Title: Population Analysis of the Finger Millet Blast Pathogen Magnaporthe oryzae in Eastern Africa

Name: Taiwo Adewale Shittu

Population Analysis of the Finger Millet Blast Pathogen *Magnaporthe oryzae* in Eastern Africa

Taiwo A. Shittu

Ph.D.

2018

UNIVERSITY OF BEDFORDSHIRE

Population Analysis of the Finger Millet Blast Pathogen *Magnaporthe oryzae* in Eastern Africa

Taiwo A. Shittu

Ph.D.

A thesis submitted to the University of Bedfordshire

In partial fulfilment of the requirement for the degree of Doctor of Philosophy

November 2018

Declaration

I, Taiwo Adewale Shittu declare that the material presented in this thesis is my own work and has not been presented for another degree. Materials provided through collaborations or the methodologies adopted from previous studies are fully acknowledged and referenced appropriately.

Signed: Date: 26/11/2018

Abstract

The main aims of the investigation were to develop an in-depth understanding of the genetic diversity, population structure and evolutionary relationships as well as to assess the sexual reproductive capability of the finger millet blast (FMB) pathogen *Magnaporthe oryzae* in Eastern Africa. A set of 300 *M. oryzae* isolates collected during 2000 – 2017 from key finger millet growing districts in Kenya, Uganda, Tanzania and Ethiopia were utilised in this study.

Two novel molecular markers designated HyP1 and HyP2 were developed in this study and two known phylogenetic markers ITS (internal transcribed spacer) and HIS4 (histone 4 gene) were identified by bioinformatic analysis. Single- and multi-locus analysis provided a clear assessment of the FMB pathogen genotype diversity and distribution pattern. At the regional level in Eastern Africa, ITS and HIS4 revealed 7 - 9 genotypes, whereas HyP1 and HyP2 identified 80 - 85 genotypes reflecting their high resolution. Multi-locus sequence (MLS) analysis revealed 207 genotypes displaying a continuous genetic variation pattern of the FMB pathogen populations in Eastern Africa. Bayesian and reticulate network analyses distinguished the vast majority of genotypes into two sub-populations (designated as Group A and B), which were geographically clustered. Diagnostic PCR revealed the presence of a high proportion of *M. oryzae* isolates containing the Grasshopper (*grh*) repeat element in Ethiopia and Tanzania (e.g. 85 %).

Reference genome assemblies have been established for two *M. oryzae* isolates representing the sub-populations identified. Genome resequence data has been developed for sixteen isolates representing the genotype diversity. Comparative analysis provided novel insights into the genomic architecture and evolutionary relationships in the FMB pathogen. Genomic regions and/or genes, putatively isolate specific have been identified. Phylogenomic analysis revealed monophyletic nature of the FMB pathogen in Eastern Africa and Asia suggesting a common origin. Genome-wide single nucleotide polymorphism (SNPs) ranges broadly corresponded to the sub-populations identified. Complete *grh* sequence has been defined and the presence of at least two versions of the element in the FMB pathogen in Eastern Africa has been shown.

Mating type specific PCR assay revealed high proportions of the two mating types MAT 1-1 (56 %) and MAT 1-2 (44 %) in the contemporary population of the FMB pathogen in Eastern Africa and also in the four countries surveyed, albeit at variable levels. Mating culture assays established a high proportion of fertile isolates (60 %) and revealed the dominance of male sexual behaviour followed by hermaphrodite and female isolates. The emerging pattern is indicative of a decrease in the fertility status as well as the level of hermaphrodites and females. Integrated assessment of the mating type and fertility data along with the high genotype diversity and their continuous variation pattern observed is strongly suggestive of a mixed reproductive behaviour including episodic sexual reproduction.

The new knowledge and resources generated contribute to the advancement of current understanding of the finger millet blast pathogen biology providing a framework for the effective utilization of host resistance in Eastern Africa as well as a strong platform for further research advances in the field.

Acknowledgments

I sincerely appreciate my director of study, Professor S. Sreenivasaprasad for his inestimable guidance, scholarly inputs and consistent encouragement throughout this research work. Also, my appreciation goes to other members of my supervisory team, Dr. Thamarai Janganan, and Dr. Nick Worfolds for their academic support when required. I want to say a Big 'Thank you' to Dr. S. Muthumeenakshi, for her valuable contributions which enhanced the quality of this work. My gratitude goes to the Federal Government of Nigeria, BMGF, and Department of Life Sciences/iBEST of the University of Bedfordshire for their financial support to execute this research successfully. I would also like to appreciate all the members of the PEARL II project team led by Dr. S.M. de Villiers for the collection and isolation of samples used for this research work. I would sincerely like appreciate Dr. R. Baroncelli, Dr. A. Armitage, Dr. L. Baxter, Dr. Steven Dodsworth, and Dr. A. Standing for various bioinformatic training provided. My special thanks are also extended to all the past and present academic and technical staffs of the School of Life Sciences for creating a friendly working environment. Lovely thanks to Nathaniel, Dr. M. Onigbaro, Ogechukwu (Dearest), Bandar, Haroon, Ingrid, Guri, Daniel, Emma, Ben, Lukasz, Elais, Andrew stallion and Denise for their support through the Ph.D. journey. My thumb is up to all my wonderful friends in Luton (Too numerous to mention). My deepest appreciation goes to my biological parent, Chief and Mrs. M. A Shittu and siblings - Kehinde, Olanrewaju and Abimbola for their undying love and unwavening support. Finally, I will like to uniquely appreciate -Wemimo for her unflinching and understanding over the years. Above all, my endless appreciation goes to the Bishop of my soul, God Almighty for his love, guidance, provision and good health through this journey. To God be the glory, Phil. 1:6.

Publications/ Presentations

Almiman, B. F., **Shittu, T. A.**, Muthumeenakshi, S., Baroncelli, R. and Sreenivasaprasad, S. 2018. Genome sequence of the mycotoxigenic crop pathogen *Fusarium proliferatum* strain ITEM 2341 from date palm. Microbiology Resource Announcements 7: e00964-18.

Shittu, T. A. 2018. Finger millet blast disease management: development of key resources to contribute to food security in East Africa. Winter conference 2018: Research with Impact. University of Bedfordshire, UK.

Shittu, T.A., Baroncelli, R., Armitage, A., Kaleta, L., Takan, J.P., de Villiers, S., Worfold, N., Muthumeenakshi, S. and Sreenivasaprasad, S. 2017. Deciphering the adaptive evolutionary patterns associated with biogeographic diversity in the finger millet blast pathogen Magnaporthe oryzae: Development of a reference genome sequence of a fungal isolate from East Africa. NCBI Deposition (Accession Number: PHFK00000000).

Shittu, T. A., Almiman, A., Muthumeenakshi, S., Worsfold, N. and Sreenivasaprasad, S. 2017. Population analysis of the finger millet blast pathogen *Magnaporthe oryzae* in East Africa using molecular markers. Molecular biology of plant pathogens conference, BSPP conference, 29th – 30th March 2017, Durham University, UK, pp. 82.

Shittu, T. A., Almiman, A., Muthumeenakshi, S., Takan, J. P., Worsfold, N. and Sreenivasaprasad, S. 2017. Understanding the pathogen distribution and evolution to develop resources and strategies for effective management of the finger millet blast disease in East Africa. Food security biosecurity & trade the role of plant health, BSPP Conference, 12 -13 September 2016, University of Oxford, UK. pp 62.

Shittu, T. A. 2015. Population dynamics of the finger millet blast pathogen *Magnaporthe oryzae* in East Africa. Early career biologist network event. Society of biology conference. GSK. 29th June 2015.

LIST OF ABBREVATIONS

AAs	Amino acids
ACT	Actin
AFLP	Amplified Fragment Length Polymorphism
AVR	Avirulence
ASTRAL	Accurate Species Tree Algorithm
BAM	Binary alignment map
Bb	Billion bases
bp	Base pair
BI	Broad Institute
BLAST	Basic Local Alignment Search Tool
BMGF	Bill & Melinda Gates Foundation
Bp	Base pair
BCF	Binary counterpart format
BSV	Bootstrap support values
BUSCO	Benchmarking Universal Single-Copy Orthologs
BWA	Burrows- Wheeler Aligner
CAZymes	Carbohydrate active enzymes
CDS	Coding sequence
cDNA	Complementary DNA
DDH	DNA-DNA hybridisation
DFID	Department for International Development
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
ETS	External transcribed spacer

ETS	Effector-triggered susceptibility
ETI	Effector-triggered immunity
EU	European Union
FAO	Food and Agriculture Organisation of the United Nations
FMB	Finger millet Blast
FUNYBASE	Fungal Phylogenomic Database
gDNA	Genomic DNA
grh	*Grasshopper*
HIS4	Histone 4
HR	Hypersensitive response
ISSR	Inter Simple Sequence Repeat
IGS	Intergenic spacer
ICRISAT	International Crops Research Institute for the Semi-Arid Tropics
InDel	Insertion or deletion
ITS	Internal Transcribed Spacer
JGI	Joint Genome Institute
Kb	Kilobase
Kg	Kilogram
L	Litre
LINE	Long Interspersed Nuclear Element
LTR	Long Terminal Repeat
LSU	Large subunit
LRR	Leucine rich repeat
LTRs	Long terminal repeats
M	Million

MAMPs	Microbial associated molecular patterns
Mb	Million bases
mg	Milligram
M. oryzae	*Magnaporthe oryzae*
MLS	Multi-locus sequence
MGR	Magnaporthe Grisea Repeat
MAT	Mating type gene
MCMC	Markov Chain Monte Carlo
MEGA	Molecular Evolutionary Genetics Analysis
MP	Mate Pair
NARES	National Agricultural Research and Extension Services
ng	Nanogram
NCBI	National Center for Biotechnology Information
NGS	Next Generation Sequencing
NB	Nucleotide binding
NRPS	Non-ribosomal peptide synthetase
ORF	Open reading frame
PAMPs	Pathogen associated molecular patterns
PCR	Polymerase Chain Reaction
PDB	Potato dextrose broth
PDA	Potato dextrose agar
PE	Paired End
1% PDB	PDB made with 1% the recommended weight of PDB powder
PPV	Posterior Probability Value
PKS	Polyketide synthase

PRRs	Pattern recognition receptors
psi	Pounds per square inch
PTI	PAMP-triggered immunity
QTL	Quantitative traits loci
R	Resistance
RFLPs	Restriction Fragment Length Polymorphisms
RAPD	Random Amplified Polymorphism DNA
rep-PCR	Repetitive-based Polymerase Chain Reaction
Rpm	Rotation per minute
REMAP	Retrotransposon Microsatellite Polymorphism
RAxML	Randomised axelerated maximum likelihood
RNA	Ribonucleic acid
rRNA	ribosomal RNA
SAM	Sequence alignment/map format
SSR	Simple Sequence Repeat
SSU	Small subunit
SCAR	Sequence Characterised Amplified Region
SBS	Sequencing by Synthesis
SSCP	Single Strand Conformation Polymorphism
SINE	Short Interspersed Nuclear Elements
SNPs	Single nucleotide polymorphisms
SMRT	Single molecule real time
spp	Species
TEs	Transposable Elements
TEF	Translation elongation factor

UK	United Kingdom
USA	United States of America
VCF	Variant call format
X	Coverage
N50	Point at which 50 % of contigs are present in contigs of size N or greater
K2+G	Kimura 2-parameter (K2) model with a discrete gramma distribution (G)
K2+G+I	kimura 2-parameter model using a discrete gramma distribution (G) with 5 rate categories that contain certain fraction of sites and evolutionary invariable (I)
HKY+G	Hasegawa-Kishino-Yano (HKY) model with a discrete gramma distribution (G)
T92	Tamura 3-parameter (T92) model with a discrete gramma distribution (G)

Symbol

^{O}C	Degree Centigrade
%	Percentage
µg	Microgram
µl	Microliter

Dedication

This work is dedicated to my loving mother, Mrs. Sarah Shittu.

Table of Contents

List of Figures

List of Tables

Chapter 1

1.0 Introduction

1.1 Research context

There is a growing concern globally about the impact of plant diseases on food and nutrition security particularly in the context of human and animal health, economic growth and environmental sustainability (Ronald, 2011; Fisher *et al*., 2012; Fisher *et al*., 2016). At present there are over 815 million people living with hunger and by 2050 the world population is expected to be around 9 billion. Increasing demand for more food production combined with the threat of climate change pose a growing stress on the agro-ecosystems.

Crop diseases contribute to losses in yields enough to feed an additional 1 billion people annually (Strange and Scott, 2005; Chakraborty and Newton, 2011; Savary *et al*., 2012), and the use of chemical control is being discouraged by various governments (e.g. European Union) due to its adverse effects on people and the environment including beneficial species. Pathologists, breeders and other stakeholders concerned with food security often recommend the use of designated resistant varieties as a durable, cost-effective and environmentally friendly control measure (Anderson *et al*., 2004; Godfray *et al*., 2010; Ronald, 2011; Wakelin *et al*., 2018). However, many pathogens possess the adaptive capacity to survive, reproduce and overcome control measures including resistant crop varieties. In-depth understanding of the pathogen population diversity and evolution in different production systems will help to evaluate the risk and develop effective control

measures including development and deployment of suitable resistant varieties (Croll and Laine, 2016; Depotter *et al*., 2017). This is critically important to achieve food and nutritional security whilst keeping our environment safe (Beddington, 2010).

1.2 Research problem and outline of the thesis

Finger millet is a major staple for millions of people in Eastern, Central and Southern Africa. In Asia, the crop is also widely grown in parts of India. In a global context, it is highly nutritious and gluten-free compared to other more widely consumed cereals such as rice and wheat (Shobana *et al*., 2013). However, blast disease caused by the fungus *Magnaporthe oryzae* (*M. oryzae*) is a major constraint affecting the crop at all stages leading to serious grain losses, which if saved are adequate to feed millions of people annually (Zhang *et al*., 2016). Genetic and pathogenic diversity of *M. oryzae* has been well documented from various parts of the world where the disease is highly prevalent (e.g. in rice crops). In a widely cited review, blast disease especially on rice has been ranked as the number one disease caused by fungal plant pathogens in the world impeding global food security (Dean *et al*., 2012). Previous work by Sreenivasaprasad and co-workers (UK-DFID funded projects) developed the initial knowledge of the blast pathogen populations on finger millet utilising samples collected between 2000-2004 in Uganda and Kenya. Blast disease control has been identified as a key entry point for sustainable production and utilisation of finger millet in Eastern Africa (e.g. Takan *et al*., 2004; Lenne *et al*. 2007; Takan *et al*., 2012).

The present study focuses on the i) identification and/or development of sequence-based molecular markers to investigate the blast pathogen genetic diversity to identify the genotypes and assess their phylogenetic relationships, ii) development of genome sequence resources from representative isolates and comparative analysis to decipher genome-level differences reflective of pathogen evolution, and iii) evaluation of the sexual reproductive capability of pathogen populations prevalent in key finger millet production locations to enable an assessment of the potential of sexual reproduction and recombination to contribute to pathogen evolution. In this investigation, a total of 300 *M. oryzae* isolates from finger millet production systems in Eastern Africa were characterised including 76 isolates representing a collection from 2000-2004 (historic isolates/population), and 224 isolates from 2015- 2017 (contemporary isolates/ population).

Rationale for the research, background and the literature pertinent to key areas, and the aims and objectives of the study are presented as part of the Introduction in Chapter 1, and the Materials and Methods used in the research are presented in Chapter 2. The results chapters have been organized as follows: Population analysis and phylogenetic relationships based on single- and multi-locus sequence data including the *Grasshopper* element screening in Chapter 3, Comparative genomic analysis of representative genotypes in Chapter 4, and the sexual reproductive capability of contemporary populations based on molecular and mating assays in Chapter 5. The final Chapter 6 provides an overarching discussion of the key findings of the research followed by a Conclusion and Future Perspectives section.

1.3 Finger millet: background, production, economic importance and cultivation constraints

1.3.1 Historical background and distribution of finger millet

Finger millet, *Eleusine coracana* L. Gaertn (Figure 1.1) is a seeded annual cereal crop that belongs to the grass family, *Poaceae* and sub-family *Chloridoideae* (Soreng *et al*., 2015). Finger millet originated from the highlands of Ethiopia to Uganda where it was first domesticated and has been cultivated in the region for more than five thousand years and thereafter introduced to other parts of the world notably to the Asian continent specifically India is well documented (Hilu *et al*., 1979; de Wet *et al*., 1984).

Historically, two distinct races of finger millet are recognised, the African highlands race and the Afro-Asiatic lowland race (Mathur *et al*., 2012). Finger millet is one of the most important staple food crops in many developing countries but often neglected due to its labour-intensive cultivation from planting to production. In Africa, it is mostly cultivated in countries such as Uganda, Ethiopia, Tanzania, Kenya, Zaire, Sudan, Eritrea, Malawi, Zambia, Madagascar, Zimbabwe, Nigeria, Rwanda and Burundi and it is widely grown in parts of India and other parts of the Asian continent including China, Japan, Sri Lanka and Malaysia as shown in Figure 1.2 (Mathur *et al*., 2012). Finger millet is also known as African millet whereas in some countries, it is known by local names. For example, in India, it is called Mandua or Ragi, Tellebun in Sudan and Bulo in Uganda (National Research Council-USA, 1996). Finger millet is a tetraploid with a total genome composition of AABB. It has a basic chromosome number of 9 (2n = 4x = 36) and a total genome size of 1, 196 Mb (Hittalmani *et al*., 2017; Hatakeyama *et al*., 2017).

Figure 1.1. Mature healthy finger millet crop in Ethiopia with panicles (seed head)

Image provided by Dr Kassahum Tesfaye at the Institute of Biotechnology, University of Ethiopia, Addis Ababa as part of ongoing research collaboration.

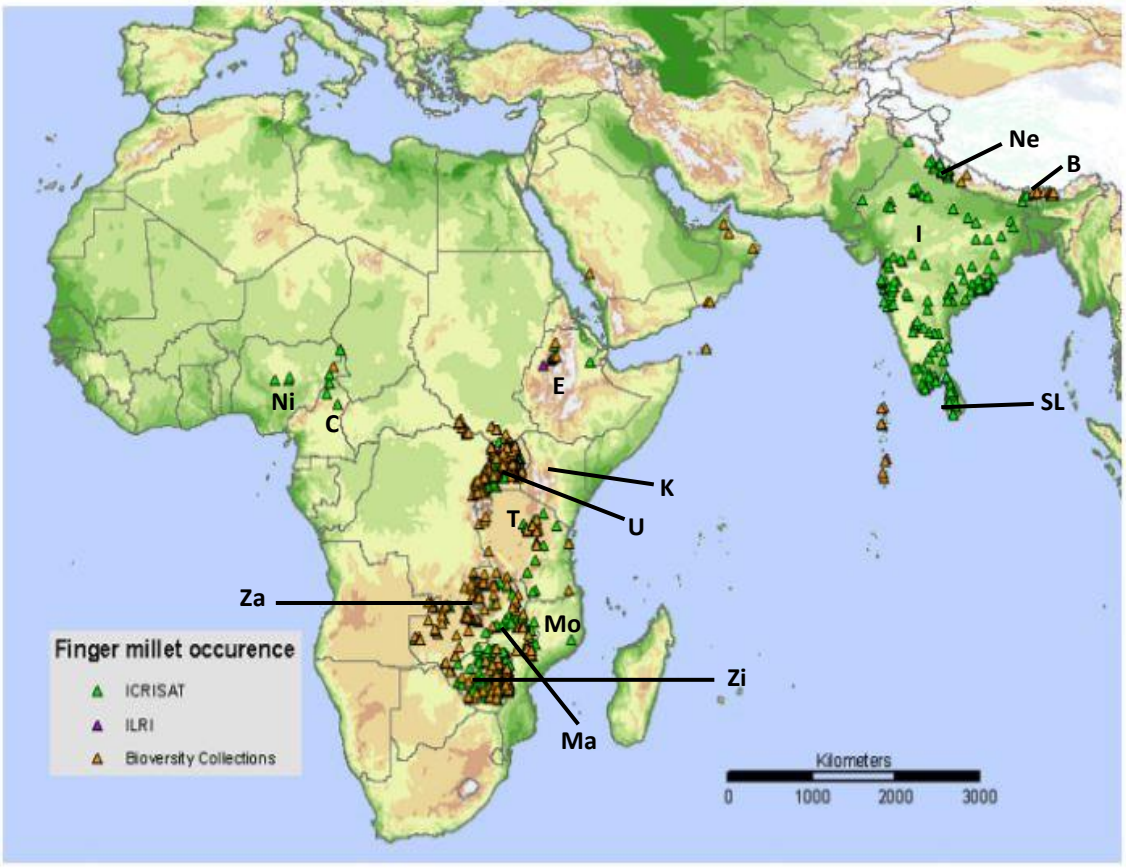

Figure 1.2. Distribution of finger millet based on the information in a global database

The map shows important finger millet growing countries in Africa and some parts of South Asia. The countries include B-Bhutan, C -Cameroon, E-Ethiopia, I-India, Ma-Malawi, Mo-Mozambique, Ne-Nepal, Ni-Nigeria, S-Sri Lanka, T-Tanzania, Za-Zambia, and Zi-Zimbabwe. ICRISAT- International crops research institute for semi-arid tropics; ILRI-International livestock research institute. The map only includes locations based on the sampling by the three organisations. However, it is also widely grown in some other countries such as Kenya (K). The map was adopted from Mathur *et al*., 2012.

1.3.2 Production, processing and utilisation of finger millet

Among the widely produced cereals in the semi-arid regions of the world, finger millet ranks third after sorghum (*Sorghum bicolor*), and pearl millet (*Pennisetum glaucum*) as reported by various authors (Barbeau and Hilu, 1993; Upadhyaya *et al*., 2007; De Villiers *et al*., 2015). Crucially, the crop ranks as the second most important cereal crop after maize (*Zea mays*) in Uganda and Tanzania (Oryokot, 2001; Wanyera, 2005a; Kisandu *et al*., 2005), third in Ethiopia after maize and wheat (Mulualem and Melak, 2013; Cochrane, 2014; Kinfe *et al*., 2017) and fourth in Kenya after maize (*Zea mays*), wheat (*Triticum aestivum*) and sorghum (*Sorghum bicolor*) as has previously been reported (Oduori, 1993; Oduori and Kanyenji, 2005). This crop is majorly grown by local farmers in the arid and semi-arid regions of Africa and Asia often as a mono crop, or intercropped with other cereals, legumes and vegetables (Purseglove, 1972; National Research Council-USA, 1996). It is a fast-growing cereal crop that reaches maturity within three to six months and in some cases as short as around 45 days (Dida and Devos, 2006). Grains from finger millet are processed by milling and sieving, soaking and cooking, fermentation, popping, malting, puffing, flaking, and decortication (Shobana *et al*., 2013; Saleh *et al*., 2013; Wafula *et al*., 2018; Ramashia *et al*., 2018). Like some other cereals, finger millet grains are crushed in a roller mill into crude flour, which is then utilised as porridge and for other food products. In Africa, it is prepared as food in the form of bread and porridge, and is an important diet for pregnant women, nursing mothers, infants and children (Oryokot, 2001). In India, it is used to prepare a wide range of food products including vermicelli noodles, sweet mixes, soups and papads (Shobana *et al*., 2013).

1.3.3 Nutritional and Health values of finger millet

Finger millet is rich in protein and carbohydrates in the form of non-starchy polysaccharides as well as minerals such as iron and calcium (Swami *et al*., 2013; Shobana *et al*., 2013; De Villiers *et al*., 2015). Of all cereals, finger millet has the highest amount of calcium and potassium (Shobana *et al*., 2013). It has high dietary fibre content, minerals, sulphur and a well-balanced amino acid profile compared to white rice. In addition, it is a good source of methionine, cysteine, lysine and tryptophan (Oduori and Kanyenji, 2005). Evidence from experimental trials using animals as a model revealed that foods from finger millet have a low glycaemic index, which is good for diabetic patients as it lowers the blood glucose and cholesterol levels (Rajasekaran *et al*., 2004, Shobana *et al*., 2013).

As a result of its nutritional values, consumption of finger millet products is encouraged and promoted by the Government of Uganda (AIDS projects) to support people suffering from Human Immunodeficiency Virus/Acquired Immunodeficiency Syndrome (HIV/AID) and the nation's healthier food school feeding programmes (Wanyera, 2005b). Its products have been recognised and recommended to the public as a dietary supplement for HIV positive patients in some parts of East Africa (Sreenivasaprasad *et al*., 2005). It has also been highlighted as a key crop to manage micronutrient deficiency, a major challenge in infant and young children and pregnant women in Sub-Saharan African countries such as Kenya and Uganda (Lenne *et al*., 2007).

Additionally, in the East African region, apart from being consumed as food, it is also a source of income generation for many households, most importantly empowering women finger millet cultivation (Oduori and Kanyenji, 2005). It can be sold directly as grain or brewed to produce local beer, generating significant income. Its straws serve as good fodder

for cattle and the fields are used for post-harvest grazing. Finger millet can be stored for a long period under good conditions compared to other agricultural crops, for example maize (Oduori and Kenyenji, 2005). These attributes along with high nutritional values and excellent storage qualities make it a potential food security crop (Oduori and Kanyenji, 2005, Wanyera, 2005b, Kisandu *et al*., 2005; Shobana *et al*., 2013).

1.3.4 Constraints to finger millet production

Finger millet cultivation is faced with some constraints that continue to lower its production. One of the major constraints is the overall cultivation process involved, which is highly labour intensive from land clearing, planting, weeding, harvesting through to processing (National Research Council-USA, 1996; Owere *et al*., 2014). Although there is no important insect pest that has so far been reported on finger millet, it is affected by some economically important diseases that drastically reduce its production. Fungal pathogens that cause diseases on finger millet include blast (*Magnaporthe oryzae*), rust (*Puccinia substriatia*), downy mildew (*Sclerospora graminicola),* seedling and leaf blight (*Helminthosporium nodulosum*), Cercospora leaf spot (*Cercospora pennisetic*) Cylindrosporium leaf spot (*Cylindrosporium species*) and tar spot (*Phyllachora eleusines*). Finger millet is also susceptible to bacterial blight (*Xanthomonas campestris* pv *eleusines*). Of all these biotic constraints, blast is considered the most damaging and economically significant disease of finger millet (Panwar *et al*., 2011; Owere *et al*., 2014).

Blast disease is highly destructive and causes more than 50 % reduction in yield when the panicle is infected. In 2005, losses of 10 - 90 % were recorded in a field in Uganda. In a similar report, 64 % loss was recorded in Kenya and nearly 100 % losses were reported in

parts of India (Kihoro *et al*., 2013, Saleh *et al*., 2014). In Kenya, a survey targeting farmers revealed that 72 % of the farmers did not engage themselves in any other socioeconomic work apart from farming and 24 % had opted out of finger millet production due to losses from blast disease (Kihoro *et al*., 2013). Blast disease has been identified as a key factor affecting further expansion of finger millet production in the East African region. An effective management of this disease will provide an entry point for fighting malnutrition and poverty in the region (Sreenivasaprasad *et al*., 2005; Owere *et al*., 2014). In addition, blast disease affects other major staple food crops such as rice and wheat that are paramount to food security and has been considered the most economically important pathogenic fungal disease in the world (Dean *et al*., 2012).

1.4 Blast disease

Blast disease is caused by the fungus *Magnaporthe oryzae* (*M. oryzae*) B. Couch (anamorph: *Pyricularia oryzae* Cavara); synonym *Magnaporthe grisea* (*M. grisea*) (Hebert Barr) (Ou, 1985; Couch and Kohn, 2002; Zhang *et al*., 2016). The fungal species has been classified into two distinct groups (Figure 1.3), *M. oryzae* and *M. grisea* based on molecular analysis involving some ortholog genes that were able to distinguish the evolutionary relationship of pathogens affecting major cereal crops including some grasses and other hosts typified by *Digitaria* species, respectively (Couch and Kohn, 2002; Klaubauf *et al*., 2014; Luo *et al*., 2015). These two species are indistinguishable in conidium, perithecium and ascospore morphology (Klaaubauf *et al*., 2014). Strains affecting cereal crops are now classified as *M. oryzae*, while strains affecting some wild hosts and certain grasses are classified as *M. grisea*. Consequently, the name *M. oryzae* is used throughout this thesis for the blast

pathogen isolates associated with finger millet and other cereal crops. *M. oryzae* belongs to the genus Magnaporthe, family Magnaporthaceae, order Magnaporthales, class Sordariomycetes within the phylum Ascomycota (Luo *et al*., 2015; Zhang *et al*., 2016). Other species in this genus include *M. poae*, *M. salvinii* and *M. rhizophilia* (Klaubauf *et al*., 2014; Luo *et al*., 2015). *M. oryzae* is a hemibiotrophic filamentous ascomycete fungus (Koeck *et al*., 2011). Hemibiotrophic fungal plant pathogens require living tissue for a phase during the infection process to survive and complete their life cycle. Hemibiotrophic fungi include major pathogens that cause huge crop losses, threatening global economy and food security (Park *et al*., 2009; Koeck *et al*., 2011).

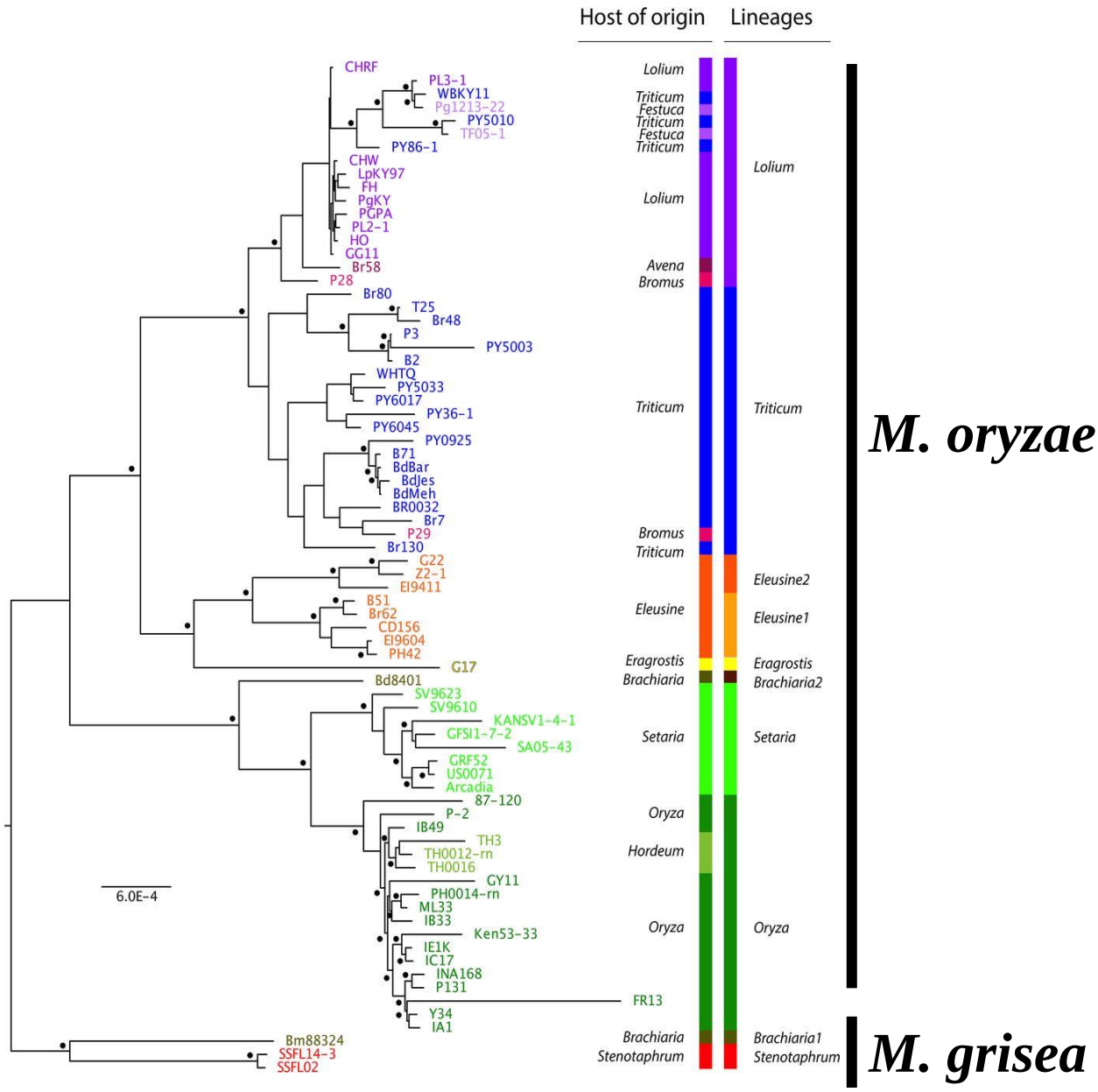

Figure 1.3. Phylogenetic tree illustrating the distinction of *Magnaporthe oryzae* from *M. grisea*

Phylogenetic tree based on the maximum likelihood analysis of the concatenation of 2682 orthologous coding sequences extracted from the genome of 76 isolates from diverse hosts. Nodes with bootstrap support of >90% are indicated by dots (100 bootstrap replicates). This tree was adopted from Gladieux *et al*., 2018.

1.4.1 Host range and specificity of the blast pathogen *Magnaporthe oryzae*

Blast disease occurs wherever the host plants are grown in the world under favourable conditions (Talbot, 2003). It has been reported in more than 85 countries of the world particularly on rice (Figure 1.4), including developed and developing nations (Roumen *et al*., 1997; Zhang *et al*., 2016). *M. oryzae* can infect and cause disease on a wide range of host plants of the Poaceae family (Talbot, 1995 and 2003) including finger millet and rice (*Oryza sativa*), which standout as the most important hosts (Ou, 1985; Talbot, 2003; Onaga *et al*., 2015), wheat (*Triticum aestivum*) (Kohli *et al*., 2011; Maciel *et al*., 2014), oat (*Avena strigosa*) (Urashima and Silva, 2011; Marangoni *et al*., 2013), maize (*Zea mays*), barley (*Hordeum vulgare*) (Hyon *et al*., 2012), foxtail millet (*Setaria italica*), wild millet (*Eleusine indica*) (Takan *et al*., 2012), pearl millet (*Pennisetum glaucum*), and ginger (Devulapalle and Suryanarayanan, 1995). The fungus also infects many weed grasses that can be used in the production of herbal formulations as well as animal feed. Some of the grasses include *Secale cerereale* (rye grass), *Lolium perenne* (perennial rye-grass), *Cynodon dactylon*, *Dactylotenium aegyptium, Digitaria horizontalis*, *Pennisetum purpureum*, *Echinochola colonum*, *Leersia hexandra*, and *Brachiaria indica* (Disthaporn, 1994; Skamnioti and Gurr, 2009; Kohli *et al*., 2011; Takan *et al*., 2012).

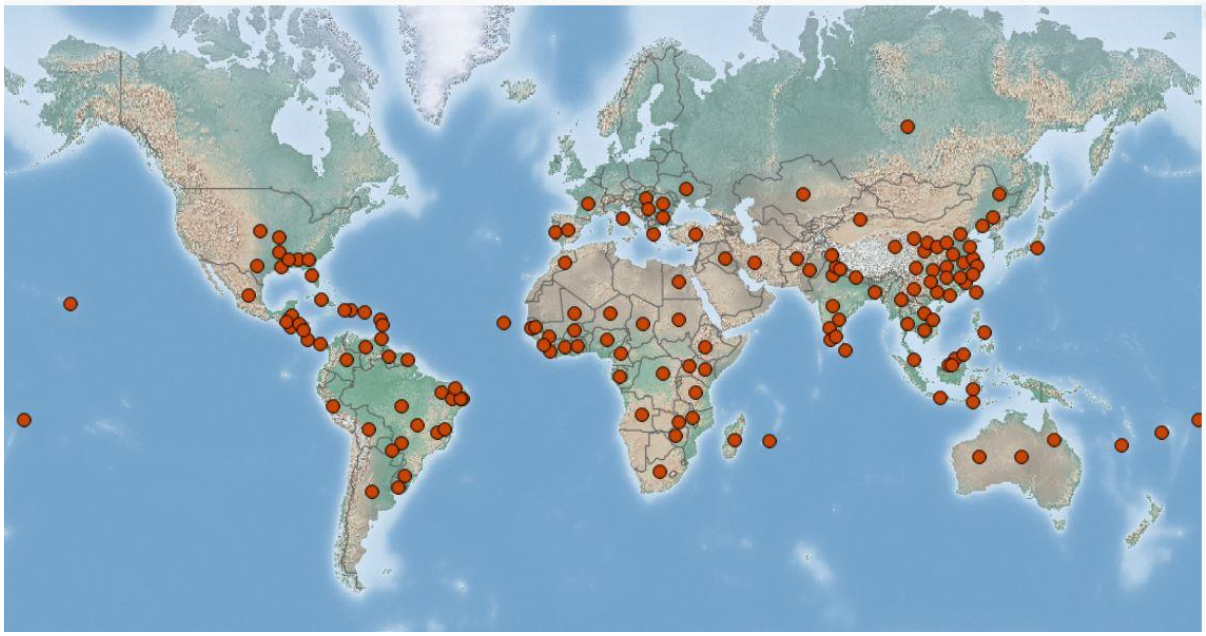

Figure 1.4. World distribution of blast disease

The red dots show the countries and regions where blast diseases have been reported. The map was adopted from the CABI website on the *M. oryzae* datasheet.
https://www.cabi.org/isc/datasheet/46103

1.4.2 Blast disease symptoms

M. oryzae attacks finger millet and other crop hosts at different stages of development from seedling stage to the panicle formation by forming lesions on the infected plant parts. These include leaves, leaf collars, neck, panicles (Figure 1.5), seeds, pedicels and even the roots of the susceptible hosts (Ghazanfar *et al*., 2009; Takan *et al*., 2012; Babu *et al*., 2013). The most obvious symptoms of this disease appear on the leaves and the neck (De Datta, 1981) as seen in Figure 1.5. The centre of the lesion has dull grey green or pale green with a dark brown outer rim that appears soaked and the centre gradually becomes grey or almost straw colour. The lesions on the neck or on the nodes of the panicles near the base of the panicle are the most striking symptoms and the destructive form of the disease leading to yield loss (De Datta, 1981). The appearance of blast lesions on the leaves of the susceptible

host plants depends largely on the age of the host plant, the environmental condition and the level at which the host varieties can resist the disease.

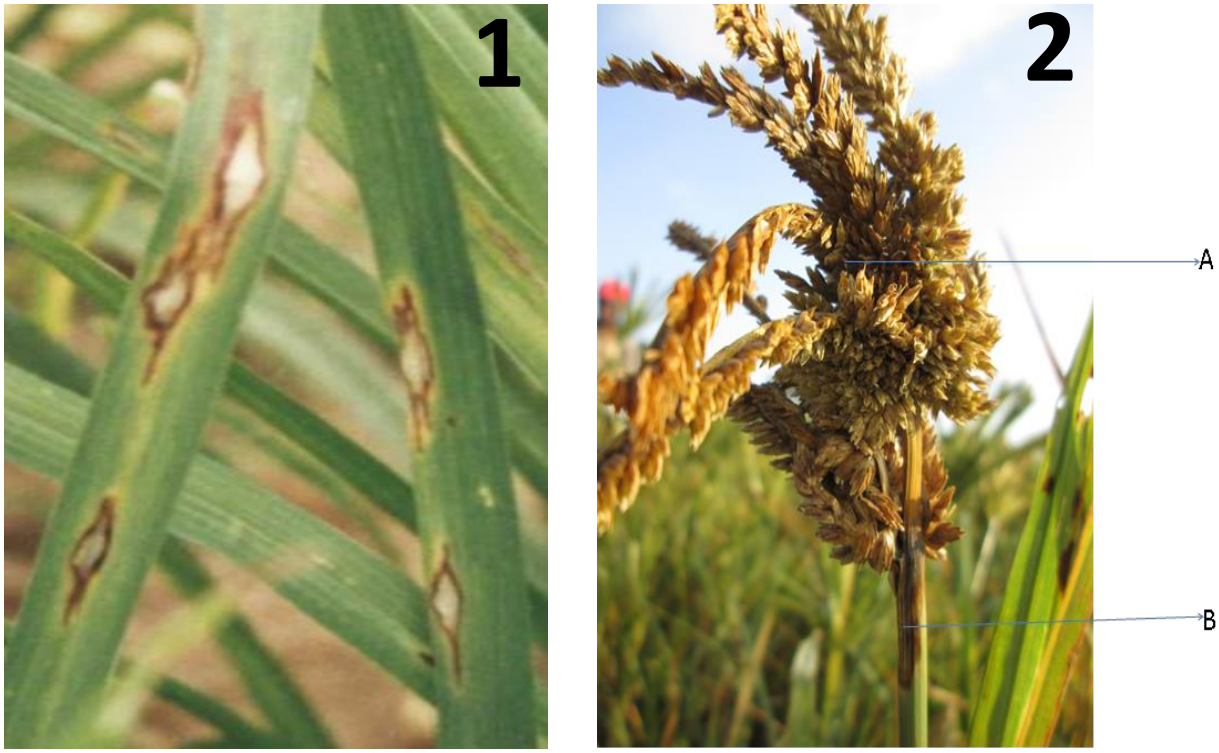

Figure 1.5. Blast disease symptoms caused by *M. oryzae* on finger millet

Lesions on the leaf is known as leaf blast [1], on the head as head blast (A) and on the neck as neck blast (B) [2].

1.4.3 Mode of transmission and infection cycle of *M. oryzae*

The conidia of the fungus *M. oryzae* are primarily transmitted through seeds (Tanaka *et al*., 2009) and through rain splash or plant-to-plant contact (Talbot, 2003; Chen *et al*., 2013a). In Africa and India, seed is considered as the primary source and mode of transmission (Afouda *et al*., 2009; Sere *et al*., 2011; Khanum *et al*., 2009). There are five recognised steps in the

fungal infection process. The process includes attachment to the plant (leaf) surface, germination on the plant surface, formation of infection structures, penetration of the host and colonisation of the host tissue (Schafer, 1994). The mechanism of *M. oryzae* infection process for finger millet is not yet available, but it is thought to be similar to the process reported on rice that has been well studied (e.g. Wilson and Talbot, 2009; Oses-Ruiz *et al*., 2017; Fernandez and Orth, 2018; Sakulkoo *et al*., 2018) as presented in Figure 1.6. After successful attachment of the spore to the surface of the leaf, the germ tube emerges across leaf surface and the germ tube then forms an appressorium (a specialised infection structure). The appressorium then penetrates the leaf cuticle by rupture using enormous turgor pressure followed by invasive growth. Successful colonisation of the leaf produces disease lesions from where the fungus sporulates and spreads to new parts of plant (Liu and Dean, 1997; Ribot *et al*., 2008; Chen *et al*., 2013a). However, the overall infection process largely depends on the environmental and developmental cues (Ribot *et al*., 2008). The infection stages involve certain genes, proteins, and structures and the function of these components in relation to the blast pathogen and rice host interactions have been well studied (Hamer *et al*., 1988; Bourett and Howard, 1990; Talbot, 1995; Talbot *et al*., 1996; DeZwaan *et al*., 1999).

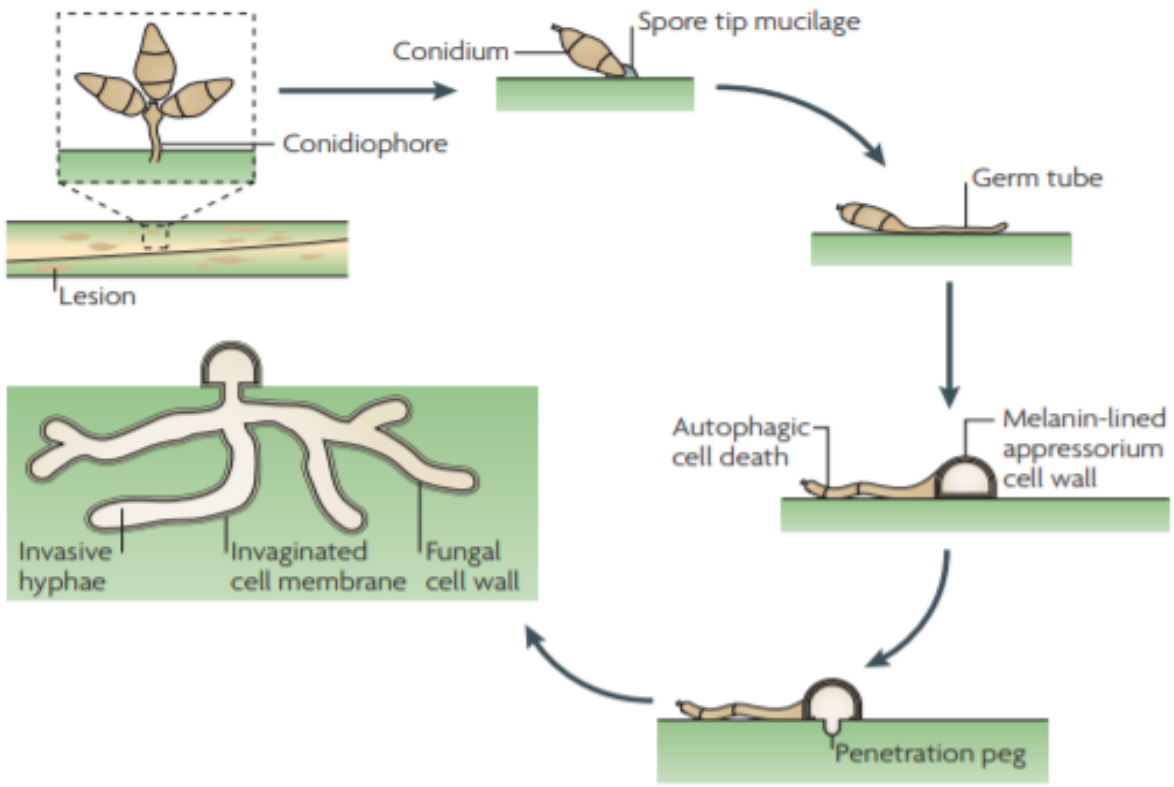

Figure 1.6. The infection cycle of the fungus *M. oryzae* as described with the rice blast system

The *M. oryzae* infection cycle starts when a three-celled conidium lands on the rice leaf surface and attaches itself to the hydrophobic cuticle before initiating germination. The conidium germinates and produces a narrow germ tube, which subsequently hooks at its tip, and then differentiates into an appressorium. The appressorium becomes melanised, generates internal turgor pressure and a penetration peg is formed at the base, which subsequently punctures the cuticle allowing entry into the rice epidermis. The appressorium matures and the conidium collapses and dies in a programmed process, involving autophagy. Invasion of rice tissue occurs by means of bulbous invasive hyphae that invaginate the rice plasma membrane and colonise epidermal cells. Disease lesions occur between 72 - 96 hours after infection with sporulation occurring under humid conditions. Spores are produced in the aerial conidiophores and spread to new rice plants by wind or dewdrop splashes. The image was adopted from Talbot and Wilson, 2009.

1.4.4 Environmental factors favouring the disease

The pathology of the blast fungus on finger millet is yet to be studied in detail and most of the information available is from blast disease on rice with some details for other cereal crops and grasses. The availability of free water as dew favours all the vital processes of the blast fungus disease cycle, such as spore production, release, adhesion, germination and host infection (Ou, 1980). Free moisture on the surface of the host plant for up to 24 hours was essential for *M. oryzae* infection on rye-grass and rapid expansion of the lesion on the surface of the host plant (Trevathan *et al*., 1994). Another report on the rye-grass blast showed that increase in the duration of leaf wetness at temperatures from $20^{O}C$ to $29^{O}C$ increased blast incidence and severity (Uddin *et al*., 2003). Besides dew, some common meteorological conditions that favour severe epidemics of rice blast include long periods of drizzling rain, lack of sunshine, moderate temperature, slow wind and high humidity, but heavy rain could be a limiting factor (Calvero, 1994; Suzuki, 1975). Temperature has however been identified as the most important determinant of the disease and affects the interaction between the host and the pathogen. In fact, it is a critical factor in the blast disease triangle of the environment, host and pathogen (Scholthof, 2007). Another important factor favouring the disease is the nitrogen level applied to the host plant. Higher levels of nitrogenous material either present in soil or applied as fertilizer lead to increased incidence of blast and predispose the plant to the disease (Osuna-Canizalez *et al*., 1991; Long *et al*., 2000). It has been reported that favourable weather conditions increased the epidemic of the blast fungus *M. oryzae* on finger millet (Patel and Tripathi, 1998). Continuous wetness and high humidity over 16 to 24 hours favour maximum infection of rice and finger millet at $22^{O}C$ - $30^{O}C$ (Andersen *et al*., 1947; Patel and Tripathi, 1998).

