninck and co-workers (2012) recently developed a qPCR assay to detect and quantify *Cercospora beticola* in sugar beet cultivars. These authors utilised a Taqman® probe and associated primers to detect the *C. beticola* calmodulin gene in infected sugarbeet material. Fungal DNA levels were normalised to levels of an endogenous sugar beet DNA sequence. As was the case with our study, these authors found a significant correlation between fungal DNA levels and leaf lesion area, as well as a significant difference between *C. beticola* levels in susceptible and resistant sugar beet cultivars. Their assay is of particular value in glasshouse screening of sugar beet in which the inoculum is limited to *C. beticola*. Although the authors state that the qPCR assay did not amplify from two other *Cercospora* species, it was not tested in field experiments. In contrast, we have demonstrated our qPCR assay to be specific and

effective in maize field-grown material (Figures 2, 4 and 5), conditions under which screening for quantitative resistance to this disease is most desirable.

Our qPCR assay was sensitive, and could detect the target pathogen in as little as 5 pg DNA for both *C. zeae-maydis* and *C. zeina* (Figure 2.4). This represents an improvement in sensitivity compared to other DNA based detection methods such as dot blots that were able to detect 160 pg of DNA from *Sporisorium reiliana*, the cause of head smut in maize (Xu et al., 1999). Alternative approaches that are not based on DNA are less sensitive, such as ELISA-based assays (Ward et al., 2004) or measurement of the fungus-specific lipid ergosterol, which has been used for quantification of *Cercospora kikuchii* in soybean (Xue et al., 2006). Ergosterol is present in many fungal species, and thus its use is limited to controlled inoculations. In contrast, our qPCR assay is specific to *Cercospora* spp. from maize, and therefore can be applied in field experiments.

A detection limit of 5 pg for both *C. zeae-maydis* and *C. zeina* in our qPCR assay corresponds to approximately 125 genome equivalents. This is based on the assumption that both species have a genome size of ~40 Mb. This was the average genome size for ten of twelve sequenced Dothidiomycete genomes (March 2011; www.osti.gov/bridge/purl.cover.jsp?purl=/1012481-8TJMwv/). Nicolaisen et al (2009) achieved a lower detection limit of 2.5 genome equivalents (0.1 pg) for *Fusarium* spp. in maize using primers to the elongation factor 1 alpha gene. A TaqMan® PCR assay based on *cpsD* from the maize pathogen *Pantoea stewartii* was able to detect 1 pg of purified *P. stewartii* DNA, the equivalent of 10^4 colony forming units per millilitre (Tambong et al. 2008). We were unable to compare the sensitivity of our qPCR assay with that for *C. beticola*, as De Coninck et al. (2012) based their detection limits on ΔCt values calculated by subtracting the Ct value of an endogenous sugar beet control from the Ct value of the *C. beticola* calmodulin gene.

C. zeae-maydis and *C. zeina* are closely related species, which both cause GLS in maize, but occur in different geographic regions. We have recently confirmed that *C. zeina* remains the causal agent of GLS of maize in southern Africa and that absence of *C. zeae-maydis* is highly likely (Meisel et al. 2009). Since occurrence of both *Cercospora* species in one field is rare and has only been reported for some locations in Ohio and New York in the U.S.A. (Wang et al. 1998), it is not crucial that the quantitative assay presented here cannot quantify *C. zeina* and *C. zeae-maydis* separately within a sample. Sequence analysis of the 164 bp *cpr1*-amplicon

revealed nine bp differences between the two *Cercospora* species (Figure 2.1). These sequence differences allowed us to employ melt curve analyses to distinguish between species, and could be utilised to detect a mixed infection within a maize leaf (Table 1; Figure 2.7). There was only one base pair difference between the *cpr1* amplicon from *C. zeina* and *Cercospora* sp. CPC 12062, a distinct isolate from maize in South Africa (Crous et al., 2006) (Figure 2.3), and thus these two species could not be distinguished based on Tm (Table 1; Figure 2.7). It is not known if this isolate is a pathogen on maize, however it would be quantified by the qPCR assay if present.

Resistance to grey leaf spot disease is generally quantitative in nature (Balint-Kurti et al., 2008), and it has been suggested that an understanding of the molecular basis of quantitative disease resistance requires exploitation of maize genetic diversity combined with improved phenotyping approaches (Poland et al., 2009). The qPCR assay reported here represents a tool that can be applied in this type of "precision phenotyping" approach for assessment of both QTL and association mapping populations of maize for quantitative resistance to GLS (Rafalski 2010). Thus, this *Cercospora*-specific qPCR assay was further employed for QTL mapping of fungal content of leaves from a field setting (see Chapter 3) to demonstrate its potential future application.

We have demonstrated the use of a qPCR assay based on the *cpr1* gene to quantify GLS disease in maize. In this manner, we were able to show good correlation between *C. zeina* DNA levels and GLS lesion area on maize leaves in the glasshouse and field. Because of its specific, precise, cost-effective and fast nature, the assay presented here will be a useful tool in grey leaf spot disease management, breeding programmes and plant-pathogen interaction studies, and thus mitigate the threat of this pathogen to food security in maize producing countries of the world.

2.5 References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*, 403-410.

Balint-Kurti, P.J., Wisser R., & Zwonitzer, J.C. (2008). Use of an advanced intercross line population for precise mapping of quantitative trait loci for gray leaf spot resistance in maize. *Crop Science*, *48*, 1696-1704.

Bluhm, B., Dhillon, B., Lindquist, E., Kema, G., Goodwin, S., & Dunkle, L. (2008). Analyses of expressed sequence tags from the maize foliar pathogen *Cercospora zeae-maydis* identify novel genes expressed during vegetative, infectious, and reproductive growth. *BMC Genomics, 9*, 523.

Brunelli, K. R., Dunkle, L. D., Sobrinh, C. A., Fazza, A. C. & Camargo, L. E. A. (2008). Molecular variability in the maize grey leaf spot pathogens in Brazil. *Genetics and Molecular Biology*, *31*, 938-942.

Chen, H., Lee, M.H., Daub, M.E., & Chung, K.R. (2007). Molecular analysis of the cercosporin biosynthetic gene cluster in *Cercospora nicotianae*. *Molecular Microbiology 64*, 755-770.

Crous, P.W., Groenewald, J.Z., Groenewald, M., Caldwell, P., Braun, U. & Harrington, T.C. (2006). Species of *Cercospora* associated with grey leaf spot of maize. *Studies in Mycology*, *55*, 189-197.

De Coninck, B.M.A., Amand, O., Delaur, S.L., Lucas, S., Hias, N., Weyens, G., Mathys, J., De Bruyne, E. & Cammue, B.P.A. (2012). The use of digital image analysis and real-time PCR fine-tunes bioassays for quantification of *Cercospora* leaf spot disease in sugar beet breeding. *Plant Pathology*, *61*, 76–84

Dunkle, L.D. & Levy, M. (2000). Genetic relatedness of African and United States populations of *Cercospora zeae-maydis*. *Phytopathology*, *90*, 486-490.

Goodwin, S.B., Dunkle, L.D. & Zismann,V.L. (2001). Phylogenetic analysis of *Cercospora* and *Mycosphaerella* based on the internal transcribed spacer region of ribosomal DNA. *Phytopathology*, *91*, 648-658.

Latterell, F. M. & Rossi, A. E. (1983). Gray leaf spot of corn: A disease on the move. *Plant Disease 67*, 842-847.

Meisel, B., Korsman, J., Kloppers, F.J. & Berger, D.K. (2009). *Cercospora zeina* is the causal agent of grey leaf spot disease of maize in southern Africa. *European Journal of Plant Pathology*, *124*, 577-583.

Möller, E.M., Bahnweg, G., Sandermann, H. & Geiger, H.H. (1992). A simple and efficient protocol for isolation of high molecular weight DNA from filamentous fungi, fruit bodies, and infected plant tissues. *Nucleic Acids Research*, *20*, 6115-6116.

Nicolaisen, M., Suproniene, S., Nielsen, L.K., Lazzaro, I., Spliid, N.H. & Justesen, A.F. (2009). Real-time PCR for quantification of eleven individual *Fusarium* species in cereals. *Journal of Microbiological Methods*, *76*, 234-240.

Okori, P., Fahleson, J., Rubaihayo, P. R., Adipala, E. & Dixelius, C. (2003). Assessment of genetic variation among East African *Cercospora zeae-maydis*. *African Crop Science Journal* 11, 75-85.

Poland, J.A., Balint-Kurti, P.J., Wisser, R.J., Pratt, R.C. & Nelson RJ (2009). Shades of gray: the world of quantitative disease resistance. *Trends in Plant Science*, *14*, 21-29.

Rafalski, J.A. (2010). Association genetics in crop improvement. *Current Opinion in Plant Biology*, *13*, 174-180.

Shim, W.B. & Dunkle, L.D. (2002). Identification of genes expressed during cercosporin biosynthesis in *Cercospora zeae-maydis*. *Physiological and Molecular Plant Pathology*, *61*, 237-248.

Shim, W.B. & Dunkle, L.D. (2005). *Malazy,* a degenerate, species-specific transposable element in *Cercospora zeae-maydis*. *Mycologia*, *97*, 349-355.

Tambong, J.T., Mwange, K.N., Bergeron, M., Ding, T., Mandy, F., Reid, L.M. & Zhu, X. (2008). Rapid detection and identification of the bacterium *Pantoea stewartii* in maize by TaqMan real-time PCR assay targeting the *cpsD* gene. *Journal of Applied Microbiology*, *104*, 1525-1537.

Van Den Brink, H.M., Van Zeijl, C.M.J., Brons, J.F., Van den Hondel, C.A.M.J.J., & Van Gorcom, R.F.M. (1995). Cloning and Characterization of the NADPH Cytochrome P450 Oxidoreductase Gene from the Filamentous Fungus *Aspergillus niger*. *DNA and Cell Biology, 14*, 719-729.

Waalwijk, C., Koch, S., Ncube, E., Allwood, J., Flett, B., de Vries, I. & Kema, G. (2008). Quantitative detection of *Fusarium* spp. and its correlation with fumonisin content in maize from South African subsistence farmers. *World Mycotoxin Journal, 1*, 39-47.

Wang, J., Levy, M. & Dunkle, L.D. (1998). Sibling species of *Cercospora* associated with gray leaf spot of maize. *Phytopathology*, 88, 1269-1275.

Ward, E., Foster, S.J., Fraaije, B.A. & McCartney, H.A. (2004). Plant pathogen diagnostics: immunological and nucleic acid-based approaches. *Annals of Applied Biology, 145*, 1-16.

Ward, J. M. J, Birch, E. B. & Nowell, D. C. (1994). Grey leaf spot on maize. Pietermaritzburg, South Africa. Coordinated extension, Maize in Natal, Cedara Agricultural Development Institute.

Ward, J.M.J, Stromberg, E.L., Nowell, D.C. & Nutter, Jr., F.W. (1999). Gray leaf spot. A disease of global importance in maize production. *Plant Disease*, *83*, 884-895.

White, T.J., Bruns, T., Lee, S. & Taylor, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: PCR Protocols: a guide to methods and applications (eds M.A.Innis *et al.*), pp. 315-322. Academic Press, California, USA.

Xu, M.L., Melchinger, A.E. & Lnbberstedt, T. (1999). Species-specific detection of the maize pathogens *Sporisorium reiliana* and *Ustilago maydis* by dot blot hybridization and PCR-based assays. *Plant Disease*, *83*, 390-395.

Xue, H.Q., Upchurch, R.G. & Kwanyuen, P. (2006). Ergosterol as a quantifiable biomass marker for *Diaporthe phaseolorum* and *Cercospora kikuchii*. *Plant Disease, 90*, 1395-1398.

Zhu, X., Reid, L.M. & Woldemariam, T. (2002). First report of gray leaf spot caused by *Cercospora zeae-maydis* on corn in Ontario, Canada. *Plant Disease*, *86*, 327.

Chapter 3

Comparative analysis of QTL derived from fungal biomass and whole plant disease phenotype in the GLS maize pathosystem

JN Korsman[1], P Tongoona[2], F Middleton[3], T Schmidt[1], M Carstens[1], BG Crampton[1], DK Berger[1]

[1]Department of Plant Science, Forestry and Agricultural Biotechnology Institute (FABI), Faculty of Natural and Agricultural Sciences, University of Pretoria, South Africa
[2]African Centre for Crop Improvement, University of KwaZulu-Natal, Pietermaritzburg, South Africa
[3]PANNAR SEED (Pty) Ltd, PO Box 19, Greytown, South Africa

The format of the European Journal of Plant Pathology was adhered to for this chapter.

I performed the gDNA extractions, SSR genotyping on agarose gels, genetic map construction, qPCR, QTL mapping of qPCR data, lesion area phenotype data and GLS disease score data, carried out all analysis unless otherwise stated and prepared the chapter. P Tongoona organised and implemented the Baynesfield field trial, provided the GLS scores from the field trial and implemented mixed models in SAS. F Middleton provided SSR genotyping expertise on polyacrylamide gels for certain SSR markers at the PANNAR reseach facilities. T Schmidt carried out the digital image analysis in Assess 2.7. M Carstens helped with the WinQTL Cartographer software. BG Crampton provided intellectual input. DK Berger provided intellectual input, funding and leaf photographs for digital image analysis.

Abstract

Grey leaf spot is a destructive foliar disease of maize caused by *Cercospora zeae-maydis* and *Cercospora zeina*. Studies mapping QTL for grey leaf spot disease resistance employ whole plant disease rating systems based on visual assessments which can be biased. We aimed to identify QTL involved in limiting *C. zeina* growth *in planta* using a qPCR assay and lesion area data, and to determine the overlap of these with previously identified GLS resistance QTL. Leaves of the CML444 × SC Malawi maize RIL population were photographed and harvested from a field trial in KwaZulu-Natal, South Africa. *C. zeina* was quantified in the leaf tissue using a qPCR assay and digital image analysis. The data obtained were used for QTL mapping. Three QTL were detected from the qPCR data on chromosomes one, five and six, and one QTL for lesion area in the same genomic region on chromosome six. Two of these QTL, located in bin 1.10 and bin 6.06/6.07, overlapped with GLS resistance QTL identified from whole plant disease scores of the same maize population. The bin 6.06/6.07 QTL appears to be novel since it does not correspond to any QTL from other studies. The bin 1.10 and bin 5.05/5.06 QTL were located in regions similar to GLS resistance QTL mapped in other populations. Interestingly, qPCR assay and lesion area data (sampled at 104 dap) had higher correlations with later disease ratings than with earlier disease ratings. This was borne out by the fact that QTL mapped from both qPCR and lesion area data coincided with QTL mapped from later disease ratings. This suggests that these scoring methods may be indicative of disease potential. The qPCR assay or digital image analysis of GLS lesions therefore have utility as alternative and less subjective methods than whole plant GLS disease scoring. These methods could be employed for identifying QTL that may elicit their effect at different time points in the disease cycle.

Keywords:

Cercospora, grey leaf spot, lesion area, maize, qPCR, QTL

Abbreviations:

dap: days after planting

GLS: grey leaf spot

lsm: least square means

QTL: quantitative trait locus/loci

RIL: recombinant inbred line

3.1 Introduction

Grey leaf spot (GLS) is an economically important foliar disease of maize caused by two related species *Cercospora zeae-maydis* and *Cercospora zeina* (Wang et al. 1998; Crous et al. 2006). Both species are present in the USA and Brazil (Wang et al. 1998; Dunkle and Levy 2000; Goodwin et al. 2001; Brunelli et al. 2008) while only C. *zeina* has been found in Africa and China (Meisel et al. 2009; Okori et al. 2003; Liu and Xu 2013).

GLS can have disastrous effects on maize crop yield when conditions are favourable for the disease. In South Africa yield losses of up to 60% have been recorded (Ward et al. 1997). Disease management practices for prevention of GLS are fungicide sprays, conventional tillage and breeding of resistant hybrids (Ward et al. 1999). The most practical and cost effective of these for subsistence farmers in Africa would be resistant hybrids. Quantitative resistance within resistant hybrids is more robust and durable and therefore more desirable than single genes conferring resistance (Lindhout 2002; Parlevliet 2002).

A number of studies looking for GLS disease resistance quantitative trait loci (QTL) have been carried out in the USA (Saghai-Maroof et al. 1996; Clements et al. 2000; Gordon et al. 2004; Balint-Kurti et al. 2008), in Southern Africa (Lehmensiek et al. 2001; Gordon et al. 2004; Berger et al. 2014), in Brazil (Juliatti et al. 2009; Pozar et al. 2009) and in China (Zhang et al. 2012; Xu et al. 2014). Most of these studies have employed a 1-9 or 1-5 scale for GLS disease rating, but a percent leaf affected area score has also been used (Clements et al. 2000; Gordon et al. 2004). The 1-9 scale usually takes into account the progression of the disease symptoms as they move up the plant (Munkvold et al. 2001) which follows the way in which the disease progresses. The conidia are wind dispersed either from old crop residue or nearby infected plants and only germinate under highly humid conditions (Ward et al. 1999). The upper leaf canopy ensures a high relative humidity for the initial infection by conidia which germinate on the lower leaves. The disease symptoms then spread from the bottom leaves upward.

The percent leaf affected area is related to the lesions on the plant at a given height. GLS lesions are rectangular, being defined by the major leaf veins. They are tan in colour becoming slightly grey when conidiation occurs (Latterell and Rossi 1983). A single lesion can be up to 60 mm long and 2-4 mm wide, depending on the leaf veins. Single lesions can coalesce forming larger tan or grey areas on the leaf. The disease has a latent period of 14 to 28 days

from infection to sporulation which occurs on mature lesions, with more resistant maize plants having latent periods closer to 28 days (Ward et al. 1999).

After entering the cells via the stomata, *C. zeina* conidia failed to establish infection in cells of a resistant maize hybrid more often than in a susceptible hybrid. In addition, the growth rate of the fungus was also slowed in a resistant hybrid in a microscopy study (Lyimo et al. 2013). A method of scoring plants for GLS, which is focussed more on fungal development and growth than on symptoms and the disease progress in plants, may identify loci more specifically targeted at the plant-fungal interaction rather than the more multifactorial visual estimation of disease progress.

Visual assessments of plant diseases have been shown to have biases (Sherwood et al. 1983), and can influence detection of some QTL and allele effects at QTL (Poland and Nelson 2011). De Coninck et al. (2012) suggested using lesion area assessment by digital image analysis and quantitative PCR (qPCR) to quantify pathogen biomass. Mutka and Bart (2015) review a number of quantitative image based techniques for high throughput phenotyping of the effect of pathogens on plants. These techniques include imaging in the visual, electromagnetic and thermal spectrum, as well as chlorophyll measurements.

Berger et al. (2014) mapped seven QTL for resistance to GLS, two of which were novel QTL, using a visual, whole plant scoring system. The study was carried out over three seasons and five locations within KwaZulu-Natal, South Africa. Four GLS resistance QTL were mapped from whole plant disease scores in the Baynesfield 2009 field trial where leaf material was sampled for further phenotyping.

The qPCR assay was available for quantifying *C. zeina in planta* (Korsman et al. 2012; Chapter 2). We thus aimed to apply this quantification assay, together with digital analysis of GLS lesion images, to identify QTL in the *Zea mays* genome involved in limiting the proliferation of the foliar pathogen *C. zeina*. In addition we wanted to determine whether these QTL overlap with GLS resistance QTL previously identified through visual disease assessment in the same population.

3.2 Materials and methods

3.2.1 Field trials and maize germplasm

A population of maize recombinant inbred lines (RIL, F7:S6) was used in this study. These RILs were produced at CIMMYT from a cross of the inbred lines CML444 from CIMMYT (Mexico) and SC Malawi from Zimbabwe (Messmer et al. 2009). The 145 RILs were planted in Baynesfield KZN (GPS coordinates: -29.763111, 30.33818), where natural infection of *C. zeina* occurs, in a randomized block design with three replicate rows of ten plants per row on 10 December 2008. The field was not sprayed with fungicide and the plants were inoculated with year-old, powdered *C. zeina* infected maize material placed in the whorls at the five to seven leaf stage (Berger et al. 2014).

3.2.2 GLS disease scores

The RILs in the field trial were scored for GLS disease on a scale of 1-9 (Munkvold et al. 2001), where 1 is resistant or no disease present, 2 is scarce lesions on lower leaves, 3 is the presence of lesions on the lower leaves with no lesions above the ear, 4 includes scarce lesions above the ear, 5 is the presence of symptoms on upper leaves and some lower leaves dead, 6 is many lesions above the ear, 7 is when the lower leaves are dead and there are many lesions on the upper leaves, 8 is the progression to coalesced lesions on the upper leaves and 9 is most leaves dead. Scores were determined from visual examination of rows of approximately 10 plants.

3.2.3 Determination of pathogen content in leaves

3.2.3.1 Molecular quantification of *C. zeina* by qPCR assay

Leaf samples for genomic DNA (gDNA) extraction were collected from 100 RILs in each of the three replicate blocks at 104 days after planting, in March 2009. These 100 RILs were selected from the population of 145 RILs using the "distant pair design" of Fu and Jansen (2006). This design selects pairs of individuals with the most dissimilar genotypes ensuring a selection of RILs representative of the population. Only 100 RILs were analysed due to cost constraints. For each replicate of each of the 100 RILs one leaf segment per plant was sampled from the first leaf above the ear for two individual representative plants in a row of ten plants. The leaf segments for photography and sampling were demarcated by placing a 120 mm long CD cover over the leaf and drawing lines with a black marker on either side of the CD cover, across the breadth of the leaf. Photographs of the leaf segments were taken for digital image analysis before sampling by cutting off the leaf segments along the marked lines. The two leaf

segments per row were frozen together in liquid nitrogen in the field, transported to the lab on dry ice and stored at -80ºC.

The leaf segments were ground with a mortar and pestle and gDNA from approximately 100 mg of the homogenized maize leaves was extracted using a CTAB method (Doyle and Dickson 1987; Saghai-Maroof et al. 1984). The CTAB buffer was modified with 2 % (w/v) polyvinylpyrrolidone prior to use. Extracted gDNA was dissolved in 1 × TE buffer (10 mM Tris-HCl, 1mM EDTA, pH 8.0). gDNA concentration was determined with a NanoDrop™ 1000 spectrophotometer (Thermo Scientific, Wilmington, USA) and aliquots diluted to a concentration of approximately 10 ng/µl in nuclease free water for qPCR.

Quantitative PCR amplifications were performed as described previously (Korsman et al. 2012; Chapter 2). Both *cpr1* and *gst3* fragments were amplified in three technical replicates for all samples of each of the three biological replicates (from the replicate blocks) to quantify the *C. zeina* gDNA and the maize gDNA present. The qPCR technical replicates were averaged to obtain values for each of the biological replicates which were analysed individually. Samples were analysed in 384 well pates on a LightCycler® 480 instrument (Roche Diagnostics). A separate standard curve was calculated for each plate. The same gDNA dilutions of known concentration were used to produce all standard curves. Regression equations were calculated from the graphs in LibreOffice version 4.0.2.2 and used to determine the amount of *C. zeina* and maize gDNA in each sample. The amount of *C. zeina* gDNA was normalised to the amount of maize gDNA, resulting in amount of fungus in a sample being expressed as ng *C. zeina* gDNA/mg maize gDNA. The ng *C. zeina* gDNA/mg maize gDNA measurements for the RIL population were transformed by $\log_{10}$ conversion, similar to the $\log_{10}$ conversion of aflatoxin data used to find QTL in maize (Paul et al. 2003).

3.2.3.2 Digital image analysis

Photographs of leaf sections in replicate block one taken in the field before sampling, at 104 days after planting, with a Nikon D90 digital camera, were analysed using ASSESS 2.7 software (L. Lamari, American Phytopathological Society, St. Paul, MN, USA) as previously described (De Coninck et al. 2012). The lesion area measurements were expressed as lesion area/leaf area. The average measurement of the two leaf segments was calculated for each RIL and transformed by $\log_{10}$ conversion as with the fungal qPCR assay.

3.2.4 Genetic map

The genetic map of the CML444 × SC Malawi RIL population (Messmer et al. 2009) was used in this study with some amendments. Gap regions of 20cM or greater were identified for potential addition of extra simple sequence repeat (SSR) markers. SSR markers in these regions were selected from the Maize Genetics and Genomics database (MaizeGDB) (Lawrence et al. 2005). Three leaves per RIL were pooled for DNA extraction, which was done using a CTAB method (Saghai-Maroof et al. 1984; Doyle and Dickson 1987). The selected SSR markers were tested on the parents and ten randomly selected RILs for polymorphism in the RIL population. A total of 145 RILs were genotyped with the new SSR markers, once non-germinating RILs, off types and heterozygous lines had been discarded. Genotyping was carried out using polyacrylamide or agarose gel electrophoresis. The genetic map data previously used (Messmer et al. 2009) was kindly supplied by J-M Ribaut (CIMMYT, Mexico). The additional SSR marker genotypes were added to the map using Kosambi's mapping function in MapManager QTX software (Manly et al. 2001). The markers added to the map were bnlg1811, bnlg615, umc1111, phi073, bnlg1449, bnlg1108, umc1720, bnlg105, dupssr10, umc1155, umc1572, bnlg2191, umc1413, umc1424, umc1562, umc1170, bnlg1375, umc1137, umc1337. The final genetic map was named QMap 2.0 and parts thereof were graphically displayed using MapChart (Voorrips 2002). Qmap 2.0 was used in a parallel study to map QTL for whole plant disease scores in the same population across eleven environments (Berger et al. 2014).

3.2.5 QTL Analysis

QTL analysis of the datatsets in this chapter were carried out as described in Berger et al. (2014). Specifically, QTL in the CML444 × SC Malawi RIL population were identified using QMap 2.0 and the Composite Interval Mapping (CIM) (Zeng 1994) function in QTL Cartographer 2.5_011 (Wang et al. 2012). The standard model, walk speed of 1 cM, window size of 10 cM, and both forward and backward regression analysis options were used. One thousand permutations with $\alpha = 0.05$ (95% confidence) were run to determine the LOD threshold at which to specify the presence of a QTL for each trait (Churchill and Doerge 1994). The 2-LOD intervals were used to define the QTL regions for QTL which had peaks above the determined LOD threshold.

Data used for QTL analysis of the qPCR assay of *C. zeina* content in leaves of each of the 100 sampled RILs across all three biological replicates were the least square means (lsm) of

log_{10}(ng *C. zeina* gDNA/mg maize gDNA), abbreviated as log_{10}(ngCz/mgZm), calculated by mixed model analysis. In addition, QTL analysis of the qPCR assay data of *C. zeina* content in leaves for each separate biological replicate block of the 100 sampled RILs was determined using log_{10}(ngCz/mgZm) values. Data used for QTL analysis of lesion area was log_{10}(lesion area/leaf area) for the 100 RILs photographed in biological replicate one. Data used for QTL analysis of whole plant GLS disease scores at each of the four separate rating time points (92, 99, 109 and 146 dap) were the average of the GLS disease scores of the three replicate blocks for each of the full set of 145 RILs at a particular time point. In the previous study (Berger et al. 2014), QTL for GLS disease scores from the full set of 145 RILs were mapped using the lsm value for each RIL calculated by mixed model analysis to account for variation between the three replicate blocks. GLS disease score data used to calculate the lsm value for each RIL were the three replicate block values, each calculated by taking the average scores from the four time points.

3.2.6 Statistical analysis

Statistical analyses including one way ANOVA with repeated measures, Tukey's multiple comparison test and Pearson correlations with an associated paired t-test were performed in the GraphPad Prism 5.04 software package (GraphPad Software Inc.). The lsm value of *C. zeina* content (log_{10}(ngCz/mgZm)) from the qPCR assay for each of the 100 RILs was determined from the data from the three replicate blocks by applying a mixed model approach using blocks as random effects and RIL genotypes as fixed effects using the PROC GLM procedure of SAS 9.3 (SAS Institute, Cary, NC).

3.3 Results

3.3.1 GLS disease assessment and quantitative phenotyping by qPCR assay

The field trial from which leaf samples for fungal quantification were collected showed a range of values for the average GLS disease symptom ratings from 1.5 to 8.3 with an average of 5.8 on the 1-9 rating scale. A range of GLS severity on field plants can be visualised in Figure 3.1. Other maize foliar diseases were not visibly present, possibly due to additional inoculum of *C. zeina* added at the five to seven leaf stage of plant growth ensuring that *C. zeina* had an advantage over other foliar pathogens. The trial was rated for GLS disease severity of all 145 RILs on four dates (92, 99, 109 and 146 days after planting) and leaf segments of 100 RILs were collected for molecular quantification on day 104 after planting (between the second and third ratings) from all three replicate blocks. These leaf segments were used for *C. zeina* quantification via the qPCR assay and digital photographs of leaves from replicate block one were used for image analysis (see below).

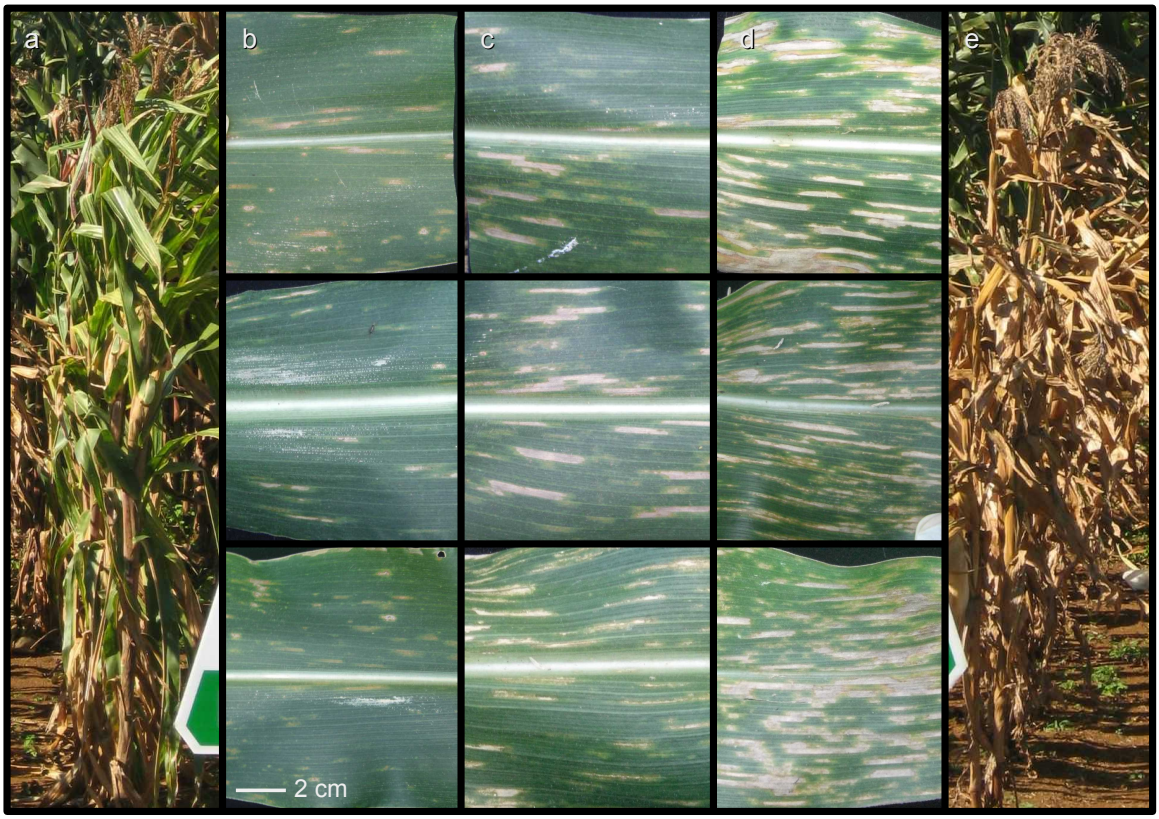


Figure 3.1: The range of GLS severity in the field. A row of healthy maize plants with few symptoms is pictured in (a) which would be scored as a 1 or a 2 on the 1-9 scale. A range of GLS lesion densities on individual maize leaves from mild (b), to intermediate (c), to severe (d). A row of severely infected maize plants (e) which would be scored as an 8 or a 9 on the 1-9 scale. Photos in (b)-(d) courtesy of Prof DK Berger.

3.3.2 Frequency distributions

For the GLS disease scores, the average score of the replicate blocks for the first rating event at 92 days after planting (dap) has normal distribution (Figure 3.2 a). The subsequent averages of the ratings for each replicated block (99, 109 and 146 dap) were increasingly skewed towards higher scores and thus started deviating from normality over time (Figure 3.2 b, c and d). The qPCR assay and lesion area scores appeared to have negatively exponential distributions. These were $\log_{10}$ transformed to normalise them (Figure 3.2 e and f, g and h). Quantitative traits controlled by numerous genes or loci, such as disease resistance, are often assumed to be normally distributed (Churchill and Doerge 1994; Doerge et al. 1997) and a number of methods for the detection of QTL rely on the assumption of normality. The graphs of the $\log_{10}$ transformed values are more symmetrical and do appear normally distributed.

Transgressive segregation can be expected with a quantitative trait (Tanksley 1993). This was visible in the data as a number of the RILs exhibited more extreme phenotypes than the parental lines for the GLS scores, the *C. zeina* quantification in the leaves, and the lesion area on the leaves (Figure 3.2 and Table 3.1).

Table 3.1: Transgressive segregation of RIL population

	min RIL score	CML444	SC Malawi	max RIL score
GLS score lsm	1.5	4.5	5.25	8.3
qPCR assay lsm $\log_{10}$(ngCz/mgZm)	0.013	2.510	11.006	79.263
lesion area $\log_{10}$(lesion/leaf)	0.005	0.060	0.148	0.258

3.3.3 Sources of variation in disease quantification methods

A one way ANOVA with repeated measures was performed on the qPCR assays of the three replicate blocks in GraphPad Prism. The differences between the replicate blocks of the field trial were found to be statistically significant (P value = 0.001) for the qPCR assay (Table 3.2). Tukey's multiple comparison test indicated that replicate block 1 and 3 were different (P value < 0.0005), but there was not significant support for differences between block 1 and 2, and block 2 and 3. Due to the variation of the qPCR assay data between the replicate blocks, a mixed model was applied to obtain the least square mean values for the qPCR assay data. Additionally, the qPCR assay data for each block was analysed individually.

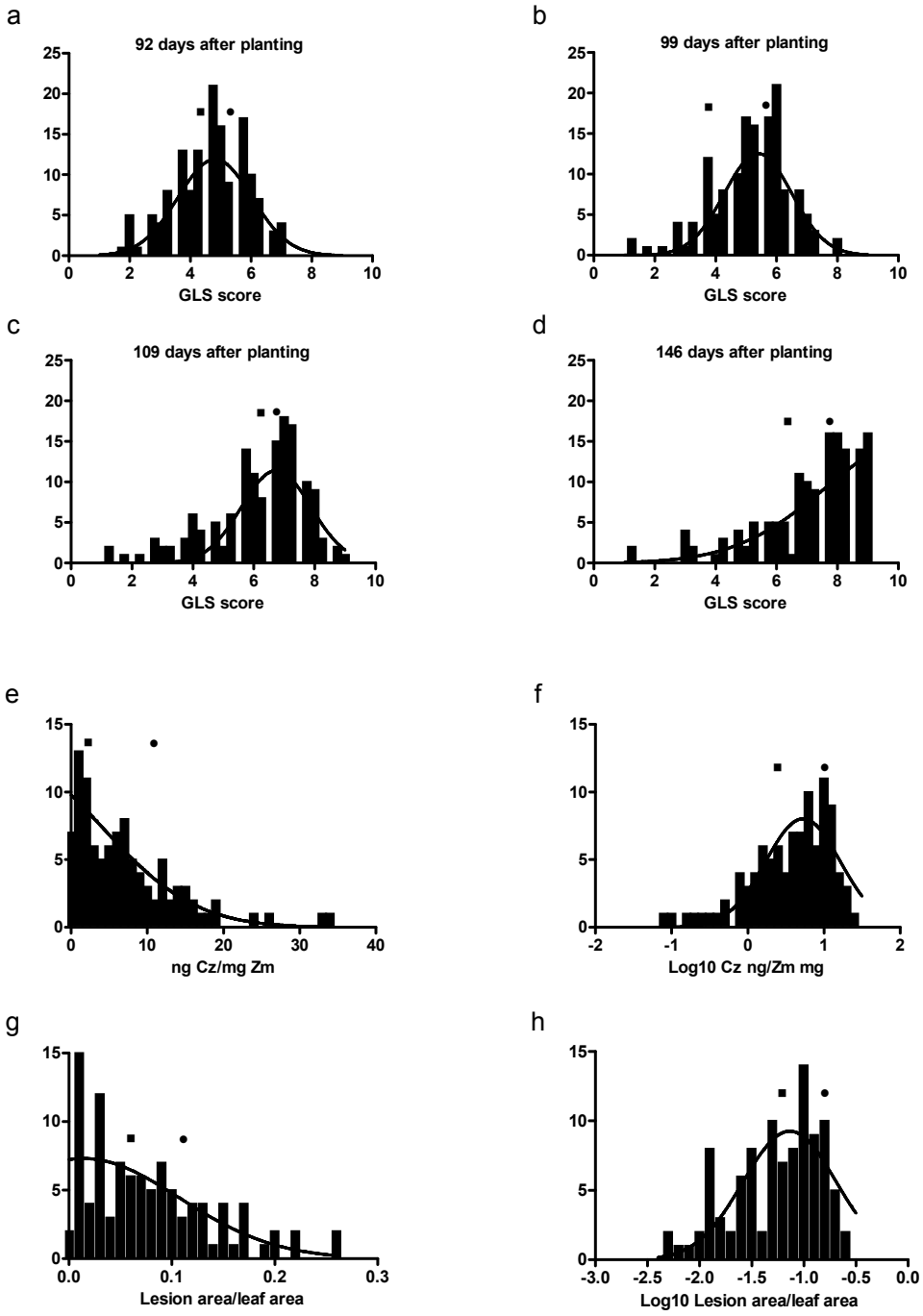


Figure 3.2: Frequency distributions of GLS ratings of rows of the 145 RILs from the Baynesfield 2009 field trial, the qPCR assay scores and the lesion area scores from the leaf sections harvested from 100 RILs of the Baynesfield 2009 field trial. The frequency distributions of the scores from 145 RILs at 92, 99, 109 and 146 days after planting are shown in graphs (a), (b), (c) and (d) respectively. The frequency distributions of the 104 days after planting qPCR assay scores on 100 RILs before $\log_{10}$ transformation are shown in graph (e), and those after $\log_{10}$ transformation in graph (f), as well as lesion area measurements before (g) and after (h) $\log_{10}$ transformation. Resistant and susceptible parental line scores are indicated by a square and a circle respectively. Normal distributions were fit to each of the graphs by a least squares method of nonlinear regression and are superimposed on the histograms.

Table 3.2: Analysis of variance of *C. zeina* quantification by the qPCR assay.

	Sum of squares	Degrees of freedom	Mean square	F value	P value
Replicate blocks	1.842	2	0.9211	7.211	P = 0.0010
Genotype	95.93	97	0.9889	7.742	P < 0.0001
Residual	24.78	194	0.1277		
Total	122.5	293			

3.3.4 Correlation of GLS disease assessment and *C. zeina* quantification values

The GLS disease scores, *C. zeina* quantification data from the qPCR assay and the lesion area data were compared to determine how well they correlate with each other (Table 3.3). All correlation values were highly significant (P value < 0.0001). The GLS score lsm data correlated well with the GLS scores for the replicate blocks (0.92-0.93). The GLS scores between replicate blocks also correlated well, the lowest correlation of 0.77 being between block one and two. The qPCR assay lsm data correlated well with the qPCR assay data from the individual replicate blocks (0.88-0.90). The GLS disease score lsm had a good correlation with the qPCR assay lsm (0.87). The lesion area scores for block one correlated well with the qPCR assay block one scores (0.80) as well as the GLS scores for block one (0.79). The GLS disease scores of the replicate blocks of the field trial also had a good positive correlation when compared to the corresponding block's qPCR assay scores (0.74, 0.82 and 0.85).

Table 3.3: Comparison of GLS disease data as separate replicates (blocks) for 100 RILs sampled. Pearson correlation coefficients between different types of GLS disease data are shown in the lower left, with corresponding P values in the upper right section. The comparisons of corresponding blocks are in bold.

	qPCR block 1	qPCR block 2	qPCR block 3	qPCR lsm	Lesion area	GLS block 1	GLS block 2	GLS block 3	GLS lsm
qPCR assay block 1[a]		1.06e-015	1.70e-016	1.21e-037	8.11e-024	1.40e-018	1.21e-012	3.12e-015	2.10e-018
qPCR assay block 2[a]	0.70		4.87e-014	1.94e-034	3.30e-016	2.12e-017	1.25e-028	3.94e-018	1.43e-025
qPCR assay block 3[a]	0.71	0.67		7.24e-034	5.02e-014	4.47e-016	4.93e-015	1.55e-025	1.67e-021
qPCR assay lsm[a]	0.90	0.89	0.88		7.41e-025	2.96e-024	2.86e-024	9.58e-027	6.92e-032
Lesion area block 1[b]	**0.80**	0.71	0.67	0.81		1.01e-022	2.67e-014	3.32e-015	1.13e-020
GLS score block 1[c]	**0.74**	0.73	0.70	0.81	**0.79**		2.30e-029	7.68e-036	0
GLS score block 2[c]	0.64	**0.85**	0.69	0.81	0.67	0.77		2.38e-036	0
GLS score block 3[c]	0.69	0.74	**0.82**	0.83	0.69	0.82	0.82		0
GLS score lsm	0.74	0.82	0.78	**0.87**	0.77	0.93	0.92	0.93	

[a] Measured in $\log_{10}$(ngCz/mgZm)
[b] Measured in $\log_{10}$(lesion area/leaf area)
[c] Average of four time points for replicate field block

Furthermore, the GLS disease score lsm and the GLS scores from the separate time points at which GLS was scored, the qPCR assay data and the lesion area data were compared to determine how well they correlate (Table 3.4). All correlation values were highly significant (P value < 0.0001). The GLS score lsm was compared to the four rating time points (92, 99, 109 and 146 dap), and they correlated well with correlation coefficients between 0.92 and 0.98. The four rating time points (92, 99, 109 and 146 dap) were compared to the qPCR assay and lesion area data (104 dap). The aim was to determine whether the closest time points (99 and 109 dap) for the ratings correlated better with the qPCR assay and lesion area data from 104 dap than those from the further time points (92 and 146 dap). The Pearson correlation coefficients between the GLS rating at 109 dap and the qPCR assay and lesion area (104 dap) were 0.84, and 0.75, respectively. However, qPCR assay and lesion area had better correlation with the 146 dap GLS rating with correlation coefficients of 0.88 and 0.76, respectively (Table 3.4).

Table 3.4: Comparison of time of disease assessment between different GLS disease data. Pearson correlation coefficients between GLS scores for different rating times, qPCR assay, lesion area and the lsm of the GLS scores are shown in the lower left, with corresponding P values in the upper right section.

	GLS 92 dap	GLS 99 dap	GLS 109 dap	GLS 146 dap	qPCR lsm 104 dap	Lesion area 104 dap	GLS score lsm
GLS rating 92 dap[a]		0	0	5.98e-030	1.65e-021	1.03e-014	0
GLS rating 99 dap[a]	0.88		0	0	2.19e-022	1.62e-016	0
GLS rating 109 dap[a]	0.88	0.94		0	3.41e-028	3.60e-019	0
GLS rating 146 dap[a]	0.77	0.83	0.90		1.22e-032	3.28e-020	0
qPCR assay lsm 104 dap[b]	0.78	0.79	0.84	0.88		7.41e-025	6.92e-032
Lesion area 104 dap[c]	0.68	0.71	0.75	0.76	0.81		1.13e-020
GLS score lsm	0.92	0.96	0.98	0.93	0.87	0.77	

[a] Average of three replicate field blocks
[b] Measured in $\log_{10}$(ngCz/mgZm)
[c] Measured in $\log_{10}$(lesion area/leaf area)

3.3.5 Genetic map

A genetic linkage map for the RIL population obtained from Messmer et al. (2009) was supplemented with an additional 19 SSR markers in regions of low marker density. A subset of the SSR markers is shown in Figure 3.3, displaying their polymorphic alleles. The map was reconstructed using the Kosambi mapping function in MapManager QTX. Markers that did not map to the expected region or distorted the map were discarded. Marker inversion is more

likely in small populations (Da Costa e Silva et al. 2007), thus markers which were closer than 5 cM to another marker were removed to reduce possible distortion of the map. The final map used here was designated QMap 2.0 and consisted of 167 markers based on data for 145 RILs with a total map size of 1862 cM (Figure 3.4).

