An investigation of *Fusarium verticillioides* infection in maize using physiological and molecular approaches

Humaira Lambarey

Supervisor: Dr. Suhail Rafudeen
Co-Supervisor: Dr. Shane Murray

Thesis submitted in fulfilment of the requirements for the degree of Master of Science in the Department of Molecular and Cell Biology, Faculty of Science, University of Cape Town, South Africa.

February 2017

Declaration

I hereby declare that all the work presented in this thesis is my own and that it has not been previously submitted at any other university for any other degree.

I know the meaning of plagiarism and declare that all of the work in this thesis, save for that which is properly acknowledged, is my own.

Humaira Lambarey
February 2017

Acknowledgments

Firstly and most importantly, I have to thank my creator, for without his infinite mercy and grace it would not have been possible for me to complete this degree.

To my supervisors, Dr M.S Rafudeen and Dr. S. Murray, thank you for your guidance, support, advice, understanding and encouragement. Thank you for always making time, for always having a moment to spare and for believing in my capabilities.

Financial assistance towards the funding of this research by the National Research Foundation (NRF) is hereby acknowledged. I would also like to thank the NRF (grant-holder linked bursary) for funding two years of my Master's.

Thank you to the CPGR for conducting my RNA-seq experiment. In addition, I would also like to thank Cornel Ghiban from the DNA learning center for helping me troubleshoot when I experienced problems during my RNA-seq analysis in DNA Subway. To all my friends and lab colleagues (both past and present), thank you for your friendship, motivation, support, encouragement and words of advice. A special thank you to Hawwa Gabier for all the help and for keeping me sane both in and out of the lab during this stressful time. To Michaella Hulley and Hawwa Gabier, thank you for taking the time and effort to read and edit these chapters.

Finally, I would like to thank my family, in particular my parents and siblings. Thank you for your constant support, love, fun and laughter, encouragement, thank you for listening to all my complaints (especially, Wafiqah) and for believing in me when I did not believe in myself. I am forever grateful, appreciate all you do for me and love you all dearly.

Abbreviations:

ANOVA	Analysis of Variance
bp	base pairs
cDNA	Complementary DNA
CAT	Catalase
Ct	Cycle threshold
CPGR	Centre for Proteomic and Genomic Research
CV^2	Co-efficient of variation
DEGs	Differentially expressed genes
DEPC	Diethypyrocarbonate
FDR	False Discovery Rate
FER	*Fusarium* ear rot
FPKM	fragments per kilobase of exon per million fragments mapped
GOI	Gene of interest
GO	Gene Ontology
GR	Glutathione reductase
Hpi	Hours post infection
IARC	International Agency for Research on Cancer
PDA	Potato Dextrose Agar
mRNA	messenger RNA
MAS	marker assisted selection
MS	Murashige and Skoog
NBT	nitroblue tetrazolium
NCBI	National Center for Biotechnology Information
PR	pathogenesis related
QTL	quantitative trait loci
RT-Qpcr	quantitative Real-Time PCR
RG	Reference gene
RIN	RNA integrity number
RNA-seq	RNA-sequencing
ROS	Reactive oxygen species
rpm	revolutions per minute
SEA	Singular Enrichment Analysis
SOD	Superoxide dismutase
T_A	Annealing temperature

Figures and Tables:

Figures:

Tables:

Abstract

Maize (*Zea mays* L.) is an important staple food crop in sub-Saharan Africa providing food security to millions of people. *Fusarium verticillioides* is an important fungal pathogen that infects maize and causes 'Fusarium Ear Rot' which decreases maize kernel yield and quality. In addition, the fungus produces mycotoxins which contaminate the kernel and upon ingestion have negative health consequences for both people and livestock. To this date, there is still no African maize line completely resistant to infection by *F. verticillioides*.

In this study, an African maize line, *Zea mays* CML144, was infected with *F. verticillioides* using a soak-seed inoculation method and grown for two weeks under controlled conditions. Analysis of the morphological characteristics showed that compared to the control (mock-infected) maize, infected maize seedlings displayed signs of stunting with leaves shorter & thinner while roots were shorter and displayed visible signs of rotting. Control and infected maize plants were also characterised physiologically and biochemically. Electrolyte leakage experiments were conducted on the meristem regions of the plants after week one and two of infection and showed that leakage increased over time in both control and infected samples with no significant difference observed between the two groups. Biochemical characterisation by analysing superoxide dismutase (SOD), catalase (CAT) and glutathione reductase (GR) antioxidant enzymes showed an increase after the two weeks of infection, indicating a defense response by the plants in response to infection by the fungal pathogen.

RNA-sequencing, the main aim of this study was conducted on control and infected plants after two weeks of infection to identify differentially expressed genes (DEGs) involved in *F. verticillioides* infection. The Illumina NextSeq 500 platform was used to sequence the transcriptome and quantify changes in gene expression. Analysis of the RNA-seq data using the Tuxedo suite of protocols revealed significant DEGs that were both up- and down-regulated in the infected samples compared to the control. Data analysis was conducted using the DNA subway online bioinformatics tool and these results were compared to those obtained using a separate analysis which also incorporated the Tuxedo suite of protocols.

Bioinformatic analysis on the RNA-seq DEGs were performed using the agriGO analysis tool which revealed three significant Gene Ontology (GO) terms for both the up- and down-regulated genes, respectively, with the 'response to stimulus' GO-term (within the down-regulated genes) being of specific interest. Other GO-terms included response to chemical stimulus, carbohydrate metabolic process and ion bonding, which also played a role in the defense response when plants were

infected by the fungal pathogen. Quantitative Real-Time PCR was performed on five DEGs that were either up- or down-regulated in response to *F. verticillioides* infection to validate RNA-seq data as well as the GO-analysis results. Quantitative Real-Time PCR was also used as a pre-validation (before RNA-seq) on *shrunken-1*, a down-regulated gene found in a previously conducted study. We observed that in response to infection by *F. verticillioides*, expression of *shrunken-1* was down-regulated, however, this was not shown to be significant (p>0.05).

The results in the current study and the identification of the genes in *Zea mays* CML144 responding to fungal infection will aid in the goal to develop a maize line completely resistant to *F. verticillioides* in Africa and in particular South Africa. This would provide improved food security and minimise health risks to the population in the long term.

To our knowledge, this is the first study investigating *F. verticillioides* infection in the African maize line *Zea mays* CML144 using the soak-seed inoculation method.

Table of contents:

Chapter 1

Literature review

1.1 Introduction

Maize (*Zea mays* L.) is an important food source in the diets of many people in sub-Saharan Africa (Fandohan *et al*., 2003; Rheeder *et al*., 2009; Small *et al*., 2012). It is a crop of economic importance and is currently the third most traded crop after wheat and rice (Pereira *et al*., 2011; Maschietto *et al*., 2016).

In general, maize is used mainly as a food source, however, it is also important for animal feed and other processes which include biofuels like ethanol (du Plessis, 2003; Pereira *et al*., 2011). In contrast to developed countries, maize is a staple food crop in developing countries where ~200 million people rely on its production (du Plessis, 2003). In South Africa, maize is the most important cereal crop with ~8 million tons produced in South Africa per year alone (du Plessis, 2003; Pereira *et al*., 2011).

Maize is a crop that usually grows in areas with warm weather conditions but is also highly adaptable to various environmental conditions (du Plessis, 2003; Pereira *et al*., 2011). However, like many other crops it can be attacked by various pests, parasites and pathogens with the invasion by various fungal pathogens and insects being significantly important (Fandohan *et al*., 2003; Pereira *et al*., 2011). Invasion by pathogens causes the plants to suffer from many diseases and also affects the growth and development of the plant which in turn affects the yield leading to economic losses (Pereira *et al*., 2011).

1.2 *Fusarium* and maize interactions

Maize is susceptible to a variety of fungal genera which include but are not limited to *Alternaria*, *Aspergillus*, *Bipolaris* and *Fusarium* spp. (Fandohan *et al*., 2003; Aiyaz *et al*., 2015). Among these fungal pathogens, *Fusarium* and *Aspergillus* are two of the major genera of fungi usually infecting maize at various stages of its life cycle (Fandohan *et al*., 2003; Zorzete *et al*., 2008). *Fusarium* species are most commonly found in soils and are considered field fungi. Studies observing the interactions of *Fusarium* spp. and maize mainly involve *Fusarium verticillioides* and *Fusarium graminearum* (Fandohan *et al*., 2003; Lanubile *et al*., 2015). *F. verticillioides* is the most common species isolated in most maize crops infected by fungal pathogens (Munkvold and Desjardins, 1997; Fandohan *et al*., 2003).

Fungal pathogens are classified into three central categories, these include biotrophs, hemibiotrophs and/or necrotrophs (Oliver and Ipcho, 2004). Biotrophs are pathogens that feed on and attain nutrients from living plant tissue whereas necrotrophs kill the tissue and then feed off the dead remains. Hemibiotrophs are pathogens that can behave as both biotrophs and necrotrophs depending on the surrounding environment (Glazebrook, 2005). *F. verticillioides* also known as *Fusarium moniliforme* is a systemic endophyte that can either be found as a hemibiotrophic pathogen or a symptomless biotroph depending on the specific environment it is found in and is both a saprophyte (lives on dead/decaying matter) and parasite of maize (Pereira *et al*., 2011). This means that hemibiotrophic *F. verticillioides* can infect living tissues as biotrophs then after a certain period of time are capable of causing death to the same tissue and then become saprotrophs (Lanubile *et al*., 2015).

Fusarium verticillioides is the most common fungus infecting maize and is almost always found in maize fields at the time of harvest (Marasas, 2001; Oren *et al*., 2003). The presence of *F. verticillioides* is usually associated with dry, warm weather conditions, these conditions combined with drought or precipitation cause further stress on the plant and expose maize to infection by the fungus (Wu *et al*., 2011; van Rensburg *et al*., 2015). In South Africa, infection usually occurs in North West province, the Northern Cape and certain areas in the Free State where average temperatures range from 29-32^{o}C (van Rensburg *et al*., 2015).

1.3 *Fusarium* ear rot

Species of *Fusarium* are capable of causing many diseases in maize which include root, stalk and ear rot as well as seedling diseases (Munkvold and Desjardins, 1997). *Fusarium* sp. can lead to the development of two types of ear rot in maize, namely red ear rot (*Gibberella* ear rot) or pink ear rot (*Fusarium* ear rot) (Logrieco *et al*., 2002; Miedaner *et al*., 2010; Qiu *et al*., 2015). In particular, the main species of *Fusarium* causing pink ear rot is *F. verticillioides*, other *Fusarium* species which include *F. proliferatum* and *F. subglutinans* also cause Fusarium ear rot (FER); however, these have been isolated less frequently from diseased ears (Logrieco *et al*., 2002; Venturini *et al*., 2015).

Infection by this fungus is associated with all of the stages in the growth cycle of the plant with disease symptoms ranging from asymptomatic to severe rotting after infection. It can infect maize systemically through infected seeds (with conidia and/or mycelia inside or on the surface of the seed), wounding (through insects) or through the silk channel (Munkvold and Desjardins, 1997; Fandohan *et al*., 2003; Oren *et al*., 2003; Lanubile *et al*., 2014). In the field, infection through the silk channel has shown to be the most common route of infection (Lanubile *et al*., 2010; Kant *et al*.,

2012). FER results when *F. verticillioides* enters the silk channel through the ear and spreads within the ear. Penetration and colonisation by *F. verticillioides* occurs after the conidia reach the silks, this causes biochemical reactions within the plant as host resistant mechanisms are triggered (Campos-Bermudez *et al*., 2013). When the spores colonise tissue in the silk, fungal hyphae begin growing from the point of attachment to the kernels (Kant *et al*., 2012). The specific pathway of infection in the maize plants depends on which *Fusarium* spp. is dominating and the insects present within a particular location (Mesterházy *et al*., 2012). The most common routes of infection can be seen in Figure 1.1.

Symptoms of FER vary depending on plant genotype, the severity of the disease as well as the surrounding environments where the plants are being grown (Lanubile *et al.,* 2010; Lanubile *et al.,* 2014). Typical FER is identified by either individual or a group of infected kernels that are scattered around the ear region as well as the presence of fungal growth identified by a whitish-pink to lavender colour (Lanubile *et al*., 2010; Miedaner *et al*., 2010). However, *F. verticillioides* is also capable of causing symptomless infection making it difficult to detect infection in the field (Munkvold and Desjardins, 1997).

Other factors, besides the effect of various climatic conditions, contribute to the development of FER; these include kernel damage (by insects and pests), the susceptibility of the host and differences in agricultural practices (Venturini *et al*., 2015).

When maize is contaminated by *Fusarium* spp., it does not only become unsuitable for consumption by both animals and humans but also results in the production of mycotoxins (Fandohan *et al*., 2003). Mycotoxins are secondary metabolites that are found to contaminate agricultural food products in the field during harvest, through unsuitable storage practices, during food processing and can occur at any time until consumption (Zorzete *et al*., 2008; Rheeder *et al*., 2009). Examples of mycotoxins include aflatoxins, fumonisins and ochratoxin A (Rheeder *et al*., 2009).

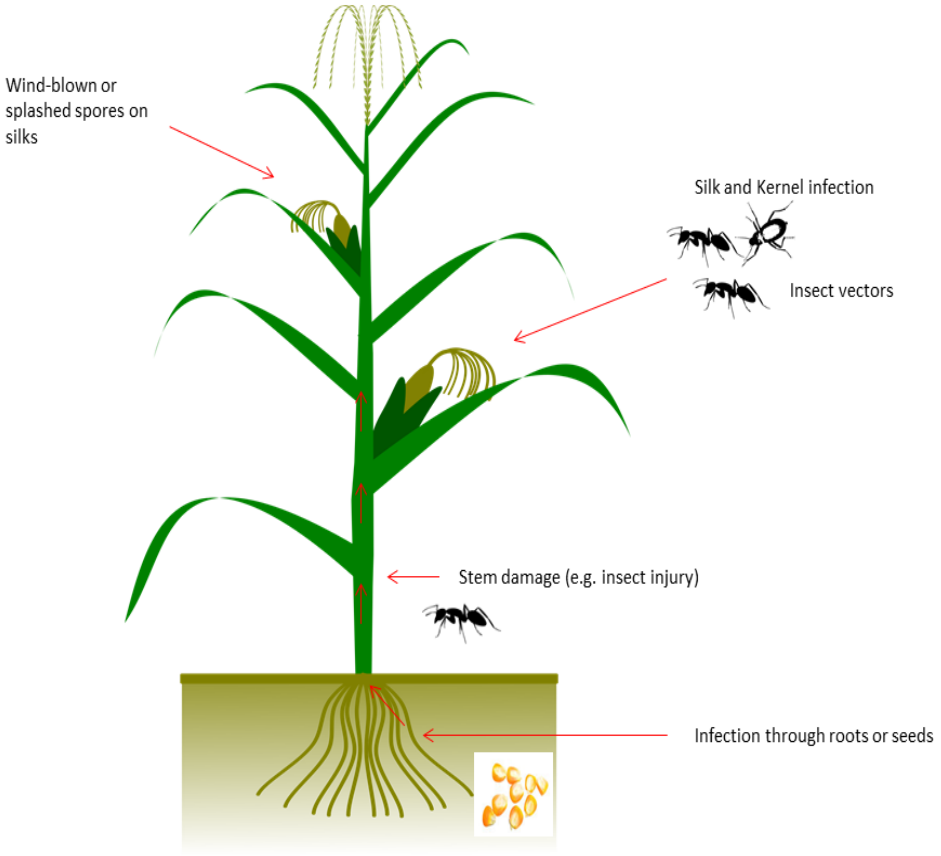

Figure 1.1: Pathways of *Fusarium verticillioides* infection that causes *Fusarium* ear rot in maize (adapted from Munkvold and Desjardins, 1997).

1.4 *Fusarium verticillioides* and the effects of fumonisins

Fumonisins are a class of mycotoxins. Their chemical structure and biological activity was identified in South Africa in 1988, where they were found to consist of a long hydrocarbon chain hydroxylated with added amino, tricaboxylic acid and methyl groups (Marasas, 2001; Fandohan *et al*., 2003; Mwalwayo and Thole, 2016). Fumonisins are produced by a number of *Fusarium* spp. which includes *F. verticillioides*, the species of fungus that is one of the highest producers of fumonisins (Fandohan *et al*., 2003). Fumonisins are heat-stable mycotoxins that are able to withstand temperatures of up to 150 ^{o}C and are therefore not completely removed during food preparation (Fandohan *et al*., 2003; Picot *et al*., 2013).

There are four series of fumonisins, namely A, B, C and P with over 20 known fumonisins analogues (Logrieco *et al*., 2002; van Rensburg *et al*., 2015). Fumonisin B is the most common fumonisin associated with maize and consists of three forms namely FB_1, FB_2 and FB_3, with FB_1 the most prevalent (van Rensburg *et al*., 2015). Other mycotoxins produced by *Fusarium* spp. include

deoxynivalenol and zearalenone (Pechanova and Pechan, 2015). Contamination of maize by fumonisins does not only cause losses economically but also poses a threat to handlers and processors of these crops (Mesterházy *et al*., 2012). In humans, consumption of *Fusarium* contaminated maize has been associated with oesophageal cancer in many regions including South Africa with the International Agency for Research on Cancer (IARC) classifying fumonisin B_1 as a Group 2B carcinogen (IARC, 2002; Logrieco *et al*., 2002; Fandohan *et al*., 2003; Zorzete *et al*., 2008; van Rensburg *et al*., 2015). Recently, a working group of experts convened by the IARC found upon reviewing the effects of fumonisins that they also contribute to stunting in children (Wild *et al*., 2016). These toxins are also a health concern to animals where they have been associated with a range of diseases (Marasas, 2001; Fandohan *et al*., 2003; Zorzete *et al*., 2008; van Rensburg *et al*., 2015).

The presence of fumonisins has been detected in both asymptomatic and symptomatic maize resulting in certain countries imposing regulations for the amount of fumonisins allowed in both foods and feeds for human and animal safety (Oren *et al*., 2003; Mesterházy *et al*., 2012). At present, South Africa does not have any regulations or monitoring systems in place with respect to fumonisin levels, increasing the risk to consumers (Marasas, 2001; van Rensburg *et al*., 2015). Marasas (2001) recorded a 1.2 and 354.9 µg/kg body weight/day probable intake of fumonisins in urban and rural South African maize consumers, respectively. Individuals in urban areas consume commercial maize whereas individuals in rural areas consume mouldy home-grown maize thus resulting in a high intake of fumonisins (Marasas, 2001).

The colonisation by fungi and the production of mycotoxins, more specifically fumonisins are influenced by both abiotic and biotic stress factors. Some important abiotic factors include water activity, pH, various nutrients and temperature (Zorzete *et al*., 2008; Picot *et al*., 2010). The main abiotic factors are temperature and water activity where increased water availability is associated with an increased fumonisin production in turn leading to an increased fungal growth. In addition, pH and the carbon to nitrogen ratio are also key factors in regulating fumonisin production (Picot *et al*., 2010).

There are presently no maize cultivars in southern Africa that are completely resistant to FER neither are there any appropriate fungicides available to control the infection by fungal pathogens (Small *et al*., 2012). It is thus of very high importance that a strategy to control and manage infection by this fungus be found in order to avoid the occurrence of FER and the accumulation of fumonisins.

Literature does suggest that there are two components that are important for resistance to *Fusarium* and that is firstly the resistance of the plant to initial penetration by the pathogen and secondly the resistance by the host to stop the spread of the pathogen in the plant tissue (Lanubile *et al*., 2010). This should form a base for further research in the development of strategies to avoid FER and fumonisin accumulation in the field.

1.5 Artificial methods of studying maize - *Fusarium* interactions

Natural infection by fungal pathogens in a real-world agricultural context is not always consistent making it difficult to study and develop strategies to avoid/resist infection. This is due to the many external abiotic and biotic variables that can influence maize infection. There are various artificial inoculation methods to mimic infection in the field while controlling external parameters. Some of the methods include the toothpick method, bubble breeding and pin-bar inoculation method (Mesterházy *et al*., 2012). The toothpick inoculation method has been modified since being introduced and involves inserting a toothpick overgrown with mycelium into a hole made using a needle in the hypocotyl of the plant (Keeling, 1982). With the pin-bar inoculation, pins are dipped into a conidial suspension and pressed through the husks and into the kernels (King and Scott, 1982). Oren *et al*. (2003) used a seed and a soil inoculation method, for the seed inoculation the seeds are soaked in a spore suspension whereas the soil inoculation involved mycelium inoculated onto maize leaves (Oren *et al*., 2003).

The use of artificial infection methods has many advantages, which includes the fact that the *Fusarium* spp. or any other fungal pathogen being studied is known. This is unlike the natural infection process where various fungal spp. could be infecting the plant at the same time. With artificial infection, the time and the route of infection are also known making it a more reliable source to measure disease severity (Mesterházy *et al*., 2012).

1.6 Plant defense strategies

Plant-pathogen recognition is the first step in the defense response which sets off the transcription of upstream defense signalling systems, however, the precise upstream mechanisms employed by maize is still not known (Fountain *et al*., 2015). It has been found that plants have two defense strategies when invaded by pathogens such as fungi which is based on an innate immunity response (Jones and Dangl, 2006). Firstly, on the extracellular surface (plasma membrane) of the host, pathogen recognition receptors recognise pathogen-associated molecular patterns (PAMPs) leading to PAMP-triggered immunity (PTI). The induction of PTI leads to reactive oxygen species (ROS) production, transcription of genes that respond to the pathogen, callose deposition as well as the

reinforcement of the cell wall at the infection sites. However, certain pathogens are able to overcome PTI and deploy effectors contributing to their virulence in the plant. The second defense strategy involves the activity of R-proteins which recognises these effector molecules and results in effector-triggered immunity (ETI). This response is more specialised and powerful than PTI and acts mostly inside the cell (cytoplasm), it results in disease resistance and often leads to a hypersensitive response (cell death) (Jones and Dangl, 2006; Lanubile *et al*., 2014). This description of the innate immune response was developed for biotrophic and hemibiotrophic pathogens (Jones and Dangl, 2006).

When plants are invaded by pathogens such as fungi, it results in changes within the plants which can be detected at both the transcriptional and translational level. Infected plants display increased mRNA levels and also increased levels of protein synthesis. The result of this is an increase in defense-related genes, enzymes and proteins (Zeilinger *et al*., 2015). These include pathogenesis-related proteins like chitinases, zeamatins as well as proteins related to minimising ROS (superoxide dismutase, catalase etc.) (Fountain *et al*., 2015). The plant defense response to invasion by the pathogen is also the activation of the hypersensitive response, an accumulation of phytoalexins (secondary metabolites protecting plant tissues after biotic stress) and the enhancement of enzyme activities (Schmelz *et al*., 2011; Zeilinger *et al*., 2015). The point of these responses are to stop the spread of the invading pathogen (Zeilinger *et al*., 2015).

Proteins associated with the response to pathogens:

Pathogenesis-related proteins

In terms of maize resistance, PR proteins are the largest group of proteins and they are inducible by a number of pathogens through the signalling of jasmonic acid, ethylene or salicylic acid. They are either accumulated at the site of infection or spread out through the plant (Pechanova and Pechan, 2015). Some PR proteins can also be found expressed constitutively within completely healthy plants during its growth and development stages. Among the most important PR proteins are chitinases and β-1,3-glucanases, others include proteins P21, P23 and zeamatin. These proteins are constitutively expressed and indicate that they are important not only after pathogen attack but also for normal plant processes (Lanubile *et al*., 2010; Pechanova and Pechan, 2015).

Reactive oxygen species scavengers

One of the main events that occur in plants is the recognition of the pathogen after the initiation of infection which results in the production of ROS. ROS can result in damage or even cell death in the host, therefore the detoxification of these species is important for the cell to survive. Plants control

ROS by scavenging systems which are able to detoxify ROS and maintain a balance within the host (Pechanova and Pechan, 2015). Some of the key detoxifying enzymes in this scavenger system include superoxide dismutase, catalase, glutathione reductase and ascorbate peroxidase (Lanubile *et al*., 2010; Pechanova and Pechan, 2015). The induction of these enzymes is important for the activation of downstream defense responses after infection has occurred.

Secondary metabolism proteins

When plants are attacked by pathogens they also produce certain organic compounds in order to protect themselves. These are produced through secondary metabolism usually via the alkaloid, isoprenoid and phenylpropanoid pathways. In particular, the phenylpropanoid pathway supplies a variety of phenolics which possess a wide range of defensive properties (Pechanova and Pechan, 2015). Phenolic compounds are antioxidants that inhibit mycotoxin production; they defend plants from the invading pathogen through direct contact with the fungus by acting as a barrier or in some cases reinforcing the structural components of the plants (Picot *et al*., 2013; Zerbo *et al*., 2014). In maize, terpenoids and benzoxazinoids have been studied for their roles in protecting the crop against biotic stresses (Huffaker *et al*., 2011; Schmelz *et al*., 2011).

Proteins in carbohydrate metabolic pathways

In terms of primary metabolic processes, carbohydrate metabolism is usually affected the most during the infection process. Proteins involved here include those enzymes involved in the TCA cycle, pentose phosphate pathway, glycolysis, ATP biosynthesis etc. (Pechanova and Pechan, 2015). However, further studies need to be conducted to identify the specific role of carbohydrate metabolism in the maize-pathogen interaction.

Proteins involved in protein synthesis, folding and stabilisation

Proteins involved in the synthesis of other proteins or in the folding and stabilisation process after infection include the eukaryotic translation initiation factor 5A, chaperonins, cyclophilin, heat-stress proteins and peptidylprolyl *cis*-trans isomerases. These proteins are usually up-regulated under stress (Lanubile *et al*., 2010; Lanubile *et al*., 2012a; Pechanova and Pechan, 2015).

It is important to obtain a comprehensive view of plant defense responses by examining the maize-fungal interaction through the use of platforms involving genomics, transcriptomics, proteomics and metabolomics (Sekhon *et al*., 2013).

The study of maize in response to pathogen attack through the use of various technologies, like microarrays, has led to the discovery of specific genes/proteins and also a better understanding of

the overall maize defense response to pathogen attack (Marioni *et al*., 2008; Picot *et al*., 2010; Cox *et al*, 2010; Sekhon *et al*., 2013).

For brevity, this literature review will only focus on the use of next-generation sequencing, specifically RNA-seq, as an approach to study the maize-fungal interaction.

1.7 Next-generation sequencing

Understanding the changes occurring in the maize transcriptome upon fungal infection is important as RNA transcripts form the basis of many and/or are implicated in many biological processes which includes the expression of defense genes (Yang and Kim, 2015).

The next-generation RNA-sequencing technique has become the method of choice for transcriptomic studies, especially in the study of plant-fungal interactions (Marioni *et al*., 2008; Cox *et al*., 2010; Sekhon *et al*., 2013; Seyednasrollah *et al*., 2015). It provides a platform to view in great detail the process of RNA transcription occurring in a cell at a specific point in time (Kim *et al*., 2013). The first step in the RNA-seq analysis process is the mapping of the RNA sequence reads against a reference genome or transcriptome (Kim *et al*., 2013; Van Verk *et al*., 2013, Yang and Kim, 2015). The purpose of the reference genome is to provide the location of the analysed RNA reads as the main purpose of RNA-seq is to reconstruct the transcripts or genes present in the cells of origin (Kim *et al*., 2013). The alignment to the genome is important for the analysis process because it provides the information on splice variants (Van Verk *et al*., 2013). The RNA-seq workflow includes the following steps: the pre-processing of raw data, read alignment to the reference genome, reconstruction of the transcriptome, expression quantification and the analysis of differential expression (Yang and Kim, 2015).

RNA-seq allows the accurate and quantitative measurement of gene expression allowing the detection of gene expression differences (Marioni *et al*., 2008). In addition, it allows the detection of novel genes or transcripts present and enables alternative splicing to be studied (Marioni *et al*., 2008; Trapnell and Salzberg, 2009; Wilhelm and Landry, 2009). It also has a better resolution and higher reproducibility than microarrays; however, one of the limitations associated with RNA-seq is the estimation of abundance at both the gene and transcript level (Yang and Kim, 2015). Protein levels are also better estimated using RNA-seq than microarrays, indicating that by using RNA-seq there is a better correlation between mRNA and protein (Fu *et al*., 2009).

With the introduction of RNA-seq, a range of bioinformatics tools were also required to be developed in order to process the output of the sequencing data into a format to observe changes in

gene expression at the quantitative level (Van Verk *et al*., 2013). In terms of quantification and transcript assembly at the gene level, the Tuxedo protocol is the most common software used and includes different programs such as TopHat, Cufflinks and Cuffdiff (Trapnell *et al*., 2012; Yang and Kim 2015). When differential gene expression is analysed, statistical methods are important because RNA-seq data are discrete and derived from read counts as compared to microarray data which is based on fluorescence intensity (Van Verk *et al*., 2013). For this purpose, different software packages have been developed to test for differential expression and include DESeq, edgeR, SAMseq and Cuffdiff amongst others (Van Verk *et al*., 2013; Seyednasrollah *et al*., 2015; Yang and Kim, 2015). Seyednasrollah and co-workers (2015) demonstrated that no single software produces the same results and they demonstrated large differences using the various bioinformatic pipelines (Seyednasrollah *et al*., 2015).

Currently, there are still some problems associated with multireads and estimating the abundance of splice variants; however, with the continued use of this technology, it is hoped that there will be improvements in sorting out these problems. Next in the next-generation sequencing circuit is third generation sequencing also known as single-molecule sequencing, this new generation of sequencing will have the potential to sequence complete transcripts reducing computing time required as with second generation sequencing (Van Verk *et al*., 2013). The fact that RNA-seq technologies are expanding rapidly, it is expected that new and improved bioinformatics tools will become available in order to test differential expression under various conditions (Yang and Kim, 2015).

