A GENOMICS APPROACH TO THE DEVELOPMENT OF MOLECULAR DIAGNOSTIC TOOLS FOR BANANA *XANTHOMONAS* WILT, GRAY LEAF SPOT OF MAIZE AND CASSAVA BROWN STREAK DISEASE.

A RESEARCH THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF DOCTOR OF PHILOSOPHY IN BIOCHEMISTRY OF THE UNIVERSITY OF NAIROBI

2014

By
Eric Maina Magembe
H80/81141/09

Declaration

I, Eric Maina Magembe, duly declare that this is my original work and has not been presented for a degree in any other university or any other award.

Eric Maina Magembe (M.Sc.)
Department of Biochemistry
University of Nairobi
P. O. Box 30197-00100, Nairobi, Kenya
Signature ..... Date .....

We confirm that the work reported here was carried out by the above named candidate and is submitted with our approval as supervisors.

Prof. Dong-Jin Kim
International Institute of Tropical Agriculture (IITA)
c/o International Livestock Research Institute
P. O. Box 30709-00100, Nairobi, Kenya
Signature ..... Date .....

Prof. P.W Kinyanjui
Department of Biochemistry
University of Nairobi
P. O. Box 30197-00100, Nairobi, Kenya
Signature ..... Date .....

Prof. J.C Mukuria
Department of Biochemistry
University of Nairobi
P. O. Box 30197-00100, Nairobi, Kenya
Signature ..... Date .....

Through,
Chairman, Department of Biochemistry, University of Nairobi.
P. O. Box 30197-00100, Nairobi, Kenya
Signature ..... Date .....

Dedication

I dedicate this thesis to my family: my son Leroy Maina and my parents Peter Magembe and Elizabeth Okorah, for their constant support and unconditional love. I love you all dearly.

Acknowledgements

I wish to express my sincere gratitute to Professor D.J Kim who recognised my potential and offered me an opportunity to undertake my PhD studies under his IITA project on Genomics approaches for Molecular Diagnostics of the crop diseases in Africa. His knowledge, guidance and the useful comments during my research are highly appreciated. I am greatful to my University of Nairobi supervisors, Profesor P.W Kinyanjui and Professor J.C Mukuria, for their continous support during my PhD study. Thank you for your continous guidance and help during the research and write up of my thesis.

I am very grateful to the Norman E. Borlaug Leadership Enhancement in Agriculture Program (LEAP) who offered me a fellowship to conduct part of this study at Texas A&M University under the mentorship of Professor Won-Bo Shim.

I thank my collaborators at the Kenya Agricultural Research Institute (KARI) during the conduct of my research especially Dr Kinyua of KARI-NARL.

A good support system is important to surviving and staying sane in PhD Study. I am very grateful for the support I received from my labmates, Gurdeep, Veronica Ogugo, Samuel Kariuki, Inosters Nzuki, Nelson Ndegwa, Mtakai Ngara and Titus Mureithi. I am grateful for Biosciences East and Central Africa (BecA-Hub) at ILRI for the lab facilities that enabled me conduct this research. I thank all my friends (old and new) at BecA for providing support and encouragement that I needed.

I will forever be thankful to my family for allowing me to take this journey. It has been very challenging being a student and a father at the same time. I thank my parents, Peter Magembe and Elizabeth Okorah, who have been a great encouragement during this journey. Most especially, I thank my dad for his great support both moraly and financialy which is greatly appreciated.

General abstract

Banana *Xanthomonas* wilt (BXW), gray leaf spot (GLS) disease of maize and cassava brown streak disease (CBSD) are very destructive crop plant bacterial, fungal and viral diseases, respectively. Conventional methods currently being used for detection of these pathogens are inadequate. To address this, specific and rapid molecular diagnostic tools have been developed for the detection of these pathogens. A combination of sequencing and bioinformatics approaches were used to design diagnostic PCR primers for the pathogens. The variability of intergenic regions and unique sequences were utilized in an attempt to design specific primers for PCR detection of BXW and GLS of maize. Sequencing was done for the whole genome of *Xanthomonas campestris* p.v *musacerium* (*Xcm*) for BXW and coat protein region for cassava brown streak virus for CBSD. This was the first genome sequence for Kenyan *Xcm*. A bioinformatics approach was utilised to design PCR diagnostic primers from intergenic regions and unique sequences for the pathogens. For more effective diagnosis of GLS of maize, unique EST sequences and Malazy gene, a gypsy-like transposable element that is differentially expressed in the two *Cercospora* groups of *Cercospora zeae-maydis* group I and C. *zeina* group II were utilized. Species specific primers developed were characterised and validated to be specific for the diagnosis of BXW, GLS of maize and CBSD. These molecular diagnostic tools can be used to detect these diseases before symptoms are visualized or spread and to identify new emerging pathogen varieties. These tools can also be used in detecting pathogens in suckers, seeds, cuttings and other plant materials used for propagation and in marker assisted selection for disease resistant lines.

Contents

List of Tables

List of Figures

Abbreviations and Acronyms

AIDS - Acquired Immunodeficiency Syndrome
BLAST - Basic Local Alighnment Search Tool
BSA - Bovine Serum Albumin
BXW - Banana *Xanthomonas* Wilt
C_t - Cycle Threshold
CBSD - Cassava Brown Streak Disease
CBSUV - Cassava Brown Streak Ugandan Virus CBSV - Cassava Brown Streak Virus
CDS - Coding Sequence
CFU - Colony Forming Units
CMD - Cassava Mosaic Disease
CMR - Comprehensive Microbial Resource
CP - Coat Protein
CTAB - Ctyltrimethylammonium Bromide
CVYV - Cucumber Vein Yellowing Virus
DEPC -Diethylpyrocarbonate
DNA - Deoxyribonucleic acid
DTT - Dithiothreitol
EB - Ethidium Bromide
EDTA - Ethylene Diamine Tetraacetic Acid
ELISA - Enzyme Linked Immunosorbent Assay
EMBL - The European Molecular Biology Laboratory
emPCR emulsion Polymerase Chain Reaction
EST - Expressed Sequence Tags
FAME - Fatty Acid Methyl Esters
FERA - Food and Environment Research Agency
GLS - Grey Leaf spot
GO - Gene Ontology
HIV - Human Immunodeficiency Virus
IC- PCR - Immunocapture Polymerase Chain Reaction
IF - Immunofluorescence
IFPRI- International Food Policy Research Institute
IITA - International Institute of Tropical Agriculture

JCVI - J. Craig Venter Institute
KARI - Kenya Agricultural Research Institute
KEPHIS - Kenya Plant Health Inspectorate Service
MPC - Magnetic Particle Collector
NCBI - National Center for Biotechnology Information
NCPPB - National Collection of Plant Pathogenic Bacteria
ORF - Open Reading Frame
PAbs - Polyclonal Antibodies
PCR - Polymerase Chain Reaction
PFU - Plaque Forming Unit
PVP - Polyvinylpolypyrrolidone
QTL - Quantitative Trait Loci
RNA - Ribonucleic acid
RT-PCR - Reverse Transcriptase PCR
SDS - Sodium dodecyl sulphate
SNP - Single Nucleotide Polymorphism
SPMMV - Sweet Potato Mild Mottle Virus
SqVYV - Squash Vein Yellowing Virus
TBE - TTris-borate-EDTA
YDC - Yeast Dextrose Chalk
YPGA - Yeast Peptone Glucose Agar

Chapter 1

GENERAL INTRODUCTION AND LITERATURE REVIEW

1.1 Background

Accurate identification and early detection of pathogens is the cornerstone of disease management in many crops. Early identification of the causative agent of disease is paramount in order to recognize the pathogen, and implement regulations involving control and quarantine. Pathogen detection is also important in many areas of plant pathology research, for example in: studies of disease epidemiology; yield loss disease relationships; and in designing new strategies for disease control. Pathogen or disease diagnosis is therefore, fundamental to virtually all aspects of plant pathology. Conventional methods to disease diagnostics often rely on interpretation of visual symptoms. This may be followed by laboratory identification, using for example selective media and microscopy, to confirm the diagnosis. In some cases these methods are still the cheapest, simplest and most appropriate. They can however lead to problems in identification, resulting in incorrect interpretation, diagnosis and ultimately treatment measures. The methods rely on experienced, skilled laboratory staff, the ability of the organism to be cultured, are time consuming, non-quantitative, prone to contamination and error and in the case of plant and medical pathology often delay treatment. There is a need to use more generic techniques that can be taught quickly and easily to relatively unskilled staff. Methods that involve culturing can often take days or weeks to complete and this is not acceptable when rapid, high throughput diagnosis is required. The results are not always conclusive, e.g. where similar symptoms can

be caused by different pathogens or physiological conditions. Closely related organisms may be difficult to discriminate on the basis of morphological characters alone. It may also be necessary to discriminate between populations of the same pathogen that have specific properties, e.g. fungicide resistance, toxin production or differences in virulence. Traditional methods may not be sensitive enough (e.g. where the detection of presymptomatic infection is needed). Molecular detection techniques can generate accurate results rapidly enough to be useful for disease management decisions. New, rapid screening methods are being developed and increasingly used in all aspects of disease diagnostics. These methods include immunological methods, DNA/RNA probe technology and polymerase chain reactions (PCR) technology. The first molecular detection technology for plant pathogens was antibody based, in particular monoclonal antibodies and enzyme-linked immunosorbent assay (ELISA) (Clark and Adams., 1977; Köhler and Milstein, 1975). This approach could identify and detect pathogens much more rapidly (Hampton et al., 1990; Miller and Martin, 1988). Then came DNA-based technologies, such as the polymerase chain reaction (PCR) (Mullis and Faloona, 1987). These techniques have several potential advantages over conventional diagnostic methods in that they are more accurate, faster and do not require a high level of plant pathology expertise and interpretation of the results. PCR is rapid, highly specific and can be used to detect minute quantities of pathogen DNA from environmental samples before symptoms occur, therefore, allowing implementation of early control methods. With improved DNA extraction methods from environmental samples available, and with adequate controls, PCRs for diagnostics can be run routinely. PCR technology can also provide very accurate quantitative data supplying the necessary additional information required for control and quarantine decisions, and for assessing how effective pathogen agents are in the case of biological control. The ability to design PCR primers to target specific regions of DNA has lead to a greater

understanding of pathogen ecology, pathogen-plant interactions, pathogen-pest interactions and pathogen-pathogen interactions. The DNA Microarray technology, originally designed to study gene expression and generate single nucleotide polymorphism (SNP) profiles, is currently a new and emerging pathogen diagnostic technology, which in theory, offers a platform for unlimited multiplexing capability. It is viewed as a technology that fundamentally alter molecular diagnostics. The fast growing databases generated by genomics and biosystematics research provides unique opportunity for the design of more versatile, high-throughput, sensitive and specific molecular assays which will address the major limitations of the current technologies and benefit plant pathology. However, while the specificity and sensitivity of detection of pathogens are greatly improved and pathogen detection is becoming simpler and faster, there are still major challenges, technical and economic nature, which need to be addressed to ensure the emergence of reliable detection system for routine applications. In developing a tool for pathogen detection, issues such as detection specificity and sensitivity are very important. In addition multiplexing, quantification and cost effectiveness are increasingly becoming important features of a diagnostic technology. There is also a growing need for a field deployable portable rapid detection system that provides the capability for pathogen testing and identification in the field.

1.2 Antibody based techniques for pathogens detection

Antibodies can be generated that recognize specific antigens associated with a given plant pathogen they can be used as the basis of a diagnostic tool (Werres and Steffens, 1994; Dewey and Thornton, 1995; Barker, 1996; Dewey et al., 1997; Torrance, 1998). The principal aim of an immunodiagnostic assay is to detect or quantify the binding of the diagnostic antibody with the target antigen. Polyclonal antibodies (PAbs) have been used successfully for detecting plant pathogens, especially viruses, but they are

not always sufficiently specific. Monoclonal antibodies are more specific and generally slow to produce (Torrance, 1995). A method of producing antibodies known as phage display, (McCafferty et al., 1990) has been used to develop immuno-diagnostic assays for plant pathogens. The technique uses libraries of functional fragments of antibody molecules which have been amplified by PCR from a number of different animal species, including humans. The antibody fragments contain the heavy and light chain variable domains of antibodies and hence the regions of hyper-variable sequence which bestow specificity to the antibodies. Immuno-diagnostic assays have been developed using this technique for the bacterial pathogen *Ralstonia solanacearum* Race 3 which causes brown rot of potatoes (Griep et al., 1998) and for a number of plant viruses, including blackcurrant reversion associated virus (Susi et al., 1998), potato virus Y (Boonham and Barker, 1998) and beet necrotic yellow vein virus (Griep et al., 1999). The development of such assays for fungal phytopathogens may be hampered by problems such as identification of antibodies that would give the required specificity when challenged with a complex antigen mixture such as mycelial washings or spores. However, a phage display library has been used to obtain recombinant antibodies to surface-exposed epitopes from both germlings and soluble mycelial components of *Phytophthora infestans* (Gough et al., 1999) and for a precursor of melanin from *Alternaria alternata* (Carzaniga et al., 2002). Use of phage display to isolate species-specific antibodies has a number of advantages over polyclonal and monoclonal antibodies; the method is generally quicker and cheaper, it does not require the use of animals, and it is easy to sustain an indefinite supply of valuable antibodies.

There are a number of different ways of detecting antibody/antigen binding, but often these involve coupling the antibody to an enzyme that can be used to generate a colour change when a substrate is added. ELISA is the most commonly used diagnostic technique that uses antibodies (Clark and Adams., 1977; Voller and Bidwell,

1985). It involves an enzyme-mediated colour change reaction to detect antibody binding. Plant pathogens have been detected *in situ* by squashing plant tissues directly on to membranes followed by immunodetection, that is, tissue printing or squash blot systems (Gwinn et al., 1991; Shine and Comstock, 1993). Immunofluorescence (IF) microscopy is another method that allows in situ localisation of the pathogen. Plant samples are applied to microscope slides in thin tissue sections and fixed. Detection is achieved by conjugating a fluorescent dye to the specific antibody (direct IF) or to a molecule that detects the specific antibody (indirect IF). This method has been used to identify specific plant pathogen spores on microscope slides (Kennedy et al., 1999). However, it is not suitable for high throughput screening and it requires specialised equipment (an ultraviolet microscope) and relatively skilled personnel to perform it. Many antibody-based assays have been developed for detection of plant viruses (Sward and Eagling, 1995; Torrance, 1995) and to a lesser extent fungi and bacteria (Dewey and Thornton, 1995; Spire, 1995; Dewey et al., 1997). The tests work well with viruses, but unfortunately it is often difficult to achieve the required specificity for more complex organisms such as bacteria and fungi.

1.3 Nucleic acid based methods for pathogens detection

1.3.1 Probes and polymerase chain reaction (PCR)

Nucleic acid-based methods (using probes and/or PCR) have increasingly been used in to develop diagnostic assays for plant pathogens (Schots et al., 1994; Martin et al., 2000; Ward, 1994). These methods, particularly those based on PCR, are potentially very sensitive and highly specific. However, since detectable DNA may be obtained from dead cells, some workers have opted to start from RNA, since this may more accurately reflect viable pathogen material. This type of assay is also commonly used

in the detection of plant viruses (Waterhouse and Chu, 1995), most of which have RNA genomes. An alternative, isothermic amplification technique, Nucleic Acid Sequence Based Amplification (NASBA), has also been used for amplification from RNA (Bentsink et al., 2002; van Beckhoven et al., 2002; Cook, 2003).

Probes have also been used in the detection of plant pathogens. Probes are single stranded DNA or RNA molecules that have been labelled with a reporter molecule such as a radioactive isotope, an enzyme or a fluorescent dye. They bind to complementary DNA/RNA sequences on the target samples which can then be detected depending on the type of reporter molecule used. Prior to the introduction of PCR, nucleic acid based diagnostics generally involved the use of probes.

In recent years most molecular diagnostic assays have used PCR because of its greater sensitivity, simplicity and speed. However, probes are now increasingly used in conjunction with PCR to produce diagnostic protocols that are more sensitive, more specific or simpler, than PCR alone (Williams et al., 2001; Knoll et al., 2002). PCR is a very sensitive technique, only small amounts of DNA are needed (e.g. from a single spore (Lee and Taylor, 1990). Sometimes nested PCR is used to improve the sensitivity and/or specificity of the assay (Foster et al., 2002). This involves two consecutive PCR reactions, the second one using primers that recognise a region within the PCR product amplified by the first set.

Another approach is to use immunocapture PCR (IC-PCR). Antibodies immobilised on the surface of a microtitre plate or microfuge tube are used to capture the pathogen, which is then detected using PCR. IC-PCR can improve the sensitivity and specificity of the assay and reduce problems with inhibitors in the sample (Jacobi et al., 1998; Shamloul and Hadidi, 1999; Hartung et al., 1996).

1.3.2 Competitive PCR

A method, called competitive PCR, has been developed to quantify target DNA using standard PCR technology (Nicholson et al., 1996, 1998). A single primer pair, in the same reaction, is used to simultaneously amplify target DNA and a known amount of competitor DNA (a piece of DNA with annealing sites for the primers but producing a differently sized product). The amount of target DNA can be estimated by comparing the relative amounts of target and competitor product produced (Gilliland et al., 1990). However, a carefully constructed competitor target is needed for each PCR assay.

1.3.3 Multiplex PCR

Multiplex PCR, a PCR variant which is designed to amplify multiple targets by using multiple primer sets in the same reaction, has been applied in many tests. Multiplex PCR can be used where there is a need to detect several pathogens simultaneously. The assay uses several PCR primers in the same reaction, which saves time and reduces costs. Multiplex PCR assays can be tedious and time consuming to establish requiring lengthy optimization processes (Elnifro et al., 2000). Among the drawbacks of such variant PCR assays are that the sensitivity is decreased enormously and the number of different targets to be amplified in one assay is limited (Bamaga et al., 2003). Moreover, the dynamic range of the target present in the sample to be tested is not always reflected in the outcome of the test. That is targets that are present in very low amounts will most of the time not amplify in contrast to those that are present abundantly. The products from the different targets need to be different sizes to ensure that they can be distinguished from one another and they must all be amplified efficiently using the same PCR conditions (Henegariu et al., 1997). Multiplex PCR has been used to develop an assay that can detect four foliar wheat pathogens (*Septoria tritici, Stagonospora nodorum, Puccinia striiformis* and *Puccinia recondita* (Fraaije

et al., 2001).

1.3.4 Real-time PCR

The process of quantifying target DNA has recently been simplified considerably with the advent of real-time PCR. This method avoids the usual need for post-reaction processing, as the amplified products are detected by a built-in fluorimeter as they accumulate. This is done by using non-specific DNA binding dyes (e.g. Sybr Green) or fluorescent probes that are specific to the target DNA (Wittwer et al., 1997). The principle underlying real-time PCR is that the larger the amount of target DNA present in the sample being tested, the quicker the reaction progresses and enters the exponential phase of amplification. The amount of PCR amplicon produced at each cycle is measured, using the fluorescent dyes or probes, and for each sample tested the cycle threshold (C_t) is calculated. This is the cycle number at which a statistically significant increase in fluorescence is detected. The C_t increases with decreasing amounts of target DNA. A calibration curve relating C_t to known amounts of target DNA is constructed and used to quantify the amount of initial target DNA in an unknown sample. In addition to simplifying quantification, real-time PCR has a number of other advantages over conventional PCR. It is faster and a higher throughput is possible. Post reaction processing is unnecessary, eliminating the risk of carryover contamination. It can be more specific than conventional PCR, if a specific probe is used in addition to the two specific primers. The high specificity of the probes used mean that it is good for detecting SNPs. The real-time PCR offer better multiplexing possibilities, however, multiplexing is still limited by the availability of dyes emitting fluorescence at different wavelengths. Thus, detection of more than few pathogens is currently not possible using these systems.

1.3.5 Microarray

The DNA microarray technology originally designed to study gene expression and generate SNP profiles, is currently a new and emerging pathogen diagnostic technology which in theory, offers a platform for unlimited multiplexing capability. The principle of microarray is the hybridization of fluorescently labelled sequences or targets to their complementary sequences spotted on solid surface, such as glass slides, serving as probes. Tens of thousands of such DNA probes can be spotted in a defined and addressable configuration on the glass slide forming the chip. The unlimited capability for simultaneous detection of pathogens makes microarrays to be an approach with a potential capacity of detecting all relevant pathogens of a specific crop. Development of microarrays for diagnostic applications is a recent history. In plant pathology the method was applied for identifying oomycete, nematode, bacterial and fungal DNA from pure cultures (Fessehaie et al., 2003; Lévesque et al., 1998; Lievens et al., 2005; Uehara et al., 1999). However, for application in practice, pathogens should be detected from environmental samples (plants, soil, etc.). Recently the possibilities of parallel detection of pathogens from such environments were shown (Lievens et al., 2004, 2005; Nicolaisen et al., 2005). In contrast with studies using pure cultures, microarray-mediated analysis from environmental samples presents several challenges that must be addressed (Franke-Whittle et al., 2005; Zhou and Thompson, 2002).

1.3.6 MassTag PCR

Visual disease symptoms are infrequently specific for single pathogens; thus, assays are needed that allow multiple agents to be simultaneously considered. Current multiplex assays employ gel-based formats in which products are distinguished by size, fluorescent reporter dyes that vary in color, or secondary enzyme hybridization assays. Gel-based assays are reported to detect 28 different targets with sensitivities of 2-100

PFU or <1-5 PFU, depending on whether amplification is carried out in a single or nested format, respectively (Fan et al., 1998; Coiras et al., 2004). Fluorescence reporter systems achieve quantitative detection with sensitivity similar to that of nested amplification; however, their capacity to simultaneously query multiple targets is limited to the number of fluorescent emission peaks that can be unequivocally resolved. At present, up to 4 fluorescent reporter dyes can be detected simultaneously (Vet et al., 1999; Verweij et al., 2004). Multiplex detection of up to 9 pathogens has been achieved in hybridization enzyme systems; however, the method requires cumbersome postamplification processing (Grndahl et al., 1999). To address the need for sensitive multiplex assays in diagnostic molecular microbiology, a new PCR platform has been developed (Briese et al., 2005) in which microbial gene targets are coded by a library of 64 distinct Masscode tags (Qiagen Masscode technology, Qiagen, Hilden, Germany). A schematic representation of this approach is shown in Figure 1.1. Microbial nucleic acids (RNA, DNA, or both) are amplified by multiplex reverse transcription (RT)-PCR using primers labeled by a photocleavable link to molecular tags of different molecular weight. After removing unincorporated primers, tags are released by UV irradiation and analyzed by mass spectrometry. The identity of the microbe in the clinical sample is determined by its cognate tags.

Oligonucleotide primers are designed in conserved genomic regions to detect the broadest number of members for a given pathogen species by efficiently amplifying a 50 to 300 bp product. A software program is used to cull sequence information from GenBank, perform multiple alignments, and maximize multiplex performance by selecting primers with uniform melting temperatures and minimal cross-hybridization potential. Primers, synthesized with a 5' C6 spacer and aminohexyl modification, are covalently conjugated by a photocleavable link to Masscode tags (Qiagen Masscode technology) (Kokoris et al., 2000; Lukhtanov et al., 1995). Masscode tags have a modular structure, including a tetrafluorophenyl ester for tag conjugation to primary

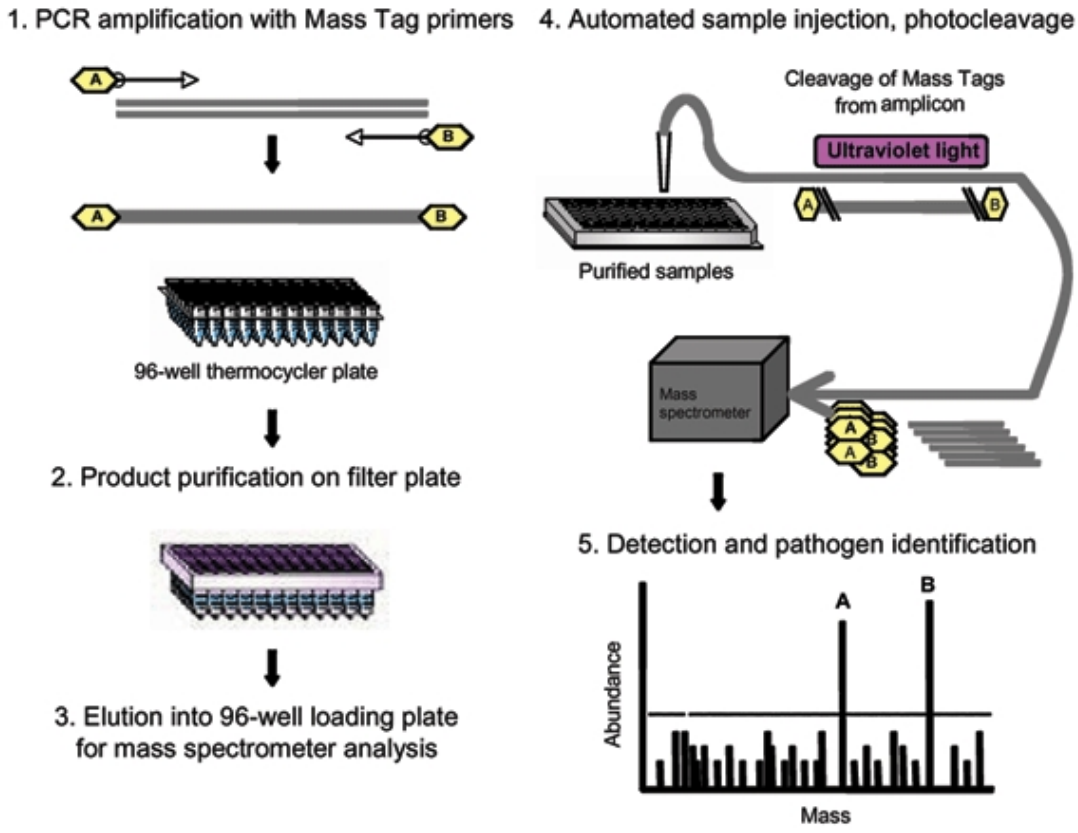

Figure 1.1: Schematic representation of Mass Tag polymerase chain reaction (PCR) (Briese et al., 2005)

amines; an o-nitrobenzyl photolabile linker for photoredox cleavage of the tag from the analyte; a mass spectrometry sensitivity enhancer, which improves the efficiency of atmospheric pressure chemical ionization of the cleaved tag; and a variable mass unit for variation of the cleaved tag mass (Kokoris et al., 2000; Venkatesan and M., 1996; Abdel-Baky et al., 1993; Saha et al., 1993). A library of 64 different tags has been established. Forward and reverse primers in individual primer sets are labeled with distinct molecular weight tags. Thus, amplification of a microbial gene target produces a dual signal that allows assessment of specificity. Mass Tag PCR is a sensitive and specific tool for molecular characterization of microflora. The advantage of Mass Tag PCR is its capacity for multiplex analysis. Although the use of degenerate primers may reduce sensitivity, the limit of multiplexing to detect specific targets will likely be defined by the maximal primer concentration that can be accommodated in a PCR mix. Analysis requires the purification of product from unincorporated primers and mass spectroscopy. Although these steps are now performed manually, and mass spectrometers are not yet widely distributed in developing countries laboratories, the increasing popularity of mass spectrometry in biomedical sciences and the advent of smaller, lower-cost instruments could facilitate wider use and integrated instrumentation. In addition to developing additional pathogen panels, our continuing work is focused on optimizing multiplexing, sensitivity, and throughput. Potential applications include differential diagnosis of infectious diseases, blood product surveillance, forensic microbiology, biodefense and plant disease diagnosis

1.4 Pathogen quantification

Although new, rapid detection and identification technologies are becoming available for various pathogens, pathogen quantification remains to be one of the main challenges in the disease management of many crops. Quantification of a pathogen upon

its detection and identification is an important aspect as it can be used to estimate its potential risk regarding disease development, establishment and spread of inoculum and economic loss. In addition it provides information for well informed disease management decisions. PCR is ideal for detection of small amount of the target but one of its limitations has been quantification. Three PCR variants namely limiting dilution PCR, kinetic PCR and competitive PCR (Hermansson and Lindgren, 2001) have been used for quantitative analysis of DNA. However, all are based on end point measurements of the amount of DNA produced which makes estimation of initial concentration of DNA and quantification rather problematic. Even the microarray technology has limitations with respect to microbial quantification in complex environmental samples due to the fact that microarray hybridization signals could vary depending on target abundance and hybridization efficiency (Wu et al., 2001). In other words, a low abundance target with high genetic similarity to a microarray probe might produce a stronger hybridization signal compared with a higher abundance target that has low similarity to the same microarray probe. Slide to slide variations from a particular probe, same hybridization condition and amount of DNA was also reported (Franke-Whittle et al., 2005) leading to a speculation that variation from the printing of probes from slide to slide may contribute to such differences. On the other hand, efforts are underway towards adding a quantitative aspect in the array technology (Lievens et al., 2005; Rudi et al., 2003)

1.5 Detection specificity and sensitivity in diagnostics

Sensitivity and specificity are numeric measures of effectiveness of a detection system (Peruski et al., 2003; Malorny et al., 2003). Diagnostic specificity is defined as a measure of the degree to which the method is affected by non target components present in a sample, which may result in false positive responses. Diagnostic sensitivity is defined as a measure of the degree to detect the target pathogen in the sample, which may

result in false negative responses (Malorny et al., 2003). Too low sensitivity often leads to false negatives. Thus, a high degree of diagnostic accuracy is characterized by the ability to detect, true and precisely the target micro organism from a sample without interference from non target components. The high degree of sensitivity of molecular methods made presymptomatic detection and quantification of pathogens possible.