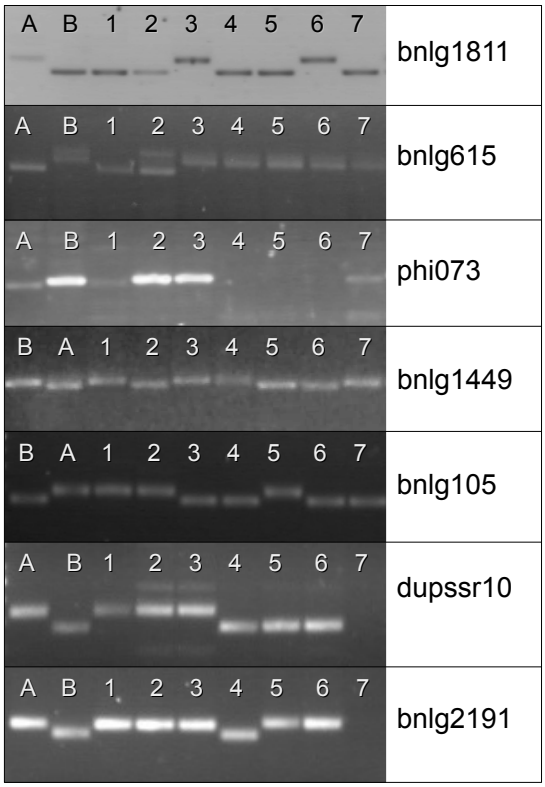


Figure 3.3: SSR markers, polymorphic in the RIL population, used to fill regions of low marker density on the original genetic linkage map for the RIL population. Marker names are listed on the right hand side of the agarose gel pictures. "A" represents the allele from CML444 (the more resistant parent), and "B", the allele from SC Malawi (the more susceptible parent), lanes 1 to 7 represent the alleles present in representative RILs at the relevant loci.

3.3.6 Identification of QTL for qPCR assay and lesion area

The QTL described here are named according to the chromosome they are situated on, then "Cz" if they were identified from the qPCR assay and "Lesion_area" for the lesion area QTL. Those identified from GLS scores at various time points are named according to the chromosome they are situated on and their time points, and those on chromosome 9 have a C or S at the end to distinguish between QTL with resistance alleles from CML444 or SC Malawi (Table 3.5).

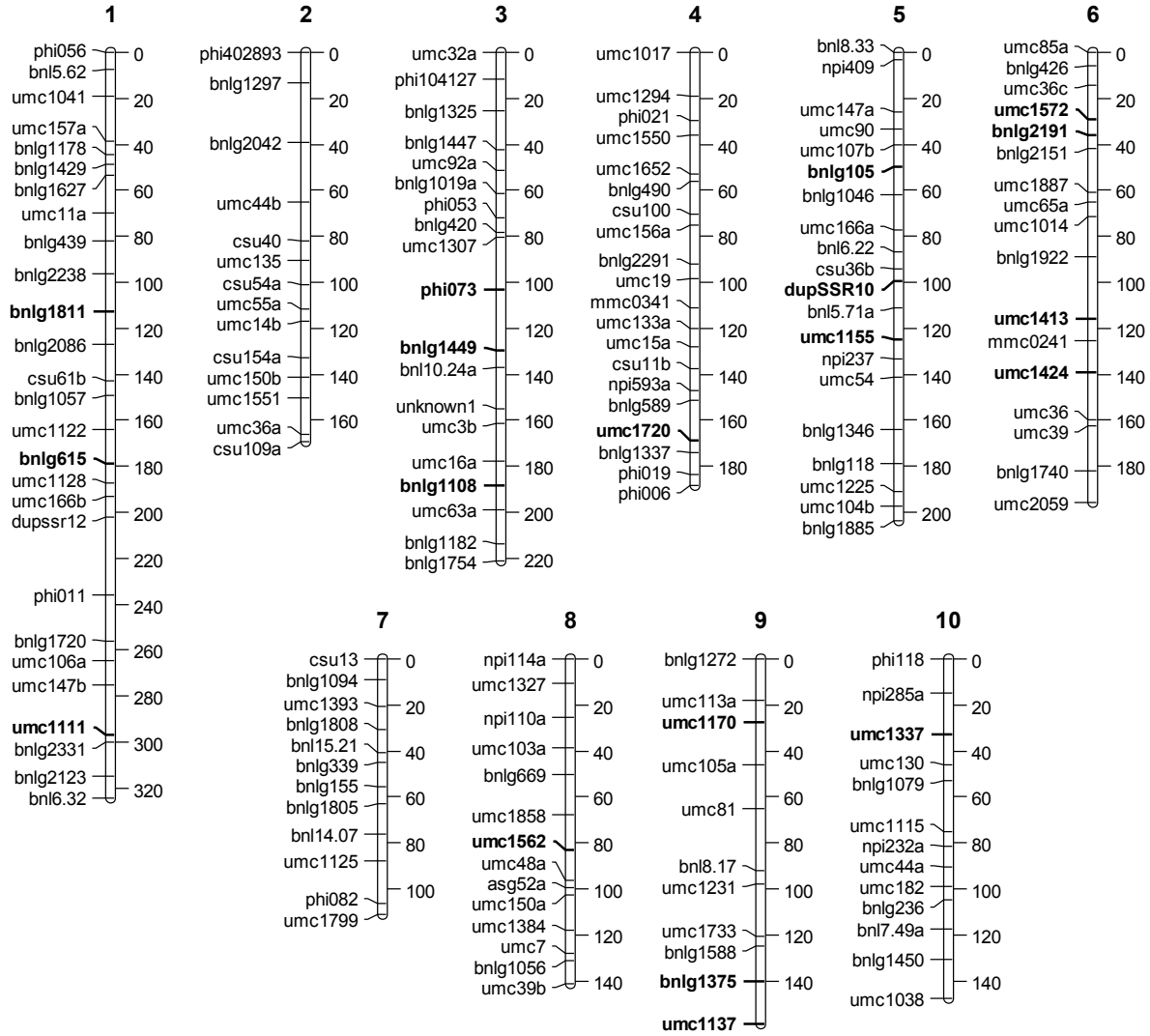


Figure 3.4: QMap 2.0, the genetic map based on the CML444 × SC Malawi RIL population. It consists of 167 markers based on data from 145 RILs with a total map size of 1862 cM. Markers added to the map of Messmer et al. (2009) are in bold.

Four GLS resistance QTL based on the lsm of the GLS scores of three biological replicate blocks, which were the average of four rating time points, were previously identified from the same population of 145 RILs in the same environment (Berger et al. 2014), namely 3H_GLS, 6H_GLS, 9H_GLS and 10H_GLS. Here we used the lsm of the scores from the qPCR assay performed on 100 RILs from the three replicate blocks, and the lesion area measurements of 100 RILs from replicate block one as traits for QTL mapping using the genetic map, Qmap 2.0 described above, as a framework.

Composite interval mapping in WinQTL Cartographer successfully detected three QTL for the qPCR assay lsm (least square means of three replicate blocks or biological replicates) on chromosomes 1, 5 and 6, named 1Cz, 5Cz and 6Cz, respectively. Additional QTL were detected using the qPCR assay data from the individual replicate field blocks, namely 3Cz_1.1, 3Cz_1.2, 6Cz_1, 6Cz_3, 9Cz_2.1, 9Cz_2.2 and 10Cz_2. QTL derived from the GLS data from the four separate rating time points were 3_92_dap, 3_99_dap, 6_109_dap, 6_145_dap, 9_92_dap_C, 9_92_dap_S, 9_99_dap_S, 9_109_dap_S and 10_146_dap.

Only 6Cz from the qPCR assay lsm QTL overlapped with a GLS resistance QTL, 6H_GLS (bin 6.06/6.07). The GLS resistance QTL 3H_GLS, 9H_GLS and 10H_GLS do not coincide with qPCR assay QTL, and 1Cz (bin 1.10) and 5Cz (bin 5.05/5.06) do not coincide with GLS resistance QTL from the same environment. The resistant allele source for the overlapping 6Cz and 6H_GLS are the same, the allele associated with lower pathogen content as measured by the qPCR assay and the allele associated with resistance is derived from SC Malawi (see Table 3.5). Only one QTL, 6Lesion_area, was detected using the lesion area phenotype. This QTL on chromosome 6 also overlaps with 6Cz and 6H_GLS, with SC Malawi as the source of the allele associated with smaller lesion area.

To determine whether the lack of overlap between disease resistance QTL and qPCR assay QTL was due to the time of sampling, the qPCR assay QTL were compared to QTL mapped from the averages of the individual timepoints (92, 99, 109, 146 dap) across the three replicate field blocks using 145 RILs. QTL 6_109_dap and 6_146_dap, the later timepoints, overlapped with 6Cz and 6H_GLS. QTL 9_92_dap_S, 9_99_dap_S and 9_109_dap_S overlap with 9H_GLS and 9Cz_2.2. The allele associated with resistance is from SC Malawi for all these QTL. The 9_92_dap_C QTL overlaps with 9Cz_2.1, with the resistance allele from CML444. In addition, these QTL are in the same position as QTL mapped from other field trials using GLS scores from the same RIL population (Berger et al., 2014). There is also overlap between the QTL on chromosome ten: 10_109_dap, 10_146_dap, 10Cz_2 and 10H_GLS all coincide. QTL 3_92_dap and 3_99_dap are found in the same region of chromosome three, but do not coincide with any other QTL observed in this RIL population. The overlap of the QTL can be visualized in Figure 3.5

Table 3.5: QTL identified from *C. zeina* qPCR assay and lesion area phenotypic data

QTL name[a]	Phenotype mapped[b]	Chr[c]	R^2 (%)[d]	LOD1 interval[e]	LOD2 interval[f]	LOD Score[g]	Peak marker[h]	Peak bin[i]	Additive effect[j]	R allele source[k]
1Cz	qPCR assay	1	11.0	265.2 - 282.5	264.4 - 287.2	3.15	umc147b	1.10	-0.198	CML444
3H_GLS	GLS score	3	8.5	29.6 - 46.2	25.3 - 50.1	2.95	bnlg1447	3.02/3.03	-0.409	CML444
3_92_dap	GLS score 92 dap	3	12,1	171.1 - 187.8	163.9 - 188.2	3,53	umc16a	3.07	-0,429	CML444
3_99_dap	GLS score 99 dap	3	13,2	170 - 186	165.3 - 187.6	4,27	umc16a	3.07	-0,478	CML444
3Cz_1.1	qPCR assay block 1	3	13.5	199.3 - 213.6	190.6 - 213.6	3.13	umc63a	3.09	-0.245	CML444
3Cz_1.2	qPCR assay block 1	3	13.0	213.6 - 219.6	213.6 - 219.6	3.31	bnlg1182	3.09	-0.240	CML444
5Cz	qPCR assay	5	12.7	121.7 - 136.9	118.8 - 146.2	3.32	umc1155	5.05	-0.235	CML444
6Cz	qPCR assay	6	19.0	134.4 - 154.9	129 - 159.5	3.01	umc1424	6.06	0.260	SC Malawi
6Cz_1	qPCR assay block 1	6	22.0	140.6 - 156.8	132 - 159.4	3.51	umc1424	6.06	0.318	SC Malawi
6Cz_3	qPCR assay block 3	6	19.3	135.9 - 154.3	131.3 - 159.4	3.33	umc1424	6.06	0.281	SC Malawi
6Lesion_area	lesion area	6	22.7	140.8 - 156.8	131.6 - 161.4	3.07	umc1424	6.06	0.210	SC Malawi
6H_GLS	GLS score	6	13.0	142.7 - 160.5	139.2 - 164.3	3.29	umc36	6.06/6.07	0.504	SC Malawi
6_109_dap	GLS score 109 dap	6	17,6	143.9 - 160	140.3 - 162.1	4,26	umc1424	6,06	0,684	SC Malawi
6_146_dap	GLS score 146 dap	6	25,6	146.3 - 160	143.8 - 161.8	5,19	umc1424	6,06	0,869	SC Malawi
9H_GLS	GLS score	9	7.6	115.0 - 122.8	110.7 - 124.5	3.40	umc1733	9.06	0.398	SC Malawi
9Cz_2.1	qPCR assay block 2	9	13.2	87.2 - 101	80.4 - 105.8	3.74	bnl8.17	9.04	-0.254	CML444
9_92_dap_C	GLS score 92 dap	9	9,5	78.4 - 97.3	71.7 - 97.6	3,20	bnl8.17	9.04	-0,401	CML444

QTL name[a]	Phenotype mapped[b]	Chr[c]	R^2 (%)[d]	LOD1 interval[e]	LOD2 interval[f]	LOD Score[g]	Peak marker[h]	Peak bin[i]	Additive effect[j]	R allele source[k]
9Cz_2.2	qPCR assay block 2	9	10.4	119.8 - 123.9	118.5 - 128.9	3.27	umc1733	9.06	0.233	SC Malawi
9_92_dap_S	GLS score 92 dap	9	17,6	119.8 - 122.9	119.8 - 123.9	6,50	umc1733	9.06	0,573	SC Malawi
9_99_dap_S	GLS score 99 dap	9	10,3	115.9 - 122.2	112.5 - 123.5	4,40	umc1733	9.06	0,436	SC Malawi
9_109_dap_S	GLS score 109 dap	9	9,7	115 - 122.7	111.4 - 124.3	3,99	umc1733	9.06	0,526	SC Malawi
10H_GLS	GLS score	10	13.5	114.0 - 127.9	110.9 - 129.8	4.53	bnl7.49a	10.06/10.07	-0.509	CML444
10Cz_2	qPCR assay block 2	10	10.3	112.6 - 123.5	107.7 - 127.7	3.15	bnl7.49a	10.06/10.07	-0.215	CML444
10_109_dap	GLS score 109 dap	10	9.4	113.8 - 126.8	110.6 - 129.7	3.47	bnl7.49a	10.06/10.07	-0.500	CML444
10_146_dap	GLS score 146 dap	10	17,1	115.3 - 128.7	111.4 - 129.8	4,75	bnl7.49a	10.06/10.07	-0,695	CML444

[a]QTL name: First number indicates maize chromosome. Cz indicates QTL is derived from qPCR assay, Lesion_area indicates QTL is derived from lesion area, 92, 99, 109 or 146 dap indicates QTL was derived from rating time point, an "H_GLS" behind the chromosome number indicates QTL derived from GLS score lsm described in Berger et al. (2014). The qPCR assay QTL derived from an individual replicate field block have an additional "_1", "_2", or "_3" to indicate the block, when there are two of these QTL on one chromosome they are indicated by a ".1" or ".2" at the end. To discriminate between multiple QTL on chromosome 9 for the GLS time points, C or S has been added to represent the source of the resistance allele.
[b]The type of data used to map the QTL
[c]maize chromosome
[d]Phenotypic variance explained by the QTL
[e]The cM range defining the 1-LOD interval
[f]The cM range defining the 2-LOD interval
[g]Log of odds (LOD) value at QTL peak
[h]QMap 2.0 marker closest to QTL peak
[i]Chromosome bin location of QTL (1-LOD interval) based on markers present on the IBM2005 neighbours frame map
[j]Additive effect of the QTL based on the 1-9 scale, or the ng of *C. zeina* DNA/mg *Z. mays* DNA, or the lesion area/leaf area.
[k]Resistance allele source: The allele associated with resistance or lower fungal biomass or fewer lesions from CML444 or SC Malawi

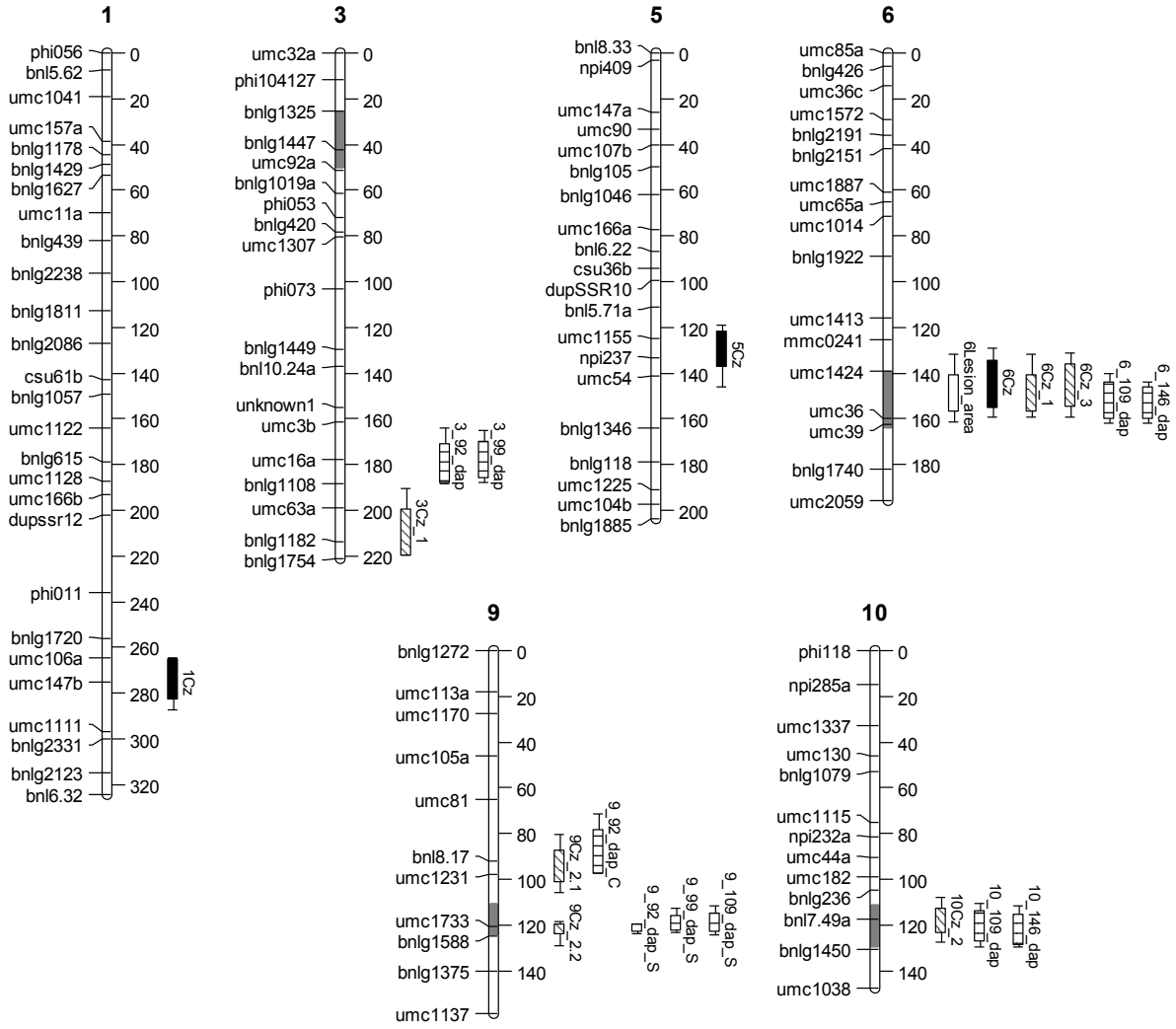


Figure 3.5: Partial genetic linkage map of the CML444 × SC Malawi RIL population indicating QTL. The chromosomes from QMap 2.0 (Figure 3.4) on which QTL were identified are shown. Marker names are indicated on the left of the chromosome and cM length in indicated on the right. QTL locations are shown by bars on the right of the cM intervals. The QTL for GLS resistance from the Baynesfield 2009 environment (Berger et al., 2014) are indicated as shaded areas on the chromosome. The QTL identified from lesion area is indicated by the white bar. The black bars show the QTL identified from the least square means of the qPCR assay of the replicate blocks and the diagonally lined bars show the QTL from the qPCR of the individual replicate blocks. The horizontal lined bars show the QTL from the four field ratings (at 92, 99, 109 and 146 dap). The 1-LOD and 2-LOD intervals around the QTL peaks are depicted by the bars and the lines respectively. Note that 3Cz_1 QTL region at the bottom of chromosome 3 is in fact two QTL with intervals that meet.

3.4 Discussion

The aims of this study were to find QTL in the maize genome, which may be involved in limiting *C. zeina* growth and to determine the overlap of these QTL with previously identified GLS resistance QTL from the same population (Berger et al. 2014). We used the *C. zeina* quantification qPCR assay and lesion area measurements as phenotypes for mapping these QTL. Another aim was to determine whether these assays could be used as a less biased approach than whole plant disease scores for the identification of GLS resistance QTL.

The Baynesfield 2009 field trial, from which the data used here was collected, exhibited a range of symptoms in a roughly normal distribution with transgressive segregation beyond the parental values (Figure 3.2). This is consistent with GLS resistance being a quantitative trait in this population. The deviations from normality may be due to the population size. As the number of individuals increases, the distribution of the phenotype is likely to become increasingly normal. The disease progression and timing of the GLS disease ratings and sampling for the qPCR assay may have had an effect on the frequency distribution. The distribution will progress from being skewed towards low scores early in the infection to being skewed towards high scores at later time points (Figure 3.2 a-d). If the 1-9 GLS disease rating system used was able to continue past 9 (which is an almost dead plant and thus further scoring is not possible) on the scale, the frequency distribution may continue to resemble a normal distribution with the whole curve shifting to the right. Thus detection limits may prevent the phenotypic data from having distributions that conform to normality.

Although the RIL population does not have normal frequency distributions for all the traits investigated here (GLS disease score, *C. zeina* biomass as quantified by the qPCR assay and lesion area), they do all exhibit transgressive segregation (Table 3.1) as expected (Tanksley 1993). Certain RILs will have inherited loci from both parents that contribute to resistance and will thus be more resistant than either parent, the converse also applies to susceptibility. Interaction between loci (epistasis) may also contribute to more extreme phenotypes as an interaction between two loci, one of which may not be present in a parent, may lead to greater or lesser resistance.

Looking at all the phenotypes measured for the different traits, i.e. the GLS scores, the qPCR assay and the lesion area of the RIL population, the correlation between the traits is strong. Correlation between GLS score lsm and the GLS replicate blocks is very good as expected,

since the lsm is derived from the GLS scores at different times for the replicate blocks. The qPCR assay lsm and the qPCR assay scores from the replicate blocks also correlate very well. The qPCR assay lsm and GLS score lsm correlate very well, as the qPCR measures the fungal biomass which plays a role in the severity and progression of the disease, with higher levels of fungal biomass resulting in greater symptom expression. Although bias may be incorporated into the GLS scoring system, it will be small enough not to result in drastic inconsistencies as experienced plant pathologists carried out scoring on replicate blocks at multiple times. This imparts some confidence in the hypothesis that the same QTL may be detected using the different methods of phenotyping. The lesion area correlates well with both the qPCR assay scores and the GLS scores from the same replicate block. Overall, correlation coefficients between GLS score, qPCR assay and lesion area were all above 0.7 for the same replicate blocks from the field (Table 3.3). If one looks at the correlations between the GLS score lsm and the GLS scores at different rating time points (92, 99, 109 and 146 dap), all of them correlate well (Table 3.4). Looking at the GLS scores at the different time points and the qPCR assay and lesion area taken at 104 dap (between 99 and 109 dap), the qPCR assay and lesion area correlate best with the 146 dap time point. The normalised frequency distributions (Figure 3.2 f and h) of the qPCR assay and lesion areas are skewed slightly to the right as with the later time points of the GLS scores. A plant with a relatively small fungal load at 104 dap will not show the extent of damage at 146 dap as a plant with a large fungal load at 104 dap. Thus the qPCR assay and lesion areas may pre-empt the GLS scores which are based on impressions of whole rows of plants, and they may be a good representation of disease potential.

Resistance is often assumed when plants do not show symptoms to a pathogen. This is true when the level of the pathogen is low and the level of symptoms is low. An alternative is tolerance, where the pathogen load may be high, but there are few symptoms or little reduction in fitness. However, the qPCR assay scores and disease scores do correlate indicating that there is not a large amount of pathogen in a plant with few symptoms, thus there is no evidence of disease tolerance as opposed to resistance (Van der Linden et al. 2013) in the population studied. We have not evaluated yield, which may indicate another form of tolerance if a plant with a high disease score can still produce a high yield which is the ultimate fitness for a crop (Ward et al. 1999).

Given that the different ways of scoring the disease or the pathogen have strong positive correlations, we hypothesised that similar QTL could be mapped from the qPCR analysis and lesion area scores as the whole plant disease scores. Thus the QTL controlling the "pathogen content" traits were investigated. Four QTL for GLS disease resistance in this population in the same environment were previously reported along with a number of other QTL found in different environments (Berger et al. 2014). Here, three QTL were detected for the qPCR assay (1Cz, 5Cz and 6Cz), and one QTL for lesion area (6Lesion_area).

Fungal biomass measured by qPCR assays has previously been found to be associated with QTL and to correlate well with other disease scoring methods (Oliver et al. 2008; Chung et al. 2010). A qPCR assay for quantifying fungal biomass of *Exserohilum turcicum* on greenhouse-grown maize has been described and used to determine the effect of the presence or absence of a QTL on the amount of fungal biomass (Chung et al. 2010). The absence of a resistance QTL for Northern Corn Leaf Blight (caused by *E. turcicum*) was associated with increased fungal biomass as measured by qPCR at 9 days post infection in a glasshouse trial. The levels of fungal DNA also had an inverse correlation to resistance levels of plants in the field (Chung et al. 2010). A study of the correlation of fungal biomass measured by qPCR, visual disease analysis and yield as methods of assessment of quantitative disease resistance of wheat to its pathogen *Stagonospora nodorum* showed high correlations between all methods (Oliver et al. 2008). No other studies have been found that use the fungal biomass determined by qPCR assays to map QTL. This study therefore serves as a proof of concept that quantification of fungal biomass by qPCR assay can be used to map QTL involved in resistance to fungal pathogens.

The lesion area method of scoring GLS was successfully used to identify one QTL, namely 6Lesion_area. Digital image analysis of a downy mildew of grapevine, which quantified white sporulating area on leaves, has also been used to map QTL successfully. The QTL detected by the image analysis coincided with QTL from spore measurement using a cell counter (Peressotti et al. 2011). In an analysis using both scanned leaves and disease scores for GLS, the scanned leaves did not yield any QTL that corresponded to the QTL mapped from the visual scores (Zhang et al. 2012). The authors did not report the nature of the frequency distribution of the data and whether it was transformed to conform to a normal distribution as was necessary here. Normally distributed data may be important for detecting reliable QTL.

Sherwood et al. (1983) described a bias in visually estimated percent leaf area of diseased orchardgrass leaves, specifically an overestimation of smaller infected areas of leaves. In a more recent study on maize, visual assessments by different raters did not correlate well but still produced the same QTL (Poland and Nelson 2011). Scoring diseased leaf area by means of digital image analysis or molecular quantification by qPCR assay of pathogen content in leaves overcomes the bias in visual estimations and the variability of different raters. We have shown that these methods can be used to identify QTL which co-localise with QTL observed with other methods of GLS scoring.

The 1-9 scale for scoring GLS resistance which is based on infected foliar area and progression up the plant (Munkvold et al. 2001) is the most common method used to rate the disease. It is a relatively fast and cheap method (if man-hours and experience are not taken into account), and is unlikely to be replaced by a molecular analysis of pathogen content of leaves in the near future as DNA extraction and qPCR are time consuming and expensive if not automated. In the case of digital image analysis, the time taken to capture images for an entire field is far greater than the time taken to rate the same field. The image quality of photos taken non-destructively in the field can be variable due to light conditions and taking photos in the field is not an easy task in the cramped spaces between plants. The choice of one or two representative leaves from a row may also result in a form of bias. Taking leaves back to the lab to be photographed or scanned under controlled conditions may increase image quality, but it is a destructive process. Image assessment does not distinguish between different disease symptoms, or other leaf blemishes. The qPCR assay will detect only *C. zeina* in leaf samples and once qPCR assays for other foliar diseases are available it can be multiplexed for detection of a range of diseases. It does however require a destructive sampling method.

A number of GLS disease resistance QTL have been identified over numerous maize populations in different environments across continents. A meta analysis of these has been conducted (Berger et al. 2014). The QTL identified from the qPCR assay lsm representing the fungal biomass within the leaves are located on chromosomes 1, 5 and 6. QTL 1Cz in bin 1.10 occurred in the same region as a GLS resistance QTL observed in the same RIL population, but identified from a different field trial or environment, namely Redgates 2010 and Hildesheim 2009 (Berger et al. 2014). It also co-localised with another GLS resistance QTL identified previously in a different maize population but reported in the same bin (bin 1.10) (Zhang et al. 2012).

While 5Cz, the QTL on chromosome 5 in bin 5.05-5.06, was not identified from GLS scores in the same population, it does correspond to QTL from three other populations. It co-localises with GLS resistance QTL that was previously detected between mmc0282 (bin 5.05) and bnlg1847 (bin 5.06) (Lehmensiek et al. 2001) and between bnl5.71 (bin 5.05) and umc54 (bin 5.06) (Saghai-Maroof et al. 1996) and between asg71 (bin 5.05) and csu440 (bin 5.06) (Clements et al. 2000). These studies were carried out in different geographical regions encompassing the probability of both *C. zeina* and *C. zeae-maydis* infections.

The 6Cz QTL in bin 6.06 on chromosome 6 was also identified using the lesion area phenotype as 6Lesion_area. This QTL region was the only one identified by all three kinds of disease or pathogen scoring systems, namely the qPCR assay for fungal biomass, the lesion area and the GLS disease score from the same field trial. It is also a strong effect QTL (R^2 is high, 6Cz R^2 = 19.0%; 6Lesion_area R^2 = 22.7%; 6H_GLS R^2 = 13.0%, see Table 3.5). A QTL for early disease development in a neighbouring chromosomal region with flanking markers npi373 (bin 6.02) and umc46 (bin 6.05), was previously identified (Clements et al. 2000) in a study possibly based on *C. zeae-maydis* infection. The markers flanking this QTL exhibit a similar irregularity in that they that should be far apart according to their bin positions (bin 6.02 and 6.05), yet they map closely, as is the case with umc1424 (bin 6.01) and umc36 (bin 6.06) on Qmap 2.0. The bin location of the left hand flanking markers as given at MaizeGDB, is far from the position of the QTL on the maps. It is thus not possible to determine whether the QTL from the two studies overlap.

The mixed model approach used to determine the lsm values for the qPCR assay accounts for the random effects of the replicate blocks, thus the QTL identified from the qPCR lsm values should be the most reliable. One would think that the qPCR assay data from the separate replicate blocks in the field would overlap with the QTL mapped from the lsm values. Surprisingly, the QTL identified from the qPCR assay data from the three individual replicate blocks do not co-localise with 1Cz and 5Cz. QTL 6Cz is corroborated by 6Cz_1 and 6Cz_3. The other individual replicate block qPCR assay QTL do correspond to GLS resistance QTL identified in the same RIL population, some in the same, and some in different environments (Berger et al. 2014). The QTL 3Cz_1.1 and 3Cz_1.2 correspond to 3F_GLS from the Ukalinga 2009 field trial, 9Cz_2.1 corresponds to 9A1_GLS and 9F1_GLS from the Redgates 2008 field trial and 9Cz_2.2 and 10Cz_2 correspond to 9H_GLS and 10H_GLS from the same Baynesfield field trial.

The overlap of 6Cz with QTL mapped from the later individual rating time point GLS scores (6_109_dap and 6_146_dap) indicate that the qPCR assay results may be indicative of disease potential. The QTL 10Cz_2 identified from qPCR assay results of replicate block two, also overlaps with later time point QTL, namely 10_109_dap and 10_146_dap. The correlation of the qPCR assay and lesion area data with the 146 dap time point mentioned earlier, also suggest that these scoring methods may be indicative of disease potential.

In conclusion, the GLS resistance QTL found in this study were also identified in other studies or in similar chromosomal regions to QTL present in other studies. They are not novel QTL, but were detected using data from a qPCR assay which measures fungal biomass within *C. zeina* infected leaves. This can be considered a novel scoring method for *C. zeina*. It is also a method for scoring the *C. zeina* pathogen, rather than GLS disease (the pathogen is only one contributing factor to the disease phenotype, another contributing factor being the environment). The qPCR assay and lesion area quantification methods may be useful tools for QTL identification as their associated QTL coincide with the traditional GLS disease resistance QTL. It would be interesting to automate the system and expand it to include multiple environments to possibly investigate QTL involved in the course of pathogen accumulation in leaves over time.

3.5 References

Balint-Kurti, P. J., Wisser, R., & Zwonitzer, J. C. (2008). Use of an advanced intercross line population for precise mapping of quantitative trait loci for gray leaf spot resistance in maize. *Crop Science*, *48*, 1696–1704.

Berger, D. K., Carstens, M., Korsman, J. N., Middleton, F., Kloppers, F. J., Tongoona, P., & Myburg, A. A. (2014). Mapping QTL conferring resistance in maize to gray leaf spot disease caused by *Cercospora zeina*. *BMC Genetics*, *15*, 60.

Brunelli, K. R., Dunkle, L. D., Sobrinho, C. A., Fazza, A. C., & Camargo, L. E. A. (2008). Molecular variability in the maize grey leaf spot pathogens in Brazil. *Genetics and Molecular Biology*, *942*, 938–942.

Chung, C.-L., Longfellow, J. M., Walsh, E. K., Kerdieh, Z., Van Esbroeck, G., Balint-Kurti, P., & Nelson, R. J. (2010). Resistance loci affecting distinct stages of fungal pathogenesis: use of introgression lines for QTL mapping and characterization in the maize--*Setosphaeria turcica* pathosystem. *BMC Plant Biology*, *10*, 103.

Churchill, G. A., & Doerge, R. W. (1994). Empirical threshold values for quantitative trait mapping. *Genetics*, *138*, 963–971.

Clements, M. J., Dudley, J. W., & White, D. G. (2000). Quantitative trait Loci associated with resistance to gray leaf spot of corn. *Phytopathology*, *90*, 1018–25.

Crous, P. W., Groenewald, J. Z., Groenewald, M., Caldwell, P., Braun, U., & Harrington, T. C. (2006). Species of *Cercospora* associated with grey leaf spot of maize. *Studies in Mycology*, *55*, 189–197.

Da Costa e Silva, L., Cruz, C. D., Moreira, M. A., & de Barros, E. G. (2007). Simulation of population size and genome saturation level for genetic mapping of recombinant inbred lines (RILs). *Genetics and Molecular Biology*, *30*, 1101–1108.

De Coninck, B.M.A., Amand, O., Delaur, S.L., Lucas, S., Hias, N., Weyens, G., Mathys, J., De Bruyne, E. & Cammue, B.P.A. (2012). The use of digital image analysis and real-time PCR fine-tunes bioassays for quantification of *Cercospora* leaf spot disease in sugar beet breeding. *Plant Pathology*, 61, 76–84.

Doerge, R. W., Zeng, Z., & Weir, B. S. (1997). Statistical issues in the search for genes affecting quantitative traits in experimental populations. *Statistical Science*, *12*, 195–219.

Doyle, J. J., & Dickson, E. E. (1987). Preservation of plant samples for DNA restriction endonuclease analysis. *Taxon*, *36*, 715–722.

Dunkle, L. D., & Levy, M. (2000). Genetic relatedness of African and United States populations of *Cercospora zeae-maydis*. *Phytopathology*, *90*, 486–490.

Fu, J., & Jansen, R. C. (2006). Optimal design and analysis of genetic studies on gene expression. *Genetics*, *172*, 1993–1999.

Goodwin, S. B., Dunkle, L. D., & Zismann, V. L. (2001). Phylogenetic analysis of *Cercospora* and *Mycosphaerella* based on the internal transcribed spacer region of ribosomal DNA. *Phytopathology*, *91*, 648–658.

Gordon, S. G., Bartsch, M., Matthies, I., Gevers, H. O., Lipps, P. E., & Pratt, R. C. (2004). Linkage of molecular markers to *Cercospora zeae-maydis* resistance in maize. *Crop Science*, *44*, 628–636.

Juliatti, F. C., Pedrosa, M. G., Silva, H. D., & Silva, J. V. C. (2009). Genetic mapping for resistance to gray leaf spot in maize. *Euphytica*, *169*, 227–238.

Korsman, J., Meisel, B., Kloppers, F., Crampton, B., & Berger, D. (2012). Quantitative phenotyping of grey leaf spot disease in maize using real-time PCR. *European Journal of Plant Pathology*, *133*, 461–471.

Latterell, F., & Rossi, A. (1983). Gray leaf spot of corn: a disease on the move. *Plant Disease*, *67*, 842–847.

Lawrence, C. J., Seigfried, T. E., & Brendel, V. (2005). The maize genetics and genomics database. The community resource for access to diverse maize data. *Plant Physiology*, *138*, 55–58.

Lehmensiek, A., Esterhuizen, A., van Staden, D., Nelson, S., & Retief, A. (2001). Genetic mapping of gray leaf spot (GLS) resistance genes in maize. *Theoretical and Applied Genetics*, *103*, 797–803.

Lindhout, P. (2002). The perspectives of polygenic resistance in breeding for durable disease. *Euphytica*, *124*, 217–226.

Liu, K.-J., & Xu, X.-D. (2013). First report of gray leaf spot of maize caused by *Cercospora zeina* in China. *Plant Disease*, *97*, 1656.

Lyimo, H. J. F., Pratt, R. C., & Mnyuku, R. S. O. W. (2013). Infection process in resistant and susceptible maize (*Zea mays* L.) genotypes to *Cercospora zeae-maydis* (Type II). *Plant Protection Science*, *49*, 11–18.

Manly, K. F., Cudmore, Jr., R. H., & Meer, J. M. (2001). Map Manager QTX, cross-platform software for genetic mapping. *Mammalian Genome*, *12*, 930–932.

Meisel, B., Korsman, J., Kloppers, F. J., & Berger, D. K. (2009). *Cercospora zeina* is the causal agent of grey leaf spot disease of maize in Southern Africa. *European Journal of Plant Pathology*, *124*, 577–583.

Messmer, R., Fracheboud, Y., Bänziger, M., Vargas, M., Stamp, P., & Ribaut, J.-M. (2009). Drought stress and tropical maize: QTL-by-environment interactions and stability of QTLs across environments for yield components and secondary traits. *Theoretical and Applied Genetics*, *119*, 913–30.

Munkvold, G. P., Martinson, C. A., Shriver, J. M., & Dixon, P. M. (2001). Probabilities for profitable fungicide use against gray leaf spot in hybrid maize. *Phytopathology*, *91*, 477–484.

Mutka, A. M., & Bart, R. S. (2015). Image-based phenotyping of plant disease symptoms. *Frontiers in Plant Science*, *5*, 1–9.

Okori, P., Fahleson, J., Rubaihayo, P., Adipala, E., & Dixelius, C. (2003). Assessment of genetic variation among East African *Cercospora zeae-maydis*. *African Crop Science Journal*, *11*, 75–85.

Oliver, R. P., Rybak, K., Shankar, M., Loughman, R., Harry, N., & Solomon, P. S. (2008). Quantitative disease resistance assessment by real-time PCR using the *Stagonospora nodorum*-wheat pathosystem as a model. *Plant Pathology*, *57*, 527–532.

Parlevliet, J. E. (2002). Durability of resistance against fungal, bacterial and viral pathogens; present situation. *Euphytica*, *124*, 147–156.

Paul, C., Naidoo, G., Forbes, A., Mikkilineni, V., White, D., & Rocheford, T. (2003). Quantitative trait loci for low aflatoxin production in two related maize populations. *Theoretical and Applied Genetics*, *107*, 263–270.

Peressotti, E., Duchêne, E., Merdinoglu, D., & Mestre, P. (2011). A semi-automatic non-destructive method to quantify grapevine downy mildew sporulation. *Journal of Microbiological Methods*, *84*, 265–271.

Poland, J. A., & Nelson, R. J. (2011). In the Eye of the Beholder: The effect of rater variability and different rating scales on QTL mapping. *Phytopathology*, *101*, 290–298.

Pozar, G., Butruille, D., Silva, H. D., McCuddin, Z. P., & Penna, J. C. V. (2009). Mapping and validation of quantitative trait loci for resistance to *Cercospora zeae-maydis* infection in tropical maize (*Zea mays* L.). *Theoretical and Applied Genetics*, *118*, 553–564.

Saghai Maroof, M. A., Yue, Y. G., Xiang, Z. X., Stromberg, E. L., & Rufener, G. K. (1996). Identification of quantitative trait loci controlling resistance to gray leaf spot disease in maize. *Theoretical and Applied Genetics*, *93*, 539–46.

Saghai-Maroof, M. A., Soliman, K. M., Jorgensen, R. A., & Allard, R. W. (1984). Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. *Proceedings of the National Academy of Sciences of the United States of America, 81*, 8014–8018.

Sherwood, R., Berg, C., Hoover, M., & Zeiders, K. (1983). Illusions in visual assessment of *Stagonospora* leaf spot of orchardgrass. *Phytopathology*, *73*, 173–177.

Tanksley, S. (1993). Mapping polygenes. *Annual Review of Genetics*, *27*, 205–233.

Van der Linden, L., Bredenkamp, J., Naidoo, S., Fouché-Weich, J., Denby, K. J., Genin, S., Marco, Y., & Berger, D. K, (2013). Gene-for-gene tolerance to bacterial wilt in *Arabidopsis*. *Molecular Plant-Microbe Interactions*, *26*, 398–406.

Voorrips, R. E. (2002). MapChart: Software for the graphical presentation of linkage maps and QTLs. *Journal of Heredity*, *93*, 77–78.

Wang, J., Levy, M., & Dunkle, L. D. (1998). Sibling species of *Cercospora* associated with gray leaf spot of maize. *Phytopathology*, *88*, 1269–1275.

Wang, S., Basten, C. J., & Zeng, Z. B. (2012). Windows QTL Cartographer 2.5. Raleigh, NC.: Department of Statistics, North Carolina State University.

Ward, J. M. J., Laing, M. D., & Rijkenberg, F. H. J. (1997). Frequency and timing of fungicide applications for the control of gray leaf spot in maize. *Plant Disease*, *81*, 41–48.

Ward, J. M. J., Stromberg, E. L., Nowell, D. C., & W, N. (1999). Gray leaf spot: a disease of global importance in maize production. *Plant Disease*, *83*, 884–895.

Xu, L., Zhang, Y., Shao, S., Chen, W., Tan, J., Zhu, M., Zhong, T., Fan, X., & Xu, M. (2014). High-resolution mapping and characterization of *qRgls2*, a major quantitative trait locus involved in maize resistance to gray leaf spot. *BMC Plant Biology*, *14*, 230

Zeng, Z. (1994). Precision mapping of quantitative trait loci. *Genetics*, *139*, 1457–1468.

Zhang, Y., Xu, L., Fan, X., Tan, J., Chen, W., & Xu, M. (2012). QTL mapping of resistance to gray leaf spot in maize. *Theoretical and Applied Genetics, 125*, 1797–808.

Chapter 4

Candidate gene discovery using a bulked segregant approach to expression profiling

JN Korsman[1], A van der Walt[2], NA Olivier[3], BA Meisel[4], P Tongoona[5], BG Crampton[1], DK Berger[1]

[1]Department of Plant Science, Forestry and Agricultural Biotechnology Institute (FABI), Faculty of Natural and Agricultural Sciences, University of Pretoria, South Africa
[2]Centre for Proteomic and Genomic Research, Cape Town, South Africa
[3]Department of Plant Science, ACGT Microarray facility, Faculty of Natural and Agricultural Sciences, University of Pretoria, South Africa
[4]Monsanto, South Africa (Pty) Ltd, Bryanston, South Africa
[5]African Centre for Crop Improvement, University of KwaZulu-Natal, Pietermaritzburg, South Africa

The format of the European Journal of Plant Pathology was adhered to for this chapter.

I selected the RILs for the bulked segregant analysis, performed the sample preparation and processed the microarray, annotated the genes identified by expression profiling, and carried out all analyses unless otherwise stated and prepared the chapter. DK Berger provided intellectual input, funding and leaf photography. BG Crampton provided intellectual input. P Tongoona organised and implemented the Baynesfield field trial. A van der Walt did the RNA-seq analysis via the Galaxy pipeline performed at the Centre for Proteomic and Genomic Research (CPGR). NA Olivier helped with microarray processing and data analysis. B Meisel helped with RNA extractions.