1.8 Strategies to combat *Fusarium* infection (breeding for resistance)

Even though many strategies to avoid *Fusarium* and fumonisin contamination in maize have been proposed, there is still not one method that has been found to be completely reliable and safe for the environment (Venturini *et al*., 2015). Some of the strategies to avoid fungal infection include good agricultural practices, pesticides, appropriate planting dates, manipulating the plant's gene pool, tillage, managing irrigation & fertilisation, intercropping and rotation (Pereira *et al*., 2011; Small *et al*., 2012). These agricultural practices alone have not been sufficient when the environmental conditions favour *Fusarium* and the accumulation of fumonisins (Small *et al*., 2012). The development of a resistant maize genotype has been difficult due to the lack of the complete understanding of factors that are important in *F. verticillioides* infection and the accumulation of fumonisins (Campos-Bermudez *et al*., 2013).

With the infection of maize by various genera of fungi, maize farmers usually resort to the use of fungicides to reduce the infection and mycotoxin contamination. These fungicides, however, could also pose a health risk for human and animals. Their continued uses on crops also result in an increased risk of fungicide resistance and many of these fungicides lack specificity to specific pathogens (Moreno *et al*., 2005; Aiyaz *et al*., 2015). As an alternative, there has been a focus on the use of beneficial microbes (biocontrol agents) which are applied as seed treatments. These agents possess certain antagonistic properties that allow the promotion of plant growth while suppressing both abiotic and biotic stresses (Aiyaz *et al*., 2015). Different genera of bacteria have shown the potential of increasing crop yields with phosphate-solubilising and rhizosphere-associated bacteria being some of the microorganisms showing potential in certain crop species which include maize. These microorganisms may improve yield but need to be accompanied by good agricultural practices which include proper harvest and storage (Pereira *et al*., 2011). Small *et al*. (2012) looked at enhancing plant resistance to FER and the accumulation of fumonisins in susceptible South African commercial maize using resistance elicitors such as β-amino butyric acid and methyl jasmonate. None of the tested elicitors were able to consistently reduce FER and fumonisin accumulation, however, optimisation of the application of these elicitors as well as other factors like dosage, timing and frequency of the application of the elicitors could show different results (Small *et al*., 2012).

The study by Venturini *et al*. (2015) examined the role of the presence of flavonoids in the kernel pericarp in maize. The presence of flavonoids acts as a barrier against fungal infection, in the kernel pericarp region. This is done by reducing the mycelial progress from the infected kernels to the healthy, intact kernels. The results of this study suggested a potential role of flavonoid pigments in the pericarp which allows the reduction of infection and the accumulation of fumonisins. They do state, however, similar to many other potential strategies, flavonoids do not act alone in providing resistance to *Fusarium* (Venturini *et al*., 2015). Picot *et al*. (2010) suggested that phenolic acid, whose role was explained previously, is also a possible candidate to reduce fumonisin contamination.

The use of seed constitutive defenses such as *PR* genes and genes linked to the protection from oxidative stress could be used to develop biomarkers; these markers would be developed from candidate genes linked to resistant maize genotypes (Maschietto *et al*., 2016). Transformation of plants by means of genetic engineering has also become one of the strategies used to improve the resistance in plants which includes the use of genes that encode antifungal proteins such as PR proteins amongst others (Moreno *et al*., 2005; Kant *et al*., 2012). These genes are usually under the control of a promoter and are expressed when under pathogen attack (Moreno *et al*., 2005).

The identification of quantitative trait loci (QTL) is one of the prerequisites to understanding the mechanism of resistance to FER and also important for marker-assisted selection (MAS). One of the most efficient ways to combat infection by *F. verticillioides* and reduce the damage from FER is to breed for resistance. This would require inoculation trials in multiple environments to improve hybrid resistance which would entail the involvement of MAS. One efficient way to develop this resistance is to identify the resistant genes/QTL for MAS (Chen *et al*., 2012a). Chen *et al*. (2012a) found three QTL's in a BT-1 line (an inbred line resistant to FER) with the largest resistance effect shown on chromosome 4, the resistance seen here lays a good foundation for future QTL work. Important factors required to breed for resistance would be to conduct experiments in different environmental conditions and to use strains/isolates that are stable and highly aggressive (Miedaner *et al*., 2010; Chen *et al*., 2012a). By discovering efficient biomarkers and QTL's that are constant across maize populations, it will aid in the strategy to breed for resistance to FER instead of relying on agronomic practices (Maschietto *et al*., 2016). In a recent study, Maschietto *et al*. (2017) found both QTL's and candidate genes that could contribute to *F. verticillioides* resistance. This will aid in selection of maize genotypes with reduced disease severity and decreased mycotoxin accumulation. These candidate genes may also be further characterised for MAS (Maschietto *et al*., 2017).

In summary, genetic engineering appears to be one of the best strategies to minimise infection by *F. verticillioides* by creating maize lines with improved resistance to FER. It is also important that such FER resistance in maize be stable across generations and does not compromise kernel yield and quality. The use of high-throughput RNA-sequencing technologies is therefore extremely useful as it allows for a better understanding of the interactions between the fungus and maize upon infection at the transcriptomic level (Picot *et al*., 2010).

1.9 Aim of this study

The main aim of this study was to determine the effect of *F. verticillioides* on the susceptible African maize line CML144 by looking at overall gene expression after infection by this fungal pathogen. This aim was achieved by performing whole-transcriptome RNA-sequencing on both mock-infected (hereafter known as the control) and *F. verticillioides* infected (hereafter known as infected) maize plants after two weeks of growth in controlled conditions. Infection was achieved using an artificial method of inoculation, namely the soak-seed method prior to RNA extraction and sequencing. RNA-seq results were supported by performing morphological, physiological and biochemical experiments conducted prior to analysis and validated by performing Gene Ontology analysis as well as quantitative Real-Time PCR on significant differentially expressed genes.

The purpose of the RNA-seq analysis was to discover significant differentially expressed genes after the infection with *F. verticillioides* specifically with important roles in the defense response. In the long term, the main aim of this study is to develop a transgenic African maize line with antifungal properties and an increased resistance to fungal infection. Finding these significant differentially expressed genes would also help in understanding the role of these genes in the interaction between the fungus and maize plant against what is known in literature. Not many studies have been conducted in maize lines that are specific to Africa or specifically South Africa and thus this study is important to help find strategies to develop an African maize line resistant to *F. verticillioides*.

To our knowledge, this is the first RNA-seq study on the African maize line CML144 using the soak-seed artificial inoculation method of infection and grown in tissue culture conditions. The work in this study will form as a foundation for future studies based on *F. verticillioides* infection in maize.

Objectives for this study:

- To infect an African maize line CML144 with *F. verticillioides*
- To perform physiological and biochemical experiments on control and infected maize
- To perform whole-transcriptome RNA-seq after two weeks of fungal infection using the Illumina NextSeq 500 platform
- To determine the changes in gene expression between control and *F. verticillioides* infected plants
- To perform Gene Ontology analyses on differentially expressed genes (up- and down-regulated)
- To perform quantitative Real-Time PCR of selected differentially expressed genes for validation

Chapter 2

Physiology, Morphology and Biochemical assays

2.1 Introduction

Plants have various defense responses to both biotic and abiotic stresses (Ferrigo *et al*., 2015). The nature of the plant's physiological and biochemical changes in response to infection by *F. verticillioides* may provide some insight into plant defense responses and susceptibility under pathogen attack. In addition, pathogen attack and associated stress may cause changes in the physical appearance or developmental processes of the plant and thus studying the morphological characteristics in response to infection may also provide other useful information.

One way of estimating plant damage in response to stress is by measuring electrolyte leakage from the cells. This technique indicates the extent of cell membrane damage. Although typically used to assess the damage caused by abiotic stress, it has also been used to assess damage in plants in response to biotic stresses (Ádám *et al*., 2000; Bajji *et al*., 2002).

The attack and colonisation of plants by fungal pathogens causes the activation of plant defense responses in order to prevent or limit the invasion. This plant resistance can be characterised by reactive oxygen species (ROS), which are formed by the incomplete reduction of oxygen to water resulting in the formation of superoxide, hydrogen peroxide and hydroxyl radicals (Magbanua *et al*., 2007; Kumar *et al*., 2009). The accumulation of ROS above threshold levels leads to oxidative damage of fungal and/or plant cells (Demidchik *et al*., 2014). Hydrogen peroxide, because of its relative stability, has received the most attention because it also has a role as a signalling molecule to regulate certain processes such as plant-pathogen interactions (Mhamdi *et al*., 2010). In order to mitigate against uncontrolled damage caused by ROS, there is often an increase in antioxidants and antioxidant enzymes which target ROS (Panda., 2012). Superoxide dismutase (SOD), catalase (CAT) and glutathione reductase (GR) are but a few antioxidant enzymes that are known to play a role in defense against various stresses (Kumar *et al*., 2009).

It was shown that infection of wheat leaves by the fungal pathogen, *Pyricularia oryzae*, resulted in an overall increase in antioxidant enzymes and electrolyte leakage in both the susceptible and partially resistant wheat cultivars (Debona *et al*., 2012). Therefore, conducting electrolyte leakage and antioxidant enzyme assays in this study might provide insight into the effects of *F. verticillioides* infection in maize plants.

In this chapter, the aim was to observe the morphological characteristics, changing physiology and biochemistry of mock-infected and infected *Zea mays* (CML144) samples under controlled conditions. *Zea mays* (CML144) is a maize line susceptible to Fusarium ear rot as characterised by Okello *et al*. (2006) and was therefore chosen in this study with the idea that the regulation of resistance genes would be enhanced compared to a resistant maize line. Physical properties of the whole maize plant, including the roots, were analysed. In relation to this, electrolyte leakage and antioxidant enzyme assay experiments were also performed. This was to determine the overall effect of biotic stress on maize plants grown under conditions different to that of plants grown in the field.

2.2 Materials and Methods

2.2.1 Surface sterilisation of maize seeds:

Zea mays (CML144) seeds were transferred to a sterile 50ml sterilin containing 20ml of 100% ethanol (EtOH). This was incubated for one min at room temperature and then inverted for 15 sec. The EtOH was decanted and replaced with 20ml of 50% JIK (active ingredient w/v 3.5% sodium hypochlorite). The seeds were inverted briefly and allowed to sit in solution for 15 min, thereafter it was inverted briefly again and the JIK was decanted. The seeds were then washed five times with sterile water and left submerged in water until used.

2.2.2 Fungi:

The *F. verticillioides* (Sacc. Nirenberg) strain, MRC 826, used in this study was provided by Pannar Seed (Pty) Ltd. The fungus was maintained at 30^0C and sub-cultured weekly on Potato Dextrose Agar (PDA) (Lab M Limited, United Kingdom) or stored in 20% glycerol at -70°C.

2.2.3 Maize seed infection:

The sterilised maize seeds were infected with *F. verticillioides* via artificial inoculation according to Oren *et al*. (2003) with minor modifications. For inoculation, the PDA plates containing the 6-7 day old fungus were flooded with 20ml of 2% (v/v) Tween20 solution. Conidiospores were scraped off the PDA plate and a haemocytometer was used to obtain a concentration of 1x 10^3 conidiospores/ ml for infection. The seeds for infection were placed in the spore suspension while seeds for the control were suspended in 2% Tween20 solution only and then incubated at 30°C for 30 min with shaking. The seeds (control and infected) were dried in a laminar flow hood on filter paper overnight and placed on Murashige and Skoog (MS) (Highveld biological (Pty) Ltd, South Africa) media in glass jars under sterile conditions. The planted seeds in the MS media were incubated under controlled

conditions at 28°C (16hr light, 8hr dark; light intensity of 140µmol/m^2/s and 60% humidity) for a period of 14 days (maize V3 leaf stage).

2.2.4 Morphology and leaf measurements:

The morphology of the maize plants (control and infected) which included the colour of the leaves, size of stem, root lengths as well as the number of leaves and roots were observed after the 14 days of incubation.

Leaf measurements from the collar (where the leaf attaches to the stem) to the tip were also taken from leaves excised from the whole maize plant after week one and two of both control and infected plants.

2.2.5 Electrolyte leakage:

The CM100 conductivity meter (Reid & Associates, South Africa) was used to determine membrane damage after *F. verticillioides* infection of one and two-week old control and infected maize plants. Samples from the meristem area of the shoots across 6-8 plants were cut into ~1cm in diameter circular disks. The wells of the conductivity meter were filled with 2ml of Millipore water and the leaf disks were submerged into the respective wells. Electrolyte leakage was measured at 60 sec intervals for a period of 30 min. Immediately thereafter, samples from both time points were placed in foil packages and dried at 70°C for ~48 hrs. The foil packages were then placed in a desiccator for 10 min and the weights of the leaf disks were obtained. To determine the rate of leakage, the values obtained from the conductivity meter were plotted on a straight line graph to obtain the gradient. The rate of electrolyte leakage in the plants was then calculated using the following equation:

$$\text{Electrolyte leakage } (\mu S.\ min^{-1}.gdw^{-1}) = \text{rate of leakage/ dry weight}$$

Electrolyte leakage experiments on control and infected plants were carried out thrice for biological validation.

2.2.6 Antioxidant assays:

The antioxidant activity of superoxide dismutase, glutathione reductase and catalase were assayed using a spectrophotometric approach. Enzyme extractions and the assays measuring enzyme activity were performed according to Bailly *et al*. (1996) with minor modifications. Approximately 0.25g of maize leaf tissue for control and infected plants, after week one and week two of infection, were ground in liquid nitrogen to a fine powder. A 2.0ml volume of extraction buffer (0.1M phosphate buffer, pH 7.8, 2mM DTT, 0.1mM EDTA, 1.25mM PEG4000) was added to each sample, thoroughly mixed and then centrifuged for 15 min at 11500rpm (Beckman Coulter, South Africa). The

supernatant was passed through a PD-10 Sephadex column (GE Healthcare, USA) which was equilibrated with three washes of 0.1M phosphate buffer pH7.8 as per manufacturer's instructions. Once the supernatants had passed through the column, the extracts were eluted with 2.5ml of 0.1M phosphate buffer (pH7.8). The flow-through was collected for analysis of the various antioxidants and quantified for total protein concentration.

Total protein concentration of samples were measured according to Bradford's method (1976) using BSA as a standard (Quick Start™ Bradford Protein Assay) and carried out as per manufacturer's instructions. The total protein concentration was measured using the MultiSkan Ex (Thermo Fisher Scientific, USA) plate reader at an absorbance of 595nm.

Superoxide dismutase (SOD, EC 1.15.11)

To measure SOD activity, the procedure by Bailly *et al*. (1996) was followed with minor modifications. The reaction was performed in a microplate reader with reduced volumes and the reaction per well contained: 0.1M phosphate buffer (pH 7.8), 1.3µM riboflavin, 13mM methionine, 63µM nitroblue tetrazolium (NBT) and 50µl extract/water to make up a final volume of 300µl. A microplate containing the mixture was placed under a fluorescent lamp (~300 watts) for 15 min; an identical plate kept in the dark was used as the blank for the experiment. After illumination, the absorbance was measured at 560nm using the MultiSkan GO plate reader (Thermo Fisher Scientific, USA). SOD specific activity was expressed as units SOD (mg protein)$^{-1}$. Here, one unit of SOD represents the enzyme activity which inhibited the photo-reduction of NBT to blue formazan by 50%.

Calculation for superoxide dismutase:

$$\text{Enzyme \% inhibition} = \frac{[(\text{Abs control}) - (\text{Abs sample})]}{\text{Abs control}} \times 100$$

$$\text{Total enzyme activity} = \frac{\text{\% inhibition}}{\text{50\% X protein concentration (mg)}}$$

Catalase (CAT, EC 1.11.1.6)

Catalase activity was measured according to Claiborne (1985); the reaction was carried out at 25˚C in a UV plate and each well contained: 50mM phosphate buffer (pH 7.0), 37.5mM H_2O_2 and 13.33µl extract to make up a final volume of 200µl. CAT activity was measured every second for 5 min, results for the assay were expressed as µmol H_2O_2 catalysed (g protein.sec)$^{-1}$, where a decrease in absorbance of H_2O_2 was observed at 240nm.

Calculation for catalase:

Enzyme activity = (slope (dA/min) X final volume)/ (0.0436 x sample volume)

Enzyme activity per gram = $\frac{\text{(Enzyme activity)x (volume of buffer used for extraction)}}{\text{weight of tissue used in extraction}}$

*0.0436: Millimolar extinction co-efficient of H_2O_2 at 240 nm

Glutathione reductase (GR, EC 1.6.4.2)

Glutathione reductase activity was determined according to Bailly *et al*. (1996); activity was measured at 25˚C every minute for a period of 20 min at 340nm in a microplate by observing the rate of NADPH oxidation. The reaction in each well contained: 0.1M phosphate buffer (pH7.8), 3mM $MgCl_2$, 10mM oxidized glutathione (GSSG), 0.5mM NADPH and 100µl extract to make up a final volume of 200µl. GR activity was expressed as µmol NADPH oxidised (mg protein)$^{-1}$.

Calculation for glutathione reductase:

Enzyme activity = $\frac{\text{(Rate)x (volume of buffer used for extraction)}}{\text{(extract volume) x (extinction co}-\text{efficient of NADPH)}}$

*6.22: Millimolar extinction co-efficient of NADPH at 240 nm

SOD, CAT and GR extract activities were measured from three individual plant samples of the control and infected samples after week one and two of infection; the seed infections and assays were repeated three times for biological validation.

2.2.7 Statistical analysis

All statistical tests and graphs presented in this chapter were generated using GraphPad Prism software version 6 (GraphPad Software Incorporation, 1992-2007). Two-way ANOVA as well as Bonferroni post-tests were performed for the leaf measurements and electrolyte leakage experiments in control and infected samples after week one and two of infection whereas t-tests were performed for the antioxidant assay experiments.

2.3 Results and Discussion

2.3.1 Morphology and leaf measurements

Figure 2.1: Maize plants after two weeks of growth in MS media. Control and Infected plants grown in bottles with MS media (A & B). Control and Infected maize showing leaves, stems and roots (C & D). An enlarged view of Control and Infected maize seeds, respectively (E & F). Red arrow indicates fungus covering infected maize seed.

Maize control and infected seeds were grown in MS media as shown above (Figure 2.1, A & B). A final concentration of $1x10^3$ conidiospores/ml was chosen to infect maize seeds after a range of concentrations were tested (data not shown). By observing the phenotypical characteristics of the plants, a higher concentration ($\geq 1x10^5$ conidiospores/ml) inhibited any growth of the maize plant. In addition, the higher fungal concentrations caused the entire seed to be covered in fungus and resulted in a fungal lawn on the media thereby inhibiting maize germination.

The morphological characteristics of the maize plants were observed after two weeks of infection to determine whether infection with *F. verticillioides* had any effect on the growth of the plants. The MS media was clear thereby enabling the growth of the whole plant including the roots to be visible. A summary of these characteristics can be seen in Table 2.1.

Table 2.1: Summary of the morphological characteristics of control and infected plants after two weeks of growth in MS media

Uninfected (Control)	**Infected (*F. verticillioides*)**
3-4 green leaves, turning brown at tips	3-4 green leaves, turning brown at tips and fungus covering the fourth leaf
All leaves long	Leaves shorter (stunted growth)
Roots (>20) are variable in length, white, thin	Roots (<10) are short, brown, purple & white
Stems are thin, green (slight purple)	Stems are thin, more purple

Both the control and infected plants had 3-4 leaves attached and were showing signs of colour loss and wilting after the two weeks of growth in MS media. However, there were also very distinct differences between these plants where infected plants, when compared to the control, were much greener, had shorter and thinner leaves which were starting to curl inwards and their stems were more purple in colour. Furthermore, infected leaves also had fungal growth on the fourth leaf and had roots that were shorter with visible signs of rotting (brown in colour) (see Table 2.1).

The reason for both the control and infected plants showing signs of colour loss and wilting at two weeks could be due to the decrease in nutrients available in the media after that period or the build-up of various gases in the glass bottles.

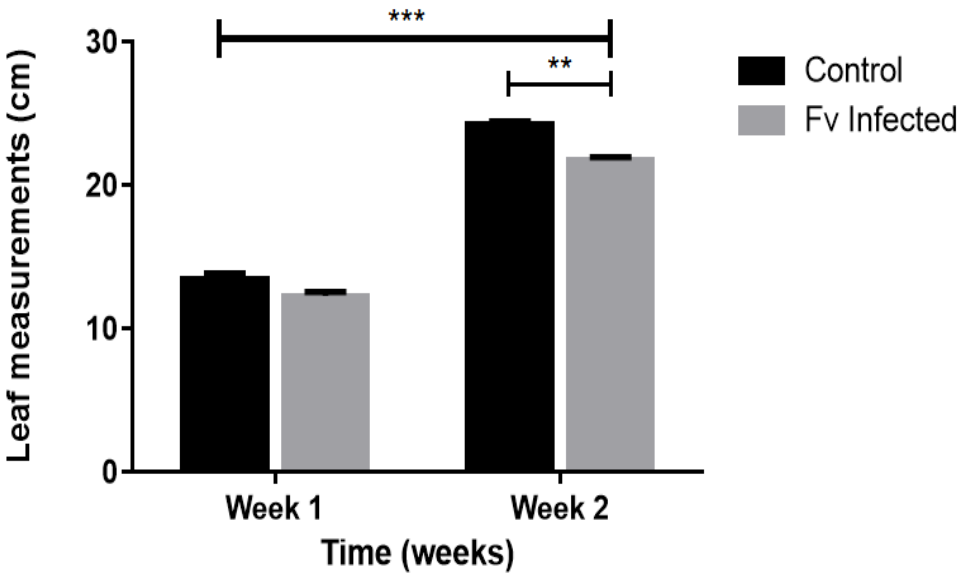

Figure 2.2: The mean leaf measurements of maize plants after week one and two of infection in control and infected plants ($n \geq 6$). Measurements were taken from the three independent infection experiments and then averaged. Error bars indicate standard error of the mean (SEM) with the level of significance indicated by the asterisk (*) where ** $p<0.01$ and *** $p<0.001$ after performing a two-way ANOVA test.

As the size difference between the control and infected leaves were very distinct, leaf measurements were also taken after week one and two of infection to observe the change in leaf growth over that period. Leaves from the same maize plant were used for measurements after both week one and two with the leaves being excised at the time of measurement. Leaves were excised from the whole plant and measurements were taken from the point at which the leaf attaches to the stem up to the tip of the leaf. As shown in Figure 2.2, the lengths of the infected leaves were shorter than the control leaves with the decrease in growth being significantly greater after week two compared to after week one of infection. Overall, over the two-week period, leaf growth was reduced in the infected plants compared to the control plants ($p<0.0001$).

The morphological characteristics and leaf measurements indicate that infection with *F. verticillioides* has negative effects on the growth of the plants over the respective growth period.

2.3.2 Electrolyte leakage

To obtain an estimation of membrane permeability from different regions of maize leaves, conductivity measurements were recorded after fungal stress. The tip, middle and meristem regions of control and infected leaves were subjected to a conductivity meter after week one and two of infection. Electrolyte leakage results from the tip and middle regions of the maize leaves after both week one and two gave variable readings across biological replicates for both control and infected plants and were thus not reported in this section. Leakage from the meristem region of the leaf, the

region which would be the youngest part of the leaf (lowest part of the leaf), gave the most uniform results in terms of biological repeats and is shown in Figure 2.3.

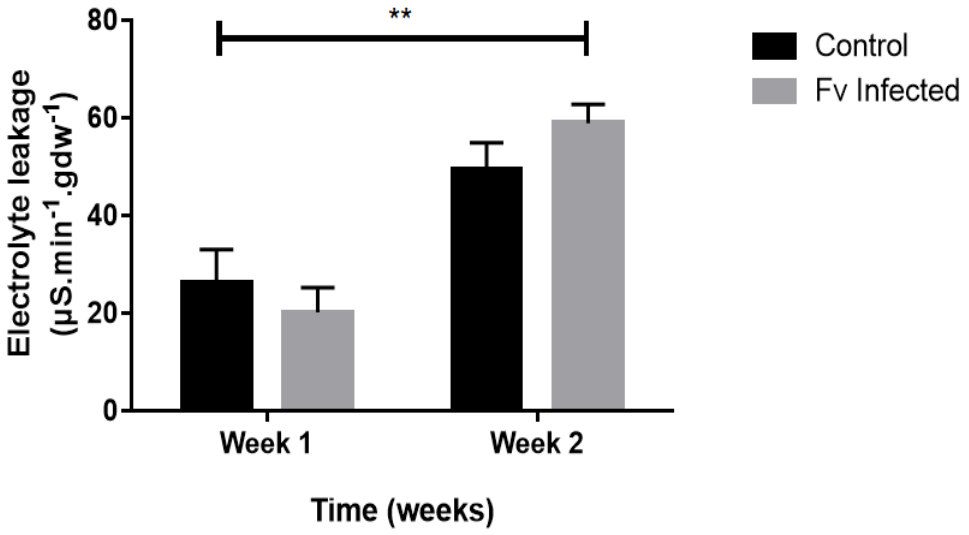

Figure 2.3: Electrolyte leakage measurements from the meristem region of maize leaves after week one and two of infection in control and infected plants ($n \geq 3$). Error bars indicate standard error of the mean (SEM) with the level of significance indicated by the asterisk (*) where ** $p<0.01$ after performing a two-way ANOVA test.

After week one of infection, electrolyte leakage in the infected maize meristem regions were lower when compared to the control, but not at a statistically significant level. The slightly higher electrolyte leakage in the control could be due to biological variation as each maize plant was grown independently in tissue culture medium. After week two of infection, the electrolyte leakage rate for the infected samples (~60µS.min^{-1}.gdw^{-1}) were higher than the control samples (~50µS.min^{-1}.gdw^{-1}) indicating an increase in leakage after *F. verticillioides* infection, however, again this difference was not significant which could also be attributed to the variation in the samples.

When comparing week one of infection of both control and infected plants to week two, there was a significant increase ($p=0.0015$) in electrolyte leakage and this could be attributed to depletion of nutrients, an increase in stress on the plants due to the build-up of various gases during the growth over the two week period, the restriction of space in the bottles as the plants grew or the natural process of ageing of the plants.

Although there are no studies on electrolyte leakage on maize in response to fungal stress, in particular to *F. verticillioides*, a study on wheat in response to infection by the fungus, *Pyricularia oryzae*, found a significant increase in electrolyte leakage between the uninfected and infected

plants. This was in both the susceptible and partially resistant line, 48 and 72 hours post-infection (hpi), respectively; however, leakage was less distinct in the resistant line (Debona *et al*., 2012).

Furthermore, a study observing *F. graminearum* infection in wheat found that an increase in mycotoxin concentration caused an increase in electrolyte leakage in both the susceptible and resistant lines (Miller and Ewen, 1997). The electrolyte leakage result seen with *F. verticillioides* infected maize over the two week period is therefore similar to what was observed in this wheat study.

2.3.3 Antioxidant assays

One of the roles of ROS upon invasion by pathogens is to generate a defense response targeted at the pathogens which is often accompanied by an increase in antioxidant enzymes (Doulis *et al*., 1997; Kumar *et al*., 2009). SOD, GR and CAT are some of the major enzymes involved in the ROS scavenger system with the SOD enzyme constituting the first line of defense (Liu and Huang, 2000; Prochazkova *et al*., 2001; Debona *et al*., 2012).

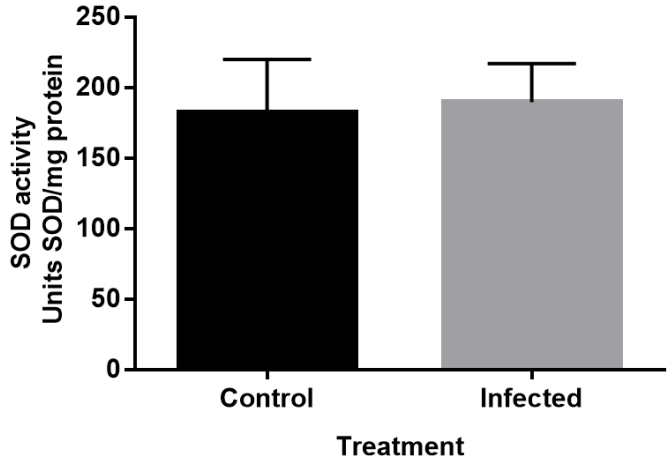

Figure 2.4: Superoxide dismutase (SOD) activity in the leaves of control and infected maize after week two of *F. verticillioides* infection. The enzyme activity is displayed as specific activity (Units SOD/mg protein) measured at a wavelength of 560nm with the assay being performed on three biological replicates (n=3) from both control and infected samples. Error bars indicate standard error of the mean (SEM).

The enzymatic activity of SOD, GR and CAT were measured in maize plants at week one and week two, respectively, after infection with *F. verticillioides*. However, the enzyme activity after week two has only been reported here (Figure 2.4, 2.5 and 2.6). This is because week two showed the greatest response after infection through the electrolyte leakage and morphology experiments. From this point onwards only results for week two have been shown and was the time point chosen for RNA-sequencing (Chapter 3).

Upon analysis, SOD activity in both the control and infected maize plants had increased after two weeks of infection, however, the results showed only a small difference between the two conditions. Both GR and CAT activity also showed an increase in activity in the control and infected maize plants after infection. However, in contrast to the activity seen for SOD, there was a distinct difference in activity between the control and infected plants for both GR (p= 0.086) and CAT (p > 0.05) after infection with *F. verticillioides*. This change in activity was shown not to be significant due to biological variation in the activity of the respective antioxidants that were measured within the individual samples. This is because each maize biological will most likely have differences in quantitative and qualitative antioxidant responses to pathogen attack. Furthermore, this response will depend on the degree of *F. verticillioides* infection in the seeds and may also include other factors that may have either increased or delayed pathogenicity. Even though there was not much change in SOD activity, the difference in CAT and GR activity suggests that there was a response elicited by the plant towards the invading pathogen.