1.5 Control of the blast disease

In view of the high economic importance of this pathogen, various methods are being adopted to effectively manage the disease. The limited information available on finger millet blast management is discussed below, in the context of a wealth of research carried out on other cereal crops such as rice. The disease is mainly managed through cultural methods, planting of resistant varieties and use of fungicides although a limited level of biological control has been tested.

1.5.1 Cultural practices

In East Africa, the traditional practice of managing the finger blast differs from farm to farm, location to location and country to country. Planting of blast free certified seeds, manipulation of planting time and moderate amount of fertilizer application have been found useful by some farmers (Finckh, 2008; Obilo *et al*., 2012; Sood and Babu, 2016). A combination of organic and inorganic fertilizer at one quarter of the recommended rate is found to be less expensive and has been reported to minimise blast infection (Dass *et al*., 2013; Mgonja *et al*., 2013; Das, 2017). Managing seed rates, increasing the spacing and weeding of the finger millet fields two or three times a season to eliminate alternate weedy hosts are known to reduce finger millet blast disease levels (Lenne and Thomas, 2006; Oduori, 2008). Crop rotation also aids blast disease management. Intercropping system is another form of cultural practice that has been effectively used to manage blast disease (Mudita *et al*., 2008). For example, finger millet planted with pigeon peas, produced good yields. This result was attributed to differences in growth cycles and increased competition for nitrogen (Adipala *et al*., 1994).

1.5.2 Chemical control

Fungicides are frequently and widely used to manage fungal diseases of crops worldwide (Pooja and Katoch, 2014). Blast disease is commonly managed by fungicides especially on rice worldwide, particularly in Japan and China (Chen *et al*., 2013b; Mew *et al*., 2004). However, fungicides are less used to manage blast disease in Latin America, South and East Asia and Africa (Mew *et al*., 2004). Commonly used fungicides for the management of blast disease particularly on rice include benomyl, tricyclazole, mancozeb, carbendazim, pyroquilon, azoxystrobin, zineb, propiconazole, edifenphos, isoprothiolane, tolprocarb, fenoxanil, blasticidin, hinosan and ferimzone (Yashoda *et al*., 2000; Chen *et al*., 2013b, Hamada *et al*., 2014; Chen *et al*., 2015; Lewis *et al*., 2016; Mew *et al*., 2004; Srivastava *et al*., 2017; FRAC 2018). A combination of fungicides such as carbendazim with mancozeb (systemic and contact) has proved to be very effective in providing resistance against blast disease on rice (Yashoda *et al*., 2000; Ghazanfar *et al*., 2009; Miah *et al*., 2018). Fungicides are applied directly to the host plant by spraying or through seed treatment which later protect other parts of the plant. These chemicals act by inhibiting and block processes such as appressorial melanization, mitochondrial respiration and mitosis in *M. oryzae* (Kurahashi, 2001; Kunova *et al*., 2013; Chen *et al*., 2013b; Kunova *et al*., 2014; Chen *et al*., 2015). Some fungicides also interfere with the enzymatic processes essential to melanin biosynthesis (e.g. Kunova *et al*., 2014). However, the effectiveness of these chemicals are determined by factors such as the composition of the fungicides, time and method of application, stability of the fungicides, the level of disease and the rate of emergence of fungicide resistant strains (Skamonioti and Gurr, 2009). The Fungicide Resistance Action Group (FRAC, 2018),

categorised *M. oryzae* as one of the pathogens capable of evolving resistance to fungicides within a short period. A series of breakdown of some of the fungicides by the blast pathogen has been reported in China and Brazil (Sawada *et al*., 2004; Zhang *et al*., 2009; Chen *et al*., 2013b; Castroagudin *et al*., 2015). Although the use of fungicides has proven to be very effective in managing this disease, the use of synthetic fungicides continues to be discouraged and some chemicals have been banned due to their negative effects on the ecosystem (Zarandi *et al*., 2009).

1.5.3 Biological control

In view of the environmental concerns associated with chemical fungicides, efforts continue to be made to develop alternative control measures also in the context of peoples' health (Law *et al*., 2017). Biological control offers scope as a practical and economic alternative for the management of plant diseases including blast (Lis-Balchin and Deans, 1997; Manidipa *et al*., 2013). Most of the microbes that have been successfully used in controlling plant diseases employ mechanisms such as competition, suppression, antibiosis, mycoparasitism, hypervirulence and induced resistance (Karthikeyan and Gnanamanickan, 2008a; Singh *et al*., 2016; Negi *et al*., 2017; Law *et al*., 2017). Some commonly used biological agents (bio-inoculant) to control fungal crop diseases include *Pseudomonas* species (Karthikeyan and Gnanamanickan, 2008a; Dorjey *et al*., 2017*), Bacillus* species (Karthikeyan and Gnanamanickan, 2008a; Prasanna Kumar *et al*., 2017)*, Trichoderma* species (Kumar *et al*., 2016; Waghunde *et al*., 2016) and *Streptomyces* species (Zarandi *et al*., 2009; Li *et al*., 2011; Boukaew and Prasertsan, 2014). For example, finger millet, foxtail millet and rice seeds treated with *Pseudomonas fluorescens* Pf7-14 showed good response and suppressed blast

infection at high percentage of 80-88% and reduced rice blast severity by 21 to 68.5 % (Karthikeyan and Gnanamanickan, 2008a). Similar reports from two different research groups in Iran and Egypt, showed effectiveness of *Streptomyces* species to antagonise blast disease on rice plant (Zarandi *et al*., 2009; Khalil *et al*., 2014). Additionally, studies have shown the capability of bio-inoculants to increase the root and shoot growth as well as flowering and maturity of the crop plants through symbiotic association (Karthikeyan and Gnanamanickan, 2008a; Singh *et al*., 2016; Tamreihao *et al*., 2016; Negi *et al*., 2017; Amruta *et al*., 2018). Nonetheless, weather condition is a major determinant for the effectiveness of a biological control agent as unfavourable conditions can render it ineffective. Furthermore, the possibility of the pathogen developing resistance rendering the treatment non-effective and non-economic needs to be carefully considered (Vasudevan *et al*., 2002; Suprapta, 2012; Gopalakrishnan *et al*., 2014).

1.5.4 Host resistance

Host resistance is the ability of a host plant to hinder or resist colonization by a pathogen (Robinson, 1969). Generally, the host plant relies on the innate immunity of each cell and systematic signals emanating from the infection site introduced by the pathogens (Jones and Dangl, 2006). Host plants do this using two modes of resistance (Jones and Dangl, 2006; Cook *et al*., 2015). The first mode recognises and responds to molecules common to many classes of microbes including non-pathogens and utilises transmembrane pattern recognition receptors (PRRs) that respond to slowly evolving microbial or pathogen-associated molecular patterns (MAMPs or PAMPs) (Jones and Dangl, 2006; Cook *et al*., 2015).

On the other hand, the second mode recognises and responds to pathogen virulence factors either directly or indirectly through their effects on host targets. This mode acts predominantly inside the cell, using the polymorphic nucleotide binding (NB)-leucine rich repeat (LRR) protein products encoded by most R genes, and are broadly characterised based on the NB and LRR domains (Jones and Dangl, 2006; Cook *et al*., 2015). Examples of this second mode are the gene for gene (direct) and guard (indirect) hypotheses (Jones and Dangl, 2006; Cook *et al*., 2015).

A number of conceptual models of the plant immune system has been developed and described over the last 2 decades (e.g. Nurnberger and Brunner, 2002; Chisholm *et al*., 2006; Postel and Kemmerling, 2009). Although these models vary, they are all grounded in the observation that the plant innate immune system is largely controlled by receptors that identify invasion by the pathogens (Cook *et al*., 2015). Among these models, the zig-zag model has been identified as the versatile tool that explained most host pathogen interactions and also reconciles the previously described gene for gene concept with the recognition of general elicitors (Jones and Dangl, 2006; Cook *et al*., 2015; Keller *et al*., 2016). The zig-zag model is based on principles underlying 4 phases (Figure 1.7). In phase 1, PAMPs or MAMPs are recognised by the PRRs and this results in PAMP-triggered immunity (PTI). In phase 2, successful pathogens deploy effectors that contribute to pathogen virulence. Effectors can interfere with PTI, which results in effector-triggered susceptibility (ETS). In phase 3, a given effector is specifically recognised by one of the NB-LRR proteins, resulting in effector-triggered immunity (ETI). The recognition is either indirect (guard hypothesis), or through direct NB-LRR recognition of an effector. The ETI accelerates and amplifies PTI, resulting in disease resistance and this is usually via hypersensitive cell death response (HR)

at the infection site. In phase 4, natural selection drives pathogens to avoid ETI either by shedding or diversifying the recognised effector gene, or by acquiring additional effectors that suppress ETI. The natural selection in the host results into a new R specificity so that ETI can be triggered again.

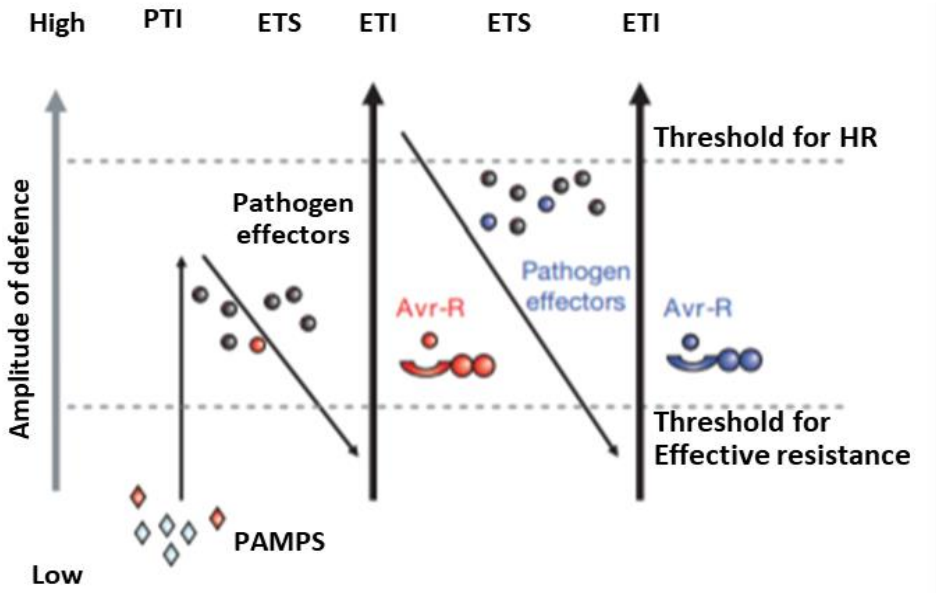

Figure 1.7. Schematic diagram of the zig-zag model

The model illustrates the quantitative output of the plant immune system based on the four phases. In phase 1, host plant detects microbial/pathogen-associated molecular patterns (MAMPs/PAMP, in red diamonds) through pattern recognition receptors (PRRs) to trigger PAMP-triggered immunity (PTI). In phase 2, successful pathogens deliver effectors that interfere with PTI, which results in effector-triggered susceptibility (ETS). In phase 3, an effector (in red) is recognised by an NB-LRR protein activating effector-triggered immunity (ETI), leading to an amplification of PTI and hypersensitive response (HR). In phase 4, pathogen isolates evolve to gain new effectors to suppress ETI. Natural selection favours new plant NB-LRR alleles that can recognise one of the newly acquired effectors, resulting again in ETI. The image was adopted from Jones and Dangl, 2006.

Several reports have recommended the use of resistant seed varieties as the most advantageous and environmentally friendly method for blast management on rice (e.g. Bonman, 1992; Sere *et al*., 2011; Miah *et al*., 2017). Two major types have been recognised

and used in rice resistance development programs, these are: qualitative resistance (also referred to as vertical or complete resistance) and quantitative resistance (also referred to as horizontal or partial resistance) (Marone *et al*., 2013; Pilet-Nayel *et al*., 2017). While qualitative resistance is conferred by the host resistance genes (R genes), which recognise and interact with the pathogen avirulence genes (AVR genes), quantitative resistance is mediated by quantitative traits loci (QTL) that confer durable non-race specific resistance. Generally, qualitative resistance is pathogen race-specific and its durability is limited by strong selection pressure imposed on the pathogen (Marone *et al*., 2013; Pilet-Nayel *et al*., 2017). More than 70 genes and 347 QTLs involved in *M. oryzae* - rice interactions have been identified using a range of molecular techniques (Koide *et al*., 2009; Debnath *et al*., 2018).

Planting disease-resistant varieties is recognised as the most economic and effective means of combating *M. oryzae* causing blast on finger millet (Babu *et al*., 2013). Less information is available on the varieties of finger millet that are resistant to blast compared to other crops such as rice. In Uganda, pure line selection was used to identify and release finger millet varieties such as Serere-1, Gulu-E, PESE I, Engeny, SEREME I, SEREMI II, SEREMI III as resistant to head blast (Esele and Odelle, 1992; Esele and Odelle, 1995, Aru *et al*., 2015). In finger millet, although some reports of genetically engineered resistance to leaf blast exist (Latha *et al*., 2005; Ignacimuthu and Ceasar, 2012), there is only limited information on the R gene – AVR gene mediated resistance (Babu *et al*., 2015). In this context, it is vital to gain an in-depth understanding of the finger millet blast pathogen population biology and genetics in key cropping locations including Eastern Africa – the primary centre of finger millet diversity as well as India – the secondary centre of the crop diversity.

1.6 Population biology of the blast pathogen *M. oryzae*

The blast pathogen *M. oryzae* is known to breakdown host resistance after a relatively short period, which has become a major challenge in combating the disease on rice (Talbot, 1995; Talbot, 2003). Therefore, a wealth of research has focused on the population biology of the rice blast pathogen using molecular and biological assays in order to gain insight into the mechanisms involved in host resistance breakdown. This has been attributed to the genetic diversity of the *M. oryzae* populations prevailing in a geographic location also linked to wide spread occurrence of transposable elements and repetitive sequences in the genome, and the reproductive behaviour (Correa-Victoria and Zeigler, 1993; Urashima *et al*., 1999; Park *et al*., 2010; Islam *et al*., 2016). Intensive efforts into the genome sequencing of an array of *M. oryzae* isolates has provided a platform to gain further insights into the critical aspects of the pathogen genetics and evolution (Dean *et al*., 2005; Xue *et al*., 2012; Chen *et al*., 2013a; Dong *et al*., 2015; Gladieux *et al*., 2018).

1.6.1 Reproductive biology: distribution of mating types and fertility of *M. oryzae*

The fungus *M. oryzae* possesses the ability to reproduce asexually and sexually under specific conditions (e.g. Viji *et al*., 2001; Saleh *et al*., 2012a). In nature, asexual reproduction is common and plays a significant role in the multiplication and dispersal of the blast pathogen (Duncan *et al*., 1998). There is no general consensus as yet on whether *M. oryzae* reproduces sexually in nature. Sexual heterothallism of this fungus was first demonstrated early on (Hebert, 1971). Being heterothallic, sexual reproduction requires crossing between individuals of opposite mating types, which are determined by the idiomorphic forms of the mating type (*MAT*) locus. The mating type gene of *M. oryzae* has been cloned and

sequenced using a genomic subtraction technique. The *MAT1-1* and *MAT1*-2 genes are approximately 2.5 and 3.5 kb in length, respectively (Kang *et al*., 1994). These genes have been used as markers in subsequent population analysis studies (e.g. Viji and Gnanamanickan, 1998; Consolo *et al*., 2005; Takan *et al*., 2012).

The mating type of each strain of *M. oryzae* can be determined by conventional approach of crossing with a standard tester of known mating type (Notteghem and Silue, 1992) or by molecular analysis such as mating type-specific PCR (e.g. Takan *et al*., 2012). The presence of both mating types in a geographic location alone is not a clear evidence for active sexual reproduction (Couch *et al*., 2005). Sexual reproduction assays carried out *in vitro* showed that at least one of the strains must be female fertile (Saleh *et al*., 2012a, Saleh *et al*., 2012b). In this fungus, successful cross between strains of opposite mating types MAT 1-1 and MAT 1-2 should lead to the production of the fruiting body structures perithecia with viable ascospores (Coppin *et al*., 1997; Viji and Gnanamanickan, 1998). Fertility in *M. oryzae* field isolates can range from total sterility (inability to mate with any other strain), through female sterility (ability to mate only as a male), to full fertility (hermaphrodite with the ability to mate as both male and female) as shown previously (Itoi *et al*., 1983; Valent *et al*., 1991; Notteghem and Silue, 1992; Viji and Gnanamanickan, 1998).

Kumar *et al*. (1999) investigated the rice blast pathogen populations in the Indian Himalayan foothills – a centre of origin of the crop in South Asia, and reported the occurrence of both mating types MAT 1-1 and MAT 1-2 and female fertile strains capable of sexual reproduction. It is important to highlight that the occurrence of female fertile strains in rice production systems is not common. Recent work has provided further evidence for the occurrence of sexual reproduction in rice pathogenic strains of *M. oryzae* in South East Asia

(Saleh *et al*., 2012a). However, *M. oryzae* isolates that infect rice are most commonly female sterile (Saleh *et al*., 2012b), whereas finger millet blast isolates and the grass isolates were female fertile and hermaphrodites (e.g. Takan *et al*., 2012).

The distribution of the mating types and the fertile strains of *M. oryzae* appears to depend on the host and geographic location also linked to the centre of origin and domestication of the crop (Saleh *et al*., 2012b; Takan *et al*., 2012; Onaga *et al*., 2015; Mwongera, 2018). Presence of both MAT 1-1 and MAT 1-2 isolates, although skewed to some extent, associated with rice cultivars has been reported in India, China, Vietnam, and Thailand, where the crop has been grown for thousands of years (Viji and Gnanamanickan, 1998; Mekwatanakarn *et al*., 1999; Le *et al*., 2010; Samanta *et al*., 2014; Imam *et al*., 2014). Separate studies in Africa showed the dominance of MAT 1-2 (71 %) isolates on rice in West Africa (Takan *et al*., 2012), whereas MAT 1-1 isolates were more prevalent in East Africa (76 %) on rice (Onaga *et al*., 2015; Mwongera, 2018). Whereas in Korea, Brazil and Argentina, where rice cultivation is relatively recent, isolates of only one mating type were found among collections ranging from 100 to 500 (Consolo *et al*., 2005; Park *et al*., 2008; Peixoto *et al*., 2014; D'Avila *et al*., 2016). Further, within Brazil, both mating types MAT 1-1 and MAT 1-2 among *M. oryzae* associated with wheat and the potential for sexual reproduction and recombination has been strongly suggested (Maciel *et al*., 2014). Likewise, wide distribution of mating types MAT1-1 and MAT1-2 isolates and dominance of fertile isolates of *M. oryzae* were observed with finger millet in Kenya and Uganda (Takan *et al*., 2012). This indicates a strong potential for sexual reproduction and recombination, necessitating further detailed investigations in key finger miller production locations in Eastern Africa.

1.6.2 Repetitive DNA elements in the blast pathogen *M. oryzae*

Repetitive DNA including transposable elements, interspersed repeats and telomeres are abundant in many eukaryotic organisms (Wostemeyer and Kreibich, 2002; Chadha and Sharma, 2014). These regions of the genome are more variable compared to other parts, which might be linked to positive selection pressure (Thon *et al*., 2006). Blast pathogen *M. oryzae* is rich in repetitive DNA elements reflecting approximately 9.7 % of the genome and the majority of the repetitive elements are retrotransposons comprising of 8 families (Dean *et al*., 2005; Thon *et al*., 2006). Magnaporthe grisea repeat (MGR) was the first repetitive-DNA family identified in the genome of *M. oryzae* distributed among all the chromosomes (Hamer *et al*., 1989; Dean *et al*., 2005). MGR sequence is a retrotransposon element that resembles those in other Ascomycetes. Two members of the MGR were initially identified, namely MGR583 and MGR586, which were highly conserved among the rice blast isolates compared to isolates from non-rice hosts (Hamer *et al*., 1989). MGR586 in DNA fingerprinting revealed polymorphic patterns capable of distinguishing *M. oryzae* isolates from rice blast into various lineages (Levy *et al*., 1991; Farman *et al*., 1996b). Further studies showed that the MGR586 contained the Pot3 retrotransposon and possessed an open reading frame, which showed 23 % and 36 % amino acid identity with transposons Pot2 within *M. oryzae* and Fot1 from *Fusarium oxysporium*. However, these 3 elements have little or no similarity at the DNA sequence level (Farman *et al*., 1996b). Furthermore, absence of the MGR586 sequence in grass-infecting isolates, but presence in low copy in *M. oryzae* isolates from other cereal crops were shown (Hamer *et al*., 1989; Farman *et al*., 1996b). Most of the retrotransposons are transposable elements (TEs), which have the ability to move from one chromosome to another in the genome (Chardha and Sharma, 2014; Jardim

et al., 2015), playing a major role in the genome variation of the *M. oryzae* (Zhang *et al*., 2013). These transposable elements are classified into two major groups. Class I TEs transpose via RNA intermediates by a 'copy and paste' mechanism and Class II TEs transpose directly from DNA to DNA by a 'cut and paste' mechanism (Kang *et al*., 2001; Chadha and Sharma, 2014). The class I transposons in *M. oryzae* are variable, ranging from retro elements, long terminal repeat (LTR) retrotransposons also known as the gypsy class, and non-LTR retrotransposons. The non-LTR retrotransposons include two types namely long interspersed nuclear element (LINE), and short interspersed nuclear elements (SINE).

It is well recognised that transposable elements (TEs) are associated with insertion, deletion and genomic rearrangements as well as other functions including organisation of chromatin structure, replication and transcription and RNA processing and stability, and importantly the ability to mobilise other sequences in the genome (Weiner *et al*., 1986; Deininger *et al*., 1992). After the initial discovery of MGR elements (e.g. Hamer *et al*., 1989), several transposable elements have been identified in the genome of *M. oryzae* and these have been cloned and characterised as detailed below (Table 1.1). The accumulation of different repeat elements in *M. oryzae* populations has been suggested to result in changes in the genome architecture that are fundamental for the emergence of new genotypes with potentially altered pathogenicity traits (Dobinson *et al*., 1993). Wide distribution of the transposable elements such as MGR586, *Pot*2, Mg-SINE and MGR583 among *M. oryzae* isolates from different hosts, suggests that these are old elements, which arose in or invaded the fungus at very early stages of its evolution (Dobinson *et al*., 1993; Eto *et al*., 2001). Conversely, limited distribution of the LTR retrotransposons, *Grasshopper* (*grh*) and MAGGY among *M. oryzae* isolates from some hosts only, suggests that these elements are

relatively new, which may have been acquired through horizontal transfer events (Eto *et al*., 2001).

In the context of finger millet blast, *grh* element is of considerable interest as it was almost entirely found in *M. oryzae* isolates infecting *Eleusine* species initially and also as these isolates were reported to occur only in certain regions of the world (Dobinson *et al*., 1993). Further work reported the presence of the *grh* element among the *M. oryzae* isolates from rye grass (Kusaba *et al*., 2006) and rice (Mahesh *et al*., 2016). The regions where *M. oryzae* isolates containing the *grh* element were found include parts of Asia (India, Nepal and Japan), West Africa (Mali and Burkina Faso) and East Africa (Uganda and Kenya). However, their frequency of occurrence in these locations varied considerably (Dobinson *et al*., 1993; Kusaba *et al*., 2006; Takan *et al*., 2012). Presence of the *grh* element in rice blast isolates correlated to gene flow between the rice and non-rice *M. oryzae* isolates (Mahesh *et al*., 2016). Transposable DNA elements such as MGR586, MGR583, *Pot*2, MAGGY and Grasshopper have contributed greatly to the understanding of the population diversity pattern of *M. oryzae* from different hosts (Farman *et al*., 1996a; Nitta *et al*., 1997, Don *et al*., 1999; Dioh *et al*., 2000; Tosa *et al*., 2004; Sawada *et al*., 2004, Park *et al*., 2010) as detailed in section 1.6.4.

Table 1. 1 Repetitive DNA elements identified in the genome of the blast pathogen *M. oryzae*

Name of the Transposable element	Class	Class Name	Plant/Crop Host	Total length (bp)	Reference
Retro5	I	Retroelement	Rice/*Setaria*/ rye grass	7623	Farman, 2002
Retro6	I	Retroelement	Rice	5705	Farman, 2002
Retro7	I	Retroelement	Rice	6210	Farman, 2002 Dean *et al*., 2005
Grasshopper (*grh*)	I	LTR[a](gypsy)	*Eleusine* sp., Annual rye grass	5233	Dobinson *et al*., 1993 Kusaba *et al*., 2006
MAGGY	I	LTR (gypsy)	Rice and *Setaria spp*	5638	Farman *et al*., 1996a
MGLR-3	I	Gypsy	Rice	6304	Kang *et al*., 2001
Pyret	I	Gypsy	Rice	7250	Nakayanshiki *et al*., 2001
Inago 1	I	Gypsy	Rice	NS**	Sanchez *et al*., 2011
Inago 2	I	Gypsy	Rice		Sanchez *et al*., 2011
MGL or MGR 583	I	Non LTR (LINE[b])	Rice/Non-rice	5977	Hamer *et al*., 1989
MGSR I	I	Non LTR(SINE[c])	Rice/Non-rice	1111	Sone *et al*., 1993
MG-SINE	I	Non LTR (SINE)	Rice/Non-rice	472	Kachroo *et al*., 1995
MGR586	II	Terminal inverted repeat	Rice	NS**	Hamer *et al*., 1989
*Pot*2	II	Terminal inverted repeat	Rice	1857	Kachroo *et al*., 1994
*Pot*3	II	Terminal inverted repeat	Rice	1860	Farman *et al*., 1996b Kang *et al*., 2001
*Pot*4	II	Terminal inverted repeat	Rice	1858	Dean *et al*., 2005 Rehmeyer *et al*., 2006
Occan	II	Terminal inverted repeat	Rice	2686	Kito *et al*., 2003 Sone *et al*., 2013
MoTeRs	I	Non LTR	Rice/Non-rice	1886	Starnes, 2013 Farman, 2007

[a] long terminal repeat, [b]long interspersed nuclear element, [c]short interspersed nuclear element.

**NS Not specified

1.6.3 Avirulence (*AVR*) genes in the blast pathogen *M. oryzae*

In the gene-for-gene crop-pathogen systems, knowledge of the pathogen *AVR* genes is crucial to the development of host resistance. It is well documented that *AVR* genes encode secreted proteins, which are specifically recognised by the host resistance (R) genes. Recognition of the *AVR* gene products by the plant surveillance system mediated by R genes triggers the host defence mechanism leading to hypersensitive response (HR) and pathogen localisation (Dodds and Rathjen, 2010; Kutcher *et al*., 2010; Bialas *et al*., 2017; Upson *et al*., 2018). To date, more than 40 *AVR* genes have been genetically mapped in the genome of the rice blast pathogen *M. oryzae* (Korinsak *et al*., 2018). Of these, 12 *AVR* genes (Table 1.2) have been cloned and characterised with reference to their corresponding R genes in the host plant (Ray *et al*., 2016; Terauchi *et al*., 2016; Wang *et al*., 2017a). The majority of the cloned *AVR* genes encode small proteins (<200 aa) except *AVR-Pita* and *ACE1* (Wang *et al*., 2017a). *M. oryzae AVR* genes are defined based on the protein structure such as glycine-rich e.g. *AVR-Piz-T* and *AVR-Pi54* (Li *et al*., 2009; Zhang *et al*., 2013; Ray *et al*., 2016), Zinc metalloprotease e.g. *AVR-Pita* (Orbach *et al*., 2000), polyketide synthase (PKS) and non-ribosomal peptide synthetase (NRPS) hybrid e.g. *ACE*1 (Bohnert *et al*., 2004).

Availability and comparative analysis of the genome sequence data of an array of *M. oryzae* isolates is beginning to provide new insights into the genomic occurrence and structural characteristics of the loci containing the *AVR* genes. For example, *AVR* genes are commonly flanked on one or both sides by repeat elements e.g. Pot3 transposable element flanking the *AVRPiz-t* at the 5' end. Likewise, *AVR* genes commonly occur close to the chromosome ends at the telomere regions, which are rich in repetitive sequences (Rouxel and Balesdent, 2017; Wang *et al*., 2017a). The regions at which *AVR* genes are located seem to undergo

rapid change and/or rearrangements leading to the diversity or changes in the pathogen effector repertoire enabling them to escape recognition by the plant R genes (Pais *et al*., 2013; Sanchez-Vallet *et al*., 2018). Similarly, transposable elements have been implicated in the loss or gain or disruption of *AVR* genes in *M. oryzae* through positive selective pressure (Yoshida *et al*., 2009; Dai *et al*., 2010; Wang *et al*., 2017a). In rice-blast pathogen interactions, it has been established that in the genomes of various *M. oryzae* isolates the *AVR* genes can be either present or absent or exist in different allelic forms (Yoshida *et al*., 2009; Wu *et al*., 2014; Sirisathaworn *et al*., 2017). The presence or absence or existence of allelic forms of *AVR* genes has been analysed using PCR based methods and BLAST analysis (Couch *et al*., 2005, Yoshida *et al*., 2009; Sirisathaworn *et al*., 2017). In the finger millet blast system, *AVR* – R gene interactions remain an unexplored area. Only recently, an initial report of the existence of 9 *AVR* genes in *M. oryzae* isolates from Eleusine has become available (Wang *et al*., 2017a). The genomic resources being established for specifically selected *M. oryzae* isolates representing the finger millet blast pathogen population diversity in Eastern Africa would serve as a sound platform to gain further insights into this key area.

Table 1.2 Details of *AVR* genes cloned from *M. oryzae* isolates associated with rice blast

***AVR* gene**	**Isolate source**	**Protein size (No. AAs)**	**Chromosome**	**Reference**
PWL1	Rice blast	147	2	Kang *et al*., 1995
PWL2	Rice blast	145	2	Sweigard *et al*., 1995
AVR1-Co39	Rice blast	89	1	Farman and Leong, 1998; Ribot *et al*., 2013
AVR-Pita	Rice blast	224	3	Khang *et al*., 2008
ACE1	Rice blast	4035	1	Bohnert *et al*., 2004
AVR-Pia	Rice blast	85	5 or 7	Miki *et al*., 2009; Yoshida *et al*., 2009
AVR-Pii	Rice blast	70	7	Yoshida *et al*., 2009
AVR-Pik/km/kp	Rice blast	113 (5 alleles) A-E	1	Yoshida *et al*., 2009; Wu *et al*., 2014
AVR-Piz-T	Rice blast	108	7	Li *et al*., 2009
AVR-Pi9	Rice blast	91	7	Wu *et al*., 2015
AVR-Pib	Rice blast	75	3	Zhang *et al*., 2015
AVR-Pi54	Rice blast	153	4	Devanna *et al*., 2014 Ray *et al*., 2016

AAs, Amino acids

1.6.4 Blast pathogen diversity and population structure

The continuing threat posed by *M. oryzae* to global food and nutrition security is well recognised. Research over the last two and half decades globally has focused on unravelling the blast pathogen diversity and population structure to gain new insights into its biology and evolution. The aim clearly is to identify effective strategies for the development and deployment of resistant varieties. Most of these efforts have focused on the rice blast system due to its economic importance globally.

Pathogen population structure can be defined as the level of variation observed among individuals reflecting their evolutionary history and the potential to evaluate the evolutionary relationships within and between sub-populations (McDonald and McDermott, 1993; Leung *et al*., 1993; Chen *et al*., 1995; McDonald, 1997). Variation within pathogen populations is mainly assessed based on the phenotype (e.g. morphology or toxin production) or pathotype (e.g. host or cultivar specificity) or genotype (DNA data) or a combination of these characteristics. In view of the huge variation known in *M. oryzae* including spontaneous changes during regular sub-culturing, phenotypic characterisation particularly based on morphological appearance has major limitations and is not widely pursued. The conventional method widely used with the blast pathogen *M. oryzae* particularly in the context of rice is a pathotyping assay. This is carried out on a set of differentials each carrying a known resistance gene. This method provides insight into the pathogen virulence diversity and the pathotype structure of the population. However, this method is resource intensive and is also strongly influenced by environmental variables and the host genetic background. In addition, these assays also require strong pathology expertise to avoid subjective scoring during experimentation. However, pathotyping is an

invaluable method that is commonly employed to understand the pathogen diversity including the *AVR* - R gene interaction patterns in rice blast. This approach is particularly facilitated by the recent development of mono-isogenic lines each containing a known R gene (Mutiga *et al*., 2017; Farman *et al*., 2017; Perello *et al*., 2017).

Advances in molecular technologies and the need to answer fundamental questions relating to the population structure of the blast pathogen has led to the development and application of various approaches. The aim is to use these techniques to unravel DNA sequence diversity embedded within the genome of the pathogen (Babujee and Gnanamanickam, 2000). The range of molecular techniques that have been used to study the population patterns of *M. oryzae* can be broadly categorised as DNA band-based methods (DNA profiling or DNA fingerprinting) and DNA sequence-based methods (DNA barcoding).

The DNA profiling methodologies used to characterise *M. oryzae* isolates include DNA hybridisation e.g. restriction fragment length polymorphisms (RFLPs), and PCR-based methods (Table 1.3). A key example of the DNA hybridisation method is MGR586 fingerprinting originally developed by Levy *et al*. (1991) to define lineages of the rice blast pathogen populations. A range of PCR-based methods such as random amplified polymorphism DNA (RAPD), repetitive DNA-based polymerase chain reaction (Rep-PCR), amplified fragment length polymorphism (AFLP), microsatellite or simple sequence repeat (SSR) analysis, sequence characterised amplified region (SCAR), single strand conformation polymorphism (SSCP) and retrotransposon microsatellite polymorphism (REMAP) have been used to study the molecular genetic diversity of *M. oryzae* from different hosts and diverse geographic locations (Couch *et al*., 2005; Chadha and Gopalakrishna, 2007; Sere *et al*., 2007;

Silva *et al*., 2009; Takan *et al*., 2012; Maciel *et al*., 2014; Motlagh *et al*., 2015; Onaga *et al*., 2015). DNA profile methods are based on scoring the presence or absence of similarly sized/positioned bands (McDonald, 1997). The PCR-based techniques operate on different principles, some of which do not require prior knowledge of DNA sequence to design the primers e.g. RAPD and AFLP (McDonald, 1997; Xu, 2006). For example, AFLP analysis showed positive resolution among the population of *M. oryzae* isolates associated with Tall fescue St. Augustine grass in Georgia (Tredway *et al*., 2005) and finger millet and related weed hosts in Kenya and Uganda (Takan *et al*., 2012). Conversely, methods such as Rep-PCR require the DNA sequence data of a specific repetitive element to design bespoke primers (George *et al*., 1998).

With the advances in DNA sequencing technologies, direct sequencing of PCR products has been used extensively to analyse the intra- and inter- species differences in diverse fungal species including *M. oryzae* (Couch and Kohn, 2002; Xu, 2006; Sucher *et al.,* 2012). Direct sequencing of PCR products only requires small quantities of DNA and offers high resolution and very little or no ambiguity at manageable costs making it an effective approach to investigate fungal pathogen genetic diversity and phylogenetic relationships relatively rapidly (Hibbett *et al*., 2016, Robert, 2016). Availability of a large number of DNA sequences of various marker genes or regions from diverse fungal species in the public databases serves as a significant resource. The most commonly used markers are from the ribosomal RNA (rRNA) gene block that comprises highly conserved genes and variable inter-genic regions (White *et al*., 1990; Schoch *et al*., 2012). This includes the 18S rRNA gene (small subunit), 28S rRNA gene (large subunit), internal transcribed spacer (ITS) and inter-genic spacer (IGS). Markers from the rRNA gene block have been used extensively to characterise

fungal phylogenetic relationships and the ITS maker is recognised as a fungal universal DNA barcode (Schoch *et al*., 2012). To achieve higher levels of resolution as well as to account for the different evolutionary rates among different parts of the genome, various protein-coding genes have also been used to develop molecular genetic markers to investigate the pathogen population diversity and structure. Commonly used markers include housekeeping genes such as actin (ACT), translation elongation factor 1-α (TEF1), β-tubulin (TUB), histone (HIS3), calmodulin (CAL) and RNA polymerase II (RPB1) glyceraldehyde-3-phospate dehydrogenase (gpd) in various fungi including *M. oryzae* (e.g. Klaubauf *et al*., 2014). Complementing the housekeeping gene markers, functional genes such as MAT genes and MPG1 gene have also been used with *M. oryzae* to characterise populations prevailing in diverse biotic and abiotic environments in different geographic locations (Couch *et al*., 2005). For example, markers from the Actin, Calmodulin and Beta-tubulin genes have been used to distinguish *M. oryzae* and *M. grisea* species associated with various cereal crops and grasses (Couch and Kohn, 2002; You *et al*., 2012).

Characterisation of *M. oryzae* isolates from different geographic locations using a range of molecular techniques has been effective in revealing the population diversity and structure, particularly from rice blast (Table 1.3). A clear lineage-based structure of *M. oryzae* populations has been deciphered in the United State of America (USA), Columbia, Iran, West Africa (Ghana, Burkina Faso, Nigeria, and Cote d'Ivoire), Philippines and Europe. Some key trends in diversity and distribution pattern of the blast pathogen lineages have emerged from these studies (Table 1.3). The length of rice cultivation and the pattern of use of rice varieties had a major influence in shaping the lineage diversity. For example, in Africa, Europe and USA, where the history of rice cultivation ranges between 200 to 500 years, only

a limited number of lineages were identified (Levy *et al*., 1991; Levy *et al*., 1993; Roumen *et al*., 1997; Takan *et al*., 2012; Mwongera, 2018). Conversely, in parts of Asia with thousands of years of rice cultivation history, a huge diversity in blast pathogen lineages has been recorded as in the case of India and Thailand (Kumar *et al*., 1999; Mekwatanakarn *et al*., 1999; Correll *et al*., 2000). Furthermore, analysis of rice blast populations from the same geographic locations but different periods revealed changes in pathogen lineages. For example, in Korea two lineages identified among isolates collected during 1990s were different from lineages detected in 1970s (Park *et al*., 2008). Similarly, in Europe (rice blast in France, Spain, Italy, Portugal and Hungary), initially 5 lineages were reported during 1990s (Roumen *et al*., 1997), but a report by Lara-Alvarez *et al*. (2010) identified an additional lineage distinct in terms of the virulence spectrum. In some geographic locations a lack of clear lineage-based structure and, continuous genotypic variation has been observed with hosts such as finger millet and rice. For example, the *M. oryzae* populations have shown continuous genotypic variation on both rice and finger millet in East Africa (Takan *et al*., 2012; Onaga *et al*., 2015). Whether sexual reproduction and recombination plays a role in these situations or is it due to the cultivation history including the host genetic background and any relationship to environmental conditions requires further detailed investigations. With a key food and nutrition security crop such as finger millet, our understanding of the blast pathogen population diversity, structure and evolution remains limited.

Table 1.3 Population structure of *M. oryzae* associated with different hosts and geographic locations based on the use of molecular techniques

Country	Period of collection	Host	No of Lineages	Molecular Technique Used	No of samples used	Reference
USA	1959-1985	Rice	8	RFLP (MGR 586)	42	Levy *et a*l., 1991
Colombia	1988-1990	Rice	6	RFLP (MGR 586)	151	Levy *et al*., 1993
Korea	1981-2000	Rice	_*	RFLP (MGR 586, MAGGY)	6,315	Park *et al*., 2003
Korea	1999	Rice	2	RFLP (MGR 586, MAGGY)	254	Park *et al*., 2008
Iran /Uruguay	NK	Rice	4	AFLP	87	Taheri and Irrannejad, 2014
Georgia	1999-2000	Tall fescue, St. Augustine grass	5	AFLP	948	Tredway *et a*l., 2005
Japan	2002-2003	Rice	2	Rep-PCR (Pot2)	1173	Suzuki *et al*., 2007
Spain	1999-2003	Rice	6	RAPD	186	Lara-Alvarez *et al*., 2010
North-Western Himalayan India	NK	Rice	5	RAPD	48	Rathour *et al*., 2004
North India	2010-2011	Rice	5	RAPD	NS	Srivastava *et al*., 2014
West Africa	1996-1997	Rice	9	RFLP (MGR 586)	305	Takan *et al*., 2012
East Africa	2000-2002	Finger millet	_*	AFLP	328	Takan *et al*., 2012
East Africa	2009-2011	Rice	_*	AFLP	88	Onaga *et al*., 2015
Malaysia	2000-2014	Rice	4	RAPD/ ISSR	35	Abed-Ashtiani *et al*., 2016
Karnataka,	2012-	Rice	6	DNA (ITS)	72	Jagadeesh *et al*., 2018a

India	2014			sequencing		
Karnataka, India	2012-2014	Rice	5	Rep-PCR (Pot2)	72	Jagadeesh *et al*., 2018b

NK- Not known *- Continuous genetic variation or lack of distinctive lineages **NS- Not specified

RFLP- Restriction fragment length length polymorphisms; AFLP -Amplified fragment length polymorphisms; RAPD – Random amplified polymorphisms DNA; rep-PCR - Repetitive DNA-based polymerase chain reaction; ISSR- Inter simple sequence repeat.

1.6.5 Genome sequencing of the blast pathogen *M. oryzae*

Fredrick Sanger's DNA sequencing by chain termination with dideoxynucleotides in 1970s (Sanger *et al*., 1977a; Sanger *et al*., 1977b) laid the basis for capillary sequencer-based automated high-throughput sequencing in early 1990s (Swerdlow *et al*., 1990; Hunkapiller *et al*., 1991). This led to the first genome sequencing of eukaryotic fungal species *Saccharomyces cerevisiae* (Goffeau *et al*., 1996) laying the foundation for the genome sequencing of diverse fungal systems including *M. oryzae* (Wood *et al*., 2002; Galagan *et al*., 2003; Martinez *et al*., 2004; Dean *et al*., 2005; Nierman *et al*., 2005, Machida *et al*., 2005; Goodwin *et al*., 2011). *M. oryzae* was the first plant fungal pathogen genome to be sequenced enabling its emergence as a key model system to investigate host-pathogen interactions (Dean *et al*., 2005). The *M. oryzae* isolate 70-15 used was a laboratory strain developed from crosses of a rice blast isolate Guy11 and an isolate from weeping-love grass (Chao and Ellingboe *et al*., 1991; Dean *et al*., 2005). The genome size of *M. oryzae* isolate 70-15 is ~ 38 Mb with 7 chromosomes and 11,109 protein coding genes (Dean *et al*., 2005). This landmark breakthrough provided initial insights into the adaptability of the blast pathogen to invade and colonise the host and processes related to the breakdown of host resistance. Some of the conceptual ideas to emerge from this work include identification of diverse

secreted proteins (subsequently designated effectors) involved in extracellular perception and signal transduction and genomic regions that are capable of generating genetic variation even in the absence of sexual reproduction, and specifically adapted regulatory pathways controlling infection related development (Dean *et al*., 2005).

Rapid developments in the next generation sequencing (NGS) technologies over the last 10 years has revolutionised the entire field of fungal genome biology. NGS technologies such as Roche 454, Illumina (solexa), Ion torrent, SOLiD, PacBio and Oxford Nanopore have enabled hitherto unimaginable levels of throughput and cost effectiveness (Schuster, 2007; Shendure and Ji, 2008; Van Dijk *et al*., 2014a; Han *et al*., 2014). High quality of the starting DNA and the DNA fragment libraries is critical to genome sequencing by NGS technologies along with proper curation of the sequence reads and their assembly using programs such as SPAdes and Velvet (Shendure and Ji, 2008; Linnarsson, 2010; Van Dijk *et al*., 2014b; McCormack *et al*., 2013, Head *et al*., 2014). The quality of the genome assembly is reflected by the number of contigs, length of the contigs, size of the assembled genome and the N50. N50 is the point at which 50% of the genome is contained in contigs of size N or greater and the N50 value is expected to be greater than the median gene size to perform the gene prediction (Yandell and Ence, 2012). Comparative genomic analysis based on the predicted gene set, genome synteny, single nucleotides polymorphism (SNPs), and the core and flexible parts of the genome enables a whole range of fundamental research questions to be addressed.

Presently, more than 200 genomes of *M. oryzae* from diverse hosts and geographic locations have been deposited in various international databases such as the National Centre for Biotechnology Information (NCBI), Joint Genome Institute (JGI) and Broad

Institute (BI). Nearly 70 % of these genomes are from rice blast isolates. Various researchers are actively pursuing genome level comparative and functional analysis to decipher the infection biology of *M. oryzae* including the identification of novel genes, effectors and AVR genes as well as their function (Xue *et al*., 2012; Chen *et al*., 2013a; Dong *et al*., 2015; Gowda *et al*., 2015, Chiapello *et al*., 2015). More recent efforts are beginning to focus on the phylogenomic analysis of the population biology of *M. oryzae*. Investigation of genome sequences of 100 rice blast pathogen isolates across the globe, identified 3 major genetically distinct clades that are distributed globally (Zhong *et al*., (2018). Further, this study confirmed Asia as the centre of origin of *M. oryzae* and that the populations in Asia generate diversity through recombination as has previously been suggested (Saleh *et al*., 2012a; 2014). Another study based on the genome sequences of 76 *M. oryzae* isolates from different hosts and geographic locations identified 6 lineages mainly associated with specific hosts (Gladieux *et al*., 2018). Two of these lineages were restricted to China, whilst the remaining 4 spread both within China and to other geographic locations such as India, Burundi and USA (Gladieux *et al*., 2018). Based on a small number of isolates, this study has suggested the association of two lineages of *M. oryzae* with *Eleusine* species. This observation is consistent with an early report by Tanaka *et al*. (2009) based on the ITS marker. These studies have suggested that the two lineages have evolved independently. This is a fundamentally important question that needs to be further addressed in the context of Eastern Africa and India as the primary and secondary centres of diversity of finger millet.

1.7 Rationale, Aim and Objectives

It is well recognised that crop fungal pathogens with high level of genetic variation are more likely to adapt rapidly to diverse biotic and abiotic environments and overcome disease control measures including resistant varieties than pathogens with limited genetic variation (e.g. McDonald and McDermott, 1993). Cereal blast pathogen *M. oryzae* is well known for its ability to infect more than 50 hosts in diverse geographic locations and ranks as one of the most important crop fungal diseases (Dean *et al*., 2005). A significant body of research carried out with the rice blast system has clearly established the varying levels of the genetic and pathogenic diversity displayed by the pathogen populations particularly in relation to the cropping history in different parts of the world (e.g. USA, Europe, India and Thailand) (Mekwatanakarn *et al*., 2000; Correll *et al*., 2000) and further research is continuing in other parts of the world on rice as well as wheat (e.g. Kenya rice (Mwongera, 2018); Bangladesh and Brazil wheat (Malaker *et al*., 2016; Cruz and Valent, 2017). With other economically important crops such as finger millet, our understanding of the blast pathogen biology, genetics and evolution still remains limited. Previous work has provided an initial assessment of the finger millet blast pathogen populations in Uganda and Kenya (Takan *et al*., 2012). International efforts with the rice blast system have shown that a clear understanding of the blast pathogen diversity and continued monitoring are critical for effective management of the blast disease.

In this context, the main aim of this research is to gain an in-depth understanding of the finger millet blast pathogen genetic diversity, population structure and evolutionary relationships specifically in Eastern Africa, which is the centre of origin and domestication of the crop. This will be achieved through the following objectives:

1. To identify and/or develop DNA sequence-based molecular markers to characterise the pathogen genetic diversity in space and time by using MO isolates representing historical and contemporary populations. Use the single- and multi-locus data to identify pathogen genotypes, their distribution pattern and phylogenetic relationships.
2. To investigate the pattern of occurrence of the *Grasshopper* repeat element, mainly associated with the Asian pathogen isolates, among the finger millet blast pathogen populations in Eastern Africa to enable an insight into the transcontinental movement of the host and the pathogen.
3. To develop a reference genome sequence(s) for the finger millet blast pathogen and carry out genome resequencing of selected genotypes using next generation sequencing technology. These resources will be used to perform comparative analysis to identify genome-level differences including gene content and single nucleotide polymorphisms (SNPs) associated with the pathogen diversity as well as to better understand pathogen evolution.
4. To determine the sexual reproductive capability of the contemporary populations by initially identifying the mating type of each isolate using a diagnostic PCR analysis. Based on the known mating type data of individual isolates perform crosses with known tester isolates to assess the fertility status. Utilise the mating type distribution pattern and the fertility status data to assess the potential for sexual reproduction and recombination.