Abstract

The response of maize to the foliar pathogen, *Cercospora zeina,* the causal agent of grey leaf spot in South Africa is investigated here by expression profiling a pooled bulk of resistant RILs and comparing it to a pooled bulk of susceptible RILs. We aimed to identify candidate genes that are differentially expressed between resistant and susceptible RILs during later stages of grey leaf spot infection of maize. Another aim was to determine if these genes were located within previously identified QTL regions. Two pools of bulked RILs were used for expression profiling using both microarray and RNA-seq technologies in a modified version of a bulked segregant analysis. Differentially expressed genes and associated gene ontologies found in the bulks were compared. The genes identified here were manually associated with the QTL to identify candidate genes within the QTL regions. Cell death was identified as a possible strategy in the resistant RILs, whereas the susceptible RILs showed an enrichment of GO terms linked to responses to biotic stress. We found some genes with good support as candidates for playing a role in the resistance response, such as a resistance to *Peronospora parasitica* protein 13 (RPP13)-like gene, a powdery-mildew-resistance gene o (*mlo*) gene, and two leucine rich repeat receptor like kinases. Additional candidate genes that co-localised with QTL regions are a malectin-containing receptor like kinase and an enhanced disease resistance 1 gene. These are likely to be constitutivly expressed at higher levels in the resistant interaction, although expression profiling of a no-disease sample was not included in the experimental design. There were many genes involved in pathogen defence with higher expression in the susceptible plants. This could be explained by the fact that they have a larger amount of pathogen to combat while the resistant RILs most likely defeat the fungus during early infection stages.

Keywords:

bulked segregant analysis, *Cercospora zeina*, grey leaf spot, maize, plant pathogen interaction

Abbreviations:

GLS: grey leaf spot

GO: Gene ontology

RIL: Recombinant inbred line

R bulk: pooled bulk of resistant RILs

S bulk: pooled bulk of susceptible RILs

4.1 Introduction

The response of maize to the foliar pathogen, *Cercospora zeina,* which causes grey leaf spot (GLS) is investigated here by expression profiling of a pooled bulk of resistant recombinant inbred lines (RILs) (R bulk) and comparing this to a pooled bulk of susceptible RILs (S bulk). All RILs were exposed to the same *C. zeina* pressure in the field. With this we hoped to identify genes that are differentially expressed between resistant and susceptible RILs during later stages of GLS infection and further identify biological pathways, which may be involved in defence against GLS.

Gene expression studies using microarrays, and lately RNA-seq, usually provide long lists of genes with differential regulation between samples. The identification of many genes and pathways can be expected from expression profiling of plant-pathogen interactions, whether the resistance phenotype is quantitative (polygenic) or due to a single resistance gene, as genes in upstream and downstream pathways are also detected. Investigating plant-pathogen interactions using microarrays has been reviewed by Lodha and Basak (2012) and Wan et al. (2002). Pooling samples on a microarray can be useful if biological variation is high compared to technical variation (Kendziorski et al. 2003). The biological variation can be considered high in the maize RIL population used, as the RILs are homozygous, but have a different allele at, on average, half their loci when compared to another RIL from the same population.

In order to identify possible candidate genes underlying the quantitative trait loci (QTL) discussed in Chapter 3, a modified version of bulked segregant analysis was implemented. Bulked segregant analysis based on RNA-seq data has been used to map and clone a gene from a segregating maize population (Liu et al. 2012). In the study the authors were concerned with mapping the gene using RNA-seq identified SNPs rather than looking at expression analysis. Another two studies have used bulked segregant analysis based on differential gene expression analysis from microarray experiments in the yeast *Saccharomyces cerevisiae* and in potato (Brauer et al. 2006; Kloosterman et al. 2010).

This study does not use the classical bulked segregant analysis method (Michelmore et al. 1991). This classical method associates the two bulks' phenotypes with two versions of a polymorphic molecular marker in the region of the genome where the trait is encoded. All other markers will not associate with the phenotypes, as they will be mixed within each bulk. The above-mentioned bulked segregant analysis studies (Brauer et al. 2006; Kloosterman et al.

2010; Liu et al. 2012) were based on one major effect QTL each. A number of lesser effect QTL, as present in the bulked RILs used here, will have a confounding effect on the association of the markers or genome regions to expression phenotypes of genes. As there are more than one QTL from the same parent, there will be more than one region on the genome that the genes will associate with. Bulked segregant analysis usually has one region that associates with resistance, or trait of interest. At this region the markers or alleles will be the same while the rest is a mix in the bulks.

A candidate gene has previously been found in maize for GLS resistance by associating a SNP in a glutathione S-transferase gene with multiple disease (including GLS) resistance, in a maize association mapping panel (Wisser et al. 2011). A maize wall-associated kinase conferring quantitative resistance against a head smut fungus was identified after fine mapping a head smut resistance QTL and sequencing the underlying region between flanking markers, followed by functional analysis of the most likely candidate (Zuo et al. 2015). Bypassing such association mapping or fine mapping in the identification of a candidate gene would save time and money.

The main aim of this study was to identify candidate genes for resistance and susceptibility mechanisms. Genes differentially expressed between the resistant and susceptible plants may be considered as candidate genes in the defence mechanisms of maize against *C. zeina*. Additionally we aimed to validate the differential expression of genes by comparing genes identified in microarray and RNA-seq analyses carried out on the same samples.

Another aim was to determine if these genes are located within QTL regions, which may indicate that an allelic difference in the gene or promoter may contribute to the QTL effect. Due to the quantitative nature of GLS resistance a number of loci are expected to play a role in the resistance phenotype. The genes identified here were manually associated with the QTL from chapter 3 to identify candidate genes within the QTL regions.

4.2 Materials and methods

All reagents and chemicals were purchased from Sigma-Aldrich (Johannesburg, South Africa), unless otherwise stated.

4.2.1 Biological material

The CML444 × SC Malawi recombinant inbred maize population described in chapter 3 was used again here. Naturally infected leaf material, with additional powdered *C. zeina* infected leaf material placed in the whorls at the five to seven leaf stage, was collected at 104 days after planting from three replicate blocks (i.e. three biological replicates) of the 2008/2009 field trial in Baynesfield, KZN. For each of the 12 RILs (chosen as descibed below) from each replicate block one leaf segment of 120 mm length per plant was sampled from the first leaf above the ear for two individual representative plants from a row of ten plants. The two leaf segments per row were frozen together in liquid nitrogen in the field, transported to the lab on dry ice, stored at -80ºC and later used for RNA isolation.

4.2.2 Choice of RILs for expression profiling

The GLS disease scores for all the 145 RILs from the Baynesfield 2009 field trial were provided by Prof P. Tongoona (UKZN). The RILs were sorted according to the average GLS score from the Baynesfield 2009 season. Those with the twenty highest and twenty lowest scores were considered for expression profiling. The choice of RILs was further based on the presence of the resistance QTL in the genomes identified for this RIL population in the Baynesfield 2009 field trial in chapter 3 (1Cz, 3Cz_1, 5Cz, 6Cz, 9Cz_2.1, 9Cz_2.2 and 10Cz_2, based on qPCR, and 6Lesion_area, based on the lesion area), and in Berger et al. (2014) (3H_GLS, 6H_GLS, 9H_GLS and 10H_GLS, based on the whole row GLS disease rating). The presence of the QTL was determined by the marker data from Qmap2.0 (Figure 3.5) of the peak QTL marker in agreement with the flanking marker or markers (Table 4.1). Whether the marker represented a more resistant or susceptible allele was determined by the parental genotype associated with the additive effect of the QTL (see Table 3.5), which is based on the phenotype scores from the GLS ratings and qPCR assay. A negative value indicated that the CML444 genotype (“A”) represented a resistance allele, and a positive value indicated that an SC Malawi genotype (“B”) represented a resistance allele. CML444 is the more resistant parent and therefore has low GLS scores or less *C. zeina* DNA in a leaf and SC Malwi has high GLS scores or more *C. zeina* DNA in a leaf. The RILs have mixed “A” and “B” genotypes, with blocks of “A” and “B” alternating throughout the genomes. Where the RILs’ low (resistant)

scores are associated with the "A" genotype in the QTL mapping, the additive values are negative with the resistance allele originating from CML444. Where low scores are associated with the "B" genotype, the additive values are positive with the resistance allele originating from SC Malawi. If there was not sufficient agreement between the peak and flanking markers, or there was missing data, the allele was considered undetermined for that QTL. Six of the most resistant RILs, with the highest number of resistance alleles for the QTL, and six of the most susceptible RILs with the highest number of susceptible alleles for the QTL, were selected (Table 4.1).

4.2.3 RNA extractions

Total RNA was extracted separately for the three biological replicates collected from the three replicate blocks in the field for each of the twelve RILs. One biological replicate consisted of the leaf material harvested from two plants of a single row that were pooled together and ground to a powder in liquid nitrogen. RNA was extracted from frozen leaf material using QIAzol® Lysis Reagent (QIAGEN, Hilden, Germany). One ml of QIAzol® was added to approximately 100 mg maize leaf tissue and incubated at 65ºC for 5 minutes with occasional agitation. The leaf tissue was pelleted by centrifugation at 12 000×g for 10 minutes at 4ºC, the supernatant transferred to a new tube and 300 µl of chloroform added. The mixture was vortexed and incubated at room temperature for 3 minutes, followed by centrifugation at 10 000×g for 15 minutes at 4ºC to separate the phases. The upper aqueous phase containing the RNA was transferred to a new tube and the RNA precipitated by the addition of half a volume of isopropanol and half a volume of 0.8 M sodium citrate, 1.2 M sodium chloride. The tube was inverted several times and incubated at room temperature for 10 minutes after which the RNA was collected in the pellet by centrifugation at 10 000×g for 10 minutes at 4ºC. The RNA pellet was washed with 75% ethanol and air dried before being dissolved in Nuclease free water. Any remaining DNA was removed by treatment with RNase-Free DNase (QIAGEN) in solution according to the manufacturer's instructions. RNeasy® Mini Kit (QIAGEN) was then used to cleanup the RNA, which was eluted in Nuclease free water. The concentration and quality of the RNA was analysed with the NanoDrop™ 1000 spectrophotometer (Thermo Scientific, Wilmington, USA) as well as on an Agilent 2100 Bioanalyser (Agilent Technologies, Santa Clara, USA). The RNA was stored at -80ºC. One microgram of total RNA from each of the six resistant RILs (R bulk) and each of the susceptible RILs (S bulk) were pooled for each of the three biological replicates.

Table 4.1: A table showing the marker data for the QTL of the six resistant and six susceptible bulked RILs. The RIL genotype numbers are listed on the left, and the average GLS scores on the right with the average GLS scores for the bulks in bold. The QTL name, as in chapter 3 or Berger et al. (2014), with the flanking and peak (bold) marker names below, head the columns containging the marker data for each RIL. The marker data is green if inherited from the parental genotype associated with the resistance allele for the QTL and brown if from the genotype associated with the susceptible allele. A represents the CML444 genopyte and B represents that of SC Malawi. The association of the marker genotype with resistance or susceptibility is based on the additive effect determined during QTL mapping and stated in the bottom row.

	1Cz[a]			**3H_GLS[b]**			**3Cz_1[a] & 3Cz_2[a]**				**5Cz[a]**				**6H_GLS[b] & 6Cz[a] & 6Lesion_area[a]**				**9Cz_2.1[a]**			**9H_GLS[b] & 9Cz_2.2[a]**		**10Cz_2[a]**				
RIL	umc106a	**umc147b**	umc1111	bnlg1325	**bnlg1447**	umc92a	bnlg1108	**umc63a**	**bnlg1182**	bnlg1754	bnl5.71a	**umc1155**	npi237	umc54	mmc0241	**umc1424**	**umc36**	umc39	umc81	**bnl8.17**	umc1231	**umc1733**	bnlg1588	bnlg236	**bnl7.49a**	bnlg1450	Ave GLS score	
387	A	A	B	A	A	A	A	A	A	A	B	A	A	A	A	B	B	B	B	B	B	B	B	A	A	B	1.51	
327	A	A	A	B	-	A	A	A	B	A	B	A	A	A	A	A	A	A	A	A	A	B	B	A	A	A	2.56	
366	B	B	A	B	B	B	A	A	A	A	A	A	A	A	A	B	B	B	A	A	A	B	B	A	-	-	2.56	**2.9**
125	A	A	A	A	A	A	B	B	A	B	A	-	A	A	A	A	B	B	B	B	B	B	B	A	A	A	3.38	
166	B	B	B	B	B	B	B	B	B	B	B	A	A	A	B	-	B	B	A	A	-	B	A	A	A	A	3.49	
381	A	A	A	A	A	B	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	A	-	A	3.82	
31	B	B	-	B	B	A	-	A	A	A	A	A	A	A	B	A	A	A	A	A	A	A	A	A	B	B	6.87	
259	B	-	-	B	B	B	-	-	A	A	B	-	-	A	A	-	A	A	B	B	-	A	A	A	-	-	6.95	
362	A	A	A	A	B	B	A	A	A	B	A	B	B	B	A	-	A	A	B	B	B	B	A	B	B	A	6.95	**7.4**
349	B	B	B	A	A	A	B	B	-	-	A	A	B	B	B	B	A	A	-	-	-	A	A	B	-	B	7.67	
329	B	B	A	B	B	B	B	B	B	B	A	B	B	B	A	B	B	B	B	B	B	A	A	A	B	B	7.87	
284	A	A	B	B	B	B	A	B	-	A	B	B	B	B	B	B	A	B	A	A	A	A	A	A	B	A	8.08	
	-0.2			-0.41			-0.24 & -0.24				-0.23				0.50 & 0.26 & 0.21				-0.25			0.40 & 0.23		-0.22			← Additive effect	

[a] QTL from Chapter 3, based on subset of 100 RILs
[b] QTL from Berger et al. (2014), based on 145 RILs

4.2.4 Microarray

4.2.4.1 Experimental design

Direct comparison was used between the resistant and susceptible bulks with three independent biological replicates on different arrays (Naidoo et al. 2005). A technical replicate of the first biological replicate was included, due to the availability of four arrays on one Agilent maize 44 K microarray slide (Agilent ID 016047) (Agilent Technologies, Santa Clara, USA) (Coetzer et al. 2011) (see Figure 4.1).

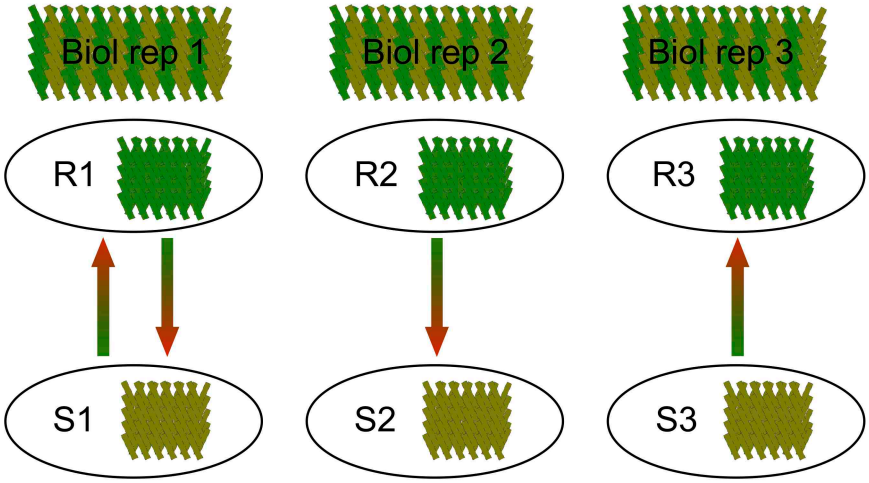


Figure 4.1: Experimental design of the microarray experiment. Three biological replicates from three replicate blocks of the Baynesfied 2009 field trial were used in the experiment. The ovals represent pools of RNA from six resistant (R) or six susceptible (S) maize lines. The arrows represent individual arrays their heads indicate Cy5 labelled samples and the tails indicate Cy3 labelled samples.

4.2.4.2 Sample preparation

Microarrays require a relatively large quantity (usually over a microgram) of input RNA. This was not possible for this experiment, as much of the RNA was destined for RNA-seq and other experiments as well. Thus linear sample amplification was employed to produce amplified RNA (aRNA). One microgram of the bulked total RNA was amplified for microarray hybridisation with the Ambion® Amino Allyl MessageAmp™ II aRNA Amplification Kit (Life Technologies, Johannesburg, South Africa) according to the manufacturer's instructions. First strand cDNA synthesis was carried out on 1 µg of total RNA. The T7 Oligo(dT) primer, First strand buffer, Ribonuclease (RNase) inhibitor, dNTP mix and Reverse transcriptase were provided in the kit. Second strand synthesis was carried out for two hours with the manufacturer's Second Strand Buffer, dNTP mix, DNA Polymerase and RNase H at 16°C. The cDNA was purified using a silica spin-column before *in vitro* transcription for the synthesis of

the aRNA. The *in vitro* transcription was carried out with the cleaned cDNA, amino allyl UTP (aaUTP) and UTP in a 3:1 ratio, ATP, CTP, GTP, T7 Reaction Buffer and T7 Enzyme Mix. The reaction was allowed to continue at 37°C for 14 hours, after which the aRNA was purified using a silica spin-column. The aRNA was eluted in 100 µl of RNase-free water, preheated to 50°C for optimal elution. The aRNA concentration was determined using the NanoDrop™ 1000 spectrophotometer (Thermo Scientific) and an aliquot of 4 µg aRNA was dried in a vacuum centrifuge at 45°C.

For the labelling reaction the aRNA aliquots were dissolved in 5 µl of $NaHCO_3$ buffer by occasional tapping of the tubes, and incubation at room temperature for 20 minutes. Fluorescent Cy3 or Cy5 (dissolved in DMSO) were coupled to the aRNA during an incubation at room temperature for two hours in the dark with occasional gentle agitation. Post-dye coupling RNA purification was carried out using an RNeasy® MinElute® cleanup kit (QIAGEN) according to the manufacturer's instructions, with two aRNA elution steps of 14 µl each in Nuclease free water. The amount of dye incorporated into the aRNA was measured in pmol/µl using the NanoDrop™ 1000 spectrophotometer (Thermo Scientific). The RNA concentration and the 260 nm/280 nm absorbance ratio was also determined.

4.2.4.3 Hybridisation and data capture

For hybridisation, two sets of labelled aRNA (100 pmol each of Cy3 and Cy5) were diluted in Blocking agent and Fragmentation buffer (Agilent) and incubated at 60°C to fragment the RNA for 30 minutes. GEx Hybridization Buffer HI-RPM (Agilent) stopped the fragmentation reaction and ensured the correct environment for hybridisation. The samples were loaded onto an Agilent 4 × 44 K gasket slide and the Agilent 44 K maize microarray, which comprises printed 60mer nucleotides (Hughes et al. 2001), was assembled onto the gasket slide. The Agilent 44 K maize microarray consists of 42 034 gene probes, targeted at about 39 000 individual transcripts (Ma et al. 2008), of which 23 668 are targeted to sense gene models (Coetzer et al. 2011). Hybridisation was carried out at 65°C with continual rotation at 10 rpm in an Agilent hybridisation oven for 17 hours.

The array was sequentially washed in Gene Expression Wash Buffer 1 and 2 (Agilent) in sterile 50 ml centrifuge tubes. The array and gasket slide were disassembled in wash buffer 1 and the array moved to a second tube of wash buffer 1 at room temperature for 1 minute with gentle

rolling. The array was then washed in pre-warmed wash buffer 2 at 37°C for 1 minute. The array was dried in a 50 ml centrifuge tube by centrifugation at 2000×g for 2 minutes.

The array was scanned with a GenePix® 4000B scanner (Molecular Devices, Sunnyvale, USA; previously Axon Instruments) at a 5 µm resolution. Two tiff files were produced, one for each dye.

4.2.5 Microarray data analysis

The spot intensities were determined in the GenePix® Pro 6.1 (Molecular Devices) software. A GenePix Array List (.gal) file with a grid template for the 4 × 44 K array was applied and manually adjusted to fit the array. Circular features were automatically delimited by the software with feature movement limited to a maximum of 20 µm. The array was checked for features that were not already detected as atypical by the software. These features were manually flagged for removal from the analysis along with the automatically flagged atypical features. Feature information and intensities were exported for analysis in the limma (linear models for microarray data) package (version 2.14.7) in the R software environment (version 2.7.2) (http://cran.r-project.org/). The background correction method employed was an adaptive foreground/ background correction using method="normexp" and offset=50 (Ritchie et al. 2007). Within array normalization was carried out using the global lowess (locally weighted scatterplot smoothing) method (Smyth and Speed 2003). Between array normalization was performed using the aquantile method. The dye swap technical replicates of replicate 1 were incorporated into the design as one biological replicate in the design (-1,1,0,0), along with biological replicate 2 (0,0,-1,0), and biological replicate 3 (0,0,0,1), which were dye swaps of each other for balancing any dye effects. A linear model was generated using only the probe spots (isGene), thus the control spots were excluded. The R bulk vs S bulk $\log_2$ fold changes for the microarray probes were calculated and the associated transcripts with significant differential expression were identified with an empirical Bayes method, which adjusts the p-value to account for the multiple comparison false discovery rates (Smyth 2004).

The expression values were analysed in Microsoft Excel or LibreOffice Spreadsheets. The values for the probes were sorted by the adjusted p-value, and those without statistical support, i.e. an adjusted p-value >0.05, and a fold change of between 1.5 and -1.5 were filtered out of the analysis.

4.2.6 RNA-seq expression profiling

RNA-seq was carried out at Beijing Genomics Institute (BGI), which utilises Illumina technology. The same samples used for the microarray analysis were sequenced. The output was 50 bp paired end read lengths.

The RNA-seq analysis was performed in Galaxy (http://galaxyproject.org/) (Goecks et al. 2010) using the Tuxedo suite. TopHat version 1.2 (Trapnell et al. 2009) was used for alignment or mapping of the sequences to the B73 reference maize genome (B73 RefGen_v2) and Cufflinks version 1.1.0 (Trapnell et al. 2010) was used for transcript assembly and to estimate abundance levels. Flagstats (part of Samtools version 0.1.14 – http://samtools.sourceforge.net) was used to obtain mapping statistics from the Tophat BAM files. Transcripts within samples were normalised with the compatible hit normalisation strategy and only transcripts mapping to known genes were used to calculate fragments per kilobase of transcript per million fragments mapped (FPKM) values. The bias detection and correction algorithm was run to improve the accuracy of transcript abundance calculation. Cuffdiff version 1.1.0, part of Cufflinks, was used for the calculation of differential expression between the resistant and the susceptible biological replicate samples (Trapnell et al. 2012). The compatible hits normalisation strategy and fragment bias correction were run under Cuffdiff as well. Cuffdiff uses a two-tailed student's t-test to calculate p-values for the changes in transcript abundance. This was then adjusted using the Benjamini-Hochberg correction for multiple testing resulting in a Q-value for a FDR of 5%.

4.2.7 Annotation

The Maize microarray annotation database (Coetzer et al. 2011) was used to assign gene names and functions to the Agilent array probes, as well as to the genes identified using RNA-seq that were represented on the array. Genes not represented on the array were annotated using the best nucleotide blast (megablast) hits from NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) based on the coding sequence from the B73 RefGen_v3 gene model information at MaizeGBD (http://www.maizegdb.org/gene_model.php). If megablast failed to provide descriptive hits, protein blast (blastp) based on the predicted amino acid sequence was used if it provided a description. Alternatively the *Arabidopsis* best hit (from http://www.arabidopsis.org/) or the Rice best hit (from http://rice.plantbiology.msu.edu/) obtained from the B73 RefGen_v3 gene model information at MaizeGBD (http://www.maizegdb.org/gene_model.php) based on the Phytozome annotations (Schnable et al. 2009) were used.

MapMan version 3.5.1R2 (Thimm et al. 2004) was employed to gain a visual overview of the differentially expressed genes putatively involved in defence against *C. zeina*. The mapping file based on the B73 maize genome, available from the MapMan site (http://mapman.gabipd.org/web/guest/mapmanstore), was altered to contain gene names instead of transcript names so that they corresponded to the gene names in the RNA-seq results. Cellular compartment localisation of proteins was determined using Plant-mPLoc (Chou and Shen 2010) at http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/ and WoLF PSORT (Horton et al. 2007) at http://www.genscript.com/psort/wolf_psort.html in agreement with literature where available. Protein domains were predicted using NCBI's conserved domain database (Marchler-Bauer et al. 2015) at http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml or InterPro scan (Jones et al. 2014) at http://www.ebi.ac.uk/interpro/search/sequence-search. The TMHMM web server (Krogh et al. 2001) was used to predict transmembrane regions in proteins at http://www.cbs.dtu.dk/services/TMHMM-2.0/.

4.2.8 Gene ontologies

Gene ontology (GO) terms for genes of interest were obtained from the GO Analysis Toolkit and Database for Agricultural Community (AgriGO) (http://bioinfo.cau.edu.cn/agriGO/analysis.php) (Du et al. 2010). For over-representation analysis, AgriGO's Singular Enrichment Analysis tool was used with the hypergeometric statistical test using the Hochberg (FDR) multiple testing adjustment method and the Plant GO slim gene ontology type. The *Zea mays* ssp V5a was used as the supported species in AgriGO for comparison for the RNA-seq analysis. A customised analysis was used for the microarray, where the genes of interest were compared to only the genes present on the Agilent 4 × 44 K array. The customised input consisted of the probe names and GO terms associated with the gene models that the probes represent.

4.2.9 Association of genes with QTL

The QTL regions investigated were those for this RIL population in the Baynesfield 2009 field trial in chapter 3, i.e. the fungal load QTL (1Cz, 3Cz_1, 5Cz, 6Cz, 9Cz_2.1, 9Cz_2.2 and 10Cz_2), the lesion area QTL (6Lesion_area) and the GLS resistance QTL (Berger et al. 2014) (3H_GLS, 6H_GLS, 9H_GLS and 10H_GLS).

The LOD2 intervals of the QTL were compared to a "look up table" described in Christie (2014) that estimates the bp position on the physical map (B73 RefGen_v2) for every 2 cM

interval on the genetic map (Qmap2.0). The B73 RefGen_v2 bp positions of the differentially expressed genes from both the microarray and RNA-seq analysis with probes on the array were obtained from http://maizearrayannot.bi.up.ac.za/ (Coetzer et al. 2011). The positions of the differentially expressed genes from the RNA-seq analysis, which did not correspond to probes on the array, were obtained from http://www.maizegdb.org/. The bp positions of the differentially expressed genes were compared to each QTL's LOD2 interval's estimated bp positions.

4.3 Results

4.3.1 Expression profiling of resistant and susceptible bulked RILs

The RIL population planted at Baynesfield in 2008, which was scored for GLS resistance and sampled in 2009 for *C. zeina* quantification (chapter 3) and expression profiling (this chapter) showed good GLS symptoms without visual evidence of other stresses (data not shown, but visible in Figure 4.2). Six of the most resistant and six of the most susceptible RILs were identified based on their GLS disease ratings and marker data in QTL regions (Table 4.1). The resistant RILs with genotype numbers 387, 327, 366, 125, 166 and 381, with an average disease score of 2.9, were used as the R bulk and the susceptible RILs with genotype numbers 31, 259, 362, 349, 329 and 284, with an average disease score of 7.4, were used as the S bulk. The differences in their disease phenotypes can be visualised in Figure 4.2. The S bulk has many more visible lesions than the R bulk.

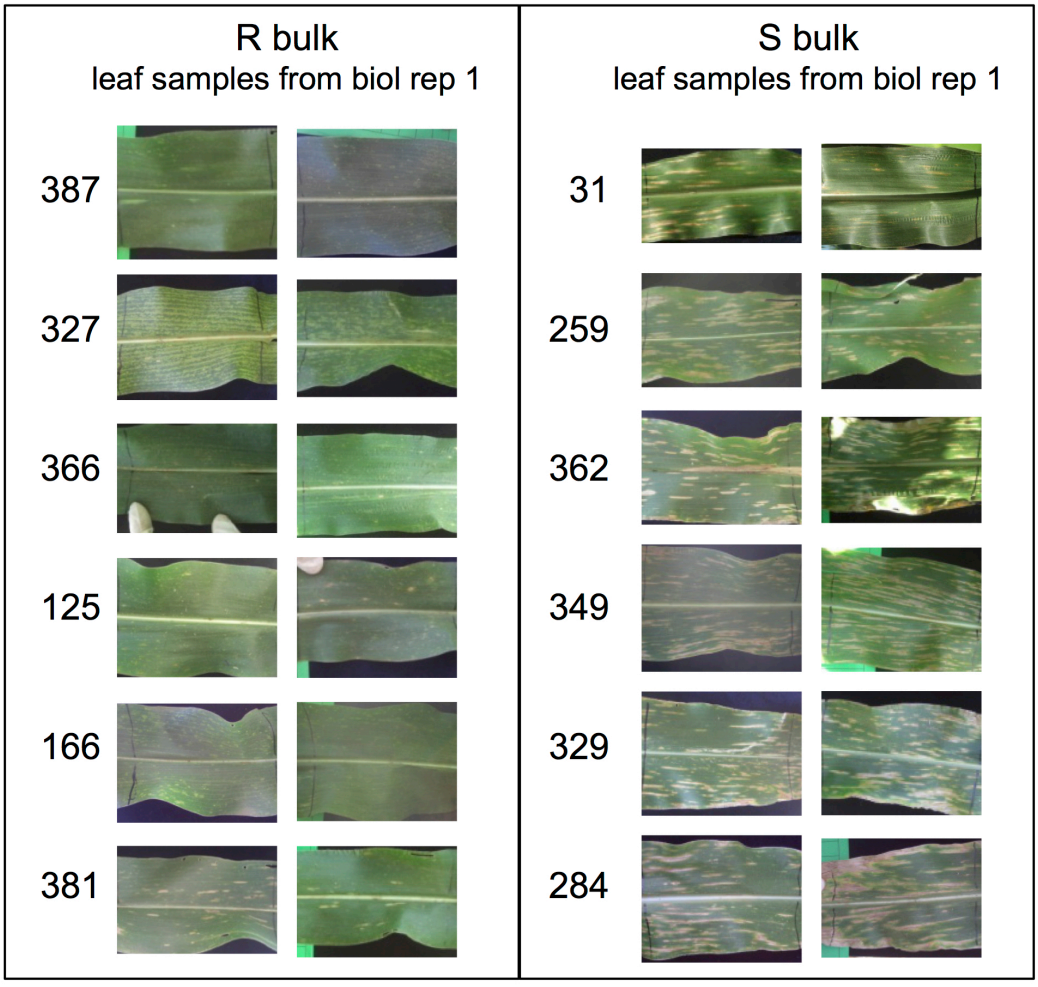


Figure 4.2: Leaf samples of the six resistant RILs and six susceptible RILs used in the bulks, from the first biological replicate. Photos by Prof DK Berger.

RNA was sucessfully extracted from the leaf material of these RILs and was used for the microarray and RNA-seq expression profiling experiments. Quality control of the RNA was carried out at CPGR on an Agilent bioanalyser. All the samples passed the QC process and the average RIN number was 7.7. For a visual comparison of the quality of total RNA isolated from the Baynesfield samples and control glasshouse samples a denaturing RNA gel was used, see Figure A1 in Appendix A. The RNA of the six resistant RILs from each biological replicate were pooled, as was the RNA of the six susceptible RILs for each biological replicate.

4.3.1.1 Microarray analysis

The R bulk and S bulk of each biological rep were compared in a microarray experiment (Figure 4.1). Quality control data for the microarray experiment is listed in table A1 in Appendix A. The mean of the Cy5 (red channel) probes was between 2957 and 3343 flourescent units, while the mean of the Cy3 (green channel) probes was between 2443 and 4533 flourescent units. The green channel usually had a higher flourescence than the red channel, except for the third biological replicate array, where the red channel had a higher flourescence. Overall, the microarray data was of sufficient quality and thus used for analysis. The differences in green and red channels for each array were normalised using a global lowess normalisation. The slight effect of the lowess normalisation on the log ratio vs the intensity is addressed in the MA plots of the four arrays is shown in Figure A2 in Appendix A. The normalisation adjusts the differences in Cy3 and Cy5 intensity and brings more spots into symmetry around M=0, where M represents the log ratio, i.e. it assumes that most probes will be associated with genes with unchanged expression levels. Aquantile normalisation was used to normalise the probe data between the arrays using dye swaps to balance any dye effects.

In order to find genes with differential expression between the R bulk and S bulk a probe list of differentially expressed genes was generated using a linear model in the limma package (version 2.14.7) (Ritchie et al. 2015). An output table with probe names, $\log_2$ fold change and an adjusted p-value was produced. The list was sorted by adjusted p-value, and cut off at adj $p = 0.05$. The $\log_2$ fold change was converted to the actual fold change for both R/S and S/R to find genes with higher expression in the R bulk as well as genes with higher expression in the S bulk. One hundred and ninety one probes showed statistically significant (adjusted p-value < 0.05, or false discovery rate of 5%) higher expression in the resistant bulk compared to the susceptible bulk (R/S). The highest fold change in the R/S top table was 5.2 and the lowest was 1.6. Of these, 46 probes had ambiguous (hits to more than one gene model) or no

gene hits as determined in the Maize microarray annotation database (Coetzer et al. 2011). These probes were removed from the analysis leaving 145 gene models. The genes with the ten highest differences in expression levels between the R bulk and the S bulk (R/S) are listed in Table 4.2. The top 100 gene lists are provided in Appendix B, Table B1. Two hundred probes showed statistically significant higher expression in the susceptible bulk compared to the resistant bulk (S/R). The highest fold change in the S/R top table was 3.5 and the lowest was 1.6. There were 49 probes with ambiguous or no gene hits which were removed from the analysis leaving 151 gene models. The genes with the ten highest differences in expression levels between the S bulk and R bulk (S/R) are listed in Table 4.3. The top 100 gene lists are provided in Appendix B, Table B2. The $\log_2$ fold change of these probes can be visualised in a volcano plot shown in Figure A4(a) in Appendix A.

Table 4.2: Top ten genes with the highest expression in the resistant bulk compared to the susceptible bulk (R/S) as determined by microarray analysis.

Gene	Sequence description	FC[a] R/S	adjP	Accession no.	description source
GRMZM2G373522	dehydrin	5.16	0.003	NP_001105327	BLAST
GRMZM5G859316	expansin precursor	4.42	0.006	LOC_Os03g44290.1	Rice[b]
GRMZM2G055698	Putative Rhomboid homologue	3.65	0.006	LOC_Os03g02530.1	Rice
GRMZM2G076972	hypothetical protein	3.23	0.006	NP_001145623	BLAST
GRMZM2G103771	mitochondrial import inner membrane translocase subunit Tim17	3.05	0.011	LOC_Os01g19770.1 /ADK88900	Rice/BLAST
GRMZM2G034843	hydroxyproline-rich glycoprotein family protein	2.86	0.009	LOC_Os09g16280.1	Rice
GRMZM2G429842	Subtilisin homologue	2.78	0.008	LOC_Os09g26920.1	Rice
GRMZM2G117164	Homeobox-transcription factor 41 (hb41)	2.73	0.008		MaizeGDB[c]
GRMZM2G144504	RHO guanyl-nucleotide exchange factor 7	2.66	0.022	AT5G02010.1	BLAST
GRMZM2G154747	AWPM-19-like membrane family protein	2.61	0.008	LOC_Os07g24000.1	Rice

[a] FC: Fold change (R bulk/S bulk)
[b] Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/)
[c] MaizeGDB (http://maizegdb.org/)

Table 4.3: Top ten genes with the highest expression in the susceptible bulk compared to the resistant bulk (S/R) as determined by microarray analysis.

Gene	Sequence description	FC[a] S/R	adjP	Accession no.	description source
GRMZM2G427815	peroxidase precursor putative expressed	3.53	0.006	LOC_Os07g48030.1	Rice[b]
GRMZM2G311036	benzoxazinone synthesis10 (bx10)	2.9	0.023		MaizeGDB[c]
GRMZM2G098577	eukaryotic initiation factor iso-4F subunit p82-34	2.89	0.008	LOC_Os04g42140.1	Rice
GRMZM5G842645	Putative lysine decarboxylase family protein	2.73	0.016	AT5G06300.1	TAIR[d]
GRMZM2G050450	transferase family protein	2.72	0.008	LOC_Os06g49660.1	Rice
GRMZM2G071390	Cupin domain containing protein/germin-like protein	2.63	0.008	LOC_Os12g05860.1 /ACJ64505	Rice/BLAST
GRMZM2G158097	hypothetical protein	2.53	0.009	NP_001167975.1	BLAST
GRMZM2G021598	OTU-like cysteine protease family protein	2.45	0.012	LOC_Os09g31280.1	Rice
GRMZM2G001572	unknown	2.42	0.012	ACF82444.1	TAIR
GRMZM2G136508	amino acid permease family protein	2.38	0.011	LOC_Os04g35540.1 /NP_001148156	Rice/BLAST

[a] FC: Fold change (S bulk/R bulk)
[b] Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/)
[c] MaizeGDB (http://maizegdb.org/)
[d] TAIR (https://www.arabidopsis.org/)

4.3.1.2 RNA-seq analysis

The same resistance and susceptible bulks (three biological replicates each) were analysed by RNA-seq. The quality of the RNA-seq experiment was measured by a quality score plot (Figure A3 in Appendix A) which indicated that the quality of all the base calling across the read lengths obtained were reliable and thus used for further analysis. The statistics for the reads and mapping are summarised in Table A2, in Appendix A. The number of reads for a sample that were mapped uniquely in pairs to the B73 reference maize genome ranged from 26 214 016 reads to 37 703 692 reads, which correspond to 51.7% to 55.1% of the mapped reads, respectively. These were used by Cufflinks for transcript assembly and to calculate FPKM values. The average FPKM values of the three biological reps of the R bulks were compared to the average FPKM values of the three biological reps of the S bulks to identify and validate genes with differential expression between two bulks.

Data was obtained for approximately 24 000 genes per sample. The statistics for these genes are summarised in Table A3 in Appendix A. There were 278 genes that showed statistically significant higher expression levels in the resistant bulk relative to the susceptible bulk (R/S),

with fold changes ranging between 30.6 and 1.4. A list of genes with the ten highest differences in expression levels between the R bulk and S bulk (R/S) is provided in Table 4.4. In the susceptible bulk relative to the resistant bulk (S/R), there were 1 072 genes with higher expression, with fold changes between 7.1 and 1.4. The ten genes with the highest differences in expression levels between the S bulk and R bulk (S/R) is provided in Table 4.5. An additional list of eight genes only detected in the susceptible bulk (FPKM=0 in resistant bulk, thus no ratio between the bulks can be calculated) is given in Table 4.6. These genes could be considered to have "higher" expression in the susceptible bulk, or they could represent the presence of genes in the susceptible RILs and deletions in the resistant RILs. The top 100 gene lists are provided in Appendix B, Tables B3 and B4. The $\log_2$ fold change of these genes can be seen in a volcano plot shown in Figure A4(b) in Appendix A.

Table 4.4: Top ten genes with the highest expression in the resistant bulk compared to the susceptible bulk (R/S) as determined by RNA-seq analysis.

Gene	Sequence description	FC[a] R/S	Q value	Accession no.	description source
GRMZM2G020556	hypothetical protein	30.6	0.0000	DAA50064.1	BLAST
GRMZM2G020514	hypothetical protein	26.8	0.0000	NP_001144140.1	BLAST
GRMZM2G092804	loricrin isoform X3	23.8	0.0000	XP_008654907.1	BLAST
GRMZM2G026364	hypothetical protein	18.2	0.0000	NM_001176560.1	BLAST
GRMZM2G062527	GASR5 - Gibberellin-regulated GASA/GAST/Snakin family protein	17.1	0.0001	LOC_Os05g31280.1	Rice[b]
GRMZM2G109127	uncharacterized transcript	16.5	0.0000	XM_008658471.1	BLAST
GRMZM2G087558	uncharacterized transcript	15.1	0.0001	XM_008658471.1	BLAST
GRMZM5G822449	hypothetical protein	12.4	0.0003	EU976680.1	BLAST
GRMZM2G036861	chorismate synthase 2 chloroplast precursor	9.8	0.0000	LOC_Os03g14990.1	Rice
GRMZM2G547449	uncharacterized transcript	9.6	0.0002	XM_008658471.1	BLAST

[a] FC: Fold change (R bulk/S bulk)
[b] Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/)

Table 4.5: Top ten genes with the highest expression in the susceptible bulk compared to the resistant bulk (S/R) as determined by RNA-seq analysis.

Gene	Sequence description	FC[a] S/R	Q value	Accession no.	description source
GRMZM5G881353	apomucin-like	91.7	0.0168	XM_008666551.1	BLAST
GRMZM2G376743	low-temperature-induced 65 kDa protein-like	101.9	0	XP_008644725.1	BLAST
GRMZM2G061450	uncharacterized protein	103.4	0	NP_001142850.1	BLAST
GRMZM2G383125	DNA binding	127.2	0	AT3G47680.1	TAIR[b]
GRMZM2G041039	hypothetical protein	131.8	0.0121	ACG48669	BLAST
GRMZM2G377613	transcription factor HBP-1b	141.9	0.0055	LOC_Os05g48650.1	Rice[c]
GRMZM2G371375	late embryogenesis abundant protein group 3	174.7	0	LOC_Os04g52110.1	Rice
GRMZM2G417954	nine-cis-epoxycarotenoid dioxygenase5 (nced5)	176.9	0		MaizeGDB[d]
GRMZM2G174192	UDP-Glycosyltransferase superfamily protein	185.6	0	AT5G49690.1	TAIR
GRMZM2G136748	uncharacterized	334.1	0.0372	XR_562557.1	BLAST

[a] FC: Fold change (S bulk/R bulk)
[b] TAIR (https://www.arabidopsis.org/)
[c] Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/)
[d] MaizeGDB (http://maizegdb.org/)

Table 4.6: Genes only expressed in the Susceptible bulk, as determined by RNA-seq analysis.

Gene	Chr	Sequence description	description source	Accession no.	Possible role
GRMZM2G042438	5	uncharacterized	BLAST	XM_008647975.1	-
GRMZM2G112238	6	Jacalin-like lectin domain containing protein	Rice[a]	LOC_Os12g14440.1	carbohydrate binding; signalling; SA/JA responsive
GRMZM2G148964	6	classical arabinogalactan protein 1-like	BLAST	XM_008651216.1	Programmed cell death; found in papillae
GRMZM2G151390	3	uncharacterized	BLAST	XM_008675395.1	-
GRMZM2G151430	3	uncharacterized	BLAST	XM_008677439.1	-
GRMZM2G306371	7	uncharacterized	BLAST	XM_008675394.1	-
GRMZM2G355846	1	zinc finger family protein	Rice	LOC_Os03g04890.1	transcription regulation
GRMZM2G704475	9	late embryogenesis abundant protein 1, putative	Rice	LOC_Os03g20680.1	osmotic stress protection

[a] Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/)

4.3.2 Over representation analysis

Over representation analysis was carried out on the GO terms associated with the genes identified as having relatively higher expression in the resistant bulk (145 genes for the microarray and 278 genes for the RNA-seq) or in the susceptible bulk (151 genes for the microarray and 1072 for the RNA-seq) in order to determine which cellular pathways or functions are involved in the RILs' defence pathways.

The sets of genes from the microarray and RNA-seq analyses and were investigated separately. No significantly enriched GO terms were found in the genes from the R bulk microarray, but the RNA-seq highlighted six GO terms, four under biological process: death, cell death, metabolic process and cellular amino acid and derivative metabolic process, and two under molecular function: catalytic activity and kinase activity (Figure 4.3). The lists of genes associated with the significantly enriched GO terms can be seen in Figures B1 and B2 in Appendix B. In the susceptible bulk, the genes from the RNA-seq data had eight enriched GO terms. In the biological process set there was response to biotic stimulus, metabolic process, catabolic process, cellular amino acid and derivative metabolic process, lipid metabolic process and carbohydrate metabolic process. In the cellular component set there was extracellular region and in the molecular function set there was catalytic activity (Figure 4.4). The lists of genes associated with the significantly enriched GO terms can be seen in Figures B3, B4 and B5 in Appendix B. The microarray genes were associated with two enriched GO terms in the molecular function category: transporter activity, which was not identified in any other set, as well as catalytic activity (Figure 4.4). The lists of genes associated with the significantly enriched GO terms can be seen in Figure B6 in Appendix B.