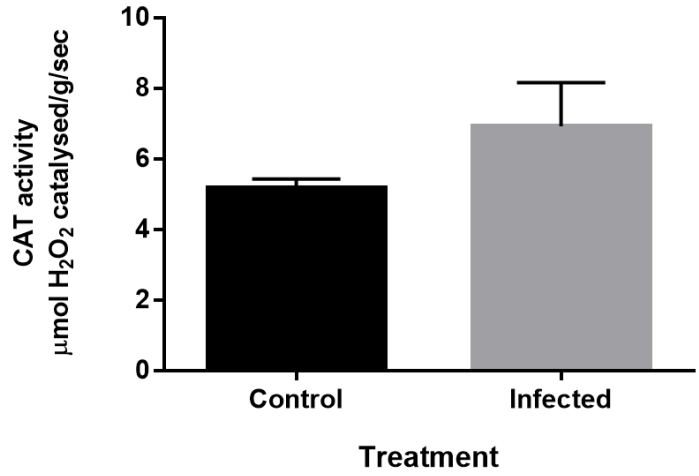

Figure 2.5: Catalase (CAT) activity in the leaves of control and infected maize after week two of *F. verticillioides* infection. The enzyme activity is displayed as specific activity (µmol H_2O_2 catalysed/g/sec) and measured at a wavelength of 240nm with the assay being performed on three biological replicates (n=3) from both control and infected samples. Error bars indicate standard error of the mean (SEM).

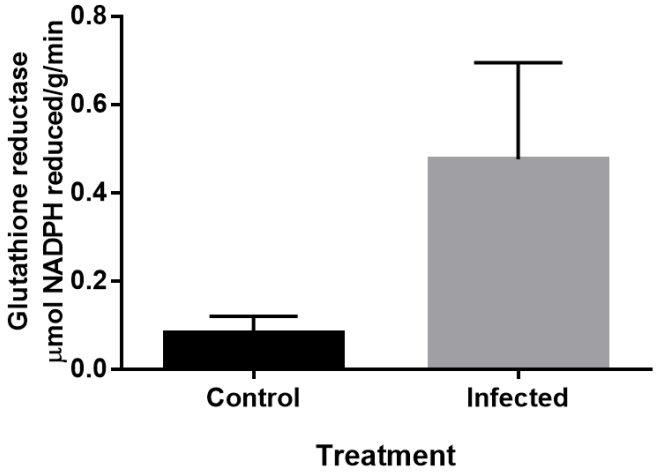

Figure 2.6: Glutathione reductase (GR) activity in the leaves of control and infected maize after week two of *F. verticillioides* infection. The enzyme activity is displayed as specific activity (µmol NADPH reduced/g/min) and measured at a wavelength of 340nm with the assay being performed on three biological replicates (n=3) from both control and infected samples. Error bars indicate standard error of the mean (SEM).

In comparison to the current study, a study by Kumar *et al*. (2009) examined maize plants infected with *F. verticillioides* and showed that there was no significant increase in CAT activity in the shoots. The study also showed a 2.5 and 8-fold significant increase in GR and SOD activity compared to control plants. When examining the roots, CAT activity in maize plants infected with *F. verticillioides* showed a significant 43-fold increase compared to control plants with no significant increase in GR and SOD activity (Kumar *et al*., 2009).

Other studies on maize infected with a *F. verticillioides* strain and other *Fusarium* strains found similar or contrasting results. In these studies, antioxidant enzyme assays were performed on maize seeds 72 hpi, where results showed an increase in GR and a significant increase in CAT activity in infected compared to control plants, however, SOD varied across these experiments (Lanubile *et al*., 2012b; Lanubile *et al*., 2015; Maschietto *et al*., 2016). These results are similar to that found in the current study, with the exception of a decrease in SOD by Maschietto *et al*. (2016), even though the mode of infection, maize line and fungal strains were not the same.

It is important to note that the level of antioxidant enzyme activity by a plant to fungal infection is variable and this can be seen across many studies including those mentioned previously. Additionally, the same maize line may respond differently to various fungal pathogens (Lanubile *et al*., 2012b; Lanubile *et al*., 2015). Comparing the results from the current study to the above mentioned papers proves that different responses can be detected at various regions of the maize plant and that each maize line or various modes of infection might result in completely different or similar responses by the plant to pathogen attack.

The research in the current study was performed on an African maize line, CML144, on which the knowledge of this maize line is fairly limited. More research needs to be conducted by not only looking at antioxidant enzyme activity but other biochemical and physiological assays as well in order to confirm the responses by this maize line to pathogen attack.

2.4 Conclusion for this chapter

In this chapter, morphological, physiological and biochemical experiments were completed in order to determine the effect of *F. verticillioides* infection on maize plants**.** After the two weeks of infection, maize plants showed visible signs of infection which included stunting amongst various other characteristics with a statistically significant difference seen in leaf measurement size over the two week period. Physiological and biochemical experiments showed an increase in activity over the two week growth period; however, the difference between control and infected plants at each time point was not significant. This could have been due to biological variation in the maize plants as they were all grown individually in different bottles of media, the various levels of infection in the different plants or due to the different responses to the pathogen by the individual plants.

Future experiments would comprise of assays done at other time points which include 24 hrs after infection; it would also be of benefit to include a maize line like B73, which has been used in a number of studies in order to have a reference to interpret the results. In addition, it would also be useful to perform both the physiological and biochemical experiments on the roots since the roots seem to be the most affected area of the plant (externally) when using the artificial infection method of this nature.

In summary, successful infection of the plants is evident by looking at the morphological characteristics of the plants even though *F. verticillioides* infection can be symptomless (see Chapter 1). Assay experiments showed changes between control and infected plants, however, without statistical significance at the specific time points which could be due to biological variation within the samples.

Chapter 3

RNA-sequencing and Gene Ontology analysis

3.1 Introduction

Analysing gene expression within the transcriptome is used to determine various mechanisms that regulate cellular processes in plants as well as other species (Van Verk *et al*., 2013). With the continuous development of high-throughput sequencing, there has been an improvement in messenger ribonucleic acid (mRNA) expression profiling methods (Davidson *et al*., 2011). RNA-seq is a next-generation sequencing technology that allows whole-transcriptome profiling and measures the levels of gene expression for all genes present in a genome in a precise manner (Davidson *et al*., 2011; Makarevitch *et al*., 2015).

A typical RNA-seq experiment requires mRNA that is converted to cDNA to form the basis of a library. The DNA fragments in the library are then sequenced with the appropriate next-generation sequencing technology (Van Verk *et al*., 2013). Sequencing reads are then quantified either by alignment to an already sequenced genome or through *de novo* assembly (Yang and Kim, 2015).

In this study, RNA-seq was used to investigate changes in gene expression in maize leaves infected with *F. verticillioides* compared to control (mock-infected) leaves. To our knowledge, there have been no previous studies using whole-transcriptome RNA-seq to study changes in gene expression in an African maize cultivar after infection with *F. verticillioides* using the soak-seed method of inoculation (see Chapter 2). In a study by Lanubile *et al*. (2014), RNA-seq was performed on maize seed material from genotypes with contrasting resistance to *F. verticillioides* (susceptible and resistant) and infected plants using the pin-bar method of inoculation.

The RNA-seq experiment for the current study was performed at the Centre for Proteomic and Genomic Research (CPGR, Cape Town, South Africa) with analysis in Linux using the Tuxedo suite of protocols (Trapnell *et al*., 2012).

The gene expression results from the RNA-seq data obtained in this study was also analysed independently using the bioinformatics platform, DNA Subway, an iPlant Collaborative (recently renamed Cyverse) (http://dnasubway.iplantcollaborative.org/). RNA-seq analysis is run in a workflow known as the Green Line in order to determine differential gene expression. The Green Line in DNA Subway also incorporates the Tuxedo suite of protocols (Williams *et al*., 2014).

For analysis using the Tuxedo suite of protocols, reads are aligned in TopHat to the B73 reference genome. Mapping to the reference genome improves sensitivity and accuracy and also adds a

significant amount of speed to the analysis pipeline. In some cases, when reads do not align correctly (when aligned reads extends a few bases into introns); TopHat corrects this by re-aligning the reads to adjacent exons (Trapnell *et al*., 2012; Kim *et al*., 2013).

After alignment using TopHat, transcript assembly is then performed using Cufflinks with the output revealing transcript isoforms identified by the alignment of the RNA-seq reads to the reference genome. Cufflinks uses the mapped read information from TopHat for transcript assembly. It also allows for abundance estimation, which is based on transcript coverage together with the distribution of fragment length. Abundance is reported as fragments per kilobase of exon per million fragments mapped (FPKM). Cuffdiff, the next step of the Tuxedo protocol, also uses the mapped reads from TopHat to determine differentially expressed genes or transcripts. This is done by calculating the expression levels and then testing the statistical significance of the changes between the conditions being tested. The Cuffdiff results can then be visualised through cummeRbund which plots abundance and differential expression into figures and plots (scatter plots, volcano plots etc.) (Trapnell *et al*., 2012; Yang and Kim, 2015).

RNA-seq studies generate large amounts of data and the output result is typically a list of genes between two groups being compared; in this case mock-infected and infected samples. Making biological sense of these results with their associated pathways and responses is challenging and for this specific reason, various annotation tools exist in order to facilitate interpretation of the sequencing data. The Gene Ontology (GO) database is a useful resource for annotating these large gene lists and includes three classification systems for gene products namely biological function, cellular component and molecular process (Kestler *et al*., 2008; Lohse *et al*., 2014).

Many GO analysis tools exist; these include but are not limited to: Blast2GO, REVIGO, FatiGO, Gorilla and agriGO (Supek *et al*., 2011). For the purpose of this study, we decided to use agriGO (Du *et al*., 2010). AgriGO is an enrichment tool that employs Gene Ontology as the annotation resource, with a focus on agricultural species such as maize. The enrichment analysis approach used in agriGO is a singular enrichment analysis (SEA) method, a traditional and widely used strategy for enrichment analysis. AgriGO also includes statistical testing and works on a more powerful server, making it faster, more flexible, robust and very simple to use (Du *et al*., 2010).

In this chapter, the objective was to extract RNA from control (mock-infected) and *F. verticillioides* infected maize leaves after two weeks of infection and perform quality control tests of the extracted samples in order to perform RNA-seq analysis. The main aim thereafter was to determine whether maize genes are differentially expressed in maize leaves between control and infected samples and

to annotate the data for interpretation. Quality control of RNA quantity and quality was assessed using nanodrop, bioanalyzer and qubit measurements. Data generated from the RNA-seq experiment was analysed with DNA Subway using the Tuxedo suite of protocols and validated against the analysis by CPGR who also used the Tuxedo suite for analysis. RNA-seq analysis allowed identification of differentially expressed genes (DEGs) between the control and infected samples after *F. verticillioides* infection. The GO-analysis using agriGO was based on DEGs found between the DNA subway analysis and the analysis conducted by the CPGR. Lastly, the RNA-seq results of the current study were compared to the study by Lanubile *et al*. (2014), though this study used a different method of *F. verticillioides* infection and different maize genotypes as compared to the current study.

3.2 Materials and Methods

3.2.1 Maize seed infection:

Sterilised maize seeds were infected with *F. verticillioides* via artificial inoculation as described in Chapter 2. The planted control and infected seeds in the MS media were incubated under controlled conditions at 28°C (16hr light, 8hr dark; light intensity of 140µmol/m^2/s and 60% humidity) for a period of 14 days (maize V3 leaf stage).

3.2.2 RNA extraction:

At the maize V3 leaf stage, the control and infected maize leaves were harvested in liquid nitrogen and RNA was extracted. All of the material used for the RNA extractions was double-autoclaved and the solutions for extraction were prepared using diethylpyrocarbonate (DEPC) water. Using a mortar and pestle, leaves were ground in liquid nitrogen until a fine powder was achieved. Approximately 500mg of the ground tissue was transferred to sterile 2ml eppendorfs containing 1ml of TRI Reagent (Zymo Research, USA). Samples were inverted carefully and then vortexed at room temperature for 10-15 min. The samples were incubated at room temperature for 5 min allowing complete dissociation of nucleoprotein complexes. Thereafter, 200µl of chloroform was added to each tube, and the samples were inverted 30 times and incubated for 3 min at room temperature. The samples were then centrifuged for 15 min at 4°C at 14000rpm (Vision Scientific). The upper aqueous phase was transferred to a new eppendorf tube containing 1ml isopropanol; samples were inverted, incubated for 10 min at room temperature and then centrifuged for 15 min at 4°C at 14000rpm. The supernatant was discarded and the pellets were washed with 1ml of cold 75% (v/v) EtOH and then centrifuged for 5 min at 4°C at 14000rpm.The supernatant was discarded and the pellet was air-dried for 5 min. The RNA pellet was resuspended in 100µl of DEPC water through incubation at 55°C

for 5 min. The samples were quantified using the NanoDrop ND-1000 (NanoDrop Technologies, USA) and then screened by gel electrophoresis on a 1.2% (w/v) agarose gel containing Ethidium Bromide (EtBr) at 100 volts (V) for 30 min. The gel was visualised using the ChemiDoc™ XRS system with Image Lab™ software (Bio-Rad). The RNA samples were stored at -70°C until further use.

3.2.3 DNase treatment and RNA purification:

Each RNA extraction sample was DNase treated as per manufacturer's instructions (Thermo Fisher Scientific) and incubated for one hr at 37°C. The samples were then purified as follows: to each eppendorf, an equal volume of phenol:chloroform-isoamyl alcohol was added (25:24:1). The samples were mixed by inversion and then centrifuged for 10 min at 4°C at 14000rpm. The upper aqueous phase was then removed and placed into a new eppendorf tube. A 0.1x volume of 3M sodium acetate (pH 5.2) was added to each tube and mixed by inversion. This was followed by a 5x volume of 100% EtOH, with the samples mixed by inversion and then centrifuged for 10 min at 4°C at 14000rpm. The supernatant was discarded and the pellet was allowed to dry for $\geq$15 min. The pellet was then resuspended in 20µl of DEPC water.

To remove TRI reagent and phenol contaminants from the samples, a method by Krebs *et al*. (2009), was used with minor modifications. To the RNA samples, 200µl of water-saturated 1-butanol was added, mixed thoroughly and centrifuged at 10000rpm for 30 sec. The organic upper layer was carefully removed and discarded; a thin layer of 1-butanol was left in the tube to ensure that no sample was lost. This was repeated six times in order to remove phenol and TRI reagent. Next, 200µl water-saturated diethyl ether was added to the samples, mixed thoroughly and centrifuged at 10000rpm for 30 sec. The organic upper layer was removed and discarded, leaving a thin layer behind to ensure that no sample was lost. This step was repeated once and the tubes were left open in the fume hood for 15 min in order for the diethyl ether to evaporate.

The RNA was quantified as previously described and the integrity was assessed on a 1.2% EtBr stained agarose gel by electrophoresis. The RNA was analysed further on a Qubit 2.0 fluorometer and on an RNA-6000 Nano chip using the Agilent 2100 BioAnalyzer (see next section).

3.2.4 Quality control of RNA-sequencing samples:

Samples submitted for RNA-sequencing

A total of six maize RNA samples isolated from control (3) and infected (3) plants were submitted for RNA-sequencing (RNA-seq) at the Centre for Proteomic and Genomic research (CPGR). The RNA samples were extracted and quantified as described in section 3.2.3. At the CPGR, samples were re-

analysed for the purpose of quality control; this included using the NanoDrop ND 1000 (Thermo Fisher Scientific) to assess the purity and quality of the samples, Qubit® 2.0 Fluorometer for absolute quantification and the BioAnalyzer RNA 6000 Pico Assay Kit (Agilent) to assess RNA integrity using the RIN values.

Sequencing of samples using the NextSeq 500/550 Mid Output Kit v2 was carried out using the NextSeq 500 sequencer instrument (Illumina).

Library preparation

Total RNA (~1µg) was treated with Ribo-Zero rRNA Removal kit (Plant) and then purified using the RNeasy® MinElute® Cleanup kit (Qiagen). Using the ScriptSeq™ Complete kit (Plant Leaf) and ScriptSeq™ Index PCR primers (Illumina), indexed libraries were prepared. The size of the libraries was profiled using the BioAnalyzer High Sensitivity Assay kit (Agilent) and then quantified using the Qubit HS DNA assay kit (Thermo Fisher Scientific).

Template preparation and sequencing

The prepared libraries were diluted to 10nm before equimolar amounts were pooled. This pooled library was then diluted to 4nm using 2N NaOH and 200mM Tris HCL, pH7. This was further diluted to 20pM using pre-chilled hybridization buffer and then to a final loading dilution of 1.6pM. A 1% PhiX sequencing control was spiked in the pool library and then samples were loaded onto the NextSeq 500/550 cartridge.

3.2.5 RNA-sequencing data analysis:

RNA-seq data was analysed using the online bioinformatics tool, DNA Subway, further described as Protocol 1. These results were then validated against another analysis conducted by the CPGR which will be described as Protocol 2 (Figure 3.1).

3.2.5.1 Analysis using Protocol 1:

Reads from the RNA-seq experiment received from CPGR were analysed using the DNA Subway's Green Line RNA-seq workflow, which incorporates the Tuxedo suite of protocols (Trapnell *et al*., 2012). The paired-end reads received from the CPGR were aligned to the *Zea mays* B73 AGPv2 genome.

Reads were processed using the FastX toolkit, where reads with lower quality scores (<20) were filtered out. Individual reads were aligned to the reference genome using TopHat version 2.0.11 to

determine intron/exon boundaries and expression levels were calculated using CuffDiff version 2.1.1 (p<0.01) to determine differential expression between the control and infected samples (Figure 3.1).

3.2.5.2 Analysis using Protocol 2:

Raw data QC

For each sample type (i.e. three control and three infected), FastQ files were generated on the NextSeq 500/550 in one single run. The quality of the raw data for each input file was assessed using FastQC (Andrews, 2010).

Trimming and filtering the reads

Raw sequencing reads were trimmed using Trimmomatic v 0.32, this algorithm includes steps for read trimming as well as filtering (Bolger *et al*., 2014). Adapters were clipped from the 3' end of raw sequencing reads, low-quality ends from the reads were trimmed as well before continuing with any downstream analysis processes. The adapters were trimmed from the 3' ends and bases were removed from the 5' ends. The bases with a quality score of <30 were clipped from both the 3' and 5' ends. Reads were processed in a 5' to 3' direction with a final length of <50 bases.

Read alignment, transcript assembly and quantification

Processed reads were aligned to the *Zea mays* B73 v3 reference genome using TopHat version 2.0.13 (Kim *et al*., 2013). Transcripts were then assembled using Cufflinks (Trapnell *et al*., 2010). Levels of gene expression were analysed using Cuffdiff to determine differentially expressed genes (DEGs) in the infected samples compared to the control samples (Trapnell *et al*., 2012) (Figure 3.1).

3.2.6 Gene annotation using Maize Microarray Annotation Database, Plant Ensembl and NCBI:

The Maize Microarray Annotation Database (Coetzer *et al*., 2011) is a resource that allows for the annotation of probes on a maize Agilent microarray. Gene ID's from the RNA-seq experiment were used to determine the functional annotation of the genes. This was performed using the B73 v2 genome as a reference with data being annotated as per Blast2GO description. Differentially expressed genes were also annotated using Plant Ensembl with the *Zea mays* B73 AGPv3 genome as the reference. For those genes not found in Plant Ensembl, the National Center for Biotechnology Information (NCBI) platform was used for annotation. Using NCBI's gene search, the unidentified genes were searched using the *Zea mays* B73 v3 genome (Figure 3.1).

3.2.7 Gene ontology analysis using agriGO

Using the Gene IDs from the RNA-seq analysis, agriGO (Du *et al*., 2010) was used to analyse gene ontology. This was a single enrichment analysis (SEA) using the *Zea mays* AGPv3.30 genome as a reference. The statistical method used in agriGO was the Fisher's Exact Test with multi-test adjustment (Yekutieli-FDR under dependency) with significance level set at $p < 0.05$ (Figure 3.1).

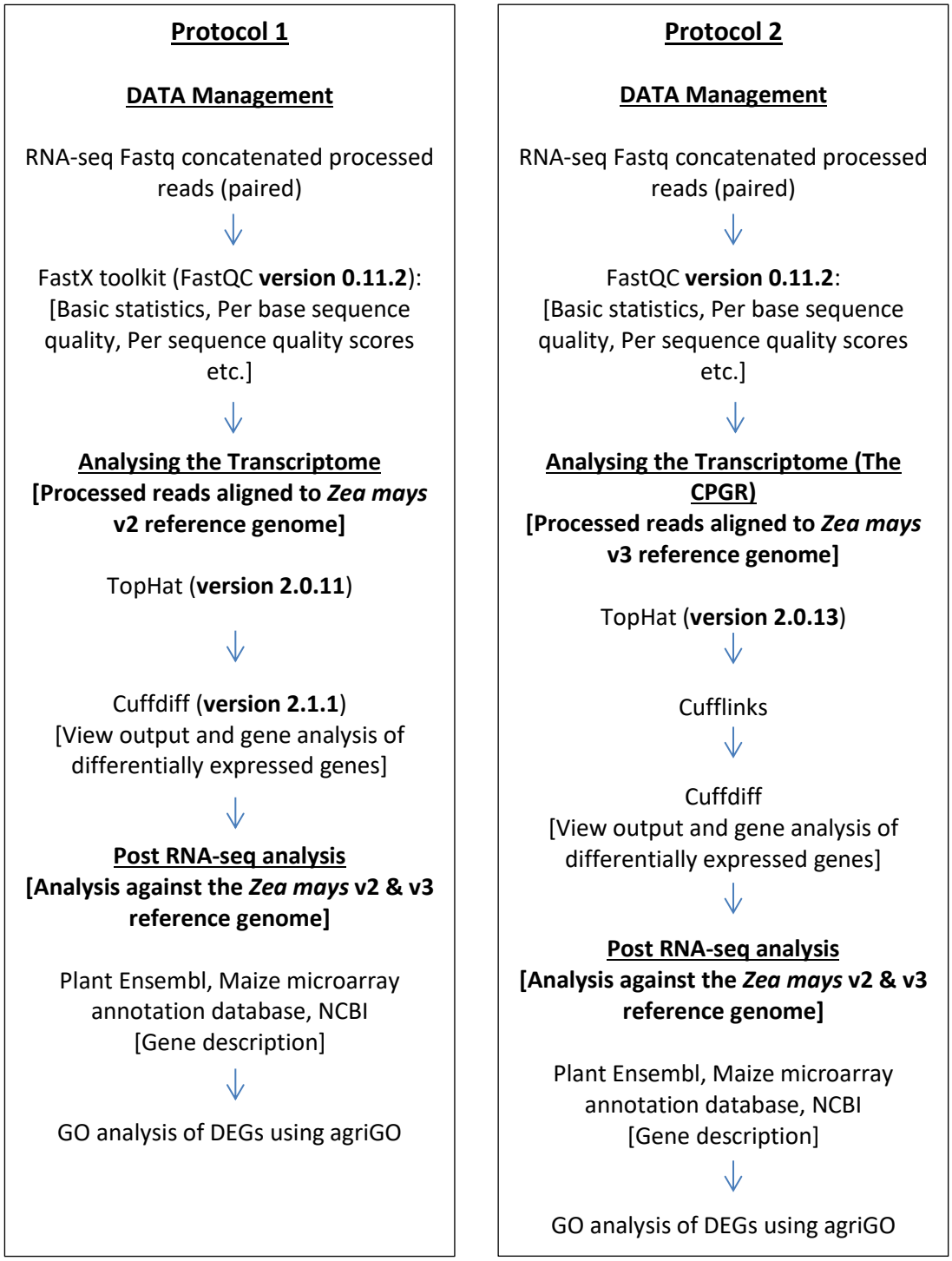

Figure 3.1: Outline of the steps involved in analysing RNA-seq data using Protocol 1 (Tuxedo analysis in DNA subway with mapping to v2 genome) and Protocol 2 (Tuxedo analysis with mapping to v3 genome; conducted by the CPGR).

3.3 Results and Discussion

3.3.1 RNA extraction and quality control

Good quality RNA was required in order to prepare for RNA-seq as well as quantitative Real-Time PCR (see Chapter 4). Maize samples from control and infected whole maize leaves that were extracted, DNase treated and purified were assessed using the Nanodrop ND-1000 (NanoDrop Technologies, USA) and visualised on a 1.2% (w/v) EtBr agarose gel (data not shown). The DNase treatment was required to remove any genomic DNA contamination and the purification was to allow the removal of any phenolic components present in the samples after the addition of TRI Reagent. Nanodrop concentration readings and the A260/230, A260/A280 ratios are important parameters for RNA-seq (Table 3.1). As seen in Table 3.1, the 260/230 ratios range from 1.72-2.07 and the 260/280 ratios range from 2.02-2.10, these ratios show that the control and infected RNA samples are of good quality with concentrations that can be used for downstream applications.

For the purpose of quality control before RNA-seq, samples sent to the CPGR were analysed using the Qubit 2.0 fluorometer and on a RNA-6000 Nano chip using the Agilent 2100 BioAnalyzer (Table 3.1). This was performed to ensure that optimal data could be generated from the output. The Qubit assay kit uses specific binding probes; this allows it to provide a more accurate reading of the RNA concentrations in the samples compared to the nanodrop readings. These results showed that the RNA samples were of good quality as seen in Table 3.1. The BioAnalyzer on the other hand was used to measure the integrity of the samples before performing the downstream experiments. Results are visualised as an electropherogram (Appendix A, Figure A1) accompanied by a RNA integrity number (RIN).

The RIN number is based on the integrity of the RNA samples and is important when conducting gene expression studies like RNA-seq. Integrity is important in studies that capture gene expression occurring in RNA samples at the point of extraction (Schroeder *et al*., 2006). A passing RIN is within the range of 7-10.

The RIN numbers of the RNA samples for RNA-seq can be seen in Table 3.1; all samples passed quality control and had a RIN $\geq$ 7 except for the infected sample 3 which just missed the threshold (RIN=6.9). However, the baseline of the electropherogram (Appendix A, Figure A1) was not raised, therefore the RNA was considered to be intact and this sample was included with the others for RNA-seq.

Table 3.1: Nanodrop, BioAnalyzer and Qubit results of control and infected RNA samples after two weeks infection for the purpose of quality control before RNA-seq analysis.

Sample name	Nanodrop			BioAnalyzer	Qubit
	ng/µl	A260/A230	A260/A280	RIN	ng/µl
Control 1	269.4	2.07	2.10	7.0	36.4
Control 2	224.6	1.82	2.02	7.4	144
Control 3	262.8	1.76	2.07	7.6	252
Infected 1	1160.5	2.05	2.10	7.0	122
Infected 2	137.4	1.72	2.04	8.1	150
Infected 3	1105.58	1.94	2.09	6.9	106

3.3.2 RNA-sequencing analysis

RNA-seq was used for whole transcriptome profiling of control and infected maize plants after two weeks of infection with *F. verticillioides*. The aim of this analysis was to investigate differential gene expression in the infected plants compared to the control. For each sample, the total RNA, as seen in Table 3.1 was subjected to sequencing using the concentrations obtained from the Qubit results and then analysed as described using DNA Subway, which integrates the Tuxedo suite of tools. These results were then compared to the analysis performed at the CPGR, the reason for this comparison was to gain confidence in results obtained in DNA Subway and to compare the analysis between two different platforms which used different versions of the B73 maize genome (v2 vs. v3, Figure 3.1).

3.3.2.1 Quality control of RNA-seq reads in Protocol 1

For each of the input RNA-seq files, quality of the data was assessed using FastQC (see Materials and Methods). After quality assessment, reads were trimmed and filtered, Figure 3.2 shows the per base quality scores of only one control and one infected sample after trimming as part of the pre-processing quality control using Protocol 1. The graph shows the quality across all bases at each position, with a higher Phred score indicating greater quality, thus the higher the score the better the base call.