Thus, the overall purpose of this research is to develop new knowledge and resources of the pathogen critically essential for the identification, development and deployment of resistance to finger millet blast in Eastern Africa.

Chapter 2

2.0 Materials and Methods

2.1 Brief description of the origin of blast pathogen isolates

This research work utilised 300 isolates of *M. oryzae* (Table 2.1A) from finger millet and associated weed hosts. The samples were collected from farmers' fields and various screening sites in major finger millet-producing regions of four Eastern Africa countries: Kenya, Uganda, Tanzania and Ethiopia (Figure 2.1 A and B). Seventy-six (76) of these isolates were selected from 328 previously characterised blast pathogen isolates collected from 7 and 15 districts in Kenya and Uganda, respectively during 2000 to 2004 (Takan, 2007; Takan *et al*., 2012) (Figure 2.1B and Table 2.1B). These 76 isolates represent the 'historical populations' in this study. Two hundred and twenty-four (224) isolates collected between 2015 to 2017 from various districts in Kenya (8), Uganda (21), Tanzania (6) and Ethiopia (24) (Figure 2.1B and Table 2.1B). These 224 isolates represent the 'contemporary populations' in this study. For the contemporary populations, infected sample collection was carried out by research partners based in these countries and isolations were carried out at the National Semi Arid Resources Research Institute (NaSARRI) laboratory, Uganda and the University of Georgia, USA as part of a PEARL (Program for Emerging Agricultural Research Leaders) project funded by Bill and Melinda Gates Foundation (BMGF), in which University of Bedfordshire is a partner with Prof Prasad S Sreenivasaprasad as Co-PI. The 300 isolates

were obtained from infected leaf, neck and panicle (head) samples of finger millet and/or weed hosts (Table 2.1A).

Table 2.1A Details of *M. oryzae* isolates used in this study

Serial No.	Isolate	Country of origin	District	Host	Plant part	Finger millet name	Improved/Landrace	Year of collection*
1	K23/123	Kenya	Busia	*Eleusine coracana*	Neck	-	-	2000
2	K5/24w	Kenya	Busia	*Eleusine coracana*	-	-	-	2002
3	K22/118	Kenya	Busia	*Eleusine coracana*	Panicle	-	-	2000
4	K4/21p	Kenya	Busia	*Eleusine coracana*	Panicle	-	-	2002
5	K8/40	Kenya	Busia	*Eleusine indica*	Panicle	-	-	2000
6	K10	Kenya	Busia	*Eleusine coracana*	Neck	P224	Improved	2016
7	K11	Kenya	Busia	*Eleusine coracana*	Neck	KENFM7	Improved	2016
8	K12	Kenya	Busia	*Eleusine coracana*	Neck	KAT FM1	Improved	2016
9	K13	Kenya	Busia	*Eleusine africana*	Neck	Wild E. Africana	Weed	2016
10	K23	Kenya	Busia	*Eleusine coracana*	Neck	-	-	2017
11	K24	Kenya	Busia	*Eleusine coracana*	Neck	-	-	2017
12	K25	Kenya	Busia	*Eleusine coracana*	Neck	-	-	2017
13	K26	Kenya	Busia	*Eleusine coracana*	Neck	-	-	2017
14	K27	Kenya	Busia	*Eleusine coracana*	Neck	-	-	2017
15	K28	Kenya	Busia	*Eleusine coracana*	Neck	-	-	2017
16	K29	Kenya	Busia	*Eleusine coracana*	Neck	-	-	2017
17	K30	Kenya	Busia	*Eleusine coracana*	Neck	-	-	2017
18	K31	Kenya	Busia	*Eleusine coracana*	Neck	-	-	2017
19	K1	Kenya	Siaya	*Eleusine coracana*	Neck	-	Improved	2016
20	K3	Kenya	Siaya	*Eleusine coracana*	Neck	-	Improved	2016
21	K6	Kenya	Siaya	*Eleusine coracana*	Neck	P224	Improved	2016
22	K7	Kenya	Siaya	*Eleusine coracana*	Neck	P224	Improved	2016
23	K8	Kenya	Siaya	*Eleusine coracana*	Neck	MS60D	Improved	2016
24	K2	Kenya	Kisumu	*Eleusine coracana*	Neck	-	Improved	2016
25	K4	Kenya	Kisumu	*Eleusine coracana*	Head	MS60D	Improved	2016
26	K5	Kenya	Kisumu	*Eleusine coracana*	Neck	-	Improved	2016
27	K9	Kenya	Kisumu	*Eleusine coracana*	Neck	MS60D	Improved	2016
28	K15	Kenya	Kisumu	*Eleusine coracana*	Neck	-	Improved	2016
29	K22	Kenya	Kisumu	*Eleusine coracana*	Neck	-	Improved	2016
30	K57/126p	Kenya	Gucha	*Eleusine coracana*	Panicle	-	-	2002

Serial No.	Isolate	Country of origin	District	Host	Plant part	Finger millet name	Improved/Landrace	Year of collection*
31	K58/128p	Kenya	Gucha	*Eleusine coracana*	Panicle	-	-	2002
32	K60/131p	Kenya	Gucha	*Eleusine coracana*	Panicle	-	-	2002
33	K55/124p	Kenya	Gucha	*Eleusine coracana*	Panicle	-	-	2002
34	K14	Kenya	Bungoma	*Eleusine coracana*	Neck	-	Improved	2016
35	K32	Kenya	Bungoma	*Eleusine coracana*	Neck	-	-	2017
36	K17	Kenya	Kakamega	*Eleusine africana*	Neck	Wild *E. africana*	Weed	2016
37	K29/164	Kenya	Suba	*Eleusine coracana*	Panicle	-	-	2000
38	K18	Kenya	Eldoret	*Eleusine coracana*	Neck	-	Landrace	2016
39	K19	Kenya	Eldoret	*Eleusine coracana*	Neck	-	Landrace	2016
40	K20	Kenya	Eldoret	*Eleusine coracana*	Neck	U15	Improved	2016
41	K21	Kenya	Marakwet	*Eleusine coracana*	Head	-	Landrace	2016
42	K64/137p	Kenya	Homabay	*Eleusine coracana*	Panicle	-	-	2002
43	K33/189	Kenya	Kericho	*Eleusine coracana*	Leaf	-	-	2000
44	K33/184	Kenya	Kericho	*Eleusine coracana*	Panicle	-	-	2000
45	K47/114p	Kenya	Kisii	*Eleusine coracana*	Panicle	-	-	2002
46	K45/112n	Kenya	Kisii	*Eleusine coracana*	Neck	-	-	2002
47	K44/111p	Kenya	Kisii	*Eleusine coracana*	Panicle	-	-	2002
48	K48/115n	Kenya	Kisii	*Eleusine coracana*	Neck	-	-	2002
49	K16	Kenya	Kisii	*Eleusine coracana*	Neck	-	Improved	2016
50	K33	Kenya	Kisii	*Eleusine coracana*	Neck	-	-	2017
51	K34	Kenya	Kisii	*Eleusine coracana*	Neck	-	-	2017
52	K35	Kenya	Kisii	*Eleusine coracana*	Neck	-	-	2017
53	K36	Kenya	Kisii	*Eleusine coracana*	Neck	-	-	2017
54	K37	Kenya	Kisii	*Eleusine coracana*	Neck	-	-	2017
55	K38	Kenya	Kisii	*Eleusine coracana*	Neck	-	-	2017
56	K39	Kenya	Kisii	*Eleusine coracana*	Neck	-	-	2017
57	K40	Kenya	Kisii	*Eleusine coracana*	Neck	-	-	2017
58	K41	Kenya	Kisii	*Eleusine coracana*	Head	-	-	2017
59	K42	Kenya	Kisii	*Eleusine coracana*	Neck	-	-	2017
60	K43	Kenya	Kisii	*Eleusine coracana*	Head	-	-	2017
61	K44	Kenya	Kisii	*Eleusine coracana*	Neck	-	-	2017
62	K45	Kenya	Kisii	*Eleusine coracana*	Neck	-	-	2017

Serial No.	Isolate	Country of origin	District	Host	Plant part	Finger millet name	Improved/Landrace	Year of collection*
63	K24/127	Kenya	Kisii central	*Eleusine coracana*	Panicle	-	-	2000
64	K28/82w	Kenya	Teso	*Eleusine coracana*	-	-	-	2002
65	K1/15	Kenya	Teso	*Eleusine coracana*	Panicle	-	-	2000
66	K9/46	Kenya	Teso	*Eleusine indica*	Panicle	-	-	2000
67	K14/74	Kenya	Teso	*Eleusine indica*	Neck	-	-	2000
68	K65/142n	Kenya	Teso	*Eleusine coracana*	Neck	-	-	2002
69	K26/76p	Kenya	Teso	*Eleusine coracana*	Panicle	-	-	2002
70	K36/98n	Kenya	Teso	*Eleusine coracana*	Neck	-	-	2002
71	K21/68n	Kenya	Teso	*Eleusine coracana*	Neck	-	-	2002
72	K15/53n	Kenya	Teso	*Eleusine coracana*	Neck	-	-	2002
73	K12/62	Kenya	Teso	*Eleusine coracana*	Neck	-	-	2000
74	K13/67	Kenya	Teso	*Eleusine coracana*	Neck	-	-	2000
75	K5/23	Kenya	Teso	*Eleusine coracana*	Neck	-	-	2000
76	K65/159w	Kenya	Alupe/Teso	*Eleusine coracana*	-	-	-	2002
77	K65/140n	Kenya	Alupe/Teso	*Eleusine coracana*	Neck	-	-	2002
78	U1	Uganda	Arua	*Eleusine coracana*	Neck	Akokonyara	Local	2015
79	U2	Uganda	Arua	*Eleusine coracana*	Neck	Anya	Local	2015
80	U3	Uganda	Arua	*Eleusine coracana*	Neck	-	-	2015
81	U42	Uganda	Arua	*Eleusine coracana*	Neck	Anya	Local	2016
82	U24	Uganda	Apac	*Eleusine coracana*	Neck	-	Local	2016
83	U25	Uganda	Apac	*Eleusine coracana*	Neck	-	-	2016
84	U26	Uganda	Apac	*Eleusine coracana*	Neck	Toodyang	Local	2016
85	D10/s71	Uganda	Apac	*Eleusine indica*	Leaf	-	-	2002
86	D10/s73	Uganda	Apac	*Digitaria scalarum*	Leaf	-	-	2002
87	D10/s63	Uganda	Apac	*Eleusine coracana*	Panicle	-	-	2002
88	D10/s77	Uganda	Apac	*Eleusine coracana*	Panicle	-	-	2002
89	U15	Uganda	Amuria	*Eleusine coracana*	Head	Ekama	Local	2016
90	U50	Uganda	Amuria	*Eleusine coracana*	Neck	Ekama	Local	2016
91	U19	Uganda	Alebtong	*Eleusine coracana*	Neck	Otopotop	Local	2016
92	U20	Uganda	Alebtong	*Eleusine coracana*	Neck	-	-	2016
93	U21	Uganda	Alebtong	*Eleusine coracana*	Neck	-	-	2016
94	U53	Uganda	Alebtong	*Eleusine coracana*	Neck	-	-	2016

Serial No.	Isolate	Country of origin	District	Host	Plant part	Finger millet name	Improved/Landrace	Year of collection*
95	U29	Uganda	Amuru	*Eleusine coracana*	Neck	Ajuku	Local	2016
96	U30	Uganda	Amuru	*Eleusine coracana*	Neck	Odyera Layelo	Local	2016
97	U54	Uganda	Amuru	*Eleusine coracana*	Neck	Ajukmanyige	Local	2016
98	U34	Uganda	Agago	*Eleusine coracana*	Neck	Awiliwili	Local	2016
99	U35	Uganda	Agago	*Eleusine coracana*	Neck	Adyang ataar	Local	2016
100	D6/s1	Uganda	Bugiri	*Eleusine coracana*	Panicle	-	-	2000
101	D8/s15	Uganda	Busia	*Eleusine coracana*	Panicle	-	-	2000
102	D9/s56	Uganda	Lira	*Eleusine coracana*	Neck	-	-	2002
103	D9/s76	Uganda	Lira	*Eleusine coracana*	Panicle	-	-	2002
104	D9/s70	Uganda	Lira	*Eleusine coracana*	Neck	-	-	2002
105	D9/s54	Uganda	Lira	*Digitaria aegytium*	Panicle	-	-	2002
106	D9/s50	Uganda	Lira	*Eleusine indica*	Panicle	-	-	2002
107	U22	Uganda	Lira	*Eleusine coracana*	Neck	-	-	2016
108	U23	Uganda	Lira	*Eleusine coracana*	Neck	-	-	2016
109	U28	Uganda	Lira	*Eleusine coracana*	Neck	Okutuwiye	Local	2016
110	D13/s5	Uganda	Katakwi	*Eleusine coracana*	Panicle	-	-	2000
111	U13	Uganda	Katakwi	*Eleusine coracana*	Neck	Etiyo	Local	2016
112	U14	Uganda	Katakwi	*Eleusine coracana*	Neck	Etiyo	Local	2016
113	D3/s24	Uganda	Mbale	*Eleusine coracana*	Panicle	-	-	2000
114	D3/s3	Uganda	Mbale	*Eleusine coracana*	Panicle	-	-	2000
115	D3/s9	Uganda	Mbale	*Eleusine coracana*	Panicle	-	-	2000
116	U58	Uganda	Mbale	*Eleusine coracana*	Neck	-	-	2016
117	D7/s6	Uganda	Kamuli	*Eleusine coracana*	Panicle	-	-	2000
118	D5/s1	Uganda	Iganga	*Eleusine coracana*	Panicle	-	-	2000
119	D12/s2	Uganda	Nakasongola	*Eleusine coracana*	Panicle	-	-	2000
120	D11/s16	Uganda	Masindi	*Eleusine coracana*	Panicle	-	-	2000
121	U43	Uganda	Masindi	*Eleusine coracana*	Neck	Kal	Local	2016
122	D2/s26	Uganda	Kumi	*Eleusine coracana*	Leaf	-	-	2000
123	D2/s14	Uganda	Kumi	*Eleusine coracana*	Panicle	-	-	2000
124	U8	Uganda	Kumi	*Eleusine coracana*	Neck	Okurut musisi	Local	2016
125	U9	Uganda	Kumi	*Eleusine coracana*	Neck	-	Local	2016
126	U10	Uganda	Kumi	*Eleusine coracana*	Neck	Omodingot	Local	2016

Serial No.	Isolate	Country of origin	District	Host	Plant part	Finger millet name	Improved/Landrace	Year of collection*
127	U16	Uganda	Kumi	*Eleusine coracana*	Neck	-	-	2016
128	U48	Uganda	Kumi	*Eleusine coracana*	Neck	Obeet	Local	2016
129	U49	Uganda	Kumi	*Eleusine coracana*	Head	-	-	2016
130	U17	Uganda	Ngora	*Eleusine coracana*	Neck	-	Local	2016
131	U18	Uganda	Ngora	*Eleusine coracana*	Neck	-	Local	2016
132	U51	Uganda	Ngora	*Eleusine coracana*	Neck	Obeet	Local	2016
133	U52	Uganda	Ngora	*Eleusine coracana*	Neck	Obeet	Local	2016
134	U27	Uganda	Gulu	*Eleusine coracana*	Neck	-	Local	2016
135	U31	Uganda	Gulu	*Eleusine coracana*	Head	Odyera	Local	2016
136	U32	Uganda	Kitgum	*Eleusine coracana*	Neck	-	Local	2016
137	U33	Uganda	Kitgum	*Eleusine coracana*	Neck	-	-	2016
138	U55	Uganda	Lamwo	*Eleusine coracana*	Neck	Agriculture	Improved	2016
139	U36	Uganda	Manafwa	*Eleusine coracana*	Neck	Namakala	Local	2016
140	U37	Uganda	Manafwa	*Eleusine coracana*	Neck	-	Local	2016
141	U38	Uganda	Manafwa	*Eleusine coracana*	Neck	-	-	2016
142	U56	Uganda	Manafwa	*Eleusine coracana*	Neck	-	Local	2016
143	U44	Uganda	Hoima	*Eleusine coracana*	Neck	Kabiriti	Local	2016
144	U46	Uganda	Hoima	*Eleusine coracana*	Neck	Bulo	Local	2016
145	U47	Uganda	Hoima	*Eleusine coracana*	Neck	Bulo	Local	2016
146	D14/s30	Uganda	Kabermaido	*Eleusine coracana*	Leaf	-	-	2002
147	D14/s27	Uganda	Kabermaido	*Eleusine coracana*	Panicle	-	-	2000
148	D1/s72	Uganda	Pallisa	*Eleusine coracana*	Leaf	-	-	2002
149	D1/s53b	Uganda	Pallisa	*Eleusine coracana*	Leaf	-	-	2002
150	D1/s11	Uganda	Pallisa	*Eleusine coracana*	Neck	-	-	2000
151	D1/s19	Uganda	Pallisa	*Eleusine coracana*	Panicle	-	-	2000
152	D1/s44	Uganda	Pallisa	*Eleusine coracana*	Panicle	-	-	2002
153	D1/s50	Uganda	Pallisa	*Eleusine coracana*	Panicle	157	-	2002
154	U11	Uganda	Pallisa	*Eleusine coracana*	Neck	Seremi 2	Improved	2016
155	U12	Uganda	Pallisa	*Eleusine coracana*	Neck	-	-	2016
156	U4	Uganda	Moyo	*Eleusine coracana*	Neck	Duhwi	Landrace	2015
157	U5	Uganda	Serere	*Eleusine coracana*	Neck	Obeet	Local	2016
158	U6	Uganda	Serere	*Eleusine coracana*	Neck	Emoru, engeny	Local	2016

Serial No.	Isolate	Country of origin	District	Host	Plant part	Finger millet name	Improved/Landrace	Year of collection*
159	U7	Uganda	Serere	*Eleusine coracana*	Neck	Seremi 2	Improved	2016
160	U45	Uganda	Serere	*Eleusine coracana*	Neck	Kabiriti	Local	2016
161	D15/s37	Uganda	Soroti	*Digitaria horizontalis*	Leaf	-	-	2002
162	D15/s47	Uganda	Soroti	*Eleusine coracana*	Neck	-	-	2002
163	D15/s6	Uganda	Soroti	*Eleusine coracana*	Panicle	-	-	2000
164	D15/s41	Uganda	Soroti	*Eleusine coracana*	Leaf	-	-	2002
165	D15/s12	Uganda	Soroti	*Eleusine coracana*	Panicle	-	-	2000
166	E11p-1-1	Uganda	Soroti	*Eleusine coracana*	Panicle	-	-	2004
167	Gup-2-1	Uganda	Soroti	*Eleusine coracana*	Panicle	-	-	2004
168	Odyp-2-1	Uganda	Soroti	*Eleusine coracana*	Panicle	-	-	2004
169	Pen-2-2	Uganda	Soroti	*Eleusine coracana*	Neck	-	-	2004
170	P665n-2-1	Uganda	Soroti	*Eleusine coracana*	Neck	-	-	2004
171	Secn-2-2	Uganda	Soroti	*Eleusine coracana*	Neck	-	-	2004
172	S1p-1-1	Uganda	Soroti	*Eleusine coracana*	Panicle	-	-	2004
173	D4/s12	Uganda	Tororo	*Eleusine coracana*	Panicle	-	-	2000
174	D4/s26	Uganda	Tororo	*Eleusine coracana*	Panicle	-	-	2000
175	D4/s41	Uganda	Tororo	*Eleusine coracana*	Neck	-	-	2002
176	U39	Uganda	Tororo	*Eleusine coracana*	Neck	-	Local	2016
177	U40	Uganda	Tororo	*Eleusine coracana*	Neck	-	Local	2016
178	U41	Uganda	Tororo	*Eleusine coracana*	Neck	-	Local	2016
179	U57	Uganda	Tororo	*Eleusine coracana*	Neck	-	Local	2016
180	T1	Tanzania	Nkasi	*Eleusine coracana*	Head	Kahurunge	Landrace	2016
181	T2	Tanzania	Nkasi	*Eleusine coracana*	Neck	Kahurunge	Landrace	2016
182	T3	Tanzania	Nkasi	*Eleusine coracana*	Neck	Kafumbata	Landrace	2016
183	T24	Tanzania	Nkasi	*Eleusine coracana*	Neck	Kahurunge	Landrace	2016
184	T26	Tanzania	Nkasi	*Eleusine coracana*	Neck	Nameka	Landrace	2016
185	T27	Tanzania	Nkasi	*Eleusine coracana*	Neck	Kahurunge	Landrace	2016
186	T28	Tanzania	Nkasi	*Eleusine coracana*	Neck	Kahurunge	Landrace	2016
187	T29	Tanzania	Nkasi	*Eleusine coracana*	Neck	Kafumbata	Landrace	2016
188	T30	Tanzania	Nkasi	*Eleusine coracana*	Neck	Kahurunge	Landrace	2016
189	T33	Tanzania	Nkasi	*Eleusine coracana*	Neck	Kafumbata	Landrace	2016
190	T4	Tanzania	Sumbawanga	*Eleusine coracana*	Neck	Kahurunge	Landrace	2016

Serial No.	Isolate	Country of origin	District	Host	Plant part	Finger millet name	Improved/Landrace	Year of collection*
191	T12	Tanzania	Sumbawanga	*Eleusine coracana*	Neck	Sanzala	Landrace	2016
192	T31	Tanzania	Sumbawanga	*Eleusine coracana*	Neck	Kahurunge	Landrace	2016
193	T32	Tanzania	Sumbawanga	*Eleusine coracana*	Neck	Kahurunge	Landrace	2016
194	T34	Tanzania	Sumbawanga	*Eleusine coracana*	Neck	Kahurunge	Landrace	2016
195	T6	Tanzania	Njombe	*Eleusine coracana*	Neck	Mbudi	Landrace	2016
196	T7	Tanzania	Njombe	*Eleusine coracana*	Neck	Mbudi/Tanzalo	Landrace	2016
197	T8	Tanzania	Njombe	*Eleusine coracana*	Neck	Mbudi/Tanzalo	Landrace	2016
198	T9	Tanzania	Njombe	*Eleusine coracana*	Neck	Mbudi/Tanzalo	Landrace	2016
199	T16	Tanzania	Njombe	*Eleusine coracana*	Neck	Mbudi	Landrace	2016
200	T18	Tanzania	Njombe	*Eleusine coracana*	Neck	Mbudi/Tanzalo	Landrace	2016
201	T20	Tanzania	Njombe	*Eleusine coracana*	Neck	Mbudi	Landrace	2016
202	T45	Tanzania	Njombe	*Eleusine coracana*	Neck	Mbudi	Landrace	2016
203	T46	Tanzania	Njombe	*Eleusine coracana*	Neck	Mbudi/Tanzalo	Landrace	2016
204	T47	Tanzania	Njombe	*Eleusine coracana*	Neck	Mbudi/Tanzalo	Landrace	2016
205	T48	Tanzania	Njombe	*Eleusine coracana*	Neck	Mbudi/Tanzalo	Landrace	2016
206	T49	Tanzania	Njombe	*Eleusine coracana*	Neck	Mbudi/Tanzalo	Landrace	2016
207	T50	Tanzania	Njombe	*Eleusine coracana*	Neck	Mbudi/Tanzalo	Landrace	2016
208	T51	Tanzania	Njombe	*Eleusine coracana*	Neck	Mbudi/Tanzalo	Landrace	2016
209	T52	Tanzania	Njombe	*Eleusine coracana*	Head	Nanga/Tanzalo	Landrace	2016
210	T53	Tanzania	Njombe	*Eleusine coracana*	Neck	Mbudi	Landrace	2016
211	T54	Tanzania	Njombe	*Eleusine coracana*	Head	Wibunda	Landrace	2016
212	T55	Tanzania	Njombe	*Eleusine coracana*	Head	Wibunda	Landrace	2016
213	T56	Tanzania	Njombe	*Eleusine coracana*	Neck	Wibunda	Landrace	2016
214	T57	Tanzania	Njombe	*Eleusine coracana*	Head	Wibunda	Landrace	2016
215	T58	Tanzania	Njombe	*Eleusine coracana*	Head	Wibunda	Landrace	2016
216	T10	Tanzania	Madaba	*Eleusine coracana*	Neck	Lihegula	Landrace	2016
217	T11	Tanzania	Madaba	*Eleusine coracana*	Neck	Lihegula	Landrace	2016
218	T14	Tanzania	Madaba	*Eleusine coracana*	Head	Lihegula/Mbude	Landrace	2016
219	T15	Tanzania	Madaba	*Eleusine coracana*	Head	Lihegula/Mbudi	Landrace	2016
220	T17	Tanzania	Madaba	*Eleusine coracana*	Neck	Lihegula/Mbude	Landrace	2016
221	T35	Tanzania	Madaba	*Eleusine coracana*	Neck	Lihegula/Mbude	Landrace	2016
222	T13	Tanzania	Momba	*Eleusine coracana*	Neck	Naupule	Landrace	2016

Serial No.	Isolate	Country of origin	District	Host	Plant part	Finger millet name	Improved/Landrace	Year of collection*
223	T22	Tanzania	Momba	*Eleusine coracana*	Neck	Naupule/Sulila	Landrace	2016
224	T23	Tanzania	Momba	*Eleusine coracana*	Neck	Naupule	Landrace	2016
225	T36	Tanzania	Momba	*Eleusine coracana*	Neck	Naupule	Landrace	2016
226	T37	Tanzania	Momba	*Eleusine coracana*	Neck	Kahurunge	Landrace	2016
227	T38	Tanzania	Momba	*Eleusine coracana*	Neck	Naupule	Landrace	2016
228	T5	Tanzania	Mbozi	*Eleusine coracana*	Neck	Maburunge	Landrace	2016
229	T19	Tanzania	Mbozi	*Eleusine coracana*	Neck	Mburunje	Landrace	2016
230	T21	Tanzania	Mbozi	*Eleusine coracana*	Neck	Inyiru(mweusi)	Landrace	2016
231	T25	Tanzania	Mbozi	*Eleusine coracana*	Neck	Yumbayumba	Landrace	2016
232	T39	Tanzania	Mbozi	*Eleusine coracana*	Neck	Kaburunge	Landrace	2016
233	T40	Tanzania	Mbozi	*Eleusine coracana*	Neck	Kaburunge	Landrace	2016
234	T41	Tanzania	Mbozi	*Eleusine coracana*	Neck	Uzelu	Landrace	2016
235	T42	Tanzania	Mbozi	*Eleusine coracana*	Neck	Maburunge	Landrace	2016
236	T43	Tanzania	Mbozi	*Eleusine coracana*	Neck	Makopera/Mburunje	Landrace	2016
237	T44	Tanzania	Mbozi	*Eleusine coracana*	Neck	Uzelu	Landrace	2016
238	E1	Ethiopia	Sire	*Eleusine coracana*	Neck	-	Landrace	2015
239	E2	Ethiopia	Wayu Tuka	*Eleusine coracana*	Neck	-	Landrace	2015
240	E18	Ethiopia	Wayu Tuka	*Eleusine coracana*	Neck	-	-	2016
241	E19	Ethiopia	Wayu Tuka	*Eleusine coracana*	Neck	-	-	2016
242	E21	Ethiopia	Wayu Tuka	*Eleusine coracana*	Neck	-	-	2016
243	E3	Ethiopia	Adet	*Eleusine coracana*	Neck	-	Landrace	2015
244	E4	Ethiopia	Diga	*Eleusine coracana*	Head	-	Landrace	2015
245	E5	Ethiopia	Diga	*Eleusine coracana*	Neck	-	Landrace	2015
246	E6	Ethiopia	Diga	*Eleusine coracana*	Head	-	Landrace	2015
247	E7	Ethiopia	Diga	*Eleusine coracana*	Neck	-	Landrace	2015
248	E20	Ethiopia	Diga	*Eleusine coracana*	Neck	-	-	2016
249	E29	Ethiopia	Diga	*Eleusine coracana*	Neck	-	-	2016
250	E32	Ethiopia	Diga	*Eleusine coracana*	Neck	-	-	2016
251	E33	Ethiopia	Diga	*Eleusine coracana*	Head	-	-	2016
252	E34	Ethiopia	Diga	*Eleusine coracana*	Neck	-	-	2016
253	E37	Ethiopia	Diga	*Eleusine coracana*	Neck	-	-	2016
254	E8	Ethiopia	Lalo Assabi	*Eleusine coracana*	Head	-	Landrace	2015

Serial No.	Isolate	Country of origin	District	Host	Plant part	Finger millet name	Improved/Landrace	Year of collection*
255	E9	Ethiopia	Lalo Assabi	*Eleusine coracana*	Neck	-	Landrace	2015
256	E30	Ethiopia	Lallo Asabi	*Eleusine coracana*	Neck	-	-	2016
257	E10	Ethiopia	Nedjo	*Eleusine coracana*	Head	-	Landrace	2015
258	E22	Ethiopia	Nedjo	*Eleusine coracana*	Head	-	-	2016
259	E23	Ethiopia	Nedjo	*Eleusine coracana*	Neck	-	-	2016
260	E24	Ethiopia	Nedjo	*Eleusine coracana*	Neck	-	-	2016
261	E25	Ethiopia	Nedjo	*Eleusine coracana*	Neck	-	-	2016
262	E12	Ethiopia	Banja	*Eleusine coracana*	Head	-	Landrace	2015
263	E38	Ethiopia	Banja	*Eleusine coracana*	Head	-	Landrace	2015
264	E39	Ethiopia	Banja	*Eleusine coracana*	Neck	-	Landrace	2015
265	E49	Ethiopia	Banja	*Eleusine coracana*	Head	-	Landrace	2015
266	E50	Ethiopia	Banja	*Eleusine coracana*	Neck	-	Landrace	2015
267	E27	Ethiopia	Bila	*Eleusine coracana*	Neck	-	-	2016
268	E13	Ethiopia	Bahir Dar Zuria	*Eleusine coracana*	Head	-	Landrace	2015
269	E14	Ethiopia	Bahir Dar Zuria	*Eleusine coracana*	Neck	-	Landrace	2015
270	E55	Ethiopia	Bahir Dar Zuria	*Eleusine coracana*	Neck	-	Landrace	2015
271	E56	Ethiopia	Bahir Dar Zuria	*Eleusine coracana*	Head	-	Landrace	2015
272	E57	Ethiopia	Bahir Dar Zuria	*Eleusine coracana*	Neck	-	Landrace	2015
273	E58	Ethiopia	Bahir Dar Zuria	*Eleusine coracana*	Neck	-	Landrace	2015
274	E15	Ethiopia	Bure	*Eleusine coracana*	Head	-	Landrace	2015
275	E59	Ethiopia	Bure	*Eleusine coracana*	Head	-	Landrace	2015
276	E16	Ethiopia	Demecha	*Eleusine coracana*	Neck	-	Landrace	2015
277	E17	Ethiopia	Angebo	*Eleusine coracana*	Neck	-	Landrace	2016
278	E26	Ethiopia	Leta Sibu	*Eleusine coracana*	Head	-	-	2016
279	E28	Ethiopia	Boji Bermeji	*Eleusine coracana*	Neck	-	-	2016
280	E31	Ethiopia	Gimbi	*Eleusine coracana*	Head	-	-	2016
281	E35	Ethiopia	Leka dulecha	*Eleusine coracana*	Neck	-	-	2016
282	E36	Ethiopia	Leka dulecha	*Eleusine coracana*	Head	-	-	2016
283	E45	Ethiopia	Leka dulecha	*Eleusine coracana*	Head	-	-	2015
284	E40	Ethiopia	Dangla	*Eleusine coracana*	Neck	-	Landrace	2015
285	E41	Ethiopia	Dure Bete	*Eleusine coracana*	Head	-	Landrace	2015
286	E42	Ethiopia	Dure Bete	*Eleusine coracana*	Neck	-	Landrace	2015

Serial No.	Isolate	Country of origin	District	Host	Plant part	Finger millet name	Improved/Landrace	Year of collection*
287	E43	Ethiopia	Dure Bete	*Eleusine coracana*	Head	-	Landrace	2015
288	E44	Ethiopia	Dure Bete	*Eleusine coracana*	Neck	-	Landrace	2015
289	E46	Ethiopia	Mandura	*Eleusine coracana*	Neck	-	Landrace	2015
290	E11	Ethiopia	Guangua	*Eleusine coracana*	Neck	-	Landrace	2015
291	E47	Ethiopia	Guangua	*Eleusine coracana*	Head	-	Landrace	2015
292	E48	Ethiopia	Guangua	*Eleusine coracana*	Neck	-	Landrace	2015
293	E51	Ethiopia	Mecha	*Eleusine coracana*	Head	-	Landrace	2015
294	E52	Ethiopia	Mecha	*Eleusine coracana*	Neck	-	Landrace	2015
295	E53	Ethiopia	Mecha	*Eleusine coracana*	Head	-	Landrace	2015
296	E54	Ethiopia	Mecha	*Eleusine coracana*	Neck	-	Landrace	2015
297	E60	Ethiopia	Ankussa-Abdo	*Eleusine coracana*	Head	-	Landrace	2015
298	E61	Ethiopia	Jabi Tana	*Eleusine coracana*	Head	-	Landrace	2015
299	E62	Ethiopia	Jabi Tana	*Eleusine coracana*	Neck	-	Landrace	2015
300	E63	Ethiopia	Qilxxu Kara	*Eleusine coracana*	Head	-	Landrace	2015
R1	G22	Japan	-	*Eleusine coracana*	-	-	-	-
R2	Guy11	French Guyana	-	*Oryza sativa*	-	-	-	-
R3	TH3	Thailand	-	*Oryza sativa*	-	-	-	-
R4	JP15	Japan	-	*Oryza sativa*	-	-	-	-
R5	BR62	Brazil	-	*Eleusine indica/africana*	-	-	-	-
R6	I-R-22	-	-	Buff mutant – laboratory	-	-	-	-
R7	4136-4-3	-	-	Laboratory strain, weeping love grass	-	-	-	-

*Isolates collected from 2000 to 2004 represent historical populations and those from 2015 to 2017 represent contemporary populations. R1 to R7 are previously characterised isolates that were utilised for various experiments in this study. - Information not available.

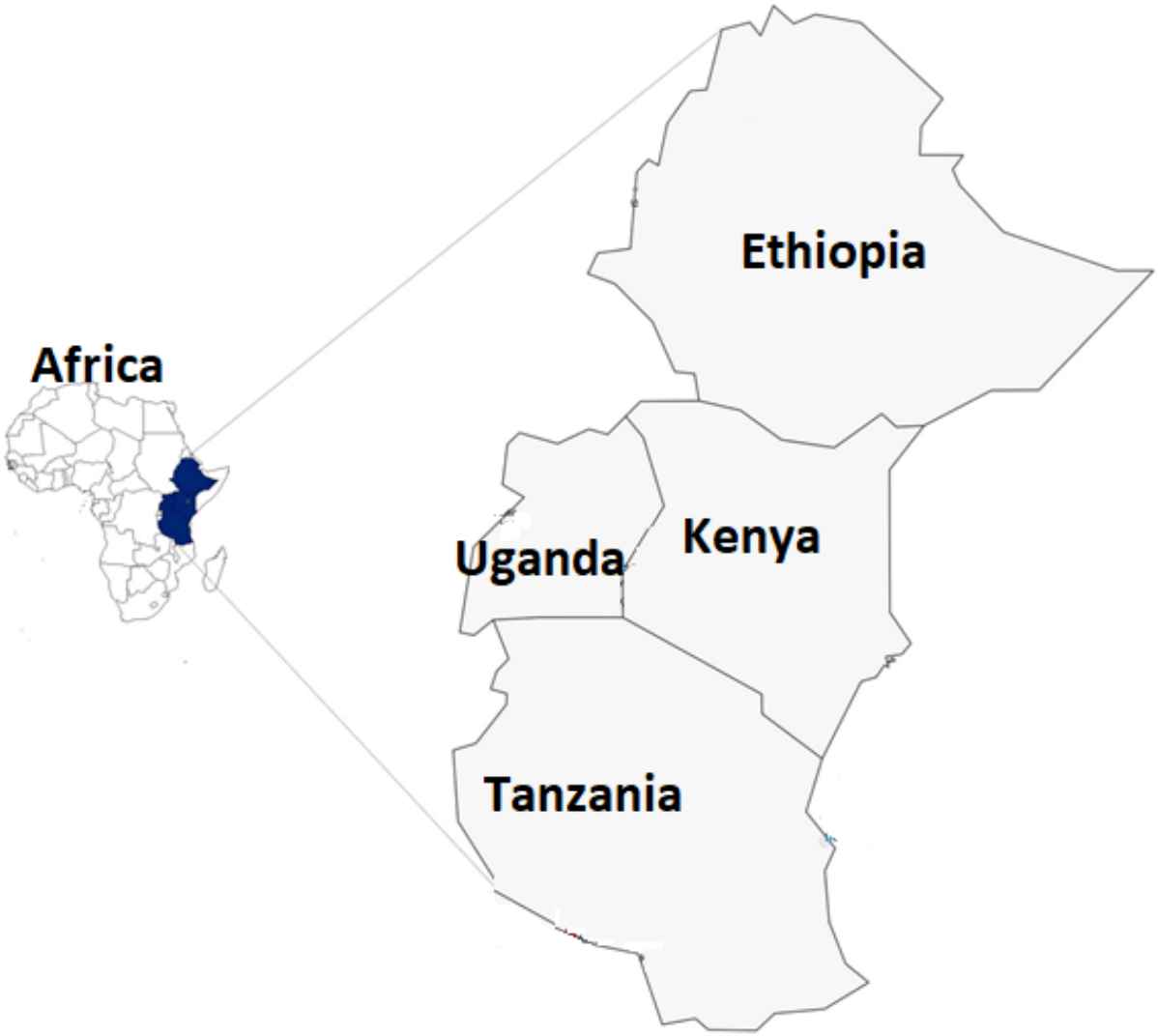

Figure 2.1A. Map of Africa showing the four countries in Eastern Africa where the samples surveyed Ethiopia, Kenya, Uganda and Tanzania

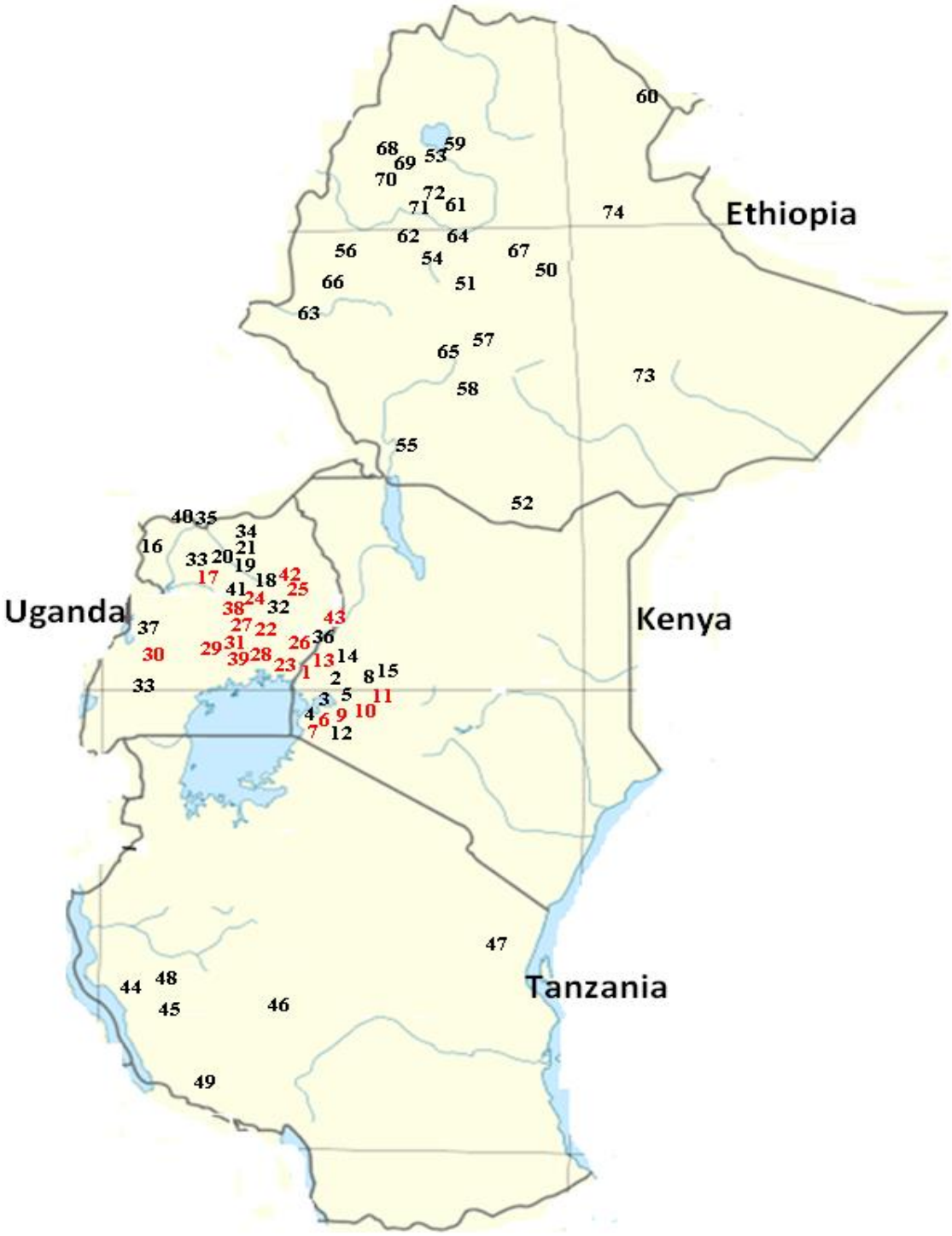

Figure 2.1B. Districts where the samples were surveyed in the four countries in Eastern Africa – Ethiopia, Kenya, Uganda and Tanzania

Numbers in Red are the districts where both historical and contemporary samples were collected and black areas where only the contemporary samples were collected. Detail of districts are available in Table 2.1B

Table 2. 1B Detail of the districts and number of *M. oryzae* isolates collected in the countries surveyed in Eastern Africa

Country	Serial No. on Map	District	No. of sample collected	Country	Serial No. on Map	District	No. of sample collected
Kenya	1	Busia	18	Uganda	27	Kamuli	1
	2	Siaya	5		28	Iganga	1
	3	Bungoma	2		29	Nakasongola	1
	4	Kisumu	6		30	Masindi	2
	5	Kakamega	1		31	Kumi	8
	6	Gucha	4		32	Ngora	4
	7	Suba	1		33	Gulu	2
	8	Eldoret	3		34	Kitgum	2
	9	Homabay	1		35	Lamwo	1
	10	Kisii	18		36	Manafwa	4
	11	Kericho	2		37	Hoima	3
	12	Kisii central	1		38	Kabermaido	2
	13	Teso	12		39	Pallisa	8
	14	Alupe Teso	2		40	Moyo	1
	15	Marakwet	1		41	Serere	4
Uganda	16	Arua	4		42	Soroti	12
	17	Apac	7		43	Tororo	7
	18	Amuria	2	Tanzania	44	Nkasi	10
	19	Alebtong	4		45	Sumbawanga	5
	20	Amuru	3		46	Njombe	21
	21	Agago	2		47	Madaba	6
	22	Bugiri	1		48	Momba	6
	23	Busia	1		49	Mbozi	10
	24	Lira	8	Ethiopia	50	Sire	1
	25	Katakwi	3		51	Wayu Tuka	4
	26	Mbale	4		52	Adet	1

Country	Serial No. on Map	District	No. of sample collected	Country	Serial No. on Map	District	No. of sample collected
Ethiopia	53	Diga	10	Ethiopia	64	Boji bermeji	1
	54	Lalo Assabi	3		65	Gimbi	1
	55	Nedjo	5		66	Leka delecha	3
	56	Guangua	3		67	Dangla	1
	57	Banja	5		68	Dure Bete	4
	58	Bahir Dar Zuria	6		69	Mandara	1
	59	Bure	2		70	Mecha	4
	60	Demecha	1		71	Ankussa-Abdo	1
	61	Angebo	1		72	Jabi Tana	2
	62	Leta Sibu	1		73	Qilxxu Kara	1
	63	Bila	1				

2.2 Media preparation and culturing of *M. oryzae* isolates

2.2.1 Preparation of solid and liquid media

M. oryzae isolates used in this study were cultured on potato dextrose agar (PDA), oatmeal agar (OMA) and potato dextrose broth (PDB) as required for the specific experiments. The media were prepared as instructed by the manufacturer and autoclaved under standard conditions of 121^{O}C for 15 minutes at 15 psi. For routine culturing on solid medium (PDA and OMA), approximately 25 ml of molten medium cooled to ~ 50^{O}C - 60^{O}C were dispensed into 9 cm diameter Petri dishes in a laminar flow bench. The medium was allowed to cool and solidify for approximately 30 minutes and stored at 4^{O}C till further use. Autoclaved PDB was stored at the room temperature and to set-up fungal cultures, appropriate volume (~25 ml) was poured into sterile 9 cm diameter Petri dishes as required.

2.2.2 Revival and sub-culture of the isolates

M. oryzae isolates used in this work were available on desiccated filter paper discs stored at -20 ^{O}C. Isolates were revived by placing the filter paper discs containing the fungal mycelium and spores on PDA plates, which were maintained for 10 days at 25^{O}C in an incubator (Panasonic, UK- MIR-154 model). Isolates were sub-cultured on fresh PDA plates for 7 days and maintained as above.

2.2.3 Liquid culture for genomic DNA extraction

Small blocks of mycelial mass with no or minimal agar were harvested from the sub-cultured plates of each isolate using a sterile loop. The fresh mycelium was introduced into 25 ml of PDB in 9 cm Petri dishes for each isolate. These cultures were maintained in a stationary

incubator at 25^{O}C for 7 days to induce growth. On day 7, the mycelial mat was harvested from the liquid medium by carefully draining the liquid. The mycelial mat was rinsed with 20 ml of distilled water twice to remove any traces of the medium residue and the liquid discarded as above. The mycelial mat for each isolate was immediately placed into a sterile 20 ml plastic container and stored at -20^{O}C till further use. The liquid waste from these processes was discarded following the standard operating procedure stipulated for the biological waste disposal, as part of the licence issued by the Animal and Plant Health Agency (APHA) of UK-Defra to hold and work with the blast pathogen *M. oryzae* isolates at the School of Life Sciences laboratories, University of Bedfordshire.

2.3 Isolation and quantification of genomic DNA

Two methods were utilised for the fungal genomic DNA extraction depending on the experimental purpose. This includes a column-based method using a commercial kit and Hexadecyltrimethylammonium bromide (CTAB) method. Prior to the DNA extraction for each isolate, frozen mycelial mat (section 2.2.3) was freeze-dried for approximately 24 hours. The freeze-dried mycelial mat for each isolate was placed in a sterile mortar; a small amount of Chelex® and liquid nitrogen or sterile acidic sand were added depending on the purpose of the experiment. The freeze-dried mycelium was ground to fine powder with a sterile pestle (for 10 to 15 minutes with acidic sand and approximately 5 minutes with liquid nitrogen). The mycelial powder for each isolate was stored at -20^{O}C till further use for the genomic DNA extraction.

2.3.1 Extraction of genomic DNA for PCR amplifications (column-based method)

The extraction of the genomic DNA through the column-based method was carried using GenElute^{TM} Plant Genomic DNA Miniprep commercial Kit (Sigma, UK-G2N350-1KT). Prior to the use of the kit, the wash solution was prepared by diluting the concentrate from the kit with absolute ethanol as directed by the manufacturer. Approximately 100 mg of the mycelial powder was transferred into a 2ml Eppendorf tube and the DNA isolation was performed following the manufacturer protocol. To elute the DNA, 100 µl of pre-warmed sterile water at 65^{O}C was added directly to the binding column and incubated at room temperature for 5 to 7 minutes. The DNA was eluted from the binding column by centrifugation at 13,000 RPM for a minute. The elution process was repeated by loading the first elute into the binding column and centrifugation at the same speed for a minute to ensure maximum yield and concentration of the DNA. After the second elution, the binding column was discarded, and the genomic DNA was stored at -20^{O}C till further use.