4.3.3 Candidate genes for resistance model

The top tables from the microarray and RNA-seq were scrutinised to find genes that are likely to play a role in the resistance response to *C. zeina*. The best candidates were found in the RNA-seq top table. The list of genes with higher expression in the R bulk (Table 4.4) includes a loricrin gene (GRMZM2G092804), which Zila et al. (2014) suggest may be involved in cell membrane function. A GASA/Snakin family protein (GRMZM2G0625527) was identified, which is a gibberellin signalling protein that may play a role in plant stress and defence (Nahirñak et al. 2012). A chorismate synthase (GRMZM2G036861) was also found, which is involved in the shikimate pathway, upstream of salicylic acid and phenlypropanoid biosynthesis.

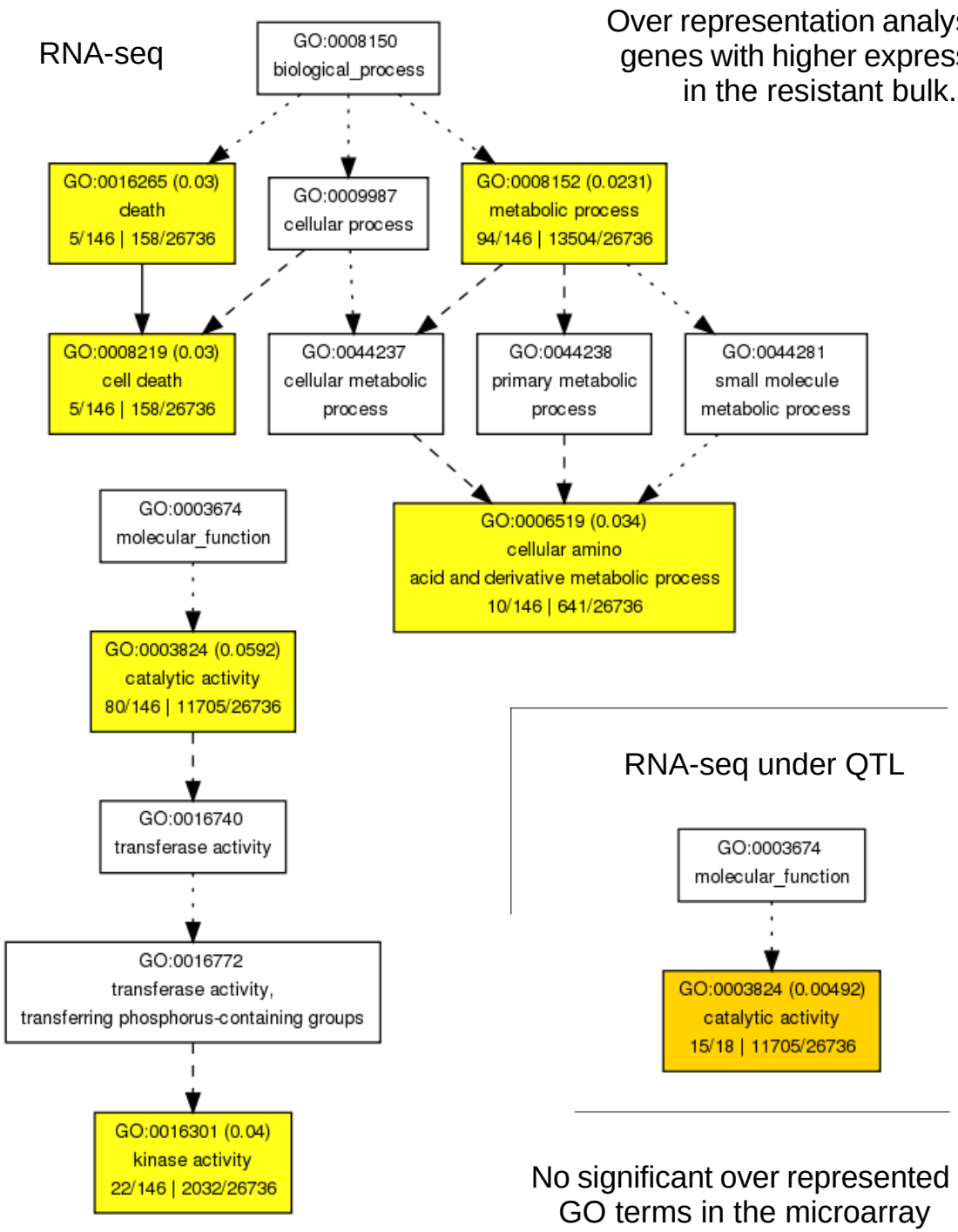


Figure 4.3: Over representation analysis of GO terms from genes with relatively higher expression in the resistant bulk. The significantly over represented GO terms identified from the genes in the RNA-seq analysis and the subset of the RNA-seq identified genes which overlap with QTL are shown in coloured blocks. The FDR is indicated in brackets, and the number of genes associated with the GO terms in the test set is indicated at the bottom of each block, followed by the proportion in the reference set. Solid lines with arrows indicate there are statistically significant over-represented GO terms in both connected boxes, whereas dashed lines indicate one box contains an over-represented GO term, and dotted lines connect boxes without over-represented GO terms.

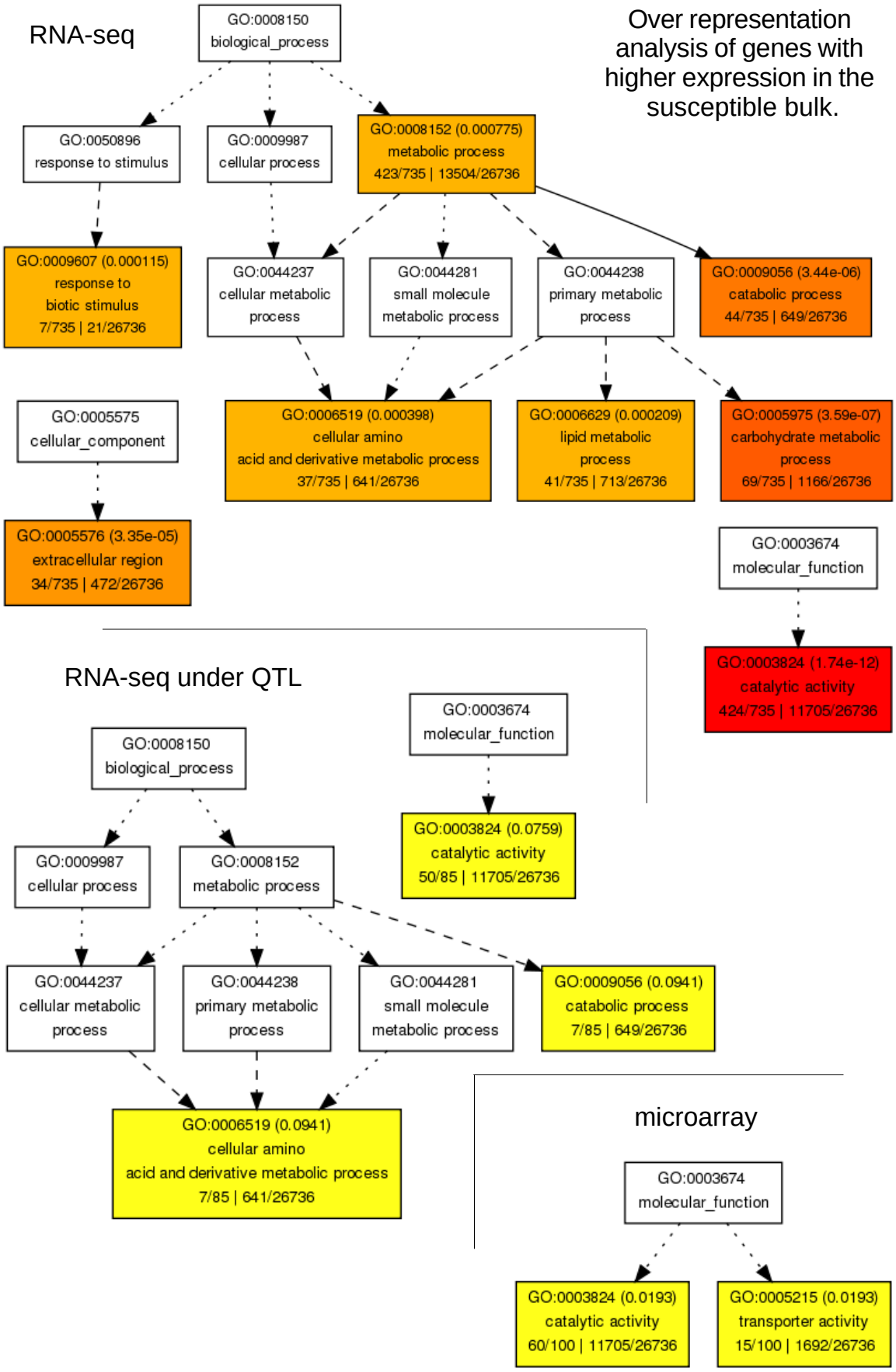


Figure 4.4: Figure legend on next page.

Figure 4.4: Over representation analysis of GO terms from genes with relatively higher expression in the susceptible bulk. The significantly over represented GO terms identified from the genes in the RNA-seq analysis, the subset of the RNA-seq identified genes which overlap with QTL, and the genes from the microarray analysis are shown in coloured blocks. The FDR is indicated in brackets, and the proportion of genes associated with the GO terms in the test set is indicated at the bottom of each block, followed by the proportion in the reference. Solid lines with arrows indicate there are statistically significant over-represented GO terms in both connected boxes, whereas dashed lines indicate one box contains an over-represented GO term, and dotted lines connect boxes without over-represented GO terms.

Looking at genes with lower fold changes than those listed in the top ten (see 100 top genes in Table B3, Appendix B), there are a number of genes of interest for defence such as a cysteine-rich RLK receptor-like kinase (GRMZM2G000633), a sphingosine kinase (GRMZM2G071145), ras-related protein Rab11B (GRMZM2G020661), which would all be involved in effector recognition or signal transduction. Additionally there is an Avirulence Induced Gene (AIG)2-like protein (GRMZM2G102912), a cation efflux protein/zinc transporter (GRMZM2G425594), a putative serine carboxypeptidase (GRMZM2G133718) and an auxin-binding protein 4 precursor (GRMZM2G064371).

The full list of differentially regulated genes from the RNA-seq analysis was entered into MapMan and mapped onto the biotic stress pathway. This gave an overall visual impression of the genes with a putative role in defence (Figure 4.5). It demonstrates that the defence response in the S bulk (red in Figure 4.5) is far greater than in that of the R bulk, similar to what the over representation analysis using AgriGO showed. Although few genes were identified in the R bulk (blue in Figure 4.5), they may provide insight into possible mechanisms of defence, thus the focus was placed on the R bulk.

The two most prominent genes identified by MapMan for defence in the R bulk are a cytoplasmic CC-NB-LRR protein (GRMZM2G074496) listed under “PR-proteins” and powdery-mildew-resistance gene o (*mlo1*) (GRMZM2G032219) listed under “signalling” (Figure 4.5). Additional R bulk genes were highlighted in the outer, light grey, section of Figure 4.5 as putatively involved in biotic stress. Three cell membrane receptor like kinases (RLKs), two with LRR receptors (GRMZM2G039665 and GRMZM2G439799) and one with a malectin receptor (GRMZM2G155837), all had serine-threonine kinase domains. Two GTPase genes (GRMZM2G101460 and GRMZM2G020661) with higher expression in the R bulk were identified to have a possible role in defence via signal transduction. A cell membrane protein

with a gamma-aminobutyric acid (GABA) receptor and a glutamate receptor (GluR) (GRMZM2G148807) was identified for potential signalling. Two lipoxygenase genes were identified in the R bulk, namely *Lox6* (GRMZM2G040095) and *Lox11* (GRMZM2G009479), these are probably found in the chloroplast. A terpene synthase (GRMZM2G049538) was found which would function in the production of terpenoids. Three defence associated transcription factors (GRMZM2G092137, GRMZM2G425920 and GRMZM2G151407) were found to have higher expression in the R bulk. Their target genes remain unknown. An enhanced disease resistance 1 (*EDR1*) gene with a protein tyrosine kinase domain (GRMZM2G044180) is likely to be involved in a MAPK signalling. Two auxin-binding proteins (GRMZM2G078508 and GRMZM2G064371) would be involved in hormone signalling. Two proteins that are likely to play a role in detoxification of the plant cells were also highlighted: a cytochrome b561 (GRMZM2G066885) and a glutathione transferase lambda 2-like protein (GSTL2) (GRMZM2G084369).

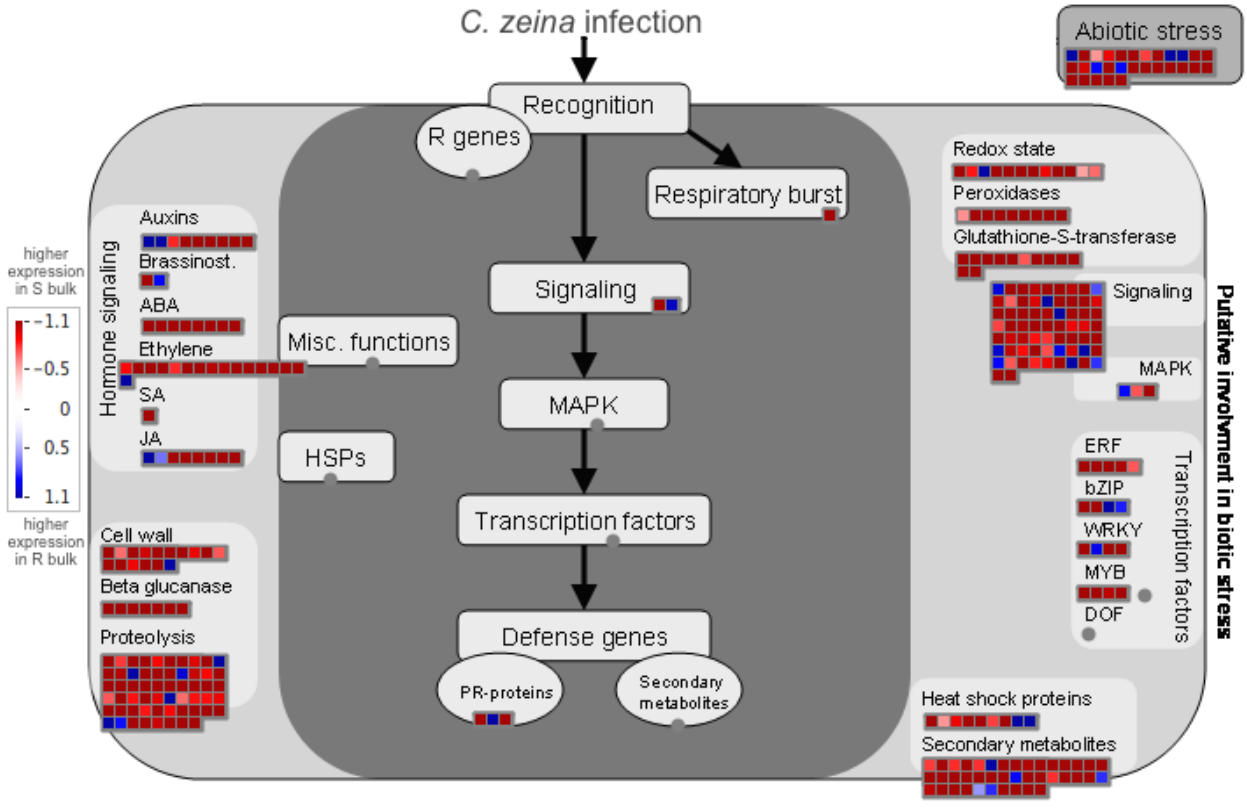


Figure 4.5: MapMan representation of the RNA-seq-identified genes involved in biotic stress. The genes with relatively higher expression in the R bulk are represented by blue, while the genes with relatively higher expression in the S bulk are represented by red. The scale of blue to red is based on the $\log_2$ fold change of the R bulk/S bulk.

There were eight genes that were only detected in the susceptible bulk (FPKM=0 in resistant bulk) (Table 4.4.). They were distributed across chromosomes 1, 3, 5, 6, 7 and 9, and thus not due to a large deletion in a parent of the RIL population. These genes showed no expression in the resistant bulk, all with a relatively low FPKM values in the susceptible bulk ranging from 4.8 to 13 FPKM, the average FPKM for the susceptible bulk was 32 FPKM. The genes include a jacalin-like lectin domain containing protein (GRMZM2G112238), an arabinogalactan protein (GRMZM2G148964), a zinc finger transcription factor (GRMZM2G355846) and a late embryogenesis abundant (LEA) protein (GRMZM2G704475).

The extended list of genes with higher expression in the susceptible bulk contained a large number of defence-related genes. There were a number of genes associated with drought resistance (mostly LEA proteins), cupin domain containing proteins, RLKs and PR proteins (beta 1,3 glucanase, chitinase, endochitinase), see Table B4 in Appendix B.

4.3.4 Candidate genes for QTL

This analysis was carried out to determine whether candidate genes with differential expression between the two bulks could be found for the QTL identified in Berger et al. (2014) and in chapter 3. The physical positions on the B73 RefGen_v2 genome of the genes identified above, from the microarray and RNA-seq analyses, were compared to the positions of the QTL on the physical map as put forward by Christie (2014). The proportion of genes with higher expression in either of the bulks, which are likely to physically fall under the QTL regions, is approximately the same as the proportion of the genome covered by the QTL. The QTL regions under investigation here cover approximately 9.5% of the physical map. Of the genes from RNA-seq analysis with higher expression in the susceptible bulk, 7.7% fall under the QTL investigated. In the resistant bulk, 8.6% fall under the QTL. Thus the differentially expressed genes seem to be distributed fairly evenly over the genome, and not in “hotspots” under the QTL identified in Chapter 3 and in the Baynesfield 2009 QTL (Berger et al. 2014).

The full list of genes from the microarray analysis falling under the QTL is available in Appendix B, Table B5; the genes from the RNA-seq analysis falling under the QTL is available in Appendix B, Table B6. The genes with support from both technologies that coincide with the QTL are listed in Table 4.7. The two genes from the resistant bulk are the chorismate synthase mentioned previously and an ubiquitin-specific protease, neither of which are proteins commonly described in pathogen defence.

Table 4.7: Differentially expressed genes observed in both the microarray and RNA-seq analysis that coincide with QTL.

Gene	QTL	Sequence description	Array Fold change R/S[a]	adj. p-value	RNA-seq Fold change R/S[a]	Q-value	Sequence description source	BLAST/TAIR/Rice Acc No.
GRMZM2G036861	9H_GLS	chorismate synthase 2	1.60	0.0498	9.82	0.0000	BLAST/Rice	NP_001148583/ LOC_Os03g14990.1
GRMZM2G128934	10Cz_2	ubiquitin-specific protease family C19-related protein	1.83	0.0294	2.81	0.0002	TAIR	AT1G78880.1
Gene	QTL	Sequence description	Array Fold change S/R[b]	adj. p-value	RNA-seq Fold change S/R[b]	Q-value	Sequence description source	BLAST/TAIR/Rice Acc No.
GRMZM2G170692	5Cz	phenylalanine ammonia-lyase	2.09	0.0193	4.11	0	BLAST	NM_001174615.1
GRMZM2G170016	5Cz	cytochrome b5-like Heme/Steroid binding domain containing protein	1.79	0.031	2.39	0	Rice	LOC_Os02g42740.1
GRMZM2G001332	6Cz/Lesion	uncharacterized	1.85	0.0316	4.73	0.0021	BLAST	XR_555125.1
GRMZM2G098875	9Cz_2.2	glutamate decarboxylase	1.66	0.04	14.38	0	BLAST/Rice	XM_004985025.1/ LOC_Os03g13300.1
GRMZM2G356618	9Cz_2.2	uncharacterized	2	0.0216	56.22	0	BLAST	XR_559425.1
GRMZM2G046601	9Cz_2.2	glutamine synthetase	2.2	0.0485	8.23	0	Rice	LOC_Os03g12290.1

[a] Fold change of R bulk/S bulk (higher expression in R bulk)

[b] Fold change of S bulk/R bulk (higher expression in S bulk)

Over representation analysis was carried out on the subset of GO terms associated with the genes within the QTL boundaries with relatively higher expression in the R bulk or in the S bulk to determine whether there would be a greater enrichment of GO terms related to biotic stress within the boundaries of the QTL regions. In the list from the genes with higher expression in the R bulk the only significant GO term was catalytic activity in the molecular function subset (Figure 4.3 and Figure B7 in Appendix B for the gene list). In the list from the genes with higher expression in the S bulk the significant GO terms were catalytic activity in the molecular function subset and catabolic process, cellular amino acid and derivative metabolic process in the biological process subset (Figure 4.4 and Figures B8 and B9 in Appendix B for the gene lists).

The genes identified from the MapMan analysis (Figure 4.5) were examined to find those falling within the genomic regions associated with the QTL. Those that do are the malectin RLK gene (QTL 10Cz_2), the gene encoding the chloroplastic GTPase (QTL 1Cz), the *EDR1* gene (QTL 9Cz_2.1) and the cytochrome b561 gene (QTL 5Cz) (Figure 4.5). Two more genes involved in secondary metabolite production are of interest: a coumarate-CoA ligase 4 gene (GRMZM2G055320) situated in the genomic region of 9Cz_2.1 and a leucoanthocyanidin reductase gene (GRMZM2G097854) in 10Cz_2. The ubiquitin-specific protease mentioned above and located in the 10Cz_2 region was also identified by Mapman, as well as a ubiquitin-protein ligase (GRMZM2G152919) in 9Cz_2.2/9H_GLS.

4.3.5 Microarray and RNA-seq comparison

The microarray and RNA-seq analysis were compared in an attempt to validate candidate genes. The overlapping genes of the two analyses were compared. The Spearman rank correlation of the overlapping genes of the microarray and RNA-seq analysis is 0.34 when considering both the fold changes in both R/S and S/R with a value above 1.5 and an FDR of < 5%. If all genes with fold changes above 1.5, but ignoring the FDR, are taken into account the correlation value rises to 0.60. This weak correlation is partially due to the large number of genes with opposite expression ratios for the two technologies, i.e. negative $\log_2$R/S values in the RNA-seq analysis whereas the microarray shows positive $\log_2$R/S values, as can be seen in Figure 4.6. These correlations are weaker than those found in comparisons of microarray and RNA-seq by Marioni et al. (2008) and Kogenaru et al. (2012), but similar to that found by ’t Hoen et al. (2008). The increase in correlation coefficient when genes with non-statistically significant fold changes are included is the inverse to that found by Kogenaru et al. (2012).

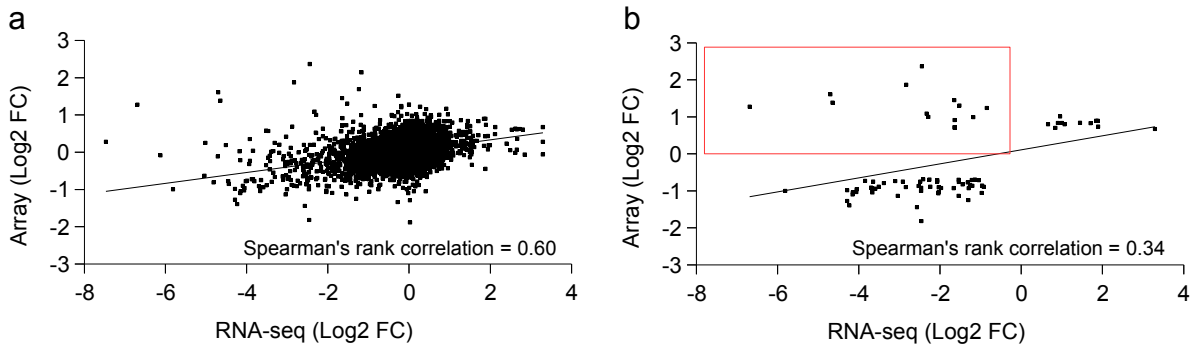


Figure 4.6: Scatter plots of the relative expressions measured in $\log_2$ fold change (FC) of the R bulk/S bulk compared between the RNA-seq (x-axis) and the microarray (y-axis) analysis. In (a) all detected genes are shown, while in (b) only statistically significant genes, with a FDR cut off of 5%, are shown. The upper left quadrant of (b) is boxed in red highlighting the genes which have positive $\log_2$R/S ratio in the microarray but a negative $\log_2$R/S ratio in the RNA-seq analysis.

The overlap of GO terms associated with the 145 genes from the R bulk from the microarray and the 278 genes from the RNA-seq was investigated, as well as the overlap of GO terms from the S bulk (151 genes from the array, and 1027 genes from the RNA-seq analysis). There is a relatively large overlap between the sets of GO terms from the microarray and RNA-seq for both the R bulk and the S bulk (Figure 4.7). This suggests that similar pathways are being detected by both technologies. The main overlap of GO terms between the resistant and susceptible bulks is in catalytic activity as well as cellular amino acid and derivative metabolic processes. The differences in the pathways between the bulks, which is highlighted by the GO terms, lies in death/cell death in the resistant bulk (Figure 4.3) and the response to biotic stimulus in the susceptible bulk (Figure 4.4).

The differentially expressed genes observed in the microarray and RNA-seq analyses were compared. The overlap of genes with higher expression in the R bulk totalled 12 genes, which is less than 10% of the genes observed in the microarray analysis (Figure 4.8). The overlap was higher for the genes with higher expression in the S bulk with 47 genes detected by both technologies (31% of the genes observed in the microarray analysis supported) as seen in Figure 4.9. The percentages of genes in the RNA-seq analysis that are supported by the microarray remain approximately the same for the resistant and susceptible bulks at 4% and 5% respectively.

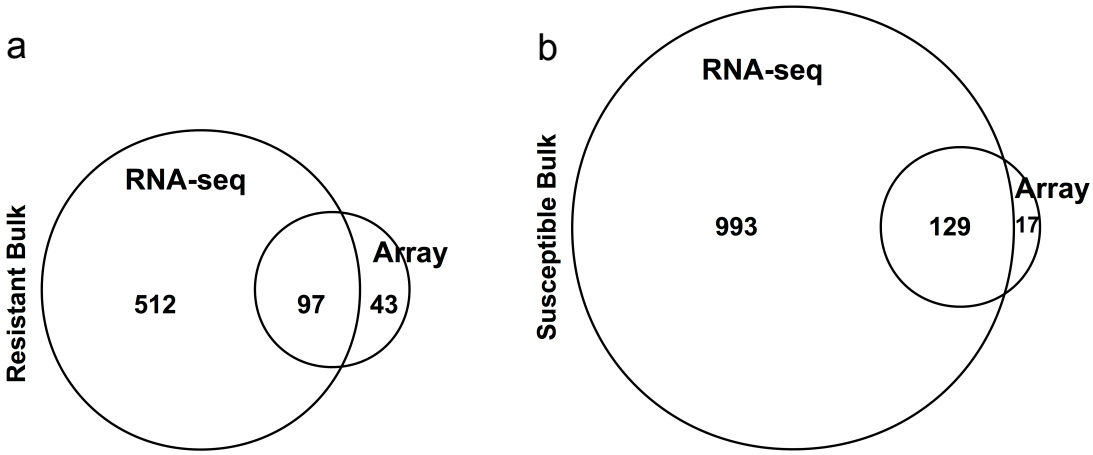


Figure 4.7: Venn diagram showing overlap of GO terms (including all biological process, molecular function, and cellular component GO terms) associated with the genes with higher expression in (a) the resistant bulk (FDR > 5%; FC R/S ≥1.5) and (b) the susceptible bulk (FDR > 5%; FC S/R ≥ 1.5) from the RNA-seq and microarray analyses.

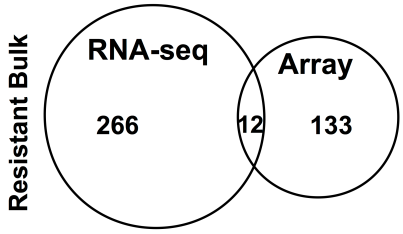


Figure 4.8: Venn diagram showing the numbers of genes observed with higher expression in the resistant bulk (FDR < 5%; FC R/S ≥ 1.5) for the different technologies and their overlap. There were 278 (266+12) genes found to have higher expression by the RNA-seq analysis and 145 (133+12) genes found to have higher expression by the microarray analysis, with an overlap of only 12 genes.

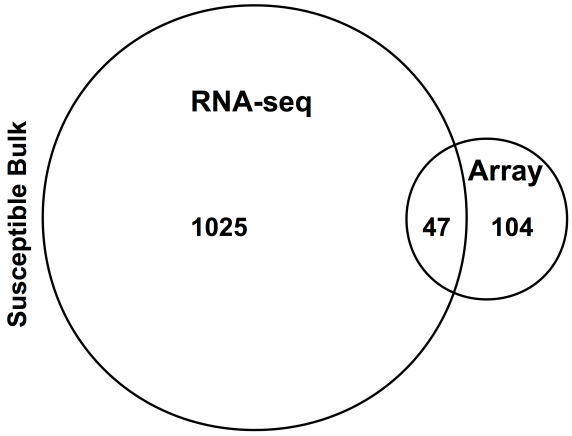


Figure 4.9: Venn diagram showing the numbers of genes observed with higher expression in the susceptible bulk (FDR < 5 %; FC S/R ≥ 1.5) for the different technologies and their overlap. There were 1072 (1025+47) genes found to have higher expression by the RNA-seq analysis and 151 (104+47) genes found to have higher expression by the microarray analysis, with an overlap of 47 genes.

The genes found to have higher expression in the resistant bulk from the microarray were not strongly supported by the RNA-seq results. Half the genes had fold changes in the opposite direction (Figure 4.6b, the data points in the upper left quadrant). There was more agreement between the two technologies on fold change direction in the susceptible bulk. The most notable gene with corroboration between technologies is a germin-like protein or a cupin domain containing protein (GRMZM2G071390; Tables 4.3 and B4 in Appendix B). These proteins have a role in stress response and development amongst other diverse functions (Bernier & Berna 2001; Dunwell et al. 2004; Breen & Bellgard 2010).

None of the genes with higher relative expression in the top ten of the resistant bulk from the microarray are supported by statistically significant genes from the RNA-seq analysis. Five of the genes have a contradictory lower expression in the resistant bulk from the RNA-seq (GRMZM2G373522, GRMZM2G055698, GRMZM2G103771, GRMZM2G117164, and GRMZM2G154747), see Tables 4.2 and B2. The phenomenon of fold changes in opposite direction in microarray and RNA-seq-based technologies was also observed by ’t Hoen et al. (2008). Five of the genes with higher expression in the susceptible bulk are, however, supported by RNA-seq data (GRMZM2G427815, GRMZM2G050450, GRMZM2G071390, GRMZM2G001572, and GRMZM2G136508), see Tables 4.3 and B2. Only one out of the top ten differentially expressed genes from the RNA-seq analysis were supported by microarray data, the chorismate synthase (GRMZM2G036861, Tables 4.4 and B5a, falls out of the top 100 for the microarray with fold change of 1.6, thus not in table B1).

Out of the eight genes detected only in the susceptible bulk in the RNA-seq (Table 4.6), only GRMZM2G112238 (the jacalin-like lectin domain containing protein) was detected in the microarray analysis, but it was shown to have higher expression in the resistant bulk of the microarray with a fold change of 2.5 (adj P = 0.021) ranked 12th highest fold change.

4.4 Discussion

In this chapter we aimed to find differentially expressed genes between resistant and susceptible maize RILs infected with GLS, using the available RNA-seq results and the microarray results to co-validate each other. The main aim was not to compare the technologies of microarray and RNA-seq as such comparisons have been done (Marioni et al. 2008; 't Hoen et al. 2008; Malone and Oliver 2011), but the comparison was unavoidable in the process. We also endeavoured to suggest candidate genes underlying the QTL in chapter 3 and Berger et al. (2014) by associating the physical position of the genes with the QTL.

Although the overlapping GO terms suggested many similar pathways from the RNA-seq and microarray data, few specific genes were in fact validated as differentially expressed between the R bulk and S bulk. The microarray probes for the Agilent 4 × 44K array were designed from EST data from B73 and other maize varieties prior to the B73 genome sequence (Coetzer et al., 2011). Maize is known to have a high level of polymorphism (Tenaillon et al. 2001). Kirst et al. (2006) showed the effect of maize lines and their high level of polymorphism on probe binding and signal. The sequences from RIL bulks are derived from CML444 or SC Malawi and may not be exact matches to the probes, resulting in differences in binding efficiencies for the different sequences. As the RILs are inbred they are mostly homozygous and contain either the CML444 allele or the SC Malawi allele at each genomic region. If one of the parental alleles has higher homology to the microarray probe it will bind with higher efficiency than the other allele present in the population. This may explain the false impressions of higher or lower expression. The differences in sensitivity of the technologies may also contribute to the failure to validate some genes. Sequencing depth of RNA-seq is far greater than microarrays, thus RNA-seq is more sensitive and can detect lower abundant genes (Marioni et al. 2008; 't Hoen et al. 2008). Agilent 60-mer microarray probes require a 23 nucleotide length in order for 90% of the probes to bind with their matching transcripts (Poulsen et al. 2008). Thus, if a probe and transcript do not have a matching sequence longer than 23 nucleotides the transcript will likely fail to be detected. The mapping of the RNA-seq data to the genome allows for mismatches between the reads and the genome, numerous and overlapping reads give confidence to the RNA-seq transcript sequence. The transcript sequence comprises mapped reads over the whole or most of the gene, thus it is more likely to be detected than by one or two microarray probes if there are sequence differences between lines. For these reasons the RNA-seq was relied on more heavily for selection of candidate genes.

Microarray is better suited to comparing different treatments, or organs of the same genotype as in Marioni et al. (2008), rather than comparing different genotypes due to the possible sequence differences in the transcript target sequences. As long as the transcript sequence differences are not too diverse, RNA-seq can be used to compare transcript levels in different genotypes. RNA-seq also takes into consideration splice variants (Trapnell et al. 2010, 2012), but this does not occur in a microarray experiment (Marioni et al. 2008) possibly also confounding gene validation. RNA-seq can also be used to identify candidate genes with sequence differences such as SNPs between the two alleles (Liu et al. 2012) as there is not necessarily an expression difference, but a functional difference due to sequence is also possible .

Two bulks of six RILs each were compared in order to find gene expression differences between resistant and susceptible lines, and found a number of differentially expressed genes under each QTL. Kloosterman et al. (2010) suggest using multiple bulks, i.e. more than one resistant and susceptible bulk, each comprising a nuber of different resistant or susceptible RILs. This will reduce the number of differentially expressed genes that are not associated with the trait of interest.

The two genes present in the overlap of the microarray and RNA-seq with higher expression in the R bulk, and falling within QTL regions are a chorismate synthase in QTL 9H_GLS and a ubiquitin-specific protease in 10Cz_2 (Table 4.4). Chorismate synthase is involved in production of chorismate leading to a number of different pathways, a few of which are involved in pathogen defence. Chorismate synthase produces chorismate from 5-enolpyruvylshikimate 3-phosphate; the chorismate feeds into various pathways to produce the aromatic amino acids (Phenylalanine, Tyrosine and Tryptophan) and salicylic acid. Phenylalanine can be further modified to produce flavonoids and anthocyanins, and is involved in lignin biosynthesis (Maeda and Dudareva 2012). Tryptophan is a precursor of camalexin (Glawischnig et al. 2004), a phytoalexin involved in plant defence (Glazebrook and Ausubel 1994). Salicylic acid is an important signalling molecule in the hypersensitive response and systemic acquired resistance (Hammond-Kosack and Parker 2003). Chorismate produced by chorismate synthase is converted to isochorismate by isochorismate synthase, which is in turn converted to salicylic acid by pyruvate lyase. Isochorismate is required for production of salicylic acid for systemic acquired resistance (Wildermuth et al. 2001). There was no evidence of increased expression of any isochorismate genes in the gene lists, but the upstream production of chorismate synthase under *C. zeina* infection may ensure that chorismate supply

is not a rate-limiting factor in the pathway. Chorismate synthase has been shown by gene silencing to prevent penetration of a biotrophic powdery mildew, *Blumeria graminis*, into Barley (Hu et al. 2009). The ubiquitin-specific protease is involved in deubiquitination of proteins. Whereas ubiquitination plays a large role in pathogen defence in plants (Marino et al. 2012), deubiquitination has not been well characterised in plant defence. However, Ewan et al. (2011) have shown that an ubiquitin-specific protease in *Arabidopsis* negatively regulates the hypersensitive response to a *Pseudomonas* pathogen.

The above genes are comparable to cis-eQTL, where the genomic element controlling the expressed gene is located in the same genomic region. Trans eQTL occur where the genomic element controlling the expressed gene is not located in the same region (e.g. a trascription factor, Hansen et al. 2008; Kliebenstein 2009). This means that not all genes related to QTL have to co-localise with the QTL, therefore the entire RNA-seq top table was investigated for candidate genes involved in *C. zeina* resistance.

In the RNA-seq top table of the resistant bulk there is a GASA/Snakin family protein which may play a role in plant defence through hormone signalling (Nahirñak et al. 2012) and has antibacterial and antifungal activity (Segura et al. 1999). Such proteins have also been shown to be up-regulated after pathogen infection of potato (Berrocal-Lobo et al. 2002). Loricrin has been nominated as a candidate gene for *Fusarium verticillioides* resistance in maize in a genome wide association study (Zila et al. 2014). The chorismate synthase gene discussed above is present in this top table as well.

The extended list of genes with higher expression in the R bulk from the RNA-seq analysis, after being mapped onto the biotic stress pathway in MapMan (Figure 4.5), was compiled into a model of possible mechanisms involved in resistance against *C. zeina* (Figure 4.10). A number of proteins are putatively involved in plant defence signalling and these will be discussed in the context of the proposed model of resistance in the following sections.

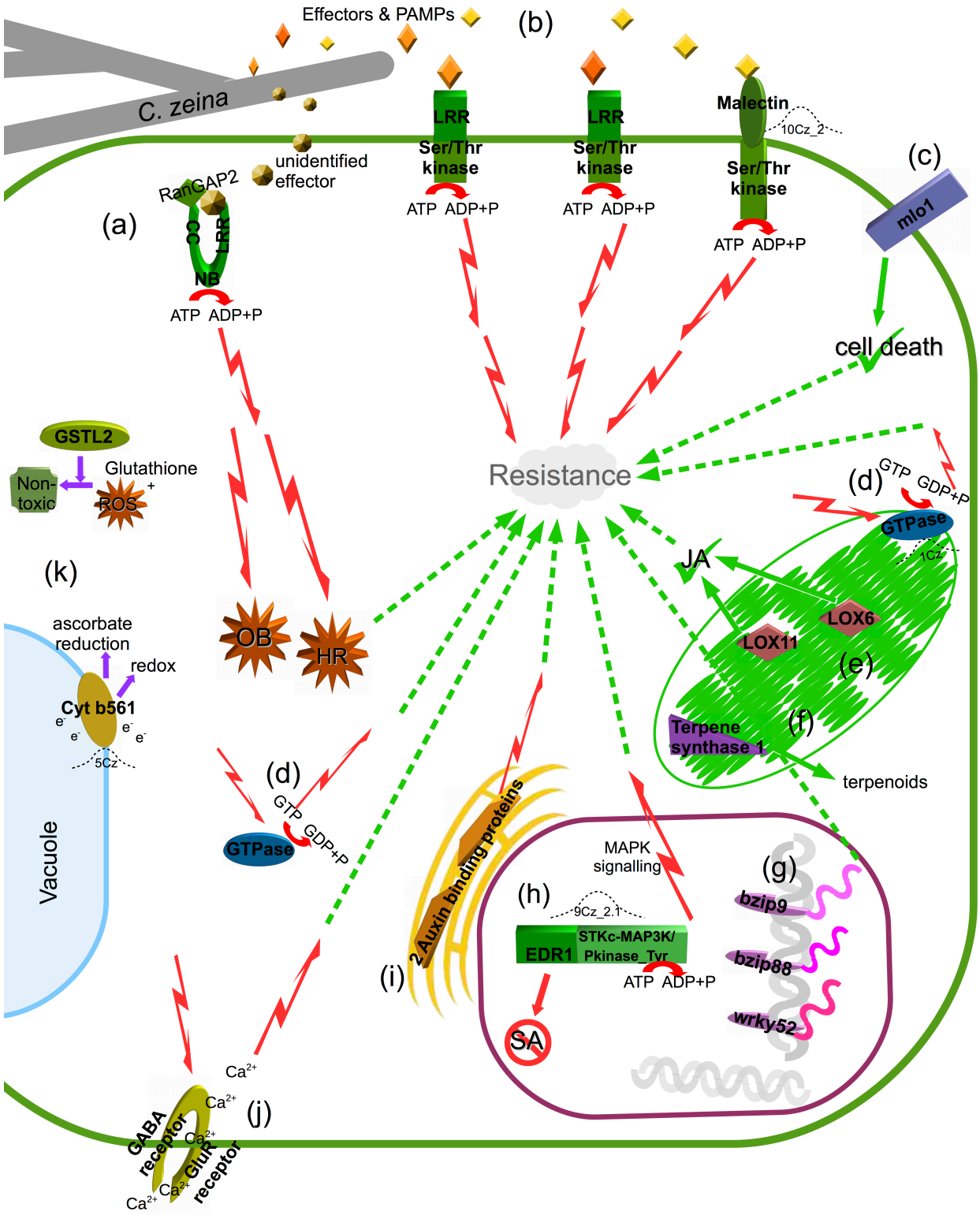


Figure 4.10: The proposed mechanism(s) of resistance based on the genes found to have higher expression in the resistant bulk compared to the susceptible bulk, based on RNA-seq analysis. (a) A cytoplasmic CC-NB-LRR protein (GRMZM2G074496), with homology to the *Arabidopsis* disease resistance protein RPP13, likely recognises an unidentified fungal effector in collaboration with an unidentified RanGAP2-like protein from maize. This RPP13-like protein is likely to start a signalling cascade (red jagged arrows) resulting in oxidative burst and a hypersensitive response. (b) Three cell

membrane RLKs, two with LRR receptors (GRMZM2G039665 and GRMZM2G439799) and one with a malectin receptor (GRMZM2G155837) and all with serine-threonine kinase domains, are possibly involved in pathogen detection and signalling. The malectin RLK gene is located within the bounds of QTL 10Cz_2. (c) A barley *mlo* defence gene homologue (GRMZM2G032219) was found to have higher expression in the R bulk and an *Mlo*-like gene (not shown) had lower expression in the R bulk. Expression of the *Mlo* gene leads to inhibition of cell death, whereas expression of the mutant or *mlo* gene prevents inhibition of cell death. (d) Two GTPase genes with higher expression in the R bulk were identified to have a possible role in defence via signal transduction. The one GTPase (GRMZM2G020661) is likely to be located in the cytoplasm with the other (GRMZM2G101460) in the chloroplast. The gene encoding the chloroplastic GTPase is located within the bounds of QTL 1Cz. (e) Two lipoxygenase genes were identified. The LOX6 (GRMZM2G040095) and LOX11 (GRMZM2G009479) proteins are probably found in the chloroplast. The lipoxygenases are found downstream of JA production. (f) A terpene synthase (GRMZM2G049538) was identified which would function in the production of terpenoids. (g) Three defence associated transcription factors (GRMZM2G092137, GRMZM2G425920, and GRMZM2G151407) were found to have higher expression in the R bulk. (h) An *EDR1* gene (GRMZM2G044180) with a protein tyrosine kinase domain is likely to be involved in a MAPK signalling cascade. EDR1 is also involved in inhibition of the salicylic acid pathway. This gene is situated within the bounds of QTL 9Cz_2.1. (i) Two auxin-binding proteins (GRMZM2G078508 and GRMZM2G064371), which would be involved in hormone signalling, are located in the endoplasmic reticulum. (j) A cell membrane protein (GRMZM2G148807) with a gamma-aminobutyric acid (GABA) receptor and a glutamate receptor (GluR) is involved in calcium homeostasis. The protein is also likely to be involved in signal transduction. (k) A cytochrome b561 (GRMZM2G066885) is involved in detoxification by reduction of ascorbate and electron transport; the gene is located within the bounds of QTL 5Cz. A glutathione transferase lambda 2-like protein (GSTL2) (GRMZM2G084369) is also involved in detoxification by conversion of ROS to non-toxic compounds by the addition of glutathione.