One of the most important advantages that molecular based detection has over conventional diagnostic detection methods is the high specificity. That is the ability to distinguish closely related organisms. The specificity of PCR, be it conventional or real-time, depends upon the designing of proper PCR primers that are unique to the target organism. Highly conserved gene regions are often the target for designing primers. Closely related microbial species often differ in a SNP to few bases in such genes. PCR allows detection of such SNPs (Papp et al., 2003). With the advancements in high throughput DNA sequencing more and more genomes of plant pathogens are sequenced and nucleotide sequence data will be available increasing the possibility for designing unique primers and probes for specific detection of pathogens.

PCR is a highly sensitive technology. However, its sensitivity is greatly affected by the presence of inhibitors which prevent or reduce amplification. A wide range of inhibitors are reported (Yang and Rothman, 2004). Although their mode of action is not clear, these inhibitors are believed to interfere with the polymerase activity for amplification of the target DNA. On the other hand, it is worth mentioning that the high sensitivity of PCR also causes one of the limitations of PCR, that is detection sensitivity exceeding threshold levels or clinical significance and false positive results from slight DNA contamination (Yang and Rothman, 2004). Hence, stringent conditions are necessary in conducting the assay and proper negative controls must be included in the test. It is also recommended to have separate dedicated areas for pre- and post PCR handling.

1.6 Justification of the study

Banana, cassava and maize are important food crops in Africa and more so in the East African region. Production of these crops is constrained by bacterial, fungal and viral diseases which have different pathosystems. Currently the most economically devastating diseases for these crops are banana *Xanthomonas* wilt (BXW), gray leaf spot (GLS) of maize and cassava brown streak disease (CBSD) caused by bacterial, fungal and viral pathogens respectively. Early and accurate diagnosis of these pathogens is a crucial component of crop-management system. Therefore there is need for development of rapid and accurate methods for diagnosis. These diseases can be managed most effectively if control measures are introduced at an early stage of disease development. Reliance on symptoms is often not adequate in this regard, since the disease may be well underway when symptoms first appear, and symptom expression can be highly variable. Recent advances in molecular biology and bioinformatics provide exciting opportunities and were applied to the development of rapid, specific, and sensitive tools for the detection of these plant pathogens.

1.7 Objectives of the study

1.7.1 General objective

To use genomics and bioinformatics approaches to develop molecular diagnostic tools for banana *Xanthomonas* wilt, gray leaf spot of maize and cassava brown streak disease.

1.7.2 Specific objectives

1. To use bioinformatics tools to analyze genome sequences of the plant pathogens and identify regions that can be targeted for more specific diagnostic tool development
2. To develop easy to use molecular diagnostics tools to detect the three pathogen diseases
3. To validate the diagnostic tools in the detection of banana *Xanthomonas* wilt, gray leaf spot of maize and cassava brown streak disease in field samples

1.8 Hypothesis

Variability of genome sequences among species can be utilized for the development of species specific PCR based diagnostic tools.

Chapter 2

DEVELOPMENT OF A MOLECULAR DIAGNOSTIC TOOL FOR BANANA *XANTHOMONAS* WILT (BXW)

2.1 Abstract

Traditional methods used to identify banana *Xanthomonas* wilt (BXW) which is caused by the bacterium *Xanthomonas campestris* pv. *musacearum* (*Xcm*) are slow and inconclusive. Molecular detection techniques, including PCR, offer a rapid and accurate method for diagnosis such that disease management decisions can be made quickly. In this study, two bioinformatics approaches were used to design specific primers for the development of a PCR-based diagnostic tool for BXW. The whole genome of an *Xcm* isolate from Kenya was sequenced and eight complete *Xanthomonas* genome sequences publicly available at the Comprehensive Microbial Resource (CMR) website of J. Craig Venter Institute (JCVI) were used for comparative sequence analysis. The first approach involved performing a multiple sequence alignment of the eight *Xanthomonas* genome sequences followed by designing primers targeting intergenic regions of the obtained consensus sequence. Primer pairs designed with this approach were tested on *Xcm* Kenyan strain followed by sequencing of the PCR products. The resulting sequences were further analysed and aligned to create specific primers for *Xcm*. In the second approach, the whole genome of *Xcm* was sequenced and identified unique sequences for *Xcm* as well as plasmid sequences that were used for diagnostic

primer design. In total, 48 *Xcm* specific primers were designed and validated using 52 strains of *Xcm* from East Africa. Seven out of the 48 primers (Xcm12, 35, 36, 38, 44, 47 and 48) showed high specificity and sensitivity. The expected PCR product was amplified with each primer set down to 10^3 CFU/ml, which approximated to a detection limit of about four cells per PCR reaction. *Xcm* was detected by PCR from symptomless banana plant tissue within 1 week following inoculation and before the development of visible symptoms. Primer Xcm38 had the highest sentivity and specificity hence it was validated on field samples. The designed primers provide a sensitive, rapid, reliable and economical tool for routine detection and identification of BXW. Further testing and validation with more strains from around the world is however required to confirm specificity. Our results indicate that primers designed through this approach could be useful in development of a PCR-based diagnostic tool for specific detection of related *Xanthomonas* and other pathogens at genus, species or pathovar levels.

2.2 Introduction

Banana (*Musa* species) is the most important food crop in Uganda, Rwanda and Burundi and is significant as a cash crop and staple food throughout the Great Lakes region of East Africa. After India, the largest producer is Uganda, which is also the largest consumer of bananas. Banana *Xanthomonas* wilt (BXW), caused by the bacterium *Xanthomonas campestris pv. musacearum* (*Xcm*), is a very destructive disease affecting all banana varieties in Africa. It is a gram-negative bacterium from the *Xanthomonas* genus and belongs to the gamma subdivision of the phylum Proteobacteria. *Xanthomonas* is a genus within the Gammaproteobacteria that includes over 20 species and hundreds of pathovars of Gram-negative rod-shaped plant-pathogenic bacteria (Vauterin et al., 1995). This genus includes causative agents of several economically important diseases.

Economic losses due to *Xcm* are estimated at $2 billion over a decade, arising from significant reductions in production (Abele and Pillay, 2007). BXW was first reported about 40 years ago in Ethiopia on Ensete species, which is closely related to banana (Yirgou and Bradbury, 1968; Yirgou and Bradbury, 1974). Outside Ethiopia, it was first reported in Uganda in 2001 (Tushemereirwe et al., 2003, 2004) and subsequently in Democratic Republic of Congo (Ndungo et al., 2004), Rwanda (Reeder et al., 2005), Tanzania, Kenya and Burundi (Carter et al., 2009). BXW endangers the livelihood of millions of farmers in the Great Lakes region of Eastern Africa. The disease is very destructive infecting all banana varieties although field observation and laboratory experiments indicate that beer banana PisangAwak (commonly known as Kayinja) is most susceptible (Tripathi et al., 2008). No banana varieties are known to be resistant to the disease and there is no chemical control effective against it. The disease is characterized by premature ripening of fruits, internal brown discoloration of fruits and vascular tissues, wilting of bracts and male buds and progressive yellowing leading

to complete wilting (Figure 2.1). Spread of BXW is mainly through infected banana planting materials, infected cutting tools, insects and browsing animals. If not detected early and controlled, BXW would spread at a rate of 8 % per annum in cooking banana plantations (Kayobyo et al., 2005), causing an estimated production loss of about 53 % over a 10-year period, equivalent to a reduction from the current 4.5 x 10^9 kg to 2.1 x 10^9 kg/year (Abele and Pillay, 2007). Once the pathogen becomes established, disease control is difficult and eradication impossible.

Figure 2.1: (A) Wilting of the banana plant showing uneven and premature ripening of the fruits, (B) When fruits are cut, the sections show unique yellowish blotches in the flesh fingers and dark brown placental scars and (C) Cross sections of diseased pseudostem reveal yellowish ooze.
Source: International Institute of Tropical Agriculture (IITA)

Accurate, affordable and reliable diagnostic tools are necessary and important for detecting the pathogen, studying disease epidemiology, and developing management measures. Early identification of the disease is also paramount in order to recognize the pathogen, and implement regulations involving control and quarantine. However, BXW has many similar symptoms to other banana bacterial diseases such as Moko, blood and bugtok diseases and fungal diseases including *Fusarium* wilt, which makes specific identification difficult. Currently, BXW is detected in plant material on the

basis of visual examination of symptoms, followed by isolation and culturing on non-selective media (Schaad et al., 2001) or semi-selective media (Tripathi et al., 2007; Mwangi et al., 2007). However, by the time the plant is showing visible symptoms, the pathogen has already established itself and spread to other plants.

For diagnosis of BXW, several pathogen characteristics have been examined, including a variety of biochemical tests, substrate utilization (BIOLOG), and fatty acid methyl esters (FAME) analysis (Aritua et al., 2006). Identification of Xanthomonads by serology (Alvarez et al., 1991), fatty acid analysis (Lazer et al., 1990), metabolic profiling (Chase et al., 1992) or SDS-polyacrylamide gel electrophoresis of proteins (Chagnon et al., 1991) is difficult as these methods are not sensitive or reliable enough to detect low levels of the pathogen. Furthermore, they tend to be time-consuming, tedious, and expensive for large-scale analyses and require further confirmation by pathogenicity tests. Lack of appropriate and reliable diagnostic tools can be a major limitation to epidemiological and ecological studies of the pathogen.

There is need for a more rapid, specific and sensitive diagnostic tool for BXW. Nucleic acid based detection methods for plant pathogens have been shown to hold great promise (Henson and French, 1993). Polymerase chain reaction (PCR) has been widely used in plant disease diagnosis including several bacterial plant diseases (Hartung et al., 1993; Coletta-Filho et al., 2006; Robène-Soustrade et al., 2006a; Trindade et al., 2007). PCR does not require isolation of pure bacterial cultures from infected tissues (Lewis-Ivey et al., 2009) and is often more sensitive, accurate, rapid, less labour intensive and more economical than conventional diagnostics (Degefu, 2008). In previous *Xanthomonas* studies, primers have targeted plasmid genes (Hartung et al., 1993), virulence genes (Mavrodieva et al., 2004), rRNA genes (Cubero and Graham, 2002; Hauben et al., 1997) or xanthan gum biosynthesis genes (Harding et al., 1987). Primer sets have been reported for various *Xanthomonas* species and pathovars. These earlier designed primers greatly improved detection and identification of *Xanthomonas*

pathogens even though some of them were not entirely specific to a particular species or pathovar. In this study a bioinformatics approach was used to design PCR diagnostic primers from unique sequences following the genome sequencing of *Xcm* and region comparison of eight complete genomes of Xanthomonads found in the Comprehensive Microbial Resource (CMR) (http://cmr.jcvi.org/tigr-scripts/CMR/CmrHomePage.cgi). The primers were tested, validated and shown to be specific for the diagnosis of BXW.

2.3 Objectives of the study

2.3.1 General objective

The main objective of this study was to use Bionformatic tools to develop PCR based diagnostic primers for *Xanthomonas campestris* pv. *musacearum* (*Xcm*)

2.3.2 Specific objectives

1. To sequence the whole genome of *Xanthomonas campestris* pv. *musacearum* (*Xcm*) and use bioinformatics tools to identify regions that can be targeted for more specific diagnostic primers

2. To perform comparative genomics of eight *Xanthomonas* genome sequences to identify unique regions to *Xcm* for the design of diagnostic primers

3. To validate the designed primers for detection of BXW in banana plant tissues from the field

2.4 Materials and Methods

2.4.1 Sample collection

A pseudostem sample from a banana plant displaying BXW symptoms was made in December 2006 from Busia county, Western Kenya. Isolation was carried out on nutrient dextrose agar prepared by adding 1 % (w/v) D-glucose to Oxoid nutrient agar before autoclaving at 121 °C for 15 minutes. Cultures were maintained on yeast dextrose chalk (YDC) agar. This medium was prepared by mixing and heat-dissolving 10 g of yeast extract, 20 g of D-glucose and 12 g of Oxoid agar in 1 litre of distilled water before adding 20 g of precipitated chalk. The mixture was then adjusted to pH 7.2 and autoclaved at 121 °C for 15 minutes. One of the isolates, which was coded as BXW-K8 (and given the NCPPB Ref. No. 20702297 after identification by fatty acid profiling at the Central Science Laboratory (currently The Food and Environment Research Agency -FERA , United Kingdom), was used for genomic sequencing.

The other strains of *Xcm* used in this study were isolated from infected banana tissues and the cultures were maintained on YPGA medium (10 g/l yeast extract, 10 g/l peptone, 10 g/l glucose, 15 g/l agar) at 4 °C. DNA of other *Xanthomonas* species., Ralstonia species. and Pseudomonas species. were obtained from the National Collection of Plant Pathogenic Bacteria (NCPPB) maintained at FERA, York, UK. The endophytic and epiphytic bacteria isolated from within and on the outside of banana plants, and identified by the University of Pretoria, were also included in the study (Table 2.1)

2.4.2 DNA extraction from a pure Cultures

The bacteria was inoculated into 50 ml L-Broth medium and allowed to grow on a rotary shaker (150 rpm) overnight at 28 °C. The bacterial cells were then harvested

Table 2.1: Bacterial isolates used in the study

Bacterium	**Host**	**Origin**	**Source**
Xanthomonas campestris pv *musacearum* (*Xcm*-KS)	banana	Kenya	This study
Xanthomonas campestris pv *musacearum* (52 isolates)	banana	Uganda	This study
Xanthomonas campestris pv *campestris* (2 isolates, Ca3, Ca5)	cabbage	Uganda	This study
Xanthomonas campestris pv *campestris* (CF12)	cauliflower	Uganda	This study
Xanthomonas vasicola pv. *Vasculorum* (NCPPB206)	Maize	South Africa	NCPPB
Xanthomonas vasicola pv. *Holcicola* (NCPPB1060)	Sorghum	Ethiopia	NCPPB
Xanthomonas arboricola pv. *Celebensis* (NCPPB1832)	*Musa* sp	New Zealand	NCPPB
Xanthomonas axonopodis pv *phaseoli* (NCPPB 381)	*Phaseolus valgaris*	Canada	NCPPB
Rastonia solanacearum (NCPPB 3969)	Musa (plantain)	venezuela	NCPPB
Xanthomonas campestris pv. *Vesicatoria* (NCPPB 881)	*Lycopersicon esculantum*	Yugoslavia	NCPPB
Xanthomonas axonopodis pv *citri* (NCPPB 409)	Citrus lemon	New Zealand	NCPPB
Xanthomonas translucens pv *translucens* (NCPPB 973)	Barley	USA	NCPPB
Xanthomonas oryzae pv *oryzae* (NCPPB 3002)	Rice	India	NCPPB
Xanthomonas hortorum pv. *Carotae* (NCPPB 230)	*Daucus carota*	USA	NCPPB
Xanthomonas axonopodis pv *phaseoli* (NCPPB 1680)	*Phaseolus valgaris*	Tanzania	NCPPB
Ralstonia eutropha (3440)	-		NCPPB
Pseudomonas syringae pv. *Mellea* (NCPPB 280)	*Nicotiana rustica*	Tanzania	NCPPB
Xanthomonas campestris pv *blepharidis* (NCPPB 1757)	*Blepharis boarhaavifolia*	India	NCPPB
Xanthomonas campestris pv *holcicola* (NCPPB 3162)	Sorghum sp.	India	NCPPB
Xanthomonas campestris pv *mangiferaeindicae* (NCPPB 490)	*Mangifera indica*	India	NCPPB
Xanthomonas translucens pv *graminis* (NCPPB 3709)	*Lolium perenne*	Norway	NCPPB
Xanthomonas vesicatoria (NCPPB 422)	*Lycopersicon escutentum*	New Zealand	NCPPB
Xanthomonas perforans (NCPPB 4322)	*Lycopersicon esculentaum*	USA	NCPPB
Xanthomonas translucens f.sp *cerealis* (NCPPB 1943)	*Hordeum vulgare*	USA	NCPPB
Xanthomonas translucens pv *arrhenatheri* (NCPPB 3229)	*Arrhenatherum elatius*	Switzerland	NCPPB
Xanthomonas arboricola pv *celebensis* (NCPPB 1630)	*Musa* sp.	New Zealand	NCPPB
Xanthomonas campestris pv. *desmodii* (NCPPB 481)	*Desmodium diffusum*	India	NCPPB
Saprophytes used: *Klebsiella* sp. (3), *Citrobacter* sp.	Banana	Uganda	This study

by centrifuging them in falcon tubes at 2700 rpm for half an hour at 4 °C. Genomic DNA was extracted from a pellet of bacterial cells from a 500 ml overnight culture in L-broth. Bacterial cells were inoculated into 5 ml of TE (0.1 mM Tris pH8, 0.1 mM EDTA) buffer to which were added 300 µl of 10 % SDS and 30 µl of 20 mg/ml proteinase K (Promega). After incubating for 1 hour at 37 °C, the lysate was extracted with an equal volume of phenol/chloroform. The upper aqueous phase was transferred to a new tube and the DNA was precipitated with 10 % volume of 3 M sodium acetate and two volumes of absolute ethanol. After washing the DNA three times with 70 % ethanol and collecting by centrigfugation, the pellet was dried and dissolved in 5ml TE. Residual RNA was removed by digestion with 50 µl of 10 mg/ml RNAse A. The DNA pellet was precipitated again and washed with 70 % ethanol. Following this, the pellet was dried on the bench for a few hours and resuspended in 2 ml TE buffer (0.1 mM Tris pH8, 1 mM EDTA).

The yield and quality of the DNA was checked by 0.8 % agarose gel electrophoresis and Nanodrop® ND-1000 spectrophotometry (NanoDrop Technologies, USA) respectively. High quality DNA was sent to the University of Oklahoma Genome Center (http://www.genome.ou.edu/) for whole genome sequencing using the Roche 454 Sequencing platform (Margulies et al., 2005).

2.4.3 Whole genome sequencing and assembly

Libraries of the genomic DNA were prepared using a modification of the protocol recommended by the 454/Roche GS-FLX manufacturer (Margulies et al., 2005) as described elsewhere (Wiley et al., 2009). Briefly, this entailed shearing the genomic DNA to 2 Kbp fragments by nebulization at -20 °C using a pressure of 30 psi for 2.5 minutes (Bodenteich et al., 1992). The sheared DNA fragments were then bound to a MinElute column (Qiagen, Valencia, CA), and washed three times by centrifugation

for 1 minute at 12,000 rpm. The DNA was eluted by rotating the column 180°, incubating for 1 minute with EB buffer and, centrifuging at 13,000 rpm for one minute. The smaller fragments were removed using SPRI (Solid Phase Reversible Immobilization) beads (Agencourt, Beverly, MA) that were prepared by vortexing 35 µl of beads and placing them on magnetic Particle Collector (MPC). After removing the supernatant, the beads were washed twice with 500 µl of 70 % ethanol. Finally, the SPRI beads were air dried and subsequently rehydrated, placed on the MPC again, and the supernatant was removed to a fresh microfuge tube.

The DNA fragments generated were end-repaired by treating them with 20 U of DNA polymerase and 20 U of T4 polynucleotide kinase after adding 5 µl of 10X end repair buffer (500 mM Tris-HCl, pH7.6, 100 mM $MgCl_2$, 10 mM DTT, and 50 µg/ml BSA), 5 µl of 10 mM rATP, 2 µl of 0.25 mM dNTPs, and 5 µl of bovine serum albumin (BSA). The reaction mixture was incubated at 12 °C for 15 minutes followed by an additional 15 minutes at 25 °C. The cleaned DNA sample was further purified on SPRI beads as described above.

Adaptors then were ligated to the cleaned DNA sample by mixing 5 µl of the DNA samples with 1 µl of 454 A and B adaptors, 20 U of DNA ligase, and 1X ligation buffer. The ligation reaction was incubated at 25 °C for 15 minutes. The ligated sample was further purified using SPRI beads as described above.

The adaptor-ligated ds DNA were filled in by treating the sample with 15 U of DNA polymerase, 800 nM dNTP mix, and 1X polymerase buffer. The reaction mixture was incubated at 37 °C for 20 minutes. Purification was performed on SPRI beads as described above. The purified DNA was amplified using the emPCR protocol, loaded onto the 454/Roche GS-FLX sequencer and sequenced using their standard protocol.

Flows from the 454 runs were assembled using Newbler, the 454 assembly software. Three different trimming lengths were used from the 454 to reduce the number of artificial contigs produced due to poor qualities at the end of the contigs. The results from

each of the three assemblies were then assembled using Phrap (developed in the lab of Phil Green at the University of Washington (http://www.phrap.org/phredphrapconsed.html)

2.4.4 Diagnostic primer design

Three bioinformatics approaches were used to select target DNA sequences for *Xcm*-specific primer design. The first approach involved designing primers based on comparative genomics of eight complete *Xanthomonas* genome sequences available on the Comprehensive Microbial Resource (CMR) website of J. Craig Venter Institute(JVCI) (http://cmr.jcvi.org/tigr-scripts/CMR/CmrHomePage.cgi). The list of the *Xanthomonas* species used are shown in Table 2.2. Primers were designed to amplify all the 203 intergenic regions. *Xanthomonas axonopodis* pv.*citri* 306 (*Xac*) was used as a reference genome in designing the intergenic primers. Flanking genes on either side of the 203 intergenic regions of *Xac* were selected for region comparison with similar or orthologous genes in the remaining seven *Xanthomonas* species from the CMR using MegAlign® software of LasergeneDNAstar® by Clustal W algorithm. Consensus sequence generated from the eight aligned sequences was used to design both forward and reverse primers using PrimerSelect program of Lasergene. Generally, primers were designed to be located on the flanking genes of target variable intergenic regions in order to amplify fragments in the range of 300-900bp.

The second approach for primer design involved utilizing the draft genome of the *Xcm* Kenyan strain (*Xcm*-KS) that was sequenced at the University of Oklahoma Genome Center

(http://www.genome.ou.edu/). The draft genome had 609 contigs and these were batched into clusters of 50 contigs. Standard batch BLAST was performed on the CMR and NCBI databases (http://www.ncbi.nlm.nih.gov). Through BLASTn analysis, regions that were found to be unique to *Xcm* were chosen and used to provide specific target

Table 2.2: List of completely sequenced *Xanthomonas* species on CMR used for region comparison.

Organism Name	**Taxon ID**	**Size**	**Sequencing Center**
Xanthomonas axonopodis pv. citri 306	190486	5.17 Mb	FAPESP Univ of Campinas Univ of Sao Paulo
Xanthomonas campestris 8004	314565	5.14 Mb	The Institute of Microbiology, China Guangxi Univ
Xanthomonas campestrispv. campestris	487884	5.07 Mb	Bielefeld University
Xanthomonas campestrispv. campestris ATCC33913	190485	5.07 Mb	FAPESP University of Sao Paulo
Xanthomonas campestrispv. vesicatoria str. 85-10	316273	5.42 Mb	Bielefeld University
Xanthomonas oryzae KACC10331	291331	4.94 Mb	MACROGEN NIAB
Xanthomonas oryzae pv. oryzae MAFF 311018	342109	4.94 Mb	NIAS, Japan
Xanthomonas oryzae pv. oryzae PXO99A	360094	5.24 Mb	J. Craig Venter Institute University of Maryland - CDCB

sequence for diagnostic primer design using PrimerSelect software of Lasergene® 7. The third approach also involved utilizing plasmid DNA sequences for diagnostic primer design. Plasmid DNA sequences were confirmed to be unique to *Xcm* by a BLASTn search performed on the CMR and NCBI using default parameters. Primers were then designed based on the unique sequences.

2.4.5 PCR conditions and sequencing of amplicons

All the designed primers were tested on the *Xcm* DNA for amplification. PCR was performed using AccuPower® PCR PreMix (www.bioneer.com) containing 1U Taq DNA polymerase, 250 of each dNTP, 10 mM Tris-HCl, 40 mM KCl, 1.5 mM $MgCl_2$ and tracking dye for electrophoresis. For all primers, initial optimization of PCR conditions involved PCR with 5 ng *Xcm* DNA, 5 pmole of each of the forward and reverse primers and water to a final volume of 10 µl. PCR amplification was done on a GeneAmp PCR system 9700 thermocycler (Applied Biosystems) using the following

program: 94 °C for 5 minutes, followed by 40 cycles of 94 °C for 20 seconds, 50 °C for 20 seconds, and 72 °C for 1 minute, followed by 72 °C for 10 minutes. The amplified PCR products were electrophoresed on a 2 % agarose gel. Primers showing multiple or no bands were further optimized by increasing or decreasing the annealing temperatures until a clear single amplicon was obtained.

For amplifications with single discrete bands on agarose gel, sequencing reactions were performed using the BIGDYE® Terminator Cycle Sequencing (Applied Biosystems). The reactions were performed in a 10 µl volume comprising of 4 µL of water; 2 µl of diluted PCR products (2.5 ng/µl); 2 µl of reverse or forward primer (5 pmoles/µl); 1.0 µl of BIGDYE and 1.0 µl of 5X sequencing buffer. The sequencing reaction was set for 45 cycles of 96 °C for 10 seconds; 55 °C for 5 seconds; and 60 °C for 4 minutes, followed by a final hold at 4 °C. PCR products were purified by ethanol precipitation according to the manufacturers instructions (Applied Biosystems). The samples were sequenced on an ABI 3730 Genetic Analyzer (Applied Biosystems). The sequencing trace files were aligned into contigs with slight modifications of the default conditions using Sequencher® software version 4.6 (Gene Codes Corporation, USA). All chromatograms and sequences were visually inspected and sequences edited, and only high quality sequences were further used.

2.4.6 PCR and multilplex PCR with specific primers

The annealing temperature for each of the specific primer pair was also optimised. The size of each PCR band was determined so that primers amplifying products of different sizes could be multiplexed. Three primer sets, Xcm27 with Xcm32 and Xcm27 with Xcm33 were multiplexed and they amplified two fragments each. Multiplex PCR was performed using AccuPower® PCR PreMix. For each multiplex reaction, initial optimization of PCR conditions involved PCR with 5 ng *Xcm* DNA, 5 pmole of Xcm27

primer, 1.25 pmole each of Xcm32, Xcm33 or Xcm38 primers and water to a final volume of 10 μl. PCR amplification and detection was done as previously described.

2.4.7 Detection limit for PCR protocol

To test the detection limit of the specific primers, ten-fold serial dilutions of *Xcm* genomic DNA in the range of 10^{-1}ng/μl to 10^{-7}ng/μl were prepared. Multiplex PCR using the optimized primers and conditions above was done with the different concentrations of DNA.