2.3.2 Extraction of genomic DNA for Illumina sequencing (column-based method)

Approximately 200 mg of the mycelial powder was transferred into 2 ml Eppendorf tubes for the DNA extraction. The DNA extraction was performed using the method referred in section 2.3.1 with a slight modification to the manufacturer protocol. The volume of the lysis solutions A and B and the precipitation solution were doubled. RNase digestion was performed using 40 µl (20-40 mg/ml concentration) at 65^{O}C as suggested by the manufacturer. The DNA elution was performed as stated in section 2.4.1 but using a pre-

warmed (65^{O}C) Tris-EDTA (TE) buffer solution (10mM Tris-HCl, 1mM disodium EDTA, pH 8.0).

2.3.3 Extraction of genomic DNA for PacBio sequencing (CTAB method)

Genomic DNA was extracted from the mycelial powder using a modified CTAB extraction protocol. Cut tips were used for pipetting throughout the process to extract higher molecular weight DNA. Approximately 300 mg mycelial powder was transferred to a sterile 15 ml Fisher tube containing 3 ml pre-warmed 2X CTAB buffer (CTAB 2 %, Tris base 100mM, EDTA 10mM and NaCl 0.7M), mixed vigorously and incubated at 65^{O}C for 30 minutes. After the incubation, 3 ml of chloroform-isoamyl alcohol 24:1 was added and gently mixed on a shaking platform for 20 minutes. The sample was centrifuged at 4^{O}C at 10,000 xg for 20 minutes, the aqueous phase was carefully transferred to a fresh 15 ml tube. The chloroform-isoamyl alcohol extraction was repeated twice to ensure maximum removal of protein and the RNA was digested with 70 µl of RNase solution (20-40 mg/ml concentration) at 37^{O}C for 45 minutes. To precipitate the nucleic, after the third chloroform-isoamyl alcohol extraction, the aqueous phase was transferred to a 15ml tube and 4 ml isopropanol was added, which is mixed by through inversion and incubated at -20^{O}C for approximately 7 hours for maximal DNA precipitation. The sample was centrifuged at 10,000 xg at 4^{O}C for 20 minutes and the filtrate discarded. The pellet was resuspended in 800 µl TE buffer till the pellet is no longer visible. The DNA solution was transferred to a 2 ml Eppendorf tube and equal volume of phenol-chloroform-isoamyl alcohol (25:24:1) was added. The sample was mixed on a shaking platform for 20 minutes at room temperature and centrifuged at 16,000 xg for 20 minutes at 4^{O}C. The aqueous phase was carefully pipetted into a fresh 2ml Eppendorf tube

and the DNA was reprecipitated with one tenth of 3M sodium acetate and 2 volumes of absolute ethanol. The sample was incubated at -20^{O}C for 10 minutes and centrifuged at 16,000 xg for 20 minutes at 4^{O}C and the filtrate discarded. The pellet was washed with 70 % ethanol and the filtrate discarded. The pellet was air dried in a sterile hood for approximately 15 minutes and resuspended in 100 µl of pre-warmed (65^{O}C) elution buffer by incubation for an hour at 4^{O}C to allow the DNA pellet to completely dissolve. The genomic DNA was stored at -20^{O}C till further use.

2.3.4 Gel electrophoresis

Agarose gels were routinely used for electrophoresis of genomic DNA (0.7 %) or PCR products (1.0 %). To prepare the gel, 0.7 or 1.0 g agarose powder (Sigma, UK- A9539) was added to 100 ml of 1X TAE buffer. The gel suspension was boiled in a microwave oven for approximately 3 minutes ensuring the agarose was fully melted. The gel solution was allowed to cool to approximately 50^{O}C and 5 µl of ethidium bromide (10 mg/ml stock; E1510, Sigma, UK) was added. Prior to this, comb was placed into the gel tray. The gel solution containing the ethidium bromide was carefully poured into a gel tray avoiding the generation of air bubbles and allowed to polymerise for approximately 30 minutes.

2.3.5 Quantification of Genomic DNA

The DNA samples were quantified using the electrophoretic and Nanodrop methods to assess the concentration, quality and integrity. In the electrophoretic method, 2 µl DNA + 6 µl of sterile water + 2 µl of 5X loading buffer was prepared for each genomic DNA sample in a 0.5 µl microfuge tube. The DNA samples and known concentrations (30 ng, 60 ng, 90 ng

and 120 ng) of Lambda DNA (λDNA) were loaded into wells of a 0.7 % agarose gel and electrophoresed for approximately 150 mins at 45 V. The DNA concentration was estimated by comparing the intensity of the DNA florescence in each sample with the known concentration of λDNA. For the Nanodrop method, the Nanodrop machine was initially calibrated using a known concentration of λDNA (1ng/µl) and the DNA samples were processed according to the manufacturer's instructions. DNA concentration of each sample was recorded using the 50ng reference value for DNA. The quality of each sample was assessed and recorded based on the 260/230 and 260/280 ratios.

2.4 Identification and screening of genetic markers used in this study

2.4.1 Identification of highly variable loci from known genetic markers

Fifty (50) genetic markers widely used for fungal diversity and phylogenetic analysis including the ITS region were initially assessed bioinformatically. This was done by identifying these genes and/or the region in the genome sequence of the *M. oryzae* reference isolate (70-15) available at the Broad institute database. The nucleotide sequence for each locus was extracted, and BLASTN analysis was performed against the genomes of nearly 50 *M. oryzae* isolates from 4 different hosts (*Oryzae sativa*, *Eragrostic curvula*, *Eleucine indica*, *and Setaria viridis*) and diverse geographic locations available at the NCBI database to identify any differences among the isolates. Further analysis of selected sequences showing differences was carried out by generating alignments using the ClustalOmega algorithm and the nucleotides differences was assessed (as number of SNPs) among *M. oryzae* isolates from the same host and between different hosts.

2.4.2 Identifying novel variable loci from the FUNYBASE (Fungal phylogenomic database)

From the 246-single copy ortholog genes contained in the FUNYBASE, more than 150 genes were selected for bioinformatics analysis. The protein sequence of *M. oryzae* gene available in FUNYBASE was downloaded and used as a query to perform a tBLASTN search against the reference genome of the *M. oryzae* isolate 70-15 databases (Dean *et al*., 2005). The nucleotide sequence of the gene was downloaded and used to perform BLASTN analysis against the genome sequences of the range of *M. oryzae* isolates as described in Figure 2.2.

2.4.3 Identifying novel variable loci from specific chromosomes

Fifty loci were randomly selected from each of the chromosomes 1, 2, 6, and 7 reported to contain large number of variable genomic regions in *M. oryzae* (Gwoda *et al*., 2015). BLASTN analysis and alignments of each gene were performed as stated in section 2.4.1. The nucleotide differences for each gene was assessed and recorded. In addition, BLASTN analysis and alignments were performed as above with some of the contigs that were randomly identified from these chromosomes. The main purpose is to evaluate the potential of the inter-genic regions.

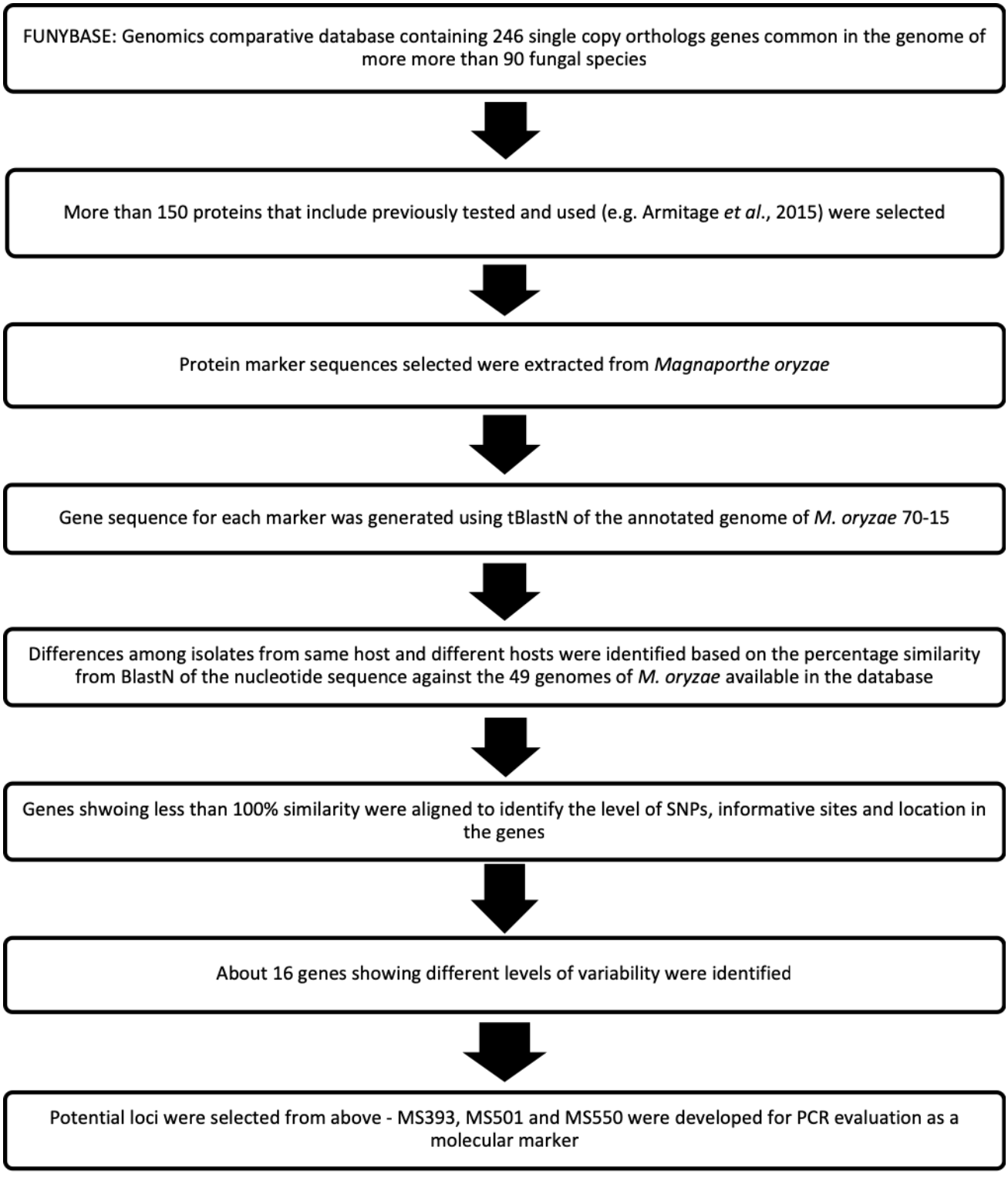

Figure 2.2: Flow diagram showing the key steps employed in the identification of a novel marker from FUNYBASE

2.4.4 Design and development of molecular markers

The loci that showed higher number of SNPs within a length of approximately 250 to 900 bp were selected. The primer sequences were designed from the 5' and 3' conserved sequences flanking the variable regions. To ensure the efficiency of new primers developed, various parameters such as the length of oligonucleotides, the G-C content and the sequence (e.g. avoiding a stretch of G) were considered. PCR products of the selected marker regions were sequenced from a set of 15 isolates chosen from previous collection. A total of 13 regions were screened (Table 2.2A). Five markers were identified from previously published work related to fungal phylogenetics including *M. oryzae*. These include ITS and Histone 4 (HIS4) as well as TEF1-α, MPG1 and NUT1 markers (Couch *et al*., 2005). Eight markers containing parts of genomic loci newly identified in this work were amplified using PCR primers designed from the *M. oryzae* genome sequence in this study. These markers include MS393, MS501a & b, MS550, HyP1, HyP2, HyP3 and HyP4 (Table 2.2A). The sequence data generated for each marker region from the 15 isolates were aligned and the number of SNPs and informative sites are recorded.

Table 2.2A. Primers for the target gene loci used in PCR with the annealing temperature and expected product size.

S/N	Genetic locus	Marker Designation	Primer F 5' -3'	Primer R 5'-3'	Annealing Temp. (^{O}C)	Expected size (bp)
1	Internal transcribed spacer	ITS1 & 4[a]	TCCGTAGGTGAACCTGCGG	TCCTCCGCTTATTGATATGC	55	600
2	Translation elongation factor	TEF1-α[a]	CATCGAGAAGTTCGAGAAGG	TACTTGAAGGAACCCTTACC	55	350
3	Hydrophobin gene	MPG1[b]	AGAAGGTCGTCTCTTGCTGC	TTCACTCAACGCTGATCGC	55	450
4	Nitrogen regulatory protein	NUT1[b]	CATGATGCACGTCAATCTGC	CTGTGTCGGTGTCGGTGTCTGAACGC	58	350
5	**Histone 4**	**HIS4[c]**	**CTCTTGAGAGCGTAGACAAC**	**CTCATCCCACATGCACTTGAC**	**62**	**400**
6	**NN***	**MS393[c]**	**GCTTTGCACTGGGTCCTAC**	**CTCTCCCTCAACACCTTCCAT**	**64**	**722**
7	**Karyopherin**	**MS501a[c]**	**GCTCAGGATGTCTTCACGCC**	**TAGCAGCTCGGTTGTCCTCT**	**66**	**745**
8	**CDP***	**MS501b[c]**	**TGGTCCCTGAGAGCGAAACT**	**TGACCATGAAACGCCAGCAC**	**60**	**751**
9		**MS550[c]**	**CTTGTGCAGCCGTCTACAGA**	**CAAGGATGGGCCTATGTCATG**	**60**	**650**
10	**Hypothetical protein**	**HyP1[d]**	**CTCACTGTTACGACCAGACAG**	**TGGTAGGTGCGTACCTGTAC**	**62**	**840**
11	**Hypothetical protein**	**HyP2[d]**	**ATCTCCTTGGTAAATGCTGCCC**	**CGACTCGGAAAAATAATGCCGC**	**66**	**552**
12	**Hypothetical protein**	**HyP3[d]**	**CGCTCGTCCTGAGCATACAGT**	**GACCCAGCATGAGCGAGTTGA**	**58**	**624**
13	**Hypothetical protein**	**HyP4[d]**	**GTTGGAGGCTCTTGTATATGCC**	**TTTGGTGTCGACGAGGAGAGAAG**	**58**	**617**

[a]Markers from Carbone and Kohn, 1999; [b]Markers from Couch *et al*., 2005; [c]Markers developed in this work from FUNYBASE;

[d]Markers developed in this work by analysing the variable regions of the chromosomes in *M. oryzae* genome sequence

*NN: N2, N2-Dimethlyguanosine tRNA methytransferase; *CDP: CDP-diacylglycerol-glycerol-3-phosphate-3-phosphatidyltransferase

Marker regions identified and new PCR primers designed in **this study** are shown in Bold;

Markers selected for further population analysis are shown in Green.

Table 2.2B. Primers used for specific targets in PCR with the annealing temperature and expected product size.

S/N	PCR Target	Designation	Primer F[a] (5' - 3')	Primer R[a] (5'- 3')	Annealing Temp. (^{O}C)	Expected size (bp)
1	*M. oryzae* specific	MG	CTGTCGTTGCTTCGGCGGGCACGC	ACGCCGGGACGATCCGAACGAGGTTC	55	400
2	Grasshopper	PES	GCGTTCGAAGCGTTGAAACAC	AGCTATATAAGCCCTAAGGTATTGC	60	1347
3	Grasshopper	PKE	CGGAATTCTTCAGTCACGGGAACACGC	TCCGAGGTGCACATGTGTGAAACGC	60	836
4	Mating type gene *MAT1-1*	MAT1-1	TGCGAATGCCTACATCCTGTACCGC	CGCTTCTGAGGAACGCAGACGACC	60	960
5	Mating type gene *MAT1-2*	MAT1-2	TCTGCTTGAAGCTGCAATACAACGG	CATGCGAGGGTGCCATGATAGGC	60	802

[a] Primers developed by Takan *et al*., 2012

2.5 PCR Amplification and purification and quantification of products

Initial optimisation of the PCR conditions for each molecular marker was carried out in 20 µl reaction volume followed by gel electrophoresis quantification of the products. Following successful amplification of the products, reactions were scaled-up to 50 µl reaction and the products were quantified. Pre-mixed 2X Biomix red (Bioline, UK) was used routinely, the mixture contains the DNA polymerase, dNTPs, $MgCl_2$, red dye which also serves as the loading buffer to achieve consistency in PCR experiments.

2.5.1 Confirmation of the isolates as *M. oryzae* using diagnostic PCR

Previously designed primers specific for *M. oryzae*, designated as MGF (forward) and MGR (reverse) (Table 2.2B) were used to screen isolates representing historical and contemporary population to confirm their identity as *M. oryzae* isolates. The primers were designed from the rRNA gene block internal transcribed spacer (ITS) region (Takan, 2007). For each isolate, 20 µl PCR contained 10 µl of 2X BioMix Red, 1 µl each of primer pair (forward and reverse), 7 µl of sterile water and 1 µl of template DNA. DNA of previously characterised isolate K23/123 and a tube with water instead of DNA were used as the positive and negative control, respectively. The amplification was performed in a thermal cycler with 35 cycles consisting of denaturation at 94^{O}C, annealing at 55^{O}C, and extension at 72^{O}C. The first cycle consisted of 94^{O}C for 2 minutes and 55^{O}C for a minute followed by 33 cycles for 94^{O}C, 55^{O}C and 72^{O}C for 30 seconds and a final cycle with extension for 5 minutes at 72^{O}C. The PCR products were electrophoresed on 1 % agarose gel stained with ethidium bromide for 60 minutes at 70 Volts. The gels were viewed under UV light and the data documented.

2.5.2 PCR reactions to amplify the product for each marker region

The details of the various primer pairs used in this initial screening and subsequent population characterisation work have been presented above (Table 2.2). For the initial PCR testing and optimisation of the various markers, 20 µl PCR were prepared containing 1 µl of template DNA, 1 µl each forward and reverse primer (20 µM stock), 10 µl 2X Biomix and 7 µl of sterile water. For negative controls with no DNA, 8 µl of sterile water was added along with all other components. Following successful optimisation and good amplification of the products, for subsequent experiments, the PCR was scaled-up to 70 µl.

With the markers selected for further work, 70 µl PCR included 3 µl DNA, 3 µl each forward and reverse primer (20 µM stock), 35 µl of 2X Biomix and 26 µl of sterile water. For the negative controls, 29 µl of sterile water was added along with all other components except DNA. PCR amplifications were performed in a thermal cycler (Bio-rad) with 34 cycles consisting of denaturation, annealing (Table 2.2A) and extension. The first cycle consisted of $94^{O}C$ denaturation, specified annealing temperature (Table 2.2A) and $72^{O}C$ extension for a minute followed by 33 cycles of denaturation, annealing and extension for 30 seconds and a final cycle with extension for 5 minutes. Following the PCR, products (5 µl of each sample) were routinely electrophoresed on 1 % agarose gel stained with ethidium bromide at 70 Volts for 60 minutes. The gels were viewed under UV light and the data documented

2.5.3 PCR amplicon purification and quantification

The PCR products in the 65 µl volume that remained for each sample after the electrophoresis were purified using a PCR product purification kit (28006, QIAGEN, UK) according to the manufacturer protocol. The purified DNA was eluted in 30 µl pre-warmed

(65^{O}C) sterile water. To quantify the purified DNA, a 10 µl mixture containing 2 µl purified amplicon, 2 µl 5X DNA loading dye and 6 µl sterile water were electrophoresed alongside a DNA quantification marker (Easyladder1, Bio-33062; Bioline, UK) on 1 % agarose gel at 70 Volts for 60 minutes. The amplicon size and concentration were estimated with reference to the DNA quantification molecular ladder mentioned above.

2.5.4 PCR screening of *M. oryzae* isolates representing contemporary populations for the Grasshopper (*grh*) repeat element

Two primer pairs designated as PES (forward and reverse) and PKE (forward and reverse) from different regions of the *grh* repeat element (Table 2.2B) were used to amplify approximately 1347 bp and 836 bp fragments, respectively. These primers were previously designed from the *grh* sequence available in the EMBL database (Takan, 2007). For each isolate, two separate PCR reactions were set-up with the PES and PKE primer pairs. Each 20 µl PCR consisted 10 µl 2X BioMix Red, 1 µl each of the forward and reverse primers (20 µM stock), 7 µl sterile water and 1 µl DNA. For these PCR experiments, DNA from the *M. oryzae* G22 from which the *Grasshopper* element was originally described (Dobinson *et al*., 1993) was used as positive control and sterile water as negative control. The thermal cycling conditions consisted of denaturation at 94^{O}C, annealing at 60^{O}C and 72^{O}C for the extension. The first cycle consisted of 94^{O}C for 2 minutes, a minute at 60^{O}C and 72^{O}C, followed by 40 cycles of 94^{O}C, 60^{O}C and 72^{O}C each for 30 seconds, and a final cycle with extension for 10 minutes at 72^{O}C. The PCR products were resolved on agarose gels and the data were recorded as described earlier.

2.6 DNA sequencing

2.6.1 Nucleotide sequencing of PCR amplicons

DNA capillary sequencing based on Sanger chemistry was carried out utilising the service providing facility at the University of Cambridge, UK. For each sequencing reaction, the required quantity of purified PCR product was provided in 10 µl as recommended. The quantity of template DNA for sequencing was calculated based on 20ng per each 100 bases of the target fragment. The sequencing primer was provided as a 10 µM stock. For each isolate, sequence data was generated routine using the forward primer. Sequence data with the reverse primer was generated for validation and/or to resolve any ambiguous bases with some isolates. BioEdit program (Hall and Carlsbad, 2011) was utilised to view the DNA sequence trace files and the data was exported into the Geneious software, v9.1.8 (Kearse *et al*., 2012) for further analysis.

2.6.2 Genome sequencing of 18 *M. oryzae* isolates using the Illumina platform

Approximately 4 µg of genomic DNA in elution buffer was used for the construction of libraries to generate Illumina genome sequence outputs. For isolates K23/123 and E34, which were used to develop the reference genomes, two libraries were prepared using the TruSeq DNA PCR-Free Illumina Paired end (PE) and the Nextera Mate-Pair (MP) kits. The DNA fragment size range was 350 to 550 bp for the PE libraries, whilst for the MP libraries, the DNA fragment size ranged from 2 to 4 kb. The two libraries were sequenced using the Illumina MiSeq 600 cycles system to generate reads of 35 to 301 nucleotide bases length from both forward and reverse fragments. This work was carried out utilising the service provision at the University of Cambridge, UK.

For the resequencing of 16 *M. oryzae* isolates representing the finger millet blast pathogen population diversity in Eastern Africa, paired end (PE) libraries was prepared using TruSeq DNA PCR-Free Illumina library kit with a fragment size range of 150 to 200 bp. The libraries were sequenced using the Illumina HiSeq 4000 system to generate 100 bases per read length from each end of the fragments. This work was carried out utilising the service provision facility at the McGill University-Genome Quebec Innovation Centre, Canada.

2.6.3 PacBio sequencing of the reference isolates K23/123 and E34

Approximately 25 µg of genomic DNA in elution buffer was used to construct libraries where the sheared DNA size selection cut-off was from 15 to 50 kb. The size selected fragments were used to prepare a single molecule real time bell (SMRT-bell) template library based on PacBio protocol. The cut-off step was repeated to remove any smaller fragments generated during the SMRT-bell library preparation to minimise the number of DNA fragments less than 15 kb when subjected to sequencing. The library was sequenced using the PacBio Sequel system. This work was carried out utilising the service provision facility at the University of Liverpool, UK.

2.7 Bioinformatic analysis of DNA sequences

2.7.1 Multiple sequence alignment of marker data

DNA sequence data generated for each marker from each isolate were exported to the Geneious (v9.1.8) and analysed with various algorithms available as plugins both for multiple sequence alignment and phylogenetic analysis, unless otherwise stated. Multiple

sequences for each marker were initially aligned using MAFFY. The alignment was visually inspected, edited and/or trimmed where required to ensure the sequence data for all isolates covers the same region (i.e. starting and ending with conserved sequences both at the 5' and 3' ends). This data was realigned using the MUSCLE to generate a best-fit alignment for each marker. These alignments were subsequently used to generate a multi-locus concatenated alignment (combined sequences) containing ITS, HIS4, HyP1 and HyP2 markers using Geneious.

2.7.2 *M. oryzae* genotype diversity and distribution

Multiple sequence alignment data were analysed using Population Analysis with Reticulate Trees (POPART) program (Leigh *et al*., 2015) to determine the number of genotypes (each representing one or more isolates with identical DNA sequence) formed based on single marker (single locus) and concatenated sequence (multi-locus) data. The alignments were exported into POPART programme and the traits were constructed based on the numbers of isolates collected from each country. The alignments and the traits files were used collectively to determine the genotypes produced per marker sequence using the median joining network. Each circle produced represents a genotype and its size corresponds to the number of isolates per circle. Within the circles, differently coloured slices reflect the number of isolates from each country. The output was edited appropriately for effective visualisation.

2.7.3 Phylogenetic analysis

Phylogenetic trees were generated both from the single locus and the multi-locus alignment data using the Bayesian method. Prior to this, multiple sequence alignments were exported to MEGA v7.0.26 (Kumar *et al*., 2016) to calculate the best-fit nucleotide substitution model for each locus. Based on the nucleotide substitution model determined in MEGA for each locus, a Markov Chain Monte Carlo (MCMC) algorithm was used to generate the phylogenetic trees with Bayesian probabilities using MrBayes plugin within Geneious (Kearse *et al.,* 2012). The MCMC analysis was performed using the best evolutionary nucleotide substitution models determined in MEGA (v7. 0.25) with 5,000,000 cycles set at 4 heated chains and burn-in length of 100,000 were run and sampled every 500 generations. Geneious was used to estimate 70 % Consensus phylogenies of the Bayesian probability for each MrBayes run where the first 25 % trees were discarded as burn-in. Phylogenetic groups or clades were identified by visual inspection of the tree topology.

2.7.4 Pre-processing of raw sequence reads, *de novo* assembly of the genome sequences and quality assessment

The quality of the raw sequences generated and the presence of any adaptor sequences were initially checked using the FASTQC (v0.11.7) program. The adaptor sequences where present in the raw data and low-quality bases with the Phred score of less than Q20 were removed using BBDuk plugin in Geneious 9.1.8 (Kearse *et al*., 2012), Trimmomatic v0.35 (Bolger *et al*., 2014) and NxTrim (O'Connell *et al*., 2015). *De novo* assemblies of the Illumina read sequences were performed using the SPAdes assembler software v3.11.1 (Bankevich *et al*., 2012) with default setting. Scaffolds less than 1 kb in size were removed. CANU

assembler v1.6 software (Koren *et al*., 2017) with default setting was used to assemble the raw data from PacBio sequencing. Due to the error rate of PacBio sequencing (Ross *et al*., 2013) the assembled genome from the CANU were polished using high-quality raw reads from Illumina PE sequencing from the corresponding isolates K23/123 and E34 using Pilon program (Walker *et al*., 2014) and the polished assemblies were merged to their corresponding genome assemblies from the SPAdes using Quickmerge program (Chakraborty *et al*., 2016) for isolates K23/123 and E34. All the assembled genomes were screened for mitochondrial sequences and other contaminant sequences, which were discarded to retain only the nuclear genome assemblies for each of the *M. oryzae* isolates selected for this study. The quality of the assembly was evaluated using the QUAST program (Gurevich *et al*., 2013) based on the standard assembly statistics such as the N50 values, number of contigs, average contig size and the total assembled genome size. In addition to this, the Benchmarking Universal Single-Copy Orthologs (BUSCO) algorithm was used to assess the presence of ortholog genes in Fungi (Simao *et al*., 2015) as part of the yardstick to measure the completeness of the genome assembly.

2.7.5 *De novo* gene prediction and gene distribution patterns in *M. oryzae* isolates

Augustus (v2.7) program (Hoff and Stanke, 2013) with *Magnaporthe grisea* as the model (species = Magnaporthe grisea, strand = both, gene model = complete) was used to predict the complete gene sets in the genomes of the *M. oryzae* isolates selected in this study. OrthoVenn program (Wang *et al*., 2015) was used to analyse the distribution pattern of the genes by comparing the total predicted protein coding gene set across various isolates representing the finger millet blast pathogen populations in Eastern Africa.

2.7.6 Genome synteny analsyis of *M. oryzae* isolates

Synteny Mapping and Analysis Program (SyMap) (Soderlund *et al*., 2006) was used to compare the assemblies of the two reference genomes developed from isolates E34 and K23/123, each representing distinct *M. oryzae* populations. Following the default setting, SyMap uses NUCmer for multiple genome alignments through a modified Smith-Waterman algorithm. This analysis was used to generate a circular map showing the region specific to each *M. oryzae* isolate (representing the *grh* positive and *grh* negative type populations) and dot-pot displays of syntenic regions between the assemblies of the two isolates.

2.7.7 Identification and analysis of Grasshopper element in the genome of FMB pathogen isolates

Previously deposited *Grasshopper* (*grh*) DNA sequence from G22 isolate available in the Genbank database was identified and downloaded (Dobinson *et al*., 1993). BLASTN analysis of the two reference genome assemblies generated in this study was performed using the downloaded *grh* element sequence (reference *grh* sequence). This analysis was used to identify the presence of full and partial DNA sequence and the level of coverage of *grh* element in the genome of the isolates. The longest DNA sequence of the *grh* element copy in the genome of the two isolates was aligned to assess the level of variation using the MUSCLE program in Geneious. Furthermore, DNA sequence of the selected copy of the *grh* element in the genome of each of these isolates was compared to the isolate G22.

2.7.8 Variant calling

To initiate the process of identifying single nucleotide polymorphisms (SNPs), curated PE sequence reads for each isolate were mapped to the genome assembly of E34 (Reference isolate containing the *grh* element) using Burrows-Wheeler Aligner (BWA) v0.7.5a (Li *et al*., 2009). Alignment files obtained from the reference mapping were filtered and sorted into the standard output format using SAMtools (v0.1.19). The alignment SAM files were converted into BAM files and sorted (Li *et al*., 2009; Li *et al*., 2009; Li, 2011). The alignment BAM files were subjected to a combination of pipeline tools including SAMtools (mpileup), BCFtools v0.1.19 (Narasimhan *et al*., 2016) and VCFtools v1.4.1 (Danecek *et al*., 2011) with default settings to remove misalignment and VCF file generated. The VCF file generated for each alignment contained both the heterozygous and homozygous SNPs, and the Indels. The VCF files were filtered to remove the heterozygous SNPs and Indels. The homozygous SNPs between the test isolate and the reference isolate were identified using a combination of tools including VCFtools, varfilter and vcfutils.pl (Danecek *et al*., 2011; Yoshida *et al*., 2016).

2.7.9 Phylogenomic analysis of BUSCO ortholog genes using the maximum likelihood method

For this analysis, alongside the 18 genome assemblies of the FMB pathogen isolates from Eastern Africa generated in this study, 9 genome assemblies of *M. oryzae* isolates from finger millet in India as well as from other hosts including rice, wheat, wild finger millet and foxtail millet available in the GenBank database were included to enable a wider comparative analysis. BUSCO analysis using Sardiomycete single copy gene sets model was carried out to predict the BUSCO genes in the genomes of all the 27 isolates. Predicted single copy core Sardiomycete genes in the assemblies were extracted and aligned to

identify the genes common in all genomes using MAFFT (Katoh and Standley, 2013). The genes found common in all the genome assemblies (isolates) with phylogenetically informative levels of variation were retained and the genes that did not show the levels of variation were trimmed using TrimAL v1.4.1 (Capella-Gutierrez *et al*., 2009). The identified genes were concatenated and various isolates were aligned and used to construct a phylogenetic tree using RAxML. A single consensus phylogeny tree was generated from the concatenated sequence data using ASTRAL v5.6.1 (Mirarab *et al*., 2014; Shekhar *et al*., 2017). The resulted tree was visualised using the R package GGtree v1.12.4 (Yu *et al*., 2016).

2.8 Mating type identification and fertility status of *M. oryzae* isolates

2.8.1 Mating type-specific PCR screening of the *M. oryzae* isolates representing the contemporary populations

Two primer pair designated as MAT1-1F (forward) and MAT1-1R (reverse) and MAT1-2F (forward) and MAT1-2R (reverse) developed previously (Table 2.2B, Takan *et al*., 2012) were used in PCR to check for the amplification of approximately 960 bp or 802 bp fragments specific for MAT1-1 and MAT1-2, respectively. These primers were used to screen 224 *M. oryzae* isolates (Table 2.1) representing the contemporary population in this study collected from different districts of Kenya, Uganda, Ethiopia and Tanzania. For each isolate, a 20 µl PCR consisted of 10 µl of BioMix Red, 1µl of each primer (20 µM stock), 1 µl of template DNA and 7 µl of sterile water. DNA from previously characterised known Tester isolates I-R-22 and JP15 were used as positive control for MAT1-1 and MAT1-2, respectively and a tube with water instead of genomic DNA was used as a negative control. The PCR thermal cycling conditions were as follows: initial denaturation at 94^{O}C for 2 minutes, annealing and

extension for a minute each at 60^{O}C and 72^{O}C, respectively followed by 41 cycles of 94^{O}C, 60^{O}C and 72^{O}C for 30 seconds each and a final extension for 5 minutes. The PCR products were routinely electrophoresed on 1 % agarose gel stained with ethidium bromide for 70 minutes at 70 Volts. The gels were viewed under a UV light and results documented using a bioimager (Biorad).

2.8.2 Fertility fitness of the known or potential tester isolates

To determine the fertility fitness of the known tester isolates of *M. oryzae* (Table 2.1), previously described procedures were used (Takan, 2007). Previously known testers included isolates TH3, I-R-22, 4136-4-3, K23/123, JP 15, BR16 and Guy11) along with a potential tester isolates D15/s47 (Table 2.1). Prior to the assay, the known tester and potential tester isolates were screened by mating type-specific PCR amplified to reconfirm their mating type. Tester isolates of opposite mating type were paired by placing small plugs of mycelium from the edge of 7 to 8-day old cultures and placed on OMA approximately 4 cm apart. The plates containing paired isolates were wrapped with parafilm and incubated at 27^{O}C ± 2^{O}C under 12 hrs light and 12 hours dark for 8 days for the isolates to grow towards each other known as the interaction zone. On the 8th day, the incubator conditions were changed to 20^{O}C $^{\pm}$ 2^{O}C and continuous white light. The plates were monitored at 5-day intervals for the presence of perithecia. The black line formed at the interaction zone, indicate perithecia, which were further observed under a stereomicroscope to confirm the perithecia production and the presence of asci and ascospores approximately on the 29th day after crossing. The isolates were designated as male or female or hermaphrodite according to the previously established nomenclature (Itio *et al*., 1983).

2.8.3 Compatibility and fertility status of contemporary isolates

Characterised isolates representing the contemporary population in this study were crossed with the selected testers of opposite mating type. Mating type of the new isolates representing the contemporary populations was determined by the mating type-specific PCR. The testers used in the crosses with new isolates included MAT1-1 isolates TH3, I-R-22 and 4136-4-3 and MAT1-2 isolates K23/123, and D15/s47. The crosses were performed on OMA as stated above. The sexual reproductive behaviour of the isolates was designated according to Itio *et al*. (1983). In female fertile isolates, perithecia were formed by the new isolates, while in male fertile (female sterile) isolates, perithecia were formed only by the tester isolates. When the perithecia were formed by both the new isolates and the tester isolates, the new isolates were designated as hermaphrodites (i.e. they were able to function as both female and male). The number of perithecia formed on the culture surface was estimated to assess the relative degree of fertility of each new isolate. The presence or absence of asci and ascospores was determined by picking up ~ 15 perithecia and placing them in a drop of water on a glass slide and squashing under a glass cover slip. The slide was mounted and observed on a stereomicroscope. For each batch of the experimental set-up, the tester isolates of the opposite mating types were crossed as positive control and to monitor the formation of perithecia with asci and ascospores.

Chapter 3

3.0 Genetic diversity of the finger millet blast pathogen in Eastern Africa

3.1 Introduction

Finger millet blast caused by the fungus *M. oryzae* is an endemic disease found in locations wherever the crop is grown. The severity of this disease is such that it accounts for a considerable proportion of finger millet production losses sufficient enough to feed millions of people annually (Strange and Scott, 2005; Lenne *et al*., 2007; Babu *et al*., 2013a). Apart from finger millet, blast disease also infects other cereal crops such as rice and wheat and ranks among the most significant plant diseases in the world threatening food sustainability (Talbot, 2003; Dean *et al*., 2012). With its distinctive capability to attack finger millet at all stages of growth from seedlings to maturity, the blast disease invasion can lead to total crop loss (Talbot, 2003; Ryan, 2016). However, given the economic and nutritional importance of finger millet to millions of people in developing countries, particularly Africa and Asia, it is crucial to develop sustainable control measures that are environmentally friendly and cost effective to the farmers. Planting of resistant cultivars has been recommended as the most effective control measure in curbing the menace of the blast disease as the crop is predominantly grown by subsistence farmers who cannot afford other methods such as expensive fungicides (US National Research Council, 1996). However, research with the rice blast system shows that the pathogen can breakdown host resistance within a relatively short period (Talbot *et al*., 1993; Talbot, 2003). Knowledge about the pathogen population

structure and its pattern of distribution are critical in the identification, development, and deployment of resistant varieties.

In understanding the population diversity and distribution pattern of fungal pathogens, various molecular techniques such as RFLP, AFLP and DNA sequencing are widely used to assess the potential of genes and transposable elements to identify and resolve the evolutionary relationships in diverse species including *M. oryzae*. In fungi, the nuclear encoded ribosomal RNA (rRNA) gene block consists of the 18S (known as the small subunit, SSU), 5.8S and 28S (known as the large subunit, LSU) rRNA genes and the spacers in between the genes (White *et al*., 1990). The rRNA gene block is present in multiple copies in the genome and the size and copy number vary from species to species (Poczai and Hyvonen, 2010). The genes are separated by two internal transcribed spacers (ITS-1 and ITS-2) and flanked by two external transcribed spacers (5'-ETS and 3'-ETS). The rRNA gene blocks are separated by non-transcribed DNA commonly known as the intergenic spacer (IGS) region (see Figure 3.1) containing a number of repeated sequences, which boost the promoter activity (Poczai and Hyvonen, 2010). The rRNA gene block has been used for fungal diagnostics and assessing phylogenetic relationships over a wide range of taxonomic levels for more than 25 years (e.g. White *et al*., 1990; Schoch *et al*., 2012). The 18S gene is less variable in fungi, while the 28S gene shows a good resolution at certain species level (Schoch *et al*., 2012). The ITS region covers the ITS-1, 5.8S gene and the ITS-2, which are flanked by the 18S gene at the 5' end and the 28S at the 3' end as shown in Figure 3.1 (White *et al*., 1990; Schoch *et al*., 2012; Koljalg *et al*., 2013; Nilsson *et al*., 2017). This region is found to be conserved across the fungal kingdom enabling high amplification success

using standard primers to identify high levels of inter-specific variation and sometimes also at intra-specific level (e.g. Schoch *et al*., 2012; Underhill and Iliev, 2014).

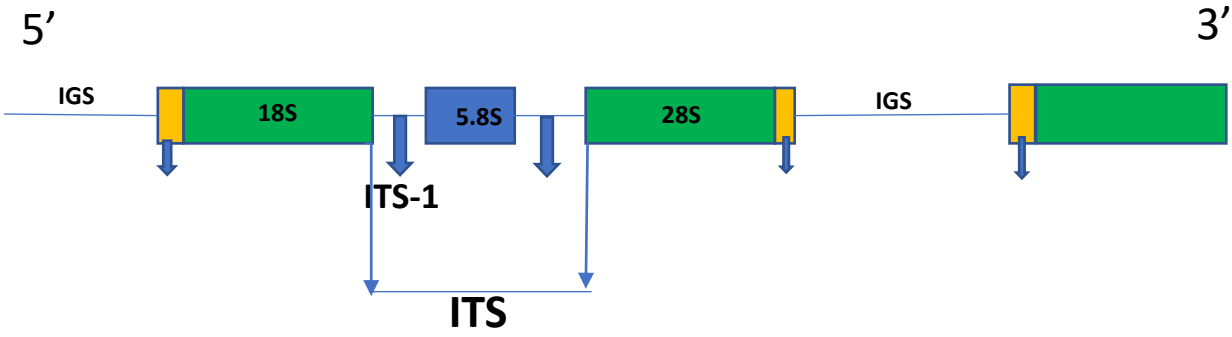

Figure 3.1 Diagrammatic representation of the nuclear encoded ribosomal RNA (rRNA) gene block

The ‘green’ and ‘blue’ boxes represent the 18S/28S and 5.8S genes, respectively; ‘yellow’ boxes represent the external transcribed spacers (ETS); thin lines represent the internal transcribed spacers (ITS) 1 and 2 and the intergenic spacer (IGS).

Over 300, 000 full length sequences of the ITS region have been deposited in the gene banks covering more than 14, 900 fungal species making the region one of the widely used genetic markers (Koljalg *et al*., 2013). The ITS region has been proposed as the universal fungal DNA barcode (Schoch *et al*., 2012; Koljalg *et al*., 2013) and this region of the nuclear rRNA gene block evolves rapidly and may be distinct among species within a genus or among populations within a species (White *et al*., 1990; Nilsson *et al*., 2008). ITS-based genetic markers have been used extensively in the classification and population study of both pathogenic and non-pathogenic fungi to understand their evolutionary relationships (e.g. White *et al*., 1990; Sreenivasaprasad *et al*., 1996; Talhinhas *et al*., 2002). In one of the studies of 52 *M. oryzae* isolates associated with the *Eleusine* species from diverse geographical locations, ITS region sequences distinguished two groups and it has been

suggested that these groups evolved independently (Tanaka *et al*., 2009). One of the groups was close to *M. oryzae* isolates from *Eragrostis* species while the other was close to isolates from *Triticum* in Brazil and isolates from *Lolium* in Japan (Tanaka *et al*., 2009). Similarly, ITS sequences were utilised to discriminate between *M. oryzae* isolates from *Eleusine* species and *Digitaria* species in East Africa (Takan 2007; Takan *et al*., 2012). However, researchers have accentuated some limitations in the resolution of the rRNA markers compared to the protein coding gene markers (Bruns *et al*., 1992; Geiser *et al*., 2006; Hofstetter *et al*., 2007; Schoch *et al*., 2009).

Protein coding genes have been identified to be more efficient and superior compared to the rRNA genes for cataloguing relationships at various taxonomic levels (Hofstetter *et al*., 2007; Schoch *et al*., 2009). Research has shown that intron rich protein coding genes generally offer better resolution of closely related species and populations compared to the rRNA gene block markers (e.g. Geiser *et al*., 2006). They have been used extensively in the identification, classification and phylogenetic analysis of fungal species (e.g. Glass and Donaldson, 1995; Couch and Kohn, 2002; Tanabe *et al*., 2004). Protein coding genes could be housekeeping genes such as translation elongation factor, actin, histone, beta-tubulin, calmodulin, and the largest subunit of RNA polymerase II or functional genes such as mating type (MAT, Kang *et al*., 1994; Ramirez-Prado *et al*., 2008) and hydrophobin (MPG1) genes particularly to analyse populations associated with different hosts and/or inter-species relationships (Talbot *et al*., 1993; Couch *et al*., 2005; Inoue *et al*., 2016). For example, molecular markers designed from some housekeeping and functional genes have been used for the phylogenetic analysis and separation of *M. oryzae* and *M. grisea* (Couch and Kohn, 2002, You *et al*., 2012; Takan *et al*., 2012) and mainly for inter-specific relationships in other

fungi such as *Fusarium* species (Donaldson *et al*., 1995), *Colletotrichum* species (Baroncelli *et al*., 2015).

Increase in the number of fungal genome sequences has revolutionised the possibilities for the identification and development of novel loci for diagnostics, diversity at the inter- and intra-species levels and phylogenetic relationships of many species (Aguileta *et al*., 2008). Applying resources from about 30 fungal genomes representing Ascomycetes, Basidiomycetes, Zygomycetes deposited in various databases such as NCBI, JGI, BROAD, 246 single copy ortholog genes have been identified (Marthey *et al*., 2008). These genes are available for public access in the Fungal Phylogenomic Database (FUNYBASE) as a resource to identify genes with high informative value for comparative studies and phylogenetic reconstruction (Marthey *et al*., 2008). This resource has shown to be useful to identify novel markers that provide good resolution in the population study of some fungal pathogens e.g. *Alternaria* species and *Fusarium* species (Aguileta *et al*., 2008; Townsend and Lopez-Giraldez, 2010, Feau *et al*., 2011; Walker *et al*., 2012; Armitage *et al*., 2015; Almiman, 2018). In addition, the genome data also serves as useful resources for identifying SNP-rich regions in the genome providing the prospect for developing novel molecular markers (e.g. Sun *et al*., 2013). High genetic variability often observed among *M. oryzae* isolates requires in-depth characterisation and monitoring of the populations to enable the development and deployment of the resistant cultivars. Moreover, the availability of high-resolution molecular markers is useful to distinguish populations adapted to different host plants.

3.1.2 Aim and Objectives

Investigations in this chapter aim to advance our understanding of the population structure and phylogenetic relationships of *M. oryzae* associated with finger millet production systems in Eastern Africa. Also, to evaluate the relationship between the *M. oryzae* isolates from finger millet in Africa and Asia based on the presence of the *Grasshopper* (*grh*) element in the fungal genome. The specific objectives are as follow:

1. To utilise the available molecular genetics and genomic resources to identify and/or develop molecular markers showing high variability among the *M. oryzae* isolates from a single host.

2. To assess the genetic diversity of the finger millet blast pathogen populations in Eastern Africa, representing historic and contemporary isolates, using the selected molecular markers.

3. To gain insights into the pathogen distribution pattern in space and time and the evolutionary relationships within and between the populations originating from different countries.

4. To screen the contemporary population of *M. oryzae* from finger millet and related hosts in Eastern Africa for the presence of the *grh* element.

3.2. Experimental approaches

Three hundred *M. oryzae* isolates (Table 2.1A) from four countries (Kenya, Uganda, Tanzania and Ethiopia) in Eastern Africa (Figure 2.1B) have been characterised using sequence-based molecular markers. Seventy-six isolates from Kenya and Uganda represent historic populations (2000-2004) and 224 isolates from all four countries represent contemporary populations (2015-2017) from key finger millet production locations (districts) (Table 2.1B). Genomic DNA extracted from each isolate was subjected to initial screening and validation with the *M. oryzae* diagnostic primers (MGF and MGR) (section 2.5.1). Three bioinformatic approaches were used to identify potential molecular markers with the resolution to decipher genetic diversity of *M. oryzae* isolates associated with a single host. These include i) screening known phylogenetic markers, ii) evaluation of genes in the FUNYBASE, and iii) analysis of variable chromosomes in the genome of *M. oryzae* isolates from a single host (e.g. rice) as described in Chapter 2 (sections 2.4.1 - 2.4.4). Subsequently, PCR amplification, sequencing and alignment of the 13 selected markers (Table 2.2A) were performed with 15 *M. oryzae* isolates representing the populations associated with finger millet to determine the efficacy of the markers (section 2.5.2 and 2.5.3). Out of the 13 markers initially tested, four markers (ITS, HIS4, HyP1 and HyP2) were selected to characterise the 300 *M. oryzae* isolates (section 2.5.2 and 2.5.3). The nucleotide sequence data generated were subjected to a range of analysis such as multiple sequence alignment to assess the genetic diversity (e.g. number of SNPs, parsimony informative sites and genotypes) and the phylogenetic relationships (e.g. Reticulate network trees and Bayesian posterior probability trees) allied to the host domestication, intensification and geographic diversity (section 2.6 – 2.7). Also, the contemporary isolates were screened for the *Grasshopper* (*grh*) transposable element

using a PCR-based approach with two sets of primers targeted to amplify different parts of *grh* designated as PKE and PES (section 2.5.4). For this analysis, *M. oryzae* isolate G22 from finger millet in Japan was used a reference, as the *grh* element was originally described from this isolate (Dobinson *et al*., 1993).