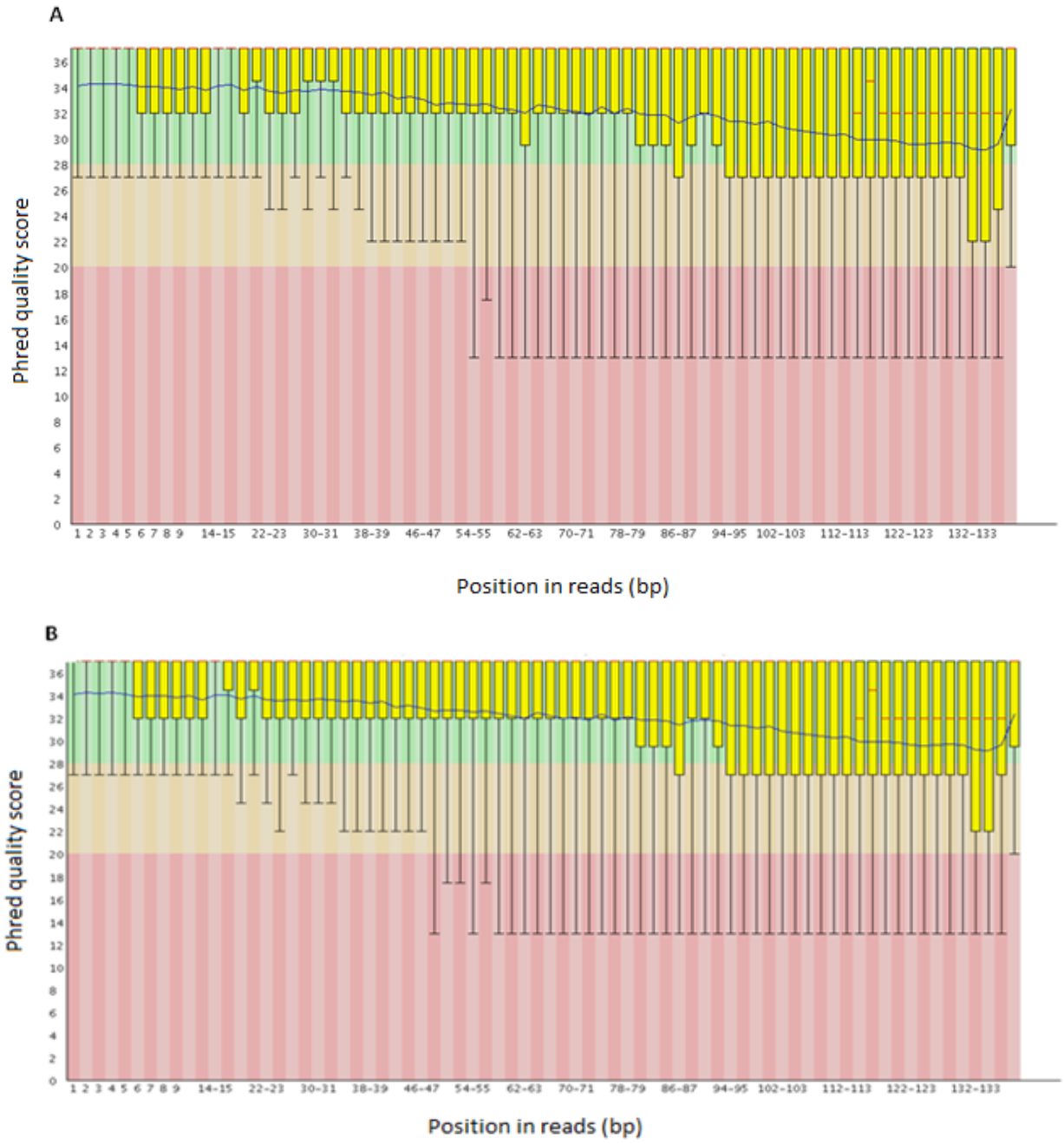

Figure 3.2: Per base quality scores of one control (A) and one infected (B) sample as part of the pre-processing quality control using Protocol 1 (FastQC in DNA Subway). Mean (blue line) and the median (red line) shows quality is above Phred quality score of 20.

Quality scoring is divided into three regions, i.e. green, orange and red. Data located in the green region of the graph indicates good quality data whereas the red region indicates data of poor quality. The RNA-seq data of the control and infected sample shown in Figure 3.2 are mostly within the green region indicating good quality data (higher Phred quality scores). Read quality graphs generated using Protocol 2 showed an identical Phred score distribution pattern as the examples shown in Figure 3.2 (data not shown).

Mapping of sequencing reads to the maize genome

TopHat was used to align the processed reads to the *Zea mays* B73 v2 reference genome for Protocol 1 and to the v3 reference genome for Protocol 2. Table 3.2 and 3.3 shows the mapping rate success of the control and infected samples to the reference genome. Overall mapping rate depicts the percentage of paired reads that aligned to the B73 reference genome.

Table 3.2: Number of reads from the control and infected samples successfully mapped to the *Zea mays* B73 v2 reference genome and overall mapping rate (%) of the reads using Protocol 1 for RNA-seq analysis.

Sample	No. of mapped reads		Overall mapping rate (%)
	Left	**Right**	
Control 1	10293607	11066641	62.6
Control 2	13072640	14726719	66.7
Control 3	11102335	12638559	67.4
Infected 1	9757579	10159711	62.6
Infected 2	10542335	11764997	65.9
Infected 3	12159647	14237213	70.3

Table 3.3: Number of reads from the control and infected samples successfully mapped to the *Zea mays* B73 v3 reference genome and overall mapping rate (%) of the reads using Protocol 2 for RNA-seq analysis.

Sample	No. of mapped reads		Overall mapping rate (%)
	Left	**Right**	
Control 1	10355561	11133885	63.0
Control 2	13131727	14801185	67.0
Control 3	11157732	12705814	67.8
Infected 1	9819472	10218423	63.0
Infected 2	10567766	11790914	66.1
Infected 3	12226508	14312632	70.8

The number of successfully mapped reads as determined using Protocol 1 was compared to the number of successfully mapped reads using Protocol 2. This comparison showed that the mapping results from both protocols were quite similar, with a <1% difference in the overall mapping rate of the paired reads between Protocol 1 and Protocol 2. The difference in mapping percentages

between the different protocols could be due to the fact that different versions of the maize genome were used for the alignment.

3.3.2.2 Quantification of differentially expressed genes (Cuffdiff)

It is important to assess the distribution of expression levels across sample types as well as replicates and it is expected that the majority of genes in both treatment groups (control and infected) will be expressed at similar levels. To assess expression, Cufdiff was used; this step of the Tuxedo protocol also uses the mapped reads from TopHat to determine any differential expression by calculating expression levels and testing the significance of the observed changes between two conditions (Trapnell *et al*., 2012). In this study, it was used to determine differential expression between the control and infected groups.

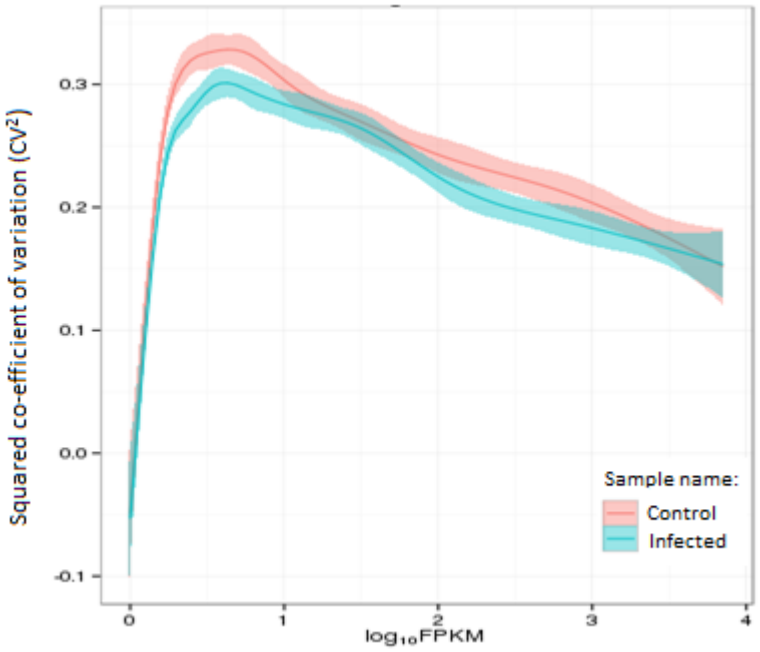

Figure 3.3: Co-efficient of variation across expression levels for both control and infected samples at the gene level using Protocol 1.

After Cuffdiff was performed, the difference in variation can also be seen by plotting the squared co-efficient of variation (CV^2) of fragments per kilobase of exon per million fragments mapped (FPKM) values against the expression level of each transcript obtained from cummeRbund. The presence of high CV^2 values would indicate that there is a lower power to detect differential expression of candidate genes between the treatments. This is due to variation between the replicates within a group being high and therefore it would be difficult to evaluate differential expression between the

groups. The CV^2 plot in Figure 3.3 shows that the variation among the samples in each experimental group is relatively low indicating that biological reproducibility of the samples is high and that differential expression tests can therefore be performed. A similar trend is seen for both the control and infected samples, with the CV^2 values decreasing as the FPKM expression levels rise.

The CV^2 plot from the analysis using Protocol 2 showed a similar trend as the analysis using Protocol 1 and can be seen in Appendix A (Figure A2). These plots provide confidence in the results between the two analysis protocols and allows candidate DEGs to be identified.

3.3.2.3 Candidate differential expression

Of all the reads mapped to the B73 v2 genome using Protocol 1, there was a total of 128 candidate genes found to be significantly differentially expressed in the infected group compared to the control group (q<0.01). The q-value is the p-value adjusted for multiple testing using false discovery rate (FDR) which is also known as the Benjamini-Hochberg method. This method is used to correct p-values calculated during multiple testing (Benjamini and Hochberg, 1995). Of these 128 candidate genes, 53 were down-regulated and 75 were up-regulated, respectively (Table 3.4).

Analysis from Protocol 2, when reads were mapped to the B73 v3 reference genome showed 156 candidate genes that were found to be significantly differentially expressed in the infected group compared to the control group (q<0.01). Of these 156 candidate genes, 50 were down-regulated and 106 were up-regulated, respectively. More genes were found to be differentially expressed using Protocol 2 compared to Protocol 1; however, the gene matches between these protocols were quite high (Table 3.4).

The significant DEGs resulting from the separate analysis performed using Protocol 1 as well as 2 can be seen in Appendix A, with the up-regulated and down-regulated genes shown in Appendix A, Table A1 & A2 and Table A3 & A4, respectively. The tables show the genes with respect to the corresponding chromosome they belong to, log2 fold change, p-values, q-values as well as the gene description as determined by the NCBI gene search, Plant Ensembl and Maize Microarray Annotation (in the form of Blast2GO description) databases.

As previously mentioned, RNA-seq data analysis was conducted using DNA Subway, with data described throughout this chapter, data was also analysed by the CPGR and served as a validation analysis. For this reason, the significant DEGs found using Protocol 1 were compared to those significant DEGs found using Protocol 2.

It should be noted that the *Zea mays* B73 v2 genome was used for RNA-seq analysis using Protocol 1 and not the v3 genome as used with Protocol 2. The reason for this being that at the time of RNA-seq analysis (April 2016), the DNA Subway RNA-seq Green Line module was stable using the v2 genome.

Table 3.4: Comparison of genes shown to be up-regulated and down-regulated after RNA-seq using Protocol 1 and Protocol 2 that incorporates the Tuxedo suite of analysis and the matching genes between these protocols.

	Genes using Protocol 1	Genes using Protocol 2	Matches between Protocol 1 and 2
Up-regulated	75	106	54
Down-regulated	53	50	29
Total	128	156	82

Upon analysis, whereby the DEGs from Protocol 1 and Protocol 2 were compared, it was found that there were 53 and 29 up- and down-regulated genes that matched between these analysis tools (Table 3.4). This comparison showed that an overall 64% of the genes found in Protocol 1 were also found in Protocol 2 (72% up-regulated and 54.7% down-regulated). The differences between these protocols could be explained by the two different versions of the B73 genome being used for analysis or due to different Tuxedo suite versions being used. The genes identified as matching in both the Protocol 1 and 2 analyses can be seen in Table 3.5 and Table 3.6.

Table 3.5: Significantly up-regulated genes matches between Protocol 1 and Protocol 2 after *F. verticillioides* infection as detected using RNA-seq with the Tuxedo suite of analysis tools and mapping to the maize B73 genome. Table shows annotation of genes as described in Plant Ensembl, NCBI and Maize Microarray Annotation database (Blast2GO) with associated enriched GO-terms.

Gene stable ID	Average log2 fold-change	Ensembl / NCBI gene description	Blast2GO description	Enriched GO-term
GRMZM2G049538	4.56	Acyclic sesquiterpene synthase	ent-kaurene synthase b	GO: 0043167 GO: 0046872 GO: 0043169
GRMZM2G119975	3.98	Uncharacterised LOC103646336	N/A	
GRMZM2G029219	3.45	Carbohydrate transporter/ sugar porter/ transporter	major facilitator superfamily antiporter/ carbohydrate transporter sugar porter transporter	
GRMZM5G874955	3.33	Uncharacterised protein	pdr-like abc transporter	
GRMZM2G062724	3.27	Uncharacterised protein	chy zinc finger family expressed	GO: 0043167 GO: 0046872 GO: 0043169
GRMZM2G026922	3.27	Hypothetical protein	acetylglutamate kinase	
GRMZM2G117971	3.25	Uncharacterised protein	pathogenesis-related protein 4	
GRMZM2G087875	3.20	Putative cytochrome P450 superfamily protein; Uncharacterised protein	cytochrome p450 family expressed	GO: 0043167 GO: 0046872 GO: 0043169
GRMZM2G143139	3.17	N/A	N/A	
GRMZM2G149422	3.07	Hypothetical protein	phi-1	
GRMZM2G154523	3.06	Patatin T5; Uncharacterised protein	N/A	

Gene stable ID	Average log2 fold-change	Ensembl / NCBI gene description	Blast2GO description	Enriched GO-term
GRMZM2G098346	3.04	Alcohol dehydrogenase 2	alcohol dehydrogenase	GO: 0043167 GO: 0046872 GO: 0043169
GRMZM5G892675	3.01	Uncharacterised protein	N/A	
GRMZM2G137861	2.99	Wall-associated receptor kinase 2-like	N/A	
GRMZM2G443728	2.95	Potassium transporter 10	potassium transporter 10	
GRMZM2G070011	2.85	Uncharacterised protein; Vignain	vignain precursor	
GRMZM2G006973	2.73	Uncharacterised protein	N/A	GO: 0043167 GO: 0046872 GO: 0043169
GRMZM2G036464	2.72	Glutamine synthetase root isozyme 4	glutamine synthetase	
GRMZM2G115451	2.63	Uncharacterised protein	neutral alkaline invertase	
GRMZM2G178546	2.61	Trehalose-phosphate phosphatase	N/A	
GRMZM2G099049	2.56	N/A	N/A	
GRMZM2G063431	2.53	N/A	N/A	
GRMZM2G427815	2.51	Uncharacterised protein	Peroxidase	GO: 0043167 GO: 0046872 GO: 0043169
GRMZM2G062531	2.50	Uncharacterised protein	c-4 sterol methyl oxidase	GO: 0043167 GO: 0046872 GO: 0043169

Gene stable ID	Average log2 fold-change	Ensembl / NCBI gene description	Blast2GO description	Enriched GO-term
GRMZM2G093826	2.49	Potassium high-affinity transporter	high-affinity potassium transporter	
GRMZM2G110504	2.46	Uncharacterised LOC100278648	hypothetical protein LOC100278648 [Zea mays]	
GRMZM2G007151	2.42	Uncharacterised protein	endomembrane-associated protein	
GRMZM2G130173	2.41	Metallothionein-like protein type 2; Uncharacterised protein	N/A	GO: 0043167 GO: 0046872 GO: 0043169
GRMZM2G477503	2.36	Uncharacterised protein	sulfolipid synthase	
AC217947.4_FG002.2	2.24	N/A	N/A	GO: 0043167 GO: 0046872 GO: 0043169
GRMZM2G026470	2.22	Soluble inorganic pyrophosphatase; Uncharacterised protein	soluble inorganic pyrophosphatase	GO: 0043167 GO: 0046872 GO: 0043169
GRMZM2G091456	2.21	Putative Uncharacterised protein	squalene expressed, squalene monooxygenase	
GRMZM2G366681	2.11	Hypothetical protein	N/A	
GRMZM2G034152	2.07	Polyamine oxidase	polyamine oxidase precursor	
GRMZM2G130149	2.07	Uncharacterised protein	myb family transcription expressed	

Gene stable ID	Average log2 fold-change	Ensembl / NCBI gene description	Blast2GO description	Enriched GO-term
GRMZM2G144097	2.04	Uncharacterised protein	Protein	
GRMZM2G125669	2.03	Alternative oxidase	alternative oxidase	GO: 0043167 GO: 0046872 GO: 0043169
GRMZM2G144083	2.02	Putative ATP dependent copper transporter	heavy metal p-type atpase	GO: 0043167 GO: 0046872 GO: 0043169
GRMZM2G034302	2.01	Uncharacterised protein	sucrose transporter	
GRMZM2G036217	1.91	Uncharacterised protein	fatty acyl coa reductase	
GRMZM2G116079	1.86	Uncharacterised protein	N/A	GO: 0043167 GO: 0046872 GO: 0043169
GRMZM2G076537	1.82	Polynucleotidyl transferase, ribonuclease H-like superfamily protein	exonuclease family protein	
GRMZM2G176433	1.80	Putative Uncharacterised protein	N/A	
GRMZM2G008247	1.78	Beta-glucosidase2	N/A	
GRMZM2G070659	1.77	Hypersensitive-induced response protein	hypersensitive-induced response protein	
GRMZM2G147243	1.72	IAA17-auxin-responsive Aux/IAA family member; Uncharacterised protein	transcription factor	
GRMZM2G099767	1.63	ATMAP70-2	N/A	
GRMZM2G057823	1.55	Fructose-bisphosphate aldolase, cytoplasmic isozyme	fructose-bisphosphate aldolase	GO: 0043167 GO: 0046872 GO: 0043169

Gene stable ID	Average log2 fold-change	Ensembl / NCBI gene description	Blast2GO description	Enriched GO-term
GRMZM2G040369	1.54	Elongation factor 2	Protein	GO: 0043167 GO: 0046872 GO: 0043169
GRMZM2G168552	1.53	Bundle sheath cell specific protein 1	N/A	
GRMZM2G020146	1.51	Uncharacterised protein	serine carboxypeptidase iii precursor	
GRMZM2G473001	1.48	Phosphoenolpyruvate carboxylase 2	phosphoenolpyruvate carboxylase	
GRMZM2G113332	1.42	Uncharacterised protein	copper chaperone	GO: 0043167 GO: 0046872 GO: 0043169
GRMZM2G141353	1.41	Uncharacterised LOC100194210	N/A	

Table 3.6: Significantly down-regulated genes matches between Protocol 1 and Protocol 2 after *F. verticillioides* infection as detected using RNA-seq with the Tuxedo suite of analysis tools and mapping to the maize B73 genome. Table shows annotation of genes as described in Plant Ensembl, NCBI and Maize Microarray Annotation database (Blast2GO) with associated enriched GO-terms.

Gene stable ID	Average log2 fold-change	Ensembl/ NCBI gene description	Blast2GO description	Enriched GO-term
GRMZM2G070172	-3.97	Uncharacterised protein	N/A	GO: 0005975 GO: 0050896 GO: 0042221
GRMZM2G468111	-3.58	Uncharacterised LOC100277849	N/A	
GRMZM2G176430	-3.19	Uncharacterised protein	sodium-dicarboxylate cotransporter	
GRMZM2G472248	-3.17	Protein induced upon tuberization	N/A	GO: 0050896
GRMZM2G004161	-2.92	Uncharacterised protein	N/A	GO: 0050896 GO: 0042221
GRMZM2G133675	-2.65	Putative HLH DNA-binding domain superfamily protein; Uncharacterised protein	amelogenin precursor like protein	
GRMZM2G125775	-2.40	AN17	arsenite inducuble rna associated protein aip-701	GO: 0050896 GO: 0042221
GRMZM2G058612	-2.39	F-box/LRR-repeat protein 3-like	N/A	
GRMZM2G124495	-2.38	Putative MYB DNA-binding domain superfamily protein; Transfactor; Uncharacterised protein	N/A	
GRMZM2G078472	-2.32	Asparagine synthetase	asparagine synthetase	

Gene stable ID	Average log2 fold-change	Ensembl/ NCBI gene description	Blast2GO description	Enriched GO-term
GRMZM2G097641	-2.27	Sucrose-phosphatase 2	sucrose phosphate synthase	GO: 0005975 GO: 0050896 GO: 0042221
GRMZM2G422955	-2.26	N/A	N/A	
GRMZM2G181081	-2.21	CIPK-like protein 1	cipk-like protein expressed	GO: 0050896 GO: 0042221
GRMZM2G358153	-2.09	Chitinase 1; Uncharacterised protein	chitinase 1	GO: 0005975
GRMZM2G147687	-2.02	Uncharacterised protein	glycosyl hydrolase family 3 n terminal domain containing expressed	GO: 0005975
GRMZM2G053669	-1.99	Asparagine synthetase	asparagine synthetase	
GRMZM2G079381	-1.86	Ferredoxin--nitrite reductase, chloroplastic	nitrite reductase	GO: 0050896 GO: 0042221
GRMZM2G173085	-1.85	Lipase/lipooxygenase, PLAT/LH2 family protein	potential zinc finger protein	
GRMZM2G121264	-1.74	Uncharacterised protein	cytochrome p450	
GRMZM2G103812	-1.70	Uncharacterised protein	selenium-binding protein	GO: 0050896 GO: 0042221
GRMZM2G154278	-1.68	Pre-mRNA-splicing factor cwc15	N/A	
GRMZM2G478568	-1.63	Nicotianamine synthase 3	nicotianamine synthase 3	
GRMZM2G024733	-1.57	Uncharacterised LOC100304285	pq-loop repeat family protein	
GRMZM5G870170	-1.57	MATE1	mate efflux family protein	GO: 0050896 GO: 0042221
GRMZM2G366659	-1.49	Putative trehalose phosphatase/synthase family protein	trehalose 6-phosphate synthase	GO: 0005975
GRMZM2G047474	-1.42	TLD-domain containing nucleolar protein	Protein	

Gene stable ID	Average log2 fold-change	Ensembl/ NCBI gene description	Blast2GO description	Enriched GO-term
GRMZM2G177077	-1.40	Glucose-6-phosphate 1-dehydrogenase	glucose-6-phosphate 1-dehydrogenase	GO: 0005975
AC214438.3_FG002.1	-FC	N/A	N/A	
GRMZM2G146004	-FC	Uncharacterised protein	N/A	GO: 0050896

*-FC: Gene is only expressed (down- regulated) in one experimental group and not the other

As part of further analysis, the DEGs in this study were also compared with the DEGs found in the study by Lanubile *et al*. (2014), where RNA-seq analysis was performed on two maize genotypes (one susceptible and one resistant line) to compare the defense responses to the pathogen using the pin-bar method of infection. This study found that both genotypes resulted in similar responses after inoculation with *F. verticillioides* with a greater induction of genes in the resistant genotype. Differential gene expression was performed using a DEseq package with the FDR threshold set at $q < 0.05$ and a log2 fold-change ≥ 1. The DEGs found for both the resistant and susceptible genotypes as well as common differentially expressed genes (genes found in both the resistant and susceptible genotype) were compared to the DEGs from the current study. What is of interest is that seven susceptible (three up-regulated and four down-regulated) and six resistant (four up-regulated and two down-regulated) line genes were found in our Protocol 1/ Protocol 2 matches list of DEGs. Analysis also showed that within the common DEG list of the Lanubile *et al*. (2014) study, there were ten up-regulated and six down-regulated genes that matched the list of genes in this study.

In total, this showed 17 up-regulated and 12 down-regulated genes in the matching gene lists of the current study that were differentially expressed in the Lanubile *et al*. (2014) study, however, only the up-regulated genes are reported and can be seen in Table 3.7. Interestingly, even though 12 down-regulated genes in the current study were found in the Lanubile *et al*. (2014) study, all of these genes except one (GRMZM2G181081/ CIPK-like protein 1) were found to be up-regulated in Lanubile *et al*. (2014) and not down-regulated. There are various reasons as to why this was observed, these include: different strains of *Fusarium*, different inoculation procedures (infections), different time points of sampling (growth stages) and the fact that different sections of the maize plant were used for RNA-seq (leaves vs. kernels). The contrast in expression seen here could also suggest that genes are regulated in a different way at various locations within the plant after *F. verticillioides* infection; however, this will need to be confirmed.

Table 3.7: Up-regulated genes from the matching DEG list in this study found to be specific to the susceptible/ resistant genotype or common to both genotypes from the Lanubile *et al*. (2014) study.

Gene stable ID	Ensembl/ NCBI gene description	Susceptible/Resistant/ Common[1]	Functional category[2]
GRMZM2G125669	Alternative oxidase	Common	Response to stress
GRMZM2G093826	Potassium high-affinity transporter	Common	Transport
GRMZM5G874955	Uncharacterised protein	Common	Transport
GRMZM2G029219	Carbohydrate transporter/sugar porter/transporter	Common	Transport
GRMZM2G036464	Glutamine synthetase root isozyme 4	Common	Metabolic process
GRMZM2G008247	Beta-glucosidase 2	Common	Cell wall
GRMZM2G034152	Polyamine oxidase	Common	Metabolic process
GRMZM2G427815	Uncharacterised protein	Common	Resistance
GRMZM5G892675	Uncharacterised protein	Common	Transport
GRMZM2G087875	Putative cytochrome P450 superfamily protein	Common	Metabolic process
GRMZM2G020146	Uncharacterised protein	Resistant	Metabolic process
GRMZM2G130173	Metallothionein-like protein type 2; Uncharacterised protein	Resistant	Unknown function
GRMZM2G026470	Soluble inorganic pyrophosphatase	Resistant	Metabolic process
GRMZM2G062724	Uncharacterised protein	Resistant	Signal transduction
GRMZM2G178546	Trehalose-phosphate phosphatase	Susceptible	Metabolic process
GRMZM2G007151	Uncharacterised protein	Susceptible	Cell component
GRMZM2G119975	Uncharacterised LOC103646336	Susceptible	Metabolic process

[1]Differentially expressed genes found in the susceptible, resistant as well as commonly expressed (found in the susceptible and resistant genotype) as found in the Lanubile *et al*., (2014) study.

[2] Differentially expressed genes belonging to their specific functional categories as per Blast2GO (found in the Lanubile *et al*., (2014) study).

For the up-regulated genes seen in Table 3.7, even though these are but a few genes, they show that there could be a particular set of genes being expressed in response to *F. verticillioides* infection despite the difference in *F. verticillioides* strains, types of infections, sampling times etc. As seen in Table 3.7, a number of genes are related to the metabolic process (41%) and transport (24%) functional categories with other categories including resistance, signal transduction, cell wall and response to stress. The metabolic process category was the largest portion represented in this list and was also most prevalent in the Lanubile *et al*. (2014) study. Changes in gene expression of these genes belonging to these specific categories indicate that the infection by *F. verticillioides* has caused a response in the plant especially a change in metabolism.

The maize line, CML144, used in the current study is classified as susceptible to ear rot (Okello *et al*., 2006). Comparing the genes found in our study to the Lanubile *et al*. (2014) study, however, showed that there were a number of genes belonging to the resistant and susceptible line as well as a number of common DEGs found mostly within the up-regulated DEGs in the current study. This does not disapprove the fact that CML144 is a susceptible line but it does indicate the importance of further research and investigation to gain more knowledge on the CML144 maize line.

3.3.3 Gene ontology analysis using agriGO

To gain a better insight into the functionality of the DEGs found by RNA-seq analysis, the matching up- and down-regulated DEGs from Protocol 1 and Protocol 2 were annotated using agriGO.

Up-regulated and down-regulated genes were annotated separately using the agriGO database and mapped to the *Zea mays* AGPv3.30 reference genome. The 54 matching genes (Table 3.5) shown to be up-regulated by RNA-seq analysis were inserted into the agriGO database and only 45 of these were annotated. Analysis revealed only three significant GO-terms ($p<0.05$, multi-test adjustment [Yekutieli- FDR under dependency]) as seen in Table 3.8. These GO-terms belonged to the molecular function (F) GO category only and not the cellular component or biological process categories. The table also indicates a description of the GO-terms, the number of genes found in the list of up-regulated genes that can be found in the specific GO-term as well as the associated *p*-value and FDR. GO-terms relating to cation, ion and metal ion binding were shown to be significantly enriched in the up-regulated genes. The genes (16/45) found in this category corresponded to all the GO-terms present in Table 3.8 and are also associated with terpene synthase activity, the defense response, response to oxidative stress, amongst many others although these GO-terms were not enriched (Table 3.5).

The up-regulated ent-Kaurene synthase B gene (GRMZM2G049538), one of the genes associated with the three aforementioned GO-terms, in particular, is important to note because ent-Kaurene is one of the intermediates in the synthesis of gibberellin and kauralexins in both fungi and plants (Kawaide *et al*., 2000; Carbon *et al*., 2008; Schmelz *et al*., 2011; Fu *et al* ., 2016). The synthesis of bioactive gibberellin involves the action of cytochrome P450 monoxygenases (also shown to be involved in kauralexin synthesis in rice). This indicates the link between the ent-Kaurene synthase gene (GRMZM2G049538) and the cytochrome p450 family expressed gene (GRMZM2G087875), another up-regulated gene associated with all the GO-terms in Table 3.8 (Hedden and Thomas, 2012; Schmelz *et al*., 2014; Fu *et al*., 2016). All the genes up-regulated within these GO-terms have in some way or another elicited a response when plants were infected with *F. verticillioides*.

Table 3.8: The three significant GO terms from the matching up-regulated resulting from the agriGO database search using *Zea mays* AGPv3.30 as the reference genome.

GO term	Ontology	Description	Number in input list	Number in BG/Ref	*p*-value[1]	FDR
GO:0043169	F	cation binding	16 /45	4135/25864	0.0011	0.028
GO:0043167	F	ion binding	16 /45	4136/25864	0.0011	0.028
GO:0046872	F	metal ion binding	16 /45	4124/25864	0.0011	0.028

[1] For statistical analysis, Fisher's Exact Test was performed with multi-test adjustment and a cut-off of 0.05 after performing a SEA.

Upon analysing the 29 down-regulated genes found in the matches between Protocol 1 and Protocol 2 (Table 3.6), only 22 of these genes were annotated in the agriGO database. Analysis revealed three significant GO-terms (p<0.05, multi-test adjustment [Yekutieli- FDR under dependency]) as seen in Table 3.9 with these GO-terms belonging to only the biological process (P) category. GO-terms relating to response to stimulus, response to chemical stimulus and carbohydrate metabolic processes were shown to be significantly enriched in the down-regulated genes. The genes belonging to these GO-terms are also involved in metabolic processes, response to abscisic acid stimulus, sucrose & starch catabolic processes, amongst other responses; although some of these GO-terms were not enriched (these genes can be seen in Table 3.6).