2.5 Results

2.5.1 *Xanthomonas campestris* pv. *musacearum* (*Xcm*) genome sequencing

Xcm-KS draft genome was assembled into 609 contigs with an average size of 8, 086bp and the largest contig sized 214,856bp and a combined total genome size of 4,924,282bp. The genome has an overall GC content of 63.1 %, with base frequencies of 18.5 %, 31.6, 31.5 % and 18.5 % for A, C, G and T respectively.

A total of 5,088 coding features were identified. These constituted: 5,023 protein coding sequences (CDS), 3 non-coding RNAs, 3 riboswitches, 2 ribozymes, 1 ribosomal RNA, 1 messenger RNA and 55 transfer RNAs. This was equivalent to 83.4 % coding potential for *Xcm*-KS genome.

Of the 5,023 CDS, 3,324 (66.18 %) could be assigned biological function(s) while 745 and 954 were hypothetical and conserved hypothetical respectively. Partial coding regions were established to be 679. These CDS had an average length of 822.02bp and GC content of 63.43 % with a standard deviation of up to 3.87 %. The full sequence of the *Xanthomonas campestris* pv. *musacearum* has been deposited at DDBJ/EMBL/-GenBank under the accession AGFQ00000000. This data (including updated annotations) also can be downloaded from URL

http://bioinformatics.iita.org/projects/banana/

In line with Rileys biological role categories (Riley, 1993) genes were classified into 17 classes that ranged from the housekeeping to the bacteria specific processes. Fifty three lipoprotein were also identified and 621 membrane proteins based on the presence of a lipoprotein sequence signature (for the candidate lipoproteins) and Y-score and S-mean score cut off of 0.36 and 0.54 respectively for the case of membrane proteins.

2.5.2 Candidate virulence and pathogenicity factors

The gum genes that encode extracellular polysaccharide (EPS) were identified. Additionally, genes involved in lipopolysacharide biosynthesis, afimbrial and fimbrial biosynthesis and regulation were identified. Of the six bacterial secretory systems, type I, II and III contribute towards virulence (Bttner and Bonas, 2010). In the *Xcm* genome, candidate genes encoding these systems were characterized. The ability to regulate expression of virulence genes, based on external stimuli such as population, oxygen levels, nutrient levels and chemicals, is an adaptive mechanism employed by pathogenic bacteria in colonizing hosts. Two-component system response regulator RpfG (NT01XV0495) and sensory/regulatory protein RpfC (NT01XV0496) which are involved in quorum sensing and virulence genes expression were identified. HrpG (NT01XV1976), which is involved in hrp gene expression regulation, was also identified.

Bacterial pathogens, through T3SS (type III secretory system), translocate a number of effector proteins into the host cell that it utilizes to modulate the hosts biological processes and ultimately establish the disease. A total of 267 effector gene sequences were downloaded from *Xanthomonas*.org resource (http://www.*Xanthomonas*.org/). Of these, 143 were found to have homology with 24 *Xcm*-KS genes based on pairwise sequence comparison. These included hypothetical proteins as well as functionally characterized genes such as HpaF and type III effector HopD1. The homologs were derived from different species with most of them belonging to the *Xanthomonas* genus but also other Proteobacteria.

2.5.3 Plasmid genes

A total of 24 *Xcm*-KS contigs had sequence similarity to 33 different whole plasmid genomes including *Xanthomonas campestris* pv. *vesicatoria* str. 85-10 plasmid

pXCV183. Contigs such as xanth.Contig337, xanth.Contig340 and xanth.Contig112 had sequence similarity of up to 100 % to portions of the plasmids thus, suggesting the presence of plasmid in *Xcm*-KS however this would require confirmatory approach.

2.5.4 Gene ontology (GO) term categorization and Blast2GO analysis

A total of 4552 out of 5023 CDS sequences had BLAST matches. *Xanthomonas campestris* registered the highest number of matchess with over 18000 hits followed by *X. oryzae* that had over 11,000 hits hence providing evidence for the close relation between these species and *Xcm*-KS. Surprisingly, *Xylella fastidiosa* and *Stenotrophomonas maltophilia* had higher number of BLAST hits than some *Xanthomonas* species such as *X.axonopodis*.

Based on GO term (Ashburner et al., 2000) annotation of the 2783 indubitably annotated *Xcm*-KS CDS, majority of the CDS were annotated as involved in catalytic activity (1,763CDS) and binding activity (1,352) in terms of GO molecular function. In terms of GO biological process, a number of CDS were assigned metabolic (1,720) and cellular processes (1,538) with other processes such as adhesion, biological regulation that a critical to the pathogens host infection strategies being represented. 1,299 GO-cellular components were observed with the largest number of the CDS found in subcellular organelles (164) and extracellular regions (16).

2.5.5 Diagnostic primer design

Three bioinformatics approaches were used to select target DNA sequences for *Xcm*-specific primer design: The first approach involved designing primers based on comparative genomics of eight complete *Xanthomonas* genome sequences (Table 2.2) available on the Comprehensive Microbial Resource (CMR) website of J. Craig Venter Institute(JVCI) (http://cmr.jcvi.org/tigr-scripts/CMR/CmrHomePage.cgi). The 5.17 Mbp genome of *Xanthomonas axonopodis* pv. *citri* 306 (*Xac*) was used as the reference genome due to its close taxonomic relatedness to *Xcm*. The total coding regions in *Xac* are 4.4Mbp or 85.71 % are coding bases and the total intergenic regions are 0.74Mbp or 14.29 %. The GC content of *Xac* is 64.77 %. The total number of protein coding genes is 4,312. Total number of intergenic regions in *Xac* was 3545 of which 612 ranged between 300-600bp. From these, 203 regions (5.7 % of intergenic regions) were found suitable for diagnostic primer design. Genome region comparisons on CMR showed that at least four out of the eight *Xanthomonas* genomes had intergenic regions that were flanked on either side by conserved or orthologous genes in synteny with the reference *Xac* genome (Figure 2.2). To reduce the possibility of primers amplifying on multiple sites, paralogous gene copies were not used to design intergenic primers.

The second approach for primer design involved utilizing the draft genome of *Xcm*-KS. DNA sequences were randomly selected from the draft genome and a BLASTn search performed using default parameters on the CMR and NCBI databases. Regions unique to *Xcm* were selected as target sequences for diagnostic primer design. PrimerSelect® software of Lasergene 7 was used to design these primers.

The third approach involved utilizing plasmid DNA sequences for diagnostic primer design. Plasmid DNA sequences were confirmed to be unique to *Xcm* by a BLASTn search performed on the CMR and NCBI using default parameters. Primers were then designed based on the unique sequences.

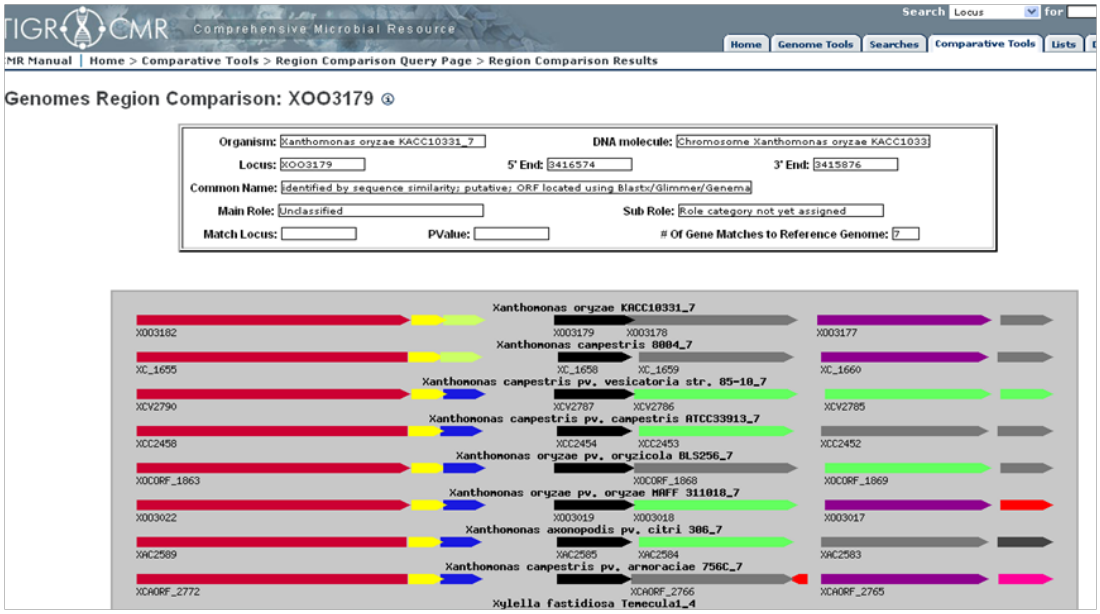

Figure 2.2: Genome region comparison of *Xanthomonas oryzae* GumB gene (locus XOO3179) across the 8 sequenced *Xathomonas* species at CMR

2.5.6 Specific primer pesign

Newly sequenced PCR products from the intergenic regions of *Xcm* were aligned with the other eight *Xanthomonas* corresponding intergenic sequences found at the CMR using MegAlign software by ClustalW algorithm. Specific primers were designed to either be located on the intergenic region unique to *Xcm* or that the 3' end of the primer was positioned on a region that is unique to *Xcm* (Figure 2.3). BLASTn analysis using default parameters was performed on sequences whose primers were designed from unique genes/sequences of *Xcm*. A BLASTn search was also performed on the newly designed specific primers themselves to ensure that there was no alternate primer binding sites in other genomes other than *Xcm*. If the newly designed primer was homologous to other organisms sequence, other than *Xcm*, an alternative specific primer was redesigned based on the above criteria. Based on these approaches, a total of 48 BXW specific primers were designed (Tables 2.3,2.4 and 2.5). Out of these, 29 primer pairs were designed from intergenic regions or the 3' end of the primer being positioned on

a region that is unique to *Xcm* (Tables 2.3), six primer pairs (Tables 2.4) from unique plasmid DNA sequences, while13 primer pairs were designed from unique sequences to *Xcm* based on BLASTn analysis on CMR and NCBI (Tables 2.5) as described in the Primer Design section above.

Figure 2.3: A multiple sequence alignment of newly sequenced fragment of *Xcm* (Xvm69) and eight other *Xanthomonas* (XAC3910 A-H) and the position of the specific forward primer.

Table 2.3: *Xcm* Primers designed by comparative genomics

No	Name	Forward Primer (5'-3')	Reverse Primer (5'-3')
1	Xcm1	CAAGCCGGGCCAGCGAAGTGC	AAAGGCGGCCGTGCTGGAAAGAGA
2	Xcm2	GCATGTGGCGGTGGTCAGTGTCC	CGGCGGCGATCCAGCAAATG
3	Xcm3	GTTGGCGCCGGATCTGTTGTTTG	GACCGGCACTGCGCTCACGAA
4	Xcm4	GTCGATTTGCCGGCGTTGGTGTAA	CGGTGCGGATCTGGTGTTGGTCA
5	Xcm5	CACCGACGGCCTGCAACACTTCAT	CGCGGATTGCATTGGTCGGTTACA
6	Xcm6	GAAGCCGGCGACGACGAAGAGAA	TTGCATTAAAGCGCCGTCCATAGC
7	Xcm7	CCCGCCCCCGTCTGCTGTGC	GCCCTGGCCGCCGTAGATTTCCTG
8	Xcm8	GCGGGTCCTGCATGGCGTTTATTC	CAAGGGCGGCGTGGTGTTCG
9	Xcm9	CCGACCCCCACCCTCAGCAACA	GCGCGCCGGCCTACAACGAC
10	Xcm10	ACGGCGCATATCGGCACATCCC	AGCGGCGCAGATCGTCGGGA
11	Xcm11	CGAAACTGCCGCCGACGATGTAA	CAGCGCCGCGGCTTTGGACTT
12	Xcm12	GCCGGCGTGCGCAACTATCTG	GCCATCCGCAAACAATCGCAACCT
13	Xcm13	TTGGAGCCGCGCCAGAACA	CATCGCGGCGCAGAGGTCTAT
14	Xcm14	GCGCCGGCCGTTGAGGTGG	GCAGCGGCGGTGGGTGTCATC
15	Xcm15	GGGATCGGGCGGGTCATTGAG	TGTGTCCGAAAAGCAGGCGTCAAC
16	Xcm16	CAGCCTCATCCGCGCCAGTGTTC	CCGCCATACCCACCGCAAACCA
17	Xcm17	CGCGGCGCTTGCAGTGATTGAC	GCACTGCGCAGACCGAAGACACG
18	Xcm18	CCGCCAGCTCCGACCCCTTGTT	CACGGCGCTCGGCTTCCAGC
19	Xcm19	CGCCCGGCAACTGCAACATCACAC	CCGCTCCGAGGGCAACACCAA
20	Xcm20	GAGATTGGCCGCCTGCAACAC	CCACCCGCACCTAGCTGACAACTG
21	Xcm21	GCCGCGCCGTAGAATAGG	GCCGAGCGTTGCGTAGTAATC
22	Xcm22	CCGGAGCGAGGGCACACT	AGACGTGCTTCTGGGCTGGC
23	Xcm23	CCTTTGGGGTAACGGACATTTTGA	AAGCCGGGATTTGGGATTGGT
24	Xcm24	TGTGCCGGTGGGTTGAGTGAAGTG	GCCCTGCGGAAGCGGATAAGGA
25	Xcm25	TGCCAAAACCATGCTGCCACTCGT	CATCGGCGTCCTGTCGGTTGTCAT
26	Xcm26	GCGTCGGCCTGTTGCTGATGG	TCCGTCCGGCGAGTGTTGTTGTTT
27	Xcm27	AACACCCATGCGCAATTTCTCTGA	CGTCGCTGGGCCGCTTCTC
28	Xcm46	TGCGGCAGGCGGGGTTTCA	GCCGATGCGCTGCCTGTCCTG
29	Xcm47	GCTGCGTAATGGGCGAGATGATGC	GCTGCCGCCGGTTTGGTTTGT

Table 2.4: *Xcm* Primers designed from *Xcm* sequenced Genome

No	Name	Forward Primer (5'-3')	Reverse Primer (5'-3')
1	Xcm34	GCATTTTGCGGGTGGGAAACATCA	TTTGAGCGGAGGGCGAACCAGAAT
2	Xcm35	GAGCGCGAGGAAACGGGGAAGT	TTGTGTTCGCCCAACCCTCTCAGT
3	Xcm36	GCTTCGGCGGAGGCGTGCTAAT	TCGGCCGGGCGAGAACTTGAA
4	Xcm37	CGCCCGGCCGAAAACGACC	TGCACGGGGATCATGCCTAACCTG
5	Xcm38	CCGCCGGTCGCAATGTGGGTAAT	CAGCGGCGCCGGTGTATTGAGTG
6	Xcm39	CGCGCCGCCGTCCACCTG	AGGCCGGCACACTCGACAACACAG
7	Xcm40	GGGCGAGCCGGGTTGGAGTTGC	TGGGCACCGGGCGCATTTACG
8	Xcm41	CGGCGGCGAGCGAGAGCAGATA	CAAAGGGGCCAGCATGAGCACACC
9	Xcm42	CGGTGGTGGCTGCGCTGTAGGTTC	GCTGTGGCGGCGCTGCTGACTA
10	Xcm43	GGATGCGGCGGCCACGGTC	TGCGCCGCCGACGTTTTGGAT
11	Xcm44	AATAGCCCGGGTGATTGTCC	AGCCGGCAGCTACGATGAG
12	Xcm45	CCGGCGGCGATGACAGC	AAGCGGAGCGACCCACCTTC
13	Xcm48	CCCGCGATCACTTCCAACAAACAC	GCTCAATCGCCGGAGGGAGAATC

Table 2.5: *Xcm* Primers designed from plasmid sequences

No	Name	Forward Primer (5'-3')	Reverse Primer (5'-3')
1	Xcm28	GCGGCACCCTCTCCCCAAAAACA	CGACCGGGAACGCTCTACGCAACA
2	Xcm29	GGGCGCCCGCACTACTGGTCAAG	GGGGCTGCGCGGTGTGGTCAT
3	Xcm30	CGCCGCTTAATAGACGCAGTGTTG	GCGCGGGTGTGGATATGGATGAT
4	Xcm31	GCCGCTGGGAAGGGAGGTCAAAGT	AAGTCGCCGCCCCTGTCATCCAC
5	Xcm32	CGGGGCCAGTTGGTCGCTTTCTC	ACTGCGACCGGCCCTCACCTGTT
6	Xcm33	AGCCGCGCCGTGCCTCTTC	ACCAGCCGCTTTTCCGTCTCCAC

92 % of the specific primer pairs (44 out of 48) gave single discrete amplification products with *Xcm* DNA. Three primer pairs (Xcm21,Xcm23 and Xcm34) did not amplify *Xcm* DNA while one primer pair (Xcm18) produced multiple bands (Figure 2.4). The specificity of the 48 primer pairs was assessed on DNA extracted from a collection of pure bacterial cultures: 52 *Xcm* strains from different geographic regions in Uganda, 23 strains of other *Xanthomonas* species, 2 strains of *Ralstonia* spp., 2 strains of *Pseudomonas* spp., and 11 saprotrophic bacterial strains isolated from healthy banana plants. There were no amplification products obtained with unrelated phytopathogenic bacteria or endophytes/epiphytes bacteria from banana using these primers (Table **??**).

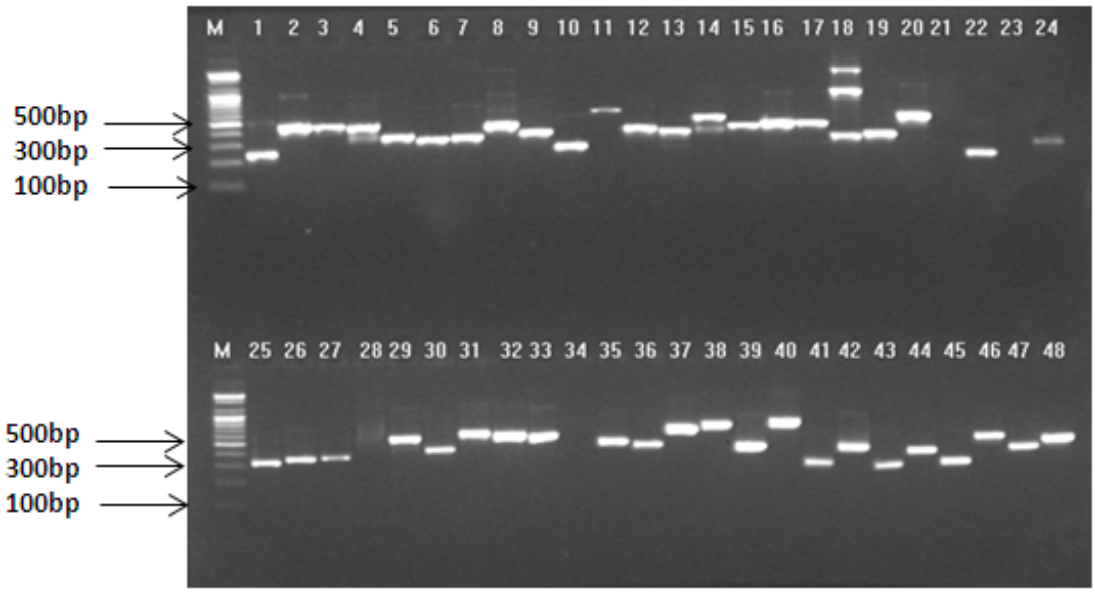

Figure 2.4: PCR products of the *Xcm* specific primers Xcm1 to Xcm48 for banana *Xanthomonas* wilt. M-100 bp molecular ladder

2.5.7 Multiplex PCR

A multiplex PCR approach was employed using three sets of primers, each set targeting to amplifying products of different sizes, to confirm that the diagnostic PCR was accurate in amplifying *Xcm*. Four primer pairs Xcm27 (amplifying a product of 250bp), Xcm32 (amplifying a product of 550bp), Xcm33 (amplifying a product of 550bp) and Xcm38 (amplifying a product of 600bp) were tested for multiplex PCR. Primer Xcm27 was multiplexed with Xcm32 and Xcm27 with Xcm33 based on the product size difference. These were then tested for DNA amount detection limit using 10-fold serial dilutions. Both primer sets Xcm27/32 and Xcm27/33 multiplex PCR gave PCR products from low DNA concentrations of 10^{-5}ng (Figure 2.5).

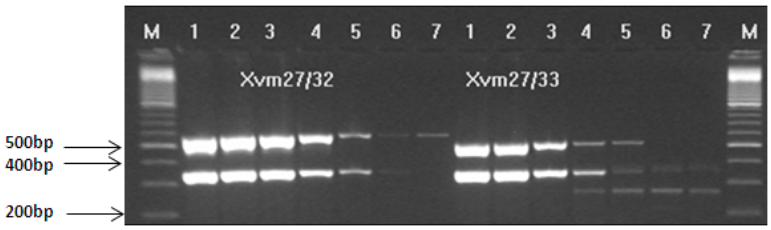

Figure 2.5: Multiplex PCR with two sets of primers showing detection limit on serial dilution of *Xcm* DNA in the range of 10^{-1}ng to10^{-7}ng. M-100 bp molecular ladder

Chapter 3

DEVELOPMENT OF DNA MARKERS SPECIFIC FOR GREY LEAF SPOT (GLS) DISEASE OF MAIZE

3.1 Abstract

A comparative genomics approach was used to develop specific primers for grey leaf spot (GLS) disease of maize caused by *Cercospora zeae maydis* (Czm) group I and *Cercospora zeina* group II. Three general approaches were used to select sequences for development of diagnostic primers for these pathogens. The first approach involved utilizing known genes in *Cercospora nicotianae* to target intergenic regions which could have sequence variation among *Cercospora* species. Out of 28 primer pairs designed by this approach, 4 primer pairs were confirmed to be specific for the two *Cercospora* groups. The second approach involved identifying unique EST sequences from the NCBI GeneBank through BLASTn analysis. Sixteen primer pairs were designed with this approach and 11 of them were specific for the two *Cercospora* groups. The third strategy was to utilize *Malazy* gene, a *gypsy*-like transposable element that is differentially expressed in the two *Cercospora* groups, i.e. PCR amplification of *pol* gene in Group I isolates (*C. zeae-maydis*) and no amplification with group II (*C zeina*). Six diagnostic primer pairs were designed with this approach but 4 primers (Czm1 - Czm4) gave good PCR products. In total 19 primer pairs were confirmed for use in PCR diagnostic of GLS of maize, 4 based on the intergenic region, 11 from

unique EST sequences and 4 from the malazy gene. A multiplex PCR assay was developed that distinguished *C. zeae maydis* and *C. zeina*. Multiplex primer sets Czm47/49 and Czm47/51 produced two bands with C. *zeina* and one band with C. *zeae-maydis*. One primer in particular, Czm9, was found to have the highest sensitivity for detection of GLS of maize disease where it was able to amplify low DNA concentration of upto 10^{-3} ng. Four specific primers pairs (Czm2, Czm39, Czm47 and Czm49) were tested for PCR amplification of 35 field samples from plants showing symptoms of GLS of maize collected from farmers fields and experimental stations from research centres in five maize growing regions in Kenya. The primers confirmed the presence of *Czm* and *Cz* in the fields. These primers need to be tested further with more *Cercospora* strains from other parts of the world to select primers for specific identification of each strain. The results from this study indicate that primers designed through these approaches could be useful in development of a PCR-based diagnostic tool for specific diagnosis of GLS of maize as well as related *Cercospora* and other pathogens at genus, species or pathovar levels.

3.2 Introduction

Maize (*Zea mays* L.), or corn, is the most important cereal crop in sub-Saharan Africa and together with rice and wheat, one of the three most important cereal crops in the world. In sub-Saharan Africa, maize is a staple food for an estimated 50 % of the population. Maize is high yielding, easy to process, readily digested, and cheaper than other cereals. It is also a versatile crop; growing across a range of agroecological zones. Despite the importance of maize, substantial economic losses from GLS of maize occur annually throughout Africa and the Western Hemisphere of USA (Ward et al., 1999; Crous et al., 2006). By the mid-1990s, GLS of maize caused significant losses in some parts of Africa and throughout the corn belt of the U.S. and it is now the most devastating foliar pathogen of maize in much of the world (Ward et al., 1999). Management of GLS of maize disease is especially difficult because commercial hybrids of maize lack effective resistance to gray leaf spot (Ward et al., 1999) and the fungus can survive between growing seasons in plant debris (de Nazareno et al., 1992).

The first incidence of grey leaf spot (GLS) of maize disease in Kenya was reported in 1995 (National Agricultural Research Laboratories (NARL) disease records, unpublished). Yield losses caused by GLS of maize were estimated to be in the range of 30-50 % in Kenya (Muriithi and Gathama, 1998) and 30-60 % in South Africa (Ward et al., 1997). Total yield loss may also occur in the absence of control measures if favourable environmental conditions prevail following successful early infection in maize crop fields (Saghai-Maroof et al., 1993). This relatively new disease is, therefore, a serious threat to the food security in Kenya, where the crop is the leading staple food for about 90 % of the population (Anonymous, 1989; Authority, 2005).

GLS of maize is a foliar disease (Figure 3.1) that is caused by two genetically distinct but morphologically similar sibling fungal species, referred to as *Cercospora zeae maydis* (Czm) group I and *Cercospora zeina* group II (Dunkle and Levy, 2000; Wang

et al., 1998) though *C.sorghi* has also been reported from maize (Crous et al., 2006). First discovered in 1924 in Illinois (Tehon and Daniels, 1925), *C. zeae-maydis* did not become an important pathogen of maize until the 1980s. *C. zeae-maydis* is dominant in the U.S.A. and occurs elsewhere in the world, while *C. zeina* occurs in Africa, the U.S.A. and possibly elsewhere (Wang et al., 1998; Dunkle and Levy, 2000; Goodwin et al., 2001).

The fungal genus *Cercospora* represents a large and diverse group of plant pathogens that are distributed worldwide and infect numerous host species. Individual species of *Cercospora* are usually host specific, but collectively they infect remarkably diverse hosts. More than 3,000 species of *Cercospora* have been named (Pollack, 1987), and they often are classified according to host association, e.g., *C. beticola* infects sugar beet (*Beta vulgaris*), *C. oryzae* infects rice (*Oryza sativa*), and *C. sorghi* infects sorghum (*Sorghum bicolor*).

Most plant-pathogenic species of *Cercospora* enter host leaves through stomata, a process facilitated in part by the ability of elongating germ tubes to sense nearby stomata and reorient their direction of growth accordingly (Beckman and Payne, 1982). Upon reaching stomata, germ tubes differentiate into multilobed infection structures similar to appressoria, from which infectious hyphae penetrate mesophyll tissues. After a period of colonization, the fungus presumably adopts a necrotrophic growth habit, leading to the formation of expanding, necrotic lesions that coalesce in severe outbreaks, leading to a significant reduction in photosynthetic tissue, defoliation, and potentially premature death of the host plant. Reproduction and formation of secondary inocula occur in colonized tissue through the production of asexual spores (conidia) that infect neighboring plants after dispersal by wind and/or rain splash. Many diseases caused by *Cercospora* species occur periodically throughout the world as epidemics singly or as components of disease complexes (Berger, 1973; de Nazareno et al., 1993; Nelson and Campbell, 1993) and for crops such as sugar beet, are major limitations for production

Figure 3.1: Grey Leaf Spot Disease of Maize. (A) Healthy growing maize leaves, (B) coalesced gray lesions on maturing leaf and (C) Blighting of leaves and entire plants due to severe gray leaf spot pressure.
Source: Tamra A. Jackson, Extension Plant Pathologist, University of NebraskaLincoln

(Cooke and Scott, 1993). Additionally, the possibility that *Cercospora* pathogens influence the distribution of plant species in natural ecosystems is a plausible but largely unexplored hypothesis. Exactly why *C. zeae-maydis* and *C. Zeina* have ascended so rapidly as pathogens of maize during the past two decades is not known, but speculation has linked the phenomenon to global climate change, the emergence of more virulent strains, and the increased practice of conservation tillage in maize production (Ward et al., 1999; Latterell and Rossi, 1983; Rosenzweig et al., 2000).