The resistance to *Peronospora parasitica* 13 (RPP13)-like protein (Figure 4.10a) is homologous to a known resistance protein in *Arabidopsis* with the CC and the LRR regions involved in recognition of a fungal Avr protein (Bittner-Eddy et al. 2000; Rose et al. 2004), most likely with the help of another unidentified maize protein which interacts with the CC region. The *Arabidopsis* RPP13 likely recognises an unidentified fungal effector in collaboration with a RanGAP2 protein (Sacco et al. 2007; Tameling et al. 2010), there is probably a similar, but unidentified protein in maize. RanGAP2 fulfils the function of a cofactor needed to help bind a viral Avr protein in the *Nicotiana benthamiana* Potato virus X pathosystem (Tameling et al. 2010; Hao et al. 2013) this results in signalling resulting in a hypersensitive response.

These proteins have been suggested to have a "bait and switch"-like mechanism where RanGAP2 binds to the CC domain of the RPP13-like protein and possibly functions as a "bait" protein or cofactor in the recognition of the (yet unknown) effector from the pathogen. The RanGAP2 "bait" brings the effector into contact with the LRR domain, the "switch", resulting in conformational changes and activation of the NB domain leading to signal transduction (Collier and Moffett 2009; Tameling et al. 2010). The NB-ARC domain is a signalling motif found in plant resistance proteins as well as in animal proteins that regulate cell death (van der Biezen and Jones 1998). This domain would be involved in an immune signalling response leading to oxidative burst and a hypersensitive response, which is usually more effective against biotrophic pathogens. According to the proposed model of resistance to GLS in this maize population, the *RPP13-like* maize gene may function as a minor R gene in amongst the QTL observed in the RIL population. The gene itself does not coincide with any QTL, but it may have a transcription factor allowing for its increased levels in the R bulk, which coincides with a QTL. Alternatively, the unidentified *RanGAP2-like* gene may be the gene associated with a QTL.

Maize *mlo1*, a barley *mlo* defence gene homologue (Figure 4.10c) had higher expression than an *Mlo*-like gene in the resistant plants. Wildtype *Mlo* is associated with susceptibility due to inhibition of cell death, whereas the mutant *mlo* does not inhibit cell death (Büschges et al. 1997). The functioning of the MLO protein may be due to Ca^{2+} dependent calmodulin binding (Kim et al. 2002). Higher expression of the *mlo* gene may facilitate a hypersensitive response by reducing the cell death inhibition mechanisms. Another cell membrane protein with a GABA receptor and GluR regions (Figure 4.10j) most likely plays a role in Ca^{2+} influx. Calcium influx can lead to cell death and can be found downstream of oxidative burst (Levine et al. 1996)

In the suggested model of resistance, RLKs in the cell membrane with serine-threonine kinase domains (Figure 4.10b) may be involved in pathogen recognition and signal transduction. Malectin RLKs have been demonstrated to be involved in resistance to *Pseudomonas syringae* bacteria and the downy mildew causing oomycete, *Hyaloperonospora arabidopsidis* (Hok et al. 2011; Chen et al. 2014) . The malectin receptor binds carbohydrate motifs leading to pathogen-triggered immunity. LRR domains tend to bind proteins or peptides and induce signalling via serine-threonine kinase domains (Böhm et al. 2014). The EDR1 protein (Figure 4.10h) is likely to be involved in a MAPK signalling cascade leading to cell death and is also involved in inhibition of the salicylic acid pathway (Frye et al. 2001). The GTPases (Figure 4.10d) are

involved in signalling. Through signalling they may regulate H_2O_2 or other ROS production, or they may regulate abscisic acid (ABA) and other hormones (Yang 2002). GTPases have been shown to be important in callose deposition and papillae formation, which prevents pathogen penetration into cells (Böhlenius et al. 2010; Ellinger et al. 2014). A cell membrane protein with a gamma-aminobutyric acid (GABA) receptor and a glutamate receptor (GluR) (Figure 4.10j) is involved in calcium homeostasis (Bouché and Fromm 2004). The protein is also likely to be involved in signal transduction.

Two *LOX* genes (Figure 4.10e), found downstream of jasmonic acid production (Vick and Zimmerman 1984), a hormone involved in defence against necrotrophs, are possibly involved in hormone signalling in the resistance response. There are also auxin-binding proteins (Figure 4.10i) which may be involved in hormone signalling. The hormone pathways suggested by the genes upregulated in the resistant plants are not entirely clear, but tend to suggest the Jasmonic acid pathway rather than the salicylic acid pathway due to the presence of the *LOX* genes (Figure 4.10e) and *EDR1* (Figure 4.10h) which is known to inhibit the salicylic acid pathway (Frye et al. 2001).

Three defence-associated transcription factors one WRKY and two bZIP transcription factors (Figure 4.10g) were found to have higher expression in the R bulk (Figure 4.6 g). Their target genes remain unknown, but are likely to be genes involved in defence (Buscaill and Rivas 2014).

A terpene synthase (Figure 4.10f) which produces terpenoids, may play a role in defence against *C. zeina*, but terpenoids are usually associated with defence against herbivores (Tholl 2006). The resistant plants leaves will be greener and less stressed, thus may be a more enticing to herbivores which would induce additional wounding responses confounding the elucidation of the resistance response to *C. zeina*.

The ROS produced by the plant cells can be toxic for the plant cells as well as the pathogen's cells. Thus the production of proteins that can detoxify the cells can prevent damage from the plant's own defence mechanisms or from the pathogen's attack mechanisms. The cytochrome b561 (Figure 4.10k), located at the tonoplast (Griesen et al. 2004), is involved in detoxification by reduction of ascorbate and electron transport (Asard et al. 2001). And GSTL2 (Figure 4.10k) is also involved in detoxification by conversion of ROS to non-toxic compounds by the addition

of glutathione, this is a different GST protein to that previously found to be involved in resistance to *C. zeina* (Wisser et al. 2011).

The biological process GO terms “death” and “cell death” were over-represented in the R bulk and not the S bulk. The cell death GO term and the candidate defence genes linked to cell death or hypersensitive response confirms the importance of cell death in the resistance response. The biological process GO term “ biotic stress” was over-represented in the S bulk and not the R bulk, probably due to the large defence response of the susceptible plants due to the large amount of fungus present in their leaves.

Moving onto the susceptible response, the genes only expressed in the susceptible bulk (Table 4.3) are distributed over different chromosomes, so they do not represent a deletion of a number of genes in the more resistant plants. A jacalin-like lectin domain containing protein is involved in carbohydrate binding, salicylic acid and jasmonic acid defence signalling pathways, and is also induced by hemi- and biotrophic fungi (Xiang et al. 2011). Although this gene was identified as only being expressed in the susceptible bulk via RNA-seq, it was also observed as being expressed in the resistant bulk in the microarray experiment. The sequence of the gene and microarray probe should be compared before any future investigation of the gene. An arabinogalactan protein is involved in signalling and programmed cell death (Gao and Showalter 1999), and is found in found in papillae formed as a barrier to pathogen entry into cells (Micali et al. 2011; Underwood 2012). It may also act as an integrity sensor or as reinforcement for the cell membrane (Humphrey et al. 2007). A zinc finger family protein would possibly participate in transcription regulation via DNA binding (Takatsuji 1998). An observed LEA protein would be involved in osmotic stress protection (Amara et al. 2014), possibly in response to desiccation of lesion areas.

In the RNA-seq top table of the susceptible bulk there are three LEA group 3 proteins which are usually produced in response to drought or salt stress and is regulated by ABA (Amara et al. 2014). The nine-cis-epoxycarotenoid dioxygenase5 plays a role in ABA production (Frey et al. 2012). A UDP-Glycosyltransferase superfamily protein is also listed. Two UDP-Glycosyltransferases have been implicated in *Arabidopsis* defence against *Pseudomonas syringae*, and influencing salicylic acid and jasmonic acid signalling with different effects on resistance (Boachon et al. 2014). Additionally, two cupin domain containing proteins are also

present in the RNA-seq top table of the susceptible bulk, these could have a wide variety of functions (Dunwell et al. 2004).

Germin-like proteins fall into the diverse cupin superfamily (Dunwell et al. 2004). They have often been characterised as oxalate oxidases that produces carbon dioxide and hydrogen peroxide, and thus may be involved in ROS pathways, some can function as superoxide dismutase which is involved in inactivating superoxide radicals found under stress conditions (Bernier and Berna 2001; Dunwell et al. 2004; Breen and Bellgard 2010). A germin-like oxalate oxidase protein in Barley was shown to be induced by powdery mildew infection (Dumas et al. 1995; Zhang et al. 1995). The germin-like protein was observed in the microarray susceptible bulk top table (Table 4.1), and thus will not confer resistance. It is most likely upregulated in response to the onslaught of the pathogen.

There are a number of pathogenesis related (PR) proteins which coincide with the QTL, but with higher expression in the susceptible bulk. Namely, thaumatin (PR5) from the microarray, three beta 1,3 glucanases (PR 2), chitinase (PR3) and an endochitinase (PR4) (Appendix B, Tables B1 and B2) (Sels et al. 2008). The inability of the higher expression of genes known for pathogen resistance in the S bulk to confer resistance could be due to timing of the gene expression, the gene sequence, the genes expressed by the fungus, e.g. *Avr* genes, rendering the plant's response ineffective. For example, chitinases have higher expression in the S bulk than in the R bulk (Apendix B, Table B2), the fungus may sidestep this host strategy by production of the Avr4 protein (chitin binding lectin) (van den Burg et al. 2006). Any actual strategies cannot be confirmed here as the fungal genes in the experimental setup have not been assayed.

The enriched GO terms indicate that cell death, possibly programmed cell death involved in the hypersensitive response, may be implicated in the resistance response, rather than a large biotic stress response as seen in the susceptible bulk. The hypersensitive response is usually associated with resistance to biotrophic pathogens (Hammond-Kosack and Parker 2003; Beckers and Spoel 2006). Such a response can occur within a few hours of pathogen invation (Morel and Dangl 1997). *C. zeina* exhibits intercellular growth after entering the plant before the life cycle becomes necrotrophic (Kim et al. 2011), thus it could be considered hemi-biotrophic and an immediate hypersensitive response followed by systemic acquired resistance may prove to be a good strategy for the plant.

The time at which sampling was carried out (i.e. when lesions had developed on most plants and were spreading) was best for observing genes expressed in the susceptible response. The S bulk did have a larger number of GO terms associated with it overall, as well as a larger number of genes with higher expression levels, especially genes associated with biotic stress, as seen in Figures 4.5. The genes with higher expression levels in the R bulk are likely to have constitutionally higher expression levels. This may be investigated in future by incorporating an expression profiling experiment including disease-free samples (e.g. fungicide sprayed field plants, or non-infected versus inoculated glasshouse grown plants). Genes in the R bulk, involved in induced resistance, would be up-regulated at the beginning infection stages. Such genes may not have been observed here because the early resistance response would have already been successful and been terminated in the R bulk, in contrast to S bulk plants which are responding to the fungal proliferation in lesions at the time of sampling. The *C. zeina* field infection is not a synchronised infection, in spite of old infected leaf material added in the plants' whorls creating a starting time point, but there are rather continual infection events happening over time. The genes involved in resistance are probably only expressed in the small areas of new infection sites and the ability to detect those genes would be diluted out by the uninfected tissue not responding to the pathogen. As the lesion areas, and thus pathogen biomass, were larger on the leaf samples of the S bulk, defence genes would therefore be more prevalent in the sample, see Figure 4.2. Most of the susceptible response is likely to occur at the lesion edges.

To conclude, the more resistant RILs may be overcoming the fungus during the early stages of infection. Some genes involved in the resistance response may not have been captured due to the late sampling time point and the fact that a field infection is a continual process. Some genes do, however, have good support as resistance candidates, such as the *RPP13-like* gene, the *mlo* gene, the LRR RLKs and the malectin containing RLK, as well as the *EDR1* gene (Figure 4.10). These are likely to exhibit higher constitutive expression in the resistant plants than in the susceptible plants. There are many genes involved in pathogen defence with higher expression in the susceptible plants. The susceptible RILs have to continue to fight the fungus from every angle while the necrotic lesions continue to grow over time.

4.5 References

Amara, I., Zaidi, I., Masmoudi, K., Ludevid, M. D., Pagès, M., Goday, A., & Brini, F. (2014). Insights into late embryogenesis abundant (LEA) proteins in plants: from structure to the functions. *American Journal of Plant Sciences*, *5*, 3440–3455.

Asard, H., Kapila, J., Verelst, W., & Bérczi, A. (2001). Higher-plant plasma membrane cytochrome b561: a protein in search of a function. *Protoplasma*, *217*, 77–93.

Beckers, G. J. M., & Spoel, S. H. (2006). Fine-tuning plant defence signalling: salicylate versus jasmonate. *Plant Biology*, *8*, 1–10.

Berger, D. K., Carstens, M., Korsman, J. N., Middleton, F., Kloppers, F. J., Tongoona, P., & Myburg, A. A. (2014). Mapping QTL conferring resistance in maize to gray leaf spot disease caused by *Cercospora zeina*. *BMC Genetics*, *15*, 60.

Bernier, F., & Berna, A. (2001). Germins and germin-like proteins: plant do-all proteins. But what do they do exactly? *Plant Physiology and Biochemistry*, *39*, 545–554.

Berrocal-Lobo, M., Segura, A., Moreno, M., López, G., García-Olmedo, F., & Molina, A. (2002). Snakin-2, an antimicrobial peptide from potato whose gene is locally induced by wounding and responds to pathogen infection. *Plant Physiology*, *128*, 951–961.

Bittner-Eddy, P. D., Crute, I. R., Holub, E. B., & Beynon, J. L. (2000). *RPP13* is a simple locus in *Arabidopsis thaliana* for alleles that specify downy mildew resistance to different avirulence determinants in *Peronospora parasitica*. *Plant Journal*, *21*, 177–188.

Boachon, B., Gamir, J., Pastor, V., Erb, M., Dean, J. V., Flors, V., & Mauch-Mani, B. (2014). Role of two UDP-Glycosyltransferases from the L group of *Arabidopsis* in resistance against *Pseudomonas syringae*. *European Journal of Plant Pathology*, *139*, 707–720.

Böhlenius, H., Mørch, S. M., Godfrey, D., Nielsen, M. E., & Thordal-Christensen, H. (2010). The multivesicular body-localized GTPase ARFA1b/1c is important for callose deposition and ROR2 syntaxin-dependent preinvasive basal defense in barley. *The Plant Cell*, *22*, 3831–3844.

Böhm, H., Albert, I., Fan, L., Reinhard, A., & Nürnberger, T. (2014). Immune receptor complexes at the plant cell surface. *Current Opinion in Plant Biology*, *20*, 47–54.

Bouché, N., & Fromm, H. (2004). GABA in plants: just a metabolite? *Trends in Plant Science*, *9*, 110–115.

Brauer, M. J., Christianson, C. M., Pai, D. A., & Dunham, M. J. (2006). Mapping novel traits by array-assisted bulk segregant analysis in *Saccharomyces cerevisiae*. *Genetics*, *173*, 1813–1816.

Breen, J., & Bellgard, M. (2010). Germin-like proteins (GLPs) in cereal genomes: gene clustering and dynamic roles in plant defence. *Functional and Integrative Genomics*, *10*, 463–476.

Buscaill, P., & Rivas, S. (2014). Transcriptional control of plant defence responses. *Current Opinion in Plant Biology*, *20*, 35–46.

Büschges, R., Hollricher, K., Panstruga, R., Simons, G., Wolter, M., Frijters, A, van Daelen, R., van der Lee, T., Diergaarde, P., Groenendijk, J., Töpsch, S., Vos, P., Salamini, F., & Schulze-Lefert, P.

(1997). The barley *Mlo* gene: a novel control element of plant pathogen resistance. *Cell*, *88*, 695–705.

Chen, C., Panzeri, D., Yeh, Y., Kadota, Y., Huang, P., Tao, C., Roux, M., Chien, S., Chin, T., Chu, P., Zipfel, C., & Zimmerli, L. (2014). The *Arabidopsis* malectin-like leucine-rich repeat receptor-like kinase IOS1 associates with the pattern recognition receptors FLS2 and EFR and is critical for priming of pattern-triggered immunity. *The Plant Cell*, *26*, 1–19.

Chou, K. C., & Shen, H. Bin. (2010). Plant-mPLoc: a top-down strategy to augment the power for predicting plant protein subcellular localization. *PLoS ONE*, *5*, e11335.

Christie, N. (2014). *Transcriptional regulation underlying the quantitative genetic response of maize to grey leaf spot disease*. University of Pretoria.

Coetzer, N., Myburg, A. A, & Berger, D. K. (2011). Maize microarray annotation database. *Plant Methods*, *7*, 31.

Collier, S. M., & Moffett, P. (2009). NB-LRRs work a "bait and switch" on pathogens. *Trends in Plant Science*, *14*, 521–529.

Du, Z., Zhou, X., Ling, Y., Zhang, Z., & Su, Z. (2010). agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Research*, *38*, W64–70.

Dumas, B., Freyssinet, G., & Pallett, K. E. (1995). Tissue-Specific Expression of Germin-Like Oxalate Oxidase during Development and Fungal Infection of Barley Seedlings. *Plant Physiology*, *107*, 1091–1096.

Dunwell, J. M., Purvis, A., & Khuri, S. (2004). Cupins: the most functionally diverse protein superfamily? *Phytochemistry*, *65*, 7–17.

Ellinger, D., Glockner, A., Koch, J., Naumann, M., Sturtz, V., Schutt, K., Manisseri, C., Somerville, S. C., & Voigt, C. A. (2014). Interaction of the *Arabidopsis* GTPase RabA4c with its effector PMR4 results in complete penetration resistance to powdery mildew. *The Plant Cell*, *26*, 3185–3200.

Ewan, R., Pangestuti, R., Thornber, S., Craig, A., Carr, C., O'Donnell, L., Zhang, C., & Sadanandom, A. (2011). Deubiquitinating enzymes AtUBP12 and AtUBP13 and their tobacco homologue NtUBP12 are negative regulators of plant immunity. *New Phytologist*, *191*, 92–106.

Frey, A., Effroy, D., Lefebvre, V., Seo, M., Perreau, F., Berger, A., Sechet, J., To, A., North, H. M., & Marion-Poll, A. (2012). Epoxycarotenoid cleavage by NCED5 fine-tunes ABA accumulation and affects seed dormancy and drought tolerance with other NCED family members. *Plant Journal*, *70*, 501–512.

Frye, C. A., Tang, D., & Innes, R. W. (2001). Negative regulation of defense responses in plants by a conserved MAPKK kinase. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 373–378.

Gao, M., & Showalter, A. M. (1999). Yariv reagent treatment induces programmed cell death in *Arabidopsis* cell cultures and implicates arabinogalactan protein involvement. *Plant Journal*, *19*, 321–331.

Glawischnig, E., Hansen, B. G., Olsen, C. E., & Halkier, B. A. (2004). Camalexin is synthesized from indole-3-acetaldoxime, a key branching point between primary and secondary metabolism in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 8245–8250.

Glazebrook, J., & Ausubel, F. M. (1994). Isolation of phytoalexin-deficient mutants of *Arabidopsis thaliana* and characterization of their interactions with bacterial pathogens. *Proceedings of the National Academy of Sciences of the United States of America*, *91*, 8955–8959.

Goecks, J., Nekrutenko, A., & Taylor, J. (2010). Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biology*, *11*, R86.

Griesen, D., Su, D., Bérczi, A., & Asard, H. (2004). Localization of an ascorbate-reducible cytochrome b561 in the plant tonoplast. *Plant Physiology*, *134*, 726–734.

Hammond-Kosack, K. E., & Parker, J. E. (2003). Deciphering plant-pathogen communication: fresh perspectives for molecular resistance breeding. *Current Opinion in Biotechnology*, *14*, 177–193.

Hansen, B. G., Halkier, B. A., & Kliebenstein, D. J. (2008). Identifying the molecular basis of QTLs: eQTLs add a new dimension. *Trends in Plant Science*, *13*, 72–77.

Hao, W., Collier, S. M., Moffett, P., & Chai, J. (2013). Structural basis for the interaction between the potato virus X resistance protein (Rx) and its cofactor ran GTPase-activating protein 2 (RanGAP2). *Journal of Biological Chemistry*, *288*, 35868–35876.

Hok, S., Danchin, E. G. J., Allasia, V., Panabières, F., Attard, A., & Keller, H. (2011). An *Arabidopsis* (malectin-like) leucine-rich repeat receptor-like kinase contributes to downy mildew disease. *Plant, Cell and Environment*, *34*, 1944–1957.

Horton, P., Park, K. J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J., & Nakai, K. (2007). WoLF PSORT: Protein localization predictor. *Nucleic Acids Research*, *35*, 585–587.

Hu, P., Meng, Y., & Wise, R. P. (2009). Functional contribution of chorismate synthase, anthranilate synthase, and chorismate mutase to penetration resistance in barley-powdery mildew interactions. *Molecular Plant-Microbe Interactions*, *22*, 311–320.

Hughes, T. R., Mao, M., Jones, A. R., Burchard, J., Marton, M. J., Shannon, K. W., Lefkowitz, S. M., Ziman, M., Schelter, J. M., Meyer, M. R., Kobayashi, S., Davis, C., Dai, H., He, Y. D., Stephaniants, S. B., Cavet, G., Walker, W. L., West, A., Coffey, E., Shoemaker, D. D., Stoughton, R., Blanchard, A. P., Friend, S. H., & Linsley, P. S. (2001). Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nature Biotechnology*, *19*, 342–347.

Humphrey, T. V, Bonetta, D. T., & Goring, D. R. (2007). Sentinels at the wall: cell wall receptors and sensors. *New Phytologist*, *176*, 7–21.

Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong,

S.,Y., Lopez, R., & Hunter, S. (2014). InterProScan 5: Genome-scale protein function classification. *Bioinformatics*, *30*, 1236–1240.

Kendziorski, C. M., Zhang, Y., Lan, H., & Attie, A. D. (2003). The efficiency of pooling mRNA in microarray experiments. *Biostatistics*, *4*, 465–477.

Kim, H., Ridenour, J. B., Dunkle, L. D., & Bluhm, B. H. (2011). Regulation of pathogenesis by light in *Cercospora zeae-maydis*: An Updated Perspective. *The Plant Pathology Journal*, *27*, 103–109.

Kim, M. C., Lee, S. H., Kim, J. K., Chun, H. J., Choi, M. S., Chung, W. S., Moon, B. C., Kang, C. H., Park, C. Y., Yoo, J. H., Kang, Y. H., Koo, S. C., Koo, Y. D., Jung, J. C., Kim, S. T., Schulze-Lefert, P., Lee, S. Y., & Cho, M. J. (2002). *Mlo*, a modulator of plant defense and cell death, is a novel calmodulin-binding protein. Isolation and characterization of a rice *Mlo* homologue. *Journal of Biological Chemistry*, *277*, 19304–19314.

Kirst, M., Caldo, R., Casati, P., Tanimoto, G., Walbot, V., Wise, R. P., & Buckler, E. S. (2006). Genetic diversity contribution to errors in short oligonucleotide microarray analysis. *Plant Biotechnology Journal*, *4*, 489–498.

Kliebenstein, D. (2009). Quantitative genomics: analyzing intraspecific variation using global gene expression polymorphisms or eQTLs. *Annual Review of Plant Biology*, *60*, 93–114.

Kloosterman, B., Oortwijn, M., uitdeWilligen, J., America, T., de Vos, R., Visser, R. G. F., & Bachem, C. W. B. (2010). From QTL to candidate gene: genetical genomics of simple and complex traits in potato using a pooling strategy. *BMC Genomics*, *11*, 158.

Kogenaru, S., Qing, Y., Guo, Y., & Wang, N. (2012). RNA-seq and microarray complement each other in transcriptome profiling. *BMC Genomics*, *13*, 629.

Krogh, A., Larsson, B., von Heijne, G., & Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *Journal of Molecular Biology*, *305*, 567–80.

Levine, A., Pennell, R. I., Alvarez, M. E., Palmer, R., & Lamb, C. (1996). Calcium-mediated apoptosis in a plant hypersensitive disease resistance response. *Current Biology*, *6*(4), 427–437.

Liu, S., Yeh, C.-T., Tang, H. M., Nettleton, D., & Schnable, P. S. (2012). Gene mapping via bulked segregant RNA-Seq (BSR-Seq). *PloS ONE*, *7*, e36406.

Lodha, T. D., & Basak, J. (2012). Plant-pathogen interactions: what microarray tells about it? *Molecular Biotechnology*, *50*, 87–97.

Ma, J., Skibbe, D. S., Fernandes, J., & Walbot, V. (2008). Male reproductive development: gene expression profiling of maize anther and pollen ontogeny. *Genome Biology*, *9*, R181.

Maeda, H., & Dudareva, N. (2012). The shikimate pathway and aromatic amino acid biosynthesis in plants. *Annual Review of Plant Biology*, *63*, 73–105.

Malone, J. H., & Oliver, B. (2011). Microarrays, deep sequencing and the true measure of the transcriptome. *BMC Biology*, *9*, 34.

Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., Geer, R. C., He, J., Gwadz, M., Hurwitz, D. I., Lanczycki, C. J., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Zheng, C., & Bryant, S. H. (2015). CDD: NCBI's conserved domain database. *Nucleic Acids Research*, *43*, D222–D226.

Marino, D., Peeters, N., & Rivas, S. (2012). Ubiquitination during plant immune signaling. *Plant Physiology*, *160*, 15–27.

Marioni, J. C., Mason, C. E., Mane, S. M., Stephens, M., & Gilad, Y. (2008). RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. *Genome Research*, *18*, 1509–1517.

Micali, C. O., Neumann, U., Grunewald, D., Panstruga, R., & Connell, R. (2011). Biogenesis of a specialized plant-fungal interface during host cell internalization of *Golovinomyces orontii* haustoria. *Cellular Microbiology*, *13*, 210–226.

Michelmore, R. W., Paran, I., & Kesseli, R. V. (1991). Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the National Academy of Sciences of the United States of America*, *88*, 9828–9832.

Morel, J. B., & Dangl, J. L. (1997). The hypersensitive response and the induction of cell death in plants. *Cell Death and Differentiation*, *4*, 671–683.

Nahirñak, V., Almasia, N., Hopp, H., & Vazquez-Rovere, C. (2012). Snakin/GASA proteins: Involvement in hormone crosstalk and redox homeostasis. *Plant Signaling and Behavior*, *7*, 1004–1008.

Naidoo, S., Denby, K. J., & Berger, D. K. (2005). Microarray experiments: considerations for experimental design. *South African Journal of Science*, *101*, 347–354.

Poulsen, L., Søe, M. J., Snakenborg, D., Møller, L. B., & Dufva, M. (2008). Multi-stringency wash of partially hybridized 60-mer probes reveals that the stringency along the probe decreases with distance from the microarray surface. *Nucleic Acids Research*, *36*. e132.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, 1–13.

Ritchie, M. E., Silver, J., Oshlack, A., Holmes, M., Diyagama, D., Holloway, A., & Smyth, G. K. (2007). A comparison of background correction methods for two-colour microarrays. *Bioinformatics*, *23*, 2700–2707.

Rose, L. E., Bittner-Eddy, P. D., Langley, C. H., Holub, E. B., Michelmore, R. W., & Beynon, J. L. (2004). The maintenance of extreme amino acid diversity at the disease resistance gene, *RPP13*, in *Arabidopsis thaliana*. *Genetics*, *166*, 1517–1527.

Sacco, M. A., Mansoor, S., & Moffett, P. (2007). A RanGAP protein physically interacts with the NB-LRR protein Rx, and is required for Rx-mediated viral resistance. *Plant Journal*, *52*, 82–93.

Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., Liang, C., Zhang, J., Fulton, L., Graves, T. A., Minx, P., Reily, A. D., Courtney, L., Kruchowski, S. S., Tomlinson, C., Strong, C., Delehaunty, K., Fronick, C., Courtney, B., Rock, S. M., Belter, E., Du, F., Kim, K., Abbott, R. M., Cotton, M., Levy, A., Marchetto, P,m Ochoa, K., Jackson, S. M., Gillam, B., Chen, W., Yan, L., Higginbotham, J., Cardenas, M., Waligorski, J., Applebaum, E., Phelps, L., Falcone, J., Kanchi, K., Thane, T., Scimone, A., Thane, N., Henke, J., Wang, T., Ruppert, J., Shah, N., Rotter, K., Hodges, J., Ingenthron, E., Cordes, M., Sgro, J., Delgado, B., Mead, K., Chinwalla, A., Leonard, S., Crouse, K., Collura, K., Kudrna, D., Currie, J., He, R., Angelova, A., Rajasekar, S., Mueller, T., Lomeli, R., Scara, G., Ko, A., Delaney, K., Wissotski, M., Lopez, G., Campos, D., Braidotti, M., Ashley, E., Golser, W., Kim, H., Lee, S., Lin, J., Dujmic, Z., Kim, W., Talag, J., Zuccolo, A., Fan, C., Sebastian, A., Kramer, M., Spiegel, L., Nascimento, L., Zutavern, T., Miller, B., Ambroise, C., Muller, S., Spooner, W., Narechania, A., Ren, L., Wei, S., Kumari, S., Faga, B., Levy, M. J., McMahan, L., Van Buren, P., Vaughn, M. W., Ying, K., Yeh, C.-T., Emrich, S. J., Jia, Y., Kalyanaraman, A., Hsia, A.-P., Barbazuk, W. B., Baucom, R. S., Brutnell, T. P., Carpita, N. C., Chaparro, C., Chia, J.-M., Deragon, J.-M., Estill, J. C., Fu, Y., Jeddeloh, J. A., Han, Y., Lee, H., Li, P., Lisch, D. R., Liu, Sa., Liu, Z., Nagel, D. H., McCann, M. C., SanMiguel, P., Myers, A. M., Nettleton, D., Nguyen, J., Penning, B. W., Ponnala, L., Schneider, K. L., Schwartz, D. C., Sharma, A., Soderlund, C., Springer, N. M., Sun, Q., Wang, H., Waterman, M., Westerman, R., Wolfgruber, T. K., Yang, L., Yu, Y., Zhang, L., Zhou, S., Zhu, Q., Bennetzen, J. L., Dawe, R. K., Jiang, J., Jiang, N., Presting, G. G., Wessler, S., R., Aluru, S., Martienssen, R. A., Clifton, S. W., McCombie, W. R., Wing, R. A., & Wilson, R. K. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science*, *326*, 1112–1115.

Segura, A., Moreno, M., Madueño, F., Molina, a, & García-Olmedo, F. (1999). Snakin-1, a peptide from potato that is active against plant pathogens. *Molecular Plant-Microbe Interactions*, *12*, 16–23.

Sels, J., Mathys, J., De Coninck, B. M. A, Cammue, B. P. A., & De Bolle, M. F. C. (2008). Plant pathogenesis-related (PR) proteins: A focus on PR peptides. *Plant Physiology and Biochemistry*, *46*, 941–950.

Smyth, G. K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology*, *3*, Article3.

Smyth, G. K., & Speed, T. (2003). Normalization of cDNA microarray data. *Methods*, *31*, 265–273.

't Hoen, P. A. C., Ariyurek, Y., Thygesen, H. H., Vreugdenhil, E., Vossen, R. H. A. M., de Menezes, R. X., Boer, J. M., van Ommen , G-J. B., & den Dunnen, J. T. (2008). Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. *Nucleic Acids Research*, *36*, e141.

Takatsuji, H. (1998). Zinc-finger transcription factors in plants. *Cellular and Molecular Life Sciences*, *54*, 582–596.

Tameling, W. I. L., Nooijen, C., Ludwig, N., Boter, M., Slootweg, E., Goverse, A., Shirasu, K., & Joosten, M. H. A. J. (2010). RanGAP2 mediates nucleocytoplasmic partitioning of the NB-LRR immune receptor Rx in the Solanaceae, thereby dictating Rx function. *The Plant Cell*, *22*, 4176–4194.

Tenaillon, M. I., Sawkins, M. C., Long, a D., Gaut, R. L., Doebley, J. F., & Gaut, B. S. (2001). Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. *mays* L.). *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 9161–9166.

Thimm, O., Bläsing, O., Gibon, Y., Nagel, A., Meyer, S., Krüger, P., Selbig, J., Müller, L. A., Rhee, S. Y., & Stitt, M. (2004). MAPMAN: A user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant Journal*, *37*, 914–939.

Tholl, D. (2006). Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. *Current Opinion in Plant Biology*, *9*, 297–304.

Trapnell, C., Pachter, L., & Salzberg, S. L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*, *25*, 1105–1111.

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.

Trapnell, C., Williams, B. A, Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., Salzberg, S. L., Wold, B. J., & Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*, *28*, 511–515.

Underwood, W. (2012). The plant cell wall: a dynamic barrier against pathogen invasion. *Frontiers in Plant Science*, *3*, 1–6.

Van den Burg, H. A., Harrison, S. J., Joosten, M. H. A. J., Vervoort, J., & de Wit, P. J. G. M. (2006). *Cladosporium fulvum Avr4* protects fungal cell walls against hydrolysis by plant chitinases accumulating during infection. *Molecular Plant-Microbe Interactions*, *19*, 1420–1430.

Van der Biezen, E. A., & Jones, J. D. (1998). The NB-ARC domain: a novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. *Current Biology*, *8*, R226–R227.

Vick, B. A., & Zimmerman, D. C. (1984). Biosynthesis of jasmonic acid by several plant species. *Plant Physiology*, *75*, 458–461.

Wan, J., Dunning, F. M., & Bent, A. F. (2002). Probing plant-pathogen interactions and downstream defense signaling using DNA microarrays. *Functional and Integrative Genomics*, *2*, 259–273.

Wildermuth, M. C., Dewdney, J., Wu, G., & Ausubel, F. M. (2001). Isochorismate synthase is required to synthesize salicylic acid for plant defence. *Nature*, *414*, 562–565.

Wisser, R. J., Kolkman, J. M., Patzoldt, M. E., Holland, J. B., Yu, J., Krakowsky, M., Nelson, R. J., & Balint-Kurti, P. J. (2011). Multivariate analysis of maize disease resistances suggests a pleiotropic

genetic basis and implicates a *GST* gene. *Proceedings of the National Academy of Sciences of the United States of America*, 7339–44.

Xiang, Y., Song, M., Wei, Z., Tong, J., Zhang, L., Xiao, L., Ma, Z., & Wang, Y. (2011). A jacalin-related lectin-like gene in wheat is a component of the plant defence system. *Journal of Experimental Botany*, *62*, 5471–5483.

Yang, Z. (2002). Small GTPases: Versatile signaling switches in plants. *The Plant Cell*, *Supplement*, S375–S389.

Zhang, Z., Collinge, D. B., & Thordal-Christensen, H. (1995). Germin-like oxalate oxidase, a H2O2-producing enzyme, accumulates in barley attacked by the powdery mildew fungus. *Plant Journal*. *8*, 139–145.

Zila, C. T., Ogut, F., Romay, M. C., Gardner, C. A., Buckler, E. S., & Holland, J. B. (2014). Genome-wide association study of *Fusarium* ear rot disease in the U.S.A. maize inbred line collection. *BMC Plant Biology*, *14*, 327.

Zuo, W., Chao, Q., Zhang, N., Ye, J., Tan, G., Li, B., Xing, Y., Zhang, B., Liu, H., Fengler, K. A., Zhao, J., Zhao, X., Chen, Y., Lai, J., Yan, J., & Xu, M. (2015). A maize wall-associated kinase confers quantitative resistance to head smut. *Nature Genetics*, *47*, 151–157.

Chapter 5

Concluding remarks

Maize is an important food crop in Africa, supplying many with their daily calorie intake. The crop is vulnerable to grey leaf spot (GLS), a foliar fungal leaf disease caused *Cercospora zeina* in Africa, which can have a significant impact on yield. Conventional measures used to protect maize crops against GLS are fungicide sprays, tillage and breeding of resistant hybrids (Ward and Nowell 1998).

In order to produce maize hybrids with durable resistance to GLS using molecular breeding techniques, one needs to know which regions of the maize genome contain loci conferring resistance. Information about the molecular mechanisms underpinning these loci will add to the knowledge of the interaction between maize and *C. zeina*.

The material chosen for this study was a maize recombinant inbred line (RIL) population developed from two sub-tropical inbred lines that were well-adapted to maize production in southern Africa, namely CML444 and SC Malawi (Messmer et al. 2009). We hypothesised that candidate genes contributing to quantitative disease resistance to GLS coincide with observed GLS resistance QTL regions on the genome. In an effort to measure disease resistance based on pathogen content, we wanted to be able to quantify *C. zeina* biomass *in planta*. Thus a sensitive and genus specific quantitative PCR (qPCR) assay was developed to measure the amount of *Cercospora* sp. DNA in infected maize leaves. We then aimed to map QTL regions associated with low *C. zeina* biomass *in planta* by using the qPCR assay to quantify the *C. zeina* within plants from a field trial scored for GLS severity. The hypothesis was that QTL derived from the qPCR assay and lesion area data from digital image analysis, reflecting the pathogen content within leaves, would overlap with QTL mapped from GLS severity scores. Lastly, we aimed to identify candidate genes that are differentially expressed between resistant and susceptible RILs during later stages of *C. zeina* infection of maize.

A qPCR assay was developed to detect and quantify *Cercospora zeae-maydis* and *C. zeina* DNA in maize leaves (Korsman et al. 2012). It is both sensitive (Figure 2.4) and specific (Figure 2.2) and can distinguish between *C. zeae-maydis* and *C. zeina* (Table 2.1). This added a

useful tool for molecular quantification of *C. zeina in planta*. It can be applied in glasshouse trials for plant-pathogen interaction studies where early disease events may be of interest; due to its sensitivity it can detect fungus before lesion development. It can also be used in field-infected plants due to the specificity of the assay; it does not amplify targets from other maize pathogens.

The use of the qPCR assay was demonstrated by using data obtained from it to map quantitative trait loci (QTL) for pathogen biomass in a maize RIL population. The same RIL population was used to map GLS resistance QTL from GLS disease scores in the same environment by Berger et al. (2014), which presented an opportunity for comparison of GLS resistance QTL to QTL for fungal biomass. Three QTL were detected using the qPCR assay data, as well as one overlapping QTL from digital image analysis data (Table 3.5; Figure 3.5). The QTL situated on chromosome six that was detected in the analyses with all data types (qPCR, lesion area, and GLS disease scores) was a strong effect QTL (Table 3.5). The QTL mapped for qPCR assay data and lesion area data also coincided with a QTL mapped using disease scores from later time points (Figure 3.5). The correlation of qPCR assay and lesion area data with later disease scores was higher than with earlier disease ratings (Table 3.4), suggesting that the qPCR and lesion area results may indicate disease potential. We can conclude that the phenotype scores from the qPCR assay and digital image analysis of lesion area may be useful in addition to whole plant disease scores in the identification of resistance QTL. The power for detection of QTL will also increase with larger sampling sizes, and this may lead to greater agreement between mapped QTL.

To our knowledge, this study reports the first QTL mapped from fungal biomass detected by qPCR data, thus the study serves as a proof of concept that fungal biomass quantification by qPCR assay can be used to map QTL involved in resistance to fungal pathogens. It would be beneficial to automate the DNA extractions and qPCR, and expand it to include multiple environments and possibly investigate QTL involved in the course of pathogen accumulation in leaves over time. It may also be useful for detecting QTL at early time points, as the qPCR assay seems to indicate disease potential, in order to fast track identification of plants possessing resistance QTL for breeding programs.

This study has revealed various candidate maize genes with putative functions in the resistance phenotype as well as candidate genes that may be responsible for the effects of the QTL

described. Two pools of bulked RILs, one resistant and the other susceptible, were used for expression profiling in a modified version of a bulked segregant analysis. The results brought to light some candidate R-genes, which had higher expression in the resistant bulk of RILs compared to the susceptible bulk of RILs. These candidate R-genes include an *RPP13-like* gene and an *mlo* gene, which have been included in a proposed model of resistance to GLS (Figure 4.10). The RPP13-like protein with CC-NBS-LRR properties and the *mlo* gene product may be the proteins responsible for early recognition of yet to be identified *C. zeina* effectors. The resistance phenotype in this maize *C. zeina* pathosystem may rely on an early hypersensitive response, suggested by the upregulation of cell death pathways in the resistant bulk.

Additional candidate genes involved in the resistance phenotype that co-localise with QTL regions are a gene encoding a malectin containing RLK, an *EDR1-like* gene, a *GTPase* gene, a *cytochrome b561* gene, and a *chorismate synthase* gene. These genes may be candidates for conferring the QTL effect. They are probably constitutively expressed as induced genes would probably not still be up regulated at the late sampling time point of the experimental design. Previously identified candidate genes conferring quantitative resistance to maize have been identified, they are a wall associated kinase conferring resistance in a major QTL for maize head smut (Zuo et al. 2015) and a GST gene implicated in multiple disease resistance in maize, including resistance to GLS (Wisser et al. 2011). Thus candidate genes involved in resistance are not necessarily always R-genes.

There were many genes involved in pathogen defence with higher expression in the susceptible plants. This is reasonable as they have to fight a larger amount of pathogen, while the resistant RILs most probably overcome the fungus during the early infection stages. The genes induced at early infection stages would not have been detected here due to sampling at a late time point. Cell death was identified as a possible strategy in the resistant RILs, whereas the susceptible RILs showed a substantial response to biotic stress, with a number of PR genes expressed (Tables B4 and B6b, Appendix B).

There were some limitations of the study that may be overcome with alternative approaches in future investigations. First, the QTL and genes identified in this study may only be relevant in the CML444 × SC Malawi maize RIL population. The maize RIL population size was small, thus the power to detect QTL was also relatively small. An alternative would be to use a

genome-wide association study or nested association mapping (Benson et al. 2015). Secondly, the qPCR assay is costly and sampling is time consuming, possibly making it prohibitive for much larger population sizes. Thirdly, the sampling time for expression profiling was more appropriate for identifying genes involved in the susceptible response. A glasshouse trial where early time point samples can be harvested would be advantageous for identifying induced genes in the early resistance response to *C. zeina*. Fourth, the study gave no information on induced or constitutive expression of candidate genes. Such information could be obtained from a comparison of infected plants with disease free control plants, or a comparison between disease free resistant and susceptible RILs. Finally, the RNA-seq and microarray data correlated poorly. Microarray analysis is better suited to comparing different treatments of one genotype, rather than comparing different genotypes due to possible sequence differences. RNA-seq is not constrained by transcript sequence difference, and it takes into consideration splice variants. RNA-seq can also be used to identify candidate genes with sequence differences between different genotypes that could result in functional differences. Microarrays only identify genes for which there is prior knowledge while RNA-seq is not limited by this.

The bulked segregant analysis approach used for expression profiling of the resistance and susceptible RILs gives an overview of the expression responses of genes in the selected QTL for the whole population. The resistance and susceptible bulks had mostly contrasting resistance and susceptible alleles at the QTL. Therefore differential expression of genes contributing to the QTL effects should have been observed, but the selection of RILs for bulking was not perfect. There were limited RILs to choose from in the small population, so if a QTL is based on the expression of two alleles the output may be confounded by the expression in the RILs containing the alternate allele to the one desired for the specific bulk. An alternate approach would be eQTL analysis (Hansen et al. 2008; Kliebenstein 2009) as carried out by Christie (2014). In spite of the problems with bulking, it proved successful as a number of genes, which coincided with QTL were identified in the resistance response (Table B6a), three of which were included in the proposed resistance model (Figure 4.10).

Future research that may confirm the roles of these candidate genes and contribute to better understanding of their mode of action would be a comparison of the sequences of the candidate genes and their promoter regions between two alleles present in the maize RIL population. Additionally, expression profiling of the genes using RT-qPCR, at early time points, under controlled infection may give an idea of the early responses of maize involved in resistance.