3.3 Results

3.3.1 Screening and validation of fungal isolates as *M. oryzae*

A total of 300 *M. oryzae* isolates (Table 2.1A) from various districts of Kenya, Uganda, Tanzania and Ethiopia as well as one reference isolate (G22) were used in this chapter. Initially, the isolates were observed for typical morphological appearance of *M. oryzae* on PDA plates (data not shown). Genomic DNA (gDNA) isolated from the cultures was screened using the diagnostic PCR with *M. oryzae*-specific primers designated as MGF (forward) and MGR (reverse) as described in Chapter 2 (Table 2.2B). In all the isolates approximately 400 bp product was amplified (e.g. Figure 3.2), which is comparable to the K23/123 isolate used as the positive control, confirming all the isolates as *M. oryzae*. K23/123 isolate has been previously characterised as *M. oryzae* and used in various molecular and biological assays in our laboratory. As expected, R16, a *Fusarium proliferatum* isolate used as a negative control in the PCR analysis did not amplify any product. Majority of the 300 *M. oryzae* isolates were collected from finger millet landraces and/or local varieties, and very few were from improved varieties (Table 2.1A).

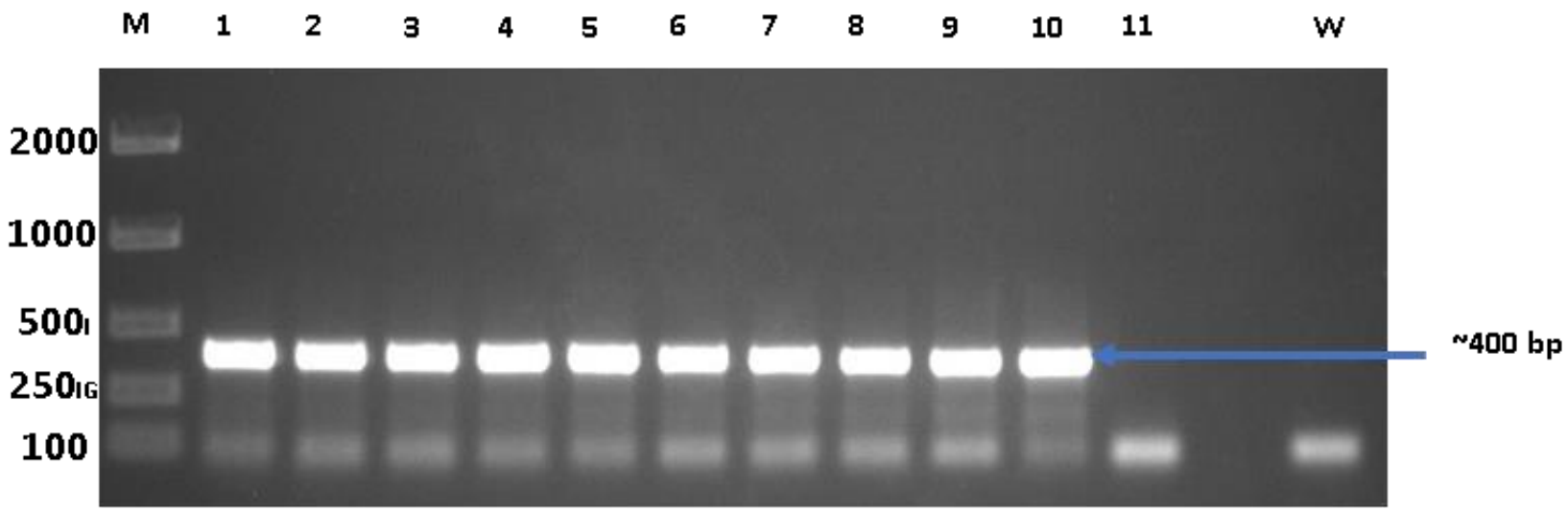

Figure 3.2. Agarose gel image showing the positive PCR product with the *M. oryzae*-specific primers

The product size for each isolate is approximately 400 bp. Lane M – Known size DNA molecular ladder (bp); Lane 1 - K1; Lane 2 - K2; Lane 3 - U1; Lane 4 – U2; Lane 5 – U3; Lane 6 -T1; Lane 7-T2; Lane 8 – E1; Lane 9- E2; Lane 10- K23/123 (Positive control, previously characterised as *M. oryzae* isolate); Lane 11- R16 (*Fusarium proliferatum* isolate as a negative control); Lane W- PCR containing water instead of genomic DNA (Negative control). Isolate details are presented in Table 2.1A.

3.3.2 Identification and development of high-resolution molecular markers

Key resources used in this context are data from known fungal phylogenetic markers, the FUNYBASE and the genome sequence data of 49 *M. oryzae* isolates from various hosts as available in NCBI at the start of this work. These were used for the three different bioinformatics-based approaches taken to identify potential molecular genetic markers as described in 3.2.

3.3.2.1. Screening of known phylogenetic markers

Fifty previously used fungal phylogenetic markers (loci) including the ITS region that has been proposed as a universal fungal barcode marker were identified from the literature. The DNA sequence data of these markers were extracted from the *M. oryzae* genomes and analysed bioinformatically to determine the level of nucleotide variability among the blast pathogen isolates from a single host (e.g. rice) and different hosts. The number of SNPs in the alignment for each locus varied from 0 to 5 for the isolates from the same host and ranged from 0 to 21 for isolates from different hosts (Table 3.1). Loci that showed a high number of SNPs among isolates from a single host include ITS, histone 3 (HIS3) and histone 4 (HIS4) ranging from 4 – 5 (Table 3.1).

Table 3.1: Number of SNPs identified for previously used fungal molecular genetic markers based on the data extracted for each marker utilising the genome sequences of 49 *M. oryzae* isolates accessed via the NCBI database

S/N	LOCUS NAME	LOCUS TAG	GENE LENGTH (bp)	CHROMOSOME	NUMBER OF SNPs WITHIN A HOST*	LENGTH OF THE VARIABLE REGION** (bp)	NUMBER of SNPs AMONG DIFFERENT HOSTS***	LENGTH OF THE VARIABLE REGION**** (bp)
1	**Internal transcribed spacer (ITS)**	**-**	**550**	**-**	**5**	**550**	**7**	**550**
2	18s rRNA biogenesis protein (RCL1)	MGG_04422	1418	2	1	1418	1-2	420
3	Actin	MGG_03982	2481	6	1	2481	6	2481
4	Glyceraldehyde-3-phosphate-dehydrogenase (GAPDH)	MGG_01084	1232	5	0	1232	2	420
5	5s RNA	MGG_20333	116	1	1	116	1	116
6	Eukaryotic translation initiation factor eIF-1	MGG_13474	1356	4	1	1356	7	936
7	Eukaryotic translation initiation factor subunit alpha	MGG_13200	2090	3	1	2090	1	2090
8	Hydrophobin like protein (MPG1)	MGG_01315	926	7	2	540	6	540
9	Chitin synthase 8	MGG_13013	6635	3	2	2000	9	6335

10	Translation initiation factor	MGG_12301	2464	2	1	2464	3-7	1020
11	Transcription elongation factor	MGG_12268	3250	2	1-3	1860	12-14	1860
12	Translation initiation factor elf- 2B subunit	MGG_12177	2281	6	1	2281	2-16	1400
13	Ras protein let-60	MGG_12077	2511	6	1	2511	5-12	1600
14	Transcription elongation factor SPT6	MGG_11491	4991	5	1	4991	2-6	4000
15	Translation initiation factor IF-2	MGG_11482	4117	5	1	4117	3-7	1000
16	Eukaryotic translation initiation factor 3 subunitA	MGG_10192	4183	2	0	4183	1-2	4183
17	Protease	MGG_09246	2813	1	2	2100	4-6	2160
18	Fungal specific transcription factor domain	MGG_09027	3778	7	0	3778	9	1000
19	Translation initiation factor e1F4E3	MGG_08170	2746	2	1	2746	1-2	2746

20	Elongation factor Tu	MGG_08162	2039	2	2	1740	4-10	1800
21	Plasma membrane ATPase	MGG_07200	4354	7	2	2940	3-7	1506
22	Eukaryotic translation initiation factor 3 subunit	MGG_07109	1802	2	1	1802	3-10	1026
23	Glutamine Synthetase	MGG_06888	2768	1	2	2768	3-8	1806
24	Calmodulin	MGG_06884	977	1	1	977	1	977
25	Histone H2A	MGG_03577	1325	4	1	1325	4	2000
26	Histone H2B	MGG_03578	1069	4	1	1069	2-4	1000
27	**Histone H3**	**MGG_01159**	**1094**	**5**	**5**	**900**	**7-12**	**1000**
28	**Histone 4**	**MGG_06293**	**1064**	**4**	**4**	**300**	**4**	**600**
29	Histone 4B	MGG_01160	860	5	2	650	2	650
30	Ras-like protein	MGG_06154	2101	3	1	2101	1	2101
31	Translation initiation factor e1F-2B subunit	MGG_06077	1916	3	1	1916	2	600
32	Eukaryotic Translation initiation factor 5B	MGG_06066	4469	3	1	4469	1	4469

33	Chitin Synthase D	MGG_06064	2702	3	1-2	700	3-6	1000
34	Chitinase 1	MGG_05533	1388	1	1	1200	5	1250
35	Chitinase	MGG_04534	4604	2	2	4600	9-12	4600
36	Chitin synthase 2	MGG_04145	4551	6	1	4511	9-11	1500
37	Elongation factor 1-alpha	MGG_03641	2606	4	1-2	300	2-6	1000
38	Chitin synthase 1	MGG_01802	4137	2	1-2	2	5	3000
39	Hydrophobin	MGG_01173	469	2	1	469	1-7	469
40	Cutinase 1	MGG_00734	871	5	0	871	0	871
41	Tubulin beta chain	MGG_00604	2498	5	0	871	2	2498
42	N-glycolase/ DNA lyase	MGG_00367	1443	5	0	1443	6-7	1000
43	DNA lyase 1	MGG_10609	1703	1	0	1703	7-8	1703
44	DNA lyase 2	MGG_02980	3836	7	1	3836	3-8	1200
45	DNA- directed RNA polymerase II largest subunit	MGG_04652	6954	2	1	6954	4	2500
46	DNA- directed RNA polymerase III subunit RPC1	MGG_04477	4700	1	1	4700	11-12	2000

47	DNA- directed RNA polymerase 1 subunit RPA1	MGG_01158	5790	5	1	5790	3	500
48	Nitrogen regulatory protein	MGG_02755	4817	7	1	4817	5	450
49	Superoxide dismutase (SOD)	MGG_00212	1305	5	1	1305	17-21	1000
50	Superoxide dismutase (SOD)	MGG_02625	903	7	1	903	3-4	600

*Number of SNPs identified among the isolates of the same host, in this context -isolates from rice only.

**Size of the variable regions (bp) among *M. oryzae* isolates from the same host;

***Number of SNPs observed among the isolates from different hosts, in this context – isolates from rice and weeping lovegrass.

****Size of the variable region (bp) among *M. oryzae* isolates from different hosts

The locus in bold showed a higher number of SNPs in distinguishing isolates from the same host. However, only ITS and histone four were selected for further screening

3.3.2.2. Screening the FUNYBASE loci for the identification of novel phylogenetic markers

Out of the 246 single copy genes available in FUNYBASE, more than 150 loci were identified and analysed using the bioinformatics approach as described in the materials and methods (section 2.4.2 and Figure 2.2). The bioinformatics analysis of the majority of the markers did not reveal any SNPS or revealed a small number of SNPs for the *M. oryzae* isolates from the same host. The number of SNPs among the isolates from different hosts was higher (Table 3.2). Among the loci analysed, 16 loci showed SNPs ranging from 1 to 12 among isolates from the same host and 2 to 57 with isolates from different hosts based on DNA sequence alignments (Table 3.2). Two loci MS393 (MGG_01448; N2, N2, - Dimethlyguanosine tRNA methytransferase) and MS501 (MGG_01449; Karyopherin) showed higher number of SNPs 7 and 12, respectively for the *M. oryzae* isolates from the same host (Table 3.2).

Table 3.2: Number of SNPs identified from some selected FUNYBASE loci utilising the genome sequences of 49 *M. oryzae* isolates

SN	FUNYBASE ID	Protein accession number	Locus tag	Locus name	Chromosome	Gene Length (bp)	Size* (bp)	No. of SNPs	Size** (bp)	No. of SNPs
1	MS277	MGG_01087	MGG_01087	Ribosome biogenesis protein TSR1	5	3984	3984	1	2100	10
			MGG_06202	GTP binding protein Bms1	3	3984	3984	1	2100	10
2	MS447	MGG_01205	MGG_01205	tRNA-dihydrouridine synthase 3	2	2874	2874	0	2100	2
			MGG_08825	tRNA-dihydrouridine synthase 1	2	2317	2317	0	2150	1-3
			MGG_01648	tRNA-dihydrouridine synthase	2	2091	2091	1	2091	1-2
3	**MS393**	**MGG_01488**	**MGG_01448**	**N2, N2- Dimethlyguanosine tRNA methytransferase**	**2**	**2229**	**900**	**7**	**1300**	**9**
4	**MS550**	**MGG_07543**	**MGG_07543**	**CDP-diacylglycerol-glycerol 3 phosphate – 3 –phoshatidyltransferase**	**3**	**2031**	**2031**	**0**	**1000**	**30-45**
5	**MS501⁺**	**MGG_01449**	**MGG_01449**	**Karyopherin**	**2**	**3307**	**2280**	**12**	**2280**	**12-14**
6	FG909	MGG_07518	MGG_07518	Centromere/ microtubule-binding protein cbf 5	3	3723	3723	1	1560	22-24
7	FG1020	MGG_03255	MGG_03255	Ubiquitin conjugation factor E4	4	4666	4666	1	2500	11-12
8	FG740	MGG_06712	MGG_ 06712	5 Methytetradropteroltrigkutamate – homocysteine-5-methytransferase	1	3299	3299	1	3000	3
9	FG533	MGG_05481	MGG_05481	Elongator complex protein 3	1	2172	2172	1	1200	15-25
10	FG975	MGG_04478	MGG_04478	Fimbrin	1	2682	2682	1	800	5
			MGG_06475	Alpha-actinin	4	3358	3358	0	1200	4
11	FG465	MGG_09222	MGG_16248	30s ribosomal protein S5	1	1354	1354	1	900	4
12	FG649	MGG_01622	MGG_01622	Uroprophyrinogen decarboxylase	2	2277	2277	1	900	7
13	FG1056	MGG_09555	MGG_09555	Double-strand break repair protein mus-23	2	2801	2801	0	2000	2
14	FG1073	MGG_06348	MGG_06348	Adenylyl-Sulfate kinase	4	1118	225	1-2	500	3
			MGG_15027	Sulfate adenylyltransferase	7	2406	2406	1	1000	16
15	MS320	MGG_07317	MGG_07317	Glutamate-cysteine ligase	2	3598	3598	1	1000	14
16	MS587	MGG_02986	MGG_02986	DNA polymerase zeta catalytic subunit	7	5892	840	2	5000	57
			MGG_08071	DNA polymerase delta catalytic subunit	2	3000	3000	0	3000	17
			MGG_06397	DNA polymerase alpha catalytic subunit	4	5302	5302	1	5302	2

*****Size of the variable regions (bp) among *M. oryzae* isolates from the same host; **Size of the variable region (bp) among *M. oryzae* isolates from different hosts. The FUNYBASE genes shown in bold were developed into molecular markers for further screening.

⁺Two variable regions were identified for the development and testing of the PCR amplification and sequencing process of potential markers

3.3.2.3 Identification of novel markers representing highly variable regions of the genome

A large number of variations have been reported to occur in chromosomes 1, 2, 6 and 7 in the genome of *M. oryzae* based on a comparative analysis of isolates B157 and MG01 from rice (Gowda *et al*., 2015). Around 50 loci were randomly selected from each of the chromosomes mentioned above from the *M. oryzae* reference genome (from strain 70-15) deposited in the BROAD institute database. These loci were analysed bioinformatically using the BLAST alignment tools. This led to the identification of a locus MGG_16180.7 that showed 13 SNPs differentiating the two rice blast isolates B157 and MG01. This locus relates to a hypothetical protein located in chromosome 1 with a gene size of 2765 bp in 70-15. The variable region spanning ~1000 to 2000 bases was used to develop the HyP1 marker. Subsequently, intergenic regions between the loci were analysed by aligning some contigs in these chromosomes. This also led to the identification of three potential marker regions (designated as HyP2 - HyP4) that showed high number of SNPs ranging from 16 to 20 among the isolates from the same host.

3.3.3 PCR amplification and sequencing of potential markers

To develop routinely usable genetic markers, initially, the most variable region was identified within each of the 13 potential marker loci identified from the three different bioinformatic approaches. The variable regions of these loci covered both exons and introns. Primers were designed from the upstream and downstream conserved sequences flanking the variable regions (Table 2.2A). The primers designed were used in PCR amplifications alongside some known markers enabling a base line comparison (Table 2.2A).

Fifteen isolates were selected from the historical population for the initial screening by PCR amplification and sequencing and the data was manually edited and aligned. The number of SNPs present and the number of parsimony informative sites (PIS) in the alignment of the 15 isolates were used as the benchmark to identify markers for further analysis (Table 3.3). Four markers including ITS, HIS4, HyP1 and HyP2 were selected and used for the characterisation of the 300 isolates representing historical and contemporary populations.

Table 3.3: DNA sequence variability data for the 13 potential markers following PCR amplification, sequencing and bioinformatics analysis in 15 *M. oryzae* isolates

S/N	Molecular markers	Expected products size	SNPs	PIS**	Indel
1	**ITS***	**600**	**4**	**4**	**0**
2	TEF1-α	400	3	2	0
3	MPG1	450	0	0	0
4	NUT1	350	2	1	1
5	**HIS4***	**400**	**4**	**4**	**0**
6	MS393	722	2	1	0
7	MS501a	650	3	1	0
8	MS501b	745	2	1	0
9	MS550	650	1	0	0
10	**HyP1***	**840**	**81**	**70**	**10**
11	**HyP2***	**550**	**49**	**36**	**2**
12	HyP3	620	16	16	3
13	HyP4	630	3	2	3

***Bold**, Markers selected for further characterisation of the 300 isolates used in the study.

**PIS, Parsimony informative site; a site that contains at least two types of nucleotides, and at least two of them occur with a minimum frequency of two among the samples as defined by Tamura K, Dudley J, Nei M & Kumar S (*2007*) *MEGA4*: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24:1596-1599.

3.3.4 Evaluation of the Genetic diversity and phylogenetic relationships among the 300 *M. oryzae* isolates

Genotypic characterisation of 300 *M. oryzae* isolates associated with finger millet and related weed hosts from key cropping location in Kenya, Uganda, Tanzania and Ethiopia in Eastern Africa was carried out using the four molecular markers. These included two novel molecular genetic markers designated as HyP1 and HyP2 developed in this study, fungal universal DNA barcode marker ITS and a housekeeping gene marker HIS4. Among the 300 isolates, seventy-six represented historic population from Kenya (32 isolates) and Uganda (44 isolates) selected from a previous collection of 328 isolates (Takan *et al*., 2012). Two hundred and twenty-four isolates represented contemporary population from Kenya (45), Uganda (58), Tanzania (58), and Ethiopia (63). Raw sequence data generated for each isolate were checked for accuracy and any ambiguities were resolved using Bioedit and Geneious software. Multiple sequence alignments were generated for each marker using MAFFY and MUSCLE softwares plugin within Geneious, manually checked and/or edited to ensure optimal alignment. Multiple sequence alignments were end trimmed to ensure uniform start and end DNA bases for each isolate. Single locus (e.g. ITS) and multi-locus (concatenated ITS, HIS4, HyP1 and HyP2) multiple sequence alignments were used in the identification of genotypes (based on sequence identity) and to decipher phylogenetic relationships of *M. oryzae* isolates within a country and across the four countries surveyed in Eastern Africa.

3.3.4.1 *M. oryzae* genotype diversity, distribution and relationships based on the ITS marker

The end trimmed sequence of the ITS genetic marker ranged from 499 to 508 bases in length. The multiple sequence alignment of the data from the 300 isolates aligned to a total length of 513 sites displaying 22 SNPs, 10 indels. Among the SNPs identified, 22 were parsimony informative sites. Seven genotypes (ITS-G1 to ITS-G7) were identified based on the nucleotide diversity of the ITS region in *M. oryzae* isolates from finger millet and nine other related weed hosts (including the wild millet *E. indica*) from different districts of Kenya, Uganda, Tanzania and Ethiopia (Table 3.4). All the genotypes identified represented more than one isolate. The overall range per genotype varied from 4 to 160 isolates that revealed identical ITS sequence among the FMB pathogen populations from Eastern Africa. Furthermore, some genotypes represented isolates from finger millet as well as related weed hosts e.g. isolates K8/40 and K13 in ITS-G1 were from wild millet *E. indica* (Table 3.4). Two major genotypes ITS-G1 and ITS-G2 represent 160 and 99 isolates, respectively and the other genotypes ITS-G3 to ITS-G7 represent from 4 to 17 isolates (Table 3.4 and Figure 3.4). Genotype ITS-G1 represents *M. oryzae* isolates from 47 districts across the four countries in Eastern Africa (Figure 3.3). This genotype also comprises isolates representing historical and contemporary populations, even from the same districts, for example Teso, Busia and Kisii in Kenya; Pallisa, Arua and Tororo in Uganda. This genotype is more dominant in Kenya (69 isolates) and Uganda (78) compared to Ethiopia (7) and Tanzania (6) as shown in Figure 3.3. Within Kenya, this genotype includes historic (27 isolates) and contemporary (42 isolates) populations and a similar pattern is evident in Uganda with 37 and 41 isolates representing historic and contemporary populations, respectively (Table 3.4).

Genotype ITS-G2 contains 99 isolates present in all four countries but clearly more dominant in Ethiopia (54 isolates) and Tanzania (40) with very few isolates in Kenya (4) and Uganda (1) as shown in Table 3.4. The limited number of isolates present in Kenya and Uganda are from the historical population. Genotype ITS-G3 contains 17 isolates from Uganda (13 isolates including historic and contemporary samples), Kenya (3) and Ethiopia (1). Genotype ITS-G4 is restricted to Tanzania (8 isolates) and Ethiopia (1), whilst Genotype ITS-G6 is restricted to Uganda (3 isolates) and Kenya (1) essentially including 4 historical isolates. Genotypes ITS-G5 (7 isolates) and ITS-G7 (4 isolates) are restricted to Uganda and Tanzania, respectively and contain only contemporary isolates from different districts.

Bayesian phylogenetic tree was constructed from the multiple sequence alignment representing the 7 genotypes (ITS-G1 – ITS-G7) and an isolate from rice used as an outgroup, with 70 % posterior probability value (PPV). The Kimura 2-parameter (K2) model using a discrete gamma distribution (G) was determined to be the best evolutionary model (K2+G) for the ITS sequence dataset. The tree clearly distinguished the Eastern African FMB isolates from the rice isolate and showed a strong relationship between the Genotypes ITS-G2 and ITS-G4 at 95.8 % PPV; whilst other genotypes did not show a clear phylogenetic relationship (Figure 3.4).

Table 3.4. Diversity and distribution pattern of the *M. oryzae* genotypes identified among the finger millet blast pathogen populations in Kenya, Uganda, Tanzania and Ethiopia based on the ITS marker sequences

Genotype	No. of isolates in genotype	Representative isolate	No. of countries present	No. of isolates per country	Isolate code	No. of district	Population type
ITS-G1	160	D1/s19	4	7 (E)	**E1, E3, E18, E32, E46, E61, E62**	6	C
				69 (K)	**K1, K2, K3, K4, K6, K7, K8, K9, K10, K11, K12, K13, K14, K15, K16, K17, K18, K19, K20, K21, K22, K23, K24, K25, K26, K27, K28, K29, K31, K30, K33, K32, K34, K35, K36, K37, K38, K39, K40, K42, K44, K45**, K1/15, K4/21p, K5/23, K5/24w, K8/40, K9/46, K12/62, K13/67, K14/74, K15/53n, K21/68n, K22/118, K23/123, K26/76p, K28/82w, K29/164, K36/98n, K45/112n, K47/114p, K48/115n, K55/124p, K57/126p, K60/131p, K64/137p, K65/140n, K65/142n, K65/159w	13	H & C
				6 (T)	**T11, T14, T20, T39, T54, T58**	3	H
				78 (U)	**U2, U4, U6, U7, U8, U9, U10, U11, U12, U16, U17, U18, U20, U21, U22, U23, U24, U25, U26, U27, U28, U32, U33, U34, U36, U37, U38, U39, U40, U41, U44, U45, U47, U48, U49, U51, U52, U57, U55, U56, U58**, D1/s19, D1/s44, D1/s53b, D1/s72, D2/s14, D2/s26, D3/s3, D3/s9, D3/s24, D4/s12, D4/s26, D4/s41, D5/s1, D6/s1, D7/s6, D8/s15, D9/s50, D9/s54, D9/s76, D10/s63, D10/s71, D10/s77, D12/s2, D13/s5, D14/s27, D14/s30, D15/s6, D15/s12, D15/s41, D15/s47, E11p-1-1*, Gup-2-1*, Odyp-2-1*, Pen-2-2*, P665n-2-1*, S1p-1-1*, Secn-2-2*	25	H & C
ITS-G2	99	D11/s16	4	54 (E)	**E2, E4, E5, E6, E7, E8, E9, E10, E11, E12, E13, E14, E15, E16, E17, E19, E20, E21, E22, E23, E24, E25,**	19	C

					E26, E27, E28, E29, E30, E31, E33, E34, E35, E36, E37, E38, E39, E40, E41, E42, E44, E45, E47, E48, E49, E50, E51, E52, E53, E54, E56, E58, E57, E59, E60, E63		
				4 (K)	K24/127, K33/184, K33/189, K44/111p	3	H
				40 (T)	**T1, T2, T3, T4, T5, T6, T7, T8, T9, T10, T12, T13, T15, T16, T18, T19, T21, T22, T23, T24, T25, T26, T27, T28, T29, T32, T41, T42, T43, T44, T45, T46, T47, T48, T49, T50, T51, T52, T56, T57**	6	C
				1 (U)	D11/s16	1	H
ITS-G3	17	D1/s11	3	1 (E)	**E43**	1	C
				3 (K)	**K5, K43, K41**	2	C
				13 (U)	**U1, U3, U5, U14, U15, U29, U30, U43, U46, U54,** D1/s11, D9/s56, D9/s70	8	H & C
ITS-G4	9	E55	2	1(E)	**E55**	1	C
				8 (T	**T33, T34, T36, T37, T38, T30, T31, T55**	4	C
ITS-G5	7	U13	1	7 (U)	**U13, U31, U35, U50, U53, U19, U42**	6	C
ITS-G6	4	D1/s50	2	1 (K)	K58/128p	1	H
				3(U)	D1/s50, D15/s37, D10/s73	3	H
ITS-G7	4	T17	1	4 (T)	**T17, T35, T40, T53**	3	C

H, historical isolates (2000-2004) and C, contemporary isolates (2015-2017); In the Isolate code, K-isolates collected from Kenya E-isolates collected from Ethiopia, T-isolates collected from Tanzania and U and D-isolates collected from Uganda; *Isolates from Uganda;

Isolates in bold and normal font are contemporary and historic collections, respectively. Further details of the isolates are available in **Table 2.1A.**

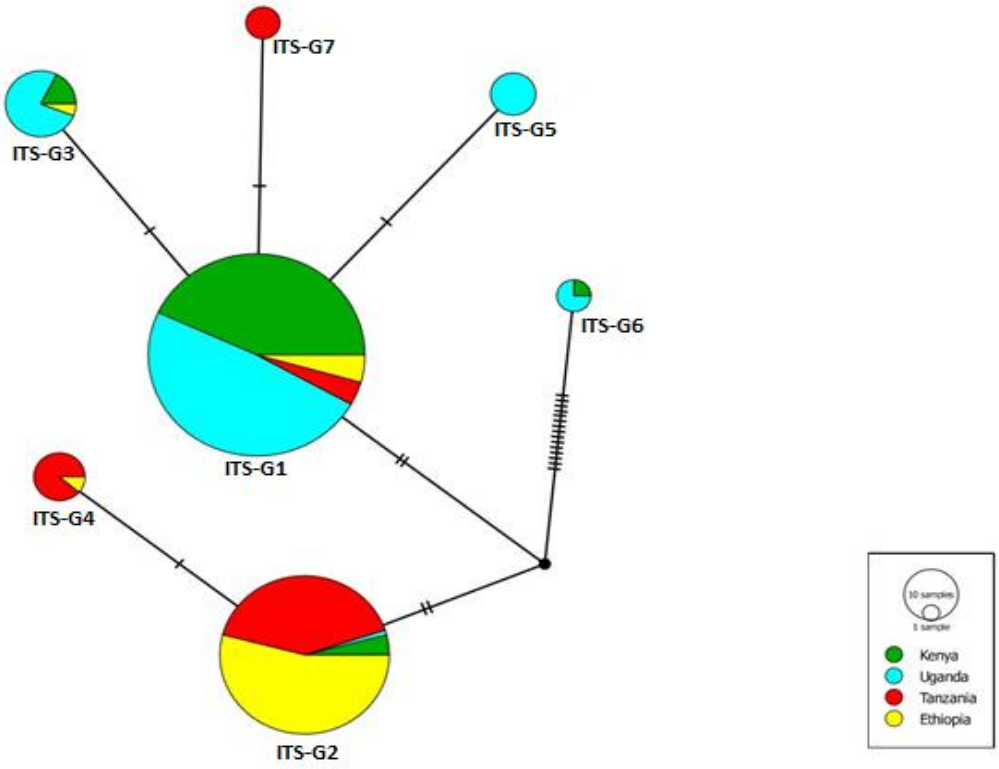

Figure 3.3: Reticulate median-joining network of 300 isolates representing the seven genotypes based on ITS marker sequences

Overall circle sizes are proportional to the number of isolates within a specific genotype. Circle slice area is proportional to the number of isolates from a country, where colours indicate the country of collection of isolates with Kenya in green, Uganda in blue, Tanzania in red and Ethiopia in yellow. The segment/hatch marks seen in the connecting lines reflect the mutation levels among the genotypes.

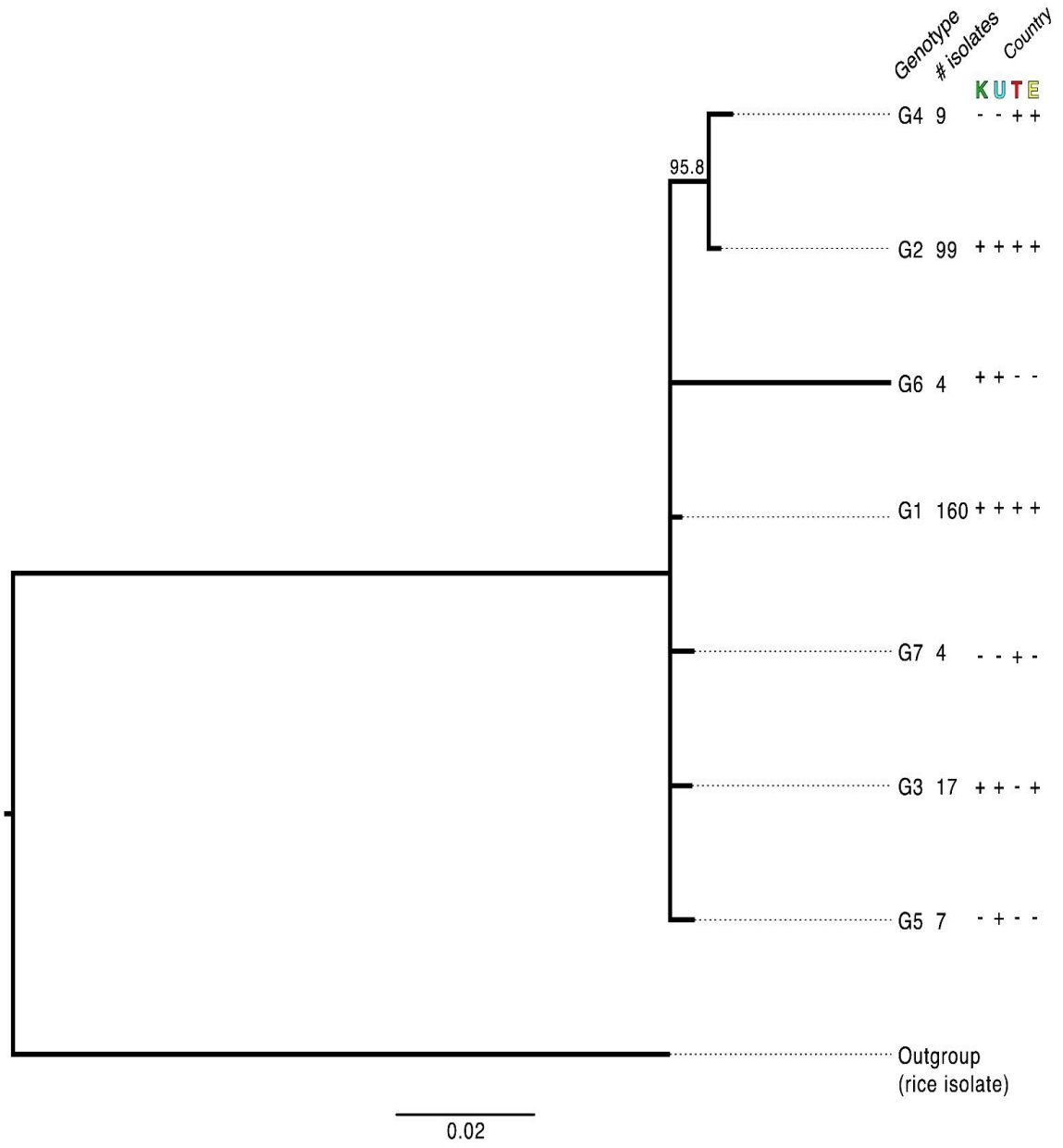

Figure 3.4: Bayesian consensus tree of the *M. oryzae* genotypes based on the ITS marker sequences

The tree shows the number of isolates per genotype and their distribution pattern across the four countries: K- Kenya, U – Uganda, T- Tanzania and E- Ethiopia; + indicates present and – indicates absent based on the isolates characterised. The number over the branch of the tree represents the posterior probability value.

3.3.4.2 *M. oryzae* genotype diversity, distribution and relationships based on the Histone 4 (HIS4) marker

The end trimmed sequence of the HIS4 genetic marker ranged from 378 to 380 bases in length. The multiple sequence alignment of the data from the 300 isolates aligned to a total length of 400 sites displaying 44 SNPs and 22 indels. Among the SNPs identified, 35 were parsimony informative sites. Nine genotypes (HIS-G1 to HIS-G9) were identified based on the nucleotide diversity of the HIS4 locus in *M. oryzae* isolates from finger millet and related weed hosts from different districts of Kenya, Uganda, Tanzania and Ethiopia (Table 3.5). Three major genotypes HIS-G1, HIS-G2 and HIS-G3 comprised 110, 103 and 58 isolates, respectively (Table 3.5 and Figure 3.5). Overall, these 3 genotypes represented 271 isolates (~90 % population), whilst other genotypes (HIS-G4 to HIS-G9) each comprised 2 to 7 isolates representing a total of 29 isolates (Table 3.5). Furthermore, some genotypes included isolates from finger millet and weed hosts (for example, isolates K8/40 and D10/s73 in HIS-G3 and isolate K13 in HIS-G5).

Genotype HIS-G1 represented isolates from all four countries, but more dominant in Kenya and Uganda containing 40 and 62 isolates, respectively with limited numbers from Ethiopia (6 isolates) and Tanzania (2 isolates). HIS-G1 isolates found in Kenya and Uganda comprised historic and contemporary populations at nearly equal level (Table 3.5). Genotype HIS-G2 comprised 103 isolates distributed across Ethiopia, Tanzania and Kenya (Figure 3.5) but clearly are more dominant in Ethiopia (50 isolates) and Tanzania (48 isolates) relative to Kenya (5 isolates) as presented below (Table 3.5 and Figure 3.5). These 5 isolates (K24/127, K33/184, K33/189, K44/111p, K48/115n) from Kenya are from historical population and collected from 3 districts (Kericho, Kisii and Kisii central). HIS-G3 mainly comprised isolates

from Kenya (21) and Uganda (31) with only 6 isolates from Tanzania. In Kenya and Uganda, isolates found in HIS-G3 represented both historical and contemporary populations. Genotypes HIS-G4, HIS-G7 and HIS-G9 contained 7, 4 and 2 isolates from Ethiopia, Kenya and Tanzania, respectively but from different districts. HIS-G5, HIS-G6 and HIS-G8 present in Kenya and Uganda only contained limited number of isolates ranging from 2 to 7 (Table 3.5).

Bayesian phylogenetic tree was constructed from the multiple sequence alignment of the 9 genotypes (Table 3.5) and a rice blast (RB) isolate used as an outgroup, with 70 % posterior probability value (Figure 3.6). The Kimura 2-parameter (K2) model using a discrete gamma distribution (G) was determined to be the best evolutionary model (K2+G) for the HIS4 sequences dataset. The phylogenetic analysis clearly distinguished the FMB pathogen isolates from the RB pathogen isolate (Figure 3.6). Eight of the FMB genotypes showed a strong phylogenetic relationship defined based on 86.9 % PPV. These genotypes comprised isolates from Ethiopia, Kenya, Tanzania and Uganda. However, HIS-G8 containing only 2 isolates from Kenya and Uganda appears phylogenetically distinct from the other genotypes (Figure 3.6).

Table 3.5. Diversity and distribution pattern of the *M. oryzae* genotypes identified among the finger millet blast pathogen populations in Kenya, Uganda, Tanzania and Ethiopia based on the HIS4 marker sequences

Genotype	No. of isolates in genotype	Representative isolate	No. of count-ries present	No. of isolates per country	Isolate code	No. of district	Population type
HIS-G1	160	D1/s19	4	6 (E)	**E1, E3, E46, E55, E61, E62**	6	C
				40 (K)	**K4, K6, K10, K12, K14, K15, K20, K21, K24, K25, K28, K29, K30, K33, K34, K39, K41, K43, K44, K45,** K1/15, K4/21p, K5/23, K5/24w, K12/62, K13/67, K14/74, K15/53n, K21/68n, K22/118, K23/123, K26/76p, K28/82w, K29/164, K36/98n, K45/112n, K47/114p, K60/131p, K64/137p, K65/140n	12	H & C
				2 (T)	**T39, T40**	1	C
				62 (U)	**U3, U5, U6, U8, U13, U16, U17, U18, U19, U20, U21, U25, U26, U28, U31, U32, U33, U34, U35, U36, U39, U41, U44, U45, U49, U50, U51, U52, U53, U55, U56, U58,** D1/s11, D1/s19, D1/s44, D1/s50, D1/s53b, D1/s72, D2/s14, D2/s26, D3/s9, D3/s24, D4/s12, D4/s26, D4/s41, D5/s1, D6/s1, D7/s6, D8/s15, D9/s76, D10/s71, D10/s77, D14/s30, D15/s6, D15/s12, D15/s41, D15/s47, E11p-1-1*, Gup-2-1*, Odyp-2-1*, S1p-1-1*, Secn-2-2*	24	H & C

HIS-G2	103	E2	3	50 (E)	**E2, E10, E11, E13, E14, E15, E16, E17, E18, E19, E20, E21, E22, E23, E24, E25, E26, E27, E28, E29, E30, E31, E32, E33, E34, E35, E36, E37, E38, E39, E40, E41, E42, E43, E44, E45, E47, E48, E49, E50, E51, E52, E53, E54, E56, E57, E58, E59, E60, E63**	20	C
				5 (K)	K24/127, K33/184, K33/189, K44/111p, K48/115n	3	H
				48 (T)	**T1, T2, T3, T4, T5, T6, T7, T8, T9, T10, T12, T13, T14, T15, T16, T18, T19, T21, T22, T23, T24, T25, T26, T27, T28, T29, T30, T31, T32, T33, T36, T37, T38, T41, T42, T43, T44, T45, T46, T47, T48, T49, T50, T51, T52, T55, T56, T57**	7	C
HIS-G3	58	D3/s3	3	21 (K)	**K1, K2, K5, K9, K11, K19, K22, K23, K27, K31, K32, K35, K36, K40, K42,** K8/40, K9/46, K55/124p, K57/126p, K65/142n, K65/159w	9	H & C
				6 (T)	**T11, T17, T20, T53, T54, T58**	2	C
				31 (U)	**U1, U2, U11, U12, U14, U15, U22, U23, U24, U27, U30, U37, U40, U43, U46, U47, U48, U54, U57**, D3/s3, D9/s50, D9/s54, D9/s56, D9/s70, D10/s63, D10/s73, D12/s2, D13/s5, D14/s27, P665n-2-1*, Pen2-2*	17	H & C
HIS-G4	7	E4	1	7 (E)	**E4, E5, E6, E7, E8, E9, E12**	3	C
HIS-G5	7	K3	2	5 (K)	**K3, K8, K13, K37, K38**	3	C
				2 (U)	**U7, U9**	2	C
HIS-G6	7	D11/s16	2	1 (K)	**K26**	1	C
				6 (U)	D11/s16, **U4, U10, U29, U38, U42**	6	H &C
HIS-G7	4	K7	1	4 (K)	**K7, K16, K17, K18**	4	C

HIS-G8	2	D15/s37	2	1 (K)	K58/128p	1	H
				1 (D)	D15/s37	1	H
HIS-G9	2	T34	1	2 (T)	**T34, T35**	2	C

H, historical isolates (2000-2004) and C, contemporary isolates (2015-2017);

In the Isolate code, K-isolates collected from Kenya E-isolates collected from Ethiopia, T-isolates collected from Tanzania and U and D-isolates collected from Uganda; *Isolates from Uganda;

Isolates in bold and normal font are contemporary and historic collections, respectively. Further details of the isolates are available in **Table 2.1A.**

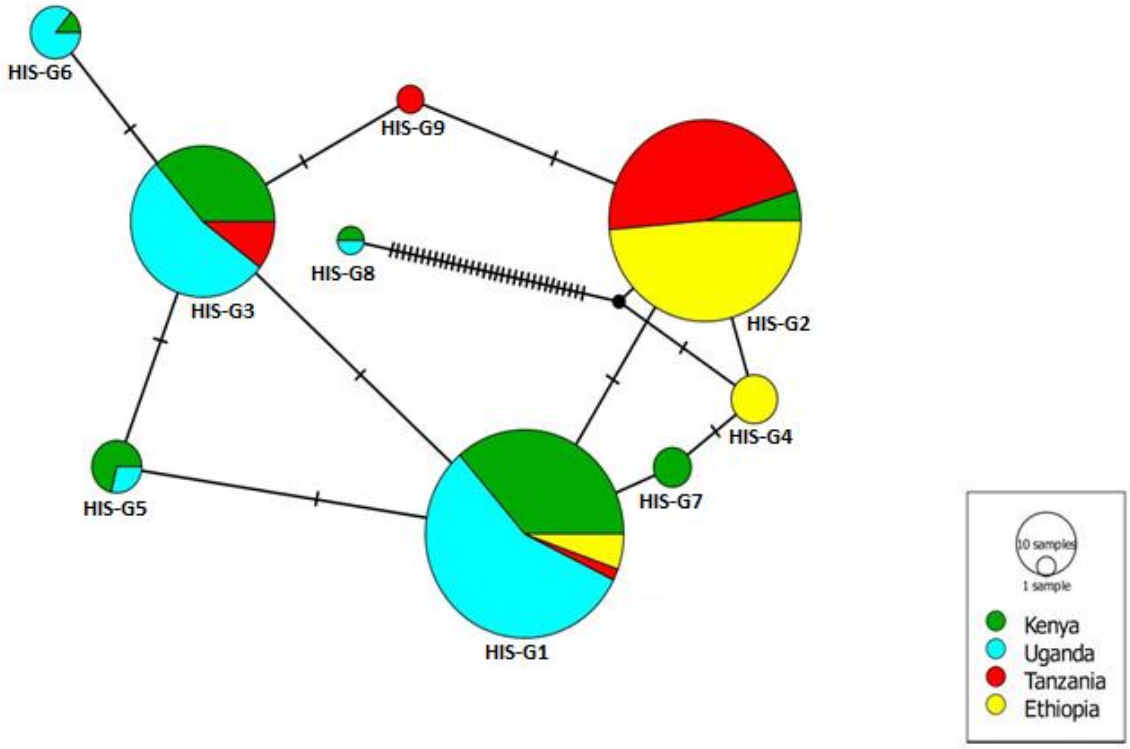

Figure 3.5: Reticulate median-joining network of 300 isolates representing the nine genotypes (HIS-G1 to HIS-G9) based on HIS4 marker sequences

Overall circle sizes are proportional to the number of isolates within a specific genotype. Circle slice area is proportional to the number of isolates from a country, where colours indicate the country of collection of isolates with Kenya in green, Uganda in blue, Tanzania in red and Ethiopia in yellow. The segment/hatch marks seen in the connecting lines reflect the mutation levels among the genotypes.

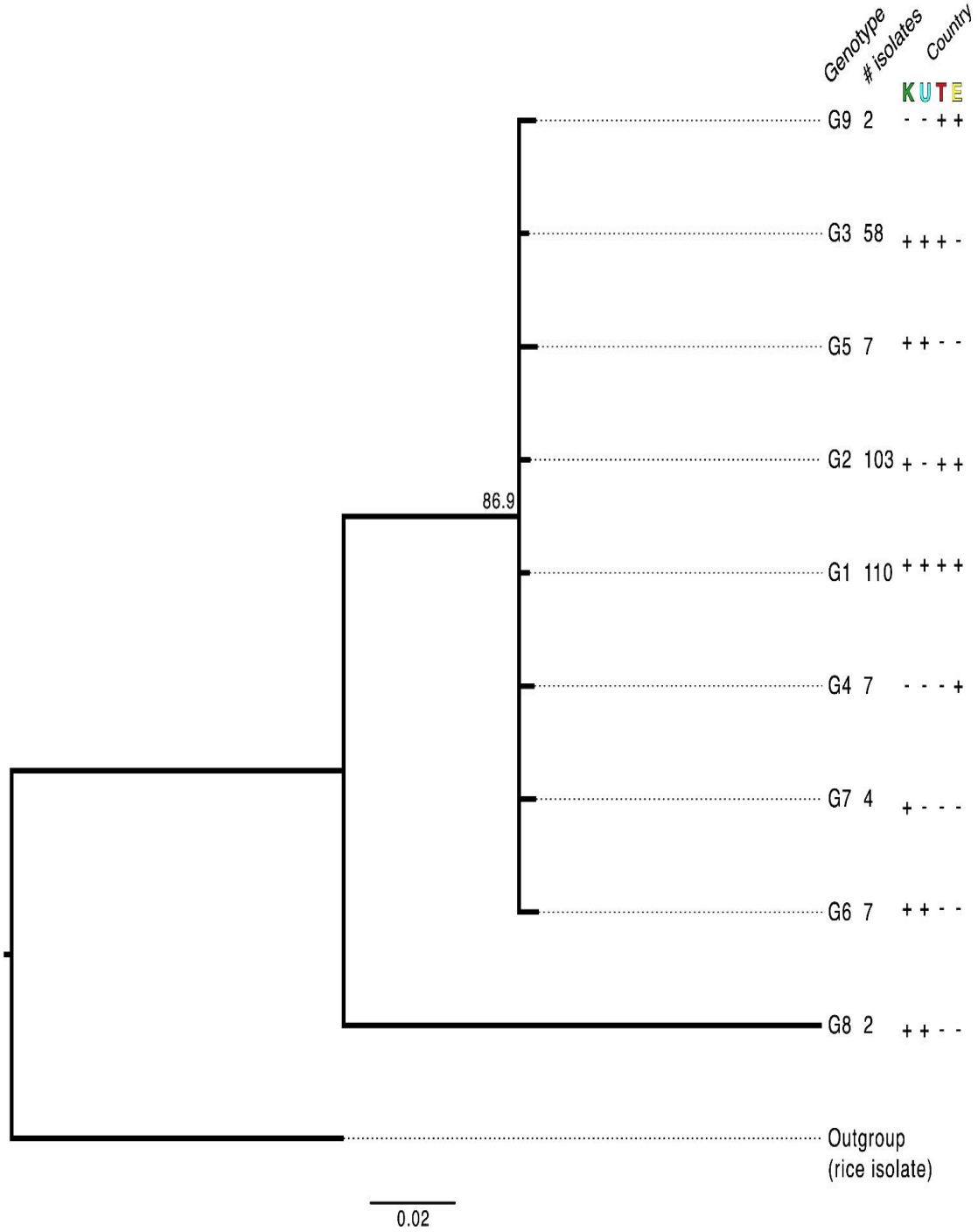

Figure 3.6: Bayesian consensus tree of the *M. oryzae* genotypes based on the HIS4 (Histone 4) marker sequences

The tree shows the number of isolates per genotype and their distribution pattern across the four countries: K- Kenya, U – Uganda, T- Tanzania and E- Ethiopia; + indicates present and – indicates absent based on the isolates characterised. The number over the branch of the tree represents the posterior probability value.