In the response to stimulus and response to chemical stimulus GO-term, a number of genes (GRMZM2G004161, GRMZM2070172, GRMZ2G125775, GRMZM2G181081) are also involved in response to salicyclic, jasmonic and especially abscisic acid stimulus. Salicyclic and jasmonic acid are involved in the biotic stress response whereas abscisic acid is involved in the abiotic stress response (a negative regulator of disease resistance). The above mentioned hormones although involved in different stress responses are known to be connected at various levels (Mauch-Mani and Mauch, 2005). In the carbohydrate metabolic process GO-term, genes were also associated with sucrose, glucose and trehalose processes. In a RNA-seq study involving arbuscular mycorrhiza and the plants *Lotus japonicas* & *Rhizophagus irregularis*, there were a number of down-regulated genes that included GO-terms related to carbohydrate metabolism. Here, fungal colonisation by the arbuscular mycorrhiza led to a decrease in starch, which indicated that the starch was broken down to provide carbohydrates required by the fungus (Handa *et al*., 2015). A similar scenario could have occurred in the maize plants in the current study when colonised by *F. verticillioides*. Another RNA-seq study investigating the mechanisms of resistance to *F. verticillioides* by maize found upon enrichment analysis that sucrose and starch metabolism was one of the most significantly enriched pathways after infection. This study did not, however, do any further analysis on this pathway (Wang *et al*., 2016). This information suggests that the fungus drives the plant to make these sugars in order to

benefit itself by using these sugars as a food source. Campos-Bermudez *et al*. (2013) also found changes occurring within the pathways related to carbon metabolism in a susceptible maize line but not a resistant maize line and state that carbohydrate involvement is an important factor that determines the outcome of interactions between the plant and pathogen. This study also adds that the pathogen attempts to manipulate carbohydrate metabolism in the plant for its own gain and that these sugars could be metabolised and converted to fungal metabolites like glycogen and mannitol (Campos-Bermudez *et al*., 2013).

When pathogens attack maize plants, carbohydrate metabolism is one of the main metabolic processes that are affected; however, studies need to be conducted in order to confirm the changes seen in carbohydrate metabolism since every pathogen creates a different response in the host. This is to determine how the changes in metabolism are linked to disease development, resistance to the fungus and the establishment of the pathogen (Pechanova and Pechan, 2015).

Table 3.9: The three significant GO terms from the matching down-regulated genes resulting from the agriGO database search using *Zea mays* AGPv3.30 as the reference genome.

GO term	Ontology	Description	Number in input list	Number in BG/Ref	*p*-value[1]	FDR
GO: 0050896	P	response to stimulus	10/22	3551/25864	0.00032	0.0059
GO: 0042221	P	response to chemical stimulus	8/22	2052/25864	0.00018	0.0059
GO: 0005975	P	carbohydrate metabolic process	6/22	1246/25864	0.00048	0.0060

[1] For statistical analysis, Fisher's Exact Test was performed with multi-test adjustment and a cut-off of 0.05 after performing a SEA.

Given these up- and down-regulated significant GO-terms, within this list of genes associated with these GO-terms, we have also found genes associated with antioxidant activity as described in Chapter 2. Within the matching up-regulated significant GO-terms (Table 3.5), the GRMZM2G427815 gene is associated with peroxidase and oxidoreductase activity which is involved in the ROS scavenger system. Peroxidase is one of the enzymes involved in scavenging ROS, the other enzymes include SOD, GR and CAT (assayed in Chapter 2), these enzymes aid in removing ROS when the host is under pathogen attack or other stresses (Kumar *et al*., 2009). This gene (GRMZM2G427815-Peroxidase, Table 3.5) is shown to be associated with all the enriched GO-terms (GO: 0043167, GO: 0046872 and GO: 0043169) in Table 3.9 and is also a gene that was found within the common gene list (Table 3.7; found to be expressed in both the resistant and susceptible genotype) from the Lanubile *et al*. (2014) study. This is an indicator that it might have an important role in the defense response in all maize tissues and genotypes. Within the down-regulated significant GO-terms, the

GRMZM2G103812 and GRMZM2G004161 genes are associated with two of the three enriched GO-terms (GO: 0050896 and GO: 0042221) and one other specific GO-term associated but not enriched within these genes involves the response to hydrogen peroxide.

Hydrogen peroxide, a ROS molecule produced after pathogen attack, is directly involved in the defense response and prompts downstream defense responses. Enzymes such as catalase, glutathione reductase (see chapter 2) and ascorbate peroxidase are involved in removing hydrogen peroxide as it has the ability to damage cells causing oxidative stress (Bailly *et al*., 1996; Magbanua *et al*., 2007). Thus the up- or down-regulation of these genes in particular as well as other defense related genes (GRMZM2G049538, GRMZM2G117971, GRMZM2G070659, GRMZM2G358153, GRMZM2G173085 etc.) suggest that there was indeed a response to *F. verticillioides* infection after the two weeks of growth.

3.4 Conclusion for this chapter

In this chapter, whole-transcriptome analysis was conducted using RNA-seq to understand and determine the effect of *F. verticillioides* on maize plants. RNA-seq was performed on six RNA samples; three control and three infected after two weeks of infection. Samples underwent a number of quality control measures in order for it to be analysed by RNA-seq. RNA-seq data analysis was conducted in DNA subway using the maize B73 v2 genome and then compared to the analysis conducted by the CPGR with the maize B73 v3 genome as the reference. Both of these analysis protocols incorporated the Tuxedo suite of analysis. The comparison between these protocols revealed differences between the two; however, this could be associated with the two different genome versions used.

The maize genome proves to be difficult to analyse as it is rich in duplicated genes, there could also be significant differences between the v2 and v3 genome, with the v4 version of the genome recently released on Plant Ensembl. This is an indication that the maize genome is not fully understood as yet and might also explain the differences in differentially expressed genes found using Protocol 1 and 2 (Davidson *et al*., 2011). It might also explain the differences seen between this study and other studies.

The DEGs found by Lanubile *et al*. (2014) that was compared to DEGs found in the current study provides a platform for further investigation with quite a number of genes found to be overlapping. A few of the up-regulated genes seemed to be associated with either the susceptible or resistant maize line, however, a larger number of genes that overlapped were genes in both the susceptible

and resistant maize lines (common genes) from the Lanubile *et al*. (2014) study. This might indicate that there are certain genes (mostly up-regulated) activated in response to pathogen attack across various maize lines, which provides a basis for future studies on these genes in order to understand the defense response when plants are under attack.

Future experiments relevant to this study would include further studies on the overlapping genes found in the research conducted by Lanubile *et al*. (2014) and the current study, performing RNA-seq analysis using the v4 genome to identify alternative splicing and possible newly characterised genes. It will also be of great interest to look at RNA-seq data at an earlier or later point of infection and expression in the roots, kernels as well as the cobs.

Upon conducting the GO analysis using agriGO on the matching genes, three up- and three- down-regulated GO-terms were found to be significantly enriched and belonged to the molecular function and biological process categories, respectively. By identifying the DEGs in these categories (with most genes belonging to more than one GO-term), it gives an indication of which genes may play an active role in the defense response to *F. verticillioides* infection. This, in turn can aid in further understanding the interaction between *F. verticillioides* and maize plants and provides a good platform for identifying better strategies to overcome the negative effects of *F. verticillioides* infection.

In summary, upon analysing RNA-seq data, DEGs were identified using both Protocol 1 and Protocol 2. Between these protocols, 54 up- and 29 down-regulated genes were found to be matching between them. The analysis and identification of these DEGs as well as the GO-analysis and comparison to the Lanubile *et al*. (2014) study provides a platform to further explore these genes in the quest for a better understanding of the plant defense response. It also serves as a good basis for future studies on *F. verticillioides* infection in maize plants.

Chapter 4

Validation of RNA-sequencing - quantitative Real-Time PCR

4.1 Introduction

Quantitative Real-Time PCR (RT-qPCR) provides a means of measuring gene expression and is highly sensitive, specific, relatively fast, provides reproducible results and is also used as a tool for the detection of pathogens (Nygard *et al*., 2007; Bustin, 2010; Chen *et al*., 2012b).

Gene expression data from microarray and RNA-seq experiments are usually validated by performing RT-qPCR analysis on various DEGs that show significant regulation in the specific study (Lanubile *et al*., 2012a, 2012b, 2014; Opitz *et al*., 2014; Handa *et al*., 2015). The changes in gene expression obtained from RT-qPCR allows for comparison to the results obtained from RNA-seq data and also confirms whether the plant activates a defense response to the fungus by changes in gene expression.

In this chapter, the aim was to validate the RNA-seq and GO analysis results using RT-qPCR with DEGs that were associated with the 'response to stimulus' GO-term chosen for analysis. Most of the genes belonging to this GO-term are associated with the defense response, with a 'response to fungus' being of specific interest. RT-qPCR was used as a pre-validation step on the *shrunken-1* gene that was observed to be differentially expressed in a previous RNA-seq study (Shu, PhD thesis, 2014) after infection with *F. verticillioides*. RT-qPCR was also used for validation of RNA-seq and GO-analysis data of the current study. From the list of differentially expressed genes, both up- and down-regulated genes were selected for quantification via RT-qPCR. Analysis was performed on mRNA extracted from biological repeats of control and infected maize leaves two weeks after infection with *F. verticillioides* and normalised against three reference genes namely *UBCE*, *Rpol* and *MEP*.

4.2 Materials and Methods

4.2.1 Plant material and RNA extraction:

In order to prepare for cDNA synthesis and RT-qPCR, control and infected plants were grown and extracted as previously described (see Chapter 2 and 3, respectively).

4.2.2 cDNA synthesis:

In order to prepare for RT-qPCR, a total of 1µg of purified RNA (extracted as described in Chapter 3) was used for cDNA synthesis. cDNA synthesis was carried out in duplicate on each RNA sample extracted from the control and infected groups. The RNA was combined with a 1:10 ratio of Oligo d(T) to random hexamers (Promega, USA) and nuclease free water. The reaction was then placed at 70˚C for 5 min and immediately placed on ice for 5 min. A mixture of 5µl reaction buffer, 2µl DNTPs, 1µl RNase inhibitor and 1µl MMLV reverse transcriptase (Promega, USA) was added to the initial reaction to make up a 25µl final volume. The tubes were then placed in the PCR machine (GeneAmp PCR system 9700, Applied Biosystems) under the following cycling conditions: 25˚C for 5 min, 40˚C for 10 min, 55˚C for 50 min, 70˚C for 15 min and then 4˚C for 1 min.

To determine successful synthesis of cDNA, samples were run on a 1.2% (w/v) agarose gel initially at 120V for 5 min and then at 100V for 20 min and visualised as previously described (see Chapter 3).

The synthesised duplicate cDNA samples were combined and used for gene expression analysis.

4.2.3 Quantitative Real-Time PCR:

Quantitative Real-Time PCR primers used for pre-validation and post-validation with reference genes

Primers for RT-qPCR were chosen from DEGs found in Protocol 1 (DNA Subway analysis) and Protocol 2 (analysis by the CPGR). Genes were chosen based on their involvement in the defense response and were associated with the response to stimulus GO-term (GO: 0050896). *UBCE*, *Rpol* and *MEP* served as reference genes (RG) and *Shrunken-1* (*Sh1)* (Shu, PhD thesis, 2014) served as the gene of interest (GOI) as part of a pre-validation before RNA-seq. For post-validation, five genes that were either up- or down-regulated after *F. verticillioides* infection as determined by RNA-seq and GO analysis were chosen, these included *Chitinase 1*, Hypersensitive induced reaction 3 (*HIR3)*, Lipoxygenase 6 (*lox6)*, *Pathogenesis related-thaumatin-like protein* (hereafter referred to as PR-th) and *Protein induced upon tuberization* (hereafter referred to as PIT). All GOI primers were designed using the Primer Premier software version 5 (Premier Biosoft) and verified using DNAman, NCBI primer blast and Genomaize (*in-silico* PCR analysis). Primers were analysed by end-point PCR followed by gel analysis to determine if the specific target DNA had been amplified. To further confirm that the primers were amplifying the correct gene, some of the genes used for post-validation were analysed and confirmed using end-point PCR with B73 DNA as a positive control. All the primers used for the analysis of gene expression are summarised in Table 4.1.

Generation of standard curves and expression analysis

Three-fold or five-fold serial dilutions from the pooled cDNA of control and infected samples were used to generate standard curves with a minimum of four points for each gene being analysed. For the *UBCE*, *Rpol*, *Sh1*, PR-th and *lox6* standard curves, a five-fold dilution series was prepared, whereas for the *MEP*, PIT, *Chitinase 1* and *HIR3* standard curve a three-fold dilution series was prepared from the pooled cDNA.

Gene expression by RT-qPCR was performed on six RNA samples from control (3) and infected (3) maize leaves. The RT-qPCR reaction consisted of KAPA SYBR fast master mix (KAPA Biosystems, South Africa), 10µM gene specific primers, 1µl cDNA and Millipore water up to 25µl was prepared. Each dilution point and sample reaction was performed in triplicate, with a no template control (NTC) and no reverse transcriptase control (pooled RNA samples) also included in each run. Control and infected cDNA samples were diluted ten-fold for the RT-qPCR reaction. The amplification was performed using the Rotor-Gene 6000 Series software (Corbett Life Science Research, Australia) with the following parameters for all genes: 95˚C for 3 min; 40 cycles of 95˚C for 10 sec, T_A (see Table 4.1) for 20 sec, and 72˚C for 5 sec.

Reactions were set-up and RT-qPCR was performed as individual runs with the respective standard curves and average expression levels calculated from three technical repeats. The standard curve allowed the determination of the cycle threshold (Ct) values, slope, R^2 values and concentrations of all the RG's and GOI's. The average of the Ct values was used for relative quantification. Relative quantification was determined using the GenEx software (MultiD, Sweden) and Excel. The efficiency (E) of the reaction was calculated according to the following equation: $E=10^{[-1/slope]}$.

Table 4.1: List of primers used for validation of RNA-seq data by RT-qPCR on control and infected maize

Gene name	Gene product	Primer sequences (Forward **(F)**/Reverse **(R)**) --> 5'-3'	Amplicon size (bp)	Annealing Temperature (T_A) ^{0}C	Reference
Reference genes					
UBCE	*Ubiquitin-conjugating enzyme*	**F:**TGCGTTAATCACGAGACAGG **R:**AATCACAAAGACAGGCAGGG	267	60	Manoli *et al*., 2012
Rpol	*DNA directed RNA-polymerase*	**F**:AGCCAAAACGCTAAAGTGGA **R**:TAAGTGACGAGCAAGGCAAA	175	58	Ma *et al*., 2006
MEP	*Membrane protein PB1A10.07c*	**F:**TGTACTCGGCAATGCTCTTG **R:**TTTGATGCTCCAGGCTTACC	203	60	Manoli *et al*., 2012
Gene of interest (pre-validation)					
Sh1	*Shrunken-1* (*sucrose-UDP glucosyltransferase 1*) **(GRMZM2G089713)**	**F:**GGAGTAGCCTGCGTTCTACG **R:**GTCAATGTGCAGGCCAGATA	136	60	Shu, PhD thesis, 2014
Gene of interest (post-validation)					
PR-th	*Pathogenesis related-thaumatin-like protein* **(GRMZM2G039639)**	**F:**GGGGTAATTCGGAGCAGC **R:**ACGAGCGGGAAGAGGTG	72	60	Self-designed
Chitinase 1	*Chitinase 1* **(GRMZM2G358153)**	**F:**GGGCTGTTCATCTGGTCG **R:**GATCTGCTGCGCCTCGGT	78	60	Self-designed
HIR3	*Hypersensitive induced reaction 3* **(GRMZM2G070659)**	**F:**GGGAGGCAGAAGCCAAGT **R:**GACGGAGAACCCCAGCAC	98	58	Self-designed
PIT	*Protein induced upon tuberization* **(GRMZM2G472248)**	**F:**GCGAACGGCGTGTTGAGC **R:**CGCACCGAGAAGACAGAA	100	58	Self-designed
lox6	*Lipoxygenase 6* **(GRMZM2G040095)**	**F:**GCCCGCCGGAAAACTGCA **R:**CTCGTAGGCGATGCTCCC	123	60	Self-designed

4.2.4 Statistical analysis:

Statistical analysis for RT-qPCR was conducted using GraphPad Prism software version 6 (GraphPad Software Incorporation, 1992-2007). Unpaired, parametric t-tests were performed on the log-transformed values obtained from the control and infected samples two weeks after infection on all DEGs used for analysis.

4.3 Results and Discussion

4.3.1 RNA extraction

As described in Chapter 3, good quality RNA was not only required for RNA-seq but for RT-qPCR as well. RNA concentrations for both control and infected RNA were >200ng/µl with good 260/230 and 260/280 ratios (>1.8) and minimal to no degradation observed on the 1.2% EtBr stained agarose gel (Figure 4.1).

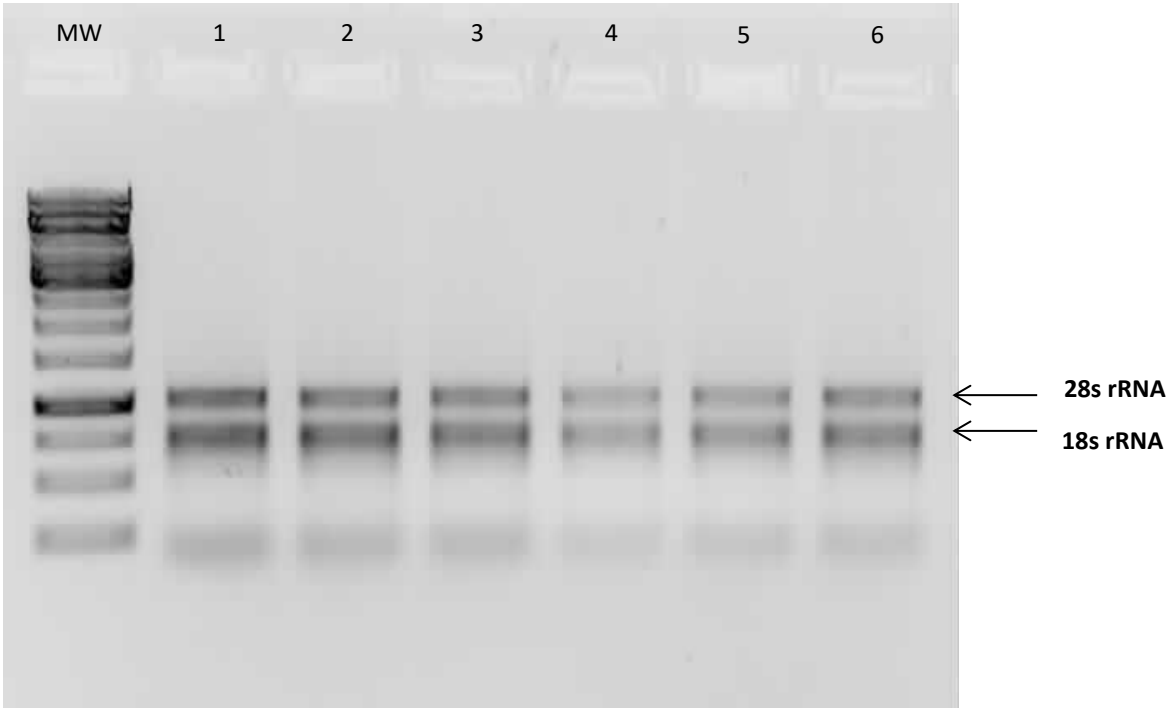

Figure 4.1: Purified RNA from control and infected maize samples after two weeks of *F. verticillioides* infection. Gel electrophoresis was performed on a 1.2% EtBr stained agarose gel for 30 min at 100V. Lane MW: 1Kb DNA ladder, Lane 1-3: Control RNA samples and Lane 4-6: Infected RNA samples.

4.3.2 cDNA synthesis

Successful cDNA synthesis was observed by a smear seen in each sample after electrophoresis on the 1.2% EtBr stained agarose gel (data not shown).

4.3.3 Gene expression analysis before RNA-seq by quantitative real-time PCR

Gene expression studies by RT-qPCR were conducted on control and infected maize plants prior to the RNA-seq experiment and served as a pre-validation step. The gene used for pre-validation was *Shrunken-1* (*Sh1*) which was analysed in the study conducted by Shu (PhD thesis, 2014). This study looked at the gene expression of *Sh1* in maize seeds during colonization by *F. verticillioides*.

Sh1 is a gene involved in sucrose synthesis and might be involved in the defense response, although it's role is not yet clear (Guglielminetti *et al*., 1996; Shu *et al*., 2015). This suggests that sugars might play a role in the defense response or that *F. verticillioides* manipulates plant metabolism to thrive within the host (Shu *et al*., 2015). In the RT-qPCR analysis by Shu (PhD thesis, 2014), the *Sh1* gene was down-regulated in infected seeds in response to *F. verticillioides* compared to the control, which corresponded to their results from their RNA-seq study (Shu, PhD thesis, 2014). They did; however, also find that *Sh1* was expressed in the embryo of uninfected kernels. In the current study, it was decided to measure *Sh1* expression in the leaves of control and infected plants after two weeks of infection. It was of particular interest to determine whether the same response would be detected in the leaves in this study as compared to the seeds in the Shu (PhD thesis, 2014) study.

The RT-qPCR results of *Sh1* in maize leaves can be seen in Figure 4.2 with R^2, slope and efficiency seen in Table 4.2. In response to infection, the expression levels of *Sh1* is shown to be down-regulated in the leaves of *F. verticillioides* infected plants relative to the control. This result was not significant but showed a similar response to that found by Shu (PhD thesis, 2014) even though the study was performed on kernels.

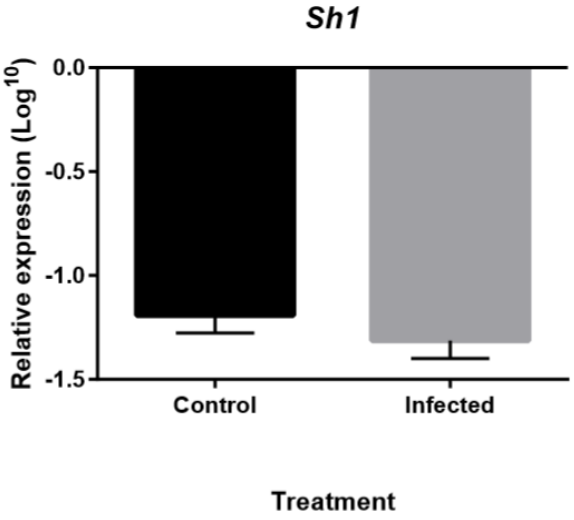

Figure 4.2: Quantitative real-time PCR analysis of *Sh1* average gene expression (n=3) in the leaves of control and infected maize plants after two weeks of infection with *F. verticillioides*. Expression was normalised to the *UBCE*, *Rpol* and *MEP* reference genes and shown as relative values (Log^{10}) with the graph generated in GraphPad prism using the t-test for statistical analysis. Error bars indicate standard error of the mean (SEM).

Analysis of RNA-seq data showed that this *Sh1* gene was not differentially expressed upon *F. verticillioides* infection (p>0.01). However, there were a number of other carbohydrate related-genes that were differentially expressed after infection with *F. verticillioides* (see Chapter 3). This gene was

also not found in the RNA-seq study that was conducted by Lanubile *et al*. (2014) (see Chapter 3). The method of fungal infection, the time points at which RNA was isolated and the maize genotype used in the respective studies could also contribute significantly to differences obtained in individual gene expression.

Table 4.2: The R^2 values, slope and efficiencies of the reference genes and the genes of interest used in this study

Gene	R^2	Efficiency	Slope
UBCE	0.99703	0.90	-3.584
Rpol	0.99660	0.85	-3.758
MEP	0.99100	0.91	-3.558
Sh1	0.99936	0.96	-3.428
PR-th	0.99735	1.06	-3.182
lox6	0.99278	1.04	-3.233
Chitinase 1	0.99615	0.95	-3.457
HIR3	0.99666	0.91	-3.554
PIT	0.99945	0.97	-3.385

4.3.4 Post RNA-seq validation by quantitative Real-Time PCR

Gene expression analysis by means of RT-qPCR was also performed to validate the RNA-seq and GO-analysis results. This was conducted on DEGs belonging to the response to stimulus category (GO term- GO: 0050896) and were selected from the Protocol 1 and Protocol 2 separate gene list (Appendix A, Table A1-A4) as well as the GO-analysis results. For expression analysis we focused on the genes involved in defense so as to validate their role upon fungal infection. The expression of five genes: *Hypersensitive induced response protein* (*HIR3*), *Lipoxygenase 6* (*lox6*), *Chitinase 1* (*Chitinase 1*), *Pathogenesis-related thaumatin-like protein* (PR-th), and *Protein induced upon tuberization* (PIT), were selected for RT-qPCR and were chosen based on their role in defense within the plant (see Table 4.1). The expression values of these GOI's after RNA-seq analysis for both Protocol 1 and 2 can be seen in Table 4.3. In both protocols genes were expressed in the same direction, however, the *lox6* and PR-th genes were only found in the analysis using Protocol 2 and not Protocol 1.

Table 4.3: GOI's for RT-qPCR with the RNA-seq expression values from Protocol 1 and 2

Gene of interest	Expression from Protocol 1 analysis	Expression from Protocol 2 analysis
HIR3	1.65	1.89
lox6	Not found	3.21
Chitinase 1	-2.02	-2.15
PR-th	Not found	-1.83
PIT	-3.23	-3.11

Upon conducting the RT-qPCR analysis, the quantitative data and melt curves for all the genes were examined. Standard curves were also generated which provided information on the efficiency of the reaction and the R^2 values. The R^2 values and efficiencies for the RGs and GOIs can be seen in Table 4.2 with results showing R^2 values >0.99000 and efficiencies within the range of 0.85-1.06.

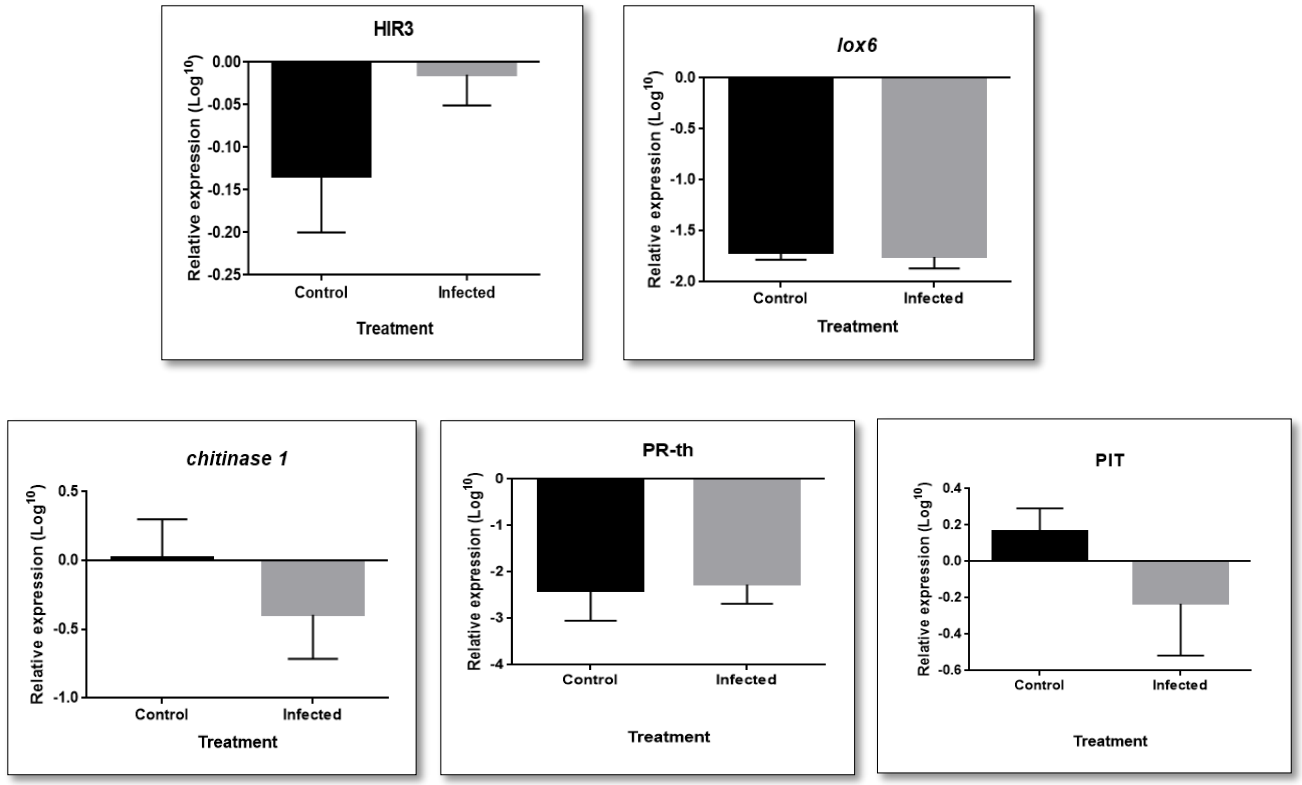

Figure 4.3: Quantitative real-time PCR analysis of *HIR3*, *lox6*, *chitinase 1*, PR-th and PIT average gene expression (n=3) in the leaves of control and infected maize plants after two weeks of infection with *F. verticillioides*. Expression was normalised to the *UBCE*, *Rpol* and *MEP* reference genes and shown as relative values (Log^{10}) with the graph generated in GraphPad prism using the t-test for statistical analysis. Error bars indicate standard error of the mean (SEM).

The RT-qPCR expression results were very similar when compared to the RNA-seq analysis results (Figure 4.3). Observing the results seen in Figure 4.3, the *HIR3* gene is down-regulated in the control group and slightly down-regulated in the infected group which suggests that there could be more *HIR3* transcripts in the infected group when compared to the control. The *lox6* gene is down-regulated in both the control and infected groups with minimal difference in mRNA expression between the two groups. The *chitinase 1* and PIT genes show a definite down-regulation in the infected group compared to the control group, whilst the PR-th gene is down-regulated in both groups with a minimal difference in expression between the two groups.