Over expression, knockdown, or knockout mutants could be generated to study specific candidate genes. *Agrobacterium* mediated transformation can be used produce over expression transformation events (Zhao et al. 2001; Ishida et al. 2007) to confirm the roles of candidate genes in the resistance response. Over expression transformation events in Hi-II maize, which has been observed by our research group to be susceptible to *C. zeina*, may confirm that higher expression levels of a candidate gene contributes to the GLS resistance response. Alternatively, knockout mutants could be used to observe the effect of the lack of a candidate gene in a plant. RNAi could be used or a public resource of maize mutants, such as the UniformMu collection (McCarty et al. 2005; Settles et al. 2007) in the W22 background, of which the GLS resistance phenotype is unknown. Candidate genes which should be investigated first would be the *RPP13-like gene* depicted in the resistance model (Figure 4.10a), the *mlo* gene (Figure 4.10c) and the *EDR1* gene (Figure 4.10h) due to the strong support for their involvement in disease resistance from other studies (Büschges et al. 1997; Bittner-Eddy et al. 2000; Frye et al. 2001; Kim et al. 2002; Rose et al. 2004).

In conclusion, this project has generated a molecular tool for quantification of *C. zeina* in maize, and the method was successfully applied in identifying QTL responsible for inhibiting *in planta* growth of *C. zeina*. This will be beneficial for early screening of GLS resistant maize material. Additionally, transcriptone analysis of bulked GLS resistant and susceptible RILs was successfully used to discover candidate GLS resistance genes. Should GMO studies prove that these genes do confer resistance to GLS, SNP markers linked to the resistant allele of the gene can be employed in molecular breeding programmes. Alternatively, the trans genes could be introgressed into local maize varieties to produce a transgenic GLS resistance GMO event to be used for breeding hybrids with greater resistance to GLS. This will increase maize yield while minimizing fungicide application.

References:

Benson, J. M., Poland, J. A., Benson, B. M., Stromberg, E. L., & Nelson, R. J. (2015). Resistance to gray leaf spot of maize: genetic architecture and mechanisms elucidated through nested association mapping and near-isogenic line analysis. *PLOS Genetics*, *11*, e1005045.

Berger, D. K., Carstens, M., Korsman, J. N., Middleton, F., Kloppers, F. J., Tongoona, P., & Myburg, A. A. (2014). Mapping QTL conferring resistance in maize to gray leaf spot disease caused by *Cercospora zeina*. *BMC Genetics*, *15*, 60

Bittner-Eddy, P. D., Crute, I. R., Holub, E. B., & Beynon, J. L. (2000). *RPP13* is a simple locus in *Arabidopsis thaliana* for alleles that specify downy mildew resistance to different avirulence determinants in *Peronospora parasitica*. *The Plant Journal*, *21*, 177–188.

Büschges, R., Hollricher, K., Panstruga, R., Simons, G., Wolter, M., Frijters, A., van Daelen, R., van der Lee, T., Diergaarde, P., Groenendijk, J., Töpsch, S., Vos, P., Salamini, F., & Schulze-Lefert, P. (1997). The barley *Mlo* gene: a novel control element of plant pathogen resistance. *Cell*, *88*, 695–705.

Christie, N. (2014). *Transcriptional regulation underlying the quantitative genetic response of maize to grey leaf spot disease*. University of Pretoria.

Frye, C. A, Tang, D., & Innes, R. W. (2001). Negative regulation of defense responses in plants by a conserved MAPKK kinase. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 373–378.

Hansen, B. G., Halkier, B. A., & Kliebenstein, D. J. (2008). Identifying the molecular basis of QTLs: eQTLs add a new dimension. *Trends in Plant Science*, *13*, 72–77.

Ishida, Y., Hiei, Y., & Komari, T. (2007). *Agrobacterium*-mediated transformation of maize. *Nature Protocols*, *2*, 1614–1621.

Kim, M. C., Lee, S. H., Kim, J. K., Chun, H. J., Choi, M. S., Chung, W. S., Moon, B. C., Kang, C. H., Park, C. Y., Yoo, J. H., Kang, Y. H., Koo, S. C., Koo, Y. D., Jung, J. C., Kim, S. T., Schulze-Lefert, P., Lee, S. Y., & Cho, M. J. (2002). *Mlo*, a modulator of plant defense and cell death, is a novel calmodulin-binding protein. Isolation and characterization of a rice *Mlo* homologue. *Journal of Biological Chemistry*, *277*, 19304–19314.

Kliebenstein, D. (2009). Quantitative genomics: analyzing intraspecific variation using global gene expression polymorphisms or eQTLs. *Annual Review of Plant Biology*, *60*, 93–114.

Korsman, J., Meisel, B., Kloppers, F., Crampton, B., & Berger, D. (2012). Quantitative phenotyping of grey leaf spot disease in maize using real-time PCR. *European Journal of Plant Pathology*, *133*, 461–471.

Messmer, R., Fracheboud, Y., Bänziger, M., Vargas, M., Stamp, P., & Ribaut, J.-M. (2009). Drought stress and tropical maize: QTL-by-environment interactions and stability of QTLs across environments for yield components and secondary traits. *Theoretical and Applied Genetics, 119*, 913–30.

McCarty, D. R., Settles, A. M., Suzuki, M., Tan, B. C., Latshaw, S., Porch, T., Robin, K., Baier, J., Avigne, W., Lai, J., Messing, J., Koch, K. E., & Curtis, H. L. (2005). Steady-state transposon mutagenesis in inbred maize. *The Plant Journal*, *44*, 52–61.

Rose, L. E., Bittner-Eddy, P. D., Langley, C. H., Holub, E. B., Michelmore, R. W., & Beynon, J. L. (2004). The maintenance of extreme amino acid diversity at the disease resistance gene, *RPP13*, in *Arabidopsis thaliana*. *Genetics*, *166*, 1517–1527.

Settles, A. M., Holding, D. R., Tan, B. C., Latshaw, S. P., Liu, J., Suzuki, M., Li, L., O'Brien, B. A., Fajardo, D. S., Wroclawska, E., Tseung, C.-W., Lai, J., Hunter, C. T., Avigne, W. T., Baier, J., Messing, J., Hannah, L. C., Koch, K. E., Becraft, P. W., Larkins, B. A., & McCarty, D. R. (2007). Sequence-indexed mutations in maize using the UniformMu transposon-tagging population. *BMC Genomics*, *8*, 116.

Ward, J. M. J., & Nowell, D. C. (1998). Integrated management practices for the control of maize grey leaf spot. *Integrated Pest Management Reviews*, *3*, 177–188.

Wisser, R. J., Kolkman, J. M., Patzoldt, M. E., Holland, J. B., Yu, J., Krakowsky, M., Nelson, R. J., & Balint-Kurti, P. J. (2011). Multivariate analysis of maize disease resistances suggests a pleiotropic genetic basis and implicates a *GST* gene. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 7339–44.

Zhao, Z. Y., Gu, W., Cai, T., Tagliani, L., Hondred, D., Bond, D., Schroeder, S., Rudert, M., & Pierce, D. (2001). High throughput genetic transformation mediated by *Agrobacterium tumefaciens* in maize. *Molecular Breeding*, *8*, 323–333.

Zuo, W., Chao, Q., Zhang, N., Ye, J., Tan, G., Li, B., Xing, Y., Zhang, B., Liu, H., Fengler, K. A., Zhao, J., Zhao, X., Chen, Y., Lai, J., Yan, J., & Xu, M. (2015). A maize wall-associated kinase confers quantitative resistance to head smut. *Nature Genetics*, *47*, 151–157.

Appendix A

Quality control for microarray and RNA-seq

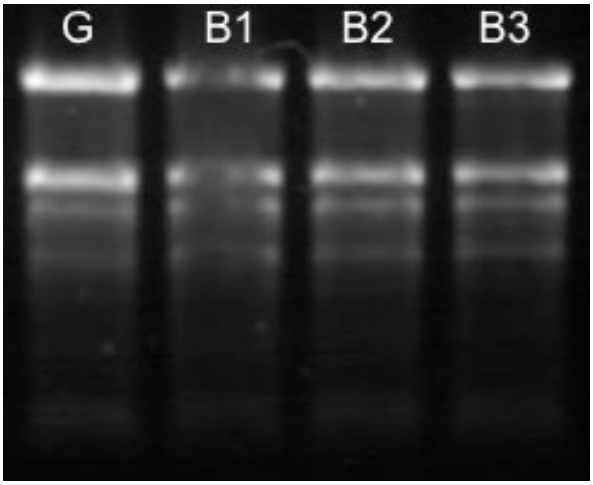


Figure A1: Denaturing agarose gel showing total RNA extracted from maize leaves from a plant grown under controlled conditions in the glasshouse (G) and three plants from the Baynesfield 2008/2009 field trial (B1 – B3).

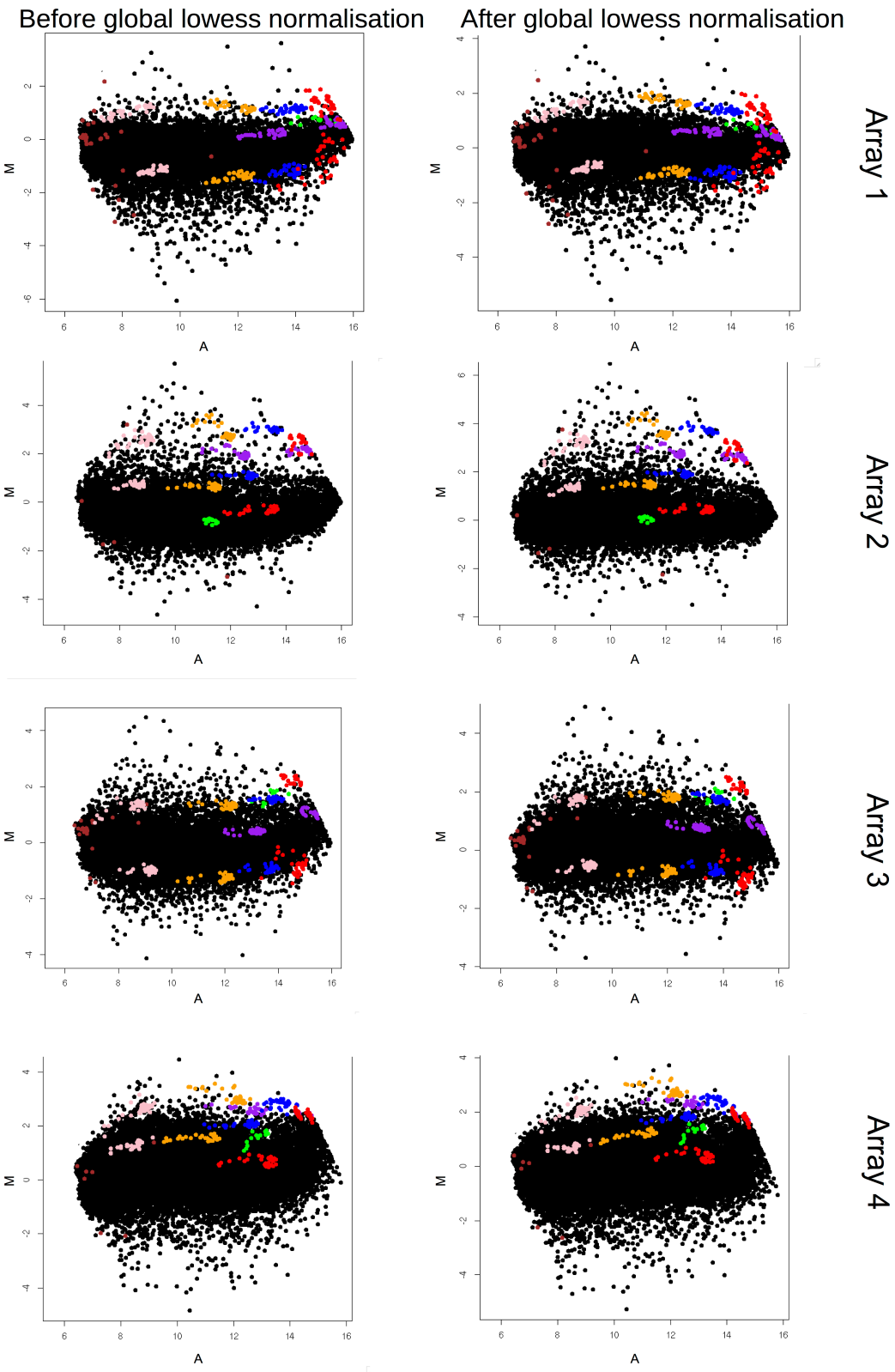


Figure A2: MA plots of the four arrays of the microarray experiment before and after within slide global lowess normalisation. M = $\log_2$R - $\log_2$G, representing the log ratio, and the intensity is represented by A = ($\log_2$R + $\log_2$G)/2. The coloured spots represent Agilent spike in controls not used in the analysis.

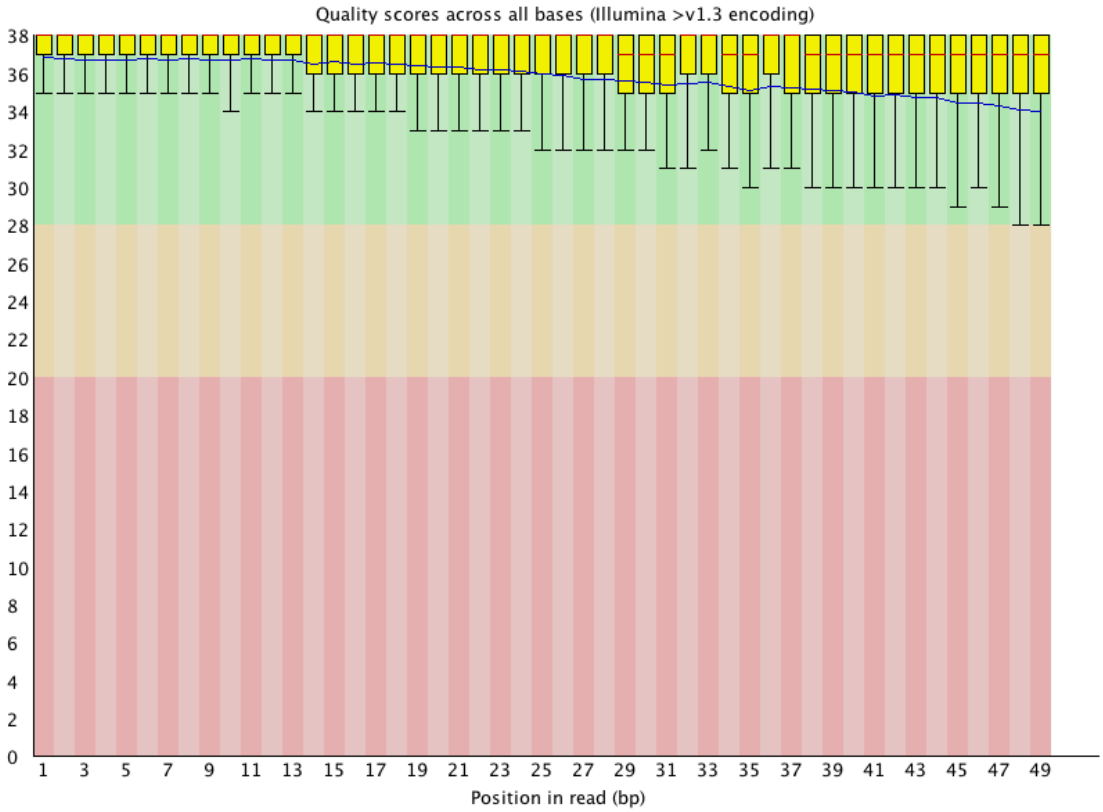


Figure A3: Sequence quality score plot for the bulked RNA-seq showing the quality scores (y-axis) per base (x-axis) of the read length. The 50 nt reads were acceptable across the read length.

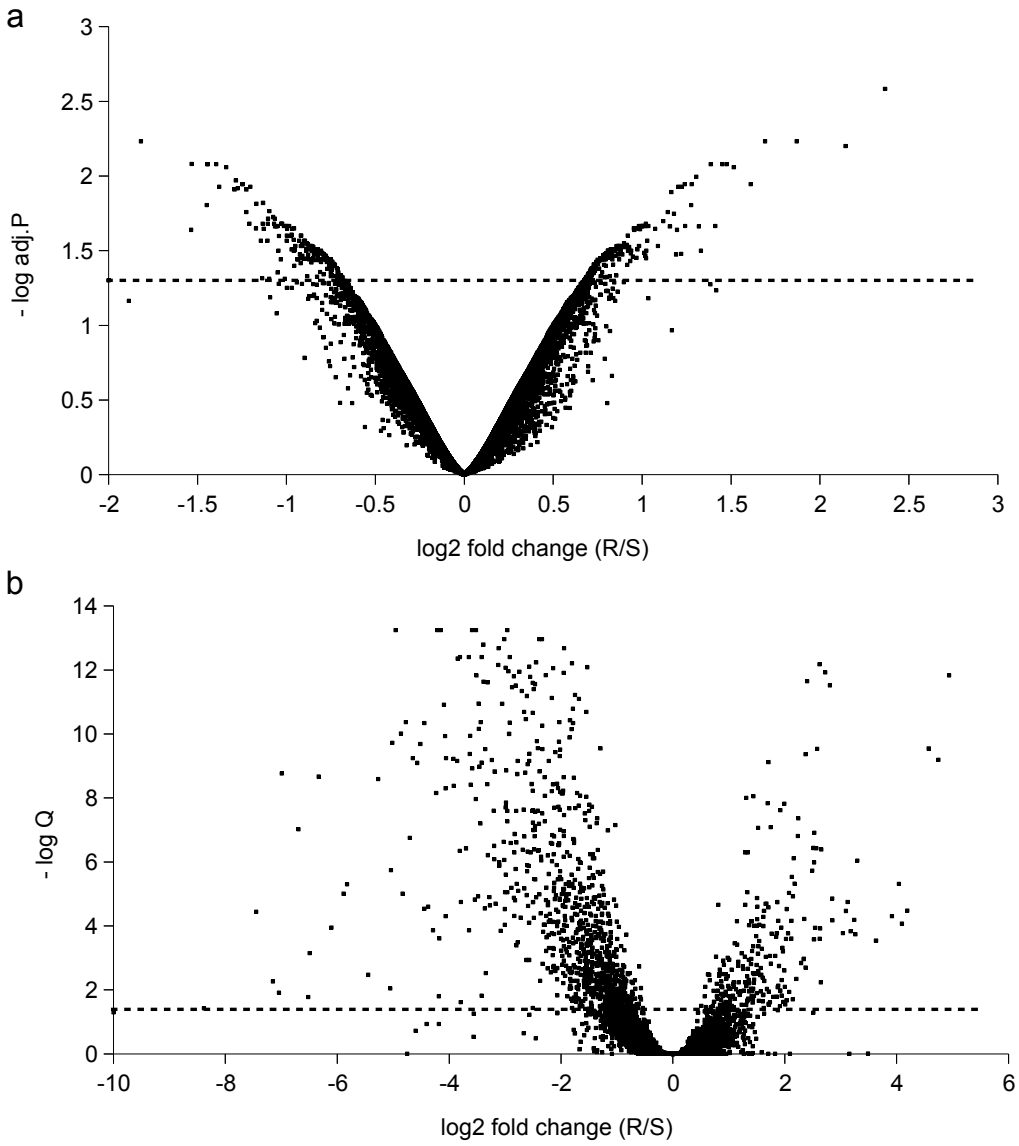


Figure A4: Volcano plots of the relative expression of genes in the R bulk compared to the S bulk from the microarray experiment (a) and the RNA-seq experiment (b). The $\log_2$ fold changes of gene expression is compared to the $-\log_{10}$ of the adjusted P value for microarray the Q value for RNA-seq. The dotted lines indicate the FDR < 5% cutoff value for statistical significance.

Table A1: QC statistics for bulked microarray analysis

	Biol rep 1 2 technical reps		Biol rep 2	Biol rep 3
	R1 cy3 S1 cy5	S1 cy3 R1 cy5	R2 cy3 S2 cy5	S3 cy3 R3 cy5
% informative probes	71.89	64.47	73.01	69.82
% background features	28.11	35.53	26.99	30.18
Variance cy5 background for all spots	26.87	29.33	34.25	24.46
Variance cy3 background for all spots	19.04	26.59	21.44	13.72
Mean cy5 probes[a]	2134	2089	2268	2343
Mean cy5 probes not flagged[b]	2957	3224	3096	3343
Mean cy3 probes[a]	2631	2936	2510	1715
Mean cy3 probes not flagged[b]	3643	4533	3425	2443
cy5 negative controls with positive signal	0.65	0.65	0.65	0.65
cy3 negative controls with positive signal	0.65	2.61	0.65	1.31
intercept of regression model[c]	0.05	0.88	0.17	1.45
gradient of regression model[d]	1.60	1.68	1.48	2.25
R-squared of regression model[e]	0.78	0.85	0.87	0.70

[a]mean [mean probe (i.e. not pos/neg controls or 'ignore' spots) less corresponding median background intensity]

[b]only for probes that are not flagged: mean [mean probe (i.e. not pos/neg controls or 'ignore' spots) less corresponding median background intensity]

[c]intercept of regression model of means of observed positive control spot intensities against expected intensities for each positive control spot-type

[d]gradient of regression model of means of observed positive control spot intensities against expected intensities for each positive control spot-type

[e]R-squared value of regression model of means of observed positive control spot intensities against expected intensities for each positive control spot-type

Table A2: RNA-seq reads obtained for samples and mapping statistics

Criteria	**Bulk_S_1**	**Bulk_S_2**	**Bulk_S_3**	**Bulk_R_1**	**Bulk_R_2**	**Bulk_R3**
Fastq single reads per file	33 865 116	32 318 237	39 371 687	35 717 600	39 830 558	37 444 703
Total paired reads	67 730 232	64 636 474	78 743 374	71 435 200	79 661 116	74 889 406
Mapped reads	54 125 199	50 760 445	59 669 044	61 805 261	72 685 956	66 728 414
% Mapped reads	79.91%	78.53%	75.78%	86.52%	91.24%	89.10%
Un-mapped reads (QC and seq)	13 605 033	13 876 029	19 074 330	9 629 939	6 975 160	8 160 992
% Un-mapped reads	20.09%	21.47%	24.22%	13.48%	8.76%	10.90%
Total reads mapped in pairs	37 629 518	35 962 526	41 438 480	42 566 090	51 041 622	45 748 756
% reads mapped in pairs	55.56%	55.64%	52.62%	59.59%	64.07%	61.09%
Total reads mapped as singles	16 495 681	14 797 919	18 230 564	19 239 171	21 644 334	20 979 658
% reads mapped as singles	30.48%	29.15%	30.55%	31.13%	29.78%	31.44%
Total reads mapped uniquely	42 818 112	39 404 691	46 673 378	51 314 681	56 802 328	54 195 105
% reads mapped uniquely	79.11%	77.63%	78.22%	83.03%	78.15%	81.22%
Total reads mapping ambiguously	11 307 087	11 355 754	12 995 666	10 490 580	15 883 628	12 533 309
% reads mapping ambiquously	20.89%	22.37%	21.78%	16.97%	21.85%	18.78%
Total reads that mapped uniquely in pairs	28 009 090	26 214 016	30 400 844	34 038 524	37 703 692	35 530 334
% reads that mapped uniquely in pairs	51.75%	51.64%	50.95%	55.07%	51.87%	53.25%

Table A3: Summarizing statistics for FPKMs calculated using Cufflinks

Description	Bulk_S_1	Bulk_S_2	Bulk_S_3	Bulk_R_1	Bulk_R_2	Bulk_R_3
genes with FPKM > 0 and status = OK	24 406	23 923	24 444	24 512	25 390	24 596
Average FPKM	8.14	8.68	8.18	7.70	6.71	7.02
STD Dev	24.45	27.23	25.53	23.99	19.07	21.32
Min FPKM	-	-	-	-	-	-
Max FPKM	853.53	1 220.31	1 123.54	1 605.39	727.19	909.50

Appendix B

GO terms and gene lists from Microarray and RNA-seq.

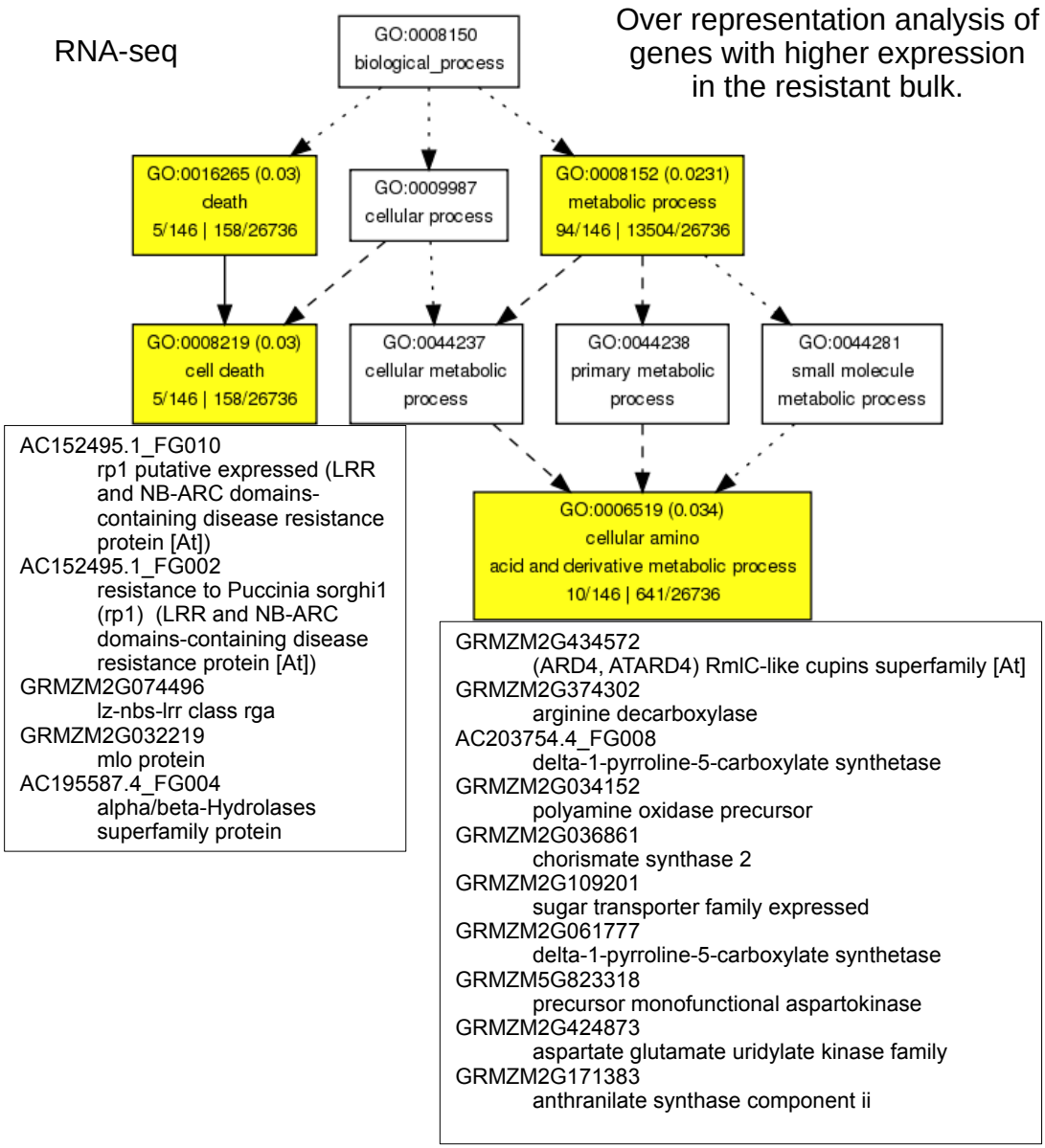
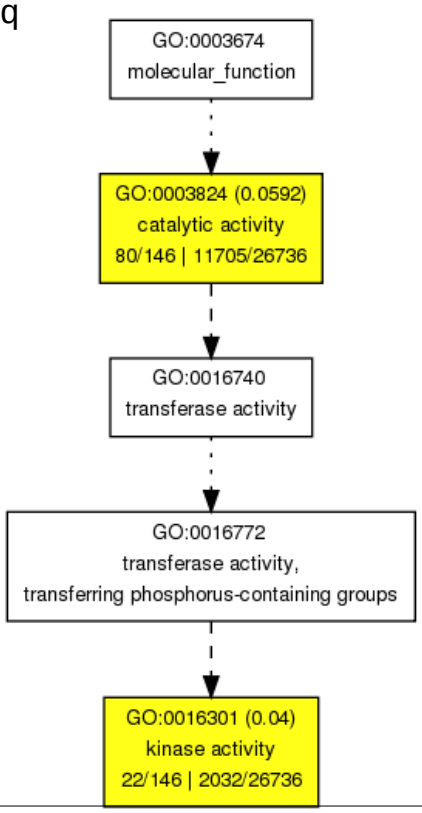


Figure B1: Over representation analysis of biological process GO terms from genes with relatively higher expression in the resistant bulk from the RNA-seq analysis. The significantly over represented GO terms are in coloured blocks. The FDR is indicated in brackets, and the number of genes associated with the GO terms in the test set is indicated at the bottom of each block, followed by the proportion in the reference set. The genes associated with the GO terms are listed in blocke below the relevant GO terms. Solid lines with arrows indicate there are statistically significant over-represented GO terms in both connected boxes, whereas dashed lines indicate one box contains an over-represented GO term, and dotted lines connect boxes without over-represented GO terms.

RNA-seq

Over representation analysis of genes with higher expression in the resistant bulk.



GRMZM2G155837	shr5-receptor-like kinase
AC203754.4_FG008	pyrroline-5-carboxylate synthetase
GRMZM2G439799	Leucine-rich repeat protein kinase family protein
GRMZM2G412601	cbl-interacting protein kinase
GRMZM2G424873	aspartate glutamate uridylate kinase family protein
GRMZM2G161094	unknown
GRMZM2G000633	af403126_1s-locus receptor-like kinase rlk14
GRMZM2G132882	Phosphofructokinase family protein [At]
GRMZM2G125001	cbl-interacting protein kinase
GRMZM2G109201	sugar transporter family expressed
GRMZM5G823318	precursor monofunctional aspartokinase
GRMZM2G149289	protein expressed
GRMZM2G014366	serine/threonine-protein kinase (Rice)
GRMZM2G121928	unknown
GRMZM2G039665	protein kinase receptor type precursor
GRMZM2G330049	cipk-like protein 1
GRMZM2G176604	BRASSINOSTEROID INSENSITIVE 1-associated receptor kinase 1
GRMZM2G113771	Protein kinase superfamily protein [At]
GRMZM2G101460	gtp-binding protein
GRMZM2G044180	enhanced disease resistance 1
GRMZM2G157727	phytochrome a
GRMZM2G702599	Leucine-rich repeat protein kinase family protein [At]

Figure B2: Over representation analysis of molecular function GO terms from genes with relatively higher expression in the resistant bulk from the RNA-seq analysis. The significantly over represented GO terms are in coloured blocks. The FDR is indicated in brackets, and the number of genes associated with the GO terms in the test set is indicated at the bottom of each block, followed by the proportion in the reference set. The genes associated with the GO terms are listed in blocks below the relevant GO terms. Solid lines with arrows indicate there are statistically significant over-represented GO terms in both connected boxes, whereas dashed lines indicate one box contains an over-represented GO term, and dotted lines connect boxes without over-represented GO terms.

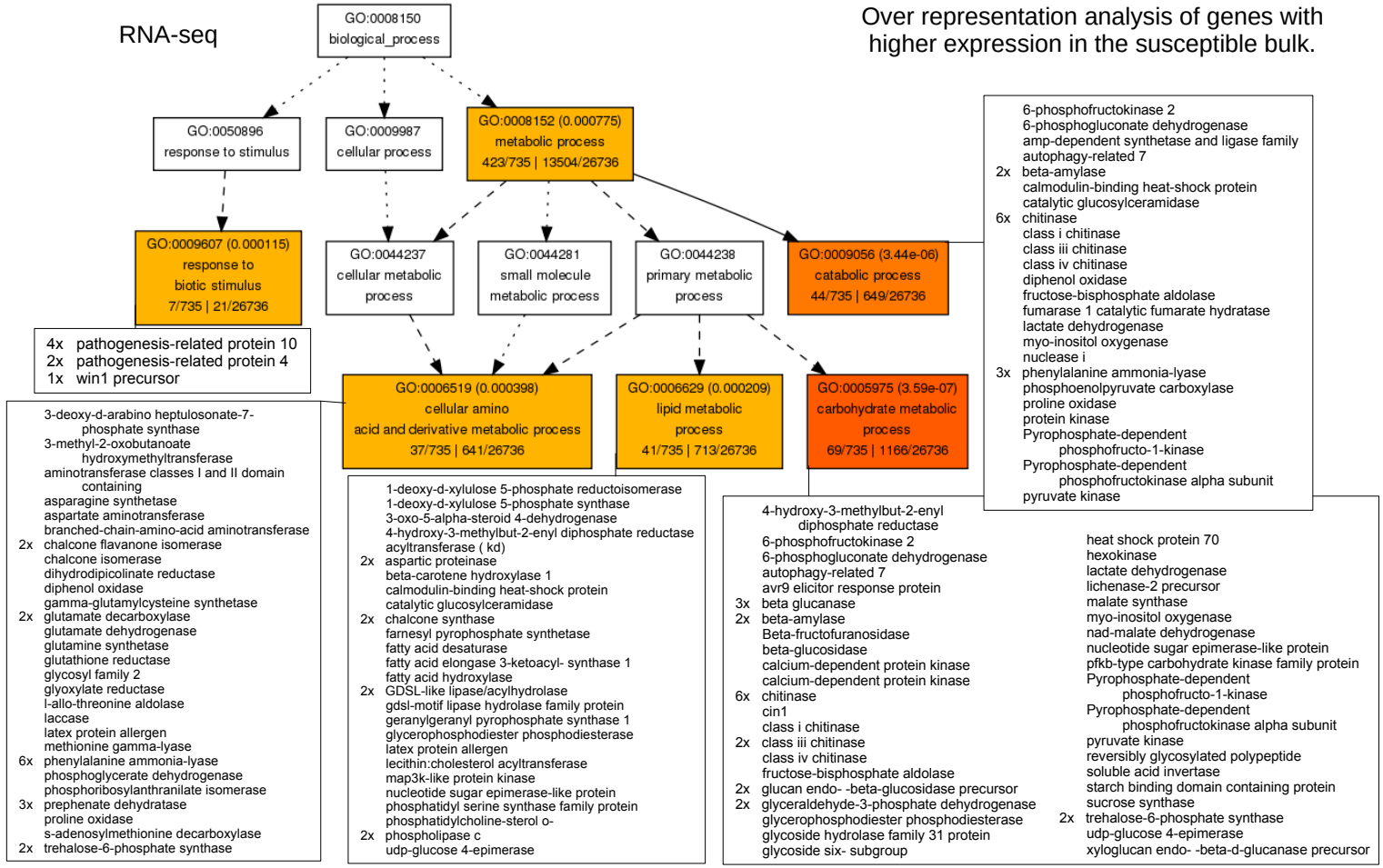
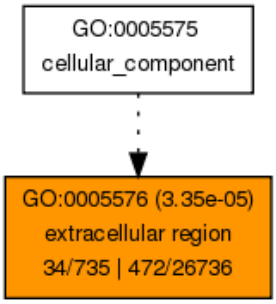


Figure B3: Over representation analysis of biological process GO terms from genes with relatively higher expression in the susceptible bulk from the RNA-seq analysis. The significantly over represented GO terms are in coloured blocks. The FDR is indicated in brackets, and the number of genes associated with the GO terms in the test set is indicated at the bottom of each block, followed by the proportion in the reference set. The annotations of the genes associated with the GO terms are listed in blocks below/next to the relevant GO terms. Solid lines with arrows indicate there are statistically significant over-represented GO terms in both connected boxes, whereas dashed lines indicate one box contains an over-represented GO term, and dotted lines connect boxes without over-represented GO terms.

RNA-seq

Over representation analysis of genes with higher expression in the susceptible bulk.



GRMZM2G090245	germin-like protein 1 [At]
GRMZM2G142709	glycosyl transferase 8 [Rice]
GRMZM2G465226	pathogenesis-related protein class i
GRMZM2G170857	RmlC-like cupins superfamily protein [At]
GRMZM2G030772	RmlC-like cupins superfamily protein [At]
GRMZM2G169033	diphenol oxidase
GRMZM2G180422	hypothetical protein LOC100280170 [Zea mays]
GRMZM2G456997	pathogenesis-related protein class i
AC177897.2_FG002	spx (syg1 pho81 xpr1) domain-containing protein zinc finger (c3hc4-type ring finger)
GRMZM2G179768	matrix metalloproteinase [At}
GRMZM2G170829	germin a
AC205274.3_FG001	pathogenesis-related maize seed protein
AC190772.4_FG011	RmlC-like cupins superfamily protein [At]
GRMZM2G026980	xyloglucan endo- -beta-d-glucanase precursor
GRMZM2G178817	RmlC-like cupins superfamily protein [At]
GRMZM2G093622	RmlC-like cupins superfamily protein [At]
GRMZM2G116520	bowman-birk type trypsin inhibitor
GRMZM2G044194	phytosulfokine 4 precursor [At]
GRMZM2G117706	bacterial-induced peroxidase precursor
GRMZM2G050768	Bowman-Birk type bran trypsin inhibitor precursor [Rice}
GRMZM2G031317	phytosulfokines 2 precursor
GRMZM2G045809	RmlC-like cupins superfamily protein [At]
GRMZM2G071390	germin-like protein subfamily 1 member 8 expressed
GRMZM2G074248	harpin inducing protein
GRMZM2G320786	Laccase 14 [At]
GRMZM2G049930	RmlC-like cupins superfamily protein [At]
GRMZM2G176798	RmlC-like cupins superfamily protein [At]
GRMZM2G156632	wound-induced serine protease inhibitor
GRMZM2G149714	RmlC-like cupins superfamily protein [At]
GRMZM2G115491	RmlC-like cupins superfamily protein [At]
GRMZM2G006676	unknown
GRMZM2G093606	RmlC-like cupins superfamily protein [At]
GRMZM2G085116	unknown
GRMZM2G148964	unknown

Figure B4: Over representation analysis of cellular compartment GO terms from genes with relatively higher expression in the susceptible bulk from the RNA-seq analysis. The significantly over represented GO terms are in coloured blocks. The FDR is indicated in brackets, and the number of genes associated with the GO terms in the test set is indicated at the bottom of each block, followed by the proportion in the reference set. The genes associated with the GO terms are listed in blocks below the relevant GO terms. Solid lines with arrows indicate there are statistically significant over-represented GO terms in both connected boxes, whereas dashed lines indicate one box contains an over-represented GO term, and dotted lines connect boxes without over-represented GO terms.

RNA-seq

Over representation analysis of genes with higher expression in the susceptible bulk.

Figure B5: Over representation analysis of molecular function GO terms from genes with relatively higher expression in the susceptible bulk from the RNA-seq analysis. The significantly over represented GO terms are in coloured blocks. The FDR is indicated in brackets, and the number of genes associated with the GO terms in the test set is indicated at the bottom of each block, followed by the proportion in the reference set. The annotations of the genes associated with the GO terms are listed in blocks below the relevant GO terms. Solid lines with arrows indicate there are statistically significant over-represented GO terms in both connected boxes, whereas dashed lines indicate one box contains an over-represented GO term, and dotted lines connect boxes without over-represented GO terms.

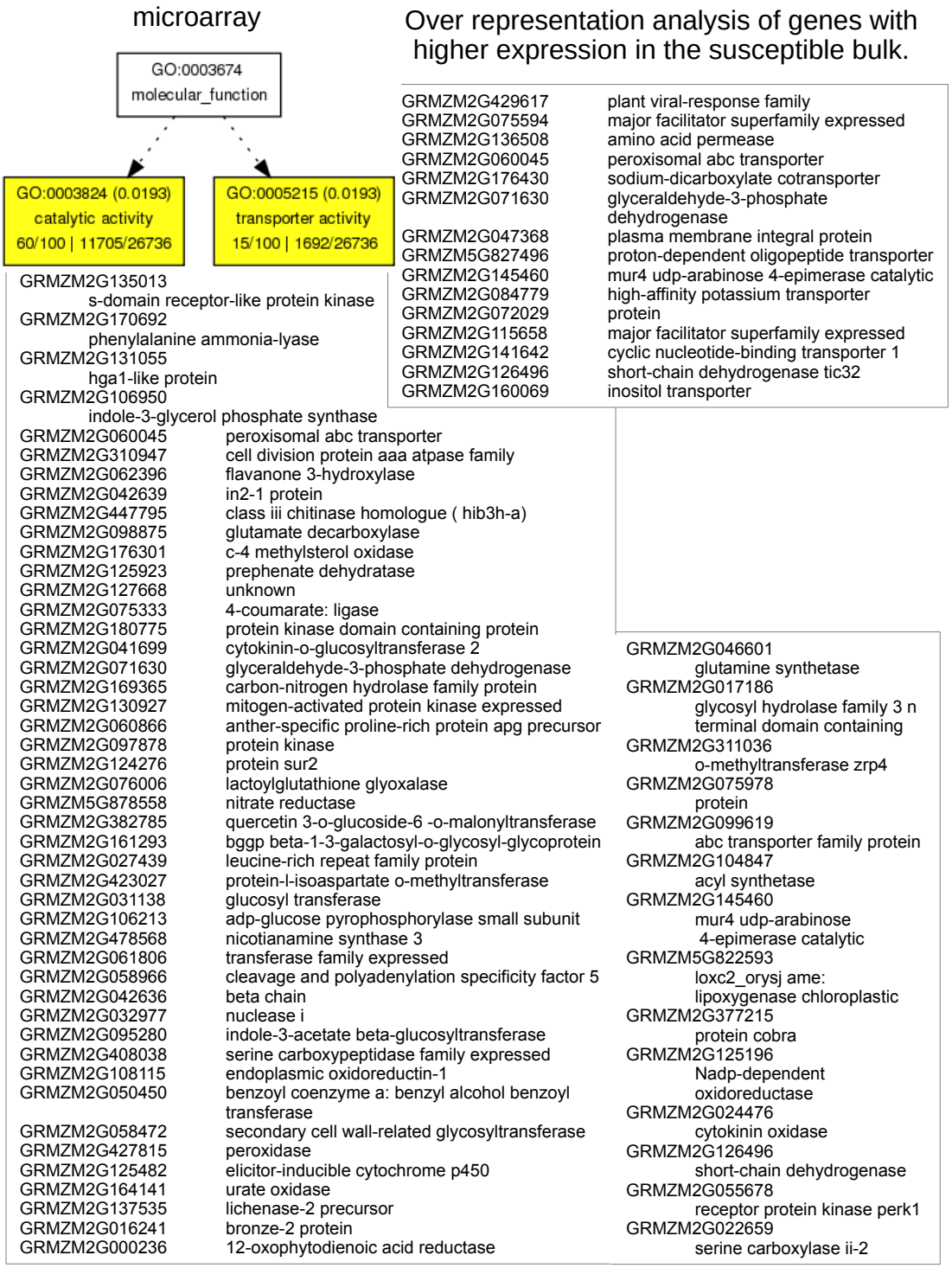
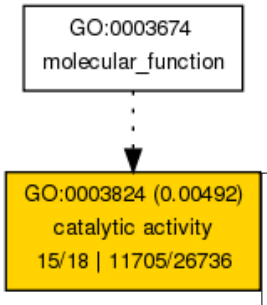


Figure B6: Over representation analysis of molecular function GO terms from genes with relatively higher expression in the susceptible bulk from the microarray analysis. The significantly over represented GO terms are in coloured blocks. The FDR is indicated in brackets, and the number of genes associated with the GO terms in the test set is indicated at the bottom of each block, followed by the proportion in the reference set. The genes associated with the GO terms are listed in blocks below/next to the relevant GO terms. Solid lines with arrows indicate there are statistically significant over-represented GO terms in both connected boxes, whereas dashed lines indicate one box contains an over-represented GO term, and dotted lines connect boxes without over-represented GO terms.

RNA-seq under QTL

Over representation analysis of genes with higher expression in the resistant bulk.