Accurate, affordable and reliable diagnostic tools are necessary and important for detecting these pathogens, studying disease epidemiology, and developing management measures. Early identification of the disease is also paramount in order to recognize the pathogen, and implement regulations involving control and quarantine.

3.3 Objectives of the study

3.3.1 General objective

The main objective of this study was to develop PCR based diagnostic primers for the specific detection of grey leaf spot disease of maize

3.3.2 Specific objectives

1. To utilize the variation in intergenic regions of *C. zeae-maydis* and *C. Zeina* and Malazy gene to design specific primers that can distinguish the two Cercospora species causing GLS of maize

2. To identifying unique available EST sequences to the two Cercospora species for designing diagnostic primers for GLS of maize

3. To validate the designed primers for detection of GLS in maize plant tissues from the field

3.4 Materials and Methods

3.4.1 Fungal strain and culture conditions

Isolates of *Cercospora zeae maydis* used in this study were obtained from naturally infected maize samples collected from farmer's fields and experimental fields from research centres in five maize growing regions of Kenya (Busia, Kakamega, Trans-Nzoia, Kitale and Coastal Kenya), representing varying agro-ecological conditions (Table 3.1). Another set of samples were obtained from USA (sample 36), Zambia (sample 37) and Kenya Agricultural Research Institute (KARI - Kakamega, samples 38 and 39). Two field surveys were undertaken in Kenya during 2007/2008 for samples 1-35. Isolations were made from lesions of naturally infected maize leaf tissues showing fungal sporulations. For non-sporulating lesions,the fungus was induced to sporulate by incubating the infected tissues. Infected leaf tissues showing distinct grey leaf spot symptoms were placed in moist chambers prepared using glass boxes and filter papers. The infected leaf tissues were incubated as described by Kinyua et al. (2010).

Water agar plates were prepared by adding 20 g of agar to 1L of water. The mixture was then sterilized in an autoclave at 121 °C at 15 psi for 15 minutes and approximately 20 ml was dispensed into each sterile Petri-dish. V-8 juice agar plates were prepared by adding 200 ml of V-8 juice agar to 20 g of agar. The mixture was topped up to 1L by adding 800 ml of sterile distilled water and the same sterilizing and dispensing procedure as for water agar was followed.

A porcelain tile was swabbed with 70 % ethanol and was laid on a bench surface. A sterile glass slide was laid on the tile and a few drops of sterile distilled water at the center of the slide. Infected materials showing fungal sporulation was placed on the stage of an ERMA binocular dissecting microscope and conidia on the synnemata

were touched with the tip of a fine moistened mounted needle without touching the host material. The conidia collected were transferred onto the water droplet on the slide, and stirred with a wire loop to form a spore suspension. A film of the spore suspension was captured by withdrawing the wire loop after stirring and streaked across the surface of water agar plates, using four strokes to distribute the spores. The plates were incubated at 25 °C under cool fluorescent light for 14 days.The resulting single spore cultures were separately bulked further using V-8 juice agar while maintaining their identity. Cultures were maintained on V8 agar in constant darkness to provide conidia for inoculations. The fungus was then grown at 24 °C on V8 agar, 0.2X potato dextrose agar (PDA), or 0.2X PDA supplemented with 10 mM ammonium phosphate. Cultures grown in constant light received 8-10 μ E m-2s-1 of illumination. To facilitate collection of fungal tissue from agar plates, conidial suspensions were inoculated onto cellophane membranes placed on the surface of the medium. Samples 38 and 39 were not cultured.

3.4.2 DNA isolation

DNA was extracted from the 39 samples (Table 3.1) using modified cetyl trimethyl ammonium bromide (CTAB) method (Doyle, 1990). For samples 1-37 (Table 3.1), mycelium was harvested and placed in 2 ml tubes. One milli litre of absolute ethanol was added and the tubes centrifuged at 1200 rpm for 10 minutes. The supernatant was discarded and the pellet transferred in clean morters. The mycelium or infected leaf (non cultured samples 38 and 39) from the field was ground into fine powder using a mortar and paste in the liquid nitrogen. The fine powder was transferred into 1.5 ml microfuge tubes and 500 μl of CTAB buffer added (2 % w/v CTAB, 1.4 M NaCl, 0.3 % 2-mercaptoethanol, 20 mM EDTA, 100 mM Tris-HCl PH 8.0). The tubes were incubated for 30 minutes at 65 °C in a recirculating water bath. Two hundrend and

Table 3.1: Grey leaf spot symptomatic samples from maize

Sample ID	Sample Name	Source (District)
1	BU36	Busia
2	BU-57a	Busia
3	BU-57b	Busia
4	KK-66	Kakamega
5	KK-79	Kakamega
6	KK-81	Kakamega
7	KK-83	Kakamega
8	KK-88a	Kakamega
9	KK-88b	Kakamega
10	KK-90	Kakamega
11	KI-01	Kitale
12	KI-05	Kitale
13	KI-20	Kitale
14	KI-22	Kitale
15	KI-24ai	Kitale
16	KI-24aii	Kitale
17	KI-24b	Kitale
18	KI-27a	Kitale
19	KI-27	Kitale
20	TN-143	Transnzoia
21	TN-144	Transnzoia
22	Czm-A	Coast
23	Czm-D1	Coast
24	Czm-D2	Coast
25	Czm-E1	Coast
26	Czm-E2	Coast
27	Czm-N	Coast
28	Czm-W1	Coast
29	Czm-W2	Coast
30	Czm-16W1	Coast
31	931a	Kakamega
32	931b	Kakamega
33	932	Kakamega
34	1031	Kakamega
35	1032	Kakamega
36	SCOH	USA
37	OYPA	Zambia
38	Infected Leaf	KARI Kakamega
39	Infected Leaf	KARI Kakamega

fifty microlitre of chloroform : Iso amyl alcohol (24:1) was added and the solution mixed by gentle inversion for 10 minutes. The tubes were centrifuged at 13000 rpm for 5 minutes. The upper aqueous phase was transferred to clean microfuge tubes and a second Chloroform : Iso Amyl Alcohol (24:1) wash was done. The upper aqueous phase was transferred to clean microfuge tubes and 2 volumes of isopropanol added. The tubes were then mixed by gentle 4-5 inversions and kept at -20 °C for 1 hour. The tubes were centrifuged at 14000 rpm for 10 minutes and the supernatant poured off and pellet was left to dry on air. Two hundrend and fifty microlitre of TE buffer (10 mM Tris, 1 mM EDTA) was added to dissolve the pellet. Half volume of 3 M sodium acetate was added followed by 2 volumes of chilled absolute ethanol. This was followed by gentle mixing until DNA precipitated. The tubes were centrifuged at 14000 rpm for 5 minutes to pellet out the DNA. The DNA pellet was washed thrice by adding 70 % ethanol and centrifuging the pellet at 14000 rpm for 5 minutes. Following this, the pellet was dried on air for a few hours and resuspended in 200 µl TE buffer (0.1 mM Tris, 1 mM EDTA). The DNA quantity and quality was checked by a Nanodrop® ND-1000 spectrophotometry (NanoDrop Technologies, USA). To assess the quality of the extracted DNA, 2 µl of each sample was run on a 0.8 % agarose gel electrophoresis in 1X TBE (Tris, boric acid, EDTA) buffer at 100V for 1 hour. The gel was pre-stained with 5 µl of 1 µg/ml ethidium bromide, and bands observed and photographed using an ultraviolet transilluminator (UVIpro software version 12.4).

3.4.3 Primer design

Three general approaches were used to select sequences for development of diagnostic primers for *Cercospora zeae-maydis*. The first approach involved utilizing known genes in *Cercospora nicotianae* (Chen et al., 2007) as a reference to conserved genes to target intergenic regions which could have sequence variation among *Cercospora*

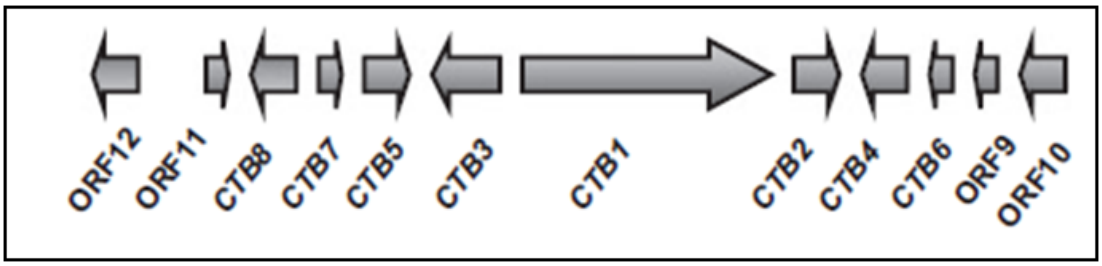

Figure 3.2: Transcription map of the cercosporin toxin biosynthesis (CTB) cluster genes in *Cercospora nicotianae* and the adjacent open reading frames (ORFs) (Chen et al., 2007).

species. The strategy involved exploiting the available cercosporin toxin biosynthesis (CTB) cluster of genes in *Cercospora nicotianae* and the adjacent open reading frames (ORFs), assuming they could be conserved or closely similar to other *Cercospora* species (Figure 3.2). BLASTn analysis was first done to confirm that these genes were only present in *Cercospora* species. Based on the order of the gene clusters (Chen et al., 2007), primer pairs were designed mainly to target intergenic/ non-coding regions and also the full-length genes in *Cercospora zeae-maydis*. In some cases, more than one primer targeting a different part of the gene was designed. Amplified products from these primers were sequenced as described below and BLASTn analysis was done for these newly sequenced fragments of Czm. Specific primers were further designed based on sequences that had no hits in the NCBI GeneBank. A standard BLASTn search was also performed on the newly designed specific primers to ensure that there were no alternate primer binding sites in other genomes other than Czm. If the newly designed primer was homologous to other organisms sequences, alternative specific primers was redesigned based on the above criteria.

The second approach involved identifying unique available EST sequences from the NCBI GeneBank through BLASTn analysis. There were a total of 19,802 EST sequences of *Cercospora zeae-maydis* available in the GeneBank. These ESTs were randomly selected and standard nucleotide BLASTn performed excluding *Cercospora*

zeae-maydis on the organism option in the search tool. This was intended to exclude the query sequence on having hits on itself and other similar sequences of *Cercospora zeae-maydis*. ESTs found to have no hits were selected on the basis of being unique to the two Cercospora species and thus utilized for primer design.

The third strategy was to utilize Malazy gene (AY170476) (Shim and Dunkle, 2005), a *gypsy*-like transposable element that is differentially expressed in the two *Cercospora* groups, i.e. PCR amplification of POL gene in Group I isolates (*Cercospora zeae-maydis*) and no amplification with group II (C. *zeina*). Malazy, a gypsy-like transposable element that is differentially expressed in the two *Cercospora* groups, was utilized for designing diagnostic primers. From previous published reports, there are primers that amplifty a 1.9kb PCR product of the POL gene in Group I isolates (*Cercospora zeae-maydis*) and no amplification with group II (C. *zeina*). PrimerSelect software of Lasergene7 DNASTAR® was used to design all primers. The primers designed using the three strategies are shown in Tables 3.2, 3.3 and 3.4.

3.4.4 PCR amplification

All primer pairs designed were used to test PCR amplification of extracted *Cercospora zeae maydis* DNA. PCR was performed using AccuPower® PCR PreMix (www.bioneer.com) containing 1U Taq DNA polymerase, 250 μM of each dNTP, 10 mM Tris-HCl, 40 mM KCl, 1.5 mM $MgCl_2$ and stabilizer and tracking dye for electrophoresis. For all primers, initial optimization of PCR conditions involved adding into the PreMix 1.0 μl of 5 ng/μl DNA, 1.0 μl of 5 pmole/μl of F/R primer and 8.0 μl of H_2O. PCR amplification was done on a GeneAmp PCR system 9700 thermocycler (PerkinElmer, Wellesley, Mass) using the following program: Initial denaturation of 94 °C for 5 minutes followed by 40 cycles of denaturation of 94 °C for 20 seconds, primer annealing temperature for 20 seconds, and extension of 72 °C for 1 minute and a final extension

at 72 °C for 10 minutes. The amplified PCR products were electrophoresed on a 2 % pre-stained ethidium bromide agarose gel to confirm amplification. Primers showing multiple or no bands were further optimized by increasing or decreasing the annealing temperatures until a clear single band was visualized.

3.4.5 DNA sequencing

PCR products which had single distinct bands on agarose gel were sequenced. Sequencing reactions were performed using the BIGDYE® Terminator Cycle Sequencing (Applied Biosystems). The reactions were performed in a 10 µl volume comprising of 2 µl of diluted PCR products (2.5 ng/µl); 2 µl of reverse and forward primer (5 pmoles/µl) done separately; 1.0 µl of BIGDYE®; 1.0 µl of 5X sequencing buffer and 4 µl of H_2O all at manufacturers recommendations. The sequencing reaction was set for 45 cycles using the program: denaturation at 96 °C for 10 seconds; annealing temperature for 5 seconds; extension of 60 °C for 4 minutes and final hold at 4 °C .

Purification of PCR products was done by the ethanol precipitation method. One microlitre of 125 mM EDTA was added to each well in the sample plate followed by 1 µl of 3 M sodium acetate. Twent five microlitresl of chilled absolute ethanol was added and the plate incubated at -20 °C for 1 hour. The plate was then centrifuged at 3,000 rpm for 30 minutes at 4 °C and then inverted and centrifuged again at 200 rpm for 2 minutes to remove the supernatant. 100 µl of 70 % ethanol was added and the plate centrifuged at 3,000 rpm for 15 minutes and inverted again and centrifuged for 2 minutes at 200 rpm. The plate was placed in a thermocycler and heated at 90 °C for 1 minute with the thermocycler's lid open to remove all residual ethanol. The sequenced products were re-suspended in 10 µl HiDi formamide® (Applied Biosystems) injection buffer. The samples were then denatured for 3 minutes at 95 °C and placed on ice.

The samples were then sequenced on ABI 3730 capillary sequencer (Applied Biosystems). The ABI generated trace files were aligned into contigs with slight modifications of the default conditions using Sequencher® software version 4.6 (Gene Codes Corporation, USA). All chromatograms and sequences were visually inspected and sequences edited. Only good quality sequences with clear electropherogrames were used further.

3.4.6 Multilplex PCR with specific primers

PCR for all specific primers was optimized like in the previous protocol to get single clear amplification products. PCR was performed using AccuPower® PCR PreMix (www.bioneer.com) as desribed above. For all primers, initial optimization of PCR conditions involved adding into the PreMix 1.0 µl of 5 ng/µl DNA, 1.0 µl of 5 pmole/µl of F/R primer and 8.0 µl of H_2O. PCR amplification was done on a GeneAmp PCR system 9700 thermocycler (PerkinElmer, Wellesley, Mass) using the following program: Initial denaturation of 94 °C for 5 minutes followed by 40 cycles of denaturation of 94 °C for 20 seconds, primer annealing of 60 °C for 20 seconds, and extension of 72 °C for 1 minute and a final extension at 72 °C for 10 minutes. The amplified PCR products were electrophoresed on a 2 % pre-stained ethidium bromide agarose gel to confirm amplification. Primers showing multiple or no bands were further optimized by increasing or decreasing the annealing temperatures until a clear single band was visualized. The size of each PCR band was determined so that specific primers amplifying products of different sizes and also targeting different regions of the genome could be multiplexed in a single reaction. Four primer pairs Czm47 (439bp), Czm48 (230bp), Czm49 (174bp) and Czm51 (205bp) were tested for Multiplex PCR. Two sets of multiplex reations were designed; Primer Czm47 with Czm49 and primer Czm49 with Czm51. The multiplex PCRs was performed using AccuPower® PCR PreMix (

www.bioneer.com) as described earlier. For each multiplex reaction, initial optimization of PCR conditions involved adding into the PreMix 1.0 µl of 5 ng/µl Czm DNA, 1.0 µl of 5 pmole/µl of Czm47 primer, 1.0 µl of 5 pmole/µl of Czm49 and 1.0 µl of 5 pmole/µl of Czm51 and 8.0 µl of H_2O. PCR amplification was done on a GeneAmp PCR system 9700 thermocycler (PerkinElmer, Wellesley, Mass) using the following program: Initial denaturation of 94 °C for 5 minutes followed by 40 cycles of denaturation of 94 °C for 20 seconds, primer annealing temperature for 20 seconds, and extension of 72 °C for 1 minute and a final extension at 72 °C for 10 minutes. The amplified PCR products were electrophoresed on a pre-stained ethidium bromide agarose gel (2 %) to confirm multiplex PCR amplification.

3.4.7 Detection limit for PCR protocol

Czm genomic DNA was serially diluted ten-fold in the range of 1 ng/µl to 10^{-7} ng/µl. The serially diluted DNA samples were amplified with the optimized specific primers using both convensional and multiplex PCR to determine the detection limit.

3.4.8 Diagnosis of field samples

Samples showing symptoms of GLS of maize collected from farmers and experimental fields from research centres in five maize growing regions in Kenya (Table 3.1) were tested to confirm presence of *Czm* or *Cz*. The other set of samples, SCOH (USA), OYPA (Zambia) and 2 field samples from Kakamega, Kenya were also included in the diagnosis. Four specific primers pairs (Czm2, Czm39, Czm47 and Czm49) were utilised for these diagnostics. Diagnostic PCR was performed using AccuPower® PCR PreMix (www.bioneer.com) as described earlier. One microlitre of 5 ng/µl DNA, 1.0 µl of 5 pmole/µl of F/R primer and 8.0 µl of H_2O were added into the PreMix. PCR amplification was done on a GeneAmp PCR system 9700 thermocycler (PerkinElmer,

Wellesley, Mass) using the following program: Initial denaturation of 94 °C for 5 minutes followed by 40 cycles of denaturation of 94 °C for 20 seconds, optimized primer annealing temperature for 20 seconds, and extension of 72 °C for 1 minute and a final extension at 72 °C for 10 minutes. The amplified PCR products were electrophoresed on a 2 % pre-stained ethidium bromide agarose gel to confirm amplification.

3.5 Results

3.5.1 PCR amplification by designed primers

Based on the three aproaches, a total of 50 primer pairs for GLS of maize were designed. In the first approach involving utilizing known genes in *Cercospora nicotianae* to target intergenic regions which could have sequence variation among *Cercospora* species, 28 primer pairs were designed (Table 3.2). Initially 23 primer pairs were designed (primers Czm5 to Czm 27) to amplify intergenic regions in C. *zeae-maydis* and C. *zeina*. BLASTn analysis of the new sequences from PCR amplification using these primers confirmed that they were unique to *Czm* or *Cz*. Five more specific primer pairs were designed to sit within these intergenic regions (primers Czm47 - Czm51) . Out of the 28 primer pairs designed, 4 primers (Czm9, 47, 49 and 51) gave PCR products with clear distinct bands (Figure 3.3).

In an attempt to find species-specific regions, a detailed *in silico* search was performed in the available EST sequence data from the NCBI gene bank based on the second approach. Through BLASTn analysis, unique sequences with hits to *Czm* only were downloaded in FASTA format and used for primer design. Sixteen primer pairs (primers Czm28 to Czm43) were designed this way (Table 3.4) and all gave PCR products (Figure 3.3).

The third strategy utilized Malazy gene, a gypsy-like transposable element that is differentially expressed in the two GLS of maize *Cercospora* groups, i.e. PCR amplification of *pol* gene in Group I isolates (*Cercospora zeae-maydis*) and no amplification with group II (*Cercospora zeina*). Six diagnostic primer pairs were designed with this approach (Table 3.3) but 4 primers (Czm1 - Czm4) gave good PCR products with distinct bands (Figure 3.3).

Table 3.2: Czm and Cz Primer Sequences designed from Intergenic Region

No	Primer Name	Foward Primer 5'-3'	Reverse Primer 5'-3'
1	Czm5	TCGGAGAAAGAAGCAGACAATC	ACTTACATACGGCAGATACTCAGC
2	Czm6	GCACGAGCGGCTTGAGTTTG	GCAGAGCCCCTTTGATTACAGAA
3	Czm7	GTGGCGAATCGTGGCGTC	AACGAGAGCGAAGAGGAACATTTG
4	Czm8	TGGAGAGCGAGCATCACTTCAG	TTCATTACAGTCGAACTCCAACACC
5	Czm9	TGCGAGATGGTTTGGGGATTGT	GCTTTGAGGAGGCTGAGACCAG
6	Czm10	ATTAGCGTTTGTGTAGCCATTCA	TCCCGATTCCGATACTGTTCTACT
7	Czm11	TGTCGCTTATCGAGGTGACTCTTG	TGCTGAGCCCGATGAAAAGGTC
8	Czm12	ACGACGGCGTAGGAAGAAGTA	GTCGATCGCGGTGGTGAT
9	Czm13	GGAGAAGGATGAAGGGGTGGAGA	AAATCAGAGCGAAGAGCAAGTGGC
10	Czm14	ATATCTGGCTCGTTTCGCTCTGAC	GTTCTGCCCGGTTCCTTGGT
11	Czm15	CGTTGCCATCGCGAATAGCAT	AGCACGCTCTCCACGCCGA
12	Czm16	CTGGGACTAGCTTGCGGACTGT	GAGGGGCTGTGGTTTGGGAGT
13	Czm17	ATGGCGTGGAGGAAAGGCTGTTAG	GCGGCGTCTGCTCGTCGTAAA
14	Czm18	TGGCGGGCATCGTTGAG	CATATGGTGGCGTTTAGCTTAGC
15	Czm19	TTCCGAAAAACGAGCTGGTGTATC	ACGCCGCCCTACTCTATTCTGTC
16	Czm20	CATGGCCGTTGACACCATTAGC	CTTCCGCAGCCATCGTTTGA
17	Czm21	GGAGACTGCCGCCACCAACT	TCGCTGGCTGGTGGAACTCA
18	Czm22	AGTTGGTGGCGGCAGTCTCC	GATGAATGCTACCGCCAAGAACG
19	Czm23	TTTGGCACCGAAGTAGTAGAACAG	CGCGATCGACGGCTGAGG
20	Czm24	CGGCGTACGAGGTGGTGGT	CAGCTCTTTGGCCTTAGTGTTGG
21	Czm25	TCATCCTTCTCCGGCAACTCCA	CCGGCTTCATTGGCTTTCGTATC
22	Czm26	CAATCCAGCCCACCAAAGTCGT	ACGCGCTGTCGACATTATCAAGA
23	Czm27	CAGGCGATGCACCCAACACAA	CCCAGCGCCGACAGGGAC
24	Czm47	GCCTTCACAAGCCGGGACTACC	GGGGATTGTTTGCCTCGTGATAGC
25	Czm48	CTCGCATAACACAGGGCAATAGC	CGTCGCTATTGCCCTGTGTTATG
26	Czm49	CCAGGGCCGCATTGAGAAAGAG	TGATGGCGTGCTGCAAGAGAAG
27	Czm50	GCGCCAGCACGAGCAATC	CTTCGGAGAAAGAAGCAGACAATC
28	Czm51	GTTTCATGTCGGTTCGGGTTTCAG	CCACTGCAGCGATCCATTACTCCT

Table 3.3: Czm and Cz Primer Sequences designed from Malazy gene

No	Primer Name	Foward Primer 5'-3'	Reverse Primer 5'-3'
1	Czm1	GGCGCCGCTGCTGCTCTC	GCGCGCTTACTCCCCGTTC
2	Czm2	GCTAGGGCAAAAGCGAGGAGC	CGCCGGAGAAACGCCTAACC
3	Czm3	GTATAGAGCGCGTCTAAAGAAAAC	GCCTCCCTTCCCCTATCG
4	Czm4	AGGCGCCTTATATTTAGTTACCG	CATTAGGCTTTTATAGCGTTAGCG
5	Czm45	AGGCGTAGCGTAGTAGAGG	GCAGTTGCTTCGTAGCGT
6	Czm46	TACGCTCTAAAGGCATAACG	GCTTGCACTTTAGCTGCG

Table 3.4: Czm and Cz Primer Sequences designed from ESTs

No	Primer Name	Foward Primer 5'-3'	Reverse Primer 5'-3'
1	Czm28	CGGCCCATTCGTGTGAGGAG	GGTGCCCGTGCCGAGTCC
2	Czm29	CGCACCCCTCCCCTTGTCAG	GACCTTCGCCGCACCTCTCG
3	Czm30	CCCGACGGGCGAAAAGACC	CCGTAGGCGGCGAAGAGACC
4	Czm31	ATTGCCACTGACGCTCCCCG	CGCCCCGCAAACCAACATTC
5	Czm32	CTCCCGCTCCTGCCTTTTGA	GATGCCCCTGGTCTGTCTGGAG
6	Czm33	TCGCCCCTCGCGAAGATAACTAC	CCGACCCCGCAGATTGTGTTG
7	Czm34	CACGCGCCTGGTCGTTGTAC	TCACGGTTGCGGAGAAGGAGAC
8	Czm35	ACCGACCCCGCAGATTGTGTTG	CGCCCCTCGCGAAGATAACTACG
9	Czm36	GATCCCGCAGCTCCCTCAAGA	ATCGACGCCCGAGCAGAATG
10	Czm37	ACCCGCGGCATGTCTTATCTG	TCATCCCGGTGTCTCTTCTGTCTC
11	Czm38	CCTGGGCCAAAGCAGCAGTG	TCGCCGCAAATTCCCCTCAC
12	Czm39	GCCGCAATTCCCCTCACATCC	TCGCAAAGCCAGGGAACATCTC
13	Czm40	CAGCCCAAGTTACAGCGAAAGTGC	CCCCTTCACCATGCCACCAG
14	Czm41	TATTTTCAGCGCACCCACCTTGTC	GCTCGTTGCGCTCTTCAGTTATCG
15	Czm42	TGGCAGACGGCGGAGACAAC	ACCACGTCCGGAATCATCGC
16	Czm43	TAGCCGCCTTCACTCTCATCACG	TCCGCCTCGTTGTATCTTTCCTTG

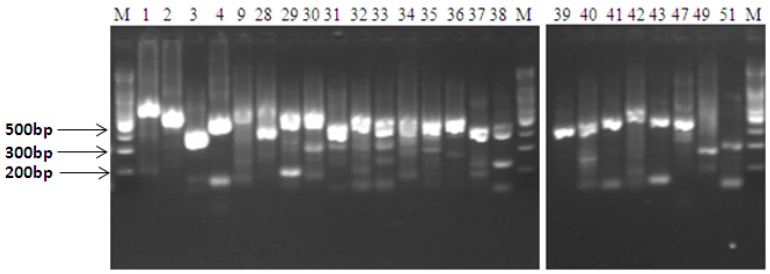

Figure 3.3: PCR products of 24 out of 50 GLS of maize primers to *Cercospora zeae-maydis* and *Cercospora zeina*. Primers Czm1, 2, 3, 4 (Malazy gene), Czm28-Czm43 (EST) and Czm9, Czm47-Czm51 (Intergenic region), M-Molecular size ladder

.

3.5.2 Sequencing of amplicons

Products from 4 primers pairs (primers Czm9, 47, 49 and 51) targeting the intergenic/ non-coding regions that gave PCR products with clear distinct bands (Figure 3.3) were selected for sequencing. They gave good quality sequences with clear electropherogrames with no BLASTn hits in the NCBI GenBank. These were therefore confirmed to be intergenic regions unique to *Cercospora zeae-maydis*. These sequences were further utilized for more specific primer design such that the primers lay within the intergenic region.

Eleven out of 16 PCR products from unique ESTs to both C. *zeae-maydis* and C. *zeina* gave good quality sequences. BLASTn analysis of these 11 sequences had hits only on *Cercospora zeae-maydis* ESTs and therefore can be confirmed to be unique to *Cercospora zeae-maydis* (Table 3.4)

Nineteen out of the 24 primer pairs that gave PCR products (Figure 3.3) were confirmed to be specific for *Czm* and *Cz* which gave PCR products of single descrete bands on agarose gel electrophoresis (Table 3.5).