In the RNA-seq analysis, the *lox6* and *HIR3* were shown to be up-regulated and the PR-th, PIT and *Chitinase 1* genes were down-regulated, with fold-changes shown in Table 4.3 for both Protocol 1 and 2. The RT-qPCR analysis for *HIR3* and *lox6* was therefore different when compared to the RNA-seq results as *HIR3* had less down-regulation in the infected leaves while there was no difference for *lox6* when comparing control with infected samples. The PIT and *Chitinase 1* genes were down-regulated during RT-qPCR; however, the PR-th showed no difference between the two groups after infection. The reasons for the difference in expression results obtained from RNA-seq and RT-qPCR analysis respectively can be due to reasons such as biological variability within the individual replicate maize plants as different RNA extraction experiments were conducted for RNA-seq and RT-qPCR experiments, gene homologues and differences in sensitivity between the two analysis platforms.

It is also of interest to note that in the microarray study conducted by Lanubile *et al*. (2012a and 2012b), other homologues of *chitinase*, *PR* as well as *HIR* genes were identified that, which are pathogenesis-related genes that were all activated in response to *F. verticillioides* infection. These genes are thought to encode putative constituents of pathways involved in disease resistance. It is possible that these genes are the same as the genes found in our study, however, it is also possible that they are different paralogues of the same genes or the same genes found on different chromosomes. These genes could not be confirmed as being the same since the Lanubile *et al*. (2012a and 2012b) study observed changes in gene expression using microarray analysis.

Conclusion for this chapter

In this chapter, RT-qPCR analysis was conducted in order to validate RNA-seq and GO analysis results (see Chapter 3). DEGs were chosen from both Protocol 1 and Protocol 2 with genes belonging mostly to the 'response to stimulus' GO-category (GO: 0050896). Analysis revealed that RT-qPCR expression results were similar to RNA-seq results with the exception of two genes (*lox6* and PR-th). The

reasons for this could be explained by variability within biological replicate samples of control and infected plants, multiple homologues of the DEG or due to differences in the sensitivity and specificity of the two assays (RNA-seq vs. RT-qPCR).

Future experiments to validate RNA-seq and GO-analysis results could include analysing more DEGs belonging to other functional categories involved in response to *F. verticillioides* infection.

In summary, RT-qPCR was used to validate five DEGs that were either up/down-regulated in response to infection by *F. verticillioides* infection, however, only three genes showed the same expression pattern as the RNA-seq data. These genes belonged to the 'response to stimulus' category which was of particular interest in our study as they are known to be involved in the defense response in plants.

Chapter 5

Conclusion and Future work

In order to understand the changes in gene expression in maize plants after *F. verticillioides* infection, an RNA-seq study was conducted to determine DEGs in the African maize line CML144. Although a number of RNA-seq studies have been done on maize in response to *F. verticillioides* infection, this study was performed on an African maize line using the soak-seed method where plants were grown on MS media in glass bottles under controlled conditions.

Before RNA-seq analysis, physiological, biochemical and morphological experiments were conducted to ensure successful infection by the fungus. Analysis of the morphological characteristics of the whole maize plant showed visible signs of infection indicating successful colonisation by *F. verticillioides* during seed germination using the soak-seed method. The physiological experiments, mainly electrolyte leakage, did not show much difference between the control and infected plants after the two weeks of infection. This could have been due to the biological variation observed in each replicate which was most likely due to the plants being grown individually in the media. These results, however, could also be due differences in infectivity of the fungus among the individual seeds or to the differences in host defense responses by the individual plants to the pathogen.

The biochemical assays focused on the SOD, CAT and GR antioxidant enzymes measured after the two weeks of infection and were conducted on the leaves. Although a difference was seen between the control and infected plants after the two-week infection period, it was not statistically significant. These enzymes are, however, expressed at different stages of plant growth and at different locations. With the CAT enzyme for example, in maize there are three catalase isoforms that are encoded on different chromosomes. These isoforms also have a high sequence similarity at both the amino acid and nucleotide levels (Magbanua *et al*., 2007; Mhamdi *et al*., 2010). It has been shown that these isoforms have distinct patterns of transcript and protein accumulation (Redinbaugh *et al*., 1990). The other enzymes which include SOD and GR also have several isoforms that are located in different regions of maize leaves (Doulis *et al*., 1997). It must be noted that the enzyme assays used in this study does not distinguish between isoforms, it is thus possible that at the time the assays were conducted, the enzymes were expressed in a different region of the plant (e.g. CAT could have been expressed in the roots and not leaves after week 1 and 2 of infection).

For any future experiments, it would be beneficial to conduct the assays at other time points which would include 24 hrs after infection; it would also be of benefit to include a maize line like B73,

which has been used in a number of studies as this would serve as a reference to interpret the observed results.

During the RNA-seq analysis, a number of genes were found to be related to peroxidase (see Chapter 3); in terms of this it would be beneficial to conduct assays to detect this enzyme after infection with *F. verticillioides*. Both the physiological and biochemical experiments should be conducted on the roots since this region seems to be the most affected after infection using the soak-seed method. The different isoforms of each enzyme should be analysed with zymograms in order to verify the results as per the specific assay performed.

RNA-seq analysis was performed at two weeks after infection on control and infected plants comprising of three biological replicates. No technical replicates were required given the specificity of this next-generation sequencing technology. Whole-transcriptome RNA-seq was the main focus of this study to determine the differential gene expression occurring in maize plants after two weeks of infection with *F. verticillioides*. Analysis of RNA-seq data was conducted in DNA subway, an analysis tool with the complete RNA-seq workflow that incorporates the Tuxedo suite of protocols. To gain confidence in the analysis results, the RNA-seq analysis conducted by the CPGR was used as a validation. The comparison between these two protocols showed a number of DEGs that were different between the two; however, this is most certainly due to the difference in genome versions used between the protocols. To improve this analysis it would be advantageous to use the v4 maize genome for analysis of differentially expressed genes, this genome has only been recently released and was not available when analysis was being conducted.

As a further validation, the DEGs found in this study were also compared to the Lanubile *et al*. (2014) study. The difference between the two studies was that there were different maize lines being investigated, a different *F. verticillioides* strain, and different modes of infection. The Lanubile *et al*. (2014) study was also conducted on maize plants grown in the field whereas the current study used maize plants grown in bottles with media. Genes from this study that were found in the DEGs list of the Lanubile *et al*. (2014) study were from the resistant and the susceptible maize genotype with quite a number of genes that were commonly expressed (included 16 genes from both the resistant and susceptible genotype). This provides a basis for future studies on these genes in order to understand the defense response when plants are under pathogen attack. Future experiments relevant to this would be to look at RNA-seq data at an earlier or later point of infection as well as to look at expression in the roots, kernels and the cobs. This will allow a better comparison to determine the genes being regulated after infection with *F. verticillioides*.

Given the sets of matching up- and down-regulated DEGs (found between Protocol 1 and 2), GO analysis was performed in agriGO to determine the enriched GO-terms. Upon conducting the GO analysis, three up- and three- down- regulated GO-terms were found to be significantly enriched and belonged to the molecular function and biological process categories, respectively. By identifying the DEGs belonging to these categories, more so with most of the genes belonging to more than one GO-term, it gives an indication of which genes may play an active role in the defense response to *F. verticillioides* infection.

Performing the RT-qPCR served as a post-validation analysis of the RNA-seq and GO-analysis results. The 'response to stimulus' GO category was of specific interest because the genes found within this category are involved in the defense response. Future experiments should, however, also focus on DEGs belonging to other categories so that a link can be made with various other responses when plants are under attack by *F. verticillioides*.

Another important part of the defense response in plants is the accumulation of PR proteins (observed to be regulated in this study as well) with many of these proteins also having antifungal activity. It would be of future interest to observe whether induced expression of selected maize PR genes (or a suite of these genes) using a transgenic approach could limit the infection by *F. verticillioides*.

Other experiments relevant to this study would be to use different routes of infection to determine whether different effects would be observed and whether similar genes are up-/down-regulated. In the current study, a two week infection period was the focal time point for most of the experiments conducted, it would be of interest to analyse different time points of infection especially after this two week period. Observing the changes at various time points would give a better view of the events or changes that take place within the plants at different stages of infection especially with respect to DEGs.

It would also be of great interest to analyse novel genes (found through the cufflinks section of the RNA-seq analysis) that were expressed after infection and determine what roles these genes play in terms of the defense response. Even though mutants of specific DEGs were researched and found (in the B73 maize line), they have not been reported here. Future work could include analysing mutant knockdowns of these genes to determine their roles with respect to *F. verticillioides* infection.

The aim of this study, to examine the changes in gene expression after *F. verticillioides* infection by whole-transcriptome RNA-seq was successful as seen in the previous chapters. This study serves as a base for all subsequent studies and helps broaden our understanding of the infection of *F.*

verticillioides in maize plants and develop strategies to not only improve maize yields but also produce healthy maize resistant to *F. verticillioides*.

As far as we know, this is the first RNA-seq study conducted on the CML144 African maize line infected with *F. verticillioides* using the soak seed method for inoculation.

References

Ádám, A. L., Galal, A. A., Manninger, K.*, et al.* 2000. Inhibition of the development of leaf rust (*Puccinia recondita*) by treatment of wheat with allopurinol and production of a hypersensitive-like reaction in a compatible host. *Plant Pathology,* **49**, 317-323.

Aiyaz, M., Divakara, S. T., Nayaka, S. C.*, et al.* 2015. Application of beneficial rhizospheric microbes for the mitigation of seed-borne mycotoxigenic fungal infection and mycotoxins in maize. *Biocontrol Science and Technology,* **25**, 1105-1119.

Andrews, S. 2010. *FastQC: A quality control tool for high throughput sequence data* [Online]. Available: http://www.bioinformatics.babraham.ac.uk/projects/fastqc [Accessed November 2016].

Bailly, C., Benamar, A., Corbineau, F.*, et al.* 1996. Changes in malondialdehyde content and in superoxide dismutase, catalase and glutathione reductase activities in sunflower seeds as related to deterioration during accelerated aging. *Physiologia Plantarum,* **97**, 104-110.

Bajji, M., Kinet, J.-M. and Lutts, S. 2002. The use of the electrolyte leakage method for assessing cell membrane stability as a water stress tolerance test in durum wheat. *Plant Growth Regulation,* **36**, 61-70.

Benjamini, Y. and Hochberg, Y. 1995. Controlling the False Discovery Rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological),* **57**, 289-300.

Bolger, A. M., Lohse, M. and Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics,* **30**, 2114-2120.

Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry,* **72**, 248-254.

Bustin, S. A. 2010. Why the need for qPCR publication guidelines?—The case for MIQE. *Methods,* **50**, 217-226.

Campos-Bermudez, V. A., Fauguel, C. M., Tronconi, M. A.*, et al.* 2013. Transcriptional and metabolic changes associated to the infection by *Fusarium verticillioides* in maize inbreds with contrasting ear rot resistance. *Plos One,* **8**.

Carbon, S., Ireland, A., Mungall, C. J.*, et al.* 2008. AmiGO: online access to ontology and annotation data. *Bioinformatics,* **25**, 288-289.

Chen, J., Ding, J., Li, H.*, et al.* 2012a. Detection and verification of quantitative trait loci for resistance to *Fusarium* ear rot in maize. *Molecular Breeding,* **30**, 1649-1656.

Chen, K., Fessehaie, A. and Arora, R. 2012b. Selection of reference genes for normalizing gene expression during seed priming and germination using qPCR in *Zea mays* and *Spinacia oleracea*. *Plant Molecular Biology Reporter,* **30**, 478-487.

Claiborne, A. 1985. Catalase activity. *CRC handbook of methods for oxygen radical research,* **1**, 283-284.

Coetzer, N., Myburg, A. A. and Berger, D. K. 2011. Maize microarray annotation database. *Plant Methods,* **7**, 1-10.

Cox, M. P., Eaton, C. J. and Scott, D. B. 2010. Exploring molecular signaling in plant-fungal symbioses using high throughput RNA sequencing. *Plant Signaling & Behavior,* **5**, 1353-1358.

Davidson, R. M., Hansey, C. N., Gowda, M.*, et al.* 2011. Utility of RNA sequencing for analysis of maize reproductive transcriptomes. *The Plant Genome,* **4**, 191-203.

Debona, D., Rodrigues, F. Á., Rios, J. A.*, et al.* 2012. Biochemical changes in the leaves of wheat plants infected by *Pyricularia oryzae*. *Phytopathology,* **102**, 1121-1129.

Demidchik, V., Straltsova, D., Medvedev, S. S.*, et al.* 2014. Stress-induced electrolyte leakage: the role of K+-permeable channels and involvement in programmed cell death and metabolic adjustment. *Journal of Experimental Botany,* **65**, 1259-1270.

Doulis, A. G., Debian, N., Kingston-Smith, A. H.*, et al.* 1997. Differential localization of antioxidants in maize leaves. *Plant Physiology,* **114**, 1031-1037.

Du Plessis, J. 2003. *Maize production*, Department of Agriculture.

Du, Z., Zhou, X., Ling, Y.*, et al.* 2010. agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Research,* **38**, W64-W70.

Fandohan, P., Hell, K., Marasas, W.*, et al.* 2003. Infection of maize by *Fusarium* species and contamination with fumonisin in Africa. *African Journal of Biotechnology,* **2**, 570-579.

Ferrigo, D., Raiola, A., Bogialli, S.*, et al.* 2015. In vitro production of fumonisins by *Fusarium verticillioides* under oxidative stress induced by H2O2. *Journal of Agricultural and Food Chemistry,* **63**, 4879-4885.

Fountain, J. C., Khera, P., Yang, L.*, et al.* 2015. Resistance to Aspergillus flavus in maize and peanut: Molecular biology, breeding, environmental stress, and future perspectives. *The Crop Journal,* **3**, 229-237.

Fu, J., Ren, F., Lu, X.*, et al.* 2016. A tandem array of ent-Kaurene synthases in maize with roles in gibberellin and more specialized metabolism. *Plant Physiology,* **170**, 742-751.

Fu, X., Fu, N., Guo, S.*, et al.* 2009. Estimating accuracy of RNA-Seq and microarrays with proteomics. *BMC Genomics,* **10**.

Glazebrook, J. 2005. Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annual Review Phytopathology,* **43**, 205-227.

Guglielminetti, L., Alpi, A. and Perata, P. 1996. Shrunken-1-encoded sucrose synthase is not required for the sucrose-ethanol transition in maize under anaerobic conditions. *Plant Science,* **119**, 1-10.

Handa, Y., Nishide, H., Takeda, N.*, et al.* 2015. RNA-seq transcriptional profiling of an Arbuscular Mycorrhiza provides insights into regulated and coordinated gene expression in *Lotus japonicus* and *Rhizophagus irregularis*. *Plant and Cell Physiology,* **56**, 1490-1511.

Hedden, P. and Thomas, S. 2012. Gibberellin biosynthesis and its regulation. *Biochemical Journal,* **444**, 11-25.

Huffaker, A., Kaplan, F., Vaughan, M. M.*, et al.* 2011. Novel acidic sesquiterpenoids constitute a dominant class of pathogen-induced phytoalexins in maize. *Plant Physiology,* **156**, 2082-2097.

IARC 2002. Some traditional herbal medicines, some mycotoxins, naphthalene and styrene. *IARC monographs on the evaluation of carcinogenic risks to humans,* **82**.

Jones, J. D. and Dangl, J. L. 2006. The plant immune system. *Nature,* **444**, 323-329.

Kant, P., Gulati, A., Harris, L.*, et al.* 2012. Transgenic corn plants with modified ribosomal protein L3 show decreased ear rot disease after inoculation with *Fusarium graminearum*. *Australian Journal of Crop Science,* **6**, 1598-1605.

Kawaide, H., Sassa, T. and Kamiya, Y. 2000. Functional analysis of the two interacting cyclase domains in ent-Kaurene synthase from the fungus *Phaeosphaeria* sp. L487 and a comparison with cyclases from higher plants. *Journal of Biological Chemistry,* **275**, 2276-2280.

Keeling, B. 1982. A seedling test for resistance to soybean stem canker caused by *Diaporthe phaseohrum* var. *caulivora*. *Phytopathology,* **72**, 807-809.

Kestler, H. A., Müller, A., Kraus, J. M.*, et al.* 2008. VennMaster: Area-proportional Euler diagrams for functional GO analysis of microarrays. *BMC Bioinformatics,* **9**, 1-12.

Kim, D., Pertea, G., Trapnell, C.*, et al.* 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology,* **14**, R36-R36.

King, S. and Scott, G. 1982. Field inoculation techniques to evaluate maize for reaction to kernel infection by *Aspergillus flavus*. *Phytopathology,* **72**, 782-785.

Krebs, S., Fischaleck, M. and Blum, H. 2009. A simple and loss-free method to remove TRIzol contaminations from minute RNA samples. *Analytical Biochemistry,* **387**, 136-138.

Kumar, M., Yadav, V., Tuteja, N.*, et al.* 2009. Antioxidant enzyme activities in maize plants colonized with *Piriformospora indica*. *Microbiology,* **155**, 780-790.

Lanubile, A., Bernardi, J., Battilani, P.*, et al.* 2012a. Resistant and susceptible maize genotypes activate different transcriptional responses against *Fusarium verticillioides*. *Physiological and Molecular Plant Pathology,* **77**, 52-59.

Lanubile, A., Bernardi, J., Marocco, A.*, et al.* 2012b. Differential activation of defense genes and enzymes in maize genotypes with contrasting levels of resistance to *Fusarium verticillioides*. *Environment and Experimental Botany,* **78**, 39-46.

Lanubile, A., Ferrarini, A., Maschietto, V.*, et al.* 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**, 1-16.

Lanubile, A., Maschietto, V., De Leonardis, S.*, et al.* 2015. Defense responses to mycotoxin-producing fungi *Fusarium proliferatum*, *F. subglutinans*, and *Aspergillus flavus* in kernels of susceptible and resistant maize genotypes. *Molecular Plant-Microbe Interactions,* **28**, 546-557.

Lanubile, A., Pasini, L. and Marocco, A. 2010. Differential gene expression in kernels and silks of maize lines with contrasting levels of ear rot resistance after *Fusarium verticillioides* infection. *Journal of Plant Physiology,* **167**, 1398-1406.

Liu, X. and Huang, B. 2000. Heat stress injury in relation to membrane lipid peroxidation in creeping bentgrass. *Crop Science,* **40**, 503-510.

Logrieco, A., Mulè, G., Moretti, A.*, et al.* 2002. Toxigenic *Fusarium* species and mycotoxins associated with maize ear rot in Europe. *European Journal of Plant Pathology,* **108**, 597-609.

Lohse, M., Nagel, A., Herter, T.*, et al.* 2014. Mercator: a fast and simple web server for genome scale functional annotation of plant sequence data. *Plant, Cell & Environment,* **37**, 1250-1258.

Ma, J., Morrow, D. J., Fernandes, J.*, et al.* 2006. Comparative profiling of the sense and antisense transcriptome of maize lines. *Genome Biology,* **7**, R22.

Magbanua, Z. V., De Moraes, C. M., Brooks, T. D.*, et al.* 2007. Is catalase activity one of the factors associated with maize resistance to *Aspergillus flavus*? *Molecular Plant-Microbe Interactions,* **20**, 697-706.

Makarevitch, I., Frechette, C. and Wiatros, N. 2015. Authentic research experience and "Big Data" analysis in the classroom: maize response to abiotic stress. *CBE-Life Sciences Education,* **14**.

Manoli, A., Sturaro, A., Trevisan, S.*, et al.* 2012. Evaluation of candidate reference genes for qPCR in maize. *Journal of Plant Physiology,* **169**, 807-815.

Marasas, W. 2001. Discovery and occurrence of the fumonisins: a historical perspective. *Environmental Health Perspectives,* **109**, 239-243.

Marioni, J. C., Mason, C. E., Mane, S. M.*, et al.* 2008. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Research,* **18**, 1509-1517.

Maschietto, V., Colombi, C., Pirona, R.*, et al.* 2017. QTL mapping and candidate genes for resistance to *Fusarium* ear rot and fumonisin contamination in maize. *BMC Plant Biology,* **17**, 1-20.

Maschietto, V., Lanubile, A., Leonardis, S. D.*, et al.* 2016. Constitutive expression of pathogenesis-related proteins and antioxydant enzyme activities triggers maize resistance towards *Fusarium verticillioides*. *Journal of Plant Physiology,* **200**, 53-61.

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

Mesterházy, Á., Lemmens, M. and Reid, L. M. 2012. Breeding for resistance to ear rots caused by *Fusarium* spp. in maize—a review. *Plant Breeding,* **131**, 1-19.

Mhamdi, A., Queval, G., Chaouch, S.*, et al.* 2010. Catalase function in plants: a focus on *Arabidopsis* mutants as stress-mimic models. *Journal of Experimental Botany,* **61**, 4197-4220.

Miedaner, T., Bolduan, C. and Melchinger, A. 2010. Aggressiveness and mycotoxin production of eight isolates each of *Fusarium graminearum* and *Fusarium verticillioides* for ear rot on susceptible and resistant early maize inbred lines. *European Journal of Plant Pathology,* **127**, 113-123.

Miller, J. D. and Ewen, M. A. 1997. Toxic effects of deoxynivalenol on ribosomes and tissues of the spring wheat cultivars *Frontana* and *Casavant*. *Natural Toxins,* **5**, 234-237.

Munkvold, G. P. and Desjardins, A. E. 1997. Fumonisins in maize: can we reduce their occurrence? *Plant Disease,* **81**, 556-565.

Mwalwayo, D. S. and Thole, B. 2016. Prevalence of aflatoxin and fumonisins (B 1+ B 2) in maize consumed in rural Malawi. *Toxicology Reports,* **3**, 173-179.
Nygard, A.-B., Jørgensen, C. B., Cirera, S.*, et al.* 2007. Selection of reference genes for gene expression studies in pig tissues using SYBR green qPCR. *BMC Molecular Biology,* **8**.
Okello, D., Manna, R., Imanyowoha, J.*, et al.* 2006. Agronomic performance and breeding potential of selected inbred lines for improvement of protein quality of adapted Ugandan maize germplasm. *African Crop Science Journal,* **14**, 37-47.
Oliver, R. P. and Ipcho, S. V. 2004. *Arabidopsis* pathology breathes new life into the necrotrophs-vs.-biotrophs classification of fungal pathogens. *Molecular Plant Pathology,* **5**, 347-352.
Opitz, N., Paschold, A., Marcon, C.*, et al.* 2014. Transcriptomic complexity in young maize primary roots in response to low water potentials. *BMC Genomics,* **15**, 1.
Oren, L., Ezrati, S., Cohen, D.*, et al.* 2003. Early events in the *Fusarium verticillioides*-maize interaction characterized by using a green fluorescent protein-expressing transgenic isolate. *Applied and Environmental Microbiology,* **69**, 1695-1701.
Panda, S. K. 2012. *Assay guided comparison for enzymatic and non-enzymatic antioxidant activities with special reference to medicinal plants*, INTECH Open Access Publisher.
Pechanova, O. and Pechan, T. 2015. Maize-pathogen interactions: an ongoing combat from a proteomics perspective. *International Journal of Molecular Sciences,* **16**, 28429-28448.
Pereira, P., Nesci, A., Castillo, C.*, et al.* 2011. Field studies on the relationship between *Fusarium verticillioides* and maize (*Zea mays* L.): Effect of biocontrol agents on fungal infection and toxin content of grains at harvest. *International Journal of Agronomy*, 1-7.
Picot, A., Atanasova-PéNichon, V., Pons, S.*, et al.* 2013. Maize kernel antioxidants and their potential involvement in *Fusarium* ear rot resistance. *Journal of Agricultural and Food Chemistry,* **61**, 3389-3395.
Picot, A., Barreau, C., Pinson-Gadais, L.*, et al.* 2010. Factors of the *Fusarium verticillioides*-maize environment modulating fumonisin production. *Critical reviews in microbiology,* **36**, 221-231.
Prochazkova, D., Sairam, R., Srivastava, G.*, et al.* 2001. Oxidative stress and antioxidant activity as the basis of senescence in maize leaves. *Plant Science,* **161**, 765-771.
Qiu, J., Xu, J., Dong, F.*, et al.* 2015. Isolation and characterization of *Fusarium verticillioides* from maize in eastern China. *European Journal of Plant Pathology,* **142**, 791-800.
Redinbaugh, M. G., Sabre, M. and Scandalios, J. G. 1990. The distribution of catalase activity, isozyme protein, and transcript in the tissues of the developing maize seedling. *Plant Physiology,* **92**, 375-380.
Rheeder, J., Shephard, G., Vismer, H.*, et al.* 2009. Guidelines on mycotoxin control in South Africa foodstuffs: From the application of the hazard analysis and critical control point (HACCP) system to new national mycotoxin regulations. *Medical Research Council Policy Brief*.
Schmelz, E. A., Huffaker, A., Sims, J. W.*, et al.* 2014. Biosynthesis, elicitation and roles of monocot terpenoid phytoalexins. *The Plant Journal,* **79**, 659-678.
Schmelz, E. A., Kaplan, F., Huffaker, A.*, et al.* 2011. Identity, regulation, and activity of inducible diterpenoid phytoalexins in maize. *PNAS USA,* **108**, 5455-5460.
Schroeder, A., Mueller, O., Stocker, S.*, et al.* 2006. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Molecular Biology,* **7**.
Sekhon, R. S., Briskine, R., Hirsch, C. N.*, et al.* 2013. Maize gene atlas developed by RNA sequencing and comparative evaluation of transcriptomes based on RNA sequencing and microarrays. *Plos One,* **8**, e61005.
Seyednasrollah, F., Laiho, A. and Elo, L. L. 2015. Comparison of software packages for detecting differential expression in RNA-seq studies. *Briefings in Bioinformatics,* **16**, 59-70.
Shu, X. 2014. *Pathogenesis and host response during infection of maize kernels by Aspergillus flavus and Fusarium verticillioides.* PhD thesis, North Carolina State University.

Shu, X., Livingston, D. P., 3rd, Franks, R. G.*, et al.* 2015. Tissue-specific gene expression in maize seeds during colonization by *Aspergillus flavus* and *Fusarium verticillioides*. *Molecular Plant Pathology,* **16**, 662-674.

Small, I., Flett, B., Marasas, W.*, et al.* 2012. Use of resistance elicitors to reduce *Fusarium* ear rot and fumonisin accumulation in maize. *Crop Protection,* **41**, 10-16.

Supek, F., Bošnjak, M., Škunca, N.*, et al.* 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. *Plos One,* **6**, e21800.

Trapnell, C., Roberts, A., Goff, L.*, et al.* 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols,* **7**, 562-578.

Trapnell, C. and Salzberg, S. L. 2009. How to map billions of short reads onto genomes. *Nature Biotechnology,* **27**, 455.

Trapnell, C., Williams, B. A., Pertea, G.*, et al.* 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology,* **28**, 511-515.

Van Rensburg, B. J., Mclaren, N., Flett, B.*, et al.* 2015. Fumonisin producing *Fusarium* spp. and fumonisin contamination in commercial South African maize. *European Journal of Plant Pathology,* **141**, 491-504.

Van Verk, M. C., Hickman, R., Pieterse, C. M. J.*, et al.* 2013. RNA-Seq: revelation of the messengers. *Trends in Plant Science,* **18**, 175-179.

Venturini, G., Toffolatti, S., Assante, G.*, et al.* 2015. The influence of flavonoids in maize pericarp on *Fusarium* ear rot symptoms and fumonisin accumulation under field conditions. *Plant Pathology,* **64**, 671-679.

Wang, Y., Zhou, Z., Gao, J.*, et al.* 2016. The mechanisms of maize resistance to *Fusarium verticillioides* by comprehensive analysis of RNA-seq data. *Frontiers in Plant Science,* **7**.

Wild, C., Miller, J. and Groopman, J. 2016. *Mycotoxin Control in Low-and Middle-income Countries*, *IARC working group report*.

Wilhelm, B. T. and Landry, J.-R. 2009. RNA-Seq—quantitative measurement of expression through massively parallel RNA-sequencing. *Methods,* **48**, 249-257.

Williams, J. DNA Subway–An educational bioinformatics platform for gene and genome analysis: DNA barcoding, and RNA-Seq. 10th World Congress on Genetics Applied to Livestock Production, 2014. Asas.

Wu, F., Bhatnagar, D., Bui-Klimke, T.*, et al.* 2011. Climate change impacts on mycotoxin risks in US maize. *World Mycotoxin Journal,* **4**, 79-93.

Yang, I. S. and Kim, S. 2015. Analysis of whole transcriptome sequencing data: Workflow and software. *Genomics & Informatics,* **13**, 119-125.

Zeilinger, S., Gupta, V. K., Dahms, T. E. S.*, et al.* 2015. Friends or foes? Emerging insights from fungal interactions with plants. *FEMS Microbiology Reviews*.

Zerbo, P., Konaté, K., Ouédraogo, M.*, et al.* 2014. Antioxidant and antifungal profiles of phenol acid rich-fractions from Sida urens L. against mycelia growth inhibiton of some *Aspergillus* and *Fusarium* species. **6**, 174-178.

Zorzete, P., Castro, R. S., Pozzi, C. R.*, et al.* 2008. Relative populations and toxin production by *Aspergillus flavus* and *Fusarium verticillioides* in artificially inoculated corn at various stages of development under field conditions. *Journal of the Science of Food and Agriculture,* **88**, 48-55.