3.3.4.3 *M. oryzae* genotype diversity, distribution and relationships based on the HyP1 marker

The end trimmed sequence of the HyP1 genetic marker was ~743 - 799 bases in length. The multiple sequence alignment of the data from the 300 isolates aligned to a total length of 874 bases showed 394 SNPs and 185 indels. Among the SNPs identified, 251 were parsimony informative sites. Eighty-five (85) genotypes were identified based on the nucleotide diversity of the HyP1 locus in *M. oryzae* isolates from finger millet and few related weed hosts from different districts of Kenya, Uganda, Tanzania and Ethiopia. Of these, 22 genotypes represented 237 isolates, with 2 to 73 isolates per genotype, which revealed identical nucleotide sequence (Table 3.6A) and these 22 are referred as shared genotypes. The other 63 genotypes were represented by single isolates (Table 3.6B, further details are available in Appendix 1) and these 63 are referred as single genotypes. Tables 3.6A and B show the distribution pattern of the *M. oryzae* genotypes with reference to various geographic locations, hosts and period of collection. Figure 3.7 based on the reticulate network analysis shows the distribution pattern of the shared genotypes across the 4 countries distinguishing them into two major groups A and B. Group A predominantly included isolates from Kenya and Uganda, whilst Group B predominantly included isolates from Ethiopia and Tanzania. Three major shared genotypes HyP1-G1, HyP1-G2 and HyP1-G3 were identified representing 61 % of the total isolates collected from the 4 countries. HyP1-G1 contained 73 isolates with at least one isolate from each country. Genotype HyP1-G1 isolates are dominant (86 %) in Ethiopia (25) and Tanzania (36), while the remaining 14 % were present in Kenya (10 isolates) and Uganda (2) as shown in Table 3.6A and Figure 3.8. Genotype HyP1-G2 comprised 72 isolates from Kenya, Uganda and Ethiopia (Table 3.6A) and is more dominant in Kenya (25) and Uganda (40) compared to Ethiopia (7). HyP1-G3

contained 38 isolates including 28 from different districts in Uganda, 9 from 3 districts in Kenya and 1 isolate from Tanzania. Genotypes HyP1-G4 to HyP1-22 included 2 or more isolates either from one or two countries. For example, genotypes 4 and 5 were found only in Ethiopia and Kenya, respectively, while genotype 6 represented a limited number of isolates from Kenya and Uganda (Table 3.6A and Figure 3.8).

Bayesian phylogenetic tree was constructed from the multiple sequence alignment of the 85 finger millet blast (FMB) genotypes (further details are available in Appendix 1) using a rice blast isolate as an outgroup, with 70 % PPV (Figure 3.8). The Hasegawa-Kishino-Yano (HKY) model using a discrete gamma distribution (G) was determined to be the best evolutionary model (HKY+G) for the HyP1 sequence dataset. Phylogenetic analysis distinguished the FMB pathogen genotypes from the outgroup RB isolate. Based on the tree topology, the *M. oryzae* genotypes were divided into 2 major groups A and B. Group A comprised 53 genotypes at 99.9 % PPV including 2 major genotypes HyP-G2 and HyP-G3. Further, Group A comprised genotypes representing isolates predominantly from Kenya and Uganda including both historical and contemporary isolates, whilst in Ethiopia and Tanzania, these genotypes were less common. For example, among the 53 genotypes in the Group A representing 183 isolates, 163 isolates were from Kenya and Uganda and only 20 isolates were from Ethiopia and Tanzania (Figure 3.8). Group B comprised 32 genotypes including one major genotype HyP-G1. Group B contained genotypes that are mostly dominant in Ethiopia and Tanzania and less common in Kenya and Uganda. For example, the 32 genotypes in Group B represented 117 isolates including 102 from Ethiopia and Tanzania and 15 from Kenya and Uganda (Figure 3.8). In addition, within the Groups A and B, some genotypes showed close

phylogenetic relationship based on high PPV, e.g. G5, G8, G10, G63, G64, G65 and G66 in Group A with 99.9 % PPV (Figure 3.8).

Table 3.6A. Diversity and distribution pattern of shared genotypes of *M. oryzae* identified among the finger millet blast pathogen populations in Kenya, Uganda, Tanzania and Ethiopia based on the HyP1 marker sequences

Genotype	No. of isolates in genotype	Representative isolate	No. of countries present	No. of isolates per country	Isolate code	No. of district	Population type
HyP1-G1	73	T56	4	25 (E)	**E2, E15, E19, E22, E26, E34, E35, E37, E38, E39, E41, E42, E44, E45, E47, E48, E49, E53, E54, E56, E57, E58, E59, E60, E63**	14	C
				10 (K)	**K14, K21, K41,** K24/127, K33/184, K33/189, K44/111p, K48/115n, K58/128p, K60/131p	6	H & C
				36 (T)	**T1, T2, T3, T6, T7, T8, T9, T10, T13, T18, T19, T22, T24, T26, T27, T28, T30, T31, T32, T33, T34, T36, T38, T41, T42, T43, T44, T45, T46, T47, T48, T49, T50, T51, T52, T56**	6	C
				2 (U)	D3/s9, D10/s73	2	H
HyP1-G2	72	U57	3	7 (E)	**E18, E27, E43, E46, E55, E61, E62**	6	C
				25 (K)	**K1, K4, K5, K7, K8, K13, K15, K22, K27, K42, K43**, K4/21p, K5/24w, K12/62, K21/68n, K22/118, K23/123, K26/76p, K28/82w, K36/98n, K55/124p, K5/23, K15/53n, K9/46, K65/159w	7	H & C
				40 (U)	**U1, U2, U4, U5, U7, U8, U20, U21, U24, U26, U32, U35, U36, U37, U39, U44, U55, U56, U57**, D2/s14, D1/s11, D3/s3, D3/s24, D5/s1, D7/s6, D9/s76, D14/s27, Pen-2-2, D1/s72, D4/s12, D9/s50, D9/s56, D9/s70, D10/s63, D10/s77, D13/s5, D14/s30, D15/s12, D15/s41, D4/s41	20	H & C
HyP1-G3	38	U58	3	9 (K)	**K2, K3, K10, K11, K12, K24, K26, K33, K35**	4	C
				1 (T)	**T35**	1	C
				28 (U)	**U3, U6, U9, U10, U11, U12, U13, U14, U17,**	17	H & C

Genotype	No. of isolates in genotype	Representative isolate	No. of countries present	No. of isolates per country	Isolate code	No. of district	Population type
					U18, U19, U22, U23, U25, U29, U30, U31, U33, U38, U42, U46, U49, U50, U51, U53, U54, U58, D15/s6		
HyP1-G4	8	E33	1	8 (E)	**E21, E23, E24, E25, E28, E29, E31, E33**	5	C
HyP1-G5	5	K57/126p	1	5 (K)	**K34, K36, K39,** K45/112n, K57/126p	2	H & C
HyP1-G6	4	Secn-2-2	2	2 (K)	K13/67, **K28**	2	H & C
				2 (U)	Secn-2-2*, D1/s19	2	H
HyP1-G7	4	P665n-2-1	2	1 (K)	K47/114p	1	H
				3 (U)	E11p-1-1*, Odyp-2-1*, P665n-2-1*	1	H
HyP1-G8	3	T54	1	3 (T)	**T20, T53, T54**	1	C
HyP1-G9	3	U48	2	1 (K)	**K23**	1	C
				2 (U)	**U27, U48**	2	C
HyP1-G10	3	K40	1	3 (K)	**K37, K38, K40**	1	C
HyP1-G11	2	E11	1	2 (E)	**E7, E11**	2	C
HyP1-G12	2	E10	1	2 (E)	**E8, E10**	2	C
HyP1-G13	2	E32	1	2 (E)	**E9, E32**	2	C
HyP1-G14	2	K32	1	2 (K)	**K31, K32**	2	C
HyP1-G15	2	E3	1	2 (E)	**E1, E3**	2	C
HyP1-G16	2	U45	2	1 (K)	**K30**	1	C
				1 (U)	**U45**	1	C
HyP1-G17	2	T40	1	2 (T)	**T39, T40**	1	C
HyP1-G18	2	U52	1	2 (U)	**U41, U52**	2	C
HyP1-G19	2	K65/140n	2	1 (K)	K65/140n	1	H
				1 (U)	D2/s26	1	H
HyP1-G20	2	E51	1	2 (E)	**E50, E51**	2	C
HyP1-G21	2	D15/s37	1	2 (U)	D1/s50, D15/s37	2	H
HyP1-G22	2	T12	1	2 (T)	**T12, T16**	2	C

H, historical isolates (2000-2004) and C, contemporary isolates (2015-2017);

In the Isolate code, K-isolates collected from Kenya E-isolates collected from Ethiopia, T-isolates collected from Tanzania and U and D-isolates collected from Uganda; *Isolates from Uganda;

Isolates in bold and normal font are contemporary and historic collections, respectively. Further details of the isolates are available in **Table 2.1A.**

Table 3.6B. Diversity and distribution pattern of single genotypes[a] of *M. oryzae* identified based on the HyP1 sequence data among the isolates associated with finger millet production in various districts in Ethiopia, Kenya, Tanzania and Uganda

Country	District	Number of Genotypes/Isolates	Finger millet	Other weed hosts[b]	Population[c]
Ethiopia	Diga	4	4	0	C
	Lallo Assabi	1	1	0	C
	Banja	1	1	0	C
	Bahir Dar Zuria	2	2	0	C
	Demecha	1	1	0	C
	Angebo	1	1	0	C
	Leka Dulecha	1	1	0	C
	Dangla	1	1	0	C
	Mecha	1	1	0	C
	Total	***13***	***13***	***0***	**C**
Kenya	Siaya	1	1	0	H
	Homabay	1	1	0	H
	Suba	1	1	0	H
	Kisumu	1	1	0	H
	Busia	2	2	0	C
	Kisii	3	3	0	C
	Kakamega	1	1	0	C
	Eldoret	3	3	0	C
	Teso	3	2	1	H
	Total	***16***	***15***	***1***	**H & C**

Tanzania	Nkasi	1	1	0	C
	Sumbawanga	1	1	0	C
	Mbozi	3	3	0	C
	Njombe	3	3	0	C
	Madaba	4	4	0	C
	Momba	2	2	0	C
	Total	***14***	***14***	*0*	**C**
Uganda	Apac	1	0	1	H
	Amuria	1	1	0	C
	Agago	1	1	0	C
	Busia	2	1	1	H
	Bugiri	1	1	0	H
	Lira	2	1	1	H & C
	Masindi	2	2	0	H & C
	Kumi	1	1	0	C
	Pallisa	2	2	0	H
	Nakasongola	1	1	0	H
	Hoima	1	1	0	C
	Soroti	3	3	0	H
	Tororo	2	2	0	H & C
	Total	***20***	***17***	***3***	**H & C**
Total	**37**	**63**	**59**	**4**	**NA**

[a] Single genotype are individual isolates with distinct HyP1 marker sequence;

[b] Isolates from other weed hosts include *Eleusine indica*, *E. africana*, *Digitaria scalarum*, *D. aegytium*, *D. horizontalis*;

[c] H - historical isolates (2000-2004), C - contemporary isolates (2015-2017);

NA - Not applicable

Further details of isolates belonging to these genotypes are available in Appendix 1.

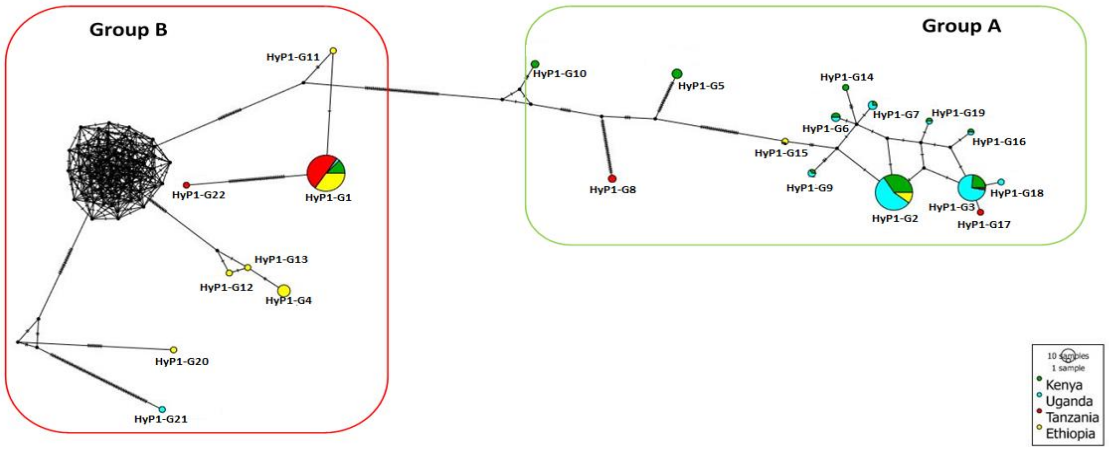

Figure 3.7: Reticulate median-joining network of 237 isolates represented by 22 shared genotypes based on HyP1 marker sequences (Table 3.6A)

Detail of isolates represented by the 22 genotypes are avilable in Table 3.6A with further detilas in Appendix 1. Overall circle sizes are proportional to the number of isolates within a specific genotype. Circle slice area is proportional to the number of isolates from a country, where colours indicate the country of collection of isolates with Kenya in green, Uganda in blue, Tanzania in red and Ethiopia in yellow. The segment/hatch marks seen in the connecting lines reflect the mutation levels among the genotypes. The reticulate analysis resulted in 2 groups with **A** containing the majority of isolates from Kenya and Uganda, and **B** containing the majority of isolates from Ethiopia and Tanzania.

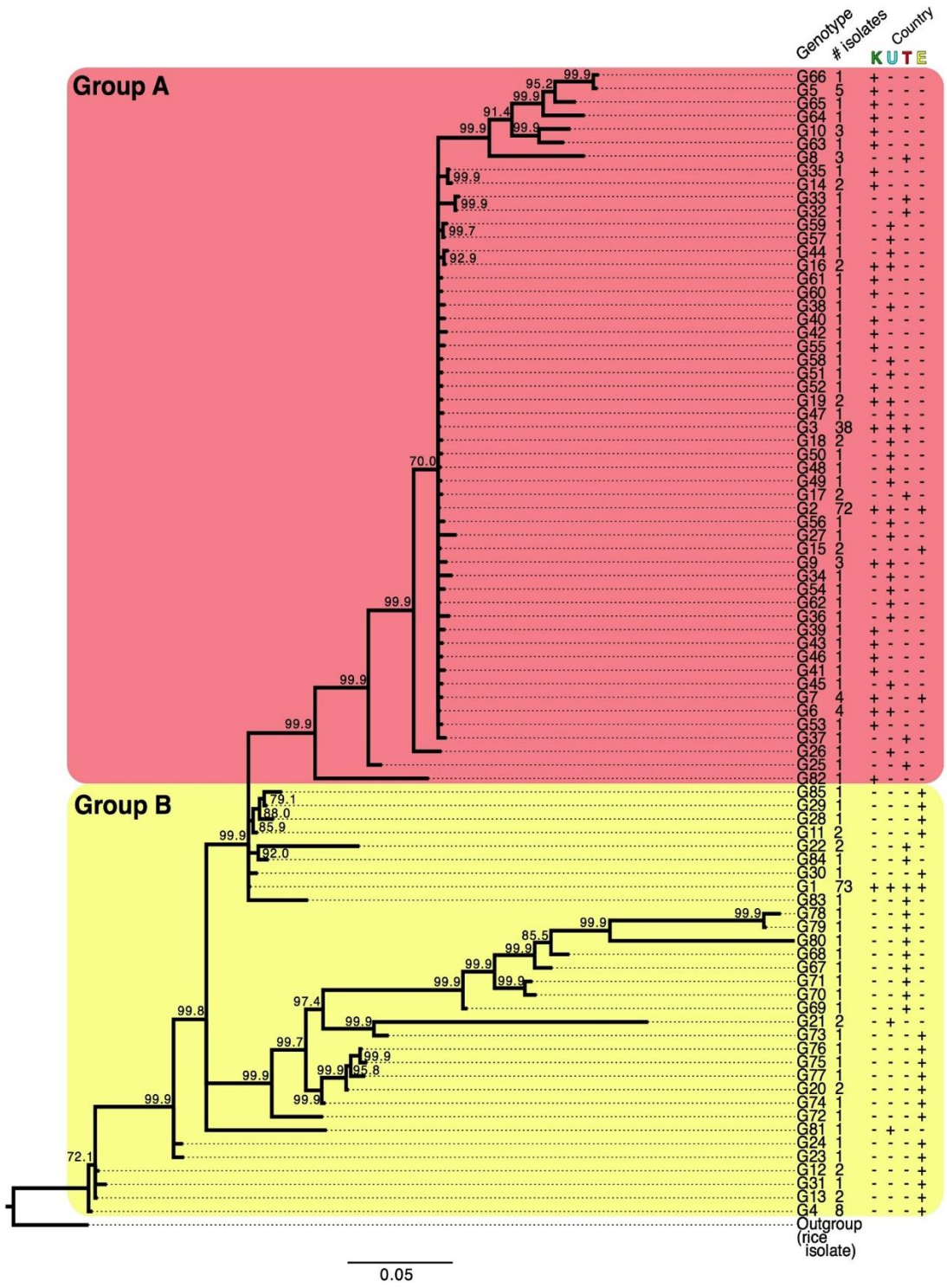

Figure 3.8: Bayesian consensus tree of the *M. oryzae* genotypes based on the HyP1 marker sequences

The tree shows the number of isolates per genotype and their distribution pattern across the four countries: K- Kenya, U – Uganda, T- Tanzania and E- Ethiopia; + indicates present and – indicates absent based on the isolates characterised. The number over the branch of the tree represents the posterior probability value.

3.3.4.4 *M. oryzae* genotype diversity, distribution and relationships based on the HyP2 marker

The end trimmed sequence of the HyP2 genetic marker was ~507 - 508 bases in length. The multiple sequence alignment of the data from the 300 isolates aligned to a total length of 508 bases and showed 133 SNPs and 5 indels. Among the SNPs identified, 77 were parsimony informative sites. The alignment distinguished the population of 300 *M. oryzae* isolates into 80 genotypes. Of these, 22 were shared genotypes representing 234 isolates with 2 to 48 isolates per genotype that revealed identical nucleotide sequence (Table 3.7A). The other 58 genotypes were represented by single isolates from 4 the countries (Table 3.7B). Tables 3.7A and B show the distribution pattern of the *M. oryzae* genotypes with respect to various geographical locations, hosts and period of collection. Five shared genotypes were identified as major genotypes (HyP2-G1 to HyP2-G5) containing nearly 50 % (148) of the total isolates. Genotypes 6 to 10 each represented 11 to 14 isolates and genotypes 11 to 22 each represented 2 to 5 isolates. Further, isolates originating from weed hosts were represented by the genotypes sharing the HyP2 marker sequence with FMB pathogen isolates, e.g. D9/s50 and D9/s54 in HyP2-G2 and K9/46 in HyP2-G5 (Table 3.7A). The rest of the genotypes HyP2-G23 to HyP2-G80 were represented by single isolates and five of these were represented by isolates originating from weed hosts (Table 3.7B, further details are available in Appendix 2).

Genotype HyP2-G1 contained 48 isolates from Ethiopia (35) and Tanzania (13) from various districts (Table 3.7A). HyP2-G2 contained 35 isolates from Uganda (20), Kenya (13) and Ethiopia (2). Genotype HyP2-G2 represented both historic and contemporary isolates from Kenya and Uganda. Nearly similar pattern was observed in HyP2-G3 comprising isolates

from Uganda (14), Kenya (7) and Ethiopia (3), but all the isolates represented by this genotype are from the contemporary collection and present in more than one district. Genotypes 4 and 5 contained 20 isolates each with HyP2-G4 representing isolates from Tanzania (17) and Kenya (3), whilst HyP2-G5 representing isolates from Kenya (11) and Uganda (9) as shown in Table 3.7A. Reticulate network analysis shows the distribution pattern of the shared genotypes across the 4 countries revealing two major groups. Group A represented genotypes predominantly from Kenya and Uganda, whilst Group B represented genotypes predominantly from Tanzania and Ethiopia (Figure 3.9).

Bayesian phylogenetic tree was constructed with the multiple sequence alignment of the 80 genotypes identified (further details are available in Appendix 2) and a rice blast isolate sequence used as an out-group, with 70 % PPV (Figure 3.10). The Tamura 3-parameter (T92) model with discrete gamma distribution (G) was determined to be the best evolutionary model (T92+G) for the HyP2 sequence dataset. Phylogenetic analysis clearly distinguished the FMB pathogen isolates from the rice blast isolate. Based on the tree topology, the *M. oryzae* genotypes were divided into 2 groups A and B. Group A comprised 59 genotypes representing 205 isolates that are mostly dominant in Kenya and Uganda (173 isolates) and 32 isolates were from Tanzania and Ethiopia (Figure 3.10). Group B with an overall 99.9 % PPV comprised 21 genotypes representing 95 isolates mainly from Tanzania and Ethiopia (88 isolates) with only 7 were isolates from Kenya and Uganda (Figure 3.10).

Table 3.7A. Diversity and distribution pattern of shared genotypes *M. oryzae* identified among the finger millet blast pathogen populations in Kenya, Uganda, Tanzania and Ethiopia based on the HyP2 marker sequences

Genotype	No. of isolates in genotype	Representative isolate	No. of countries present	No. of isolates per country	Isolate code	No. of districts	Population type
HyP2-G1	48	T57	2	35 (E)	**E2, E4, E5, E7, E11, E12, E14, E15, E19, E20, E22, E29, E33, E34, E36, E37, E38, E39, E41, E44, E45, E47, E48, E49, E50, E51, E52, E53, E54, E56, E57, E58, E59, E60, E63**	13	C
				13 (T)	**T1, T2, T3, T5, T8, T10, T22, T25, T47, T49, T51, T55, T57**	5	C
HyP2-G2	36	U36	3	2 (E)	**E46, E62**	2	C
				13 (K)	**K7, K9, K15, K18, K24, K27, K35, K42,** K5/23, K14/74, K47/114p, K60/131p, K65/140n	8	H & C
				21 (U)	**U4, U24, U25, U34, U36,** D6/s1, D2/s14, D2/s26, D3/s3, D5/s1, D9/s50, D9/s54, D9/s56, D9/s76, D10/s71, D13/s5, D14/s30, E11p-1-1*, Odyp-2-1*, Secn-2-2*	12	H & C
HyP2-G3	24	E9	3	3 (E)	**E9, E17, E27**	3	C
				7 (K)	**K1, K2, K3, K10, K11, K12, K33**	4	C
				14 (U)	**U6, U11, U12, U16, U17, U22, U23, U29, U42, U46, U49, U50, U52, U58**	10	C
HyP2-G4	20	T52	2	3 (K)	K33/184, K33/189, K48/115n	2	H
				17 (T)	**T6, T7, T9, T18, T19, T21, T24, T28, T32, T34, T36, T38, T41, T43, T44, T48, T52**	5	C
HyP2-G5	20	U54	2	11 (K)	**K5, K19, K23, K26, K39, K41, K43,** K9/46, K28/82w, K21/68n, K29/164	6	H & C
				9 (U)	**U10, U14, U27, U30, U38, U48, U54,** D1/s44, D14/s27	7	H & C

HyP2-G6	14	U31	3	4 (E)	**E3, E32, E43, E55**	4	C
				3 (K)	**K4, K45,** K1/15	3	H & C
				7 (U)	**U2, U31,** D1/s72, D15/s12, D15/s47, Gup-2-1*, P665n-2-1*	4	H & C
HyP2-G7	12	U45	2	3 (K)	**K30,** K5/24w, K65/159w	2	H & C
				9 (U)	**U15, U28, U37, U43, U44, U45**, D4/s41, D11/s16, D1/s53b,	8	H & C
HyP2-G8	12	U55	2	1 (K)	K26/76p	1	H
				11 (U)	**U3, U7, U8, U18, U26, U33, U39, U53, U55, U20**, D15/s6	9	H & C
HyP2-G9	12	U1	2	10 (E)	**E8, E10, E21, E23, E24, E25, E26, E30, E31, E42**	6	H & C
				2 (U)	**U1**, D9/s70	2	H & C
HyP2-G10	11	D7/s6	2	6 (K)	**K14**, **K25, K29**, K23/123, K36/98n, K65/142n	3	H & C
				5 (U)	**U56, U57**, D7/s6, D8/s15, D15/s41	5	H &C
HyP2-G11	5	T58	3	1 (K)	**K22**	1	C
				3 (T)	**T35, T53, T58**	2	C
				1 (U)	Pen2-2	1	H
HyP2-G12	4	T42	1	4 (T)	**T30, T33, T37, T42**	3	C
HyP2-G13	4	T16	2	1 (K)	K58/128p	1	H
				3 (T)	**T4, T12, T16**	2	C
HyP2-G14	4	U5	2	2 (K)	**K6, K40**	2	C
				2 (U)	**U5**, S1p-1-1*	2	H & C
HyP2-G15	3	K4/21p	2	1 (K)	K4/21p	1	H
				2 (U)	D3/s9, D4/s12	2	H
HyP2-G16	2	E13	1	2 (E)	**E6, E13**	2	C
HyP2-G17	2	E35	1	2 (E)	**E16, E35**	2	C
HyP2-G18	2	K13/37	1	2 (K)	**K13,** K13/67	2	H & C
HyP2-G19	2	T40	1	2 (T)	**T39, T40**	1	C
HyP2-G20	2	U13	1	2 (U)	**U9, U13**	2	C
HyP2-G21	2	D10/s63	1	2 (U)	D1/s50, D10/s63	2	H
HyP2-G22	2	K57/126p	1	2 (K)	K55/124p, K57/126p	1	H

H, historical isolates (2000-2004) and C, contemporary isolates (2015-2017);

In the Isolate code, K-isolates collected from Kenya, E-isolates collected from Ethiopia, T-isolates collected from Tanzania and U and D-isolates collected from Uganda; *Isolates from Uganda;

Isolates in bold and normal font are contemporary and historic collections, respectively. Further details of the isolates are available in **Table 2.1A.**

Table 3.7B. Diversity and distribution pattern of single genotypes[a] of *M. oryzae* identified based on the HyP2 sequence data among the isolates associated with finger millet production in various districts in Ethiopia, Kenya, Tanzania and Uganda

Country	District	Number of Genotypes/Isolates	Finger millet	Other weed hosts[b]	Populations[c]
Ethiopia	Sire	1	1	0	C
	Waju Tuka	1	1	0	C
	Boji Bermeji	1	1	0	C
	Dangla	1	1	0	C
	Jabi Tana	1	1	0	C
	Total	***5***	***5***	***0***	**C**
Kenya	Siaya	1	1	0	C
	Homabay	1	1	0	C
	Busia	4	3	1	H & C
	Bungoma	1	1	0	C
	Kisii	8	8	0	H & C
	Kisii central	1	1	0	H
	Kakamega	1	1	0	C
	Eldoret	1	1	0	C
	Marakwet	1	1	0	C
	Teso	2	2	0	H
Total		***21***	***20***	***1***	**H & C**
Tanzania	Nkasi	3	3	0	C
	Sumbawanga	1	1	0	C
	Njombe	6	6	0	C
	Madaba	4	4	0	C
	Momba	2	2	0	C
Total		***16***	***16***	***0***	**C**
Uganda	Apac	2	1	1	H
	Alebtong	2	2	0	C
	Agago	1	1	0	C

	Mbale	1	1	0	H
	Ngora	1	1	0	C
	Kitgum	1	1	0	C
	Pallisa	2	2	0	H
	Nakasongola	1	1	0	H
	Hoima	1	1	0	C
	Soroti	1	0	1	H
	Tororo	3	3	0	H & C
Total		***16***	***14***	***2***	**H & C**
Total	**31**	**58**	**55**	**3**	**NA**

[a] Single genotype are individual isolates distinct HyP2 marker sequence;

[b] Other weed host isolates include *Eleusine indica*, *E. africana*, *Digitaria scalarum*, *D. aegytium*, *D. horizontalis*;

[c] H - historical isolates (2000-2004), C- contemporary isolates (2015-2017);

NA - Not applicable;

Further details of the isolates represented by these genotypes are available in Appendix 2.

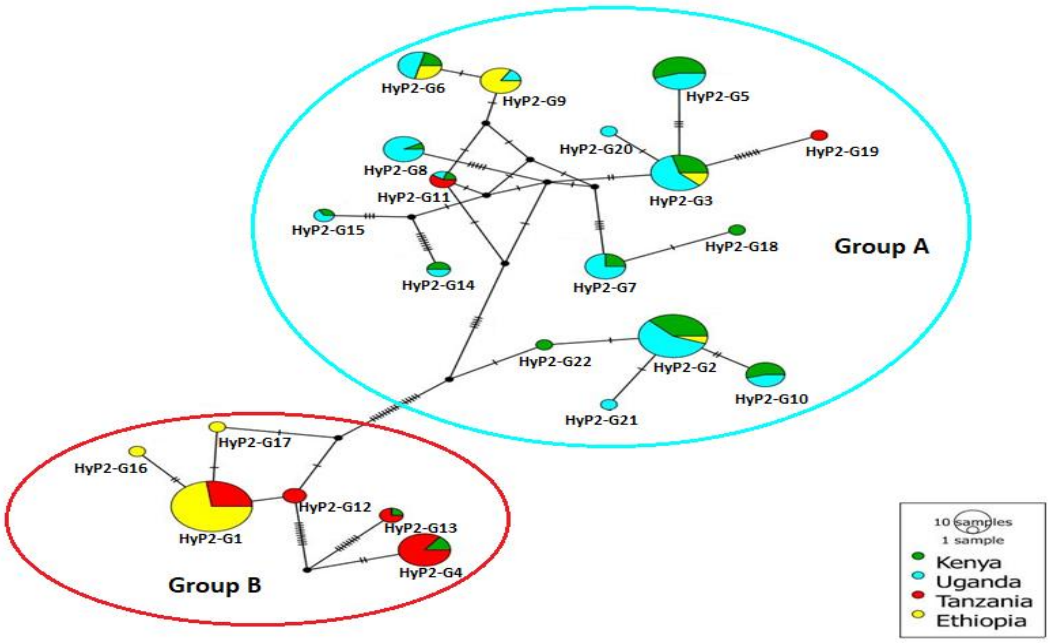

Figure 3.9: Reticulate median-joining network of 234 isolates represented by 22 shared genotypes based on HyP2 marker sequences (Table 3.7A)

Detail of isolates represented by the 22 genotypes are avilable in Table 3.7A with further detilas in Appendix 2. Overall circle sizes are proportional to the number of isolates within a specific genotype. Circle slice area is proportional to the number of isolates from a country, where Kenya is in green, Uganda in blue, Tanzania in red and Ethiopia in yellow. The segment/hatch marks reflect mutation levels between genotypes. Group **A** includes majority of isolates from Kenya and Uganda and Group **B** includes majority of isolates from Tanzania and Ethiopia.

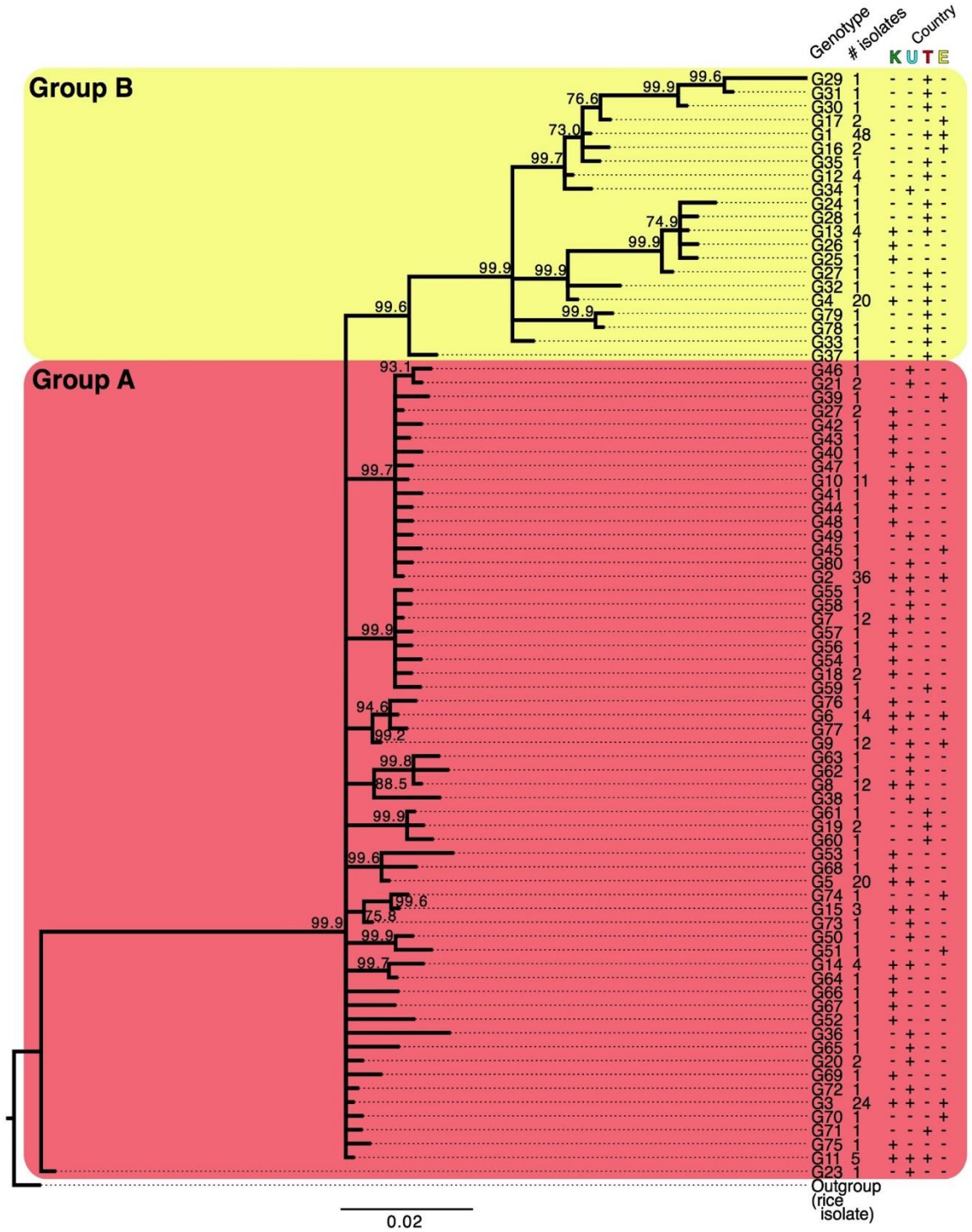

Figure 3.10: Bayesian consensus tree of the *M. oryzae* genotypes based on the HyP2 marker sequences

The tree shows the number of isolates per genotype and their distribution pattern across the four countries: K- Kenya, U – Uganda, T- Tanzania and E- Ethiopia; + indicates present and – indicates absent based on the isolates characterised. The number over the branch of the tree represents the posterior probability value.

3.3.4.5 Multi-locus sequence (MLS) analysis of genetic diversity among *M. oryzae* isolates associated with finger millet in Eastern Africa

The concatenated data of end-trimmed sequence from each marker namely HIS4, ITS, HyP2 and HyP1 generated for the 300 *M. oryzae* isolates (section 2.1) led to a total length ranging from 2135 to 2207 bases. Multiple sequence alignment of the concatenated data of the 300 isolates led to a total length of 2303 base with 560 SNPs and 178 indels. Among the SNPs identified, 375 were parsimony informative sites. The sequence variation distinguished among the 300 *M. oryzae* isolates led to the identification of 207 genotypes. Of these, 28 genotypes ML-G1 to ML-G28 were shared genotypes representing 121 isolates, with each genotype representing 2 to 26 isolates based on identical ML sequences (Table 3.8A). The remaining 179 genotypes were represented by single isolates (Table 3.8B). Nearly 60 % of the total isolates are different to each other based on the ML analysis and in some districts, all the isolates collected were different. For example, in Siaya district of Kenya all isolates were different genotypes, also in Busia district of Kenya 15 out of the 18 isolates were different genotypes (Table 3.8B). The remaining 40 % included two or more isolates with shared ML sequence data. In some instances, the isolates originated from the same geographic location, whilst others originated from different locations (Table 3.8A and Figure 3.11). Figure 3.11 shows the distribution pattern of the shared genotypes across the 4 countries. Furthermore, the network analysis showed a close relationship among the isolates from Kenya and Uganda (designated as group A), and those from Ethiopia and Tanzania (designated as group B).

ML-G1 (Table 3.8A) is the most common genotype containing 26 isolates collected from finger millet in different districts of Ethiopia (17) and Tanzania (9); ML-G2 is the second

largest genotype containing 15 isolates from Kenya (2) and Tanzania (13). Some genotypes represented isolates from one country, e.g. ML-G6, ML-G7, ML-G14. Other genotypes represented isolates from different countries, e.g. ML-G3, ML-G4, ML-G8, ML-9 in Kenya and Uganda (Table 3.8A).

Bayesian phylogenetic tree was constructed from the multiple sequence alignment of the 207-finger millet blast pathogen (FMB) with a rice blast pathogen isolate (RB) used as an outgroup, with 70 % PPV (Figure 3.12). The Kimura 2-parameter model using a discrete gamma distribution (G) with 5 rate categories that contain certain fraction of sites and evolutionary invariable (I) was determined to be the best evolutionary model (K2+G+I) for the concatenated sequence dataset. Based on the tree topology, the *M. oryzae* genotypes were divided into 2 distinct groups A and B (Figure 3.12). Eight genotypes such as K5 and K14 (Kenya), D3/s9 and D10/s71 (Uganda) and T11 (Tanzania) were distinctive and not represented in either of these groups (Table 3.8C). Group A represented 181 isolates from Eastern Africa predominantly from Uganda (97) and Kenya (66) with a limited number from Tanzania (9) and Ethiopia (9) as shown in Table 3.8C). The nucleotide differences among the isolates in Group A ranged from 0 to 3 % (Table 3.8C). Group B contained 117 isolates predominantly from Ethiopia (54) and Tanzania (48) with a limited number from Kenya (6) and Uganda (3) as shown in Table 3.8C. The nucleotide differences among the isolates in Group B ranged from 0 to 11 % (Table 3.8C).

Table 3.8A. Diversity and distribution pattern of shared genotypes of *M. oryzae* identified among the finger millet blast pathogen populations in Kenya, Uganda, Tanzania and Ethiopia based on the multi-locus marker sequences

Genotype	No. of isolates in genotype	Representative isolate	No. of countries present	No. of isolates per country	Isolate code	No. of district	Population type
ML-G1	26	T51	2	17 (E)	**E2, E15, E19, E22, E34, E37, E38, E39, E44, E48, E49, E53, E54, E58, E59, E60, E63**	12	C
				9 (T)	**T1, T2, T3, T8, T10, T22, T47, T49, T51**	4	C
ML-G2	15	T52	2	13 (T)	**T6, T7, T9, T18, T19, T24, T28, T32, T41, T43, T44, T48, T52**	4	C
				2 (K)	K33/184, K33/189	1	H
ML-G3	7	U58	2	4 (U)	**U6, U17, U49, U5**8	4	C
				3 (K)	**K10, K12, K33**	2	C
ML-G4	6	U23	2	2 (K)	**K2, K11**	2	C
				4 (U)	**U11, U12, U22, U23**	2	C
ML-G5	5	K15	3	2 (E)	**E46, E62**	2	C
				2 (U)	D5/s1, D9/s76	2	H
				1 (K)	**K15**	1	C
ML-G6	5	U55	1	5 (U)	**U8, U20, U26, U39, U55**	5	C
ML-G7	5	E31	1	5 (E)	**E21, E23, E24, E25, E31**	3	C
ML-G8	4	U24	2	3 (U)	D9/s50, D13/s5, **U24**	3	H & C
				1 (K)	**K27**	1	C
ML-G9	4	U56	2	2 (K)	K23/123, K36/98n	2	H
				2 (U)	D7/s6, **U56**	2	H & C
ML-G10	3	K4	2	1 (K)	**K4**	1	C
				2 (U)	D1/s72, D15/s12	2	H
ML-G11	3	U48	2	1 (K)	**K23**	1	C
				2 (U)	**U27, U48**	2	C
ML-G12	3	U36	2	1 (K)	K5/23	1	H
				2 (U)	D14/s30, **U36**	2	H & C

ML-G13	3	Odyp-2-1	2	1 (K)	K47/114p	1	H
				2 (U)	E11p-1-1*, Odyp-2-1*	1	H
ML-G14	3	U54	1	3 (U)	**U14, U30, U54**	2	C
ML-G15	3	U33	1	3 (U)	D15/s6, **U18, U33**	3	H & C
ML-G16	2	T16	1	2 (T)	**T12, T16**	2	C
ML-G17	2	T33	1	2 (T)	**T30, T33**	1	C
ML-G18	2	T38	1	2 (T)	**T36, T38**	1	C
ML-G19	2	E33	1	2 (E)	**E29, E33**	1	C
ML-G20	2	E51	1	2 (E)	**E50, E51**	2	C
ML-G21	2	E57	1	2 (E)	**E56, E57**	1	C
ML-G22	2	K28/28w	1	2 (K)	K21/68n, K28/82w	1	H
ML-G23	2	K65/140n	2	1 (U)	D2/s26	1	H
				1 (K)	K65/140n	1	H
ML-G24	2	U25	2	1 (U)	**U25**	1	C
				1 (K)	**K24**	1	C
ML-G25	2	K42	2	1 (U)	D3/s3	1	H
				1 (K)	**K42**	1	C
ML-G26	2	U45	2	1 (U)	**U45**	1	C
				1 (K)	**K30**	1	C
ML-G27	2	U37	2	1 (U)	**U37**	1	C
				1 (K)	K65/159w	1	H
ML-G28	2	U10	2	1 (U)	**U10**	1	C
				1 (K)	**K26**	1	C

H, historical isolates (2000-2004) and C, contemporary isolates (2015-2017);

In the Isolate code, K-isolates collected from Kenya E-isolates collected from Ethiopia, T-isolates collected from Tanzania and U and D-isolates collected from Uganda; *Isolates from Uganda;

Isolates in bold and normal font are contemporary and historic collections, respectively. Further details of the isolates are available in **Table 2.1A.**

Table 3.8B. Diversity and distribution pattern of single genotypes[a] of *M. oryzae* identified based on the multi-locus sequence data among the isolates associated with finger millet production in various districts in Ethiopia, Kenya, Tanzania and Uganda

Country	District	Number of Genotypes/Isolates	Finger millet	Other weed hosts[b]	Populations[c]
Ethiopia	Adet	1	1	0	C
	Angebo	1	1	0	C
	Banja	1	1	0	C
	Bila	1	1	0	C
	Bahir Dar Zuria	3	3	0	C
	Demecha	1	1	0	C
	Dure Bete	3	3	0	C
	Diga	6	6	0	C
	Gaungau	2	2	0	C
	Sire	1	1	0	C
	Lallo Assabi	3	3	0	C
	Leta Sibu	1	1	0	C
	Leta dulecha	3	3	0	C
	Mecha	1	1	0	C
	Nedjo	1	1	0	C
	Waju Tuka	1	1	0	C
	Boji Bermeji	1	1	0	C
	Dangla	1	1	0	C
	Jabi Tana	1	1	0	C
Total		**33**	**33**	**0**	**C**
Kenya	Siaya	5	5	0	C
	Suba	1	1	0	H
	Homabay	1	1	0	H
	Busia	9	7	2	H & C
	Bungoma	2	2	0	C
	Gucha	4	4	0	H
	Kisumu	3	3	0	C
	Kisii	15	15	0	H & C
	Kisii central	1	1	0	H
	Kakamega	1	1	0	C
	Eldoret	3	3	0	C
	Marakwet	1	1	0	C
	Teso	8	6	2	H
Total		**54**	**50**	**4**	**H & C**
Tanzania	Nkasi	3	3	0	C
	Sumbawanga	3	3	0	C

	Njombe	10	10	0	C
	Madaba	5	5	0	C
	Momba	3	3	0	C
	Mbozi	6	6	0	C
Total		**30**	**30**	**0**	**H &C**
Uganda	Apac	4	2	2	H
	Arua	4	4	0	C
	Alebtong	3	3	0	C
	Agago	2	2	0	C
	Amuria	2	2	0	C
	Amuru	1	1	0	C
	Bugiri	1	1	0	H
	Busia	1	1	0	H
	Gulu	1	1	0	C
	Mbale	2	2	0	H & C
	Masindi	2	2	0	H
	Manafwa	1	1	0	C
	Moyo	1	1	0	C
	Ngora	2	2	0	C
	Katakwi	1	1	0	C
	Kabermaido	1	1	0	H
	Kumi	3	3	0	H & C
	Kitgum	1	1	0	C
	Lira	4	3	1	H
	Pallisa	5	5	0	H
	Nakasongola	1	1	0	H
	Hoima	3	3	0	C
	Serere	2	2	0	C
	Soroti	8	7	1	H
	Tororo	6	6	0	H & C
Total		**62**	**58**	**4**	**H & C**
Total	**63**	**179**	**175**	**8**	**NA**

[a] Single genotype are individual isolates with distinct multi-locus sequence data;

[b] Isolates from other weed hosts include *Eleusine indica*, *E. africana*, *Digitaria scalarum*, *D. aegytium*, *D. horizontalis*;

[c] H - historical isolates (2000-2004), C - contemporary isolates (2015-2017);

NA - Not applicable

Further details of isolates belonging to these genotypes are available in Appendix 3.

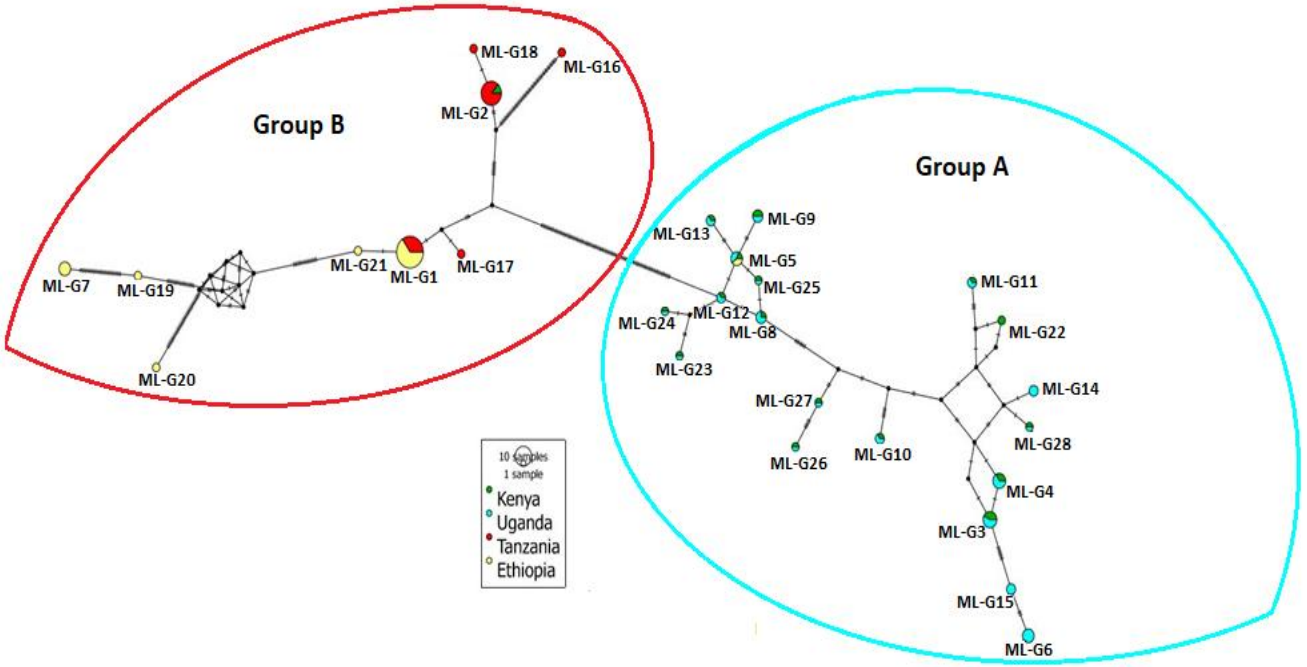

Figure 3.11: Reticulate median-joining network of 121 isolates represented by 28 shared genotypes (Table 3.8B) based on the multi-locus sequence data (HIS4, ITS, HyP2 and HyP1)

Detail of isolates represented by the 28 genotypes are avilable in Table 3.8A with further detilas in Appendix 3. Overall circle sizes are proportional to the number of isolates within a specific genotype. Circle slice area is proportional to the number of isolates from a country, where Kenya is in green, Uganda in blue, Tanzania in red and Ethiopia in yellow. The segment/hatch marks reflect mutation levels between genotypes.Group **A** includes majority of isolates from Kenya and Uganda and Group **B** includes majority of isolates from Tanzania and Ethiopia.