Table 3.5: 19 primer pairs confirmed to specific to *Czm and Cz*

No	Pime Name	Foward Primer Sequence 5'-3'	Reverse Primer Sequence 5'-3'	Source
1	Czm1	GGCGCCGCTGCTGCTCTC	GCGCGCTTACTCCCCGTTC	Malazy gene
2	Czm2	GCTAGGGCAAAAGCGAGGAGC	CGCCGGAGAAACGCCTAACC	Malazy gene
3	Czm3	GTATAGAGCGCGTCTAAAGAAAAC	GCCTCCCTTCCCCTATCG	Malazy gene
4	Czm4	AGGCGCCTTATATTTAGTTACCG	CATTAGGCTTTTATAGCGTTAGCG	Malazy gene
5	Czm9	TGCGAGATGGTTTGGGGATTGT	GCTTTGAGGAGGCTGAGACCAG	Intergenic region
6	Czm47	GCCTTCACAAGCCGGGACTACC	GGGGATTGTTTGCCTCGTGATAGC	Intergenic region
7	Czm49	CCAGGGCCGCATTGAGAAAGAG	TGATGGCGTGCTGCAAGAGAAG	Intergenic region
8	Czm51	GTTTCATGTCGGTTCGGGTTTCAG	CCACTGCAGCGATCCATTACTCCT	Intergenic region
9	Czm29	CGCACCCCTCCCCTTGTCAG	GACCTTCGCCGCACCTCTCG	EST
10	Czm30	CCCGACGGGCGAAAAGACC	CCGTAGGCGGCGAAGAGACC	EST
11	Czm32	CTCCCGCTCCTGCCTTTTGA	GATGCCCCTGGTCTGTCTGGAG	EST
12	Czm33	TCGCCCCTCGCGAAGATAACTAC	CCGACCCCGCAGATTGTGTTG	EST
13	Czm34	CACGCGCCTGGTCGTTGTAC	TCACGGTTGCGGAGAAGGAGAC	EST
14	Czm35	ACCGACCCCGCAGATTGTGTTG	CGCCCCTCGCGAAGATAACTACG	EST
15	Czm36	GATCCCGCAGCTCCCTCAAGA	ATCGACGCCCGAGCAGAATG	EST
16	Czm39	GCCGCAATTCCCCTCACATCC	TCGCAAAGCCAGGGAACATCTC	EST
17	Czm41	TATTTTCAGCGCACCCACCTTGTC	GCTCGTTGCGCTCTTCAGTTATCG	EST
18	Czm42	TGGCAGACGGCGGAGACAAC	ACCACGTCCGGAATCATCGC	EST
19	Czm43	TAGCCGCCTTCACTCTCATCACG	TCCGCCTCGTTGTATCTTTCCTTG	EST

3.5.3 Sensitivity tests

To assay for sensitivity, these primers were tested for the detection limit using 10-fold serial dilutions of DNA in the range of 1 ng/µl to 10^{-7} ng/µl. Primer Czm9 gave PCR products from low DNA concentration of upto 10^{-3} ng (Figure 3.4).

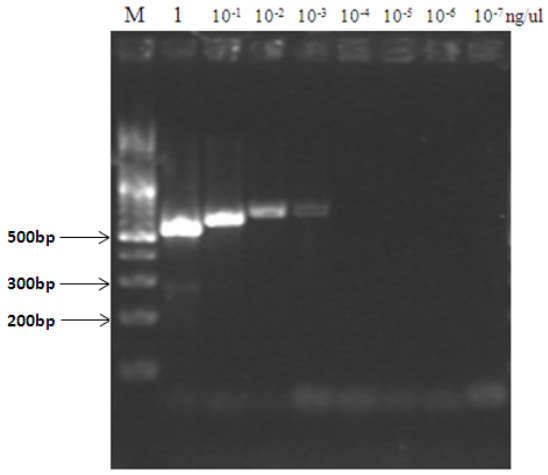

Figure 3.4: Detection limit for primer Czm9 on serial dilution of DNA in the range of 1 ng/µl to 10^{-7} ng/µl, M- 100 bp Molecular size ladder

3.5.4 Multiplex PCR

The 19 primer pairs were subsequently tested for their ability to give good multiplex PCR results, high sensitivity and act as effective diagnostic primers for *Cercospora zeae-maydis* field samples. A multiplex PCR approach was also employed to improve the specificity of PCR diagnosis. Two sets of primers that amplify two distinct products of different sizes were used for multiplex PCR. Primer sets, Czm47 was multiplexed with Czm49 while primer set Czm47 was multiplexed with Czm51. These primers were able to distinguish the two GLS of maize *Cercospora* groups isolates. Smaller fragments of 174 bp for multiplex primers Czm47/49 and a 205 bp for multiplex primers Czm47/49, serves as the positive controls and are amplified for the two *Cercospora* species, while the larger 439 bp fragment is only observed for *C. zeina* and absent for *Cercospora zeae-maydis* species (Figure 3.5). *C. zeina* samples produced the two bands i.e. samples 1, 2, 3, 37, 38 and 39 while *C. zeae-maydis* produced one band i.e. sample 36.

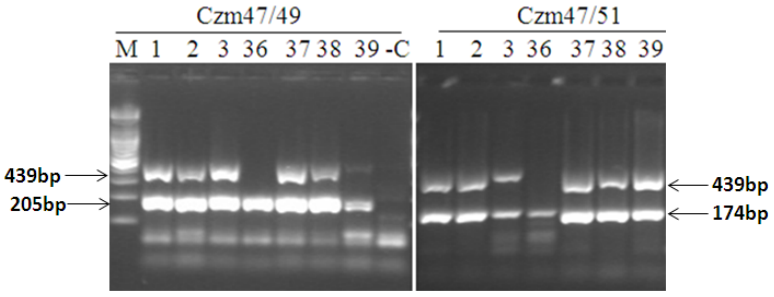

Figure 3.5: Multiplex PCR with two sets of primers Czm47/49 and Czm47/51. Samples 1-3 (DNA extracted from pure Cultures), Sample 36(SCOH *Cercospora zeae-maydis*), sample 37 (OYPA *C. zeina*), samples 38 and 39 (DNA extracted from GLS of maize infected leaf), -C (negative control) and M-100 bp Molecular size ladder.

3.5.5 Diagnosis of maize field samples

Thirty five samples showing symptoms of GLS of maize which were collected from farmers fields and experimental fields from research centres in five maize growing regions in Kenya were tested to confirm presence of *Czm* or *Cz*. Another set of samples, SCOH (USA), OYPA (Zambia) and 2 field samples from KARI-Kakamega, Kenya were also included in the diagnosis. Four primer pairs (Czm2, Czm39, Czm47 and Czm49) were utilised for these diagnostics (Figure 3.6). Primer pairs Czm39 (EST) and Czm49 (intergenic) were able to amplify the two *Cercospora* groups isolates, Group I (*C zeae-maydis*) and group II (*C. zeina*) with clear distinct bands except for samples 31-35 from unknown pathogens. Primer Czm47 amplified all samples except samples 7, 26, 28, 30, 31-35 (unknown pathogens) and 36 (SCOH known to be *C zeae-maydis*). Primer Czm2 amplified samples 7, SCOH and OYPA samples at the expected band size but did not give good PCR products with the rest of the samples. A negative control DNA extracted from non-symptomatic maize sample was also tested with these primer sets and did not give an amplification product (Figure 3.6).

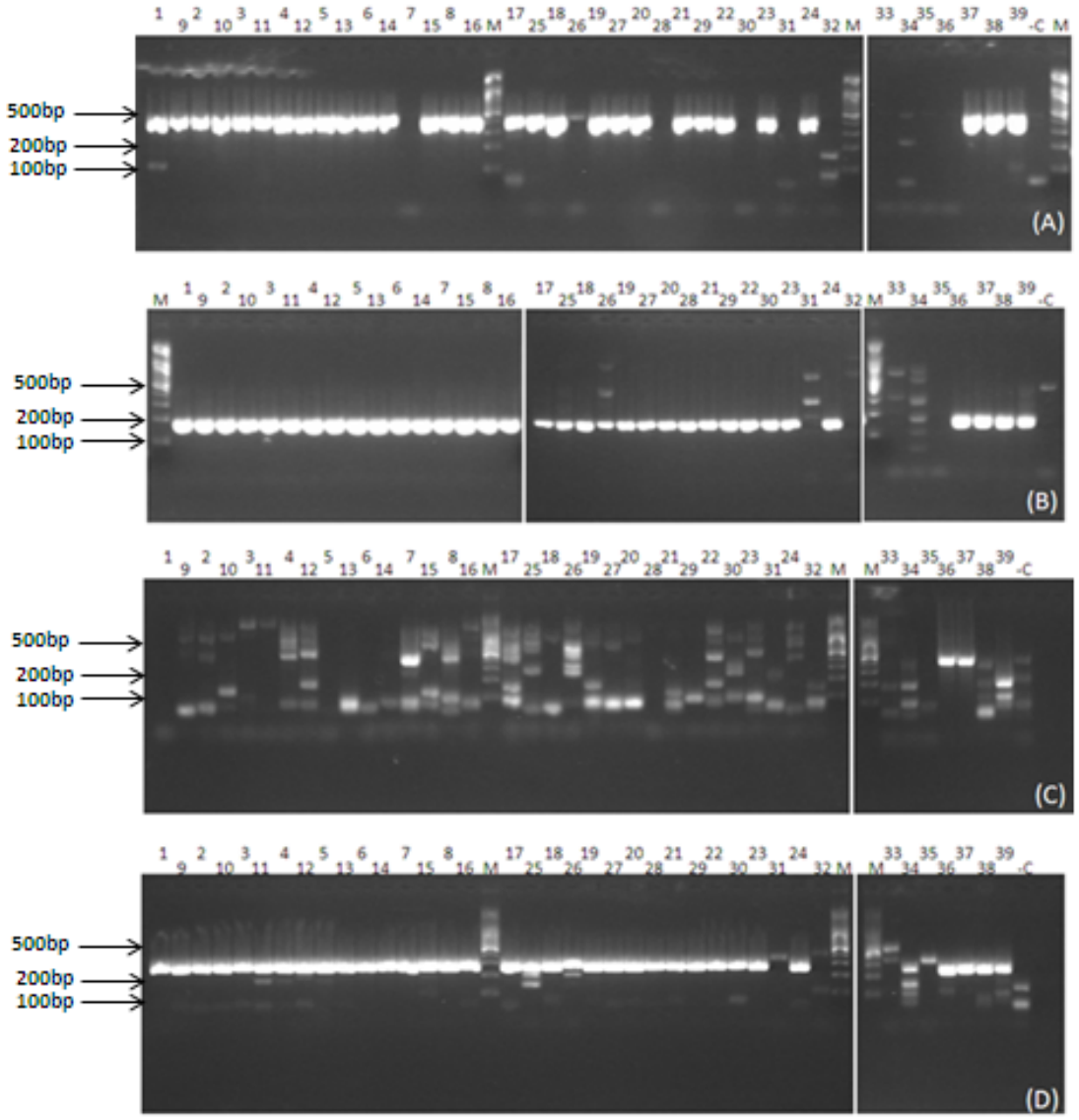

Figure 3.6: PCR amplification of field samples with four Specific primer pairs. Panel (A), Czm47 (Within Intergenic region), Panel (B), Czm49 (Within Intergenic region), Panel (C), Czm2(Malazy gene) and Panel (D), Czm39 (Unique EST). Samples 1 to 30 kenyan GLS of maize sample, samples 31-35 unknown pathogens with symptoms similar to GLS of maize, sample 36 SCOH, sample 37 OYPA, sample 38 and 39 from non-cultured infected leaf and sample -C-negative control and M-Molecular size ladder.

3.6 Discussion

3.6.1 Molecular diagnostic tool for grey leaf spot disease of maize

The increasing availability of completely sequenced genomes and expressed sequence tags (ESTs) has now made genome and transcriptiome scale comparisons possible. These comparisons offer the opportunity of exploring and designing specific PCR primers based on intergenic sequences and unique sequences. In this study unique *Cercospora zeae-maydis* EST sequences from the NCBI GeneBank were utilized in selecting regions for designing 16 Czm and Cz PCR primer pairs. Out of these, 11 primers worked well when tested with both Czm and Cz DNA. Amplicons wth single distinct bands were observed following PCR and running the products on an agarose gel electrophoresis. Sequences from these PCR products had BLASTn hits on ESTs unique to *Cercospora zeae-maydis* and C. *zeina* species.

However the primers from ESTs cannot distinguish the two groups of *Cercospora* species as shown in Figure 3.6 (D) i.e. group I isolates (*Cercospora zeae-maydis*) and group II (C. *zeina*) since they are specific to expressed genes that are common to the two species. This lead us to utilize the Malazy gene, a gypsy-like transposable element that is differentially expressed in the two *Cercospora* groups. Six primer pairs were designed based on this gene and 4 were able to give amplification products with SCOH (*Cercospora zeae-maydis*) and no amplification product with all other samples of group II (C. *zeina*). These primers can be used for the specific diagnosis of *Cercospora zeae-maydis* and not *Cercospora zeina*.

In an attempt to design more specific primers that can distinguish the two *Cercospora* groups, the transcription map of the cercosporin toxin biosynthesis (CTB) cluster genes in *Cercospora nicotianae* and the adjacent open reading frames (ORFs) were also utilized to mine intergenic sequences in both C. *zeae maydis* and C. *zeina*. Here, the

variability of intergenic regions were also targeted. Twenty eight primer pairs were designed (Table 3.2) of which 4 were found to be specific for diagnosis. BLASTn analysis on NCBI or CMR of C. *zeae maydis* and C. *zeina* intergenic sequences from PCR products of these primers did not produce close hits with any other species.

Primer Czm47 designed from intergenic regions is able to distinguish the two *Cercospora* species i.e. group I isolates (*Cercospora zeae-maydis*) and group II (C. *zeina*). Primer Czm47 is specific to C. *zeina* and did not amplify sample 36 which is a positive control for SCOH *C. zeae-maydis*). Multiplex primer sets Czm47/49 and Czm47/51 produces two bands with C. *zeina* and one band with C. *zeae-maydis* (Figure 3.5). Other molecular diagnosis protocols to distinguish the two species have been developed previously. However this require digesting the amplified PCR products with a combination of several restriction enzymes to distinguish the two species (Wang et al., 1998; Kinyua et al., 2010). Multiplex primers sets Czm47/49 and Czm47/51 can therefore be used to easily diagnose the two groups of *Cercospora* species in a single PCR reaction.

To demostrate the sensitivity of these primers, the detection limit using 10-fold serial dilutions of DNA in the range of 1 ng/µl to 10^{-7} ng/µl was tested. One primer pair Czm9, had the highest sensitivity giving a PCR product from DNA concentration of as low as 10^{-3}ng (Figure 3.4). At this concentration the maize plants are symptomless for GLS of maize and the primer can therefore be used to detect the pathogen before visual symptoms appear allowing early disease management strategies to be implimented.

3.6.2 Application of *Czm* and *Cz* specific primers

These primers can be applied for the specific identification of GLS of maize. Primer Czm49 which amplifies within the intergenic region and primer Czm39 which amplifies a unique EST were able to detect the two *Cercospora* groups isolates, Group

I (*Cercospora zeae-maydis*) and group II (C. *zeina*). These primers cannot amplify non *Cercospora* isolates (samples 31-35) as was demostrated in Figure 3.6. Primer Czm47 specific to C. *zeina* did not amplify samples 7, 26, 28, 30, and 36 (positive control for SCOH *C. zeae-maydis*). This confirms that samples 7, 26, 28 and 30 are *C. zeae-maydis* group 1 samples. Samples 31-35 from unknown pathogens were not amplified by either primers indicating they were not *Cercospora zeae-maydis* or *Cercospora zeina*.

Primers Czm2 designed from the Malazy gene (Shim and Dunkle, 2005), can be used to distinguish the two *Cercospora* groups isolates, Group I (*Cercospora zeae maydis*) which has an amplification product and group II (C. *zeina*) which does not have an amplification product (Figure 3.6). Primer Czm9 designed from the intergenic region was very sensitive since it can amplify trace amounts of pathogen DNA (up to 10^{-3} ng) and did not amplify other species (Figure 3.4). This shows that primer can be used for diagnosis of GLS of maize from infected symptomless maize plants. The specificity of this primer results are consistent with those reported by Shim and Dunkle (2005) in the two *Cercospora* groups isolates.

The approach demonstrated in this study can similarly be exploited in designing diagnostic PCR primers for other *Cercospora* species or pathovars and also for other plant, animal or human pathogens. Conserved genes present in all *Cercospora* could be targeted for designing primers to identify all *Cercospora* species. Such primers would be useful in identification of *Cercospora* at the genus level. Hence sequences from conserved or unique genes and those from intergenic regions can be used to develop primers for specific identification of *Cercospora* at genus, species or pathovar levels. Furthermore, sequence analyses within the intergenic regions could be utilized for diversity studies within the *C. zeae-maydis* and *C. zeina* species.

Chapter 4

DEVELOPMENT AND EVALUATION OF PCR DIAGNOSTIC PRIMERS FOR CASSAVA BROWN STREAK DISEASE (CBSD)

4.1 Abstract

Cassava brown streak disease (CBSD) has been a problem in the East African coast for more than 70 years and is now spreading to other areas. For the early and accurate detection of this disease, a two-step Reverse Transcriptase PCR method for cassava brown streak virus (CBSV) was developed. Eighteen diagnostic primers were designed according to the most conserved regions of six published coat protein (CP) sequences for CBSV. The primers were highly sensitive and one primer set, CBSV-R1R10, was able to detect up to 10^{-5} fold diluted CBSV cDNA sythesized from 200 ng of RNA compared to a commonly used primer set CBSV-10/11 that could only detect up to 10^{-2} fold dilution of the same cDNA. Cassava field-sampling and diagnosis procedures were also established in this study. These protocols have great potential to minimize field sample handling and contamination. The ability of the primer CBSV-R1R10 to detect infection was validated on field samples from cassava multiplication sites in Kenya, Uganda, Rwanda, Democratic Republic of Congo (DRC), Burundi and Tanzania. CBSV-R1R10 showed positive results on all symptomatic leaves for CBSD and also some symptomless plants. The diagnostic primers therefore have the potential for diagnosis of CBSV on field plant materials. They were also highly efficient for the

detection of early infections of CBSVs even before symptoms appear which will be highly valuable for epidemiological studies and for testing of symptomless plants.

4.2 Introduction

Cassava (*Manihot esculenta* Crantz) is the second most important food crop after maize (Nweke et al., 2011) in Africa. It is valued for its broad ecological adaptation, low labour requirement, ease of cultivation and high yields. It is widely cultivated because it can be successfully grown in poor soils and under conditions of marginal rainfall. It remains a reliable starchy staple crop for 200 million Africans (New Partnership for Africas Development (NEPAD) Secretariat).

Cassava provides more than half of the dietary calories for over half of the total rural and urban population in sub-Saharan Africa. The leaves, eaten as a vegetable, provide a cheap and rich source of protein, minerals, and vitamins. The nutritional value of cassava leaves is compared favorably with the composition of other foods such as soybean grain, maize grain and amaranths (Bokanga, 1994). Cassava leaves eaten in East Africa have a well balanced amino acid composition with an essential amino acid profile better than standard nutritional recommendations (West et al., 1988). About 2kg of cassava leaves are consumed weekly per person in the DRC implying that more than 5,304,000 MT of leaves are produced (Mahungu et al., 2003).

Due to its suitability in low input farming systems, cassava production has tripled in sub-Saharan Africa over the past four decades (Hillocks et al., 2002). Recent increases in the area planted have been a response to erratic and declining rainfall and its use as a replacement for cereals on exhausted soils. Africa is currently the worlds largest cassava producer with an estimated production of 110 million metric tons (MT) of fresh roots. The crop is both a major food staple and a key food resilience crop. Virtually the entire crop is produced by smallholder farmers (Spencer and Associates, 2005) on 12 million hectares (ha) (FAOSTAT, 2005). Major producing countries are Nigeria (42 million MT), Democratic Republic of Congo (DRC) (15 million MT), Ghana (10 million MT), Tanzania (7 million MT) and Uganda (5 million MT) (Figure 4.1). Women

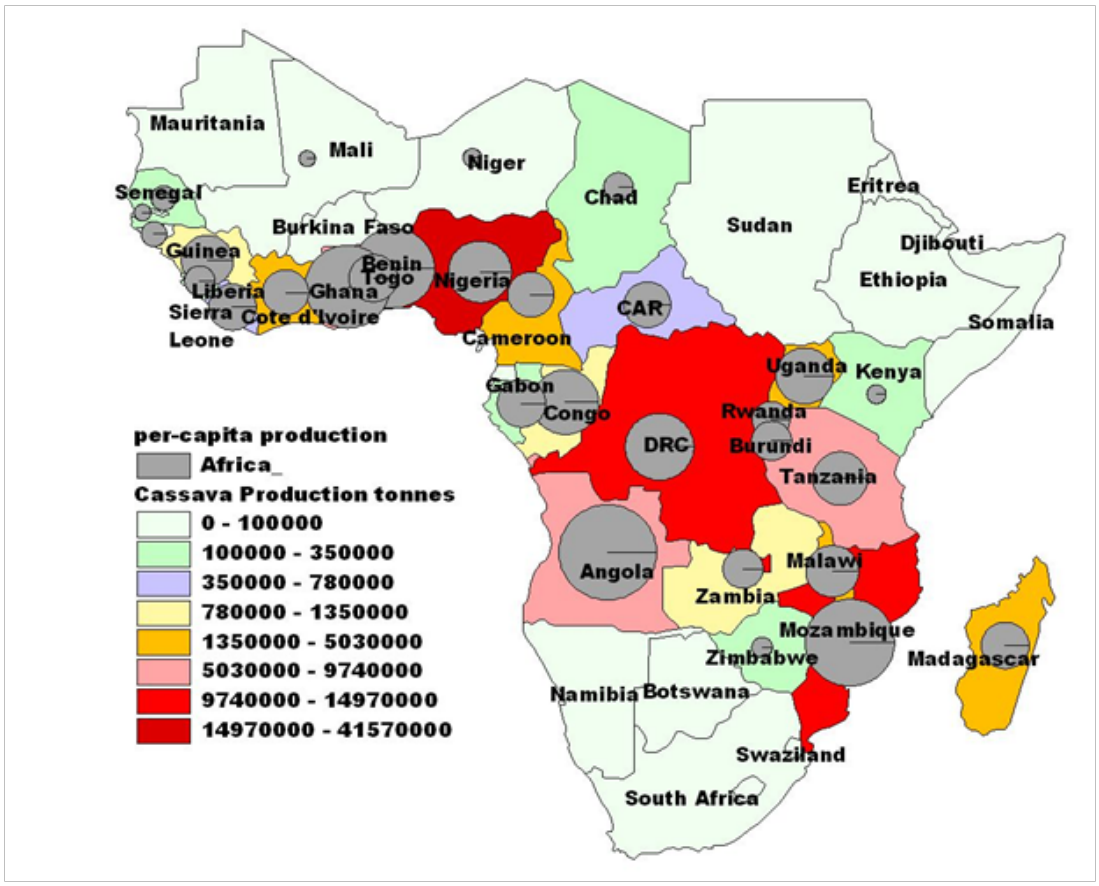

Figure 4.1: Total and per capita cassava production in sub-Saharan Africa FAOSTAT (2005)

are responsible for a large proportion of this production and for almost all of the processing and marketing activities. Over 90 % of production is for human consumption. Table 4.1 indicates the importance of cassava in East and Central Africa.

Flexibility of labour and harvest means that cassava is particularly important for the poor and vulnerable in rural communities, showing considerable benefits for communities with a high incidence of HIV/AIDS, as indicated by the International Food Policy Research Institute (IFPRI). Cassava proves financially profitable for smallholders in a wide variety of settings. It requires no purchased inputs. Its flexible planting and harvesting calendar enables households to fit in labor requirements around other obligations, making cassava one of the easiest crops to grow. Initial evidence from Zambia

Country	Area planted (ha)	Production (MT)	Per capita production (kg/yr)	Daily intake (g/capita)	% calorie intake derived from cassava
Burundi	82,000	710,000	134	228	29
DRC	1,845,510	14,974,470	435	750	73
Rwanda	116,000	781,640	113	245	32
Tanzania	660,000	7,000,000	273	378	22
Kenya	N/A	630	18.2	N/A	Up to 45
Uganda	405,000	5,031,000	300	309	28

Table 4.1: Cassava production and consumption in East and Central Africa. FAOSTAT (2005)

suggests that HIV/AIDS prevalence makes a small but positive contribution to area expansion of cassava among affected household (Nweke et al., 2011).

The ability of cassava to become a primary resource for the livelihood development of rural communities in Africa has been demonstrated in Nigeria and Ghana. In both countries, cassava production has rapidly grown in tandem with urbanization and the demand for processed cassava in the form of gari; a dried, fermented cassava product. The same process of rural processing linked to urban consumption is evident in Latin America, with the development of Farinha; and, more recently, there has been additional market development of cassava in the form of animal feed and starch production in both Asia and Latin America. In Latin America, industrial processing is now supplied from large-scale mechanized farms; but in Asia, the industrial-level processing is still supported by large numbers of smallholder farmers; a model that may be applicable in Africa.

Production methods are generally sub-optimal in much of Africa, however, and particularly in the less commercially-oriented parts of Eastern Africa. Production is characterized by use of traditional varieties and inadequate cultural practices, particularly poor quality planting material; but also sub-optimal planting densities and inadequate weed, pest and disease management. There is virtually no mineral fertilizer used on cassava and soil nutrients are seldom replenished. The resulting yields of 5-9 MT/ha, compare unfavorably with levels of 25-35 MT realized in low input on-station trials in East Africa (Fermont *et al*., 2004) and up to 80 MT/ha realized on commercial farms. These higher yields were due mainly to the use of resistant varieties, clean planting material and improved agronomic practices; demonstrating the scope for productivity enhancement through packages containing these essential elements.

Concomitant with increased cassava production has been increased pressure from a range of pests and diseases, most particularly cassava mosaic disease (CMD), but also white fly, cassava green mite, cassava mealy bug, cassava bacterial blight and, more

recently, cassava brown streak disease (CBSD). Despite the crops generally robust nature, these pests and diseases can lead to crop losses of between 10 % and 100 % (Hahn. et al., 1980; Bellotti and Ariass, 2001).

The prominent symptoms of CBSD appear on leaves with varying patterns of chlorosis. Leaf chlorosis appear in a feathery pattern, first along the margins of the secondary veins, later affecting tertiary veins and may develop into chlorotic blotches (Figure 4.2). Alternatively, the chlorosis may not be clearly associated with the veins but appear in roughly circular patches between the main veins. There is considerable variation in the expression of foliar symptoms depending on variety, growing conditions (temperature, rainfall, and altitude), age of the plant, and the virus isolate involved in causing the symptoms. Some cultivars show marked foliar symptoms but without or delayed root symptoms, and vice versa. Symptoms of the disease are more difficult to recognize in older plants as the lower leaves with prominent symptoms senesce and fall off. New leaves produced from these plants often do not show symptoms, especially at high temperatures. The roots can become constricted due to the tuber rot and stunting growth (Figure 4.2). Symptoms can be also transient when a period of active growth produces symptom-free tissues (Jennings, 1960). However, it is difficult to interpret these observations precisely because they have been made in field situations with varying agroclimatic conditions, on cassava varieties with differing virus resistance levels and crop age, and possibly infected with different virus strains, which on their own or in combinations affected symptom development.

CBSD is transmitted by whiteflies, also the vector for CMD. This pandemic has been driven by what is thought to be a novel type of whitefly vector *Bemisia tabaci* (Legg et al., 2002). Super-abundant populations of whitefly are also causing physical damage through sheer weight of numbers in parts of the pandemic-affected zone. While yield losses related to CBSD have not been extensively documented, field trails in Tanzania show that CBSD can decrease root weight by up to 70 % with the length of time

Figure 4.2: Cassava brown streak disease (CBSD) symptoms on the leaf (left) and rotting of the root (right).
Source: Biosciences for Farming in Africa (B4FA)

between appearance of leaf symptom and the development of root necrosis (rot) being a varietal characteristic. Most susceptible varieties show root rot within 6 months of planting in cases where cuttings are derived from a symptomatic mother plant. Foliar symptoms do not greatly affect plant growth although the most sensitive varieties may be stunted and defoliated. This is a marked contrast with CMD where visual diagnosis can be highly accurate. While the destructive symptoms of CBSD are root necrosis, the presence of foliar symptoms is not an indication that a cultivar suffer root necrosis (Mohammed et al., 2012).