Appendices

Appendix A:

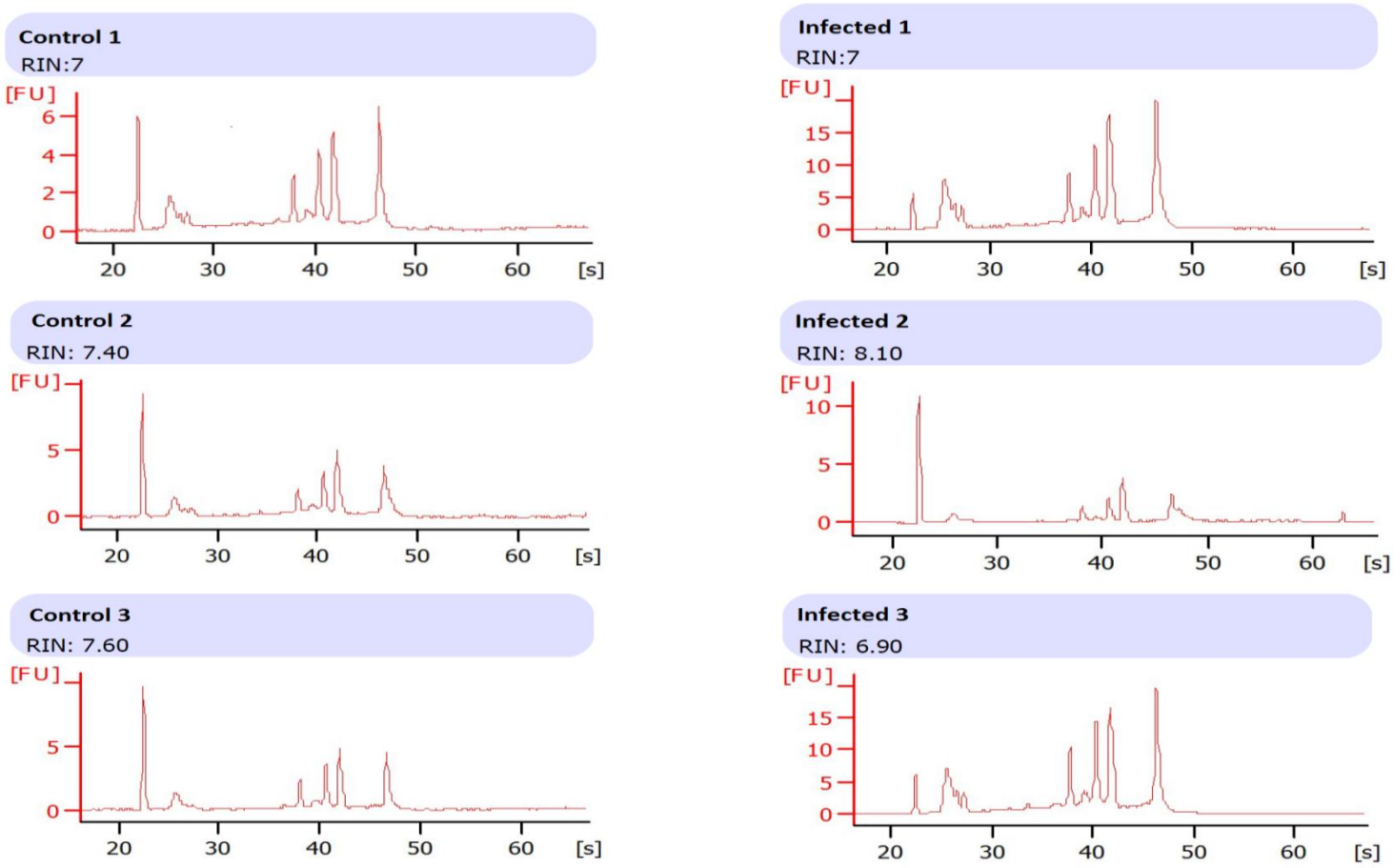

Figure A1: Electropherogram summary and RIN numbers of control and infected RNA samples using the BioAnalyzer.

Result from Protocol 2:

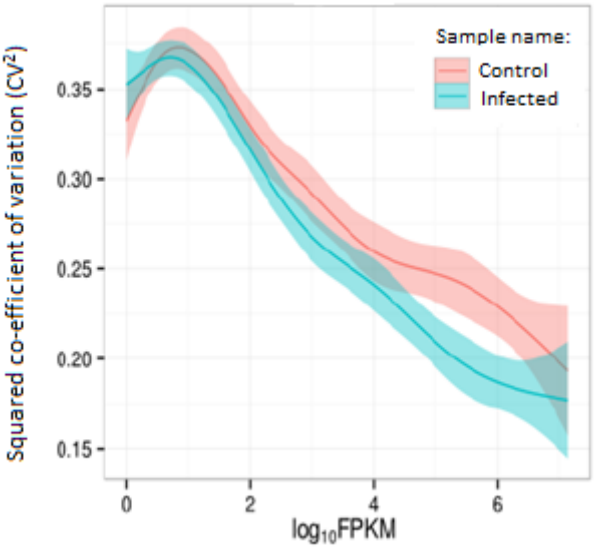

Figure A2: Co-efficient of variation across expression levels for both control and infected samples at the gene level using Protocol 2.

Table A.1: Significantly up-regulated genes from Protocol 1 after *F. verticillioides* infection as detected using RNA-seq with the tuxedo suite of analysis and mapping to the maize B73 v3 genome. Table shows annotation of genes as described in Plant Ensembl, NCBI and Maize Microarray Annotation Database (Blast2GO).

Key- **Bold**: Protocol 1 vs. Protocol 2 matches; Lanubile *et al*., 2014 genes: Susceptible genotype , Resistant genotype ; qRT-PCR genes: ; (*): Sequencing in progress; FC/-FC: Gene is only expressed (up-/down- regulated) in one experimental group and not the other

Chr	Gene Stable ID	log2 (fold-change)	*p*-value	*q*-value	Ensembl/ NCBI gene description	Blast2GO description
1	**GRMZM2G062724**	**3.19**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**chy zinc finger family expressed**
	GRMZM2G034302	**1.72**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**sucrose transporter**
	GRMZM2G130149	**2.10**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**myb family transcription expressed**
	GRMZM2G137861	**3.13**	**5.00E-05**	**0.006114**	**Wall-associated receptor kinase 2-like**	**N/A**
	GRMZM2G099767	**1.66**	**5.00E-05**	**0.006114**	**ATMAP70-2**	**N/A**
	GRMZM2G091456	**2.19**	**5.00E-05**	**0.006114**	**Putative Uncharacterised protein**	**squalene expressed, squalene monooxygenase**
	GRMZM2G366681	**2.11**	**5.00E-05**	**0.006114**	**Hypothetical protein**	**N/A**
	GRMZM2G040369	**1.57**	**5.00E-05**	**0.006114**	**Elongation factor 2**	**protein**
	GRMZM2G076537	**1.86**	**5.00E-05**	**0.006114**	**Polynucleotidyl transferase, ribonuclease H-like superfamily protein**	**exonuclease family protein**
	GRMZM2G061817	1.68	5.00E-05	0.006114	N/A	N/A
	GRMZM2G576460	2.78	5.00E-05	0.006114	N/A	cysteine proteinase
	GRMZM2G162879	2.42	5.00E-05	0.006114	N/A	N/A
2	**GRMZM2G049538**	**4.71**	**5.00E-05**	**0.006114**	**Acyclic sesquiterpene synthase**	**ent-kaurene synthase b**
	GRMZM2G125669	**1.97**	**5.00E-05**	**0.006114**	**Alternative oxidase**	**alternative oxidase**
	GRMZM2G115451	**2.63**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**neutral alkaline invertase**
	GRMZM2G144083	**2.06**	**5.00E-05**	**0.006114**	**Putative ATP dependent copper transporter**	**heavy metal p-type**

						atpase
	GRMZM2G176433	**1.84**	**5.00E-05**	**0.006114**	**Putative Uncharacterised protein**	**N/A**
	GRMZM2G119975	**3.90**	**5.00E-05**	**0.006114**	**Uncharacterised LOC103646336**	**N/A**
	GRMZM2G093826	**2.29**	**5.00E-05**	**0.006114**	**Potassium high-affinity transporter**	**high-affinity potassium transporter**
	GRMZM2G587803	1.53	5.00E-05	0.006114	N/A	N/A
	GRMZM2G394450	2.15	5.00E-05	0.006114	Beta-fructofuranosidase 1	beta-fructofuranosidase 1
3	**GRMZM2G062531**	**2.38**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**c-4 sterol methyl oxidase**
	GRMZM2G057823	**1.48**	**5.00E-05**	**0.006114**	**Fructose-bisphosphate aldolase, cytoplasmic isozyme**	**fructose-bisphosphate aldolase**
	GRMZM5G874955	**3.40**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**pdr-like abc transporter**
	GRMZM2G143139	**3.28**	**5.00E-05**	**0.006114**	**N/A**	**N/A**
	GRMZM2G029219	**3.52**	**5.00E-05**	**0.006114**	**Carbohydrate transporter/ sugar porter/ transporter**	**major facilitator superfamily antiporter/ carbohydrate transporter sugar porter transporter**
	GRMZM2G122296	2.84	5.00E-05	0.006114	Putative phosphoethanolamine N-methyltransferase	phosphoethanolamine n-methyltransferase
	GRMZM5G896496	1.62	5.00E-05	0.006114	N/A	heat shock protein 93-v
	GRMZM2G161310	1.62	5.00E-05	0.006114	Uncharacterised protein	carbohydrate transporter sugar porter transporter
	GRMZM2G050275	2.53	5.00E-05	0.006114	N/A	N/A
4	**GRMZM2G098346**	**2.96**	**5.00E-05**	**0.006114**	**Alcohol dehydrogenase 2**	**alcohol dehydrogenase**
	GRMZM2G026922	**3.28**	**5.00E-05**	**0.006114**	**Hypothetical protein**	**acetylglutamate kinase**

	GRMZM2G070011	**2.81**	**5.00E-05**	**0.006114**	**Uncharacterised protein; Vignain**	**vignain precursor**
	GRMZM2G036464	**2.64**	**5.00E-05**	**0.006114**	**Glutamine synthetase root isozyme 4**	**glutamine synthetase**
	GRMZM2G149422	**2.91**	**5.00E-05**	**0.006114**	**Hypothetical protein**	**phi-1**
	GRMZM2G116079	**1.80**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**N/A**
	GRMZM2G036217	**1.85**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**fatty acyl coa reductase**
	GRMZM2G154523	**3.10**	**5.00E-05**	**0.006114**	**Patatin T5; Uncharacterised protein**	**patatin t5 recursor***
	AC217947.4_FG002	**2.23**	**5.00E-05**	**0.006114**	**NADPH--cytochrome P450 reductase**	**N/A**
	GRMZM2G117971	**3.39**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**pathogenesis-related protein 4**
	GRMZM2G165387	2.08	5.00E-05	0.006114	Uncharacterised protein	leucine rich repeat family expressed
	GRMZM2G018424	1.60	5.00E-05	0.006114	N/A	N/A
	GRMZM2G075430	1.55	5.00E-05	0.006114	N/A	N/A
	GRMZM2G122076	1.44	5.00E-05	0.006114	Homeodomain leucine zipper protein CPHB-5; Putative homeobox DNA-binding and leucine zipper domain family protein	homeodomain leucine zipper protein cphb-5
	GRMZM5G864911	1.55	5.00E-05	0.006114	N/A	N/A
5	**GRMZM2G113332**	**1.37**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**copper chaperone**
	GRMZM2G144097	**2.07**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**protein**
	GRMZM2G020146	**1.50**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**serine carboxypeptidase iii precursor**
	GRMZM2G134903	1.63	5.00E-05	0.006114	Exonuclease; Uncharacterised protein	exonuclease family protein
6	**GRMZM2G147243**	**1.68**	**5.00E-05**	**0.006114**	**IAA17-auxin-responsive Aux/IAA family member; Uncharacterised protein**	**transcription factor**
	GRMZM2G070659	**1.65**	**5.00E-05**	**0.006114**	**Hypersensitive-induced response protein**	**hypersensitive-induced response protein**
	GRMZM2G130173	**2.38**	**5.00E-05**	**0.006114**	**Metallothionein-like protein type 2; Uncharacterised protein**	**N/A**

	GRMZM2G088501	1.97	5.00E-05	0.006114	Uncharacterised LOC100193404	N/A
	GRMZM2G158972	1.52	5.00E-05	0.006114	Putative inositol polyphosphate phosphatase (Synaptogenin-like) family protein	ipsi (inositol polyphosphate 5-phosphatase i) inositol triphosphate phosphatas inositol-polyphosphate 5-phosphatase
7	**GRMZM2G099049**	**2.68**	**5.00E-05**	**0.006114**	**N/A**	**N/A**
	GRMZM2G427815	**2.44**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**peroxidase**
	GRMZM2G473001	**1.45**	**5.00E-05**	**0.006114**	**Phosphoenolpyruvate carboxylase 2**	**phosphoenolpyruvate carboxylase**
	GRMZM2G149916	2.25	5.00E-05	0.006114	N/A	N/A
8	**GRMZM2G026470**	**2.09**	**5.00E-05**	**0.006114**	**Soluble inorganic pyrophosphatase; Uncharacterised protein**	**soluble inorganic pyrophosphatase**
	GRMZM2G477503	**2.35**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**sulfolipid synthase**
	GRMZM5G892675	**3.12**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**N/A**
	GRMZM2G063431	**2.61**	**5.00E-05**	**0.006114**	**N/A**	**N/A**
	GRMZM2G168552	**1.56**	**5.00E-05**	**0.006114**	**Bundle sheath cell specific protein 1**	**N/A**
	GRMZM2G087875	**3.12**	**5.00E-05**	**0.006114**	**Putative cytochrome P450 superfamily protein; Uncharacterised protein**	**cytochrome p450 family expressed**
	GRMZM2G007151	**2.27**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**endomembrane-associated protein**
	GRMZM2G110504	**2.41**	**5.00E-05**	**0.006114**	**Uncharacterised LOC100278648**	**hypothetical protein LOC100278648 [Zea mays]**
	GRMZM2G141353	**1.36**	**5.00E-05**	**0.006114**	**Uncharacterised LOC100194210**	**N/A**
9	**GRMZM2G178546**	**2.71**	**5.00E-05**	**0.006114**	**Trehalose-phosphate phosphatase**	**N/A**
	GRMZM2G006973	**2.78**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**N/A**
	GRMZM2G443728	**2.97**	**5.00E-05**	**0.006114**	**Potassium transporter 10**	**potassium transporter**

						10
	GRMZM2G116966	1.33	5.00E-05	0.006114	Benzoate carboxyl methyltransferase	N/A
10	**GRMZM2G008247**	**1.73**	**5.00E-05**	**0.006114**	**Beta-glucosidase 2**	**N/A**
	GRMZM2G034152	**2.00**	**5.00E-05**	**0.006114**	**Polyamine oxidase**	**polyamine oxidase precursor**
	GRMZM2G006490	1.91	5.00E-05	0.006114	N/A	N/A
	GRMZM2G173413	3.44	5.00E-05	0.006114	allantoinase [Source:Projected from Arabidopsis thaliana (AT4G04955) TAIR;Acc:AT4G04955]	protein

Table A.2: Significantly down-regulated genes of Protocol 1 after *F. verticillioides* infection as detected using RNA-seq with the Tuxedo suite of analysis and mapping to the maize B73 v3 genome. Table shows annotation of genes as described in Plant Ensembl, NCBI and Maize microarray annotation database (Blast2GO).

Chr	Gene stable ID	log2 (fold-change)	*p*-value	*q*-value	Ensembl/ NCBI gene description	Blast2GO description
1	**GRMZM2G070172**	**-3.88**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**N/A**
	GRMZM2G478568	**-1.62**	**5.00E-05**	**0.006114**	**Nicotianamine synthase 3**	**nicotianamine synthase 3**
	GRMZM2G147687	**-1.89**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**glycosyl hydrolase family 3 n terminal domain containing expressed**
	GRMZM2G053669	**-2.07**	**5.00E-05**	**0.006114**	**Asparagine synthetase**	**asparagine synthetase**
	GRMZM5G815358	-1.80	5.00E-05	0.006114	Uncharacterised LOC100274071	N/A
	GRMZM2G106928	-2.52	5.00E-05	0.006114	Superoxide dismutase [Cu-Zn]	copper zinc superoxide dismutase
	GRMZM5G857674	-1.60	5.00E-05	0.006114	N/A	N/A
	GRMZM2G066441	-1.49	5.00E-05	0.006114	Uncharacterised protein	N/A
2	**GRMZM2G125775**	**-2.36**	**5.00E-05**	**0.006114**	**AN17**	**arsenite inducuble rna associated protein aip-701**
	GRMZM2G047474	**-1.49**	**5.00E-05**	**0.006114**	**TLD-domain containing nucleolar protein**	**protein**
	GRMZM2G499582	-FC	5.00E-05	0.006114	N/A	N/A
	GRMZM2G348956	-1.39	5.00E-05	0.006114	DNA mismatch repair protein MutS, type 2	2 family protein
	GRMZM2G078480	-1.65	5.00E-05	0.006114	Cinnamoyl-CoA reductase 2-like	N/A
	GRMZM2G121264	**-1.70**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**cytochrome p450**
3	**GRMZM2G103812**	**-1.66**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**selenium-binding protein**
	GRMZM2G004161	**-3.03**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**btb and taz domain protein**
	GRMZM2G468111	**-3.46**	**5.00E-05**	**0.006114**	**Uncharacterised LOC100277849**	**N/A**

	GRMZM2G181081	**-2.27**	**5.00E-05**	**0.006114**	**CIPK-like protein 1**	**cipk-like protein expressed**
	GRMZM2G472248	**-3.23**	**5.00E-05**	**0.006114**	**Protein induced upon tuberization**	**N/A**
	GRMZM2G024733	**-1.60**	**5.00E-05**	**0.006114**	**Uncharacterised LOC100304285**	**pq-loop repeat family protein**
	GRMZM2G479423	-1.50	5.00E-05	0.006114	Aldose reductase	aldose reductase
	GRMZM5G814451	-1.40	5.00E-05	0.006114	N/A	3-cyclic-nucleotide phosphodiesterase rega
	GRMZM2G102572	-1.70	5.00E-05	0.006114	Isoamyl acetate-hydrolyzing esterase	isoamyl acetate-hydrolyzing
4	**GRMZM2G358153**	**-2.02**	**5.00E-05**	**0.006114**	**Chitinase 1; Uncharacterised protein**	**chitinase 1**
	GRMZM2G173085	**-1.86**	**5.00E-05**	**0.006114**	**Lipase/lipooxygenase, PLAT/LH2 family protein**	**potential zinc finger protein**
	GRMZM2G079381	**-1.92**	**5.00E-05**	**0.006114**	**Ferredoxin--nitrite reductase, chloroplastic**	**nitrite reductase**
*	**AC214438.3_FG002**	**-FC**	**5.00E-05**	**0.006114**	**N/A**	**N/A**
	GRMZM2G133675	**-2.74**	**5.00E-05**	**0.006114**	**Putative HLH DNA-binding domain superfamily protein; Uncharacterised protein**	**amelogenin precursor like protein**
	GRMZM2G366659	**-1.49**	**5.00E-05**	**0.006114**	**Putative trehalose phosphatase/synthase family protein**	**trehalose 6-phosphate synthase**
	GRMZM2G161198	-FC	5.00E-05	0.006114	N/A	N/A
	GRMZM2G157522	-1.57	5.00E-05	0.006114	Hypothetical protein LOC103654120	N/A
	GRMZM2G033236	-1.85	5.00E-05	0.006114	N/A	ubiquitin-activating enzyme
	GRMZM2G326911	-FC	5.00E-05	0.006114	N/A	N/A
	GRMZM2G570968	-FC	5.00E-05	0.006114	N/A	N/A
	GRMZM2G377695	-FC	5.00E-05	0.006114	N/A	N/A
5	GRMZM2G171539	-FC	5.00E-05	0.006114	N/A	mitochondrial nadh ubiquinone oxidoreductase 13kd-like subunit
6	GRMZM5G873765	-2.84	5.00E-05	0.006114	N/A	N/A

	GRMZM2G093325	-1.37	5.00E-05	0.006114	CONTAINS InterPro DOMAIN/s: Sgf11, transcriptional regulation (InterPro:IPR013246)	N/A
	GRMZM5G870170	**-1.59**	**5.00E-05**	**0.006114**	**MATE1**	**mate efflux family protein**
7	**GRMZM2G176430**	**-3.35**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**sodium-dicarboxylate cotransporter**
	GRMZM2G422955	**-2.28**	**5.00E-05**	**0.006114**	**N/A**	**N/A**
	GRMZM2G099879	-FC	5.00E-05	0.006114	N/A	N/A
8	GRMZM2G519073	-1.58	5.00E-05	0.006114	Uncharacterised protein	slt1 protein
	GRMZM2G154278	**-FC**	**5.00E-05**	**0.006114**	**Pre-mRNA-splicing factor cwc15**	**N/A**
	GRMZM2G146004	**-FC**	**5.00E-05**	**0.006114**	**Uncharacterised protein**	**N/A**
9	**GRMZM2G078472**	**-2.44**	**5.00E-05**	**0.006114**	**Asparagine synthetase**	**asparagine synthetase**
	GRMZM2G042510	-1.79	5.00E-05	0.006114	N/A	N/A
	GRMZM2G152417	-1.96	5.00E-05	0.006114	AMP-binding protein; Putative AMP-dependent synthetase and ligase superfamily protein; Uncharacterised protein	amp dependent
10	**GRMZM2G124495**	**-2.52**	**5.00E-05**	**0.006114**	**Putative MYB DNA-binding domain superfamily protein; Transfactor; Uncharacterised protein**	**N/A**
	GRMZM2G058612	**-2.40**	**5.00E-05**	**0.006114**	**F-box/LRR-repeat protein 3-like**	**N/A**
	GRMZM2G097641	**-2.25**	**5.00E-05**	**0.006114**	**Sucrose-phosphatase 2**	**sucrose phosphate synthase**
	GRMZM2G177077	**-1.40**	**5.00E-05**	**0.006114**	**Glucose-6-phosphate 1-dehydrogenase**	**glucose-6-phosphate 1-dehydrogenase**
	GRMZM2G152135	-2.12	5.00E-05	0.006114	Beta-carotene hydroxylase 1	beta-carotene hydroxylase 1

Table A.3: Significantly up-regulated genes from Protocol 2 after *F. verticillioides* infection as detected using RNA-seq with the Tuxedo suite of analysis and mapping to the maize B73 v3 genome. Table shows annotation of genes as described in Plant Ensembl, NCBI and Maize microarray annotation database (Blast2GO).

Chr	Gene stable ID	log2 (fold-change)	*p*-value	*q*-value	Ensembl / NCBI gene description	Blast2GO description
1	**GRMZM2G062724**	**3.35**	**5.00E-05**	0.004255	**Uncharacterised protein**	**chy zinc finger family expressed**
	GRMZM2G034302	**2.30**	**5.00E-05**	0.004255	**Uncharacterised protein**	**sucrose transporter**
	GRMZM2G130149	**2.04**	**5.00E-05**	0.004255	**Uncharacterised protein**	**myb family transcription expressed**
	GRMZM2G137861	**2.84**	**5.00E-05**	0.004255	**Wall-associated receptor kinase 2-like**	**N/A**
	GRMZM2G456217	2.82	5.00E-05	0.004255	Vignain	cysteine proteinase
	GRMZM2G161274	2.15	5.00E-05	0.004255	Ribonuclease 3; Uncharacterised protein	s-like rnase
	GRMZM2G073725	1.73	5.00E-05	0.004255	Alpha-1,4-glucan-protein synthase [UDP-forming]	reversibly glycosylated polypeptide
	GRMZM2G099767	**1.61**	**5.00E-05**	0.004255	**ATMAP70-2**	**N/A**
	GRMZM2G091456	**2.23**	**5.00E-05**	0.004255	**Putative Uncharacterised protein**	**squalene expressed, squalene monooxygenase**
	GRMZM2G119755	FC	5.00E-05	0.004255	Cell number regulator 7	N/A
	GRMZM2G366681	**2.12**	**5.00E-05**	0.004255	**Hypothetical protein**	**N/A**
	GRMZM2G040369	**1.52**	**5.00E-05**	0.004255	**Elongation factor 2**	**protein**
	GRMZM2G076537	**1.79**	**5.00E-05**	0.004255	**Polynucleotidyl transferase, ribonuclease H-like superfamily protein**	**exonuclease family protein**
2	**GRMZM2G049538**	**4.40**	**5.00E-05**	0.004255	**Acyclic sesquiterpene synthase**	**ent-kaurene synthase b**
	GRMZM2G125669	**2.10**	**5.00E-05**	0.004255	**Alternative oxidase**	**alternative oxidase**

	GRMZM2G115451	**2.64**	**5.00E-05**	0.004255	**Uncharacterised protein**	**neutral alkaline invertase**
	GRMZM2G144083	**1.99**	**0.0001**	0.007870	**Putative ATP dependent copper transporter**	**heavy metal p-type atpase**
	GRMZM2G062156	2.03	5.00E-05	0.004255	Uncharacterised protein	N/A
	GRMZM2G176433	**1.75**	**5.00E-05**	0.004255	**Putative Uncharacterised protein**	N/A
	GRMZM2G119975	**4.07**	**5.00E-05**	0.004255	**Uncharacterised LOC103646336**	**N/A**
	GRMZM2G040095	3.21	5.00E-05	0.004255	Lipoxygenase	lipoxygenase
	GRMZM2G093826	**2.68**	**5.00E-05**	0.004255	**Potassium high-affinity transporter**	**high-affinity potassium transporter**
	GRMZM2G106413	1.34	5.00E-05	0.004255	Uncharacterised protein LOC100282066 / wound induced protein	N/A
3	**GRMZM2G062531**	**2.62**	**5.00E-05**	0.004255	**Uncharacterised protein**	**c-4 sterol methyl oxidase**
	GRMZM2G057823	**1.63**	**5.00E-05**	0.004255	**Fructose-bisphosphate aldolase, cytoplasmic isozyme**	**fructose-bisphosphate aldolase**
	GRMZM2G022915	2.89	0.0001	0.007870	N/A	N/A
	GRMZM5G874955	**3.26**	**5.00E-05**	0.004255	**Uncharacterised protein**	**pdr-like abc transporter**
	GRMZM2G143139	**3.05**	**5.00E-05**	0.004255	**N/A**	**N/A**
	GRMZM2G402977	FC	0.0001	0.007870	N/A	N/A
	GRMZM2G029219	**3.38**	**5.00E-05**	0.004255	**Carbohydrate transporter/ sugar porter/ transporter**	**major facilitator superfamily antiporter/ carbohydrate transporter sugar porter transporter**
	GRMZM2G141665	1.67	5.00E-05	0.004255	Uncharacterised protein	syringomycin biosynthesis enzyme
4	**GRMZM2G098346**	**3.12**	**5.00E-05**	0.004255	**Alcohol dehydrogenase 2**	**alcohol dehydrogenase**
	GRMZM2G026922	**3.26**	**5.00E-05**	0.004255	**Hypothetical protein**	**acetylglutamate kinase**

	GRMZM2G070011	**2.89**	**5.00E-05**	0.004255	**Uncharacterised protein; Vignain**	**vignain precursor**
	GRMZM2G036464	**2.79**	**5.00E-05**	0.004255	**Glutamine synthetase root isozyme 4**	**glutamine synthetase**
	GRMZM2G047319	FC	5.00E-05	0.004255	Putative subtilase family protein	N/A
	GRMZM2G149422	**3.22**	**5.00E-05**	0.004255	**Hypothetical protein**	**phi-1**
	GRMZM2G041699	2.47	5.00E-05	0.004255	Cytokinin-O-glucosyltransferase 2	cytokinin-o-glucosyltransferase 2
	GRMZM2G343828	FC	5.00E-05	0.004255	Putative O-Glycosyl hydrolase superfamily protein	N/A
	GRMZM2G015295	1.60	5.00E-05	0.004255	Adenosylhomocysteinase	s-adenosyl-l-homocysteine hydrolase
	GRMZM2G116079	**1.92**	**5.00E-05**	0.004255	**Uncharacterised protein**	N/A
	GRMZM2G015419	2.01	5.00E-05	0.004255	Uncharacterised protein	lipoxygenase, lipoxygenase chloroplast precursor
	GRMZM2G036217	**1.96**	**5.00E-05**	0.004255	**Uncharacterised protein**	**fatty acyl coa reductase**
	GRMZM2G154523	**3.02**	**5.00E-05**	0.004255	**Patatin T5; Uncharacterised protein**	N/A
*	**AC217947.4_FG002.2**	**2.25**	**0.0001**	0.007870	**N/A**	**N/A**
	GRMZM2G117971	**3.11**	**5.00E-05**	0.004255	**Uncharacterised protein**	**pathogenesis-related protein 4**
5	GRMZM2G094353	3.54	5.00E-05	0.004255	Uncharacterised protein	rna-binding protein cabeza
	GRMZM2G038874	2.30	5.00E-05	0.004255	N/A	N/A
	GRMZM2G113332	**1.47**	**5.00E-05**	0.004255	**Uncharacterised protein**	**copper chaperone**
	GRMZM2G165530	2.04	0.0001	0.007870	Putative Uncharacterised protein	tetracycline transporter
	GRMZM2G144097	**2.00**	**0.0001**	0.007870	**Uncharacterised protein**	**protein**
*	AC225718.2_FG004	FC	5.00E-05	0.004255	N/A	N/A
	GRMZM2G060659	1.67	5.00E-05	0.004255	Putative Uncharacterised protein	protein
	GRMZM2G173192	FC	5.00E-05	0.004255	Uncharacterised protein	N/A
	GRMZM2G020146	**1.53**	**5.00E-05**	0.004255	**Uncharacterised protein**	**serine**

						carboxypeptidase iii precursor
	GRMZM2G130053	3.30	5.00E-05	0.004255	Cysteine protease 1	N/A
	GRMZM2G011888	1.96	5.00E-05	0.004255	Putative Uncharacterised protein	N/A
	GRMZM2G075333	3.01	5.00E-05	0.004255	Uncharacterised protein	4-coumarate: ligase
6	GRMZM2G036564	1.59	5.00E-05	0.004255	Transmembrane protein20	embryogenesis transmembrane
	GRMZM2G147243	**1.76**	**5.00E-05**	0.004255	**IAA17-auxin-responsive Aux/IAA family member; Uncharacterised protein**	**transcription factor**
	GRMZM2G124799	FC	5.00E-05	0.004255	Uncharacterised protein	N/A
	GRMZM2G070659	**1.89**	**5.00E-05**	0.004255	**Hypersensitive-induced response protein**	**hypersensitive-induced response protein**
	GRMZM5G844094	2.44	5.00E-05	0.004255	N/A	N/A
	GRMZM2G130173	**2.44**	**5.00E-05**	0.004255	**Metallothionein-like protein type 2; Uncharacterised protein**	**N/A**
	GRMZM2G100719	FC	0.00005	0.004255	N/A	N/A
7	**GRMZM2G099049**	**2.44**	**5.00E-05**	0.004255	**N/A**	**N/A**
	GRMZM2G003179	1.81	5.00E-05	0.004255	copper transporter 5	N/A
	GRMZM2G086714	1.90	0.0001	0.007870	Uncharacterised LOC103632825	plastid ppgpp synthase
	GRMZM2G050172	1.90	0.0001	0.007870	Uncharacterised LOC103632825	plastid ppgpp synthase
	GRMZM5G884407	2.07	5.00E-05	0.004255	N/A	aldehyde dehydrogenase
	GRMZM5G817559	2.65	5.00E-05	0.004255	Uncharacterised protein	protein
	GRMZM2G415529	2.33	5.00E-05	0.004255	N/A	pdr-like abc transporter
	GRMZM2G366977	FC	5.00E-05	0.004255	Equilibrative nucleotide transporter 3-like	N/A
	GRMZM2G427815	**2.59**	**5.00E-05**	0.004255	**Uncharacterised protein**	**peroxidase**
	GRMZM2G170734	FC	5.00E-05	0.004255	Chlorophyllase-2, chloroplastic-like	N/A
	GRMZM2G473001	**1.51**	**5.00E-05**	0.004255	**Phosphoenolpyruvate carboxylase 2**	**phosphoenolpyruvate carboxylase**
8	**GRMZM2G026470**	**2.36**	**5.00E-05**	0.004255	**Soluble inorganic pyrophosphatase;**	**soluble inorganic**