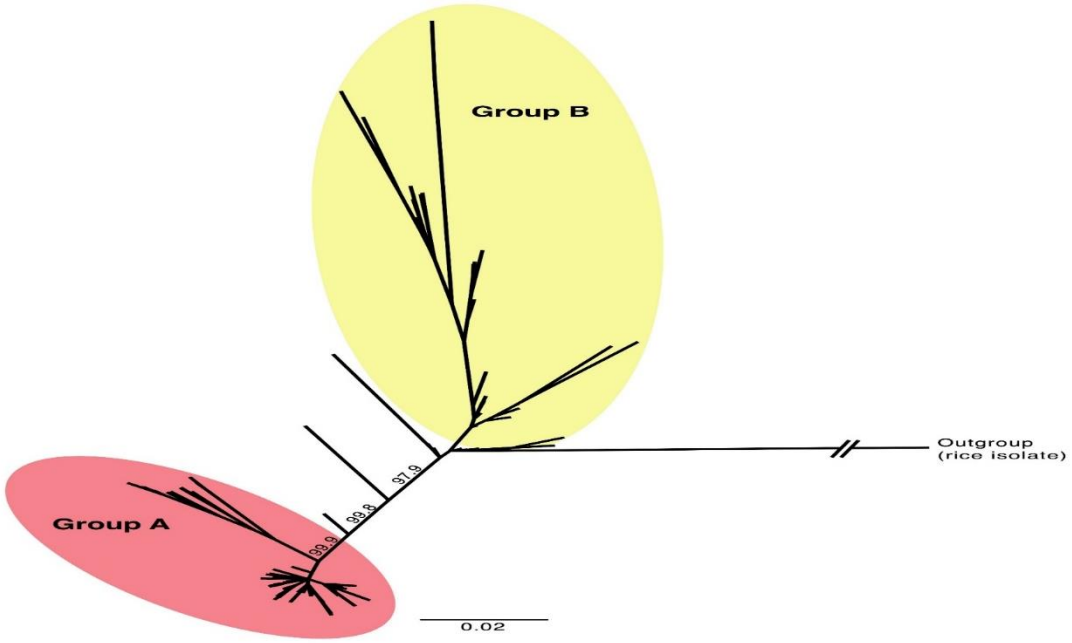

Figure 3.12: Bayesian consensus tree of the *M. oryzae* genotypes based on the multi-locus sequence data

The tree reflects the two main phylogenetic groups of *M. oryzae* isolates associated with finger millet production in Eastern Africa. Also, some isolates were not represented in either group (the three lines between the two-coloured circles). The number next to the branch of the tree represents the posterior probability value. Isolates in each group are presented in Table 3.8C and further details of isolates are available Appendix 3.

Table 3.8C Diversity and distribution of the phylogenetic groups based on the multi-locus sequence data from HIS4, ITS, HyP2 and HyP1 markers

Group	Total isolates per group	Isolates+	Country	Total isolates per country	Nucleotide differences (%)**
A	181	D10/s77, **U21**, Pen-2-2, D4/s12, S1p-1-1*, **U5,** D1/s19, D1/s11, **U7, U32, U8, U20, U26, U39, U55, U51, U53**, D15/s6**, U33, U38**, D1/s72, D15/s12**, U18, U3**, D9/s70, **U1**, P665n-2-1***, U2,** Gup-2-1*, D15/s47, D1/s53b, **U45**, D1/s44, **U11, U12, U22, U23, U9, U50, U19, U13, U16, U6, U17, U49, U58, U31**, D4/s26, **U28**, **U15**, D11/s16, **U43, U47, U44, U37,** D4/s41, **U41, U52, U29, U42, U14, U30, U54, U10, U34**, D5/s1, D9/s76, D3/s24, **U27, U48,** D9/s50, D13/s5**, U24**, D14/s27, **U46**, D2/s14, D8/s15, D6/s1, **U35**, D12/s2, D10/s63, **U25, U4**, D2/s26, **U40,** D14/s30, **U36**, D9/s56, D3/s3, D9/s54, E11p-1-1*, Odyp-2-1*, **U57**, Secn-2-2*, D15/s41, **U56**, D7/s6	Uganda	97	0 - 3
		K1, K1/15, K2, K3, K4, K4/21p, **K5,** K5/23, K5/24w, **K7, K8,** K8/40, **K9**, K9/46, **K10, K11**, **K12**, K12/62, **K13**, K13/67, K14/74**, K15**, K15/53n, **K16, K17, K18, K19, K20**, K21/68, **K22**, K22/118, **K23,** K23/123, **K24, K25, K26, K27,** K26/76p, **K28**, K28/82w, **K29**, K29/164, **K30, K31, K32, K33, K34, K35, K36,** K36/98n, **K37, K38, K39, K40, K42**, K45/112n, **K43, K44, K45,** K47/114p, K55/124p, K57/126p, K64/137p, K65/140n, K65/142n, K65/159w	Kenya	66	
		T14, T17, T20, T35, T39, T40, T53, T54, T58	Tanzania	9	
		E1, E3, E18, E27, E43, E46, E55, E61, E62	Ethiopia	9	
B	111	**E2, E4, E5, E6, E7, E8, E9, E10, E11, E12, E13, E15, E14, E16, E17, E19, E20, E21, E22, E23, E24, E25, E26, E28, E29, E30, E31, E32, E33, E34, E35, E36, E37, E38, E39, E40, E41, E42, E44, E45, E47, E48, E49, E50, E51, E52, E53, E54, E56, E57, E58, E59, E60, E63**	Ethiopia	54	0-11
		T1, T2, T3, T4, T5, T6, T7, T8, T9, T10, T12, T13, T15, T16, T18, T19, T21, T22, T23, T24, T25, T26, T27, T28, T29, T30, T31, T32, T33, T34, T36, T37, T38, T41, T42, T43, T44, T45, T46, T47, T48, T49, T50, T51, T52, T55, T56, T57	Tanzania	48	
		K24/127, K33/184, K33/189, K44/111p, K48/115n, K58/128p	Kenya	6	
		D1/s50, D10/s73, D15/s37	Uganda	3	

K*/ -isolates from Kenya (Historical), K- isolates from Kenya, E-isolates from Ethiopia, T-isolates from Tanzania, U-isolates from Uganda (Contemporary), D-isolates from Uganda (Historical). *Isolates from Uganda. ** Percentage of nucleotide sequence variation within a group. Isolates **K5, K14, K21, K41,** K60/131p (Kenya), D3/s9 and D10/s71 (Uganda), and **T11** (Tanzania) were not represented in either Group A or B. Isolates in bold and normal front are contemporary and historic collections, respectively. Further details of the isolates are available in **Table 2.1A.**

3.3.4.6 Multi-locus analysis of the genetic diversity among the *M. oryzae* isolates within Ethiopia

The concatenated data of end-trimmed sequence from each marker namely HIS4, ITS, HyP2 and HyP1 generated for isolates from Ethiopia ranged from 2175 to 2181 bases in length. Multiple sequence alignment of the data from the 63 isolates led to a total alignment length of 2121 bases showing 200 SNPs and 81 indels. Among the SNPs identified, 158 were parsimony informative sites. The alignment distinguished the 63 *M. oryzae* isolates from Ethiopia into 39 genotypes. Of these, 6 were shared genotypes representing 30 isolates, with each of the genotype representing 2 to 17 isolates that revealed identical multi-locus sequences (Table 3.9A). The remaining 33 genotypes were represented by single isolates (Table 3.9B). ET-G1 is the most distributed genotype representing 17 isolates and found in 11 districts of Ethiopia. ET-G2 isolates were found in 3 districts and the remaining shared genotypes were found in 1 or 2 districts of Ethiopia (Table 3.9A). Bayesian phylogenetic tree was constructed from the multiple sequence alignment of 63 FMB pathogen isolates from Ethiopia and a rice blast isolate used as an outgroup, with 70 % PPV (Figure 3.13). The Kimura 2-parameter model using a discrete gamma distribution (G) with 5 rate categories that contain certain fraction of sites and evolutionary invariable (I) was determined to be the best evolutionary model (K2+G+I) for the concatenated sequence dataset. Based on the tree topology, FMB pathogen isolates from Ethiopia were divided into 4 distinct groups ET1 to ET4 (Figure 3.13). Four isolates (E17, E26, E40 and E42) were distinctive and were not represented in any of the groups (Figure 3.13 and Table 3.9C). ET1 contained 29 isolates distributed in 12 districts of Ethiopia including Diga (6 isolates) and Banja (4) as presented in in Table 3.9C. ET2 contained 9 isolates distributed across 8 districts. ET3 contained 13 isolates distributed across 7 districts including Nedjo (4 isolates) and Diga (3). ET4 contained

8 isolates distributed across 6 districts including Bahir Dar Zuria (2 isolates) and Mecha. The extent of nucleotide difference within the groups varied considerably (Table 3.9C).

Table 3.9A: Diversity and distribution pattern of shared genotypes of *M. oryzae* identified among the isolates collected from finger millet in various districts of Ethiopia based on the multi-locus sequence data analysis

Genotypes	Isolate code	Total no. of isolates	District	No. of districts	Plant part	Variety	Year of collection
ET-G1	E2	17	Wayu Tuka	11	Neck	Landrace	2015
	E19		Wayu Tuka		Neck	-	2016
	E15		Bure		Head	Landrace	2015
	E59		Bure		Head	Landrace	2015
	E22		Nedjo		Neck	-	2016
	E34		Diga		Neck	-	2016
	E37		Diga		Neck	-	2016
	E38		Banja		Head	Landrace	2015
	E39		Banja		Neck	Landrace	2015
	E49		Banja		Head	Landrace	2015
	E53		Mecha		Head	Landrace	2015
	E54		Mecha		Neck	Landrace	2015
	E48		Guangau		Neck	Landrace	2015
	E60		Ankussa-Abdo Gor		Head	Landrace	2015
	E58		Bahir Dar Zuria		Neck	Landrace	2015
	E44		Dure Bete		Neck	Landrace	2015
	E63		Qilxxu Kara		Head	Landrace	2015
ET-G2	E21	5	Wayu Tuka	3	Neck	-	2016
	E23		Nedjo		Neck	-	2016
	E24		Nedjo		Neck	-	2016
	E25		Nedjo		Neck	-	2016
	E31		Gimbi		Head	-	2016
ET-G3	E29	2	Diga	1	Neck	-	2016
	E33		Diga		Head	-	2016
ET-G4	E46	2	Mandura	2	Neck	Landrace	2015
	E62		Jabi Tana		Neck	Landrace	2015
ET-G5	E50	2	Banja	2	Neck	Landrace	2015
	E51		Mecha		Head	Landrace	2015
ET-G6	E56	2	Bahir Dar Zuria	1	Head	Landrace	2015
	E57		Bahir Dar Zuria		Neck	Landrace	2015

+ Numbers in parentheses are the number of isolates represented by a genotype

- Information not available.

Table 3.9B. Diversity and distribution pattern of single genotypes of *M. oryzae* identified among the isolates collected from finger millet in various districts of Ethiopia based on the multi-locus sequence data analysis

District	Number of isolates collected*	Number of single genotypes*
Adet	1	1
Angebo	1	1
Banja	5	1
Bila	1	1
Bahir Dar Zuria	6	3
Demecha	1	1
Dure Bete	4	3
Diga	10	6
Gaungau	3	2
Lallo Assabi	3	3
Leta Sibu	1	1
Leta dulecha	3	3
Mecha	4	1
Nedjo	5	1
Waju Tuka	4	1
Boji Bermeji	1	1
Dangla	1	1
Jabi Tana	2	1
Sire	1	1
Bure	2	0
Mandura	1	0
Ankussa-Abdo	1	0
Gimbi	1	0
Qilxxu Kara	1	0
Total	**63**	**33**

***** Where the numbers are not the same in a district, the isolates belong to shared genotype presented in Table 3.9A.

Further details of isolates belonging to these genotypes are available in Appendix 4.

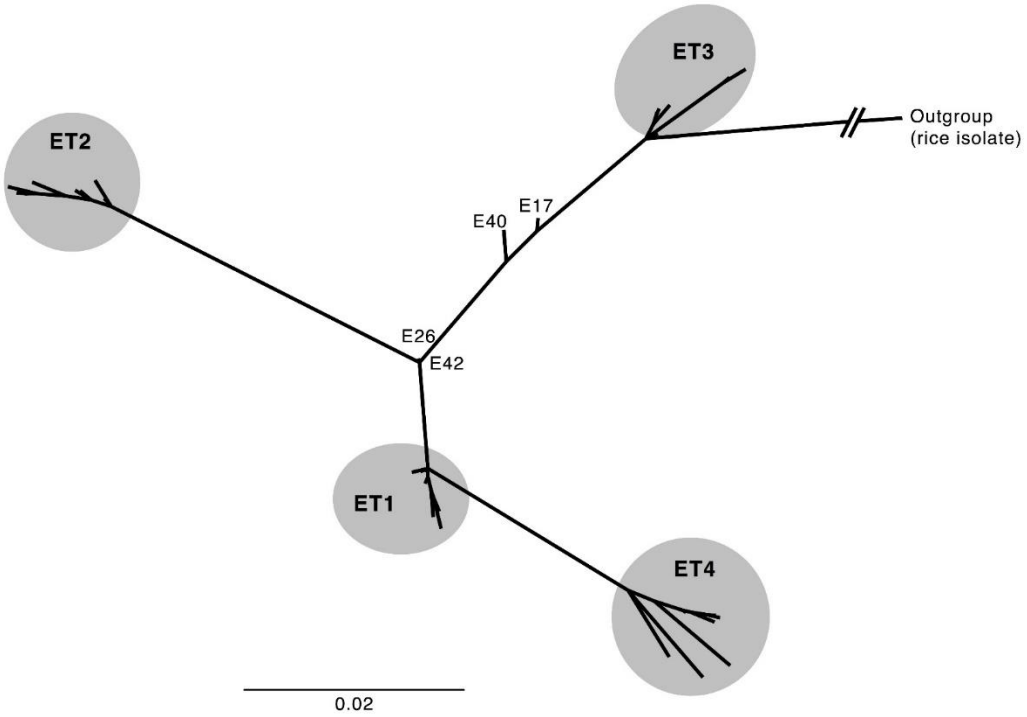

Figure 3.13: Bayesian consensus tree based on the multi-locus sequence data of the finger millet blast pathogen *M. oryzae* isolates from Ethiopia

The phylogenetic tree (based on ITS, HIS4, Hyp1 and HyP2) representing distinctive groups ET1 to ET4. Isolates belonging to each of the groups are listed in the Table 3.9C; isolates such as E17, E26, E40 and E42 were distinctive and were not included in any of the groups.

Table 3.9C Finger millet blast pathogen phylogenetic groups in Ethiopia and their distribution pattern across different districts

Group*	Number of isolates per group	Isolates**	Number of districts	Nucleotide Difference (%)[e]
ET1	29	E2, E5, E6, E7, E11, E12, E15, E19, E20, E22, E34, E35, E37, E38, E39, E41, E44, E45, E47, E48, E49, E53, E54, E56, E57, E58, E59, E60, E63	12[a]	0 - 5.04
ET2	9	E1, E3, E18, E27, E43, E46, E55, E61, E62	8[b]	0 - 0.32
ET3	13	E8, E9, E10, E21, E23, E24, E25, E28, E29, E31, E32, E33, E36	7[c]	0 - 1.2
ET4	8	E4, E13, E14, E16, E30, E50, E51, E52	6[d]	0 - 4.25

[a] Twelve districts include Ankussa-Abdo Gor, Bahir Dar Zuria, Banja, Bure, Diga, Dure Bete, Guangau, Leka dulecha, Nedjo, Mecha, Qilxxu Kara and Wayu Tuka

[b] Eight districts include Adet, Bila, Bahir Dar Zuria, Mandura, Dure Bete, Jabi Tana, Sire and Wayu Tuka

[c] Seven districts include Bahir Dar Zuria, Banja, Diga, Demecha, Lallo Asabi and Mecha

[d] Six districts include Boji Bermeji, Diga, Gimbi, Leka dulecha, Nedjo, Wayu Tuka and Lallo Assabi

[e] Percent nucleotide difference within a group

*The groups identified include isolates represented by shared (Table 3.9A) and single (Table 3.9B) genotypes; shared genotype ET-G5 is included within ET4, whilst the other five shared genotypes are included within ET1.

**Isolates E17, E26, E40 and E42 were not included in any of the groups.

3.3.4.7 Multi-locus analysis of the genetic diversity among the *M. oryzae* isolates within Tanzania

The concatenated data of end-trimmed sequence from each marker namely ITS, HIS4, HyP1 and HyP2 of the 58 *M. oryzae* isolates from Tanzania yielded a total sequence from 2167 to 2181 bases in length. Multiple sequence alignment of the data led to a total alignment length of 2317 bases showing 280 SNPs and 87 indels. Among the SNPs identified, 207 were parsimony informative sites. The alignment distinguished the 58 *M. oryzae* isolates collected from Tanzania into 35 genotypes. Of these, 5 shared genotypes represented 28 isolates with each genotype representing 2 to 13 isolates that revealed identical multi-locus sequences. TZ-G1 and TZ-G2 were the most distributed genotypes containing 13 and 9 isolates, respectively (Table 3.10A). The remaining 30 were single genotypes (Table 3.10B). Bayesian phylogenetic tree was constructed from the concatenated multiple sequence alignment of 58 FMB pathogen isolates and a rice blast isolate as outgroup (Figure 3.14). The Kimura 2-parameter model using a discrete gamma distribution (G) with 5 rate categories that contain certain fraction of sites and evolutionary invariable (I) was determined to be the best evolutionary model (K2+G+I) for the dataset. FMB pathogen isolates were clearly distinguished from the rice blast isolate. Based on the tree topology, FMB pathogen isolates from Tanzania were divided into 3 distinct groups TZ1, TZ2, and TZ3 (Figure 3.14). Isolate T11 was distinctive and was not included in any of the groups (Figure 3.14 and Table 3.10C). TZ1 contained 40 isolates with 0 to 2.1 % nucleotide difference and this group comprised isolates from 6 districts of Tanzania including Nkasi (10 isolates) and Njombe (15). TZ2 contained 9 isolates from 3 districts, Njombe (4 isolates), Madaba (3) and Mbozi (2). TZ3 contained 8 isolates from 5 districts including Mbozi (3 isolates) and Momba (2). Taken together, all the phylogenetic groups identified (TZ1, TZ2 and TZ3) were represented in

Njombe, Mbozi and Madaba districts and the 3 groups showed varied levels of nucleotide differences (Table 3.10C).

Table 3.10A: Diversity and distribution pattern of shared genotypes of *M. oryzae* identified among the isolates collected from finger millet in various districts of Tanzania based on the multi-locus sequence analysis

Genotype	Isolate code	Total no. of isolates	District	No. of districts	Plant part	Improved/Landrace	Year of collection
TZ-G1	T6	13	Njombe	4	Neck	Landrace	2016
	T7		Njombe		Neck	Landrace	2016
	T9		Njombe		Neck	Landrace	2016
	T18		Njombe		Neck	Landrace	2016
	T48		Njombe		Neck	Landrace	2016
	T52		Njombe		Head	Landrace	2016
	T19		Mbozi		Neck	Landrace	2016
	T41		Mbozi		Neck	Landrace	2016
	T43		Mbozi		Neck	Landrace	2016
	T44		Mbozi		Neck	Landrace	2016
	T24		Nkasi		Neck	Landrace	2016
	T28		Nkasi		Neck	Landrace	2016
	T32		Sumbawanga		Neck	Landrace	2016
TZ-G2	T1	9	Nkasi	4	Neck	Landrace	2015
	T2		Nkasi		Neck	Landrace	2016
	T3		Nkasi		Neck	Landrace	2016
	T8		Njombe		Neck	Landrace	2016
	T47		Njombe		Neck	Landrace	2016
	T49		Njombe		Neck	Landrace	2016
	T51		Njombe		Neck	Landrace	2016
	T10		Madaba		Neck	Landrace	2016
	T22		Momba		Neck	Landrace	2016
TZ-G3	T30	2	Nkasi	1	Neck	Landrace	2016
	T33		Nkasi		Neck	Landrace	2016
TZ-G4	T36	2	Momba	1	Neck	Landrace	2016
	T38		Momba		Neck	Landrace	2016
TZ-G5	T12	2	Sumbawanga	2	Neck	Landrace	2016
	T16		Njombe		Neck	Landrace	2016

Table 3.10B. Diversity and distribution pattern of single genotypes of *M. oryzae* identified among the isolates collected from finger millet in various districts of Tanzania

District	Total no. of isolates collected	No. of single genotypes
Nkasi	10	3
Sumbawanga	5	3
Njombe	21	10
Madaba	6	5
Momba	6	3
Mbozi	10	6
Total	**58**	**30**

Details of isolates belonging to these genotypes are available in Appendix 5.

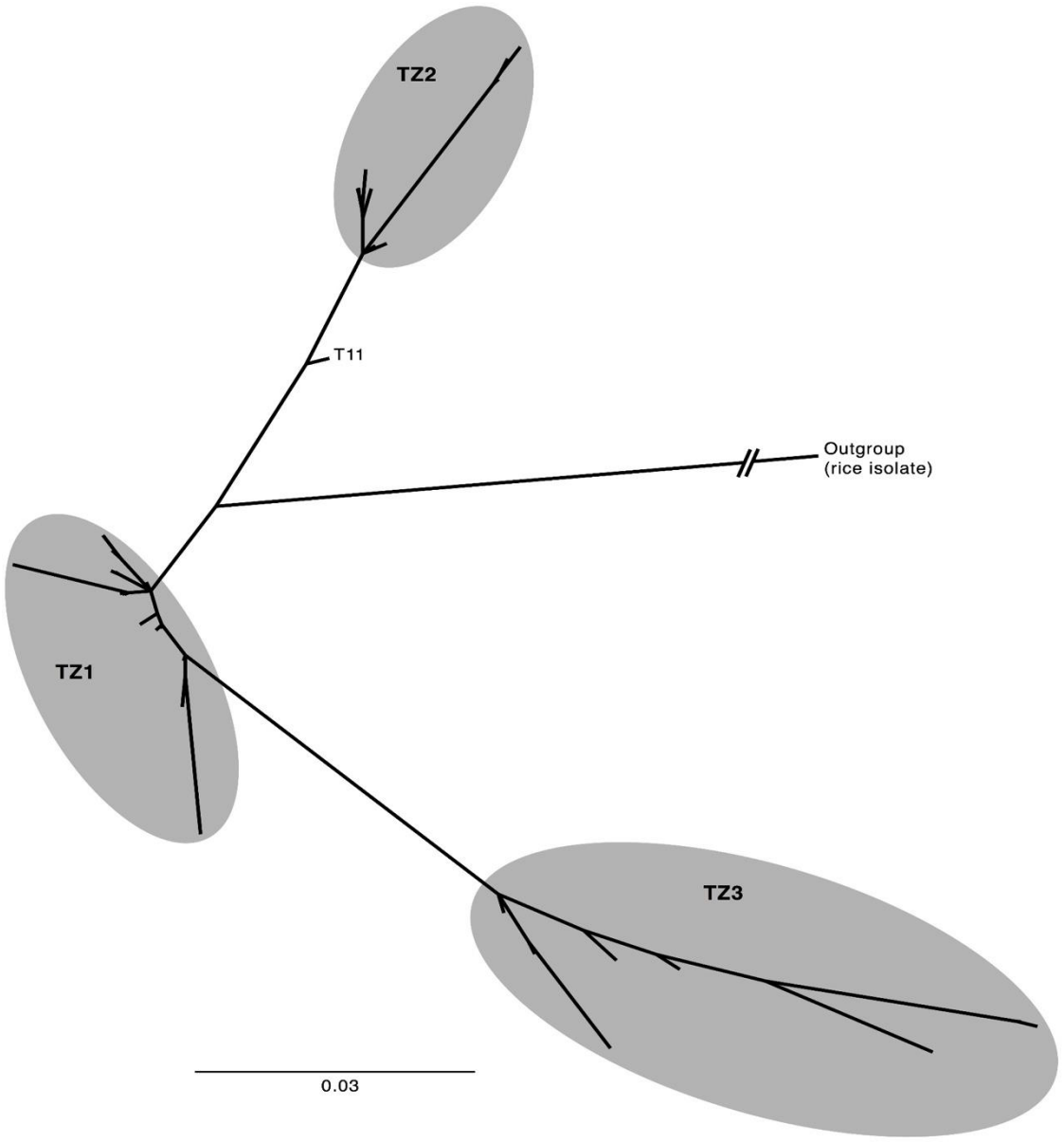

Figure 3.14: Bayesian consensus tree based on the multi-locus sequence data of the finger millet blast pathogen *M. oryzae* isolates from Tanzania

The phylogenetic tree based on ITS, HIS4, Hyp1 and HyP2 representing distinctive groups TZ1, TZ2 and TZ3. Isolates belonging to each group are listed in the Table 3.10C; Isolate T11 was distinctive and not included in any of the groups.

Table 3.10C. Finger millet blast pathogen phylogenetic groups in Tanzania and their distribution pattern across different districts

Group*	Number of isolates per group	Isolates**	Number of districts	Nucleotide difference (%)[d]
TZ1	40	T1, T2, T3, T6, T7, T8, T9, T10, T12, T13, T16, T18, T19, T22, T24, T26, T27, T28, T29, T30, T31, T32, T33, T34, T36, T38, T41, T42, T43, T44, T45, T46 T47, T48, T49, T50, T51, T52, T55, T56,	6[a]	0 – 2.1
TZ2	9	T14, T17, T20, T35, T39, T40, T53, T54, T58	3[b]	0 – 3.0
TZ3	8	T4, T5, T15, T21, T23, T25, T37, T57	5[c]	0 – 4.0

[a] Five districts include Nkasi, Njombe, Madaba, Momba, Mbozi and Sumbawanga

[b] Three districts include Madaba, Njombe and Mbozi

[c] Five districts include Sumbawanga, Mbozi, Madaba, Momba and Njombe

[d] Percentage nucleotide difference within a group

*The groups identified include isolates represented by shared (Table 3.10A) and single (Table 3.10B) genotypes; all the shared genotypes are within TZ1.

**Isolate T11 was distinctive and not included in any of the groups.

3.3.4.8 Multi-locus analysis of the genetic diversity among the *M. oryzae* isolates within Kenya

The concatenated data of end-trimmed sequence from each marker namely ITS, HIS4, HyP1 and HyP2 of the 77 *M. oryzae* isolates from Kenya led to a total of 2136 to 2181 bases in length. These 77 isolates comprised 32 isolates representing historical collection (2000 - 2004) and 45 isolates representing contemporary collection (2015 to 2017). Multiple sequence alignment of the concatenated sequence of these isolates led to a total alignment length of 2198 bases showing 280 SNPs and 78 indels. Among the SNPs identified, 160 were parsimony informative sites. The alignment distinguished the 77 *M. oryzae* isolates collected from Kenya into 71 genotypes. Of these, 5 were shared genotypes representing 2 to 3 isolates each (Table 3.11A) and the remaining 66 were single genotypes (Table 3.11B). This pattern shows that *M. oryzae* isolates in Kenya are diverse irrespective of the host and location. For example, all the isolates collected from Siaya, Gucha and Eldoret districts were different to each other based on the multi-locus sequence analysis. No dominant genotype could be identified among the *M. oryzae* isolates collected from finger millet production systems in Kenya. The evolutionary relationships of these isolates were assessed based on Bayesian analysis with Kimura 2-parameter model using a discrete gamma distribution (G) with 5 rate categories that contain certain fraction of sites and evolutionary invariable (I) chosen as the best evolutionary model (K2+G+I) and a rice blast isolate as outgroup. The analysis distinguished FMB pathogen isolates from Kenya into 3 groups KN1, KN2, and KN3. K6 isolate was not group and designated as ungroup (Figure 3.15 and Table 3.11C). KN1 contained 54 isolates (70 %) from 12 districts in Kenya including Busia (18 isolates), Kisii (6), and Teso (11). KN2 comprised 10 isolates from 6 districts including Gucha (2 isolates) and Kisii (3). KN3 contained 12 isolates from 4 districts including Kisii (9 isolates). Each of the

groups contained isolates representing historical and contemporary populations. The extent of nucleotide difference within the groups varied considerably (Table 3.11C).

Table 3.11A: Diversity and distribution pattern of shared genotypes of *M. oryzae* identified among the isolates collected from finger millet in various districts of Kenya based on the multi-locus sequence analysis

Genotype*	Isolate code	District	No. of districts	Plant part	Variety/type	Year of collection
KN-G1 (3)	K10	Busia	2	Neck	Improved	2016
	K12	Busia		Neck	Improved	2016
	K33	Kisii		Neck	-	2017
KN-G2 (2)	K2	Kisumu	2	Neck	Improved	2016
	K11	Busia		Neck	Improved	2016
KN-G3 (2)	K23/123	Busia	2	Neck	-	2000
	K36/98n	Teso		Neck	-	2002
KN-G4 (2)	K33/184	Kericho	1	Panicle	-	2000
	K33/189	Kericho		Leaf	-	2000
KN-G5 (2)	K21/68	Teso	1	Neck	-	2002
	K28/82w	Teso		-	-	2002

* Numbers in parentheses are the number of isolates represented by a genotype

- Information not available

Table 3.11B: Diversity and distribution pattern of single genotypes of *M. oryzae* identified among the isolates collected from finger millet in various districts of Kenya

Districts	Total no. isolates collected*	Number of single Genotypes*	Population
Siaya	5	5	C
Suba	1	1	H
Homabay	1	1	H
Busia	18	15	H & C
Bungoma	2	2	C
Gucha	4	4	H
Kisumu	6	5	C
Kisii	18	17	H & C
Kisii central	1	1	H
Kakamega	1	1	C
Eldoret	3	3	C
Marakwet	1	1	C
Alupe/Teso	2	2	H
Teso	12	8	H
Kericho	2	0	H
Total	**77**	**66**	

H - historical isolates (collected from 2000-2004) and C - contemporary isolates (collected from 2015 to 2017).

***** Where the numbers are not the same in a district, the isolates belong to shared genotype presented in Table 3.11A.

Details of isolates belonging to these genotypes are available in Appendix 6.

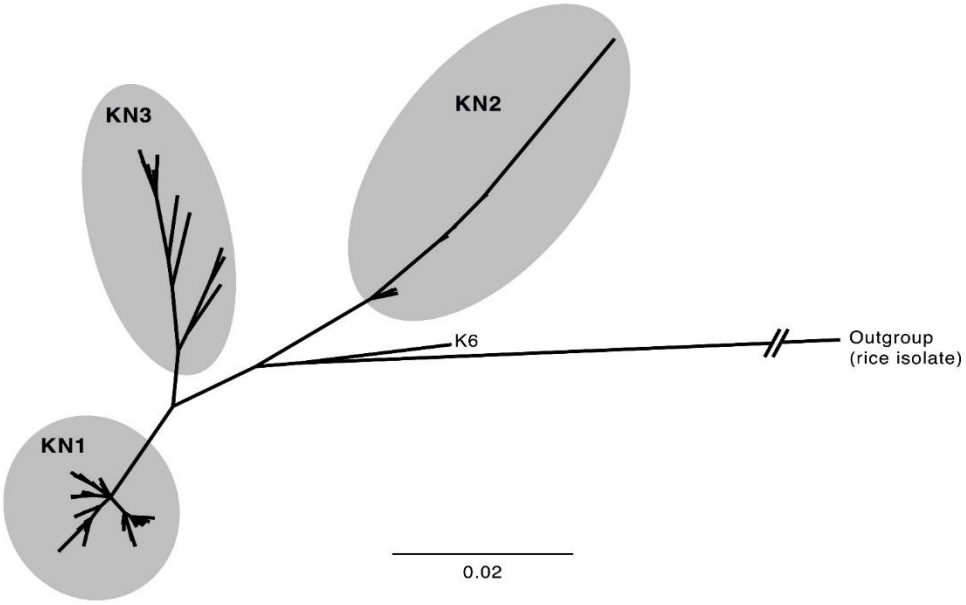

Figure 3.15: Bayesian consensus tree based on the multi-locus sequence data of the finger millet blast pathogen *M. oryzae* isolates from Kenya

The phylogenetic tree based on ITS, HIS4, Hyp1 and HyP2 representing distinctive groups KN1, KN2 and KN3. Isolates belonging to each group are listed in the Table 3.11C; Isolate K6 was distinctive and not included in any of the groups.

Table 3.11C Finger millet blast pathogen phylogenetic groups in Kenya and their distribution pattern across different districts

Group*	Number of isolates per group	Isolates**	Number of districts	Nucleotide difference (%)[d]
KN1	54	K1, K1/15, K2, K3, K4, K4/21p, K5, K5/23, K5/24w, K7, K8, K8/40, K9, K9/46, K10, K11, K12, K12/62, K13, K13/67, K14/74, K15, K15/53n, K17, K18, K19, K20, K21/68, K22, K22/118, K23, K23/123, K24, K25, K26, K26/76p, K27, K28, K28/82w, K29/164, K30, K31, K32, K33, K35, K36/98n, K42, K43, K45, K47/114p, K55/124p, K65/140n, K65/142n, K65/159w	12[a]	0 – 2.7
KN2	10	K14, K21, K41, K24/127, K33/184, K33/189, K44/111p, K48/115n, K58/128p, K60/131p	6[b]	0 -7.9
KN3	12	K16, K29, K34, K36, K37, K38, K39, K40, K44, K45/112n, K57/126p, K64/137p	4[c]	0 – 1.5

[a] Twelve districts include Eldoret, Kisii, Busia, Bungoma, Kisumu, Siaya, Teso, Gucha, Alupe/Teso, Kakamega, Eldoret and Suba

[b] Six districts include Bungoma, Gucha, Kisii, Kisii central, Kericho and Marakwet

[C] Four districts include Busia, Gucha, Homabay and Kisii

[d] Percentage nucleotide difference within a group

*The groups identified include isolates represented by shared (Table 3.11A) and single (Table 3.11B) genotypes; genotype KN-G4 is included within KN2, whilst the other four genotypes are included within KN1.

**Isolate K6 was distinctive and not included in any of the groups.

3.3.4.9 Multi-locus analysis of the genetic diversity among the *M. oryzae* isolates within Uganda

The concatenated data of end-trimmed sequence from each marker namely ITS, HIS4, HyP1 and HyP2 of the 102 *M. oryzae* isolates from Uganda yielded a total sequence of 2136 to 2207 bases in length. These 102 isolates comprised 44 isolates from historical collection (2000 - 2004) and 58 isolates from contemporary collection (2015 to 2017). Multiple sequence alignment of the data led to a total alignment length of 2231 bases showing 364 SNPs and 82 indels. Among the SNPs identified, 198 were parsimony informative sites. The alignment distinguished the isolates into 80 genotypes. Of these, 12 were shared genotypes representing 37 isolates (Table 3.12A) with each shared genotype representing 2 to 5 isolates. The remaining 68 genotypes were represented by single isolates (Table 3.12B). There is no dominant genotype identified among the *M. oryzae* isolates associated with finger millet production systems in Uganda. The majority of the isolates are different to each other, for example, all the isolates from Arua district and the majority of the isolates from locations such as Apac, Tororo and Soroti are genetically different based on the multi-locus sequence data. The evolutionary relationships of these isolates were assessed based on Bayesian analysis with Kimura 2-parameter model using a discrete gamma distribution (G) with 5 rate categories that contain certain fraction of sites and evolutionary invariable (I) chosen as the best evolutionary model (K2+G+I) with a rice blast isolate as outgroup. Based on the tree topology, FMB pathogen isolates from Uganda were divided into 2 distinct groups UG1 and UG2. Isolates D3/s9 and D10/s73 were distinctive and were not included in either of the groups (Figure 3.16). UG1 contained 97 isolates (95 %) from 28 districts of Uganda. UG1 comprised isolates representing both historical and contemporary populations and the nucleotide difference ranged from 0 to 2.9 % in this group. UG2 represented 3

isolates from Apac, Soroti and Pallisa districts with a nucleotide difference of 0 to 8.5 % (Figure 3.16 and Table 3.12C).

Table 3.12A: Diversity and distribution pattern of shared genotypes of *M. oryzae* identified among the isolates collected from finger millet in various districts of Uganda based on the multi-locus sequence analysis

Genotype+	Isolate code	District	No. of districts	Plant part	Finger millet variety/type	Year of collection
UG-G1 (5)	U8	Kumi	5	Neck	Local	2016
	U20	Alebtong		Neck	-	2016
	U26	Apac		Neck	Local	2016
	U39	Tororo		Neck	Local	2016
	U55	Lamwo		Neck	Improved	2016
UG-G2 (4)	U6	Serere	4	Neck	Local	2016
	U17	Ngora		Neck	Local	2016
	U49	Kumi		Head	-	2016
	U58	Mbale		Neck	-	2016
UG-G3 (4)	U11	Pallisa	2	Neck	Improved	2016
	U12	Pallisa		Neck	-	2016
	U22	Lira		Neck	-	2016
	U28	Lira		Neck	Local	2016
UG-G4 (3)	D9/s50	Lira	3	Panicle	-	2002
	D13/s5	Katakwi		Panicle	-	2002
	U24	Apac		Neck	Local	2016
UG-G5 (3)	U14	Katakwi	2	Neck	Local	2016
	U30	Amuru		Neck	Local	2016
	U54	Amuru		Neck	Local	2016
UG-G6 (3)	D15/s6	Soroti	3	Panicle	-	2000
	U13	Katakwi		Neck	Local	2016
	U18	Ngora		Neck	Local	2016
UG-G7 (2)	D5/s1	Iganga	2	Panicle	-	2000
	D9/s76	Lira		Panicle	-	2002
UG-G8 (2)	E11p-1-1	Soroti	1	Panicle	-	2004
	Odyp-2-1	Soroti		Panicle	-	2004
UG-G9	D14/s30	Kabermaido	2	Leaf	-	2002

(2)	U36	Manafwa		Neck	Local	2016
UG-G10	D7/s6	Kamuli	2	Panicle	-	2000
(2)	U56	Manafwa		Neck	Local	2016
UG-G11	U27	Gulu	2	Neck	Local	2016
(2)	U48	Kumi		Neck	Local	2016
UG-G12	D1/s72	Pallisa	2	Leaf	-	2002
(2)	D15/s12	Soroti		Panicle	-	2000

+ Numbers in parentheses are the number of isolates represented by a genotype

- Information not available.

Table 3.12B Diversity and distribution pattern of single genotypes of *M. oryzae* identified among the isolates collected from finger millet in various districts of Uganda

Districts	Total no. of isolates collected per district*	Number of single Genotypes*	Population
Apac	7	5	H & C
Arua	4	4	C
Alebtong	4	3	C
Agago	2	2	C
Amuria	2	2	C
Amuru	3	1	C
Bugiri	1	1	H
Busia	1	1	H
Gulu	2	1	H
Mbale	4	3	H
Masindi	2	2	H & C
Manafwa	4	2	C
Moyo	1	1	C
Ngora	4	2	C
Katakwi	3	1	C
Kabermaido	2	1	H
Kumi	8	5	H & C
Kitgum	2	1	C
Lira	8	4	H & C
Pallisa	8	5	H
Nakasongola	1	1	H
Hoima	3	3	C
Serere	4	3	C
Soroti	12	8	H
Tororo	7	6	H & C
Iganga	1	0	H
Kamuli	1	0	H
Lamwo	1	0	C
Total	**102**	**68**	**H & C**

H - historical isolates (collected from 2000-2004) and C - contemporary isolates (collected from 2015 to 2017).

*****Where the numbers are not the same in a district, the isolates belong to shared genotype presented in Table 3.12A.

Details of isolates belonging to these genotypes are available in Appendix 7.

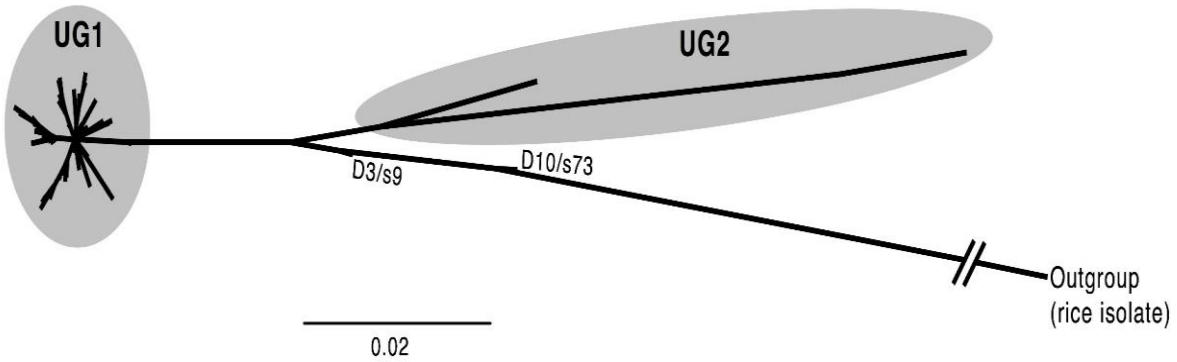

Figure 3.16: Bayesian consensus tree based on the multi-locus sequence data of the finger millet blast pathogen *M. oryzae* isolates from Uganda

The phylogenetic tree based on ITS, HIS4, Hyp1 and HyP2 representing distinctive groups UG1 and UG2. Isolates belonging to each group are listed in the Table 3.12C; Isolates D3/s9 and D10/s73 were distinctive and not included in either of the groups.

Table 3.12C Finger millet blast pathogen phylogenetic groups in Uganda and their distribution pattern across different districts

Group*	Number of isolates per group	Isolates**	Number of districts	Nucleotide difference (%)[c]
UG1	97	U1, U2, U3, U4, U5, U6, U7, U8, U9, U10, U11, U12, U13, U14, U15, U16, U17, U18, U19, U20, U21, U22, U23, U24, U25, U26, U27, U28, U29, U30, U31, U32, U33, U34, U35, U36, U37, U38, U39, U40, U41, U42, U43, U44, U45, U46, U47, U48, U49, U50, U51, U52, U53, U54, U55, U56, U57, U58, D1/s11, D1/s19, D1/s44, D1/s72, D1/s53b, D2/s14, D2/s26, D3/s3, D3/s24, D4/s12, D4/s26, D4/s41, D5/s1, D6/s1, D7/s6, D8/s15, D9/s50, D9/s54, D9/s56, D9/s70, D9/s76, D10/s63, D10/s77, D11/s16, D12/s2, D13/s5, D14/s27, D14/s30, D15/s6, D15/s12, D15/s41, D15/s47, Gup-2-1, E11p-1-1, S1p-1-1, Secn-2-2, Odyp-2-1, P665n-2-1, Pen-2-2	28[a]	0 – 2.9
UG2	3	D1/s50, D10/s71, D15/s37	3[b]	0 – 8.5

[a] All 28 districts in Uganda where samples were collected for this study

[b] Three districts include Apac, Soroti and Pallisa

[c] Percentage nucleotide difference within a group

* The groups identified include isolates represented by shared (Table 3.12A) and single (Table 3.12B) genotypes; all the shared genotypes are within UG1.

**Isolates D3/s9 and D10/s73 were distinctive and not included in either of the groups.

3.3.5 PCR screening of the contemporary populations of *M. oryzae* for the *Grasshopper* (*grh*) repeat element

Initially, the PCR conditions including the two pairs of primers available (designated as PES and PKE) for two different regions of the *grh* element (Takan *et al*., 2012) were validated using a limited number of previously characterised isolates, along with a reference isolate G22. The PCR conditions and sets of primers were effective in the amplification of the two regions of the *grh* element (data not shown).

In Figure 3.17, out of the nine contemporary isolates screened, six isolates T1 to K4 were identified as *grh* positive based on the presence of the amplicons with the PKE and PES primer pairs (~840 bp and ~1340 bp, respectively) comparable to the reference isolate G22. On the contrary, isolates K5, K6 and U1 were identified as *grh* negative based on the absence of corresponding amplicons in the PCR screening. This approach was used to screen the entire collection of contemporary isolates from Ethiopia (63), Kenya (45), Tanzania (58) and Uganda (58) for the *grh* repeat element.

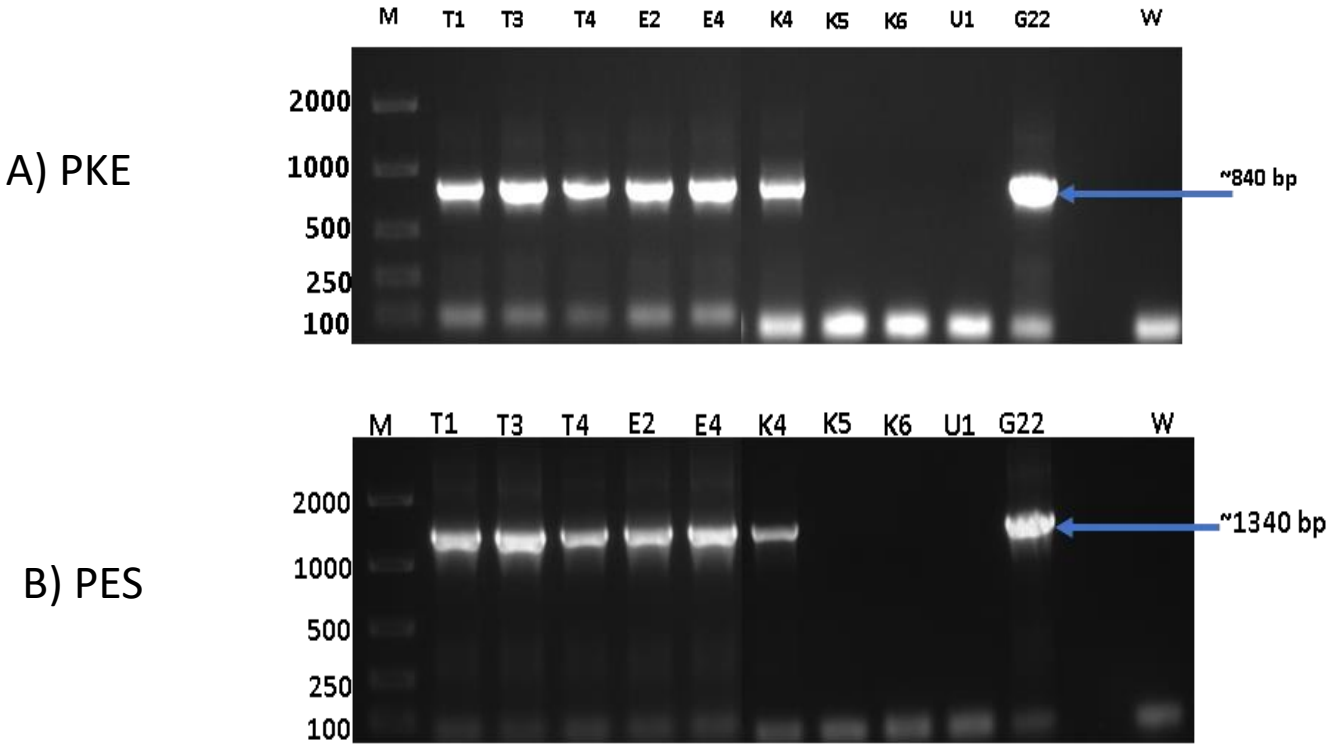

Figure 3.17. Agarose gel showing example of PCR screening of *M. oryzae* and identification of *grasshopper* (*grh*) positive and *grh* negatives isolates

In the gel image A, the PCR product is ~840 bp using PKE primers and in the gel image B, the PCR product is ~1340 bp using PKE primers. The *M. oryzae* isolates shown on the gel are: T1, T3, T4 from Tanzania; E2 and E4 from Ethiopia; K4, K5 and K6 from Kenya; and U1 form Uganda. G22 used a reference (positive control) is a previously characterised *M. oryzae* isolate from finger millet in Japan from which the *grh* repeat element was originally described (Dobinson *et al*., 1993; Takan *et al*., 2012). Lane M is a DNA molecular ladder of kwon sized fragments and Lane W is a PCR negative control reaction containing water instead of DNA. Further details of the isolates are available in Table 2.1.