CBSD is now endemic in the coastal areas of Mozambique, Tanzania and Kenya (Figure 4.3). The disease has recently appeared in the mid-altitude areas of Uganda and Tanzania, as shown in this map. cassava brown streak disease (CBSD) has also become increasingly widespread, since 2005, in parts of Africa from where it had never previously been common (Uganda and north-western Tanzania). This apparently new epidemic threatens much of the recent success realized in mitigating the impacts of the CMD pandemic.

CBSD is caused by two viruses that are sufficiently genetically distinct to be given species status (Winter et al., 2010; Mbanzibwa et al., 2011). These two strains have

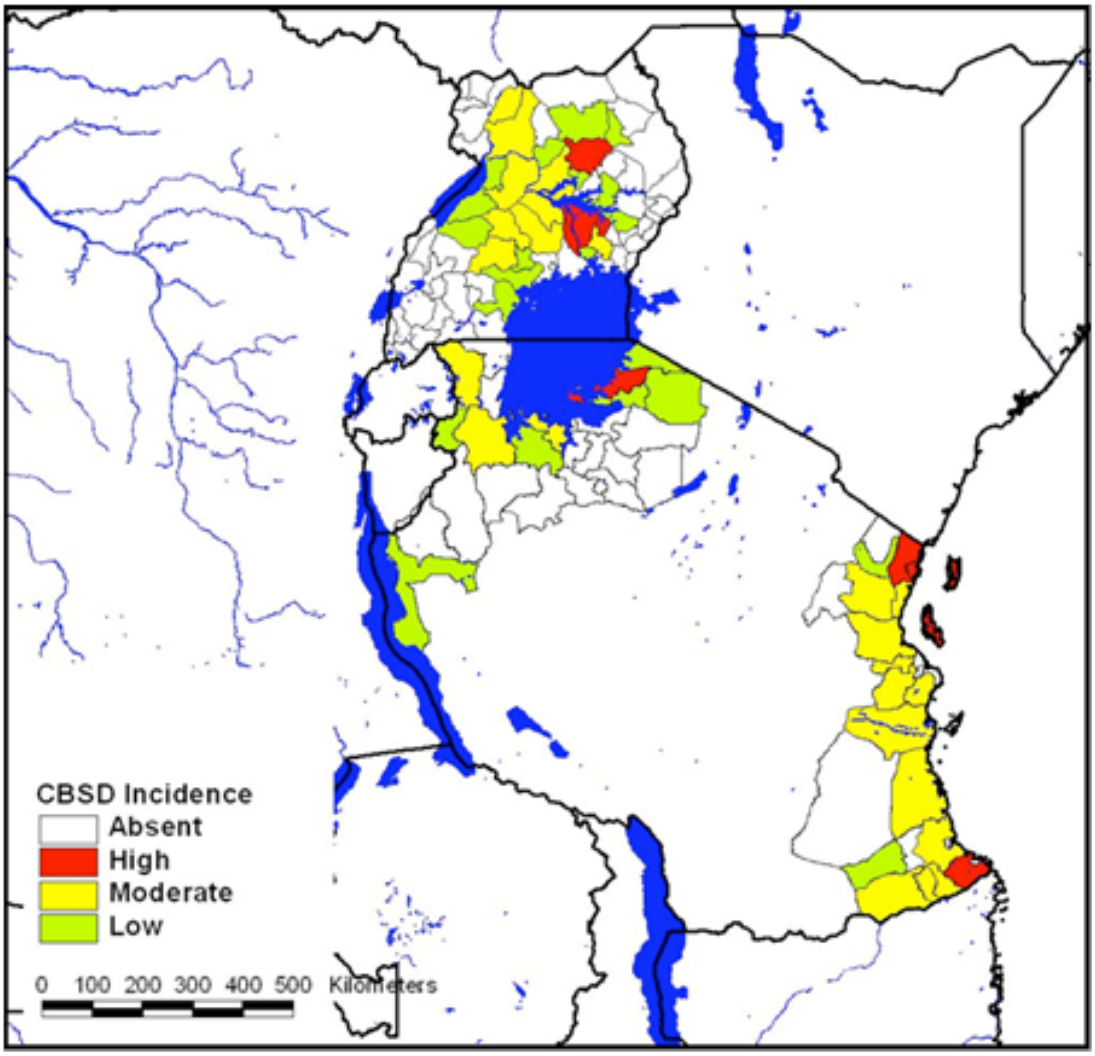

Figure 4.3: CBSD incidence in East Africa.
Source: Great Lakes Cassava Initiative (GLCI), 2012)

been described in the literature as CBSV for the Tanzanian or coastland strain and cassava brown streak Ugandan virus (CBSUV) for the Ugandan or highland strain (Mbanzibwa et al., 2009b; Monger and Foster., 2001; Winter et al., 2010). The name CBSUV has recently been approved by the International Committee on Taxonomy of Viruses, with CBSV retaining its original species nomenclature. CBSV and CBSUV supersede other CBSV strain terminologies and the two viruses are collectively termed CBSVs (Legg et al., 2011).

Cassava brown streak virus is a member of the genus Ipomovirus that belongs to the family Potyviridae, which contains the largest number (ca. 200) of positive single-stranded RNA viruses infecting plants (Fauquet et al., 2005; Maruthi et al., 2005; Monger and Foster., 2001). This virus family is divided into the genus *Bymovirus* with bipartite genomes and the genera *Ipomovirus, Macluravirus, Potyvirus, Rymovirus, and Tritimovirus*, containing monopartite viruses that encode a large polyprotein auto-proteolytically cleaved into 10 mature proteins (Figure 4.4) (Fauquet et al., 2005; Valli et al., 2008). Additionally, a small open reading frame (ORF) created by frameshifting was recently detected in the P3 protein encoding region (Chung et al., 2008). Among members of the Potyviridae, ipomoviruses are exceptional in variability of protein-encoding sequences at the 5' end of the genome (Figure 4.4). Sweet potato mild mottle virus (SPMMV) (Colinet et al., 1998) contains a single P1 serine proteinase at the polyprotein N terminus, whereas Cucumber vein yellowing virus (CVYV)(Janssen et al., 2005; Lecoq et al., 2000; Zhang et al., 2008) and Squash vein yellowing virus (SqVYV) (Li et al., 2008) contain two P1 proteinases (P1a and P1b) that are evolutionary diversified (Valli et al., 2007). In addition, SqVYV and CVYV lack the multifunctional helper component proteinase (HC-Pro) (Janssen et al., 2005; Li et al., 2008; Valli et al., 2007), which is located second in the polyprotein (Figure 4.4) in other monopartite Potyviridae members (Adams et al., 2005; Fauquet et al., 2005) and which acts as a suppressor of RNA silencing (Anandalakshmi et al., 1998; Brigneti

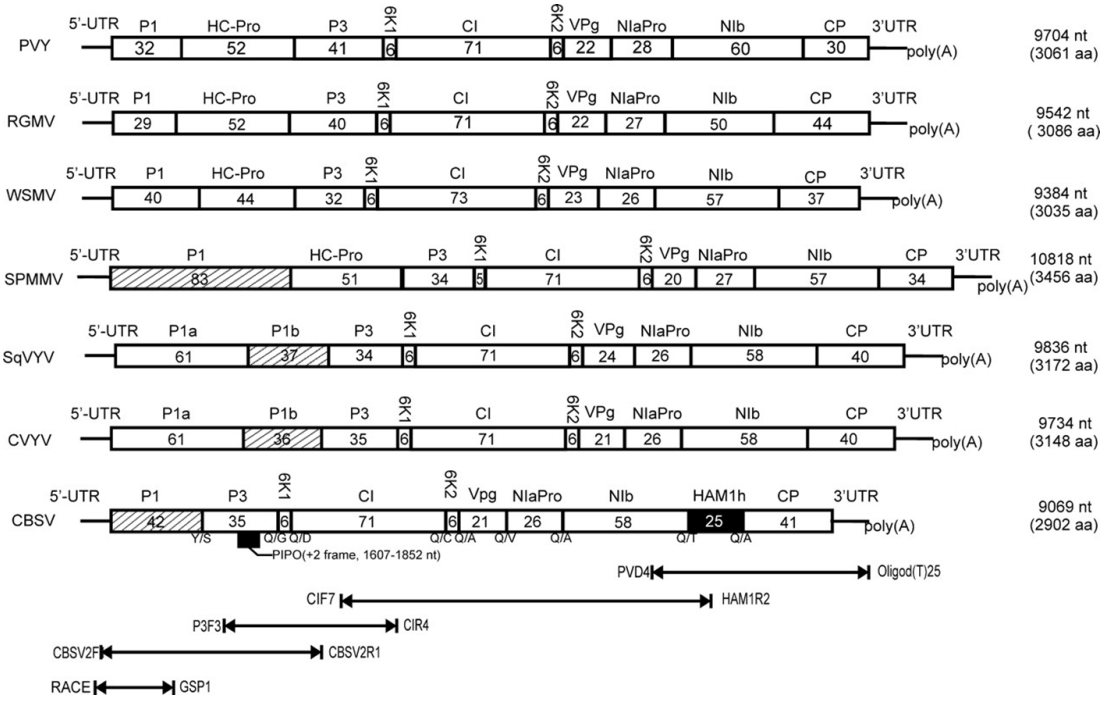

Figure 4.4: Comparison of the viral genomes in four genera of Potyviridae. (Mbanzibwa et al., 2009a)

et al., 1998; Kasschau and Carrington, 1998; Zhang et al., 2008). This function has been adopted by P1b in CVYV (Valli et al., 2008, 2006).

The common method used for detection of CBSD is based on symptoms of infected cassava plant. The method is unreliable because the symptoms are inconsistently expressed in leaf, stem and root and are difficult to distinguish from mite damage and nutrient disorders (Hillocks and DL, 2003). Also, there is no evidence of any symptom differences between CBSV and CBSUV, making differentiation of CBSVs by visual symptoms impossible. Current PCR based detection methods are also inconclusive due to unspecificity primers due to emergence of new strains. For these reasons a sensitive detection method would be a valuable tool required to accurately identify the presence of viruses for research, policy (e.g. quarantine) and planting material multiplication purposes.

4.3 Objectives of the study

4.3.1 General objective

The main objective of this study was to develop and evaluate PCR based diagnostic primers for cassava brown streak disease (CBSD)

4.3.2 Specific objectives

1. To utilize conserved regions of the available CBSV sequences in the design of specific primers for the detection of CBSD

2. To test the specificity and sensitivity of the designed primers compared to the current diagnostic primers

3. To validate the designed primers with cassava planting materials from multiplication sites

4.4 Materials and Methods

4.4.1 CBSV diagnostic primer design

By the time this study was being conducted the genome structure of CBSV was not known. Only partial coat protein (CP)-encoding sequences of coastal lowland isolates (Monger and Foster., 2001) and complete CP sequences of highland isolates from East Africa were available, revealing that these isolates belong to two phylogenetically different strains (Mbanzibwa et al., 2009a). These available CP sequences were utilized to design diagnostic primers for the specific identification of CBSV. CBSV was initially detected using the previously described CBSV10 and CBSV11 primers (Monger and Foster., 2001) but the inability of these primers to detect viruses from known infected samples prompted the development of new primers. For this purpose, better diagnostic primers were designed based on six published coat protein (CP) sequences on the NCBI (GenBank accession Numbers AY008440, AY008441, AY008442, AY007597, AF311052 and AF311053). These sequences were downloaded and assembled into contigs using the Sequencher®v4.6 program. Diagnostic primers were designed according to the most conserved regions of the six CP sequences or in such a way that the 3' end of the primer was positioned on a region that is conserved across the six sequences. Primers were designed using the PrimerSelect®software of Lasergene 7. These primers were used for diagnostis of CBSV together with previous publicly available primers (Monger and Foster., 2001).

4.4.2 RNA extraction

Leaf samples from cassava plants showing typical CBSD symptoms, 'Guzo' and 'Kiberiti Mweusi' clones representing coastal lowland isolates of CBSV, were collected from the Kenya Plant Health Inspectorate Service (KEPHIS) substation at Muguga, Kenya.

RNA was extracted based on a modified protocol as described by Lodhi et al. (1994). Approximately 0.3 g of leaf sample was ground on liquid nitrogen using mortar and pestle. The fine powder was transferred into a 2 ml microfuge tube and stored at -80 °C until use. Eight hundred microlitres of extraction buffer (2.0 % (w/v) CTAB (cetyltrimethylammonium bromide), 2.0 % PVP 40 (polyvinylpolypyrrolidone), 20 mM EDTA, 1.4 M NaCl, 100 mM Tris-HCl pH8.0) was added. The samples were incubated at 60 °C for 15 minutes. Six hundred microlitres of chloroform:Isoamyl (24:1) alcohol were added and samples mixed gently by inverting the tubes 20 to 25 times to form an emulsion. Samples were centrifuged at 13000 rpm for 10 minutes at room temperature. The upper aqueous phase was transfered to a new microfuge tube. A second chloroform:Isoamyl extraction was performed if the aqueous phase was cloudy due to the presence of PVP. Half volumes of 5 M NaCl and two volumes of cold (-20 °C) absolute ethanol was added to the aqueous solution. The samples were mixed well and incubated at -20 °C for 20-30 minutes. Samples were centrifuged at 6500 rpm for 10 minutes to pellet the nucleic acid at the bottom of the microfuge tube. The supernatant was poured off and the pellet suspended in 200 µl of 2 M LiCl. Samples were incubated at 4 °C overnight and then centrifuged at 13000 rpm for 30 minutes to pellet the RNA. The LiCl was poured off and the pellet washed with 500 µl cold (0 to 4 °C) 70 % ethanol. The ethanol was completely removed and the pellet left to dry at room temperature. The pellet was suspended in 100 µl of Diethylpyrocarbonate (DEPC) treated distilled water. The yield and quality of the RNA was checked on the Nanodrop®ND-1000 spectrophotometry (NanoDrop Technologies, USA). The extracted RNA was then diluted to 100 ng/µl and stored at -80 °C.

4.4.3 cDNA synthesis

A reverse transcriptase PCR protocol was used in which cDNA was synthesized from extracted RNA with AccuPower®CycleScript Premix RT(dT20) kit (BIONEER Corporation) containing a thermostable cyclescript reverse transcriptase in a total 20 µl reaction volume. Two microlitres of 100 ng/µl RNA was dispensed into the AccuPower CycleScript RT PreMix(dT20). The AccuPower CycleScript RT PreMix(dT20) was filled up to 20 µl with DEPC-distilled water. cDNA was synthesized in a single temperature reaction of 45 °C for 60 minutes followed by RTase inactivation at 95 °C for 5 minutes in the GeneAmp PCR system 9700 (Applied Biosystems). The synthesized cDNA was stored in -20 °C.

All cDNA was checked using a control primer set designed to the cassava ribulose 1-5-bisphosphate carboxylase (RUBISCO) small subunit precursor (rbc1 5'-CTA CTA TGG TGG CTC CGT TC-3' and rbc2 5'-CCG TTC AGT CGG AGA AAC TC-3') (Mak and Ho, 1995). The purpose of these primers was to check that the cDNA had been sythesised and prevent false negative results. The primer set generates a 619-bp product with cassava cDNA (contaminating genomic DNA can be distinguished because it produces a larger PCR product of approximately 800 bp).

4.4.4 PCR and sequencing of amplicons

For the RT-PCR detection of CBSV, the CBSV10 and CBSV11 primers (Monger and Foster., 2001) were used initially, which was followed by the use of the new primer sets designed in this study. PCR was performed in a total reaction volume of 10 µl constituting of 2.0 µl of cDNA, 1.0 µl of 5 pmole/µl of forward and reverse CBSV specific primer and 7.0 µl of de-ionized sterile H_2O. PCR amplification was done using the following program: 94 °C for 5 minutes followed by 40 cycles of 94 °C for 20 seconds, annealing temperature of primer for 20 seconds, 72 °C for 1 minutes and a

final extension at 72 °C for 10 minutes on a GeneAmp PCR system 9700 thermocycler (Applied Biosystems). PCR products were run on a 1.5 % agarose gel electrophoresis. The specificity of the newly designed primers to CBSV was confirmed by sequencing the PCR products. For amplifications with single discrete bands on agarose gel, the products were sequenced using the BIGDYE®Terminator Cycle Sequencing (Applied Biosystems). The reactions were performed in a 10 µl volume comprising of 4 µl of H_2O; 2 µl of diluted PCR products (2.5ng/µl); 2 µl of reverse and forward primer (5 pmoles/µl) done separately; 1.0 µl of BIGDYE and 1.0 µl of 5X sequencing buffer all at manufacturers recommendations. The sequencing reaction was set for 45 cycles using the program: denaturation at 96 °C for 10 seconds; annealing temperature for 5 seconds; extension of 60 °C for 4 minutes and final hold at 4 °C . Purification of PCR products was done by the ethanol precipitation method. One microlitre of 125 mM EDTA was added to each well in the sample plate followed by 1 µl 3 M sodium acetate. 25 µl of chilled absolute ethanol was added and the plate incubated at -20 °C for 1 hour. The plate was then centrifuged at 3,000 rpm for 30 minutes at 4 °C and then inverted and spun at 200 rpm for 2 minutes to remove the supernatant. Hundrend microlitres of 70 % ethanol was added and the plate spun at 3,000 rpm for 15 minutes and inverted again and spun for 2 minutes at 200 rpm. The plate was placed in a thermocycler without the lid and heated at 90 °C for 1 minutes to remove all residual ethanol. The sequenced products were re-suspended in 10 µl injection buffer (HiDi formamide, Applied Biosystems). The samples were then denatured for 3 minutes at 95 °C and placed on ice. The samples were then run on ABI 3730 sequencer (Applied Biosystems). The ABI generated trace files were aligned into contigs with slight modifications of the default conditions using Sequencher software version 4.6 (Gene Codes Corporation, USA). All chromatograms and sequences were visually inspected and sequences edited; only good quality sequences were used further. The identity of the sequences was confirmed using the BLAST search analysis.

4.4.5 Detection limit of diagnostic protocol

The cDNA sythesized from 200 ng of RNA by RT-PCR kit from Bioneer (BIONEER Corporation) was evaluated undiluted, as well as serially diluted with sterile distilled water down to 10^{-6} concentrations. To test the detection limit, 1.0 µl of cDNA was used in each 10 µl PCR reaction as described in the section 4.4.4.

4.4.6 Validation of designed primers on planting materials from cassava multiplication sites

4.4.6.1 Field sampling

The ability of the primers to detect CBSD was validated further by analyzing field samples from cassava multiplication sites in Kenya, Uganda, Rwanda, Democratic Republic of Congo (DRC), Burundi and Tanzania. Three hundred plants were sampled per field using a diamond transect (Figure 4.5). Since there were four sides of the transect, 75 plants were sampled along each side, and these plants were selected at regular intervals calculated by dividing the approximate length of the side by 75. If the length of the side to be sampled was approximately 300m, for example, the interval between each plant to be sampled was 300/75 = 4m. Sampled plants were taken from the left hand side of the transect line (i.e. on the immediate left of the sampler). For uniformity, samples were taken from the tallest shoot of the selected plant, and the central two leaflets picked from the second or third leaf, counting from the top of the selected shoot. Consecutive groups of 10 leaflets were arranged on an A4-sized sheet of newspaper such that the leaflets were not touching one another and this leaflet set comprised one sample. Second sets of ten leaflets were then placed on top of the two newspaper sheets and the same process was repeated until all 30 samples, comprising 300 leaflets from 300 plants, were collected.

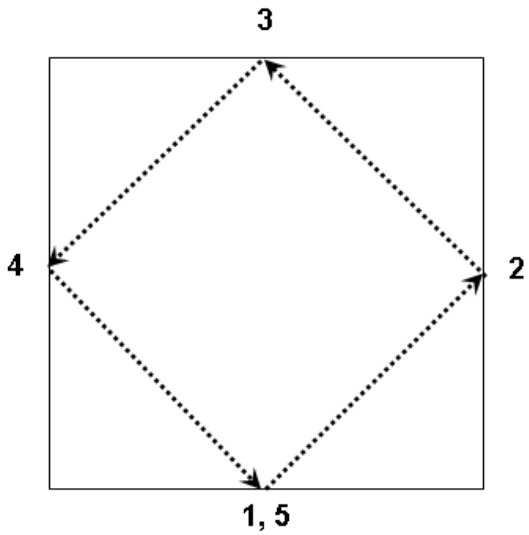

Figure 4.5: Sampling approach for cassava multiplication plots diamond transect

Following the completion of the thirtieth group of ten leaflets, and their covering with a final single sheet of A4-sized newspaper, a CBSD test data card was filled-in and placed on top of the lot and they were all stapled together along the edges, taking care not to damage the leaflets inside. Following the fixing together of the sample and data sheets by stapling, the fastened sets were placed into two herbarium presses and the presses tightened. Sheets for sampled fields were continually addeded to the presses until the presses were full. The samples were brought to the laboratory and store at 4 °C cold room.

4.4.6.2 RNA extraction

Samples for RNA extraction were removed from the cold room and brought to the lab. The samples were removed from the A4 base sized cardboard boxes and placed on a clean bench. The sheets of A4-sized newspaper containing the leaf samples were unstapled taking care not to distort the stack leaves. Ten leaf samples were pulled to

make one RNA extraction sample i.e. a total of 30 RNA prep from 300 leaflets. One punch was made out of the 10 pooled leaflets using a sterile single disk punch and the 10 discs transferred into one 1.2 ml strip tube. Samples were prepared the same way until two sets of 96 strips format were made. Two Rnase, Dnase Pyrogen free stainless steel grinding balls were placed into the strip tubes using the grinding ball dispenser. RNA was extracted using the AccuPrep® Plant RNA virus Extraction Kit (BIONEER Corporation). The extraction buffer from the kit was added into the strip tubes. The samples were then ground using the Geno/Grinder® 2000 (SamplePrep) for 10-20 minutes until they formed a uniform homogenate. The homogenate was transferred into 96 well format of the AccuPrep Plant RNA virus Extraction Kit and RNA extracted according to the manufacturer's instructions. After successful RNA extraction, the quantity and quality of extracted RNA was checked using Nanodrop® ND-1000 spectrophotometry (NanoDrop Technologies, USA). Extracted RNA samples were stored at -80 °C.

cDNA synthesis and diagnostic PCR were carried out with primer CBSV R1R10 as described previously in sections 4.4.3 and 4.4.4.

4.5 Results

4.5.1 Optimization of the RT-PCR technique

Eighteen primers were designed from the published sequences (GenBank accession No.s AY008440, AY008441, AY008442, AY007597, AF311052 and AF311053) and are illustrated in Table 4.2. These primers were designed using the PrimerSelect computer program (Lasergene, USA) and together generate products from within the coding region of the virus coat protein. The primers were used for PCR amplification of CBSVs and 22 out of the 30 primers combinations (73.3 %) gave PCR products with clear discrete bands (Figure 4.6).

One primer set, CBSV-R1R10 (357-bp product), proved most consistent in generating a distinct PCR product band, when analysed on ethidium bromide stained agarose gels. The specificity of this primer was confirmed by sequencing the RT-PCR products. The sequences of the PCR amplicons had hits on CBSV sequences on the NCBI GenBank based on BLASTn analysis. Most sequences for 'Kiberiti Mweusi' isolates were not clean and thus comparison between the two isolates was not done. However, clean sequences from the CBSVR1-R10 fragment of 'Guzo' isolate showed variation with other CBSV sequences in the NCBI. Therefore this loc can be adequately utilized for diversity analysis of CBSV isolates and more specific primers designed. There was also evidence of variation within the priming sites of the above primers a probable reason for the poor amplification and sequencing for 'Kiberiti Mweusi' isolate with most primers. More regions that are common for all isolates need to be identified for a good diagnostic primer design for all CBSV isolates.

Table 4.2: Primers used for the amplification of CBSVs

Primer name	Primer sequence (5'-3')	Reference
CBSV10	ATCAGAATAGTGTGAACTGCTGG	Monger and Foster (2001)
CBSV11	ATGCTGGGGTACAGACAAG	Monger and Foster (2001)
CBSV-F1	AGGAGAAACATAAAAAGACAAGGTTCAG	This study
CBSV-F2	GCTTCCGGAGTTGAAGTTGAGAA	This study
CBSV-F3	GGATACCGCGTTCGCAAAGAGA	This study
CBSV-F4	AAGCAATCCTAATGTATCAGAATAGTGTG	This study
CBSV-F5	CAGCCAAGCAATCCTAATGTATCAGA	This study
CBSV-F6	AAAATAAGTGGTTGGGTATTGACTTCCTA	This study
CBSV-F7	AGCTGCTCCTGGTGACGATAATAA	This study
CBSV-F8	TCTAGCGTACGTGCCTCCATCA	This study
CBSV-R1	ACAGGGTGAGGATATTGATGATTC	This study
CBSV-R2	CTTTATCCACATTATTATCGTCACCAGG	This study
CBSV-R3	GTGTGCCTTTCTTTATCCACATTATTATC	This study
CBSV-R4	CCGCGTAGTGTGCCTTTCTTTATC	This study
CBSV-R5	TCACAACCAAAATATAACAAGCTACCTAAA	This study
CBSV-R6	TTTCCTCATTGTCGGGAGAGCA	This study
CBSV-R7	TGCTTCGGCTAGGAAGTCAATACC	This study
CBSV-R8	CGGCATTTCGGGCTGCTTTTA	This study
CBSV-R9	GCTCAACAGCTCTCCACGATTTCT	This study
CBSV-R10	CGCGTAGTGTGCCTTTCTTTAT	This study

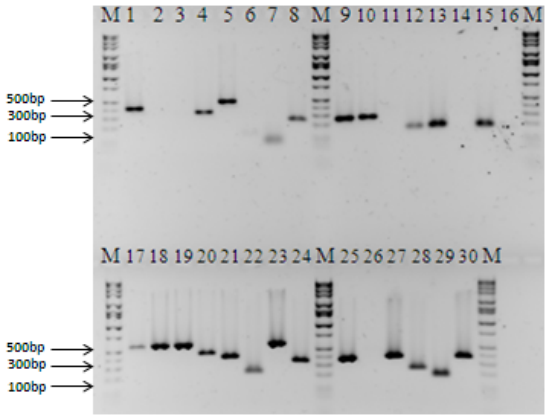

Figure 4.6: Inverted gel image of primer combinations for PCR amplifications of CBSV cDNA. 1-F1R9, 2-FR2, 3-F2R8, 4-F3R6, 5-F3R6, 6-F3R1, 7-F4R1, 8-F4R8, 9-F5R2, 10-F5R3, 11-F5R4, 12-F5R5, 13-F1R3, 14-F1R4, 15-F1R5, 16-F1R10, 17-F1R5, 18-F8R2, 19-F8R3, 20-F8R4, 21-F8R1, 22-F8R8, 23-F8R9, 24-F8R10, 25-R1R3, 26-R1R5, 27-R1R4,28-R1R9, 29-R1R8 and 30-R1R10. M - 100 bp molecular size ladder.

4.5.2 Detection limit of diagnostic protocol

A serial dilution of the cDNA generated from 200 ng RNA with the first strand Reverse Transcriptase PCR was performed (Figure 4.7). The greatest concentration of product was produced with the undiluted cDNA (lane 1). Primers CBSV-R1R10 and CBSV-F8R4 gave a higher detection limit as compared to previous primer CBSV 10/11(Monger and Foster., 2001). A product band was still visible with a 10^{-5} fold dilution of cDNA (lane 6) with Primer CBSV-R1R10 while primer CBSV-F8R4 gave amplification with a 10^{-3} fold dilution (lane 4). However, primer CBSV-10/11(Monger and Foster., 2001) was only able to give a band up to a 10^{-2} fold dilution of cDNA (lane 3).