GRMZM2G331833	ATP-dependent Clp protease ATP-binding
GRMZM2G097854	leucoanthocyanidin reductase
GRMZM2G065210	ubiquinone biosynthesis methyltransferase
GRMZM2G036861	chorismate synthase 2
AC203754.4_FG008	delta-1-pyrroline-5-carboxylate synthetase gene
GRMZM2G133718	Putative Serine Carboxypeptidase
GRMZM2G055320	4-coumarate--CoA ligase 4-like
GRMZM2G044180	serine/threonine-protein kinase EDR1-like
GRMZM2G061777	delta-1-pyrroline-5-carboxylate synthase-like
GRMZM2G130375	beta-galactosidase
GRMZM2G108103	polyphenol oxidase
GRMZM2G155837	LRR receptor-like serine/threonine-protein kinase
GRMZM2G000633	cysteine-rich Receptor-like protein kinase 10
GRMZM2G152919	ubiquitin-protein ligase
GRMZM2G177412	glutamyl-tRNA reductase

Figure B7: Over representation analysis of molecular function GO terms from genes with relatively higher expression in the resistant bulk from the subset of RNA-seq genes which fall within the borders of the QTL. The significantly over represented GO terms are in coloured blocks. The FDR is indicated in brackets, and the number of genes associated with the GO terms in the test set is indicated at the bottom of each block, followed by the proportion in the reference set. The genes associated with the GO terms are listed in blocks next to the relevant GO terms. Solid lines with arrows indicate there are statistically significant over-represented GO terms in both connected boxes, whereas dashed lines indicate one box contains an over-represented GO term, and dotted lines connect boxes without over-represented GO terms.

RNA-seq under QTL

AC214360.3_FG001	terpene synthase
GRMZM2G078500	4,5-DOPA dioxygenase extradiol
GRMZM2G170692	phenylalanine ammonia-lyase
GRMZM2G067225	allene oxide synthase1 (aos1)
GRMZM2G132706	UDP-glycosyltransferase 85A1-like
GRMZM2G110369	oxidoreductase, 2OG-Fe oxygenase family protein
GRMZM2G152417	AMP-binding enzyme
GRMZM2G396856	probable polyamine oxidase 2
GRMZM2G108125	aminotransferase y4uB
GRMZM2G086845	fructokinase-1 [Zea mays]
GRMZM2G047546	65-kDa microtubule-associated protein 3-like
GRMZM2G098875	glutamate decarboxylase 1-like
GRMZM2G035584	anthranilate N-benzoyltransferase protein 1
GRMZM2G176307	glyceraldehyde-3-phosphate dehydrogenase
GRMZM2G404702	copper-transporting ATPase RAN1-like
GRMZM2G167253	receptor-like protein kinase HAIKU2 precursor
GRMZM2G140817	putative cytochrome P450 superfamily protein
GRMZM2G065585	beta 1,3 glucanase (PR 2)
GRMZM2G097728	ypt homolog2 (ypt2)
GRMZM2G029048	phenylalanine ammonia-lyase
GRMZM2G031447	gibberellin receptor GID1L2
GRMZM2G132212	receptor-like protein kinase HSL1
GRMZM2G123107	beta-1,3-glucanase 1 (PR 2)
GRMZM2G151406	copper-transporting ATPase RAN1-like
GRMZM2G118241	Putative Serine Carboxypeptidase homologue
GRMZM2G070708	Galactosyltransferase (avr9 elicitor response protein)
GRMZM2G148387	glutaredoxin subgroup I
GRMZM2G145715	pgd2 (6-phosphogluconate dehydrogenase2)
GRMZM2G025832	flavonoid 3' hydroxylase gene
GRMZM2G170017	carbonyl reductase 1
GRMZM5G872256	galactinol synthase 3 (gols3) (osmoprotectant)
GRMZM2G129189	endochitinase PR4
GRMZM2G103342	peroxidase precursor
GRMZM2G104575	dihydrodipicolinate reductase
GRMZM2G391042	calcium-transporting ATPase 8, plasma membrane-type
GRMZM2G334660	phenylalanine ammonia-lyase
GRMZM2G094655	methylglutaconyl-CoA hydratase
GRMZM2G174975	GDSL-like lipase/acylhydrolase
GRMZM2G178645	ATPase
GRMZM2G453805	PRm3 (chitinase)
GRMZM2G125032	beta-1,3-glucanase precursor (PR 2)
GRMZM2G046601	glutamine synthetase
GRMZM2G151992	haloacid dehalogenase-like hydrolase family protein
GRMZM2G074604	pal1 (phenylalanine ammonia lyase homolog1)
GRMZM2G118037	subtilisin N-terminal Region family protein
GRMZM2G314652	bifunctional 3-dehydroquinate dehydratase/shikimate dehydrogenase
GRMZM2G093092	5-pentadecatrienyl resorcinol O-methyltransferase-like
GRMZM2G423331	naringenin 7-O-methyltransferase-like
GRMZM2G052571	glutathione S-transferase
GRMZM2G029547	cinnamoyl-CoA reductase 1-like

Over representation analysis of genes with higher expression in the susceptible bulk.

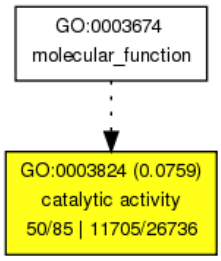


Figure B8: Over representation analysis of molecular function GO terms from genes with relatively higher expression in the susceptible bulk from the subset of RNA-seq genes which fall within the borders of the QTL. The significantly over represented GO terms are in coloured blocks. The FDR is indicated in brackets, and the number of genes associated with the GO terms in the test set is indicated at the bottom of each block, followed by the proportion in the reference set. The genes associated with the GO terms are listed in blocks next to the relevant GO terms. Solid lines with arrows indicate there are statistically significant over-represented GO terms in both connected boxes, whereas dashed lines indicate one box contains an over-represented GO term, and dotted lines connect boxes without over-represented GO terms.

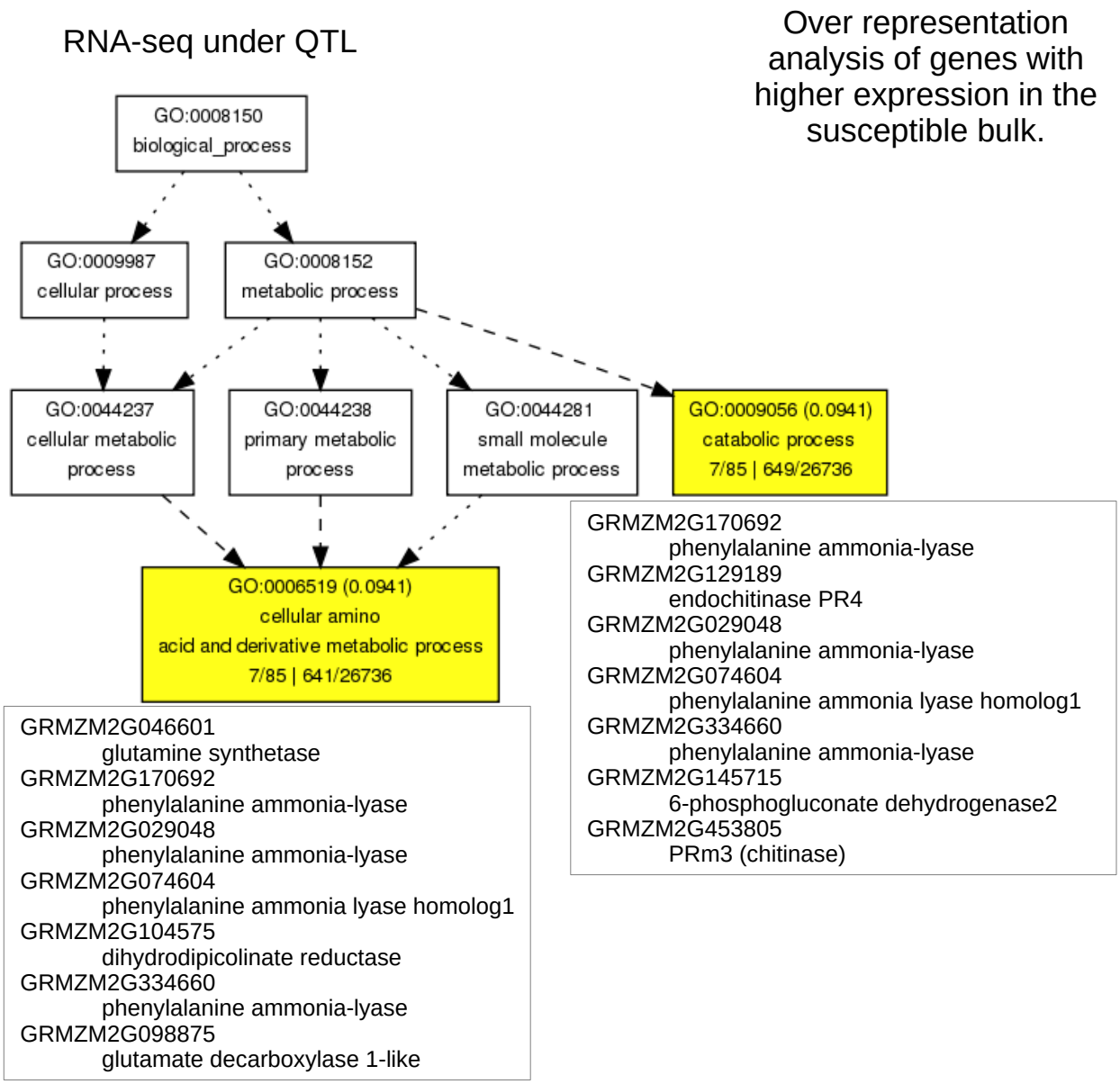


Figure B9: Over representation analysis of biological process GO terms from genes with relatively higher expression in the susceptible bulk from the subset of RNA-seq genes which fall within the borders of the QTL. The significantly over represented GO terms are in coloured blocks. The FDR is indicated in brackets, and the number of genes associated with the GO terms in the test set is indicated at the bottom of each block, followed by the proportion in the reference set. The genes associated with the GO terms are listed in blocks below the relevant GO terms. Solid lines with arrows indicate there are statistically significant over-represented GO terms in both connected boxes, whereas dashed lines indicate one box contains an over-represented GO term, and dotted lines connect boxes without over-represented GO terms.

Table B1: Top 100 genes with higher expression in the resistant bulk compared to the susceptible bulk as determined by microarray analysis.

Gene	Sequence description	FC[a] (R/S)	Adj P	Accession no.	description source
GRMZM2G373522	dehydrin	5.2	0.003	NP_001105327	BLAST
GRMZM5G859316	expansin precursor	4.4	0.006	LOC_Os03g44290.1	Rice[c]
GRMZM2G055698	Putative Rhomboid homologue	3.7	0.006	LOC_Os03g02530.1	Rice
GRMZM2G076972	hypothetical protein	3.2	0.006	NP_001145623	BLAST
GRMZM2G103771	mitochondrial import inner membrane translocase subunit Tim17	3.1	0.011	LOC_Os01g19770.1/ ADK88900	Rice/BLAST
GRMZM2G034843	hydroxyproline-rich glycoprotein family protein	2.9	0.009	LOC_Os09g16280.1	Rice
GRMZM2G429842	Subtilisin homologue	2.8	0.008	LOC_Os09g26920.1	Rice
GRMZM2G117164	Homeobox-transcription factor 41 (hb41)	2.7	0.008		MaizeGDB[d]
GRMZM2G144504	RHO guanyl-nucleotide exchange factor 7	2.7	0.022	AT5G02010.1	BLAST
GRMZM2G154747	AWPM-19-like membrane family protein	2.6	0.008	LOC_Os07g24000.1	Rice
GRMZM2G112238	Jacalin-like lectin domain containing protein	2.5	0.022	LOC_Os12g14440.1	Rice
GRMZM2G089836	invertase2 (ivr2)	2.5	0.010		MaizeGDB
GRMZM2G069911	histone H1 putative expressed	2.4	0.011	LOC_Os06g04020.1	Rice
GRMZM2G061450	hypothetical protein LOC100275244 [Zea mays]	2.4	0.016	NP_001142850	BLAST
GRMZM2G106344	hypothetical protein LOC100192725 [Zea mays]	2.4	0.022	NP_001131397	BLAST
GRMZM2G085086	LOC100285326 [Zea mays]	2.4	0.011	NP_001151691	BLAST
GRMZM2G028665	DNA-binding protein DSP1	2.3	0.012	LOC_Os01g50622.2	Rice
GRMZM2G165919	galactinol synthase 1 [Zea mays]	2.3	0.012	NP_001105748	BLAST
GRMZM2G427444	NADPH-dependent oxidoreductase	2.3	0.034	LOC_Os08g15248.1	Rice
AC203834.4_FG004	AP2 domain transcription factor	2.3	0.018	EF659468.1	BLAST
GRMZM2G073755	hypothetical protein LOC100191792 [Zea mays]	2.2	0.013	NP_001130689	BLAST
GRMZM2G059836	LOC100281644 [Zea mays]	2.2	0.022	NP_001148035	BLAST
GRMZM2G315431	ubiquitin-protein ligase [Zea mays]	2.2	0.020	NP_001147953	BLAST
GRMZM2G344630	GTP binding protein [Zea mays]	2.1	0.029	NP_001147495	BLAST
GRMZM2G012891	MTD1 [Zea mays]	2.1	0.023	NP_001148854	BLAST
GRMZM2G338160	hypothetical protein LOC100276191 [Zea mays]	2.0	0.022	ACG46856	BLAST
GRMZM2G095964	hypothetical protein SORBIDRAFT_03g007210 [Sorghum bicolor]	2.0	0.022	XP_002457433	BLAST
GRMZM2G060583	LRR and NB-ARC domains-containing disease resistance protein	2.0	0.032	AT3G14460.1	TAIR[b]
GRMZM2G043338	auxin-repressed protein	2.0	0.021	LOC_Os03g22270.1	Rice
GRMZM2G008053	Acyl-CoA synthetase/AMP-acid ligase II [Magnetospirillum magneticum AMB-1]	2.0	0.022	YP_421963	BLAST
GRMZM2G148090	hypothetical protein LOC100272340 [Zea mays]	2.0	0.033	NP_001140295	BLAST
GRMZM2G095968	LOC100285844 [Zea mays]	2.0	0.022	NP_001152206	BLAST
GRMZM2G017110	glutamate decarboxylase [Zea mays]	2.0	0.022	NP_001150761	BLAST
GRMZM2G122276	glycine-rich cell wall structural protein precursor	2.0	0.022	NP_001151740/ LOC_Os02g37490.1	BLAST/Rice
GRMZM5G817173	cytokinin oxidase 4 (cko4)	2.0	0.022		MaizeGDB

Gene	Sequence description	FC[a] (R/S)	Adj P	Accession no.	description source
GRMZM5G869299	hypothetical protein LOC100276066 [Zea mays]	2.0	0.022	NP_001143419	BLAST
GRMZM2G414813	hypothetical protein LOC100272944 [Zea mays]	2.0	0.023	NP_001140868	BLAST
GRMZM2G371462	hypothetical protein OsI_33234 [Oryza sativa Indica Group]	2.0	0.022	EAY78187	BLAST
GRMZM2G106819	hypothetical protein LOC100381528 [Zea mays]	2.0	0.029	NP_001167828	BLAST
GRMZM2G314328	cytochrome c biogenesis FN [Zea mays subsp. mays]	2.0	0.022	YP_588359	BLAST
GRMZM2G165919	galactinol synthase 1 [Zea mays]	2.0	0.031	NP_001105748	BLAST
GRMZM2G060183	hypothetical protein LOC100279532 [Zea mays]	2.0	0.035	NP_001146002	BLAST
GRMZM2G092137	hypothetical protein LOC100383382 [Zea mays]	1.9	0.022	NP_001169508	BLAST
GRMZM2G439201	elongation factor protein	1.9	0.023	LOC_Os07g46750.1	Rice
GRMZM5G837563	hypothetical protein SORBIDRAFT_05g027340 [Sorghum bicolor]	1.9	0.022	XP_002450037	BLAST
GRMZM2G091790	maturase K [Chasmanthium latifolium]	1.9	0.035	ADN86071	BLAST
GRMZM2G163849	SNF2 domain/helicase domain-containing protein-like [Oryza sativa Japonica Group]	1.9	0.033	BAD07923	BLAST
GRMZM2G111482	hypothetical protein SORBIDRAFT_09g026720 [Sorghum bicolor]	1.9	0.037	XP_002441440	BLAST
GRMZM2G346455	hypothetical protein LOC100274554 [Zea mays]	1.9	0.030	NP_001142381	BLAST
GRMZM2G309847	arabinogalactan peptide 23 precursor	1.9	0.039	LOC_Os06g21410.1	Rice
GRMZM2G024733	hypothetical protein LOC100304285 [Zea mays]	1.9	0.029	NP_001159199	BLAST
GRMZM2G063126	hypothetical protein LOC100382495 [Zea mays]	1.9	0.029	NP_001168703	BLAST
GRMZM2G106650	hypothetical protein LOC100382246 [Zea mays]	1.9	0.028	NP_001168470	BLAST
GRMZM2G322817	hypothetical protein LOC100274866 [Zea mays]	1.9	0.029	NP_001142599	BLAST
GRMZM2G148074	Homeobox-leucine zipper protein HOX1	1.9	0.030	Q40691	BLAST
GRMZM2G417107	hypothetical protein LOC100275319 [Zea mays]	1.8	0.029	NP_001142896	BLAST
GRMZM2G028685	hypothetical protein LOC100191401 [Zea mays]	1.8	0.031	NP_001130307	BLAST
GRMZM2G149422	phi-1-like phosphate-induced protein [Zea mays]	1.8	0.029	ACG26284	BLAST
GRMZM2G027120	zinc finger, C3HC4 type domain containing protein	1.8	0.036	LOC_Os03g20870.1	Rice
GRMZM2G128934	Ubiquitin-specific protease family C19-related protein	1.8	0.029	AT1G78880.1	TAIR
GRMZM2G012119	hypothetical protein LOC100272863 [Zea mays]	1.8	0.031	NP_001140788	BLAST
GRMZM2G133698	hypothetical protein SORBIDRAFT_10g001120 [Sorghum bicolor]	1.8	0.029	XP_002437709	BLAST
GRMZM2G466667	hypothetical protein SORBIDRAFT_09g023660 [Sorghum bicolor]	1.8	0.034	XP_002441277	BLAST
GRMZM2G059799	hypothetical protein LOC100304338 [Zea mays]	1.8	0.029	NP_001159248	BLAST
GRMZM2G006341	kinesin motor domain containing protein, expressed	1.8	0.030	ABF95490	BLAST/Rice

Gene	Sequence description	FC[a] (R/S)	Adj P	Accession no.	description source
GRMZM2G017266	hypothetical protein LOC100275677 [Zea mays]	1.8	0.030	NP_001143178	BLAST
GRMZM5G800598	hypothetical protein SORBIDRAFT_04g037100 [Sorghum bicolor]	1.8	0.046	XP_002453034	BLAST
GRMZM2G340656	alkaline alpha galactosidase 1 [Zea mays]	1.8	0.032	NP_001105793	BLAST
GRMZM2G084958	protochlorophyllide reductase A [Zea mays]	1.8	0.031	NP_001167683	BLAST
GRMZM5G832989	LOC100285189 [Zea mays]	1.8	0.031	NP_001151555	BLAST
GRMZM2G061932	hypothetical protein LOC100274420 [Zea mays]	1.8	0.032	NP_001142251	BLAST
GRMZM2G116086	Os01g0772600 [Oryza sativa Japonica Group]	1.8	0.031	NP_001044393	BLAST
GRMZM2G038512	lysine ketoglutarate reductase trans-splicing related 1	1.8	0.032	LOC_Os01g69050.1	Rice
GRMZM2G081541	hypothetical protein SORBIDRAFT_08g012980 [Sorghum bicolor]	1.8	0.031	XP_002442114	BLAST
GRMZM2G000812	hypothetical protein LOC100382256 [Zea mays]	1.8	0.031	NP_001168479	BLAST
GRMZM2G472167	peptide transporter PTR2 [Zea mays]	1.8	0.039	NP_001147986	BLAST
GRMZM2G060257	ATPP2-A13 [Zea mays]	1.8	0.032	NP_001149956	BLAST
GRMZM2G147974	hypothetical protein SORBIDRAFT_10g003460 [Sorghum bicolor]	1.8	0.031	XP_002437838	BLAST
GRMZM2G172214	brown planthopper-induced resistance protein 1	1.8	0.031	AAQ54305	BLAST
GRMZM2G067315	LOC100284861 [Zea mays]	1.8	0.031	NP_001151228	BLAST
GRMZM2G110504	hypothetical protein LOC100303825 [Zea mays]	1.8	0.034	NP_001158925	BLAST
GRMZM2G351977	hypothetical protein LOC100274453 [Zea mays]	1.8	0.031	NP_001142284	BLAST
GRMZM2G055273	membrane associated DUF588 domain containing protein	1.8	0.034	LOC_Os06g44610.1	Rice
GRMZM2G107226	hypothetical protein LOC100278578 [Zea mays]	1.8	0.031	NP_001145282	BLAST
GRMZM2G150688	LOC100284720 [Zea mays]	1.8	0.034	NP_001151087	BLAST
GRMZM2G058491	hypothetical protein SORBIDRAFT_03g001130 [Sorghum bicolor]	1.8	0.032	XP_002454901	BLAST
GRMZM2G142409	hypothetical protein LOC100191184 [Zea mays]	1.8	0.032	NP_001130091	BLAST
GRMZM2G105387	MADS-box transcription factor 26 [Zea mays]	1.8	0.032	NP_001148873	BLAST
GRMZM2G040095	lipoxygenase6 [Zea mays]	1.7	0.047	NP_001105976	BLAST
GRMZM2G179147	hypothetical protein LOC100383693 [Zea mays]	1.7	0.033	NP_001169802	BLAST
GRMZM2G330049	CIPK-like protein 1 [Zea mays]	1.7	0.033	NP_001152211	BLAST
GRMZM2G039841	hypothetical protein LOC100275488 [Zea mays]	1.7	0.040	NP_001143022	BLAST
GRMZM2G010406	hypothetical protein LOC100274106 [Zea mays]	1.7	0.036	NP_001141957	BLAST
GRMZM2G028535	LOC100280719 [Zea mays]	1.7	0.034	NP_001147111	BLAST
GRMZM2G401511	plant-specific domain TIGR01627 family protein [Zea mays]	1.7	0.048	NP_001151168	BLAST
GRMZM2G388855	hypothetical protein SORBIDRAFT_09g025290 [Sorghum bicolor]	1.7	0.036	XP_002440061	BLAST

Gene	Sequence description	FC[a] (R/S)	Adj P	Accession no.	description source
GRMZM2G402564	metallothionein-like protein type 4 [Hordeum vulgare subsp. vulgare]	1.7	0.039	CAD88267	BLAST
GRMZM5G831577	LOC100280792 [Zea mays]	1.7	0.034	NP_001147185	BLAST
GRMZM2G110922	CAMK_CAMK_like.11 - CAMK includes calcium/calmodulin dependent protein kinases	1.7	0.035	LOC_Os01g64970.1	Rice
GRMZM5G814451	hypothetical protein [Zea mays]	1.7	0.044	ACG28093	BLAST

[a] FC: Fold change
[b] TAIR (https://www.arabidopsis.org/)
[c] Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/)
[d] MaizeGDB (http://maizegdb.org/)

Table B2: Top 100 genes with higher expression in the susceptible bulk compared to the resistant bulk as determined by microarray analysis.

Gene	Sequence description	FC[a] (S/R)	Adj P	Accession no.	description source
GRMZM2G427815	peroxidase precursor putative expressed	3.5	0.006	LOC_Os07g48030.1	Rice[c]
GRMZM2G311036	benzoxazinone synthesis10 (bx10)	2.9	0.023		MaizeGDB[d]
GRMZM2G098577	eukaryotic initiation factor iso-4F subunit p82-34	2.9	0.008	LOC_Os04g42140.1	Rice
GRMZM5G842645	Putative lysine decarboxylase family protein	2.7	0.016	AT5G06300.1	TAIR[b]
GRMZM2G050450	transferase family protein	2.7	0.008	LOC_Os06g49660.1	Rice
GRMZM2G071390	Cupin domain containing protein/germin-like protein	2.6	0.008	LOC_Os12g05860.1/ ACJ64505	Rice/BLAST
GRMZM2G158097	hypothetical protein	2.5	0.009	NP_001167975.1	BLAST
GRMZM2G021598	OTU-like cysteine protease family protein	2.4	0.012	LOC_Os09g31280.1	Rice
GRMZM2G001572	unknown	2.4	0.012	ACF82444.1	TAIR
GRMZM2G136508	amino acid permease	2.4	0.011	LOC_Os04g35540.1/ NP_001148156	Rice/BLAST
GRMZM5G827496	hypothetical protein LOC100274435 [Zea mays]	2.3	0.017	NP_001142266	BLAST
GRMZM2G047368	plasma membrane intrinsic protein2	2.3	0.012		MaizeGDB
GRMZM2G176430	hypothetical protein LOC100383391 [Zea mays]	2.3	0.021	NP_001169517	BLAST
GRMZM2G125653	WRKY DNA-binding protein [Zea mays]	2.3	0.012	NP_001120723	BLAST
GRMZM2G353076	zinc finger homeodomain protein 1 [Zea mays]	2.2	0.015	NP_001152541	BLAST
GRMZM2G141386	hypothetical protein LOC100191956 [Zea mays]	2.2	0.027	NP_001130852	BLAST
GRMZM2G046601	glutamine synthetase	2.2	0.048	LOC_Os03g12290.1	Rice
GRMZM2G084779	Potassium transporter 5	2.2	0.021	Q5JK32	BLAST
GRMZM2G069176	ZOS3-12 - C2H2 zinc finger protein	2.2	0.015	LOC_Os03g32230	Rice
GRMZM2G016241	bronze-2 protein [Zea mays]	2.2	0.023	AAV64188	BLAST
GRMZM2G544469	-	2.2	0.027	-	BLAST
GRMZM2G018673	hypothetical protein LOC100277448 [Zea mays]	2.1	0.019	NP_001144480	BLAST
GRMZM2G053206	hypothetical protein LOC100272880 [Zea mays]	2.1	0.049	NP_001140805	BLAST
GRMZM5G808811	DAG protein [Zea mays]	2.1	0.017	NP_001150208	BLAST
GRMZM2G160069	hypothetical protein LOC100383930 [Zea mays]	2.1	0.019	NP_001170020	BLAST
GRMZM2G170692	phenylalanine ammonia-lyase	2.1	0.019	NM_001174615.1	BLAST
GRMZM2G175606	-	2.1	0.022	-	BLAST
GRMZM5G822593	lipoxygenase chloroplast precursor	2.1	0.022	LOC_Os08g39850.1	Rice
GRMZM2G058966	hypothetical protein LOC100273812 [Zea mays]	2.1	0.022	NP_001141685	BLAST
GRMZM2G130927	MAPK/ERK kinase kinase 1	2.1	0.044	AT4G08500.1	TAIR
GRMZM5G878558	nitrate reductase putative	2.1	0.032	LOC_Os02g53130.1	Rice
GRMZM2G466563	putative calmodulin-binding protein [Oryza sativa Japonica Group]	2.1	0.025	BAD03817	BLAST
GRMZM2G329069	expressed protein	2.0	0.021	LOC_Os01g66544.1	Rice
GRMZM2G170734	chlorophyllase1 (chph1)	2.0	0.021		MaizeGDB
GRMZM2G085974	hypothetical protein LOC100382861 [Zea mays]	2.0	0.021	NP_001169028	BLAST
GRMZM2G137954	hypothetical protein LOC100273706 [Zea mays]	2.0	0.049	NP_001141590	BLAST
GRMZM2G017186	periplasmic beta-glucosidase precursor	2.0	0.022	LOC_Os03g53860.1	Rice

Gene	Sequence description	FC[a] (S/R)	Adj P	Accession no.	description source
GRMZM2G356618	uncharacterized	2.0	0.022	XR_559425.1	BLAST
GRMZM2G106213	ADP-glucose pyrophosphorylase small subunit [Zea mays]	2.0	0.025	NP_001105178	BLAST
GRMZM2G164141	hypothetical protein LOC100274529 [Zea mays]	2.0	0.023	NP_001142358	BLAST
GRMZM2G001572	expressed protein	2.0	0.022	LOC_Os10g33710.1	Rice
GRMZM2G408038	serine carboxypeptidase-like 7	2.0	0.025	AT3G10450.1	TAIR
GRMZM2G125923	hypothetical protein LOC100279620 [Zea mays]	2.0	0.031	NP_001146088	BLAST
GRMZM2G041699	cytokinin-O-glucosyltransferase 2 [Zea mays]	1.9	0.027	NP_001149205	BLAST
GRMZM2G145460	NAD dependent epimerase/dehydratase family domain containing protein	1.9	0.022	LOC_Os04g52730.1	Rice
GRMZM2G059502	hypothetical protein LOC100192597 [Zea mays]	1.9	0.040	NP_001131284	BLAST
GRMZM2G109668	hypothetical protein [Zea mays]	1.9	0.031	ACG30711	BLAST
GRMZM2G300945	LOC100284874 [Zea mays]	1.9	0.029	NP_001151241	BLAST
GRMZM2G137535	lichenase-2 [Zea mays]	1.9	0.036	NP_001148461	BLAST
GRMZM2G466563	putative calmodulin-binding protein [Oryza sativa Japonica Group]	1.9	0.036	BAD03817	BLAST
GRMZM2G161293	BGGP Beta-1-3-galactosyl-O-glycosyl-glycoprotein [Zea mays]	1.9	0.027	NP_001152358	BLAST
GRMZM2G125482	flavonoid 3-monooxygenase [Zea mays]	1.9	0.034	NP_001147572	BLAST
GRMZM2G335019	hypothetical protein LOC100279036 [Zea mays]	1.9	0.038	NP_001145572	BLAST
GRMZM2G042639	protein IN2-1 [Zea mays]	1.9	0.036	NP_001105433	BLAST
GRMZM2G108115	endoplasmic oxidoreductin	1.9	0.029	NP_001148525/ LOC_Os03g52340.2	BLAST/Rice
AC234528.1_FG004	hypothetical protein LOC100277774 [Zea mays]	1.9	0.025	NP_001144730	BLAST
GRMZM2G085974	hypothetical protein LOC100382861 [Zea mays]	1.9	0.025	NP_001169028	BLAST
GRMZM2G473266	major facilitator transporter [Bacillus weihenstephanensis KBAB4]	1.9	0.029	YP_001645820	BLAST
GRMZM2G076468	hypothetical protein SORBIDRAFT_01g028810	1.9	0.029	XP_002467475	BLAST
GRMZM2G349749	hypothetical protein LOC100272901 [Zea mays]	1.9	0.028	NP_001140826	BLAST
GRMZM2G032977	LOC100282067 [Zea mays]	1.9	0.028	NP_001148452	BLAST
GRMZM2G374971	zeamatin precursor [Zea mays]	1.9	0.027	NP_001105356	BLAST
GRMZM2G001332	uncharacterized	1.9	0.032	XR_555125.1	BLAST
GRMZM2G131055	hypothetical protein LOC100279381 [Zea mays]	1.9	0.031	NP_001145867	BLAST
GRMZM2G333980	hypothetical protein LOC100279844 [Zea mays]	1.8	0.028	NP_001146269	BLAST
GRMZM2G024211	hypothetical protein LOC100193276 [Zea mays]	1.8	0.036	NP_001131893	BLAST
GRMZM2G078396	unknown [Zea mays]	1.8	0.029	ACR34433	BLAST
GRMZM2G061806	hypothetical protein LOC100383097 [Zea mays]	1.8	0.031	NP_001169239	BLAST
GRMZM2G060919	hypothetical protein OsI_08565 [Oryza sativa Indica Group]	1.8	0.029	EEC73831	BLAST
GRMZM2G053652	40S ribosomal protein S7	1.8	0.036	ACG35364/LOC_Os03g18570.1	BLAST/Rice
GRMZM2G062396	hypothetical protein LOC100191369 [Zea mays]	1.8	0.029	NP_001130275	BLAST
GRMZM2G176301	C-4 methylsterol oxidase [Zea mays]	1.8	0.029	NP_001148435	BLAST
GRMZM2G095280	indole-3-acetate beta-glucosyltransferase	1.8	0.038	ACG40098	BLAST

Gene	Sequence description	FC[a] (S/R)	Adj P	Accession no.	description source
	[Zea mays]				
GRMZM2G071630	glyceraldehyde-3-phosphate dehydrogenase, cytosolic 3 [Zea mays]	1.8	0.031	NP_001105385	BLAST
GRMZM2G456835	LOC100282352 [Zea mays]	1.8	0.030	NP_001148736	BLAST
GRMZM2G141642	hypothetical protein SORBIDRAFT_04g034650 [Sorghum bicolor]	1.8	0.030	XP_002452903	BLAST
GRMZM2G446201	hypothetical protein LOC100274970 [Zea mays]	1.8	0.032	NP_001142675	BLAST
GRMZM2G075978	unknown [Zea mays]	1.8	0.031	ACL54597	BLAST
GRMZM2G051135	endothelial differentiation-related factor 1 [Zea mays]	1.8	0.031	NP_001151413	BLAST
GRMZM2G099619	hypothetical protein LOC100279509 [Zea mays]	1.8	0.031	NP_001145981	BLAST
GRMZM2G124477	hypothetical protein SORBIDRAFT_04g035080 [Sorghum bicolor]	1.8	0.031	XP_002452927	BLAST
GRMZM2G104847	LOC100281785 [Zea mays]	1.8	0.033	NP_001148177	BLAST
GRMZM2G170016	cytochrome b5-like Heme/Steroid binding domain containing protein	1.8	0.031	LOC_Os02g42740.1	Rice
GRMZM2G429617	hypothetical protein LOC100272958 [Zea mays]	1.8	0.032	NP_001140882	BLAST
GRMZM2G015409	RING-H2 finger protein ATL5F [Zea mays]	1.8	0.031	NP_001151780	BLAST
GRMZM2G478568	nicotianamine synthase 3 [Zea mays]	1.8	0.031	BAB91326	BLAST
GRMZM2G113052	negatively light-regulated protein [Zea mays]	1.8	0.031	NP_001148326	BLAST
GRMZM2G076006	hypothetical protein LOC100191237 [Zea mays]	1.8	0.035	NP_001130143	BLAST
GRMZM2G027439	BRASSINOSTEROID INSENSITIVE 1-associated receptor kinase 1 [Zea mays]	1.8	0.031	NP_001150753	BLAST
GRMZM2G092804	loricrin [Zea mays]	1.8	0.037	NP_001147561	BLAST
AC207043.3_FG002	VQ motif family protein [Zea mays]	1.8	0.033	NP_001147304	BLAST
GRMZM2G030293	40S ribosomal protein S28	1.8	0.036	LOC_Os10g27174.1	Rice
GRMZM2G014382	hypothetical protein LOC100194239 [Zea mays]	1.7	0.036	NP_001132752	BLAST
GRMZM2G162622	THION18 - Plant thionin family protein precursor, expressed	1.7	0.033	LOC_Os01g41140.1	Rice
GRMZM2G106950	hypothetical protein LOC100382674 [Zea mays]	1.7	0.033	NP_001168869	BLAST
GRMZM5G841900	3-5 exoribonuclease CSL4 [Zea mays]	1.7	0.033	NP_001148354	BLAST
GRMZM2G024476	cytokinin dehydrogenase 4b [Zea mays]	1.7	0.034	NP_001185960	BLAST
GRMZM2G447795	xylanase inhibitor protein 1 [Zea mays]	1.7	0.045	NP_001151661	BLAST
GRMZM2G042636	tubulin beta-5 chain [Zea mays]	1.7	0.034	NP_001105458	BLAST
GRMZM2G060866	anther-specific proline-rich protein APG [Zea mays]	1.7	0.034	NP_001150585	BLAST

[a] FC: Fold change
[b] TAIR (https://www.arabidopsis.org/)
[c] Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/)
[d] MaizeGDB (http://maizegdb.org/)

Table B3: Top 100 genes with higher expression in the resistant bulk compared to the susceptible bulk as determined by RNA-seq analysis.

Gene	Sequence description	FC[a] (R/S)	Q value	Accession no.	description source
GRMZM2G020556	hypothetical protein	30.6	0.000	DAA50064.1	BLAST
GRMZM2G020514	hypothetical protein	26.8	0.000	NP_001144140.1	BLAST
GRMZM2G092804	loricrin isoform X3	23.8	0.000	XP_008654907.1	BLAST
GRMZM2G026364	hypothetical protein	18.2	0.000	NM_001176560.1	BLAST
GRMZM2G062527	GASR5 - Gibberellin-regulated GASA/GAST/Snakin family protein	17.1	0.000	LOC_Os05g31280.1	Rice[c]
GRMZM2G109127	uncharacterized transcript	16.5	0.000	XM_008658471.1	BLAST
GRMZM2G087558	uncharacterized transcript	15.1	0.000	XM_008658471.1	BLAST
GRMZM5G822449	hypothetical protein	12.4	0.000	EU976680.1	BLAST
GRMZM2G036861	chorismate synthase 2 chloroplast precursor	9.8	0.000	LOC_Os03g14990.1	Rice
GRMZM2G547449	uncharacterized transcript	9.6	0.000	XM_008658471.1	BLAST
GRMZM2G373174	uncharacterized transcript	9.5	0.000	XM_008658471.1	BLAST
GRMZM2G523549	Setaria italica NADP-specific glutamate dehydrogenase-like	9.0	0.000	XR_215235.1	BLAST
GRMZM2G568748	uncharacterized transcript	8.7	0.000	XR_557506.1	BLAST
GRMZM2G097756	uncharacterized transcript	8.5	0.000	XM_008652827.1	BLAST
GRMZM5G851617	uncharacterized transcript	8.2	0.000	JQ887800.1	BLAST
GRMZM2G388285	hypothetical protein	7.2	0.000	EU974605.1	BLAST
GRMZM2G157990	tesmin/TSO1-like CXC 5	7.2	0.000	XR_553665.1	BLAST
GRMZM2G077317	monocopper oxidase	7.0	0.000	LOC_Os08g05820.1	Rice
GRMZM2G031354	LTPL124 - Protease inhibitor/seed storage/LTP family protein	7.0	0.000	LOC_Os04g52260.1	Rice
GRMZM2G000633	(CRK10, RLK4) cysteine-rich RLK (RECEPTOR-like protein kinase) 10	6.6	0.000	AT4G23180.1	TAIR[b]
GRMZM2G113415	LOC100285146 [Zea mays]	6.3	0.000	NP_001151512	BLAST
GRMZM2G071145	sphingosine kinase [Lotus japonicus]	6.3	0.000	BAD86587	BLAST
GRMZM2G164486	-	6.3	0.006		
GRMZM2G037705	-	6.2	0.000		
GRMZM5G899119	hypothetical protein SORBIDRAFT_01g048810 [Sorghum bicolor]	6.2	0.000	XP_002465956	BLAST
GRMZM2G374302	arginine decarboxylase [Zea mays]	6.2	0.000	ACG41098	BLAST
GRMZM2G332348	protein binding protein [Oryza sativa Indica Group]	6.0	0.000	ABR26196	BLAST
GRMZM2G445139	-	5.9	0.000		
GRMZM2G465453	-	5.8	0.000		
GRMZM2G439799*	Leucine-rich repeat protein kinase family protein	5.8	0.000	AT3G47570.1	TAIR
GRMZM2G102912	AIG2-like protein [Zea mays]	5.8	0.000	NP_001150494	BLAST
GRMZM5G876112	forkhead box protein D4 [Mus musculus]	5.7	0.000	NP_032048	BLAST
GRMZM2G465860	-	5.7	0.000		
GRMZM2G114017	hypothetical protein [Zea mays]	5.6	0.000	ACG26965	BLAST
GRMZM2G153178	hypothetical protein SORBIDRAFT_-4g--234- [Sorghum bicolor]	5.3	0.000	XP_002451462	BLAST
GRMZM2G172642	LOC1--285611 [Zea mays]	5.2	0.000	NP_001151974	BLAST
GRMZM2G157858	-	5.2	0.002		
GRMZM2G413999	-	5.1	0.000		
GRMZM2G099867	-	5.1	0.001		
GRMZM2G031102	-	5.1	0.001		
GRMZM2G119407	-	4.9	0.001		

Gene	Sequence description	FC[a] (R/S)	Q value	Accession no.	description source
GRMZM2G124927	retrotransposon protein [Zea mays]	4.7	0.000	NP_001152059	BLAST
GRMZM5G844514	-	4.7	0.000		
GRMZM2G038557	-	4.6	0.005		
GRMZM2G425920*	transcription factor HY5 [Zea mays]	4.5	0.000	ACG48760	BLAST
GRMZM2G138291	-	4.5	0.000		
GRMZM2G163796	-	4.4	0.000		
AC205376.4_FG005	-	4.4	0.020		
GRMZM2G022613	hypothetical protein OsI_-3933 [Oryza sativa Indica Group]	4.4	0.000	EEC71568	BLAST
GRMZM2G575305	-	4.3	0.000		
GRMZM2G160569	hypothetical protein LOC1--274-42 [Zea mays]	4.3	0.000	NP_001141895	BLAST
GRMZM2G176585	hypothetical protein SORBIDRAFT_-1g-4881- [Sorghum bicolor]	4.2	0.012	XP_002465956	BLAST
GRMZM2G057208	-	4.2	0.000		
GRMZM5G819130	-	4.2	0.000		
GRMZM2G430455	-	4.1	0.000		
GRMZM2G305996	hypothetical protein LOC1--383269 [Zea mays]	4.1	0.021	NP_001169400	BLAST
GRMZM2G121928	-	4.1	0.000		
GRMZM2G054029	-	4.0	0.029		
GRMZM2G027510	-	4.0	0.008		
GRMZM2G319573	hypothetical protein SORBIDRAFT_-4g--3-7- [Sorghum bicolor]	4.0	0.000	XP_002453277	BLAST
GRMZM2G163582	-	4.0	0.002		
GRMZM2G518604	-	3.9	0.024		
GRMZM2G425594	cation efflux protein/zinc transporter [Ajellomyces dermatitidis ER-3]	3.9	0.001	EEQ87344	BLAST
GRMZM2G119773	LOC1--283321 [Zea mays]	3.9	0.000	NP_001149695	BLAST
GRMZM2G121176	seed specific protein Bn15D14A [Zea mays]	3.8	0.015	ACG40639	BLAST
GRMZM2G133718	Putative Serine Carboxypeptidase	3.8	0.005	LOC_Os01g06490.2	Rice
GRMZM2G132882	-	3.8	0.000		
GRMZM5G801427	-	3.8	0.000		
GRMZM2G434572	-	3.8	0.010		
GRMZM2G066961	-	3.8	0.011		
GRMZM2G504565	-	3.8	0.032		
GRMZM2G410352	hypothetical protein [Zea mays]	3.7	0.002	ACG38858	BLAST
GRMZM2G001386	-	3.7	0.007		
GRMZM2G322817	hypothetical protein LOC1--277222 [Zea mays]	3.7	0.000	NP_001144327	BLAST
GRMZM2G566149	-	3.7	0.044		
GRMZM2G112219	-	3.7	0.008		
GRMZM2G109201	LOC1--281158 [Zea mays]	3.7	0.000	NP_001147549	BLAST
GRMZM2G179342	-	3.6	0.000		
GRMZM2G156299	uncharacterized protein	3.6	0.016	XR_560607.1	BLAST
GRMZM2G168382	cytochrome P45- [Triticum aestivum]	3.6	0.000	BAB40322	BLAST
GRMZM2G362819	-	3.6	0.001		
GRMZM2G139399	-	3.5	0.017		
GRMZM2G069290	-	3.4	0.000		
GRMZM2G125304	hypothetical protein LOC1--384273 [Zea mays]	3.4	0.002	NP_001170309	BLAST
GRMZM2G027983	-	3.4	0.000		
GRMZM2G108103	polyphenol oxidase	3.3	0.001	LOC_Os01g58100.1	Rice

Gene	Sequence description	FC[a] (R/S)	Q value	Accession no.	description source
GRMZM2G020661*	GTPase/ ras-related protein Rab11B [Zea mays]	3.3	0.025	NP_001151123	BLAST
GRMZM2G311151	-	3.3	0.029		
GRMZM2G090176	-	3.3	0.011		
GRMZM2G174430	-	3.3	0.027		
GRMZM2G351977	chlorophyll a-b binding protein 2 [Zea mays]	3.3	0.000	NP_001148439	BLAST
GRMZM2G074496*	CC-NB-LRR protein/ stripe rust resistance protein Yr10	3.3	0.000	LOC_Os11g34920.1	Rice
GRMZM2G364988	aconitase2 [Zea mays]	3.2	0.000	NP_001147431	BLAST
GRMZM2G425583	-	3.2	0.009		
GRMZM2G149289	hypothetical protein LOC1--383454 [Zea mays]	3.2	0.000	NP_001169574	BLAST
GRMZM2G026322	hypothetical protein SORBIDRAFT_-1g-3434- [Sorghum bicolor]	3.2	0.001	XP_002467800	BLAST
GRMZM2G345700	-	3.2	0.000		
GRMZM5G881529	hypothetical protein SORBIDRAFT_-1g-3698- [Sorghum bicolor]	3.1	0.005	XP_002467946	BLAST
GRMZM2G064371*	auxin-binding protein 4 precursor [Zea mays]	3.1	0.000	NP_001105353	BLAST
GRMZM5G872568	-	3.1	0.037		

[a] FC: Fold change
[b] TAIR (https://www.arabidopsis.org/)
[c] Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/)
* Genes indentified as candidates in the proposed resistance model (Figure 4.10)

Table B4: Top 100 genes with higher expression in the susceptible bulk compared to the resistant bulk as determined by RNA-seq analysis.