					Uncharacterised protein	**pyrophosphatase**
	GRMZM6G198866	1.68	5.00E-05	0.004255	Metallothionein-like protein type 2	N/A
	GRMZM2G070912	1.68	5.00E-05	0.004255	Putative metallothionein family protein	N/A
	GRMZM2G477503	**2.36**	**5.00E-05**	0.004255	**Uncharacterised protein**	**sulfolipid synthase**
	GRMZM5G892675	**2.90**	**5.00E-05**	0.004255	**Uncharacterised protein**	**N/A**
	GRMZM2G063431	**2.46**	**5.00E-05**	0.004255	**N/A**	**N/A**
	GRMZM2G077054	1.37	5.00E-05	0.004255	Uncharacterised protein	nadh-dependent glutamate synthase 1 gene
	GRMZM2G173718	1.40	5.00E-05	0.004255	Uncharacterised LOC100273627	N/A
	GRMZM2G455124	FC	5.00E-05	0.004255	nitrate transporter2.	N/A
	GRMZM2G168552	**1.51**	**5.00E-05**	0.004255	**Bundle sheath cell specific protein 1**	**N/A**
	GRMZM5G875238	1.42	0.0001	0.007870	Sucrose-phosphate synthase	sucrose phosphate synthase
	GRMZM2G022958	1.92	5.00E-05	0.004255	Uncharacterised LOC100275172	N/A
	GRMZM2G087875	**3.28**	**5.00E-05**	0.004255	**Putative cytochrome P450 superfamily protein; Uncharacterised protein**	**cytochrome p450 family expressed**
	GRMZM2G007151	**2.57**	**5.00E-05**	0.004255	**Uncharacterised protein**	**endomembrane-associated protein**
	GRMZM2G054123	3.62	5.00E-05	0.004255	S-adenosylmethionine synthase	N/A
	GRMZM2G097141	5.20	5.00E-05	0.004255	N/A	N/A
	GRMZM2G110504	**2.51**	**5.00E-05**	0.004255	**Uncharacterised LOC100278648**	**hypothetical protein LOC100278648 [Zea mays]**
	GRMZM2G141353	**1.45**	**5.00E-05**	0.004255	Uncharacterised LOC100194210	N/A
9	**GRMZM2G178546**	**2.51**	**5.00E-05**	0.004255	**Trehalose-phosphate phosphatase**	**N/A**
	GRMZM2G132238	2.13	5.00E-05	0.004255	Putative metacaspase family protein	N/A
	GRMZM2G479243	7.08	5.00E-05	0.004255	Putative leucine-rich repeat receptor-like protein kinase family protein	brassinosteroid insensitive 1-associated receptor kinase 1
	GRMZM2G006973	**2.68**	**5.00E-05**	0.004255	**Uncharacterised protein**	**N/A**

	GRMZM2G035285	3.66	0.0001	0.007870	N/A	N/A
	GRMZM2G443728	**2.94**	**5.00E-05**	0.004255	**Potassium transporter 10**	**potassium transporter 10**
10	GRMZM2G147390	FC	5.00E-05	0.004255	Uncharacterised protein	N/A
	GRMZM2G034882	1.87	5.00E-05	0.004255	Uncharacterised LOC100276570	N/A
	GRMZM2G008247	**1.82**	**5.00E-05**	0.004255	**Beta-glucosidase2**	N/A
	GRMZM2G034152	**2.14**	**5.00E-05**	0.004255	**Polyamine oxidase**	**polyamine oxidase precursor**
	GRMZM2G163998	1.85	0.0001	0.007870	Uncharacterised protein; VAMP protein SEC22	N/A
scaffold_510:0-2226	GRMZM6G761998	1.89	5.00E-05	0.004255	Zinc transporter 2	N/A

Table A.4: Significantly down-regulated genes of Protocol 2 after *F. verticillioides* infection as detected using RNA-seq with the Tuxedo suite of analysis and mapping to the maize B73 v3 genome. Table shows annotation of genes as described in Plant Ensembl, NCBI and Maize microarray annotation database (Blast2GO).

Chr	Gene stable ID	log2 (fold-change)	*p*-value	*q*-value	Ensembl/ NCBI gene description	Blast2GO description
1	GRMZM2G070685	-1.86	5.00E-05	0.004255	N/A	N/A
	GRMZM2G001877	-1.34	5.00E-05	0.004255	N/A	N/A
	GRMZM2G070172	**-4.06**	**5.00E-05**	**0.004255**	**Uncharacterised protein**	N/A
	GRMZM2G039639	-1.83	5.00E-05	0.004255	Protein P21	pathogenesis-related thaumatin-like protein
	GRMZM2G478568	**-1.63**	**5.00E-05**	**0.004255**	**Nicotianamine synthase 3**	**nicotianamine synthase 3**
	GRMZM2G147687	**-2.15**	**5.00E-05**	**0.004255**	**Uncharacterised protein**	**glycosyl hydrolase family 3 n terminal domain containing expressed**
	GRMZM2G061626	-FC	0.0001	0.007870	ZFP16-2	
	GRMZM2G053669	**-1.92**	**5.00E-05**	**0.004255**	**Asparagine synthetase**	**asparagine synthetase**
2	**GRMZM2G125775**	**-2.44**	**5.00E-05**	**0.004255**	**AN17**	**arsenite inducuble rna associated protein aip-701**
	GRMZM2G121264	**-1.78**	**5.00E-05**	**0.004255**	**Uncharacterised protein**	**cytochrome p450**
	GRMZM2G047474	**-1.35**	**5.00E-05**	**0.004255**	**TLD-domain containing nucleolar protein**	**protein**
	GRMZM2G015024	-1.34	5.00E-05	0.004255	50S ribosomal protein L22, chloroplastic	N/A
3	**GRMZM2G103812**	**-1.73**	**5.00E-05**	**0.004255**	**Uncharacterised protein**	**selenium-binding protein**
	GRMZM2G166548	-2.09	5.00E-05	0.004255	N/A	N/A
	GRMZM2G004161	**-2.80**	**5.00E-05**	**0.004255**	**Uncharacterised protein**	**N/A**
	GRMZM2G468111	**-3.71**	**5.00E-05**	**0.004255**	**Uncharacterised LOC100277849**	**N/A**
	GRMZM2G181081	**-2.14**	**5.00E-05**	**0.004255**	**CIPK-like protein 1**	**cipk-like protein expressed**

*	AC194022.3_FG013.1	-2.76	5.00E-05	0.004255	N/A	N/A
	GRMZM2G472248	**-3.11**	**5.00E-05**	**0.004255**	**Protein induced upon tuberization**	N/A
	GRMZM2G024733	**-1.54**	**5.00E-05**	0.004255	**Uncharacterised LOC100304285**	**pq-loop repeat family protein**
4	**GRMZM2G358153**	**-2.15**	**5.00E-05**	**0.004255**	**Chitinase 1; Uncharacterised protein**	**chitinase 1**
	GRMZM5G845532	-FC	0.0001	0.007870		
	GRMZM2G173085	**-1.85**	**5.00E-05**	**0.004255**	**Lipase/lipooxygenase, PLAT/LH2 family protein**	**potential zinc finger protein**
	GRMZM2G079381	**-1.79**	**5.00E-05**	**0.004255**	**Ferredoxin--nitrite reductase, chloroplastic**	**nitrite reductase**
	GRMZM5G863229	-1.77	5.00E-05	0.004255	Uncharacterised protein	protein
*	**AC214438.3_FG002.1**	**-FC**	**5.00E-05**	**0.004255**	**N/A**	**N/A**
	GRMZM2G133675	**-2.56**	**5.00E-05**	**0.004255**	**Putative HLH DNA-binding domain superfamily protein; Uncharacterised protein**	**amelogenin precursor like protein**
	GRMZM2G366659	**-1.49**	**5.00E-05**	**0.004255**	**Putative trehalose phosphatase/synthase family protein**	**trehalose 6-phosphate synthase**
	GRMZM2G061126	-1.35	5.00E-05	0.004255	Hypothetical protein	N/A
5	AC212351.4_FG001.1	-1.32	5.00E-05	0.004255	Uncharacterised LOC100502277	N/A
	GRMZM2G057766	-FC	5.00E-05	0.004255	Chitinase 1	N/A
	GRMZM5G878558	-3.61	5.00E-05	0.004255	Uncharacterised protein	nitrate reductase
	GRMZM2G083788	-2.00	0.0001	0.004255	Vacuolar amino acid transporter 1-like	amino acid transporter family protein
	GRMZM2G168747	-FC	5.00E-05	0.007870	Metal transporter NRAT1-like	N/A
6	GRMZM5G847462	-FC	5.00E-05	0.004255	N/A	N/A
	GRMZM5G870170	**-1.54**	**5.00E-05**	**0.004255**	**MATE1**	**mate efflux family protein**
7	**GRMZM2G176430**	**-3.03**	**5.00E-05**	**0.004255**	**Uncharacterised protein**	**sodium-dicarboxylate cotransporter**
	GRMZM2G422955	**-2.24**	**5.00E-05**	**0.004255**	**N/A**	**N/A**
	GRMZM2G016212	-FC	5.00E-05	0.004255	N/A	N/A

8	GRMZM2G020594	-4.97	5.00E-05	0.004255	F-box domain containing protein; F-box domain containing protein isoform 1; F-box domain containing protein isoform 2; Uncharacterised protein	N/A
	GRMZM2G154278	**-3.36**	**5.00E-05**	**0.004255**	**Pre-mRNA-splicing factor cwc15**	**N/A**
	GRMZM2G146004	**-FC**	**5.00E-05**	**0.004255**	**Uncharacterised protein**	**N/A**
9	**GRMZM2G078472**	**-2.20**	**5.00E-05**	**0.004255**	**Asparagine synthetase**	**asparagine synthetase**
10	GRMZM2G455476	-2.06	0.0001	0.004255	Uncharacterised protein	white-brown-complex abc transporter family
	GRMZM2G124495	**-2.24**	**5.00E-05**	**0.007870**	**Putative MYB DNA-binding domain superfamily protein; Transfactor; Uncharacterised protein**	**N/A**
	GRMZM2G058612	**-2.39**	**5.00E-05**	**0.004255**	**F-box/LRR-repeat protein 3-like**	**N/A**
	GRMZM2G064008	-FC	5.00E-05	0.004255	N/A	N/A
	GRMZM2G097641	**-2.28**	**5.00E-05**	**0.004255**	**Sucrose-phosphatase 2**	**sucrose phosphate synthase**
	GRMZM2G177077	**-1.42**	**5.00E-05**	**0.004255**	**Glucose-6-phosphate 1-dehydrogenase**	**glucose-6-phosphate 1-dehydrogenase**
scaffold_509:407701-410547	GRMZM2G149326	-1.77	5.00E-05	0.004255	N/A	N/A

Appendix B: Quantitative Real-time PCR analysis

UBCE

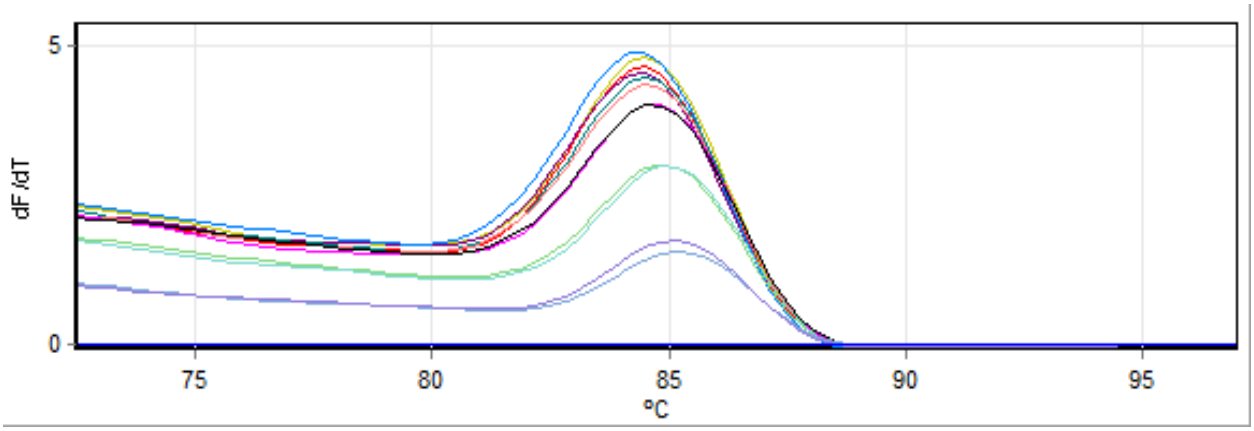

Figure B1: Melt curve analysis of the ***UBCE*** reference gene.

Rpol

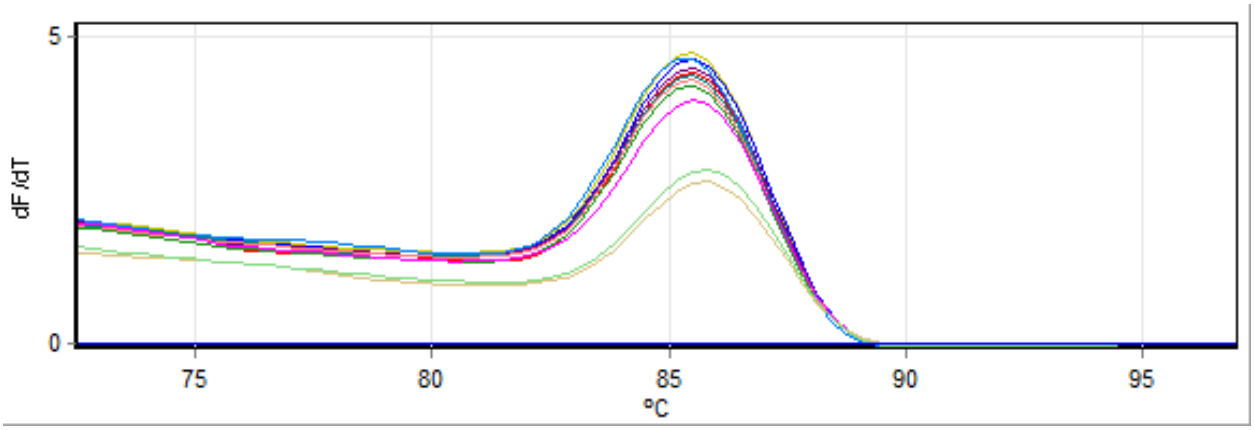

Figure B2: Melt curve analysis of the ***Rpol*** reference gene.

MEP

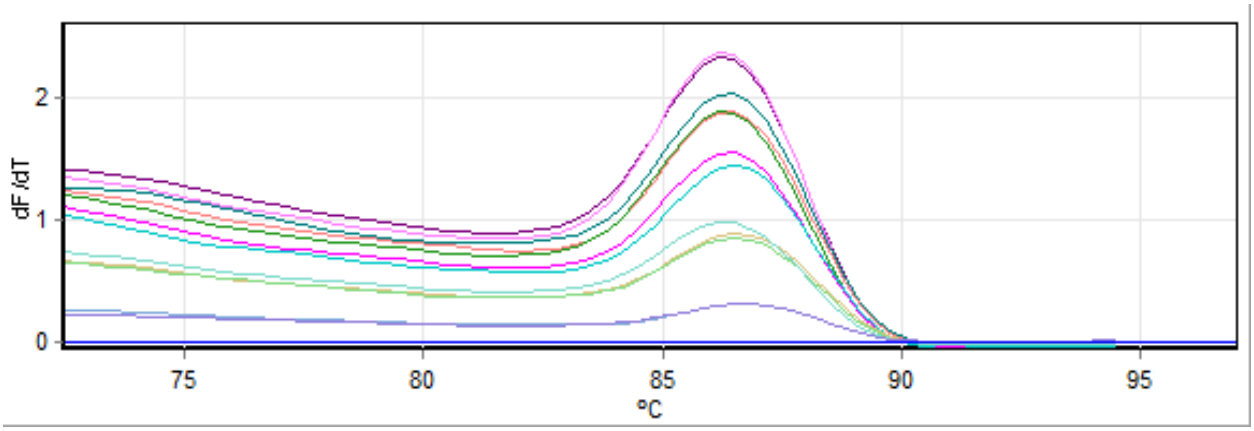

Figure B3: Melt curve analysis of the ***MEP*** reference gene.

Sh1

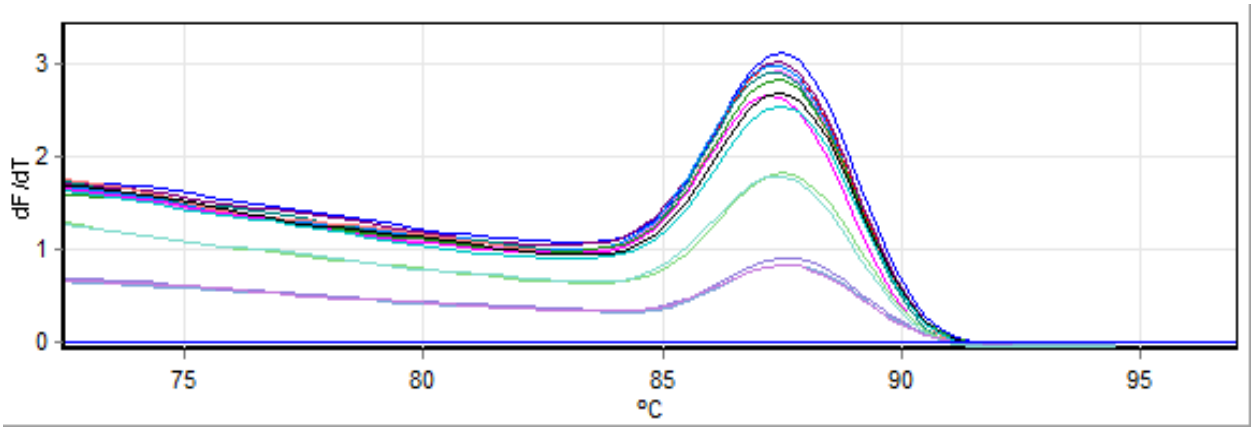

Figure B4: Melt curve analysis of the ***Sh1*** pre-RNAseq validation gene.

HIR3

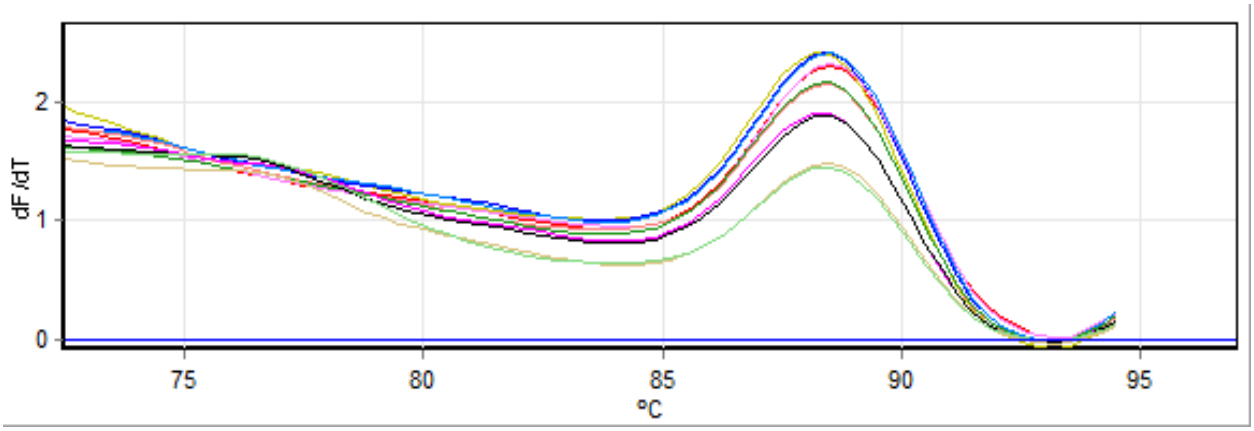

Figure B5: Melt curve analysis of the ***HIR3*** up-regulated gene.

lox6

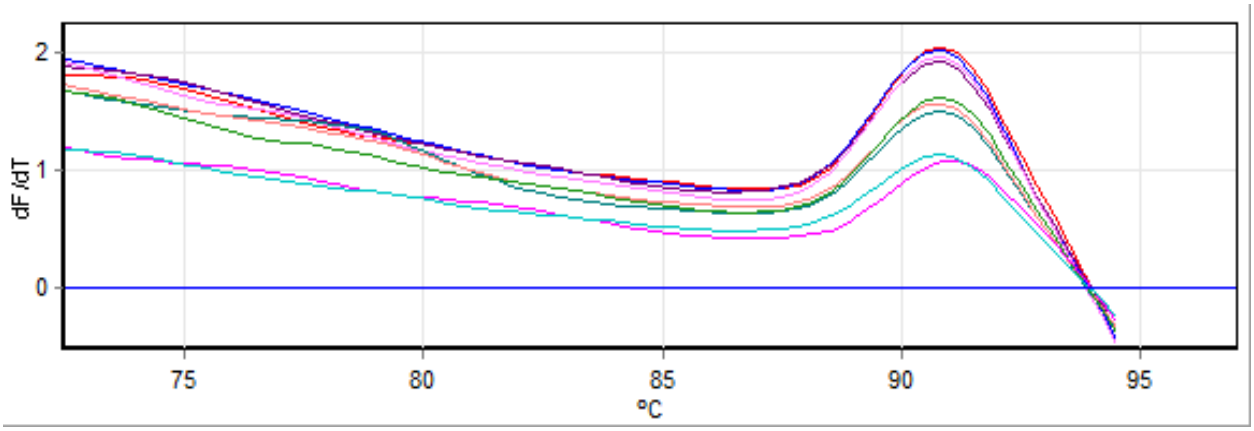

Figure B6: Melt curve analysis of the ***lox6*** up-regulated gene.

Chitinase 1

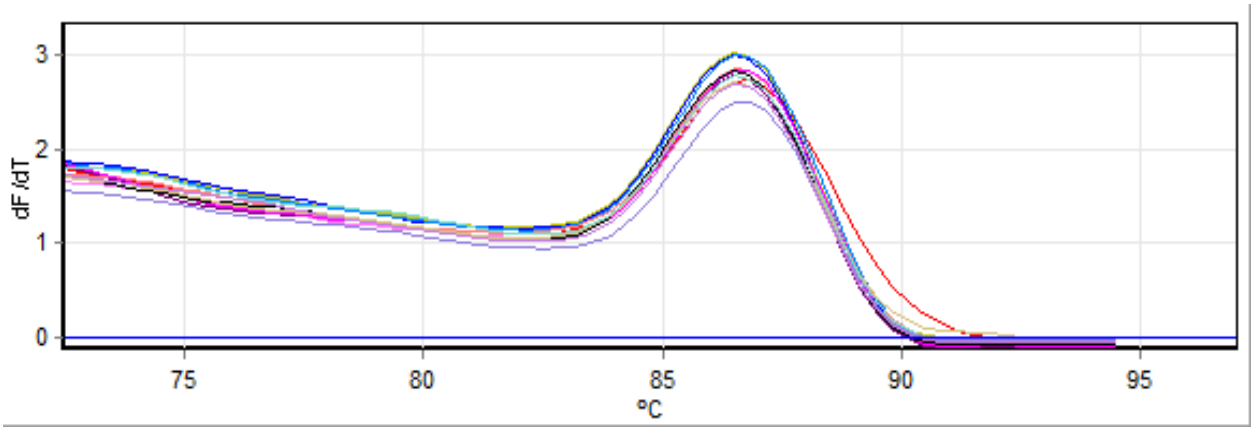

Figure B7: Melt curve analysis of the ***Chitinase 1*** down-regulated gene.

PR-th

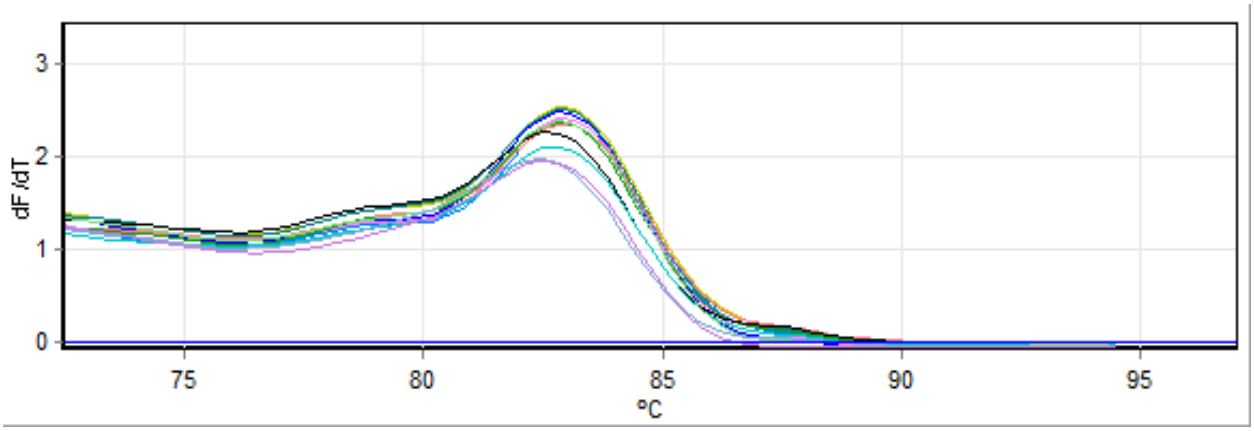

Figure B8: Melt curve analysis of the **PR-th** down-regulated gene.

PIT

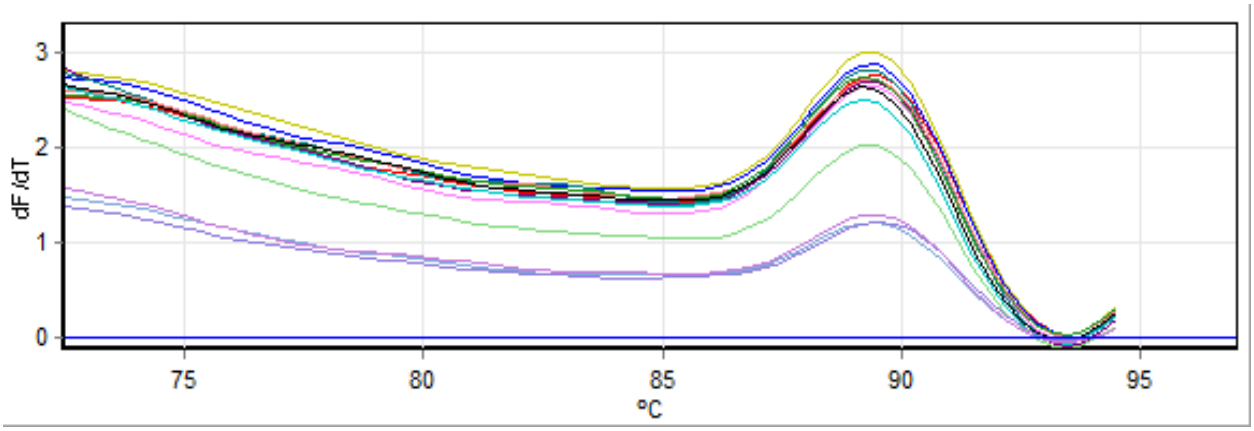

Figure B9: Melt curve analysis of the **PIT** down-regulated gene.