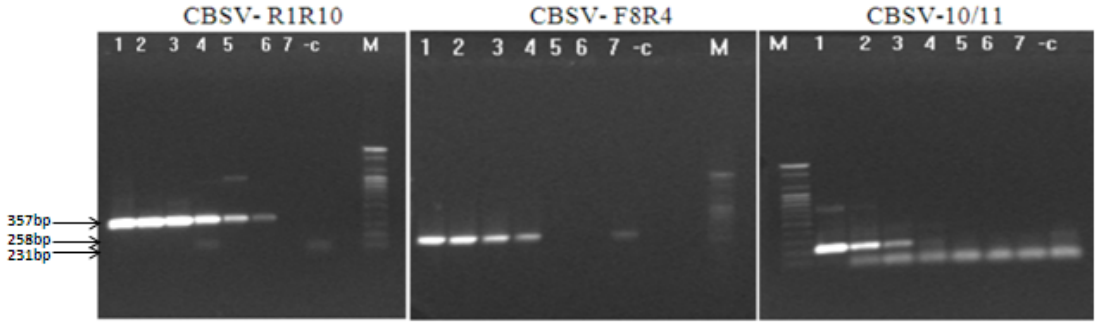

Figure 4.7: Agarose gel electrophoresis of RT-PCR test of serial dilution of the cDNA of CBSV-infected cassava (Gozo isolate) with primer set CBSV-R1R10 (357bp), CBSV- F8R4 (258bp) and CBSV 10/11 (231bp Monger and Foster 2001). Serial dilutions lane 1(undiluted cDNA), lane 2 (10^{-1}), lane 3 (10^{-2}), lane 4 (10^{-3}),lane 5 (10^{-4}),lane 6 (10^{-5}) and lane 7 (10^{-6}). -C - negative control from clean cassava, M-100bp Molecular marker

4.5.3 Validation of new primers for the detection of CBSV

The efficiency of the new primers for the diagnosis of CBSV was tested on field-collected samples from Kenya, Uganda, Rwanda, Democratic Republic of Congo (DRC), Burundi and Tanzania. Cassava leaves were collected from cassava multiplication sites for testing the presence of CBSV. Some leaves were randomly chosen because they showed atypical symptoms not previously associated with CBSV and others which were without symptoms were chosen to see if the material was virus-free or latently infected. The superiority of the new primers was observed further when the viruses were detected before symptom expression on plants in the field (Figure 4.8.)

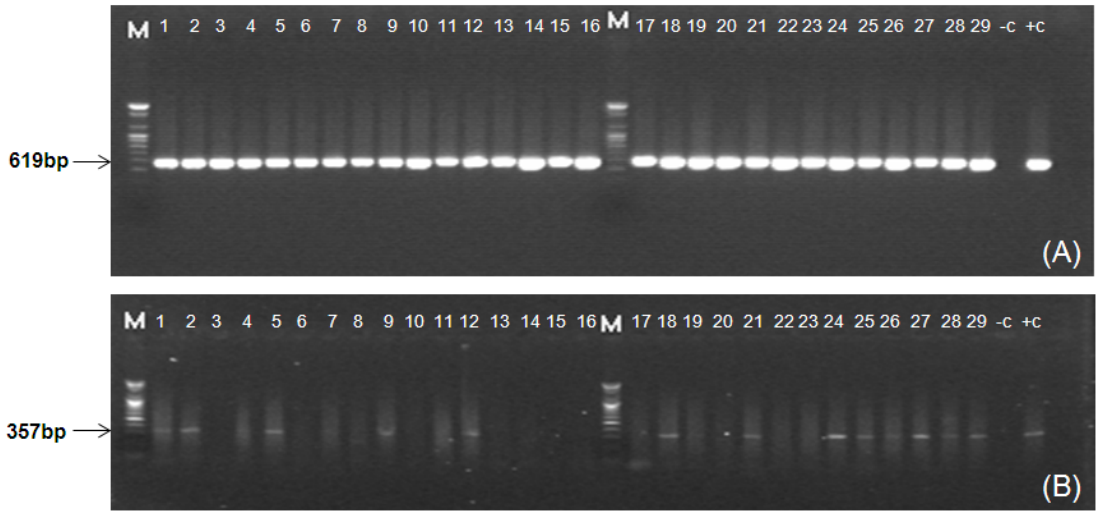

Figure 4.8: A representative gel image of diagnosis of CBSV from CBSD symptomless cassava plants sampled from field multiplication sites. Panel (A), amplification of cDNA with control RUBISCO primer rbc1/2. Panel (B), CBSV diagnosis with primer CBSVR1-R10. +c - positive control from known CBSV infected cassava plant, -c - negative control from clean cassava plant and M - 100 bp molecular size ladder.

4.6 Discussion

A robust, sensitive RT-PCR protocol and a cost-effective field-sampling procedures for CBSD diagnosis were established in this study for an economically important food crop, cassava. Several isolates of the two CBSVs have been fully sequenced, some of which cannot be detected using the previously described CBSV10 and CBSV11 primers (Monger and Foster., 2001) because of the mismatches at the 3' ends of both the primers to the corresponding viral sequences. Other protocols have also been developed for the detection of mixed infections of cassava mosaic virus (CMV) and CBSVs which are becoming an increasingly common feature, following the recent spread of CBSD into CMD endemic areas in parts of eastern Africa (Abarshi et al., 2012).

In this study coat protein sequences were utilized to design diagnostic primers for the specific identification of cassava brown streak viruses (CBSVs). Better diagnostic primers were designed based on six published coat protein (CP) sequences on the NCBI (GenBank accession Numbers AY008440, AY008441, AY008442, AY007597, AF311052 and AF311053). Eighteen primers were developed for the specific detection of CBSVs. One primer pair (CBSVR1-R10) proved to be highly sensitive and can detect up to 10^{-5} fold dilution of cDNA synthesized from 200 ng RNA as compared to primers described by Monger and Foster. (2001) which could detect only up to 10^{-2} fold dilution of same cDNA. As well as improving the efficiency of virus detection, the primers have the potential to be highly efficient for the detection of early infections by CBSVs, which will be highly valuable for epidemiological studies, quarantine testing of symptomless plants and for selecting clean (non-infected) planting material.

Since there are new and damaging outbreaks of CBSVs in the Great Lakes countries, measures are needed to minimize the impact of CBSD. These might include: regular surveillance and virus testing, avoidance of inadvertent virus spread by restricting long-distance dissemination of planting material and identification of suitable resistant

germplasm. Primer pair CBSV-R1R10 was utilized for the diagnosis of CBSVs from field-collected samples of cassava multiplication sites in Kenya, Uganda, Rwanda, Democratic Republic of Congo (DRC), Burundi and Tanzania (Figure 4.8). The new primers were highly sensitive and the viruses could be detected before symptom expression on plants in the field. The field sampling procedures for testing large numbers of samples established from this study will also minimize sample handling, greatly reducing the inherent risk of cross contamination and associated costs.

Chapter 5

GENERAL DISCUSSION, CONCLUSSION AND RECOMMENDATIONS

5.1 Discussion

Efficient laboratory diagnosis of plant pathogens is increasingly important in disease management. Methods to directly detect nucleic acids of microbial pathogens should be rapid, sensitive, and may succeed when culturing the organism fails. The use of PCR for detection of plant and animal pathogens has become widespread (Henson and French, 1993) and is most commonly used now. DNA based methods provide powerful tools to identify microorganisms with high sensitivity and specificity. Successful detection of pathogens using PCR depends upon the specificity of primers.

In the development of specific primers for diagnosis of BXW and gray leaf spot (GLS) disease of maize, intergenic regions were mainly targeted as they are more variable between species than the conserved coding regions. Although not transcribed, intergenic regions have been shown to play a role in the expression of adjacent genes and there is increasing evidence that they contain important control sequences. Correspondingly, coding regions of DNA are far less variable than intergenic regions. Such a discrepancy is expected considering that the coding regions are parts of alleles that instruct the synthesis of proteins. Intergenic regions are not transcribed into proteins and therefore they do not regulate cellular structure and function in such a manner that selection pressures do not directly influence them. The size of genes, and the related sequence complexity (for intergenic regions and unique sequences) are specific to the genus or

species and thus can be exploited for designing specific primers for diagnosis. Accordingly, intergenic size and complexity can greatly vary and therefore can be used to discriminate pathovar and strains within species (Chocholová et al., 2008). Moreover, unlike genes, that may be similar in a wide variety of species, intergenic regions show a far greater diversity or difference among species. Although sequences that serve as gene enhancers and silencers occur in intergenic regions, they do not directly contribute to protein synthesis. Correspondingly, intergenic regions do not directly contribute to the phenotype of the individual and changes in intergenic regions would not be expected to affect the fitness of the organism.

In this study the first whole genome of *Xanthomonas campestris* pv. *musacearum* Kenyan strain was sequenced and deposited at DDBJ/EMBL/GenBank under the accession AGFQ00000000. This data (including updated annotations) also can be downloaded from URL

http://bioinformatics.iita.org/projects/banana/. The available unique *Cercospora zeae-maydis* EST sequences from the NCBI GeneBank were utilized in selecting regions for designing Czm and Cz specific PCR primers.

5.2 Conclusion

In this study, simple and straight forward approaches in the development of reliable and sensitive methods for the PCR based detection of BXW, GLS disease of maize and CBSD have been described. Genome sequencing and comparative genomics, and bioinformatics were utilized to provide the specific target sequence for the diagnostics tools for these diseases.

Using a combination of sequencing and bioinformatics approaches, 44 *Xcm* specific primer pairs (Figure 2.4) were developed, tested and validated for the specific detection of BXW. Of all the primers tested, seven primer pairs (Xcm12, 35, 36, 38, 44, 47 and

48) were found to be specific to *Xcm* and can thus be used for routine diagnosis. PCR amplification was also successfully used to detect *Xcm* directly from bacterial cells without extracting DNA. Using this method it was possible to detect up to 10^3 CFU/ml of *Xcm* whole cells by any of the seven specific primers, in comparison to a detection limit of 10^4 to 10^5 CFU/ml reported by Lewis et al. (2009). These findings demonstrate the ability of the primers to be used as a diagnostic tool for *Xcm* from symptomless infected banana plants. The PCR-based diagnostic tool described is easy to set up and would be useful in *Xcm* epidemiological studies.

For the development of diagnostic protocol for gray leaf spot (GLS) disease of maize, 16 Czm and Cz PCR primers were designed from unique *Cercospora zeae-maydis* EST sequences. Out of these 11 primer pairs worked well when tested with both Czm and Cz DNA. Six primer pairs were designed from the Malazy gene and 4 out of these were able to give amplification products with (*Cercospora zeae-maydis*) and no amplification product with all other samples of group II (C. *zeina*). These can be used for the specific diagnosis of *Cercospora zeae-maydis* and not *Cercospora zeina*. Primer Czm47 designed from intergenic regions is also able to distinguish the two *Cercospora* species i.e. group I isolates (*Cercospora zeae-maydis*) and group II (C. *zeina*). Multiplex primer sets Czm47/49 and Czm47/51 produces two bands with C. *zeina* and one band with C. *zeae-maydis*. These primers can therefore be used to easily diagnose the two groups of *Cercospora* species. One primer pair Czm9, had the highest sensitivity giving a PCR product from DNA concentration of as low as 10^{-3} ng. At this concentration the maize plants are symptomless for gray leaf spot (GLS) disease of maize and the primer can therefore be used to detect the pathogen before visual symptoms appear allowing early disease management strategies to be implimented.

In this study coat protein sequences were utilized to design diagnostic primers for the specific identification of Cassava brown streak viruses (CBSVs). Better diagnostic primers were designed based on six published coat protein (CP) sequences on the

NCBI (GenBank accession Numbers AY008440, AY008441, AY008442, AY007597, AF311052 and AF311053). Diagnostic primers were designed according to the most conserved regions of the six CP sequences or in such a way that the 3' end of the primer was positioned on a region that is conserved across the six sequences. Twenty two out of the 30 primer combinations (73.3 %) of the 18 primer set designed gave good amplification with clear discrete bands. One primer set, CBSV-R1R10 proved most consistent in generating a strong PCR product band. Primers CBSV-R1R10 and CBSV-F8R4 gave a high detection limit where CBSV was detected from cDNA synthesised from 200 ng RNA. There was PCR amplification from this cDNA at 10^{-5} fold dilution for Primer CBSV-R1R10 and at 10^{-3} fold dilution for primer set CBSV-F8R4 compared to previous primer CBSV-10/11 by Monger and Foster. (2001) which only give a band up to a 10^{-2} fold dilution. The efficiency of the new primers for the diagnosis of CBSV was tested on field-collected samples from Kenya, Uganda, Rwanda, Democratic Republic of Congo (DRC), Burundi and Tanzania. The superiority of the new primers was observed further when the viruses were detected before symptom expression on plants in the field.

5.3 Recommendations

The developed molecular diagnostics tools should be further applied for better disease management, and also implemented in marker assisted selection of lines in breeding programs or genetic engineered crops for disease resistant. These tools can be used for routine diagnosis of the pathogens in the field, and in planting materials (cuttings, suckers and seeds).

The approaches demonstrated in this study could potentially be exploited for designing diagnostic PCR primers for other species or pathovars and also for other plant bacterial, fungal and viral pathogens. Conserved genes present in all pathogens of a particular

genus could be targeted for designing primers to identify all the species. Hence sequences from unique genes, those from intergenic regions and coat protein can be used to develop primers for specific identification of pathogens at genus, species or pathovar levels. Furthermore, sequence variability within the intergenic regions and coat protein for viruses could be utilized for diversity studies within species.

The full genome of *Xcm* from this study will provide insights into the physiological and biochemical processes of the pathogen particularly in the protein secretion system and the disease specificity. However, to reliably unravel these processes and subsequently utilize the primary genetic material in downstream studies, accurate annotation of the genome elements is needed. In addition to diagnostics, accurate annotation will be essential for reliable identification of candidate targets for developing resistant banana cultivars, variation studies and comparative genomics among others.

References

Abarshi, M. M., Mohammed, I. U., Jeremiah, S. C., Legg, J. P., Kumar, P. L., Hillocks, R. J., and Maruthi, M. N. (2012). Multiplex rt-pcr assays for the simultaneous detection of both rna and dna viruses infecting cassava and the common occurrence of mixed infections by two cassava brown streak viruses in east africa. *Journal of Virological Methods*, 179(1):176–184.

Abdel-Baky, S., Allam, K., and Giese, R. W. (1993). Detection of electrophore-labeled dna and albumin by gas chromatography: labile amide electrophoric release tags. *Analytical Chemistry*, 65(4):498–499.

Adams, M. J., Antoniw, J. F., and Beaudoin, F. (2005). Overview and analysis of the polyprotein cleavage sites in the family potyviridae. *Molecular Plant Pathology*, 6(4):471–487.

Alvarez, C. V., Mallo, F., Burguera, B., Cacicedo, L., Dieguez, C., and Casanueva, F. F. (1991). Evidence for a direct pituitary inhibition by free fatty acids of in vivo growth hormone responses to growth hormone-releasing hormone in the rat. *Neuroendocrinology*, 53(2):185–189.

Anandalakshmi, R., Pruss, G. J., Ge, X., Marathe, R., Mallory, A. C., Smith, T. H., and Vance, V. B. (1998). A viral suppressor of gene silencing in plants. *Proceedings of the National Academy of Sciences USA*, 95(22):13079–13084.

Anonymous (1989). Ministry of agriculture report. Technical report, Government Printer, Nairobi, Kenya.

Aritua, V., Parkinson, N., Thwaites, R., Heeney, J., Jones, D., Tushemereirwe, W., Crozier, J., Boa, E., Stead, D., and Smith, J. (17-20th July 2006.). Molecular characterization of xanthomonas campestris pv. musacearum. In *In: Proceedings of the Fourth International Bacterial Wilt Symposium, Central Science Laboratory, York, UK. Page 59 in: http://www.ibws 2006.org.uk/book*$_r$*ev*

Armengaud, J. (2009). A perfect genome annotation is within reach with the proteomics and genomics alliance. *Current Opinion in Microbiology*, 12(3):292–300.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. (2000). Gene ontology: tool for the unification of biology. the gene ontology consortium. *Nature Genetics*, 25(1):25–29.

Authority, E. P. Z. (2005). Grain production in kenya. Technical report, Export Processing Zones Authority. 20pp.

Bamaga, M. S., Wright, D. J., and Zhang, H. (2003). An assessment of a multiplex pcr assay for differentiating clinically important mycobacteria based on pnca gene variation. *Molecular and Cellular Probes*, 17(2-3):69–75.

Barker, I. (1996). Serological methods in crop protection. In *In Diagnostics in Crop Production, British Crop Protection Council Symposium Proceedings No. 65, pp. 13-22. Ed. G Marshall. Farnham, UK: British Crop Protection Council.*

Beckman, P. and Payne, G. (1982). External growth, penetration, and development of cercospora zeae-maydis and lesion development in corn. *Phytopathology*, 73:286–289.

Bellotti, A. C. and Ariass, B. (2001). Host plant resistance to whiteflies with emphasis on cassava as a case study. *Crop Protection*, 20:813–823.

Bentsink, L., Leone, G., Van Beckhoven, J., van Schijndel, H., Van Gemen, B., and Van der Wolf, J. M. (2002). Amplification of rna by nasba allows direct detection of viable cells of ralstonia solanacearum in potato. *Journal of Applied Microbiology*, 93:647–655.

Berger, R. (1973). Early blight of celery: analysis of disease spread in florida. *Phytopathology*, 63:116–165.

Bodenteich, A., Mitchell, L. G., Polymeropoulos, M. H., and Merril, C. R. (1992). Dinucleotide repeat in the human mitochondrial d-loop. *Human Molecular Genetics*, 1(2):140.

Bokanga, M. (1994). Processing of cassava leaves for human consumption. acta hort. *Acta Horticulture*, 375:203207.

Boonham, N. and Barker, I. (1998). Strain specific recombinant antibodies to potato virus y potyvirus. *Journal of Virological Methods*, 74:193–199.

Briese, T., Palacios, G., Kokoris, M., Jabado, O., Liu, Z., Renwick, N., Kapoor, V., Casas, I., Pozo, F., Limberger, R., Perez-Brena, P., Ju, J., and Lipkin, W. I. (2005). Diagnostic system for rapid and sensitive differential detection of pathogens. *Emerging Infectious Diseases*, 11(2):310–313.

Brigneti, G., Voinnet, O., Li, W. X., Ji, L. H., Ding, S. W., and Baulcombe, D. C. (1998). Viral pathogenicity determinants are suppressors of transgene silencing in nicotiana benthamiana. *EMBO Journal*, 17(22):6739–6746.

Bttner, D. and Bonas, U. (2010). Regulation and secretion of xanthomonas virulence factors. *FEMS Microbiology Reviews*, 34(2):107–133.

Carter, B., Reeder, R., Mgenzi, S., Kinyua, Z., Mbaka, J., Doyle, K., Nakato, V., Mwangi, M., Beed, F., Aritua, V., Lewis Ivey, M., Miller, S., and Smith, J. (2009). Identification of xanthomonasvasicola (formerly x. campestrispv. musacearum), causative organism of banana xanthomonas wilt, in tanzania, kenya and burundi. *New Disease Reports online [http://www.bspp.org.uk/publications/new-disease-eports/ndr.php?id=019025].*

Carzaniga, R., Fiocco, D., and Bowyer, P.O. andConnell, R. J. (2002). Localization of melanin in conidia of alternaria alternate using phage display antibodies. *Molecular Plant-Microbe Interactions*, 15:216–224.

Chagnon, A., Marlier, S., Carli, P., Vauterin, G., Yao, N., and Paris, J. F. (1991). [pericardial asbestosis. a case]. *Presse Med*, 20(43):2218–2219.

Chan, J. W. and Goodwin, P. H. (1999). The molecular genetics of virulence of xanthomonas campestris. *Biotechnology Advances*, 17(6):489–508.

Chase, Jr, C., Del Vecchio, R. P., Smith, S. B., and Randel, R. D. (1992). In vitro metabolism of glucose by bovine reproductive tissues obtained during the estrous cycle and after calving. *Journal of Animal Science*, 70(5):1496–1508.

Chen, H., Lee, M.-H., Daub, M. E., and Chung, K.-R. (2007). Molecular analysis of the cercosporin biosynthetic gene cluster in cercospora nicotianae. *Molecular Microbiology*, 64(3):755–770.

Chocholová, E., Jirk?, M., and Lukes, J. (2008). A diagnostic assay based on variable intergenic region distinguishes between leishmania donovani and leishmania infantum. *Folia Parasitol (Praha)*, 55(1):75–78.

Chung, B. Y.-W., Miller, W. A., Atkins, J. F., and Firth, A. E. (2008). An overlapping essential gene in the potyviridae. *Proceedings of the National Academy of Sciences USA*, 105(15):5897–5902.

Clark, M. and Adams., A. (1977). Characteristics of the microplate method of enzyme-linked immunosorbent assay for the detection of plant viruses. *Journal of General Virology*, 34:475–483.

Coiras, M. T., Aguilar, J. C., Garca, M. L., Casas, I., and Prez-Brea, P. (2004). Simultaneous detection of fourteen respiratory viruses in clinical specimens by two multiplex reverse transcription nested-pcr assays. *Journal of Medical Virology*, 72(3):484–495.

Coletta-Filho, H. D., Takita, M. A., Souza, A. A., Neto, J. R., Destéfano, S. A. L., Hartung, J. S., and Machado, M. A. (2006). Primers based on the rpf gene region provide improved detection of xanthomonas axonopodis pv. citri in naturally and artificially infected citrus plants. *Journal of Applied Microbiology*, 100(2):279–285.

Coletta-Filho HD, T. M. and deSouza (2005). Primers based on the rpf gene region provide improved detection of xanthomonas axonopodis pv. citri in naturally and artificially infected citrus plants. *Journal of Applied Microbiology*, 100:279–85.

Colinet, D., Kummert, J., and Lepoivre., P. (1998). The nucleotide sequence and genome organization of the whitefly transmitted sweet potato mild mottle virus: a close relationship with members of the family potyviridae. *Virus Research*, 53:187–196.

Cook, N. (2003). The use of nasba for the detection of microbial pathogens in food and environmental samples. *Journal of Microbiological Methods*, 53(2):165–174.

Cooke, D. and Scott, R. (1993.). *The Sugar Beet Crop.* London: Chapman and Hall.

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

Cubero, J. and Graham, J. H. (2002). Genetic relationship among worldwide strains of xanthomonas causing canker in citrus species and design of new primers for their identification by pcr. *Applied and Environmental Microbiology*, 68(3):1257–1264.

de Nazareno, N., Lipps, P., and Madden, L. (1992). Survival of cercospora zeae-maydis in corn residues in ohio. *Plant Disease*, 76:560–564.

de Nazareno, N., Madden, L., and Lipps, P. (1993). Characterization of gray leaf spot epidemics of maize. *Journal of Plant Diseases and Protection*, 100:410–425.

Degefu, Y. (2008). Molecular diagnostic technologies and end user applications: potentials and challenges. *Agrifood Research Finland, Biotechnology and Food Research, Seed Potato Biotechnology, Tutkimusasemantie 15, 92400. Published at http://www.smts.fi.*

Denny, Jr, F. (1995). The clinical impact of human respiratory virus infections. *American Journal of Respiratory and Critical Care Medicine*, 152(4 Pt 2):S4–12.

Devos, D. and Valencia, A. (2001). Intrinsic errors in genome annotation. *Trends in Genetics*, 17(8):429–431.

Dewey, F. M. and Thornton, C. R. (1995). Fungal immunodiagnostics in plant agriculture. *In New Diagnostics in Crop Sciences, Biotechnology in Agriculture No. 13, pp. 151-170. Eds J H Skerritt and R Appels. Wallingford, UK: CAB International.*

Dewey, F. M., Thornton, C. R., and Gilligan, C. A. (1997). The use of monoclonal antibodies to detect, quantify and visualise fungi in soils. *Advances in Botanical Research*, 24:275–308.

Dunkle, L. D. and Levy, M. (2000). Genetic relatedness of african and united states populations of cercospora zeae-maydis. *Phytopathology*, 90(5):486–490.

Elnifro, E. M., Ashshi, A. M., Cooper, R. J., and Klapper, P. E. (2000). Multiplex pcr: optimization and application in diagnostic virology. *Clinical Microbiology Reviews*, 13(4):559–570.

Fan, J., Henrickson, K. J., and Savatski, L. L. (1998). Rapid simultaneous diagnosis of infections with respiratory syncytial viruses a and b, influenza viruses a and b, and human parainfluenza virus types 1, 2, and 3 by multiplex quantitative reverse transcription-polymerase chain reaction-enzyme hybridization assay (hexaplex). *Clinical Infectious Diseases*, 26(6):1397–1402.

FAOSTAT (2005). Agriculture data. http://apps.fao.org. Technical report, FAO.

Fauquet, C. M., M. A, Mayo, J., Maniloff, U. D., and L. A, B. (2005). Virus taxonomy: the eighth report of the international committee on taxonomy of viruses. *Elsevier Academic Press, San Diego, CA*.

Fessehaie, A., De Boer, S. H., and Lévesque, C. A. (2003). An oligonucleotide array for the identification and differentiation of bacteria pathogenic on potato. *Phytopathology*, 93(3):262–269.

Foster, S. J., Ashby, A. M., and Fitt, B. D. L. (2002). Improved pcrbased assays for pre-symptomatic diagnosis of light leaf spot and determination of mating type of pyrenopeziza brassicae on winter oilseed rape. *European Journal of Plant Pathology*, 108:379–383.

Fraaije, B. A., Lovell, D., Coelho, J. M., Baldwin, S., and Hollomon, D. W. (2001). Pcr-based assays to assess wheat varietal resistance to blotch (septoria tritici and stagonospora nodorum) and rust (puccinia striiformis and puccinia recondita) diseases. *European Journal of Plant Pathology*, 107:905–917.

Franke-Whittle, I. H., Klammer, S. H., and Insam, H. (2005). Design and application of an oligonucleotide microarray for the investigation of compost microbial communities. *Journal of Microbiological Methods*, 62(1):37–56.

Gilliland, G., Perrin, S., Blanchard, K., and Bunn, H. F. (1990). Analysis of cytokine mrna and dna: detection and quantitation by competitive polymerase chain reaction. *Proceedings of the National Academy of Sciences USA*, 87(7):2725–2729.

Goodwin, S. B., Dunkle, L. D., and Zismann, V. L. (2001). Phylogenetic analysis of cercospora and mycosphaerella based on the internal transcribed spacer region of ribosomal dna. *Phytopathology*, 91(7):648–658.

Gough, K. C., Li, Y., Vaughan, T. J., Williams, A. J., Cockburn, W., and Whitelam, G. C. (1999). Selection of phage antibodies to surface epitopes of phytophthora infestans. *Journal of Immunological Methods*, 228(1-2):97–108.

Griep, R. A., van Twisk, C., and Schots, A. (1999). Selection of beet necrotic yellow vein virus specific single- chain fv antibodies from a semi-synthetic combinatorial antibody library. *European Journal of Plant Pathology*, 105:147–156.

Griep, R. A., van Twisk, C., van Beckhoven, J., van der Wolf, J. M., and Schots, A. (1998). Development of specific recombinant monoclonal antibodies against the lipopolysaccharide of ralstonia solanacearum race 3. *Phytopathology*, 88:795–803.

Grndahl, B., Puppe, W., Hoppe, A., Khne, I., Weigl, J. A., and Schmitt, H. J. (1999). Rapid identification of nine microorganisms causing acute respiratory tract infections by single-tube multiplex reverse transcription-pcr: feasibility study. *Journal of Clinical Microbiology*, 37(1):1–7.

Gwinn, K., H., C.-S. M., and B., R. B. (1991). Tissue print-immunoblot, an accurate method for the detection of acremonium coenophialum in tall fescue. *Phytopathology*, 81:747–748.

Hahn., S. K., Terry., E. R., and Leuschner, K. (1980). Breeding cassava for resistance to cassava mosaic disease. *Euphytica*, 29(3):673–683.

Hampton, R., Ball, E., and De Boer, S. (1990). *Serological methods for detection and identification of viral and bacterial plant pathogens: a laboratory manual.* American Phytopathological Society, St. Paul, Minn.