Gene	Sequence description	FC[a] (S/R)	Q value	Accession no.	description source
GRMZM2G136748	uncharacterized	334.1	0.037	XR_562557.1	BLAST
GRMZM2G174192	UDP-Glycosyltransferase superfamily protein	185.6	0.000	AT5G49690.1	TAIR[b]
GRMZM2G417954	nine-cis-epoxycarotenoid dioxygenase5 (nced5)	176.9	0.000		MaizeGDB[d]
GRMZM2G371375	late embryogenesis abundant protein group 3	174.7	0.000	LOC_Os04g52110.1	Rice[c]
GRMZM2G377613	transcription factor HBP-1b	141.9	0.005	LOC_Os05g48650.1	Rice
GRMZM2G041039	hypothetical protein	131.8	0.012	ACG48669	BLAST
GRMZM2G383125	DNA binding	127.2	0.000	AT3G47680.1	TAIR
GRMZM2G061450	uncharacterized protein	103.4	0.000	NP_001142850.1	BLAST
GRMZM2G376743	low-temperature-induced 65 kDa protein-like	101.9	0.000	XP_008644725.1	BLAST
GRMZM5G881353	apomucin-like	91.7	0.017	XM_008666551.1	BLAST
GRMZM2G434203	AP2-EREBP-transcription factor 10 (ereb10)	90.1	0.001		MaizeGDB
GRMZM2G425629	late embryogenesis abundant protein group 3	83.0	0.000	LOC_Os01g50910.2	Rice
GRMZM2G041493	hypothetical protein	80.6	0.000	EU960076.1	BLAST
GRMZM2G096475	lea protein group3 (mlg3)	69.7	0.000		MaizeGDB
GRMZM2G552956	hypothetical protein	68.8	0.000	AFW63695.1	BLAST
GRMZM5G831724	apomucin-like	59.2	0.000	XM_008666551.1	BLAST
GRMZM2G107570	DUF1264 domain containing protein	56.7	0.000	LOC_Os01g52830.1	Rice
GRMZM2G356618	uncharacterized	56.2	0.000	XR_559425.1	BLAST
GRMZM2G049930	Cupin domain containing protein	44.9	0.000	LOC_Os08g35750.1	Rice
GRMZM2G063287	Cupin domain containing protein	44.5	0.000	LOC_Os12g05860.1	Rice
GRMZM2G428040	-	43.6	0.003		
GRMZM2G361984	-	38.6	0.000		
GRMZM2G005825	-	33.2	0.009		
GRMZM2G180328	putative no apical meristem (NAM) protein [Oryza sativa Japonica Group]	32.9	0.000	AAT44250	BLAST
GRMZM2G305446	aquaporin TIP3-1 [Zea mays]	32.9	0.000	NP_001105032	BLAST
GRMZM2G067600	uncharacterized	32.4	0.000	XM_008677071.1	BLAST
GRMZM2G081458	NB-ARC domain-containing disease resistance protein	31.0	0.000	AT3G14470.1	TAIR
GRMZM2G018806	-	29.1	0.000		
GRMZM2G376061	-	28.5	0.000		
GRMZM2G094375	laccase family protein, putative [Oryza sativa Japonica Group]	28.2	0.000	ABA97328	BLAST
GRMZM2G448627	-	27.4	0.000		
GRMZM2G087875	cytochrome P450 CYP81A1 [Zea mays]	26.3	0.000	ACG29835	BLAST
GRMZM2G103771	putative stress-inducible membrane pore protein [Oryza sativa Japonica Group]	26.1	0.000	BAB20636	BLAST
GRMZM2G154747	plasma membrane associated protein [Zea mays]	25.1	0.000	ACG24926	BLAST
GRMZM2G178817	-	23.8	0.000		
GRMZM2G129761	1-Cys peroxiredoxin PER1 [Zea mays]	22.9	0.000	NP_001105998	BLAST
GRMZM2G102183	malate synthase, glyoxysomal [Zea mays]	22.0	0.000	NP_001105328	BLAST
AC233851.1_FG010	-	21.8	0.000		
GRMZM2G117378	hypothetical protein SORBIDRAFT_07g023200 [Sorghum]	21.7	0.000	XP_002445638	BLAST

Gene	Sequence description	FC[a] (S/R)	Q value	Accession no.	description source
AC205471.4_FG007	-	21.2	0.000		
GRMZM2G083855	uncharacterized	20.8	0.000	NM_001151479.1	BLAST
GRMZM2G001572	unknown [Zea mays]	19.6	0.000	ACF82444	BLAST
GRMZM2G071390	germin-like protein [Zea mays]	18.8	0.000	ACJ64505	BLAST
GRMZM2G320786	-	18.6	0.000		
GRMZM5G879851	-	18.2	0.016		
GRMZM2G169121	hypothetical protein LOC100194164 [Zea mays]	18.1	0.000	NP_001132686	BLAST
GRMZM2G053206	hypothetical protein LOC100272880 [Zea mays]	17.8	0.000	NP_001140805	BLAST
GRMZM5G878558	hypothetical protein SORBIDRAFT_04g034470 [Sorghum bicolor]	17.8	0.000	XP_002454625	BLAST
GRMZM2G383699	abscisic acid response protein [Cucumis melo]	17.1	0.000	AAL27560	BLAST
GRMZM2G175927	-	16.8	0.000		
GRMZM2G170734	hypothetical protein SORBIDRAFT_02g012300 [Sorghum bicolor]	16.7	0.000	XP_002459848	BLAST
GRMZM2G001853	hypothetical protein ARALYDRAFT_473001 [Arabidopsis lyrata subsp. lyrata]	16.7	0.000	XP_002893498	BLAST
GRMZM2G041699	cytokinin-O-glucosyltransferase 2 [Zea mays]	16.6	0.000	NP_001149205	BLAST
GRMZM2G127418	LOC100281319 [Zea mays]	16.3	0.000	NP_001147709	BLAST
GRMZM5G833332	hypothetical protein SORBIDRAFT_01g045980 [Sorghum bicolor]	15.9	0.000	XP_002465799	BLAST
GRMZM2G102550	-	15.6	0.000		
GRMZM5G805609	glucan endo-1,3-beta-glucosidase 7 [Zea mays]	15.3	0.000	NP_001149419	BLAST
GRMZM2G105348	heat shock factor protein 1 [Zea mays]	15.2	0.000	NP_001152657	BLAST
GRMZM2G036351	ZIM motif family protein [Zea mays]	15.1	0.000	NP_001149525	BLAST
GRMZM2G181135	-	15.1	0.000		
GRMZM2G542272	-	14.7	0.000		
AC212835.3_FG008	-	14.5	0.000		
GRMZM2G098875	glutamate decarboxylase 1-like	14.4	0.000	XM_004985025.1	BLAST
GRMZM5G894568	-	14.1	0.000		
GRMZM2G062724	hypothetical protein LOC100192089 [Zea mays]	14.0	0.000	NP_001130984	BLAST
GRMZM2G573083	-	13.9	0.024		
GRMZM2G166776	-	13.8	0.000		
GRMZM2G473266	major facilitator transporter [Bacillus weihenstephanensis KBAB4]	13.0	0.000	YP_001645820	BLAST
GRMZM2G466563	putative calmodulin-binding protein [Oryza sativa Japonica Group]	12.7	0.000	BAD03817	BLAST
GRMZM2G000236	12-oxophytodienoate reductase 2 [Zea mays]	12.5	0.000	ACG42962	BLAST
GRMZM2G047456	hypothetical protein LOC100382459 [Zea mays]	12.5	0.000	NP_001168671	BLAST
GRMZM2G148272	hypothetical protein LOC100274659 [Zea mays]	12.4	0.000	NP_001142456	BLAST
GRMZM2G336533	stress-induced transcription factor NAC1 [Oryza sativa Indica Group]	12.3	0.000	ACX71077	BLAST
GRMZM2G052571	glutathione S-transferase [Zea mays]	12.3	0.000	NP_001147759/LOC_Os03g39850.1	BLAST/Rice

Gene	Sequence description	FC[a] (S/R)	Q value	Accession no.	description source
GRMZM2G137535	lichenase-2 [Zea mays]	12.1	0.000	NP_001148461	BLAST
GRMZM2G006676	-	12.0	0.000		
GRMZM2G089557	-	11.6	0.000		
GRMZM2G043191	putative inositol polyphosphate 5-phosphatase [Oryza sativa Japonica Group]	11.5	0.000	AAW34241	BLAST
GRMZM2G314075	abscisic stress-ripening	11.5	0.000	LOC_Os01g72900.1	Rice
GRMZM2G131055	glycosyltransferase [Zea mays]	11.4	0.000	ACG27644	BLAST
GRMZM2G305280	unknown [Zea mays]	11.2	0.000	ACR37234	BLAST
GRMZM2G094510	hypothetical protein LOC100278059 [Zea mays]	11.1	0.000	NP_001144932	BLAST
GRMZM2G407223	hypothetical protein [Zea mays]	11.1	0.000	ACG23938	BLAST
GRMZM2G015912	-	11.0	0.000		
GRMZM2G450498	-	10.9	0.000		
GRMZM2G050234	hypothetical protein SORBIDRAFT_06g026350 [Sorghum bicolor]	10.8	0.000	XP_002448382	BLAST
GRMZM2G354909	unknown [Zea mays]	10.8	0.000	ACR38658	BLAST
GRMZM2G106622	ABA-responsive protein [Zea mays]	10.7	0.000	NP_001152088	BLAST
GRMZM2G472236	-	10.7	0.015		
GRMZM2G323888	-	10.7	0.000		
GRMZM2G023346	hypothetical protein LOC100279937 [Zea mays]	10.5	0.000	NP_001146359	BLAST
GRMZM2G447795	xylanase inhibitor protein 1 [Zea mays]	10.5	0.000	NP_001151661	BLAST
GRMZM2G457346	-	10.3	0.000		
GRMZM2G159768	-	10.2	0.000		
GRMZM5G892827	-	10.2	0.003		
GRMZM2G162505	chitinase 2 [Zea mays]	10.2	0.000	NP_001152001	BLAST
GRMZM2G169033	putative laccase [Zea mays]	10.1	0.000	NP_001105915	BLAST
GRMZM2G474755	polcalcin Jun o 2 [Zea mays]	10.0	0.000	NP_001146874	BLAST
GRMZM2G162359	-	9.9	0.000		
GRMZM2G146644	-	9.9	0.000		

[a] FC: Fold change
[b] TAIR (https://www.arabidopsis.org/)
[c] Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/)
[d] MaizeGDB (http://maizegdb.org/)

Table B5a: Differentially expressed genes with higher expression in R bulk observed in microarray analysis that coincide with QTL.

Gene	QTL	Sequence description	FC[a] (R/S)	adj. p-value	description source	BLAST/TAIR[b]/Rice[c] Accession No.
GRMZM2G128617	1Cz	transmembrane BAX inhibitor motif-containing protein	1.69	0.0417	BLAST/Rice	NP_001151352/ LOC_Os03g53400.1
GRMZM2G165709	1Cz	penicillin-binding protein 1C	1.60	0.0498	BLAST	ZP_01000924
GRMZM2G028665	3H_GLS	DNA-binding protein DSP1	2.33	0.0118	Rice	LOC_Os01g50622.2
AC203834.4_FG004	3H_GLS	AP2 domain transcription factor	2.27	0.0178	BLAST	EF659468.1
GRMZM2G338160	3H_GLS	hypothetical protein	2.05	0.0216	BLAST	ACG46856
GRMZM2G038512	3H_GLS	lysine ketoglutarate reductase trans-splicing related 1	1.79	0.0316	Rice	LOC_Os01g69050.1
GRMZM2G110922	3H_GLS	CAMK_CAMK_like.11 - CAMK includes calcium/calmodulin dependent protein kinases	1.72	0.0352	Rice	LOC_Os01g64970.1
GRMZM2G364528	3Cz_1.1	transcription factor bHLH148-like	1.66	0.0391	BLAST	XM_004969186.1
GRMZM2G122276	5Cz	glycine-rich cell wall structural protein precursor	2.00	0.0216	BLAST/Rice	NP_001151740/ LOC_Os02g37490.1
GRMZM2G117164	5Cz	homeobox associated leucine zipper	2.73	0.0083	BLAST/Rice	A2X7U1/ LOC_Os02g43330.1
GRMZM2G000829	6Cz	hypothetical protein	1.70	0.0469	BLAST	XP_002462577
GRMZM2G064679	6Cz/Lesion	hypothetical protein	1.66	0.0419	BLAST	NP_001144678
GRMZM2G388855	6Cz/Lesion	hypothetical protein	1.73	0.0361	BLAST	XP_002440061
GRMZM2G055273	9Cz_2.1	membrane associated DUF588 domain containing protein	1.77	0.0335	Rice	LOC_Os06g44610.1
GRMZM2G043338	9Cz_2.1	auxin-repressed protein	2.03	0.0208	Rice	LOC_Os03g22270.1
GRMZM2G027120	9Cz_2.1	zinc finger, C3HC4 type domain containing protein	1.83	0.0361	Rice	LOC_Os03g20870.1
GRMZM2G006341	9Cz_2.1	kinesin motor domain containing protein, expressed	1.81	0.0295	BLAST/Rice	ABF95490

Gene	QTL	Sequence description	FC[a] (R/S)	adj. p-value	description source	BLAST/TAIR[b]/Rice[c] Accession No.
GRMZM2G036861	9H_GLS	chorismate synthase 2	1.60	0.0498	BLAST/Rice	NP_001148583/ LOC_Os03g14990.1
GRMZM2G024571	9Cz_2.2	potyvirus VPg interacting protein	1.63	0.0447	Rice	LOC_Os03g11890.1
GRMZM2G128934	10Cz_2	Ubiquitin-specific protease family C19-related protein	1.83	0.0294	TAIR	AT1G78880.1
GRMZM2G107226	10Cz_2	hypothetical protein	1.77	0.0310	BLAST	NP_001145282

[a] FC: Fold change; higher expression in the R bulk compared to the S bulk
[b] TAIR (https://www.arabidopsis.org/)
[c] Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/)

Table B5b: Differentially expressed genes with higher expression in S bulk observed in microarray analysis that coincide with QTL.

Gene	QTL	Sequence description	FC[a] (S/R)	adj. p-value	description source	BLAST/TAIR[b]/Rice[c] Accession No.
GRMZM2G108115	1Cz	endoplasmic oxidoreductin	1.89	0.0294	BLAST/Rice	NP_001148525/ LOC_Os03g52340.2
GRMZM2G017186	1Cz	periplasmic beta-glucosidase precursor	2.00	0.0216	Rice	LOC_Os03g53860.1
GRMZM2G123714	1Cz	X8 domain containing protein	1.72	0.0352	Rice	LOC_Os03g54910.1
GRMZM2G031138	1Cz	glucosyl transferase	1.63	0.0442	BLAST/TAIR	AAM47589/ AT3G02100.1
GRMZM2G162622	3H_GLS	THION18 - Plant thionin family protein precursor, expressed	1.75	0.0330	Rice	LOC_Os01g41140.1
GRMZM2G126260	3Cz_1.1	auxin efflux carrier component 3a	1.71	0.0361	BLAST/Rice	Q5VP70/ LOC_Os01g45550.2
GRMZM2G075594	3Cz_1.1	major facilitator superfamily antiporter	1.69	0.0361	BLAST	ACG42557

Appendix B

Gene	QTL	Sequence description	FC[a] (S/R)	adj. p-value	description source	BLAST/TAIR[b]/Rice[c] Accession No.
GRMZM2G115809	5Cz	unknown	1.66	0.0387	BLAST	ACR33917
GRMZM2G170692	5Cz	phenylalanine ammonia-lyase	2.09	0.0193	BLAST	NM_001174615.1
GRMZM2G170016	5Cz	cytochrome b5-like Heme/Steroid binding domain containing protein	1.79	0.0310	Rice	LOC_Os02g42740.1
GRMZM2G058472	6Cz	glycosyl transferase	1.66	0.0391	Rice	LOC_Os01g65780.1
GRMZM2G002128	6Cz	MYB-like transcription factor	1.68	0.0361	BLAST/Rice	ACG30643/ LOC_Os05g37730.1
GRMZM2G001332	6Cz/Lesion	uncharacterized	1.85	0.0316	BLAST	XR_555125.1
GRMZM2G030293	9Cz_2.1	40S ribosomal protein S28	1.75	0.0361	Rice	LOC_Os10g27174.1
GRMZM2G053652	9Cz_2.1	40S ribosomal protein S7	1.83	0.0361	BLAST/Rice	ACG35364/ LOC_Os03g18570.1
GRMZM2G149798	9H_GLS	thaumatin (PR5)	1.68	0.0373	BLAST/Rice	NP_001149439/ LOC_Os03g14030
GRMZM2G098875	9Cz_2.2	glutamate decarboxylase	1.66	0.0400	BLAST/Rice	XM_004985025.1/ LOC_Os03g13300.1
GRMZM2G356618	9Cz_2.2	uncharacterized	2.00	0.0216	BLAST	XR_559425.1
GRMZM2G046601	9Cz_2.2	glutamine synthetase	2.20	0.0485	Rice	LOC_Os03g12290.1
GRMZM2G145460	10Cz_2	NAD dependent epimerase/dehydratase family domain containing protein	1.94	0.0224	Rice	LOC_Os04g52730.1
GRMZM2G127668	10Cz_2	gibberellin 20 oxidase 2	1.63	0.0477	Rice	LOC_Os04g55070.1

[a] FC: Fold change; higher expression in the R bulk compared to the S bulk
[b] TAIR (https://www.arabidopsis.org/)
[c] Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/)

Table B6a: Differentially expressed genes with higher expression in R bulk observed in RNA-seq analysis that coincide with QTL.

Gene	QTL	Sequence description	FC[a] (R/S)	Q-value	R Bulk FPKM	S Bulk FPKM	description source	BLAST/TAIR[b]/Rice[c] Accession No.
GRMZM2G010034	1Cz	Mu transposon insertion Mu1009011 flanking sequence	1.48	0.02	248.4	167.7	BLAST	FJ910835.1
GRMZM2G164233	1Cz	uncharacterized protein	2.84	0	18.2	6.4	BLAST	XM_008667206.1
GRMZM2G101460*	1Cz	GTPase	1.93	0.03	8.1	4.2	Rice	LOC_Os03g51790.1
GRMZM2G133718	3H_GLS	Putative Serine Carboxypeptidase	3.84	0.01	1.1	4.2	Rice	LOC_Os01g06490.2
GRMZM2G099382	3Cz_1.1	tonoplast dicarboxylate transporter	1.89	0.04	14.1	7.4	BLAST	NM_001154848.1
GRMZM2G130375	3Cz_1.1	beta-galactosidase	1.84	0.03	11.3	6.1	Rice	LOC_Os01g39830.1
GRMZM2G066885*	5Cz	Cytochrome b561/ Ascorbate-specific transmembrane electron transporter 2	2.44	0	90.4	37.0	BLAST	NM_001177098.1
GRMZM2G157990	5Cz	tesmin/TSO1-like CXC 5	7.20	0	4.2	0.6	BLAST	XR_553668.1
GRMZM2G013821	5Cz	Zea mays high mobility group c1	1.83	0.01	35.0	19.2	BLAST	AJ131374.1
GRMZM2G061777	6Cz/ Lesion	delta-1-pyrroline-5-carboxylate synthase-like	2.70	0	24.5	9.1	BLAST/TAIR	XR_555054.1/ AT3G55610.1
AC203754.4_FG008	6Cz/ Lesion	delta-1-pyrroline-5-carboxylate synthetase (P5CS) gene	3.07	0	12.7	4.1	BLAST/TAIR	EF620362.1/ AT3G55610.1
GRMZM2G331833	6Cz/ Lesion	ATP-dependent Clp protease ATP-binding	1.82	0.01	13.0	7.2	Rice	LOC_Os05g45750.1
GRMZM2G108103	9Cz_2.1	polyphenol oxidase	3.33	0	5.1	1.5	Rice	LOC_Os01g58100.1
GRMZM2G055320	9Cz_2.1	4-coumarate--CoA ligase 4-like	1.87	0.03	16.6	8.9	BLAST/TAIR	XM_004965627.1/ AT3G21240.1
GRMZM2G044180*	9Cz_2.1	serine/threonine-protein kinase EDR1-like	1.85	0.02	8.6	4.7	BLAST	XM_008661820.1
GRMZM2G573686	9Cz_2.1	retrotransposon	2.87	0	28.8	10.0	BLAST	U68403.1
GRMZM2G065210	9H_GLS	ubiquinone biosynthesis methyltransferase	1.79	0.03	6.1	10.9	BLAST	ACG49031

Appendix B

Gene	QTL	Sequence description	FC[a] (R/S)	Q-value	R Bulk FPKM	S Bulk FPKM	description source	BLAST/TAIR[b]/Rice[c] Accession No.
GRMZM5G851617	9H_GLS	uncharacterised	8.22	0	0.6	4.6	BLAST	JQ887800.1
GRMZM2G036861	9H_GLS	chorismate synthase 2	9.82	0	0.6	6.1	BLAST/Rice	NP_001148583/ LOC_Os03g14990.1
GRMZM2G152919	9Cz_2.2	ubiquitin-protein ligase	1.82	0.02	15.5	8.5	BLAST	NM_001154420.1
GRMZM5G887345	10Cz_2	putative ubiquitin-like-specific protease 1B	2.07	0	14.2	6.8	BLAST	XM_008665338.1
GRMZM2G128934	10Cz_2	Ubiquitin-specific protease family C19-related protein	2.81	0	10.5	3.7	TAIR	AT1G78880.1
GRMZM2G156299	10Cz_2	uncharacterized protein	3.62	0.02	4.3	1.2	BLAST	XR_560607.1
GRMZM2G031354	10Cz_2	Protease inhibitor/seed storage/LTP family protein precursor	6.97	0	86.7	12.4	Rice	LOC_Os04g52260.1
GRMZM2G155837*	10Cz_2	Malectin LRR receptor-like serine/threonine-protein kinase	2.22	0	30.4	13.7	BLAST	XM_008665377.1
GRMZM2G028902	10Cz_2	proteasome assembly chaperone 3	2.52	0.02	10.5	4.2	BLAST	XR_561104.1
GRMZM2G097854	10Cz_2	leucoanthocyanidin reductase	1.76	0.04	22.3	12.7	BLAST/Rice	EU966048.1/ LOC_Os04g53850.1
GRMZM2G000633	10Cz_2	cysteine-rich RLK (Receptor-like protein kinase) 10	6.60	0	20.4	3.1	TAIR	AT4G23180.1
GRMZM2G177412	10Cz_2	glutamyl-tRNA reductase	1.78	0	59.6	33.6	BLAST	EU953356.1

[a] FC: Fold change; higher expression in the S bulk compared to the S bulk
[b] TAIR (https://www.arabidopsis.org/)
[c] Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/)
* Genes indentified as candidates in the proposed resistance model (Figure 4.10)

Table B6b: Differentially expressed genes with higher expression in S bulk observed in RNA-seq analysis that coincide with QTL.

Gene	QTL	Sequence description	FC[a] (S/R)	Q-value	R Bulk FPKM	S Bulk FPKM	description source	BLAST/TAIR[b]/Rice[c] Accession No.
AC208897.3_FG004	1Cz	sugar transport protein 13	2.78	0	31.9	88.9	BLAST	XM_008654539.1
GRMZM2G342685	1Cz	zinc finger CCHC domain-containing protein 10	4.20	0	1.1	4.8	BLAST	NM_001279712.1
GRMZM2G174975	3H_GLS	GDSL-like lipase/acylhydrolase	3.33	0	2.4	8.1	Rice	LOC_Os01g22780.2
GRMZM2G314075	3H_GLS	abscisic stress-ripening	11.48	0	2.4	27.1	Rice	LOC_Os01g72900.1
GRMZM2G383699	3H_GLS	abscisic acid response protein [Cucumis melo]	17.10	0	1.0	16.9	BLAST	AAL27560
GRMZM2G096008	3H_GLS	heavy metal-associated domain containing protein	2.54	0.01	2.6	6.7	Rice	LOC_Os03g60480.1
GRMZM2G314652	3H_GLS	bifunctional 3-dehydroquinate dehydratase/shikimate dehydrogenase, chloroplast precursor	1.91	0.01	11.9	22.7	Rice	LOC_Os01g27750.1
GRMZM2G061527	3H_GLS	Leucine-rich repeat protein kinase family protein	3.26	0	1.7	5.7	TAIR	AT1G74360.1
AC214360.3_FG001	3H_GLS	terpene synthase	2.98	0	18.7	55.7	Rice	LOC_Os02g36220.1
GRMZM2G114850	3H_GLS	NAC domain containing protein 1	2.37	0.01	5.1	12.0	TAIR	AT1G56010.2
GRMZM2G160840	3H_GLS	MYB family transcription factor	9.22	0	0.9	8.3	Rice	LOC_Os12g37690.1
GRMZM2G045976	3H_GLS	phosphatidic acid phosphatase-related	4.16	0	1.4	6.0	Rice	LOC_Os05g21180.1
GRMZM5G888204	3H_GLS	hypothetical protein	2.49	0.01	2.3	5.8	BLAST	OsJ_02038
GRMZM2G110369	3H_GLS	oxidoreductase, 2OG-Fe oxygenase family protein	2.70	0	34.1	91.9	Rice	LOC_Os07g07410.2
GRMZM2G092169	3H_GLS	carboxyl-terminal peptidase	2.64	0	48.5	128.3	Rice	LOC_Os07g38590.1
GRMZM2G118241	3H_GLS	Putative Serine Carboxypeptidase homologue	2.73	0	16.5	45.0	Rice	LOC_Os12g15470.2
GRMZM2G174726	3H_GLS	unknown	3.50	0	2.5	8.6		
GRMZM2G103342	3H_GLS	peroxidase precursor	6.18	0	19.6	120.8	Rice	LOC_Os01g73170.1
GRMZM2G426046	3H_GLS	Calmodulin-related calcium sensor protein	4.67	0	1.0	4.8	Rice	LOC_Os01g72530.1

Appendix B

Gene	QTL	Sequence description	FC[a] (S/R)	Q-value	R Bulk FPKM	S Bulk FPKM	description source	BLAST/TAIR[b]/Rice[c] Accession No.
GRMZM2G052571	3H_GLS	glutathione S-transferase	12.25	0	0.5	6.1	BLAST/Rice	NP_001147759/ LOC_Os03g39850.1
GRMZM2G028556	3H_GLS	glutathione S-transferase	8.04	0	0.7	5.4	BLAST/Rice	ACG47313/ LOC_Os03g39850.1
GRMZM2G037452	3H_GLS	delta adaptin subunit of AP-3 [Drosophila melanogaster]	5.98	0	23.8	142.5	BLAST	AAC01743
GRMZM2G125032	3H_GLS	beta-1,3-glucanase precursor (PR 2)	5.69	0	24.3	138.0	BLAST	ADL60383
GRMZM2G065585	3H_GLS	beta 1,3 glucanase (PR 2)	3.49	0	150.1	524.5	BLAST	ACJ62647
GRMZM2G145715	3H_GLS	pgd2 (6-phosphogluconate dehydrogenase2)	1.93	0	36.0	69.5	MaizeGDB[d]	
GRMZM2G084779	3H_GLS	potassium transporter	2.92	0	41.9	122.1	Rice	LOC_Os01g70490.1
GRMZM2G173534	3H_GLS	inducer of CBF expression 2	3.96	0	2.0	8.0	Rice	LOC_Os01g70310.1
GRMZM2G129114	3H_GLS	Nucleotide-diphospho-sugar transferase family protein	2.76	0	2.8	7.7	TAIR	AT1G14590.1
GRMZM2G077809	3H_GLS	copine	1.80	0.03	12.3	22.1	Rice	LOC_Os01g68060.1
GRMZM2G086845	3H_GLS	fructokinase-1 [Zea mays]	2.56	0	3.1	8.0	BLAST	NP_001105210
GRMZM2G329069	3H_GLS	hypothetical protein	4.10	0	3.2	13.0	BLAST	NP_001145649
GRMZM2G014653	3H_GLS	NAC domain-containing protein 48 [Zea mays]	2.06	0	26.7	55.0	BLAST	ACG28360
GRMZM2G078667	3H_GLS	Disease resistance-responsive (dirigent-like protein) family protein	2.51	0	31.2	78.1	TAIR	AT2G21100.1
GRMZM2G078500	3H_GLS	4,5-DOPA dioxygenase extradiol	3.37	0	1.4	4.7	Rice	LOC_Os01g65680.1
GRMZM2G167253	3H_GLS	receptor-like protein kinase HAIKU2 precursor	1.98	0.02	3.5	7.0	Rice	LOC_Os01g65650.1
GRMZM2G074248	3H_GLS	harpin-induced protein 1 domain containing protein	6.08	0	1.3	8.0	Rice	LOC_Os01g64470.1
GRMZM2G453805	3Cz_1.1	PRm3 (chitinase)	3.20	0	221.4	708.4	BLAST	NM_001279467.1
GRMZM2G067600	3Cz_1.1	uncharacterized	32.37	0	0.5	14.7	BLAST	XM_008677071.1

Appendix B

Gene	QTL	Sequence description	FC[a] (S/R)	Q-value	R Bulk FPKM	S Bulk FPKM	description source	BLAST/TAIR[b]/Rice[c] Accession No.
GRMZM2G029547	3Cz_1.1	cinnamoyl-CoA reductase 1-like	2.93	0	1.9	5.5	BLAST	XM_008677078.1
GRMZM2G061723	3Cz_1.2	putative calmodulin-binding family protein	1.74	0.03	11.3	19.6	BLAST	NM_001152830.1
GRMZM2G070708	5Cz	Galactosyltransferase (avr9 elicitor response protein)	1.65	0.01	14.3	23.6	BLAST/Rice	EU969433.1/ LOC_Os02g35870.2
GRMZM2G132706	5Cz	UDP-glycosyltransferase 85A1-like	2.39	0.02	2.5	5.9	BLAST	XM_004952826.1
GRMZM2G325693	5Cz	uncharacterized	2.17	0.01	14.6	31.7	BLAST	NM_001176804.1
GRMZM2G027193	5Cz	YGL010w-like protein	1.83	0.04	18.9	34.5	BLAST	EU968035.1
GRMZM2G146380	5Cz	NAC transcription factor NAM-1-like	8.59	0	0.8	6.8	BLAST	XM_008647641.1
GRMZM2G176253	5Cz	peptide transporter PTR2	2.22	0	7.0	15.5	BLAST	NM_001294257.1
GRMZM2G083855	5Cz	uncharacterized	20.79	0	0.2	4.3	BLAST	NM_001151479.1
GRMZM5G882364	5Cz	heavy metal transport/detoxification protein	4.81	0	1.3	6.4	Rice	LOC_Os02g37280.1
GRMZM2G025832	5Cz	flavonoid 3' hydroxylase gene	2.20	0.01	5.1	11.1	BLAST	HQ699781.1
GRMZM2G176307	5Cz	glyceraldehyde-3-phosphate dehydrogenase	3.41	0	18.6	63.5	BLAST/Rice	X73152.1/ LOC_Os04g40950.1
GRMZM2G129189	5Cz	endochitinase PR4	2.92	0	7.1	20.8	BLAST	NM_001157282.1
GRMZM2G035584	5Cz	anthranilate N-benzoyltransferase protein 1	2.18	0	14.7	32.0	BLAST	EU957596.1
GRMZM2G148387	5Cz	glutaredoxin subgroup I	2.92	0	49.8	145.5	BLAST	NM_001165476.2
GRMZM2G074604	5Cz	pal1 (phenylalanine ammonia lyase homolog1)	4.33	0	11.7	50.7	BLAST	NM_001254868.1
GRMZM2G029048	5Cz	phenylalanine ammonia-lyase	2.16	0	31.8	68.9	BLAST	NM_001158010.1
GRMZM2G334660	5Cz	phenylalanine ammonia-lyase	3.87	0	2.4	9.1	BLAST	XM_008647730.1
GRMZM2G170692	5Cz	phenylalanine ammonia-lyase	4.11	0	2.2	9.1	BLAST	NM_001174615.1

Appendix B

Gene	QTL	Sequence description	FC[a] (S/R)	Q-value	R Bulk FPKM	S Bulk FPKM	description source	BLAST/TAIR[b]/Rice[c] Accession No.
GRMZM2G370991	5Cz	ethylene-insensitive protein	4.46	0	1.2	5.5	Rice	LOC_Os07g06130.1
GRMZM2G354777	5Cz	ATP-dependent Clp protease proteolytic subunit	1.90	0.01	30.0	56.9	BLAST	EU969883.1
GRMZM2G170016	5Cz	cytochrome b5-like Heme/Steroid binding domain containing protein	2.39	0	25.7	61.6	Rice	LOC_Os02g42740.1
GRMZM2G170017	5Cz	carbonyl reductase 1	5.86	0	5.1	30.1	BLAST	EU955516.1
GRMZM2G048728	5Cz	plastid-lipid-associated protein 2	3.17	0.03	4.0	12.6	BLAST	EU964536.1
GRMZM2G117164	5Cz	homeobox-leucine zipper protein ATHB-6	3.14	0	4.4	13.7	BLAST	NM_001158353.1
GRMZM2G135385	5Cz	cytochrome b5 isoform A-like	3.17	0	16.3	51.6	BLAST	NM_001301606.1
GRMZM2G097728	5Cz	ypt homolog2 (ypt2)	1.67	0.04	15.7	26.2	BLAST	XM_008646050.1
GRMZM2G094655	5Cz	methylglutaconyl-CoA hydratase	1.73	0.04	10.2	17.6	BLAST	NM_001157194.1
GRMZM2G398781	5Cz	uncharacterized	1.93	0.03	5.8	11.3	BLAST	XR_553676.1
GRMZM2G147014	5Cz	dehydrin COR410	2.87	0	21.9	63.0	BLAST	NM_001154006.1
GRMZM2G468585	5Cz	DUF584 domain containing protein	9.01	0	0.7	6.2	Rice	LOC_Os02g41840.1
GRMZM2G154628	5Cz	plasma membrane intrinsic protein2 (pip2d)	3.99	0	1.9	7.5	BLAST	NM_001111556.1
GRMZM2G434203	5Cz	ethylene-responsive transcription factor ERF027-like	90.06	0	0.1	4.7	BLAST	XM_008647830.1
GRMZM2G033641	6Cz	patellin-2-like	3.85	0	1.7	6.6	BLAST	XM_004961931.1
GRMZM2G000836	6Cz	TRANSPARENT TESTA 1-like	2.71	0	4.0	10.8	BLAST	XM_008651459.1
GRMZM2G118037	6Cz/Lesion	subtilisin N-terminal Region family protein	1.64	0.04	113.0	185.6	Rice	LOC_Os05g35740.1
GRMZM2G099481	6Cz/Lesion	phosphatidate phosphatase LPIN2-like	1.76	0.02	9.6	16.8	BLAST	XM_004961768.1
GRMZM2G080045	6Cz/Lesion	ammonium transporter 2	2.48	0.01	3.0	7.4	BLAST	NM_001154311.1
GRMZM2G140817	6Cz/Lesion	putative cytochrome P450 superfamily protein	1.78	0.05	11.2	20.0	BLAST	NM_001148638.1

Appendix B

Gene	QTL	Sequence description	FC[a] (S/R)	Q-value	R Bulk FPKM	S Bulk FPKM	description source	BLAST/TAIR[b]/Rice[c] Accession No.
GRMZM2G389301	6Cz/Lesion	EID1-like F-box protein 3	6.85	0	1.1	7.6	BLAST	XM_008651590.1
GRMZM2G001332	6Cz/Lesion	uncharacterized	4.73	0	1.2	5.7	BLAST	XR_555125.1
GRMZM2G029546	6Cz/Lesion	zma-miR399f-2 precursor	2.81	0	34.6	97.3	BLAST	GQ905615.1
GRMZM2G132212	6Cz/Lesion	receptor-like protein kinase HSL1	2.71	0	9.7	26.4	BLAST	XM_008651688.1
GRMZM2G126048	6Lesion_area	Nse4, component of Smc5/6 DNA repair complex	3.16	0	2.6	8.3	TAIR	AT1G51130.1
GRMZM2G126079	6H_GLS	GTPase-activating protein	2.34	0	6.6	15.4	Rice	LOC_Os05g45840.2
GRMZM2G337594	9Cz_2.1	RALFL28 - Rapid ALkalinization Factor RALF family protein precursor	2.22	0	43.8	97.1	Rice	LOC_Os06g29730.1
GRMZM2G413607	9Cz_2.1	bromodomain containing protein	2.33	0.01	7.8	18.0	BLAST	NM_001157746.1
GRMZM2G092146	9Cz_2.1	hypothetical protein	5.11	0	2.4	12.5	BLAST	NM_001149447.1
GRMZM2G041068	9Cz_2.1	hypothetical protein	2.26	0	13.3	30.1	BLAST	XM_002437243.1
GRMZM2G086841	9Cz_2.1	heat shock protein DnaJ	3.00	0	6.9	20.6	Rice	LOC_Os06g44160.1
GRMZM2G151406	9Cz_2.1	copper-transporting ATPase RAN1-like	2.38	0	6.7	15.8	BLAST	XM_008662819.1
GRMZM2G404702	9Cz_2.1	copper-transporting ATPase RAN1-like	2.12	0.03	2.8	6.0	BLAST	XM_008678071.1
GRMZM2G329869	9Cz_2.1	uncharacterized	2.84	0	2.0	5.8	BLAST	XR_554109.1
GRMZM2G093092	9Cz_2.1	5-pentadecatrienyl resorcinol O-methyltransferase-like	2.82	0	75.3	212.2	BLAST	XM_008661781.1
GRMZM2G047546	9Cz_2.1	65-kDa microtubule-associated protein 3-like	2.16	0	2.9	6.3	BLAST	XM_008661875.1
GRMZM5G806784	9Cz_2.1	SNARE associated Golgi protein	2.58	0	7.2	18.5	Rice	LOC_Os03g25750.1
GRMZM2G178645	9Cz_2.1	ATPase	1.65	0.04	58.3	95.9	Rice	LOC_Os03g58790.1
GRMZM2G159500	9Cz_2.1	GRAB2 protein (NAC domain family)	3.11	0	6.9	21.3	BLAST	NM_001154259.1
GRMZM2G704475	9Cz_2.1	late embryogenesis abundant protein 1-like	NA	0	0.0	7.1	BLAST	XM_004984480.1

Appendix B

Gene	QTL	Sequence description	FC[a] (S/R)	Q-value	R Bulk FPKM	S Bulk FPKM	description source	BLAST/TAIR[b]/Rice[c] Accession No.
GRMZM5G872256	9Cz_2.1	galactinol synthase 3 (gols3) (osmoprotectant)	2.93	0	31.3	91.8	BLAST	AY192144.1
GRMZM5G877547	9Cz_2.1	aspartic proteinase nepenthesin-1-like	5.94	0	1.2	7.3	BLAST	XM_008661976.1
AC233955.1_FG008	9Cz_2.1	SMP-30/Gluconolaconase/LRE domain protein	2.26	0.03	4.0	8.9	Rice	LOC_Os03g19452.1
GRMZM2G152417	9Cz_2.1	AMP-binding enzyme	2.15	0.01	4.5	9.6	Rice	LOC_Os03g19250.1
GRMZM2G145568	9Cz_2.1	LOB domain-containing protein 1	3.42	0	2.4	8.1	BLAST	XM_008662033.1
GRMZM2G063151	9H_GLS	initiator-binding protein	2.00	0.03	3.0	6.1	Rice	LOC_Os03g16700.1
GRMZM2G151992	9H_GLS	haloacid dehalogenase-like hydrolase family protein	3.15	0	6.8	21.5	Rice	LOC_Os03g16670.1
GRMZM2G123107	9H_GLS	beta-1,3-glucanase 1 (PR 2)	2.55	0	4.4	11.2	TAIR	AT3G57270.1
GRMZM2G031447	9H_GLS	gibberellin receptor GID1L2	2.69	0.01	2.1	5.7	Rice	LOC_Os03g15270.1
GRMZM2G104575	9H_GLS	dihydrodipicolinate reductase	2.06	0.02	5.8	12.0	Rice	LOC_Os03g14120.2
GRMZM2G147775	9H_GLS	actin-depolymerizing factor 5	6.50	0	1.9	12.3	BLAST/Rice	ACG43718/ LOC_Os03g13950.1
GRMZM2G081175	9H_GLS	senescence-associated protein	1.95	0.01	13.1	25.6	Rice	LOC_Os03g13840.2
GRMZM2G081458	9H_GLS	NB-ARC domain-containing disease resistance protein	31.04	0	0.2	6.7	TAIR	AT3G14470.1
GRMZM2G159768	9H_GLS	unknown	10.22	0	1.7	17.3		
GRMZM2G098875	9Cz_2.2	glutamate decarboxylase 1-like	14.38	0	0.5	7.0	BLAST	XM_004985025.1
GRMZM2G067225	9Cz_2.2	allene oxide synthase1 (aos1)	2.38	0	18.5	44.0	BLAST	NM_001111774.1
GRMZM2G356618	9Cz_2.2	uncharacterized	56.22	0	0.2	9.3	BLAST	XR_559425.1
GRMZM2G046601	9Cz_2.2	glutamine synthetase	8.23	0	4.5	36.8	Rice	LOC_Os03g12290.1
GRMZM2G423331	9Cz_2.2	naringenin 7-O-methyltransferase-like	4.41	0	1.6	7.0	BLAST	XM_008662995.1
GRMZM2G309899	9Cz_2.2	retrotransposon protein	2.11	0.02	3.9	8.2	Rice	LOC_Os12g05290.1

Appendix B

Gene	QTL	Sequence description	FC[a] (S/R)	Q-value	R Bulk FPKM	S Bulk FPKM	description source	BLAST/TAIR[b]/Rice[c] Accession No.
GRMZM2G050159	9Cz_2.2	membrane protein	3.90	0	2.4	9.3	BLAST	XM_008660360.1
GRMZM2G391042	10Cz_2	calcium-transporting ATPase 8, plasma membrane-type	1.94	0.02	5.7	11.1	BLAST	XM_008665346.1
GRMZM2G456547	10Cz_2	nucleoredoxin 3	4.54	0	2.4	10.9	BLAST	XM_008665355.1
GRMZM2G438202	10Cz_2	ethylene-responsive transcription factor 4-like	9.15	0	0.5	4.9	BLAST	XM_008665357.1
GRMZM2G108125	10Cz_2	aminotransferase y4uB	2.17	0	17.1	37.0	BLAST	XM_008663848.1
GRMZM2G145137	10Cz_2	uncharacterized	5.93	0	4.5	26.6	BLAST	XR_560610.1
GRMZM2G396856	10Cz_2	probable polyamine oxidase 2	2.25	0	4.5	10.2	BLAST	XM_008665391.1
GRMZM2G396846	10Cz_2	bromodomain containing protein	2.35	0	13.3	31.3	BLAST	XM_008663811.1
GRMZM5G818977	10Cz_2	putative acyl-CoA N-acyltransferases (NAT) family protein isoform 1 (GNAT42)	2.22	0	13.4	29.7	BLAST	XM_008663326.1
GRMZM2G445616	10Cz_2	uncharacterized	2.89	0	4.7	13.5	BLAST	NM_001176772.1

[a] FC: Fold change; higher expression in the S bulk compared to the R bulk
[b] TAIR (https://www.arabidopsis.org/)
[c] Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/)
[d] MaizeGDB (http://maizegdb.org/)