Harding, N. E., Cleary, J. M., Cabañas, D. K., Rosen, I. G., and Kang, K. S. (1987). Genetic and physical analyses of a cluster of genes essential for xanthan gum biosynthesis in xanthomonas campestris. *Journal of Bacteriology*, 169(6):2854–2861.

Hartung, J., P., P. O., I., V., and A., A. (1996). Rapid and sensitive colorimetric detection of xanthomonas axonopodis pv citri by immunocapture and a nestedpolymerase chain reaction assay. *Phytopathology*, 86:95–101.

Hartung, J. S., Daniel, J. F., and Pruvost, O. P. (1993). Detection of xanthomonas campestris pv. citri by the polymerase chain reaction method. *Applied and Environmental Microbiology*, 59(4):1143–1148.

Hauben, L., Vauterin, L., Swings, J., and Moore, E. R. (1997). Comparison of 16s ribosomal dna sequences of all xanthomonas species. *International Journal of Systematic Bacteriology*, 47(2):328–335.

Henegariu, O., Heerema, N. A., Dlouhy, S. R., Vance, G. H., and Vogt, P. H. (1997). Multiplex pcr: critical parameters and step-by-step protocol. *Biotechniques*, 23(3):504–511.

Henson, J. M. and French, R. (1993). The polymerase chain reaction and plant disease diagnosis. *Annual Review of Phytopathology*, 31:81–109.

Hermansson, A. and Lindgren, P. E. (2001). Quantification of ammonia-oxidizing bacteria in arable soil by real-time pcr. *Applied and Environmental Microbiology*, 67(2):972–976.

Hillocks, R. and DL, J. (2003). Cassava brown streak disease: a review of present knowledge and research needs. *International Journal of Pest Management*, 49:225–340.

Hillocks, R., Thresh., J., and Bellotti, A. (2002). *Cassava in Africa*. CABI Publishing, Wallingford, UK.

Jacobi, V., Bachand, G. D., Hamelin, R. C., and Castello, J. D. (1998). Development of a multiplex immunocapture rt-pcr assay for detection and differentiation of tomato and tobacco mosaic tobamoviruses. *Journal of Virological Methods*, 74(2):167–178.

Janssen, D., Martín, G., Velasco, L., Gómez, P., Segundo, E., Ruiz, L., and Cuadrado, I. M. (2005). Absence of a coding region for the helper component-proteinase in the genome of cucumber vein yellowing virus, a whitefly-transmitted member of the potyviridae. *Archives of Virology*, 150(7):1439–1447.

Jennings, D. L. (1960). Observations on virus diseases of cassava in resistant and susceptible varieties. ii. brown streak disease. *Empire Journal of Experimental Agriculture*, 28:261–269.

Kasschau, K. D. and Carrington, J. C. (1998). A counterdefensive strategy of plant viruses: suppression of posttranscriptional gene silencing. *Cell*, 95(4):461–470.

Katzen, F., Ferreiro, D. U., Oddo, C. G., Ielmini, M. V., Becker, A., Phler, A., and Ielpi, L. (1998). Xanthomonas campestris pv. campestris gum mutants: effects on xanthan biosynthesis and plant virulence. *Journal of Bacteriology*, 180(7):1607–1617.

Kayobyo, G., Aliguma, L., Omiat, G., Mugisha, J., and Benin, S. (2005). Impact of bxw on household livelihoods in uganda. In *Paper presented at the workshop "Assessing the impact of the banana bacterial wilt (Xanthomonascampestrispv. musacearum) on household livelihoods in East Africa", held on Dec. 20th 2005 in Kampala, Uganda.*

Kennedy, R., Wakeham, A. J., and Cullington, J. E. (1999). Production and immunodetection of ascospores of mycosphaerella brassicicola: ringspot of vegetable crucifers. *Plant Pathology*, 48:297–307.

Kinyua, Z., SMITH., J., KIBATA., G., SIMONS, S., and Langat, B. (2010). Status of grey leaf spot disease in kenyan maize production ecosystems. *African Crop Science Journal*, 18:183–194.

Knoll, S., Vogel, R. F., and Niessen, L. (2002). Identification of fusarium graminearum in cereal samples by dna detection test strips. *Letters in Applied Microbiology*, 34(2):144–148.

Köhler, G. and Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature*, 256(5517):495–497.

Kokoris, M., Dix, K., Moynihan, K., Mathis, J., Erwin, B., Grass, P., Hines, B., and Duesterhoeft, A. (2000). High-throughput snp genotyping with the masscode system. *Journal of Molecular Diagnostics*, 5(4):329–340.

Kuflu, K. M. and Cuppels, D. A. (1997). Development of a diagnostic dna probe for xanthomonads causing bacterial spot of peppers and tomatoes. *Applied and Environmental Microbiology*, 63(11):4462–4470.

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

Lazer, E. S., Wong, H. C., Wegner, C. D., Graham, A. G., and Farina, P. R. (1990). Effect of structure on potency and selectivity in 2,6-disubstituted 4-(2-arylethenyl)phenol lipoxygenase inhibitors. *Journal of Medicinal Chemistry*, 33(7):1892–1898.

Lecoq, H., Desbiez, C., Delécolle, B., Cohen, S., and Mansour, A. (2000). Cytological and molecular evidence that the whitefly-transmitted cucumber vein yellowing virus is a tentative member of the family potyviridae. *Journal of General Virology*, 81(Pt 9):2289–2293.

Lee, S. and Taylor, J. (1990.). Isolation of dna from fungal mycelia and single spores. in pcr protocols. *A guide to methods and applications M A Innis*, pages pp. 282–287. Eds D H.

Legg, J. P., French, R., Rogan, D., Okao-Okuja, G., and Brown, J. K. (2002). A distinct bemisia tabaci (gennadius) (hemiptera: Sternorrhyncha: Aleyrodidae) genotype cluster is associated with the epidemic of severe cassava mosaic virus disease in uganda. *Molecular Ecology*, 11(7):1219–1229.

Legg, J. P., Jeremiah, S. C., Obiero, H. M., Maruthi, M. N., Ndyetabula, I., Okao-Okuja, G., Bouwmeester, H., Bigirimana, S., Tata-Hangy, W., Gashaka, G., Mkamilo, G., Alicai, T., and Lava Kumar, P. (2011). Comparing the regional epidemiology of the cassava mosaic and cassava brown streak virus pandemics in africa. *Virus Research*, 159(2):161–170.

Leite P R, E. D. and RE, S. (1994). Genetic analysis of hrp –related dna sequences of xanthomonascampestris strains causing diseases of citrus. *Applied and Environmental Microbiology*, 60:1078–1086.

Lertsuchatavanich, U., Paradornuwat, A., Chunwongse, J., Schaad, N., and Thaveechai, N. (2007). Novel pcr primers for specific detection of xanthomonascitri subsp. citri the causal agent of bacterial citrus canker. *Kasetsart Journal, Natural Sciences*, 41:262–73.

Lévesque, C. A., Harlton, C. E., and de Cock, A. W. (1998). Identification of some oomycetes by reverse dot blot hybridization. *Phytopathology*, 88(3):213–222.

Lewis, I. M., Tusiime, G., and Miller, S. (2009). A pcr assay for the detection of xanthomonascampestrispv. musacearum in bananas. *Plant Disease*, 94:109–114.

Lewis-Ivey, M., Tusiime, G., and Miller, S. (2009). A pcr assay for the detection of xanthomonas campestris pv. musacearum in bananas. *Plant Disease*, 94:109–14.

Li, W., Hilf, M. E., Webb, S. E., Baker, C. A., and Adkins, S. (2008). Presence of p1b and absence of hc-pro in squash vein yellowing virus suggests a general feature of the genus ipomovirus in the family potyviridae. *Virus Research*, 135(2):213–219.

Lievens, B., Brouwer, M., Vanachter, A. C. R. C., Lévesque, C. A., Cammue, B. P. A., and Thomma, B. P. H. J. (2005). Quantitative assessment of phytopathogenic fungi in various substrates using a dna macroarray. *Environmental Microbiology*, 7(11):1698–1710.

Lievens, B., H., I.R.M., Vanachter, A., Cammue, B., and Thomma, B. (2004). Root and foot rot on tomato caused by phytophtora infestans detected in belgium. *Plant Disease*, 88:86.

Lodhi, M. A., Ye., G.-N., Weeden, N. F., and Reisch., B. I. (1994). A simple and efficient method for dna extraction from grapevine cultivars, vitis species and ampelopsis. *Plant Molecular Biology Reporter*, 12(1):6–13.

Lukhtanov, E. A., Kutyavin, I. V., Gamper, H. B., and Meyer, Jr, R. (1995). Oligodeoxyribonucleotides with conjugated dihydropyrroloindole oligopeptides: preparation and hybridization properties. *Bioconjugate Chemistry*, 6(4):418–426.

Mahungu, N. M., Bidiaka, M., and Tata, H., L. S. N. S. (2003). Cassava brown streak disease-like symptoms in democratic republic of congo. *Roots*, 8(2):8–10.

Malorny, B., Tassios, P. T., Rådström, P., Cook, N., Wagner, M., and Hoorfar, J. (2003). Standardization of diagnostic pcr for the detection of foodborne pathogens. *International Journal of Food Microbiology*, 83(1):39–48.

Manulis, S., Valinsky, L., Lichter, A., and Gabriel, D. W. (1994). Sensitive and specific detection of xanthomonas campestris pv. pelargonii with dna primers and probes identified by random amplified polymorphic dna analysis. *Applied and Environmental Microbiology*, 60(11):4094–4099.

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y.-J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Ho, C. H., Irzyk, G. P., Jando, S. C., Alenquer, M. L. I., Jarvie, T. P., Jirage, K. B., Kim, J.-B., Knight, J. R., Lanza, J. R., Leamon, J. H., Lefkowitz, S. M., Lei, M., Li, J., Lohman, K. L., Lu, H., Makhijani, V. B., McDade, K. E., McKenna, M. P., Myers, E. W., Nickerson, E., Nobile, J. R., Plant, R., Puc, B. P., Ronan, M. T., Roth, G. T., Sarkis, G. J., Simons, J. F., Simpson, J. W., Srinivasan, M., Tartaro, K. R., Tomasz, A., Vogt, K. A., Volkmer, G. A., Wang, S. H., Wang, Y., Weiner, M. P., Yu, P., Begley, R. F., and Rothberg, J. M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 437(7057):376–380.

Martin, R. R., James, D., and Levesque, C. A. (2000). Impacts of molecular diagnostic technologies on plant disease management. *Annual Review of Phytopathology*, 38:207–239.

Maruthi, M. N., Hillocks., R. J., Mtunda., K., Raya., M. D., Muhanna., M., Kiozia., H., Rekha., A. R., Colvin, J., and Thresh., J. M. (2005). Transmission of cassava brown streak virus by bemisia tabaci (gennadius). *Journal of Phytopathology*, 153:307–312.

Mavrodieva, V., Levy, L., and Gabriel, D. W. (2004). Improved sampling methods for real-time polymerase chain reaction diagnosis of citrus canker from field samples. *Phytopathology*, 94(1):61–68.

Mbanzibwa, D. R., Tian, Y., Mukasa, S. B., and Valkonen, J. P. T. (2009a). Cassava brown streak virus (potyviridae) encodes a putative maf/ham1 pyrophosphatase implicated in reduction of mutations and a p1 proteinase that suppresses rna silencing but contains no hc-pro. *Journal of Virology*, 83(13):6934–6940.

Mbanzibwa, D. R., Tian, Y. P., Tugume, A. K., Mukasa, S. B., Tairo, F., Kyamanywa, S., Kullaya, A., and Valkonen, J. P. T. (2009b). Genetically distinct strains of cassava brown streak virus in the lake victoria basin and the indian ocean coastal area of east africa. *Archives of Virology*, 154(2):353–359.

Mbanzibwa, D. R., Tian, Y. P., Tugume, A. K., Mukasa, S. B., Tairo, F., Kyamanywa, S., Kullaya, A., and Valkonen, J. P. T. (2011). Simultaneous virus-specific detection of the two cassava brown streak-associated viruses by rt-pcr reveals wide distribution in east africa, mixed infections, and infections in manihot glaziovii. *Journal of Virological Methods*, 171(2):394–400.

McCafferty, J., Griffiths, A. D., Winter, G., and Chiswell, D. J. (1990). Phage antibodies: filamentous phage displaying antibody variable domains. *Nature*, 348(6301):552–554.

Miller, S. and Martin, R. (1988). Molecular diagnosis of plant disease. annu. rev. *Phytopathology*, 26:409–432.

Mohammed, I. U., Abarshi, M. M., Muli, B., Hillocks, R. J., and Maruthi, M. N. (2012). The symptom and genetic diversity of cassava brown streak viruses infecting cassava in east africa. *Advances in Virology*, 2012:795697.

Monger, W. A, S. S. S. C. and Foster., G. D. (2001). Identification of different isolates of cassava brown streak virus and development of a diagnostic test. *Plant Pathology*, 50:50:768–775.

Mullis, K. and Faloona, F. (1987). Specific synthesis of dna in vitro via a polymerase-catalyzed chain reaction. . *Methods in Enzymology*, 155:335–350.

Muriithi, L. and Gathama, S. (1998). Gray leaf spot of maize: a new disease on increase. *Crop Protection Newsletter*.

Mwangi, M., Mwebaze, M., Bandyopadhyay, R., Aritua, V., Eden-Green, S., Tushemereirwe, W., and Smith, J. (2007). Development of a semiselective medium for isolating xanthomonascampestrispv.musacearum from insect vectors, infected plant material and soil. *Plant Pathology*, 56:383–390.

Ndungo, V., Bakelana, K., Eden-Green, S., and Blomme, G. (2004). An outbreak of banana xanthomonas wilt (xanthomonas campestris pv. musacearum) in the democratic republic of congo. *Infomusa*, 13:43–44.

Nelson, S. and Campbell, C. (1993). Disease progress, defoliation, and spatial pattern in a multiple-pathogen disease complex on white clover. *Phytopathology*, 83:419–429.

Nicholson, P., Lees, A., Maurin, N., Parry, D., and Rezanoor, H. N. (1996). Development of a pcr assay to identify and quantify microdochium nivale var. nivale and microdochium nivalevar. majus in wheat. *Physiological and Molecular Plant Pathology*, 48:257–271.

Nicholson, P., Simpson, D. R., Weston, G., Rezanoor, H. N., Lees, A. K., Parry, D. W., and Joyce, D. (1998). Detection and quantification of fusarium culmorum and fusarium graminearum in cereals using pcr assays. *Physiological and Molecular Plant Pathology*, 53:17–37.

Nicolaisen, M., Justesen, A. F., Thrane, U., Skouboe, P., and Holmstrøm, K. (2005). An oligonucleotide microarray for the identification and differentiation of trichothecene producing and non-producing fusarium species occurring on cereal grain. *Journal of Microbiological Methods*, 62(1):57–69.

Nweke, O. C., Payne-Sturges, D., Garcia, L., Lee, C., Zenick, H., Grevatt, P., Sanders, 3rd, W. H., Case, H., and Dankwa-Mullan, I. (2011). Symposium on integrating the science of environmental justice into decision-making at the environmental protection agency: an overview. *American Journal of Public Health*, 101 Suppl 1:S19–S26.

Papp, A. C., Pinsonneault, J. K., Cooke, G., and Sadée, W. (2003). Single nucleotide polymorphism genotyping using allele-specific pcr and fluorescence melting curves. *Biotechniques*, 34(5):1068–1072.

Peruski, J., Leonard, F., Perusk., and Harwood, A. (2003). Rapid diagnostic assays in the genomic biology era: detection and identification of infectious disease and biological weapon agents. *Biotechniques*, 35(4):840–846.

Poptsova, M. S. and Gogarten, J. P. (2010). Using comparative genome analysis to identify problems in annotated microbial genomes. *Microbiology*, 156(Pt 7):1909–1917.

Reeder, R., Opolot, O., Muhinyuza, J. B., Aritua, V., and Crozier, J. (2005). Presence of banana bacterial wilt (xanthomonascampestrispv. musacearum) in rwanda. *New Disease Reports online http://www.bspp.org.uk/ndr/about.asp*.

Riley, M. (1993). Functions of the gene products of escherichia coli. *Microbiology Reviews*, 57(4):862–952.

Robène-Soustrade, I., Laurent, P., Gagnevin, L., Jouen, E., and Pruvost, O. (2006a). Specific detection of xanthomonas axonopodis pv. dieffenbachiae in anthurium (anthurium andreanum) tissues by nested pcr. *Applied and Environmental Microbiology*, 72(2):1072–1078.

Robène-Soustrade, I., Laurent, P., Gagnevin, L., Jouen, E., and Pruvost, O. (2006b). Specific detection of xanthomonas axonopodispv.dieffenbachiae in anthurium (anthuriumandreanum) tissues by nested pcr. *Applied and Environmental Microbiology*, 72:1072–1078.

Rosenzweig, C., Iglesias, A., Yang, X., Epstein, P., and Chivan, E. (2000). Climate change and us agriculture: The impacts of warming and extreme weather events on productivity. *plant diseases, and pests. Boston: Center for Health and the Global Environment, Harvard Medical School*.

Rudi, K., Rud, I., and Holck, A. (2003). A novel multiplex quantitative dna array based pcr (mqda-pcr) for quantification of transgenic maize in food and feed. *Nucleic Acids Research*, 31(11):e62.

Saghai-Maroof, M., van Scoyoc, S., Yu, Y., and Stromberg, E. (1993). Gray leaf spot disease of maize: Rating, methodology and inbred line evaluation. *Plant Disease*, 77:583–587.

Saha, M., Saha, J., and Giese, R. W. (1993). 4-(trifluoromethyl)-2,3,5,6-tetrafluorobenzyl bromide as a new electrophoric derivatizing reagent. *Journal of Chromatography*, 641(2):400–404.

Schaad, N., Jones., J., and Chun, W. (2001). *Laboratory Guide for Identification of Plant Pathogenic Bacteria, Third Edition*. Australian Plant Pathology.

Schots, A., Dewey, F. M., and Oliver, R. (1994). Modern assays for plant pathogenic fungi: Identification, detection and quantification. *Wallingford: CAB International*.

Shamloul, A. M. and Hadidi, A. (1999). Sensitive detection of potato spindle tuber and temperate fruit tree viroids by reverse transcription-polymerase chain reaction-probe capture hybridization. *Journal of Virological Methods*, 80(2):145–155.

Shim, W.-B. and Dunkle, L. D. (2005). Malazy, a degenerate, species-specific transposable element in cercospora zeae-maydis. *Mycologia*, 97(2):349–355.

Shine, J. M. and Comstock, J. C. (1993). Digital image analysis system for determining tissue-blot immunoassay results for ratoon stunting disease of sugarcane. *Plant Disease*, 77:511–513.

Skovgaard, M., Jensen, L. J., Brunak, S., Ussery, D., and Krogh, A. (2001). On the total number of genes and their length distribution in complete microbial genomes. *Trends in Genetics*, 17(8):425–428.

Spire, D. (1995). Serological methods for the diagnosis of fungi. some examples to illustrate strategies for their use. *Bulletin OEPP*, 25:19–29.

Stoodley, P., Sauer, K., Davies, D. G., and Costerton, J. W. (2002). Biofilms as complex differentiated communities. *Annual Review of Microbiology*, 56:187–209.

Sulzinski, M. (2001). Differentiation of xanthomonascampestrispvs.pelargonii and hederae by polymerase chain reaction. *Journal of Phytopathology*, 149:45–49.

Susi, P., Ziegler, A., and Torrance, L. (1998). Selection of single-chain variable fragment antibodies to black currant reversion associated virus from a synthetic phage display library. *Phytopathology*, 88(3):230–233.

Sward, R. J. and Eagling, D. R. (1995). Antibody approaches to viral diagnostics. *In New Diagnostics in Crop Sciences, Biotechnology in Agriculture*, 13:171–193.

Tehon, L. and Daniels, E. (1925). Notes on the parasitic fungi of illinois: Ii. *Mycologia*, 17:240–249.

Torrance, L. (1995). Use of monoclonal antibodies in plant pathology. *European Journal of Plant Pathology*, 101:351–363.

Torrance, L. (1998). Developments in serological methods to detect and identify plant viruses. *Plant Cell, Tissue and Organ Culture*, 52:27–32.

Toth, I., Hyman, L., Taylor, R., and Birch, R. J. (1998). Pcr based detection of xanthomonascampestrispvphaseolivarfuscans in plant material and its differentiation from x.c. pv. phaseoli. *Journal of Applied Microbiology*, 85:327–338.

Trindade, L., Eder, M., Biaggioni, L., and Álvares da Silva, M. V. F. (2007). Development of a molecular method for detection and identification of xanthomonas campestris pv. viticola. *Phytopathology*, 33:16–23.

Tripathi, L., Odipio, J., Tripathi, J. N., and Tusiime, G. (2008). A rapid technique for screening banana cultivars for resistance to xanthomonaswilt. *EuropeanJournal of Plant Pathology*, 121:9–19.

Tripathi, L., Tripathi, J. N., Tushemereirwe, W., and Bandyopadhyay, R. (2007). Development of semi selective medium for isolation of xanthomonascampestrispv-musacearum from banana plants. *European Journal of Plant Pathology*, 117:177–186.

Tushemereirwe, W., Kangire, A., Smith, J., Ssekiwoko, F., Nakyanzi, M., Kataama, D., Musiitwa, C., and Karyeija, R. (2003). An outbreak of bacterial wilt on banana in uganda. *InfoMusa*, 12:6–8.

Tushemereirwe, W., Kangire, A., Ssekiwoko, F., Offord, L., Crozier, J., Boa, E., Rutherford, M., and Smith, J. (2004). First report of xanthomonas campestris pv. musacearum on banana in uganda. *Plant Pathology*, 53:802.

Uehara, T., Kushida, A. ., and Momota, Y. (1999). Rapid and sensitive identification of pratylenchus spp. using reverse dot blot hybridization . *Nematology*, 1:549–555.

Valli, A., Dujovny, G., and García, J. A. (2008). Protease activity, self interaction, and small interfering rna binding of the silencing suppressor p1b from cucumber vein yellowing ipomovirus. *Journal of Virological Methods*, 82(2):974–986.

Valli, A., López-Moya, J. J., and García, J. A. (2007). Recombination and gene duplication in the evolutionary diversification of p1 proteins in the family potyviridae. *Journal of General Virology*, 88(Pt 3):1016–1028.

Valli, A., Martín-Hernández, A. M., López-Moya, J. J., and García, J. A. (2006). Rna silencing suppression by a second copy of the p1 serine protease of cucumber vein yellowing ipomovirus, a member of the family potyviridae that lacks the cysteine protease hcpro. *Journal of Virology*, 80(20):10055–10063.

van Beckhoven, J. R. C. M., Stead, D. E., and van der Wolf, J. M. (2002). Detection of clavibacter michiganensis subsp. sepedonicus by amplidet rna, a new technology based on real time monitoring of nasba amplicons with a molecular beacon. *Journal of Applied Microbiology*, 93(5):840–849.

Vauterin, L., Hoste, B., Kersters, K. ., and Swings, J. (1995). Reclassification of xanthomonas. *International Journal of Systematic Bacteriology*, 45:472–489.

Venkatesan, H. and M., G. M. (1996). Improved utility of photolabile solid phase synthesis supports for the synthesis of oligonucleotides containing 3'-hydroxyl termini. *Journal of Organic Chemistry*, 61(2):525–529.

Verweij, J. J., Blang, R. A., Templeton, K., Schinkel, J., Brienen, E. A. T., van Rooyen, M. A. A., van Lieshout, L., and Polderman, A. M. (2004). Simultaneous detection of entamoeba histolytica, giardia lamblia, and cryptosporidium parvum in fecal samples by using multiplex real-time pcr. *Journal of Clinical Microbiology*, 42(3):1220–1223.

Vet, J. A., Majithia, A. R., Marras, S. A., Tyagi, S., Dube, S., Poiesz, B. J., and Kramer, F. R. (1999). Multiplex detection of four pathogenic retroviruses using molecular beacons. *Proceedings of the National Academy of Sciences USA*, 96(11):6394–6399.

Voller, A. and Bidwell, D. E. (1985). Enzyme immunoassays. ed. w p collins. chichester, uk: John wiley and sons. wallingford, uk: Cab international. *Alternative Immunoassays*, pages pp. 77–86.

Wang, J., Levy, M., and Dunkle, L. D. (1998). Sibling species of cercospora associated with gray leaf spot of maize. *Phytopathology*, 88(12):1269–1275.

Wang, Z., Comstock, J., Hatziloukas, E., and Schaad, N. (1999). Comparison of pcr, bio-pcr, dia, elisa and isolation on semiselective medium for detection of xanthomonasalbilineans, the causal agent of leaf scald of sugarcane. *Plant Pathology*, 48:245–52.

Ward, E. (1994). Use of the polymerase chain reaction for identifying plant pathogens. *In Ecology of Plant Pathogens*, pages 143–160.

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

Ward, J., Stromberg, E., Nowell, D., and Nutter, F. (1999). Gray leaf spot: a disease of global importance in maize production. *Plant Disease*, 83:884–895.

Waterhouse, P. M. and Chu, P. W. G. (1995). Nucleic acid based approaches to plant virus and viroid diagnostics. in new diagnostics in crop sciences, biotechnology in agriculture. *In New Diagnostics in Crop Sciences, Biotechnology in Agriculture*.

Werres, S. and Steffens, C. (1994). Immunological techniques usedwith fungal plant pathogens - aspects of antigens, antibodiesand assays for diagnosis. *Annals of Applied Biology*, 125:615–643.

West, C., Pepping, F., and Temalilwa, C. (1988). *The composition of foods commonly eaten in East Africa*. The Netherlands: Wageningen Agricultural University.

Wiley, G., Macmil, S., Qu, C., Wang, P., Xing, Y., White, D., Li, J., White, J. D., Domingo, A., and Roe, B. A. (2009). Methods for generating shotgun and mixed shotgun/paired-end libraries for the 454 dna sequencer. *Current Protocols in Human Genetics*, Chapter 18:Unit18.1.

Williams, R. H., Ward, E., and McCartney, H. A. (2001). Methods for integrated air sampling and dna analysis for detection of airborne fungal spores. *Applied and Environmental Microbiology*, 67(6):2453–2459.

Winter, S., Koerbler, M., Stein, B., Pietruszka, A., Paape, M., and Butgereitt, A. (2010). Analysis of cassava brown streak viruses reveals the presence of distinct virus species causing cassava brown streak disease in east africa. *Journal of General Virology*, 91(Pt 5):1365–1372.

Wittwer, C. T., Herrmann, M. G., Moss, A. A., and Rasmussen, R. P. (1997). Continuous fluorescence monitoring of rapid cycle dna amplification. *Biotechniques*, 22(1):130–1, 134–8.

Wu, L., Thompson, D. K., Li, G., Hurt, R. A., Tiedje, J. M., and Zhou, J. (2001). Development and evaluation of functional gene arrays for detection of selected genes in the environment. *Applied and Environmental Microbiology*, 67(12):5780–5790.

Yang, S. and Rothman, R. E. (2004). Pcr-based diagnostics for infectious diseases: uses, limitations, and future applications in acute-care settings. *Lancet Infectious Diseases*, 4(6):337–348.

Zaccardelli, M., Campanile, F., Spasiano, A., and Merighi, M. (2007). Detection and identification of the crucifer pathogen, xanthomonascampestrispv.campestris, by pcr amplification of the conserved hrp/type iii secretion system gene hrcc. *European Journal of Plant Pathology*, 118:299–306.

Zhang, X., Du, P., Lu, L., Xiao, Q., Wang, W., Cao, X., Ren, B., Wei, C., and Li, Y. (2008). Contrasting effects of hc-pro and 2b viral suppressors from sugarcane mosaic virus and tomato aspermy cucumovirus on the accumulation of sirnas. *Virology*, 374(2):351–360.

Zhou, J. and Thompson, D. K. (2002). Challenges in applying microarrays to environmental studies. *Current Opinion in Biotechnology*, 13(3):204–207.

Zimmermann, C., Hinrichs-Berger, J., Moltmann, E., and Buchenauer, H. (2004). Nested pcr (polymerase chain reaction) for the detection of xanthomonas fragariae in symptomless strawberry plants. *Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz*, 111:39–51.