NOTE: 6 isolates T1 to K4 are identified as *grh* positive based on the presence of the amplicons with the PKE and PES primer pairs comparable to the reference isolate G22; Isolates K5 to U1 are identified as *grh* negative based on the absence of corresponding amplicons in the PCR screening.

Out of the 224 isolates screened, 110 isolates yielded *grh* positive products in PCR and 114 isolates tested *grh* negative (Table 3.13). The *grh* positive isolates includes 55 from Ethiopia, 49 from Tanzania, 5 from Kenya and 1 from Uganda (Table 3.13). In Ethiopia, 55 out of the 63 isolates tested were *grh* positive and were found almost in all the districts apart from Adet, Jabi Tana, Mandura and Sire. In Tanzania, 49 out of the 58 isolates tested were *grh* positive and were found in all the districts (Table 3.13). Five *M. oryzae* isolates from Kisii, Kisumu and Siaya districts of Kenya and an isolate (U46) from Hoima district of Uganda tested *grh* positive in PCR amplification method. Majority of the *grh* positive isolates were obtained from varietal mixtures and local landraces of finger millet. Isolates K4 and K8 from Kisumu and Siaya districts in Kenya were obtained from an improved variety MS60D. Overall contemporary collection of isolates from Eastern Africa revealed a near equal distribution of the *grh* positive (49.1 %) and *grh* negative (50.9 %) *M. oryzae* isolates (Figure 3.18A) based on the PCR screening method. Based on the PCR screening analysis, *grh* positive isolates are dominant in Ethiopia and Tanzania (~86 %), while the *grh* negative isolates are dominant (~ 94 %) in Kenya and Uganda (Figure 3.18B).

Table 3.13. Distribution pattern of the *grh* positive and *grh* negative *M. oryzae* isolates collected from finger millet in Eastern Africa

Country/ Region	District	Number of isolates*	Grasshopper (*grh*) element PCR	
			Positive	Negative
East Africa	**Total**	**224**	**110**	**114**
Ethiopia	Sire	1	0	1
	Waju Tuka	4	4	0
	Adet	1	0	1
	Diga	10	9	1
	Lalo Assabi	3	3	0
	Nedjo	5	5	0
	Guangau	3	3	0
	Banja	5	5	0
	Bahir Dar Zuria	6	5	1
	Bure	2	2	0
	Demecha	1	1	0
	Angebo	1	1	0
	Leta Sibu	1	1	0
	Bila	1	1	0
	Boji Bermeji	1	1	0
	Gimbi	1	1	0
	Leka Dulecha	3	3	0
	Dangla	1	1	0
	Dure Bete	4	3	1
	Mandura	1	0	1
	Mecha	4	4	0
	Ankussa-Abdo Gor	1	1	0
	Jabi Tana	2	0	2
	Qilxxu Kara	1	1	0
	Total	**63**	**55**	**8**
Tanzania	Nkasi	10	10	0
	Sumbawanga	5	5	0
	Mbozi	10	8	2
	Njombe	21	17	4
	Madaba	6	3	3
	Momba	6	6	0
	Total	**58**	**49**	**9**
Kenya	Siaya	5	1	4
	Kisumu	6	1	5
	Busia	13	0	13
	Bungoma	2	0	2
	Kisii	14	3	11
	Kakamega	1	0	1
	Eldoret	3	0	3

	Marakwet	1	0	1
	Total	**45**	**5**	**40**
Uganda	Arua	4	0	4
	Apac	3	0	3
	Amuria	2	0	2
	Alebtong	4	0	4
	Amuru	3	0	3
	Agago	2	0	2
	Lira	3	0	3
	Katakwi	2	0	2
	Mbale	1	0	1
	Ngora	4	0	4
	Gulu	2	0	2
	Kitgum	2	0	2
	Lamwo	1	0	1
	Masindi	1	0	1
	Kumi	6	0	6
	Moyo	1	0	1
	Pallisa	2	0	2
	Manafwa	4	0	4
	Hoima	3	1	2
	Serere	4	0	4
	Tororo	4	0	4
	Total	**58**	**1**	**57**

*Total number of *M. oryzae* isolates collected during 2015 - 2017. Further information about the isolates characterised as PCR Grh+ and Grh- are presented in Appendices 4 to 7.

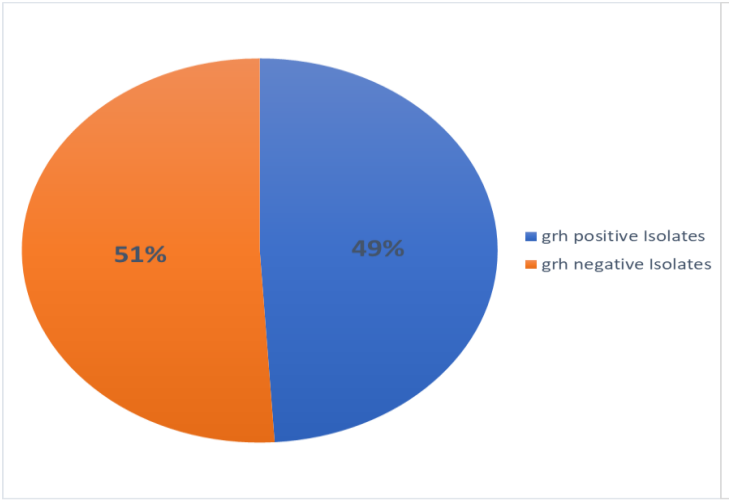

Figure 3.18A Proportion of the *grh* positive and *grh* negative *M. oryzae* isolates associated with finger millet in East Africa

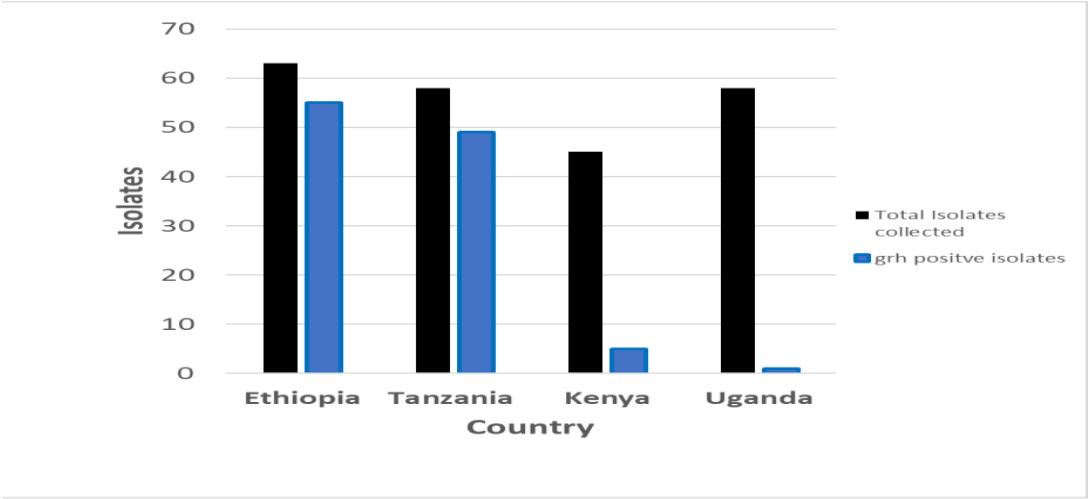

Figure 3.18B Pattern of occurrence of *grh* positive *M.oryzae* isolates in the four countries surveyed

3.4 Discussion

The work reported in this chapter utilized a distinctive collection of 300 isolates of the blast pathogen *M. oryzae*, representing historic (2000 - 2004) and contemporary (2015 - 2017) populations, from finger millet and related hosts in Eastern Africa. These isolates were from Ethiopia (63), Kenya (77), Tanzania (58) and Uganda (102) representing key cropping locations (districts). The finger millet blast (FMB) pathogen populations were analysed using a set of DNA sequence-based molecular markers identified and/or developed in this study. Isolates representing the contemporary populations were screened for the *Grasshopper* (*grh*) element, which is known to be mainly present in FMB pathogen isolates in certain geographic locations.

In the initial identification and/or development of the DNA sequence-based molecular markers, genome sequence data from 49 isolates of *M. oryzae* from various host systems including rice blast available in the NCBI database were extremely useful. This resource enabled a rapid and efficient evaluation of the level of nucleotide differences in any selected locus among *M. oryzae* isolates from a single host. To my knowledge, this is the first study to harness this resource for this specific purpose, particularly in the context of the FMB pathogen population analysis. Among the three different bioinformatics-based approaches evaluated to identify and/or develop the molecular markers, the FUNYBASE resource was not effective and the limited number of loci identified for testing by PCR and sequencing were not suitable for further development. This is in contrast to recent reports where FUNYBASE genes were an effective source to develop novel markers for major fungal pathogens such as *Alternaria alternata* and *Fusarium proliferatum* from diverse hosts (Armitage *et al*., 2015; Almiman, 2018). This is because vast majority of the single copy

orthologs represented in the FUNYBASE showed minimal or no differences among *M. oryzae* isolates reflecting the highly conserved nature of various parts of the genome in this fungus.

Evaluation of the known fungal phylogenetic markers enabled the identification of the internal transcribed spacer (ITS) and the histone 4 (HIS4) gene as markers yielding a good level of resolution with the FMB system. The Internal transcribed spacer (ITS) has been widely used as a universal DNA barcode marker for diverse range of fungal species including *M. oryzae* from rice and eleusine blast (e.g. Takan *et al*., 2012; Xu, 2016; Nilsson *et al*., 2017; Lear *et al*., 2018; Jagadeesh *et al*., 2018a). In this study, ITS sequences distinguished the 300 FMB pathogen isolates from Eastern Africa into 7 genotypes (ITS-G1 to ITS-G7) with the phylogenetic tree displaying close relationship of isolates represented by genotypes ITS-G2 and ITS-G4 isolates at 95.8% confidence level (Figure 3.4). The level of SNPs among the *M. oryzae* genotypes (Figure 3.3) was limited apart from genotype ITS-G6 that showed significant differences to others (Table 3.4). Low variability of ITS sequences among the Eastern African FMB pathogen isolates is similar to the pattern reported recently among the Kenyan rice blast pathogen isolates where 138 isolates were grouped into 5 clades and each of the clades were only separated by 3 to 4 nucleotide bases (Mwongera, 2018). Similar results have been reported for other fungal pathogens such as *Fusarium proliferatum* (Mule *et al*., 2004; Visentin *et al*., 2009; Almiman, 2018) and *Colletotrichum acutatum* (Talhinhas *et al*., 2002; Baroncelli, 2012). The HIS4 marker developed from the histone 4 locus in this study, covers both the exon and intron part of the gene (Figure 3.19). The HIS4 marker provided more resolution in differentiating FMB pathogen isolates in Eastern Africa compared to the ITS marker. The HIS4 sequence distinguished the FMB pathogen isolates into 9 genotypes (Table 3.5), among which the HIS-G8 was distinctive with 32 SNPs (Figure

3.5). The other 8 genotypes displayed close phylogenetic relationship at 86.9 % confidence value (Figure 3.6).

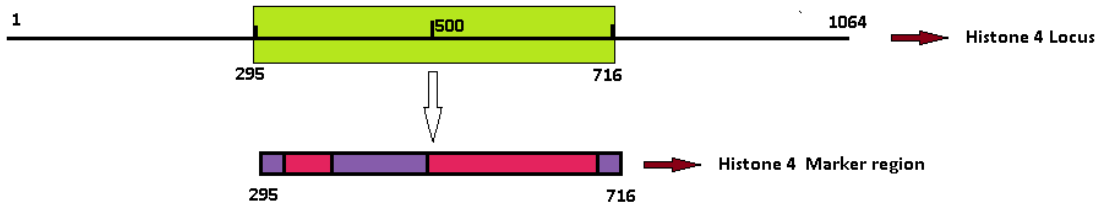

Figure 3.19. Schematic diagram of histone 4 locus and histone 4 marker region

The diagram illustrates the Histone 4 locus with 1064 bases within which the Histone 4 (HIS4) marker region (green) covering both exon (purple) and intron (red) regions spanning nucleotide bases 295 to 716 is shown.

To my knowledge, this is the first use of HIS4 gene sequences for the population analysis of *M. oryzae* from a single host, although histone 4 sequences have been used in other fungal pathogens such as *Fusarium* and *Colletotrichum* species (e.g. Glass and Donaldson, 1995; Donaldson *et al*., 1995; Talhinhas *et al*., 2002).

Intensive analysis of the variable chromosomes of *M. oryzae* led to the identification of two novel markers HyP1 and HyP2 from different genomic regions, in this study. These markers provided nearly 10-fold higher resolution in differentiating the FMB pathogen isolates compared to ITS and HIS4. It is pertinent to highlight that, markers from SNP-rich genomic regions provided a similarly higher resolution among fungal pathogen *Ustilagoinoidea virens* isolates compared to phylogenetic markers from conserved regions/housekeeping genes (Sun *et al*., 2013). Further detailed evaluation using BLAST analysis revealed that the two novel markers HyP1 and HyP2 were from regions of the *M. oryzae* genome reflecting the typical characteristics of genomic islands known to cause chromosomal rearrangements.

These genomic islands are known to be rich in pathogenicity genes including effector genes and repeat elements (Dobrindt *et al*., 2004). The HyP1 marker contains partial sequence of a hypothetical protein (HyP) and telomere. A candidate *AVR-Pita2* gene is present at the upstream of the HyP1 marker, whilst the downstream part includes telomere region (Figure 3.20). In the case of the HyP2 marker, non-LTR (MoTeR1) and LTR (Inago1) retrotransposon repeat elements flank the downstream, whilst the up-stream region (approximately 2 kb) has no homology to known sequences in *M. oryzae* and/or other fungal genomes (Figure 3.21).

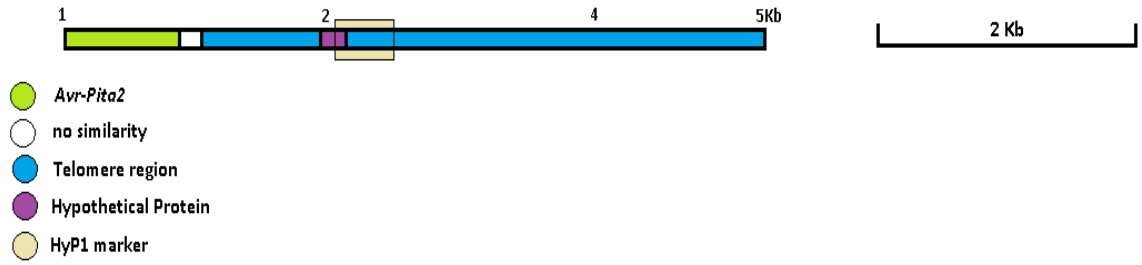

Figure 3.20. Schematic diagram representing the architecture of the genomic island harbouring the HyP1 marker in *M. oryzae*

Following the successful PCR amplification of the HyP1 marker region and subsequent sequencing, BLAST analysis was used to identify and extract a corresponding 5 kb region in the genome of the *M. oryzae* isolate K23/123 (one of the reference genomes developed in this work presented in Chapter 4). Further bioinformatic analysis enabled the determination of the genomic architecutre of this region containing the HyP1 marker sequence including the hypothetical protein (purple), telomere region (blue), uncharacterised nucleotide sequence with no similarity in the databases (white) and a candidate *AVR-Pita2* gene (green).

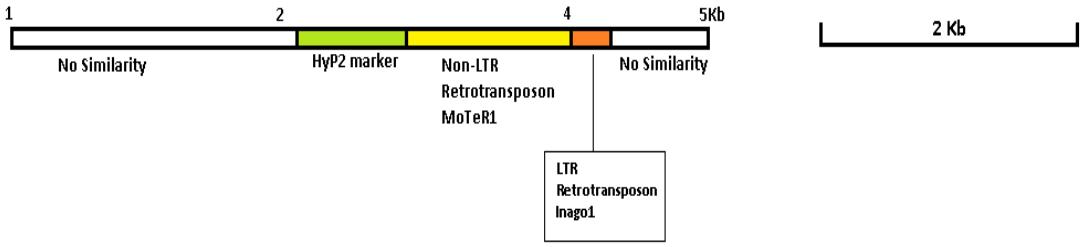

Figure 3.21. Schematic diagram representing the architecture of the genomic island harbouring the HyP2 marker in *M. oryzae*

Following the successful PCR amplification of the HyP2 marker region and subsequent sequencing, BLAST analysis was used to identify and extract a corresponding 5 kb region in the genome of the *M. oryzae* isolate K23/123 (one of the reference genomes developed in this work presented in Chapter 4). Further bioinformatic analysis enabled the determination of the genomic architecutre of this region containing the HyP2 marker sequence including uncharacterised nucleotide sequence with no similarity in the databases (white) and repeat elements MoTeR1 and Inago1, known in *M. oryzae* (Table 1.1)

Thus, the two novel markers identified (HyP1 and HyP2) represent genomic islands that are dynamic regions rich in transposable elements and repetitive sequences in the FMB pathogen *M. oryzae*. Recent and emerging research in the rice blast and other systems has highlighted that these genomic hotspots are critically important in the adaptive divergence of the fungal pathogens to diverse host and environmental conditions (Starnes *et al*., 2012; Santana *et al*., 2014; Faino *et al*., 2016; Bialas *et al*., 2018).

Among the four markers used in this study, ITS and HIS4 revealed a comparable level of resolution distinguishing the 300 FMB pathogen isolates from Eastern Africa into 7 to 9 genotypes. Likewise, the level of resolution provided by the HyP1 and HyP2 markers was comparable with 85 and 80 genotypes, respectively. This distinctive pattern of genotype diversity revealed by these two sets of markers is reflective of the two-speed nature of the *M. oryzae* genomes (Dong et al., 2015; Yoshida et al., 2016; Faino *et al*., 2016). Markers such as ITS and HIS4 represent the highly conserved areas of the genome including the house

keeping genes. Markers such as HyP1 and HyP2 represent the highly variable areas of the genome rich in transposable and repeat elements as well as effector genes including avirulence genes. Furthermore, both HyP1 and HyP2 markers distinguished the genotypes representing the FMB pathogen in Eastern Africa into two sub-populations represented by Groups A and B (Figure 3.8 and Figure 3.10), which were geographically clustered. These markers were also useful in deciphering the close relatedness of certain genotypes, e.g. within Group A, genotypes HyP1-G5, HyP1-G64, HyP1-G65 and HyP1-G66 from Kenya were closely related supported by 99.9% confidence value with the HyP1 marker. Further, Group A mainly included isolates from Kenya and Uganda, whilst in Group B, majority of the isolates were from Ethiopia and Tanzania. With the ITS and HIS4 markers, this pattern was not readily apparent, although these and similar markers such as ACT and MPG1 have been widely used in characterising various crop fungal pathogens (e.g. Couch *et al*., 2005; Jagadeesh *et al*., 2018a). This highlights the value of the novel markers HyP1 and HyP2 developed in this study in gaining an in-depth understanding of the genotype diversity, their distribution pattern and the overall population structure of the FMB pathogen in Eastern Africa.

Multi-locus sequence (MLS) analysis based on the concatenated sequences of HIS4, ITS, HyP2 and HyP1 (total 2303 bases) provided a refined assessment of the FMB pathogen populations in Eastern Africa, relative to the single locus analysis. MLS data distinguished FMB pathogen isolates in Eastern Africa into 207 genotypes, whilst the single locus phylogenies revealed 7 (ITS marker) to 85 (HyP1 marker) genotypes. MLS phylogenies also revealed some interesting trends in the distribution pattern of the genotypes in space and time. For example, 15 genotypes contained isolates from Kenya and Uganda including 4

genotypes that contained isolates representing both historic and contemporary populations in these countries (Table 3.8A). Conversely, a considerable number of genotypes (single and shared) were restricted to one country only, e.g. 33 genotypes in Tanzania and 65 in Uganda (Tables 3.8A – 3.8B). It is pertinent to mention that the MLS analysis also distinguished 2 distinct sub-populations of the FMB pathogen in Eastern Africa (Figure 3.12), corresponding to the pattern recorded with the HyP1 and HyP2 markers in single locus analysis. Furthermore, MLS analysis of country level data was also effective in distinguishing phylogenetic groups within each country ranging from 2 groups in Uganda to 4 groups in Ethiopia (Figure 3.13 - 3.16).

Overall, the MLS analysis is more informative than the single locus analysis in deciphering the genetic diversity, population structure and phylogenetic relationships of the finger millet blast pathogen *M. oryzae* in Eastern Africa. Further, the MLS analysis has also enabled the integration of sequence data from loci displaying varied evolutionary rates. This is consistent with the patterns reported in other fungal pathogens such as *Alternaria*, *Colletotrichum* and *Fusarium species* based on MLS analysis (Talhinhas *et al*., 2002; Armitage *et al*., 2015; Baroncelli *et al*., 2015; O'Donnell et al., 2015; Almiman, 2018). However, it is important to highlight the value of appropriate single locus marker-based analysis depending on the research aim and also in view of the resource efficiency. For instance, using the 9 genotypes identified by the HIS4 marker (Table 3.5) as a framework, specific sets of isolates representing various geographic locations for each genotype can be selected to analyse their virulence spectrum or aggressiveness levels in initial varietal screening assays. Likewise, in the next stage screening, any promising varieties can be tested against selected isolates represented by the wider range of genotypes identified by a marker such as HyP1.

Equally, the high resolution offered by single locus analysis with HyP1 maker would make it a suitable candidate for future monitoring of FMB pathogen population dynamics in locations where the baseline data is available.

Finger millet blast pathogen population structure at a regional level in Eastern Africa, based on the genotype diversity and distribution, shows a continuous genetic variation pattern reflecting varying degrees of genotype flow across countries. AFLP and SSR analyses also showed similar pattern of the finger millet blast pathogen in Kenya and Uganda (Takan *et al*., 2012) as well as in India (Babu *et al*., 2013b). The continuous genetic variation pattern of the FMB pathogen populations in Eastern Africa as well as in India is distinctive to the clonal lineage-based structure well documented with the rice blast pathogen in various geographic locations (Levy *et al*., 1991; Takan *et al*., 2012; Wang *et al*., 2017; Pagliaccia *et al*., 2018). Blast pathogen diversity, reflected by the number of lineages and/or haplotypes, has directly been related to the history of crop cultivation in a geographic location. For example, several reports have shown that *M. oryzae* populations were highly diverse in countries with a long history of crop cultivation as in the case of rice in Asia (Silva *et al*., 2009; Le *et al*., 2010; Saleh *et al*., 2014; Zhong *et al*., 2018; Gladieux *et al*., 2018a), and wheat in Brazil (Castroagulin *et al*., 2017). Thus, in India and China, where rice has been cultivated for thousands of years, 157 and 381 haplotypes of the rice blast pathogen *M. oryzae*, respectively have been reported (Kumar *et al*., 1999; Chen *et al*., 2006). On the contrary, in diverse locations (e.g. USA, Colombia, Europe, Iran, and West Africa), where rice cultivation is relatively recent (~200 to 500 years) dominated by a limited range of blast resistant varieties, only 2 to 8 lineages of *M. oryzae* have been reported (e.g. Consolo *et al*., 2008; Park *et al*., 2008; Takan *et al*., 2012; Wang *et al*., 2017b; Pagliaccia *et al*., 2018). A very

similar pattern is emerging with the wheat blast system and 198 *M. oryzae* genotypes have been identified in Brazil, which has a long history of wheat cultivation and viewed as the centre of domestication (Castroagulin *et al*., 2017). Conversely, in Bangladesh, where wheat blast outbreak has recently been reported, a single clonal lineage of the pathogen has been identified (Islam *et al*., 2016).

The varying patterns of association of the blast pathogen genetic diversity with the cultivation history of rice and wheat discussed above are particularly relevant to the genetic diversity and population structure of the finger millet blast pathogen in Eastern Africa. This is because, it is now well recognised that Eastern Africa - more specifically the highland stretches across Ethiopia and Uganda, is the centre of origin and domestication of finger millet with ~5000 years of cultivation history (Hilu *et al*., 1979; de Wet *et al*., 1984). Ethiopia, Kenya, Tanzania and Uganda selected for this investigation are regions of high crop diversity through traditional cultivation of landraces and/or varietal mixtures for many centuries and also the cultural exchange of genetic materials by the ancient farmers. It is also important to note recent efforts at the development and introduction of improved varieties in the region (Manyasa *et al*., 2015; De Villiers *et al*., 2015; Gimode *et al*., 2016; Tesfaye and Mengistu, 2017; Lule *et al*., 2018). These factors are likely to have strongly influenced the FMB pathogen population diversity and structure in Eastern Africa reported in this study, as proposed by Zhao *et al*. (2010) in the wider context of the centre of origin and domestication of crops and pathogen evolution.

Further, the Bayesian as well as reticulate median-joining phylogenies based on MLS analysis have distinguished the FMB pathogen in Eastern Africa into two geographically clustered sub-populations designated Groups A and B (Figures 3.11 and 3.12, Table 3.8C).

Group A mainly included isolates from Kenya and Uganda, whilst Group B predominantly included isolates from Ethiopia and Tanzania. Formal and/or informal seed exchanges and movement of planting materials between countries in the Eastern African region are known (Lenne *et al*., 2007; Louwaars and de Boef, 2012). This has been suggested as a key mechanism for the spread of crop pathogens (Tanaka *et al*., 2009; Zhao *et al*., 2010) including wheat blast (Ceresini *et al*., 2018). However, the specific pattern of geographic clustering of the two sub-populations of the FMB pathogen in Eastern Africa identified in this study is intriguing and requires further investigations of the plant genetic background, because Ethiopia and Tanzania are geographically separated by Kenya. This brings into consideration the complex history of the genetic background and movement of finger millet, which originated from and was first domesticated in Eastern Africa ~ 5000 years ago and then introduced into Asia ~3000 years ago (Hilu *et al*., 1979; de Wet *et al*., 1984).

Within finger millet, African highlands race and Afro-Asiatic lowland race are recognized and it is considered that the African highlands race was derived from *E. africana* under cultivation giving rise to the African lowland race. Subsequently, Afro-Asiatic lowland race developed from the African lowland race following its migration to India (Mathur *et al*., 2012). Cultivated finger millet *Eleusine coracana* includes two sub-species *E. coracana* subsp. *Africana*, and *E. coracana* subsp. *coracana* (Werth *et al*., 1994; Gimode *et al*., 2016). Collaborative seed exchange programmes are considered to have led to the reintroduction of *E. coracana* subsp. *coracana* back into Eastern Africa from Asia (Werth *et al*., 1994; Gimode *et al*., 2016). These two-sub species of *E. coracana* are genetically distinct (Werth *et al*., 1994; Gimode *et al*., 2016) including 4 groups within subsp. *africana* and 5 groups within subsp. *coracana* identified by genotyping (Gimode *et al*., 2016). This leads to the hypothesis

that the two sub-populations of the FMB pathogen in Eastern Africa identified in this study have co-evolved and are associated with the cultivation and material exchange patterns of the two types of finger millet across the region (Hilu *et al*., 1979; de Wet *et al*., 1984; Werth *et al*., 1994; Gimode *et al*., 2016). Ongoing work on the genetics and pathology of finger millet varieties by the project collaborators is likely to shed further light on this hypothesis. Moreover, a recent study has reported two distinct lineages of *M. oryzae* associated with *Eleusine* species based on genomic analysis of a small number of samples from different geographic locations (Gladieux *et al*., 2018). Potential occurrence of two groups of the FMB pathogen has also been suggested previously (Tanaka *et al*., 2009). Given the widely accepted view that Eastern Africa is the centre of origin of finger millet, from where it was introduced and spread to other parts of the world, how the two sub-populations of the finger millet blast pathogen in Eastern Africa identified in this study relate to the two lineages reported from wider geographic locations requires further in-depth investigations.

Finger millet blast pathogen diversity at the country level showed some interesting trends, taking into careful consideration of the sampling bias due to the varied number of isolates available from each country. In each of the four countries surveyed, MLS analysis revealed high genotype diversity. For instance, the number of *M. oryzae* genotypes present in Tanzania and Ethiopia were 35 and 39 from 58 and 63 isolates characterised, respectively (Tables 3.9A and B to 3.10A and B). Relative to these, a higher proportion of *M. oryzae* genotypes were identified in Uganda with 80 genotypes detected from 102 isolates (Tables 3.12A and B), whilst the highest proportion of genotypes was in Kenya with 71 genotypes identified from 77 isolates (Tables 3.11A and B). The higher proportion of genotypes in Uganda and Kenya reflects that the historical and contemporary populations analysed in

each of these countries are relatively distinctive. It is important to note that genotypes such as UG-G4 and UG-G10 in Uganda contained isolates from both historical and contemporary populations (Table 3.12A). Likewise, it is also pertinent to note that some *M. oryzae* genotypes represented isolates from 2 or more countries (Table 3.8A).

Bayesian analysis of the *M. oyzae* genotypes from each country enabled their clustering into a limited number of genetic groups revealing some interesting trends in the evolutionary relationship of these genotypes. For example, in Uganda, where 80 genotypes were identified, only two genetic groups were present. Among these, Group UG1 comprised ~95% of the *M. oryzae* isolates from a wide geographical area including historic and contemporary populations (Figure 3.16) with the nucleotide diversity ranging from 0 to 2.9 % (Table 3.12C). Conversely, in Ethiopia, the 39 genotypes present were clustered into 4 genetic groups (Figure 3.13) and the nucleotide diversity within Group ET1 ranged from 0 to 5.04 % (Table 3.9C). In Kenya, Group KN2 showed nucleotide diversity ranging from 0 to 7.9 % (Table 3.11C). Finger millet varietal development and consequently the grower cultivation patterns vary considerably across these countries. For example, Uganda Government has supported the development and introduction of improved varieties such as ENGENY and GULU E (Wanyera, 2005a and b) and some of these varieties have also been shared with Kenya. In Tanzania and Ethiopia, majority of the farmers are known to use landraces and varietal mixtures (Dida *et al*., 2008).

At this stage, adequate details of name and/or genetic background of the varieties with which these genotypes are associated are not available for various reasons including the growers' lack of awareness. However, as discussed above, the cultivation pattern of the finger millet material is most likely to have influenced the pathogen genotype diversity and

distribution as highlighted by the breeders and geneticists working on the characterisation of finger millet varieties in the region (Dida *et al*., 2008, Gimode *et al*., 2016).

Grasshopper (*grh*) is a repeat element present in *M. oryzae* isolates mainly associated with finger millet but previous research reported only a sporadic occurrence (e.g. only 13 isolates out of 328 screened) of this type of isolates in Eastern Africa (Dobinson *et al*., 1993; Takan *et al*., 2012). Results from this investigation revealed a very different picture with 110 out of 224 *M. oryzae* isolates collected across key finger millet production areas in Eastern Africa showing the presence of the *grh* element in a PCR based screening assay (Figure 3.18A). Results from the *grh*-screening assay have again highlighted the distinctiveness of the FMB pathogen populations from Ethiopia and Tanzania (104 *grh*+ isolates) compared to those in Kenya and Uganda (6 *grh*+ isolates; Figure 3.18B). The distribution pattern of *grh*+ isolates again is strikingly different, with wide distribution in Ethiopia and Tanzania (27 districts) and restricted occurrence in Kenya and Uganda (4 districts; Table 3.13).

Previous reports have suggested that the *grh*+ *M. oryzae* isolates in Eastern Africa might have been introduced through seed exchange programmes with Asia, where the finger millet blast pathogen isolates more commonly contain the repeat element (Dobinson *et al*., 1993; Viji *et al*., 2000; Takan *et al*., 2012). However, this study has shown, for the first time, wide spread occurrence of the FMB pathogen isolates containing the *grh* element in parts of Eastern Africa particularly in Ethiopia and Tanzania. Present results have broader significance both in terms of the differences among the countries in Eastern Africa and also in relation to the origin and spread of *M. oryzae* populations along with the movement of finger millet including research for development and trade (Kato *et al*., 2000; Tanaka *et al*., 2009). In view of these key issues, genome level differences in *M. oryzae* isolates

representing the *grh*+ and *grh*- populations of the finger millet blast pathogen populations in Eastern Africa have been further investigated in Chapter 4.

In conclusion, results reported this chapter have led to the development of new resources as well as to gain new insights of the finger millet blast pathogen genetic diversity, population structure and phylogenetic relationships as briefly summarised below. Novel markers HyP1 and HyP2 developed provide nearly 10-fold higher level of resolution of the FMB pathogen genetic diversity (80-85 genotypes) compared to the ITS and HIS4 markers identified (7-9 genotypes). An extensive assessment of the distribution pattern of the FMB pathogen genotypes at a regional level in Eastern Africa and in each of the 4 countries surveyed based on both single locus and multi-locus analysis. This is the first report of the identification of two geographically clustered sub-populations of the FMB pathogen in Eastern Africa based on both Bayesian and Reticulate Network analyses of historic and contemporary populations using 300 *M. oryzae* isolates. Finally, first report of the wide occurrence of the *M. oryzae* isolates containing the *grh* element in Eastern Africa particularly in Ethiopia and Tanzania. These exciting new observations provided the baseline for further genome-level investigations using 18 representative *M. oryzae* isolates in Chapter 4.

Chapter 4

4.0 Comparative Genomic Analysis of Finger Millet Blast Pathogen in Eastern Africa

4.1 Introduction

Genome comparison studies have become essential approaches in understanding the genetic diversity and evolutionary relationships in plant, animal and human pathogens (Li *et al*., 2014; Taylor, 2015; Baroncelli *et al*., 2016, McMullan *et al*., 2018; Gladieux *et al*., 2018). These approaches are emerging as a new gold standard for unravelling population diversity, speciation and cryptic processes at a high level of resolution allowing fine-grained epidemiological surveillance and evidence-based regulatory decisions (Gladieux *et al*., 2018). In addition, whole-genome analyses provide an avenue to gain novel insights into the genetic basis of life history traits, factors influencing the process of speciation and the signatures of speciation that are apparent from the patterns of variability in genomic architecture (Seehausen *et al*., 2014; Planet *et al*., 2017; Wolf and Ellegren, 2017). Several studies (e.g. Howlett *et al*., 2015; Teixeira *et al*., 2016; Sheppard *et al*., 2018) have shown that in-depth knowledge of the pathogen adaptive divergence is crucial for the development and deployment of effective disease control measures including resistant varieties.

Comparative genome analyses of plant pathogens have revealed uneven evolutionary rates across the genomes including the genes in the essential and conditionally dispensable (accessory) chromosomes (Dong *et al*, 2015; Wang *et al*., 2017b). These patterns have now been reported in the genomes of a range of crop pathogens such as *Magnaporthe oryzae*, *Leptosphaeria maculans*, *Zymoseptoria tritici*, *Phytophthora infestans* and *Fusarium*

graminearum (Raffaele *et al*., 2010; Stukenbrock *et al*., 2010; Grandaubert *et al*., 2014; Chiapello *et al*., 2015; Gowda *et al*., 2015; Yoshida *et al*., 2016; Wang *et al*., 2017b). These uneven evolutionary patterns led to the "two-speed genome" concept in which pathogen genomes have a bipartite architecture with the gene sparse, repeat-rich compartments serving as sites of adaptive evolution (Raffaele and Kamoun, 2012; Dong *et al*, 2015; Wang *et al*., 2017b; Faino *et al*., 2016).

This "two-speed genome" concept has been used to define the slow and fast evolving regions of the genomes. The slow speed sub-genome also referred as the core sub-genome has been recognised as essential for basic cellular functions. The fast speed sub-genome also referred as the flexible sub-genome constitutes repeat elements, mobile genetic elements and pathogenicity-related genes (Ogier *et al*., 2010; Croll and McDonald, 2012). Compared to the slow speed sub-genome, the fast speed sub-genome has a lower GC content, shorter gene length and higher variation of exon numbers. In addition, the flexible sub-genome is primarily considered as the hot spots for duplication, deletion and recombination that underlie rapid evolution through increased structural variation (Croll *et al*., 2015). Deciphering the fast speed sub-genome is crucial in understanding the pathogen evolution, adaptation and the processes involved in overcoming host immunity (Wang *et al*., 2017b). Phylogenomic investigations of pathogen populations or their representatives associated with a single host or multiple hosts offer a unique opportunity to identify the genomic features associated with host range expansion, host shifts and biogeographic relationships.

This study has focused on *M. oryzae,* causative agent of blast disease on finger millet particularly in Eastern Africa – the centre of origin and domestication of the crop.

Ascomycete fungal pathogen *M. oryzae* has been well studied in the context of rice blast disease and has emerged as a model system to understand plant-microbe interactions and co-evolution due to its economic significance and genetic tractability (Talbot, 2003; Dean *et al*., 2005). *M. oryzae* was the first plant pathogenic fungus sequenced to genome level (Dean *et al*., 2005). *M. oryzae* has now been assembled into 7 chromosomes with a total size of ~41.98 Mbp and used as a reference to generate a significant volume of genome data of the pathogen isolates. This landmark breakthrough was initially built on using Sanger capillary sequencing and subsequently NGS technologies such as sequencing by synthesis (SBS) and single molecule real time (SMRT) sequencing platforms (e.g. Xue *et al*., 2012; Chiapello *et al*., 2015; Gladieux *et al*., 2018; Zhong *et al*., 2018). Presently, genome sequence data of more than 200 isolates of *M. oryzae* from different host plants and geographical locations are available in global databases such as the National Centre for Biotechnology Information (NCBI) and the Joint Genome Institute (JGI).

Comparative analysis of genome sequences of *M. oryzae* isolates established differences in the genome architecture with an evolutionary trend that is mostly driven by the repeat-rich regions (Chen *et al*., 2013; Gowda *et al*., 2015). Another key advance is the recognition of the diversification of some effector proteins including avirulence (AVR) genes that function as factors modulating the host immune system and promote infection (Khang *et al*., 2008; Huang *et al*., 2014; Howlett *et al*., 2015). These effector proteins are not randomly distributed in the genome but tend to reside more in the regions enriched in repetitive sequences and mobile genetic elements (Rouxel and Balesdent, 2017; Wang *et al*., 2017a). Based on recent and emerging research, *M. oryzae* isolates possess a two speed genome, as recognised in some major filamentous crop pathogens, including the slow speed sub-

genome and fast speed sub-genome (e.g. Dong *et al*., 2015; Yoshida *et al*., 2016). These two sub-genomes have been analysed separately or wholly to gain insight on the evolutionary relationships in *M. oryzae* isolates causing blast disease on different hosts such as rice and wheat using various bioinformatic tools such as VCFtool, GATK, and ASTRAL (Danecek *et al*., 2011; Mirarab *et al*., 2014). Recent phylogenomic analysis of sequence data from slow and fast evolving regions distinguished 76 *M. oryzae* isolates from multiple cereals and grass hosts into distinct lineages associated with specific hosts (Gladieux *et al*., 2018). In a similar study, using whole-genome sequence data, 100 *M. oryzae* isolates associated with rice blast in various geographical locations were distinguished as 3 groups based on their phylogenetic relationships (Zhong *et al*., 2018). Based on this result, the researchers hypothesised that one of the groups might have given rise to the other two groups that emerged in parallel (Zhong *et al*., 2018). Initial phylogenetic analysis using ITS marker, and more recent research using a limited level of genome sequence data have suggested that two lineages of *M. oryzae* are associated with *Eleusine* species which includes the cultivated finger millet-*Eleusine corocana* (Tanaka *et al*., 2009; Gladieux *et al*., 2018). To my knowledge, however, in-depth and detailed genome level analysis of the finger millet blast pathogen *M. oryzae* and its biogeographic relationships have not so far been reported.

Results presented in Chapter 3 of this thesis have clearly identified 2 major Groups A and B among the finger millet blast FMB) pathogen populations in Eastern Africa including Kenya, Uganda, Ethiopia and Tanzania. Further, population level analysis of the FMB pathogen using 300 isolates and DNA sequence-based markers provided a clear picture of the pathogen genotype diversity and distribution pattern in space and time in Eastern Africa, which is the centre of origin and domestication of finger millet.

4.1.2. Aim and Objectives

The main aim of the work reported in this chapter is to gain an in-depth understanding of the genome level differences of the finger millet blast (FMB) pathogen *M. oryzae* in Eastern Africa and its relationship to blast pathogen isolates from finger millet in Asia as well as from other key hosts such as rice and wheat. The objectives are:

1. To develop the reference genome sequence for FMB pathogen isolates representing the Groups A and B.
2. To develop genome resequence data for another 16 *M. oryzae* isolates representing the biogeographic diversity.
3. To utilise the sequence data to carry out comparative analysis to gain novel insights into the genomic architecture of the FMB pathogen *M. oryzae*.
4. To identify genome level differences including the gene content and single nucleotide polymorphisms (SNPs)
5. To analyse the differences in the Grasshopper (*grh*) element amplification pattern representing the *grh* positive and *grh* negative *M. oryzae* isolates
6. To decipher the phylogenomic relationships among the blast pathogen from finger millet in Eastern Africa and Asia as well as other major hosts.

4.2 Experimental approaches

Eighteen *M. oryzae* isolates (Table 4.1) from finger millet in Eastern Africa were selected on the basis of the molecular characteristics including the genotype and *grh* positive or

negative, phylogenetic relationships, geographic location and year of collection from the 300 isolates (Table 2.1) characterised in Chapter 3. Selected isolates were cultured as described in Chapter 2 (Section 2.2.2 – Section 2.2.3). The mycelium was carefully harvested, freeze-dried and ground into a fine powder in liquid nitrogen with a sterile pestle and mortar (Section 2.3.1). Genomic DNA was extracted from the mycelial powder using the CTAB method for the preparation of genomic libraries and single molecular real-time (SMRT, PacBio) sequencing of the K23/123 and E34 isolates (Section 2.3.3). Genomic DNA extracted from mycelial powder using a column-based method (GenElute Plant Genomic DNA Miniprep kit- Sigma, UK) was used for the preparation of paired end and mate pair libraries and sequencing using Illumina sequencing platforms MiSeq600 or HiSeq4000 (Section 2.3.2). Prior to the preparation of the libraries, extracted genomic DNA was quantified using gel electrophoresis and NanoDrop methods to check the quantity, quality and integrity of each sample (Section 2.3.5). To develop the reference genome of isolates K23/123 and E34, initially, a combination of the paired end and mate pair libraries were sequenced using the Illumina MiSeq600 system (Section 2.6.2). The average length of DNA fragments used was 550 bp for the paired end and 2 to 4 kb for the mate pair libraries (Section 2.6.2). SMRT-bell libraries with 20 to 40 kb insert size were also prepared for the reference isolates and sequenced using SMRT (PacBio) sequencing technology (Sequel platform, Section 2.6.3). To generate the genome resequence data for the 16 isolates, paired end libraries were prepared with an average insert size of 100 to 150 bp and sequenced using the Illumina HiSeq 4000 system (Section 2.6.2). Details of the service providers utilised in the library preparation and sequencing processes are presented in Chapter 2.

Quality of the reads generated and elimination of the adaptor sequence were checked using the FASTQC program. Low quality bases of less than Q20 and adaptor sequences where present, were filtered using the BBDuk plugin in Geneious, Trimmomatic and NxTrim (Section 2.7.4). Curated reads were assembled with the SPAdes program for the Illumina sequencing, and the CANU program for the PacBio sequencing, with default settings (section 2.7.4). Reference assemblies integrating the Illumina and PacBio reads were optimised using different programs such as Pilon and Quickmerge to reduce nucleotide base errors to further enable accurate further analysis (section 2.7.4). The assemblies were screened for the presence of mitochondrial and any potential contamination sequences using BLASTN search and contaminated sequences were removed to have a nuclear genome for further analysis. Standard assembly statistics and quality parameters such as N50 values, number of contigs/scaffolds and the total genome size of the nuclear genome assemblies were evaluated using QUAST program. BUSCO program was used to further evaluate the genome sequence assembly quality by analysing the presence of the established set of fungal Ortholog genes (section 2.7.4). The syntenic regions between the two isolates used to develop the reference genomes were identified and mapped using SyMap (section 2.7.6). Complete gene sets were predicted using the Augustus program (v3.3) based on available data of *Magnaporthe grisea* as the model (Section 2.7.5). The distribution pattern of the protein-coding genes among the *M. oryzae* isolates was analysed using OrthoVenn program (section 2.7.5). For the SNP calling, curated sequence reads were aligned to the reference genome assembly of E34 isolate from this study with BWA (v0.7.5a). The alignment SAM files were converted and sorted into BAM files using SAMtools (v0.1.19). A combination of tools including SAMtools, BCFtools (v0.1.19) and VCFtools (v1.4.1) was used to filter the

bases in the alignment based on the quality thresholds as well as removing the heterozygous SNPs and Indels from the alignments. The final SNP data generated is based on the homozygous SNP calling (section 2.7.8).

For the phylogenomic relationship analysis, in addition to the 18 genome assemblies of the FMB pathogen isolates from East Africa generated in this study (serial numbers 1-18, Table 4.1), 9 genome assemblies available via NCBI were included (serial numbers 19-27, Table 4.1). The 9 additional genome assemblies included were from *M. oryzae* isolates from *Eleusine coracana* in Asia (India, 3 isolates), *E. indica* (1), *Setaria italic* (1), *Oryza sativa* (2), and *Triticum aestivum* (1) as well as the 70-15, a lab strain from which the first *M. oryzae* genome sequence was developed (Table 4.1). Single copy ortholog genes present in the genome assemblies were predicted and identified using BUSCO program and the sequence data of genes found common in all the assemblies were concatenated and used to construct a phylogenetic tree using the RAxML suite of programs (section 2.7.9). A single consensus tree was generated at the 80 % Bootstrap support values using the ASTRAL program and viewed using the Rpackage GGtree v1.12.4 (section 2.7.9). In addition to this, coding sequences of selected avirulence genes known in *M. oryzae* (Guy11) were searched from the genome assemblies of the 18 FMB isolates from this study using BLASTN analysis. Where present, the nucleotide sequences were aligned and the variation observed was checked for synonymous and nonsynonymous substitutions. The reference genome assemblies of isolates E34 (*grh* positive) and K23/123 (*grh* negative) were analysed using the original *Grasshopper* element described and the sequence data deposited in GenBank (Dobinson *et al*., 1993). BLASTN analysis of the two genome assemblies with the reference *grh* sequence (Dobinson *et al*., 1993) enabled the initial identification of the scaffolds showing varying

levels of coverage and identities and their detailed analysis using the options within the Geneious package (section 2.7.7).

Table 4.1: Details of the *M. oryzae* isolates selected for the genome sequencing and comparative analysis in this work

S/N	Isolate*	Host	Location	Collection Year	*Grasshopper (grh) element*	Genotype[a]	Group**	Accession No.	Reference
1	E1	*Eleusine coracana*	Sire, Ethiopia	2015	-	ML-G91	A		This study
2	E7	*Eleusine coracana*	Diga, Ethiopia	2015	+	ML-G96	B		This study
3	E17	*Eleusine coracana*	Angebo, Ethiopia	2016	+	ML-G105	B		This study
4	E34	*Eleusine coracana*	Diga, Ethiopia	2016	+	ML-G1	B		This study
5	K5	*Eleusine coracana*	Kisumu, Kenya	2016	-	ML-G128	A		This study
6	K13/67	*Eleusine coracana*	Teso, Kenya	2000	+	ML-G138	A		This study
7	K48/115n	*Eleusine coracana*	Kisii, Kenya	2002	+	ML-G173	B		This study
8	K23/123	*Eleusine coracana*	Busia, Kenya	2000	-	ML-G9	A	PHFK00000000	This study
9	K38	*Eleusine coracana*	Kisii, Kenya	2017	+	ML-G162	A		This study
10	T5	*Eleusine coracana*	Nbozi, Tanzania	2016	+	ML-G179	B		This study
11	T11	*Eleusine coracana*	Madaba, Tanzania	2016	+	ML-G180	NA		This study
12	T12	*Eleusine coracana*	Sumbawanga, Tanzania	2016	+	ML-G16	B		This study
13	T17	*Eleusine coracana*	Madaba Tanzania	2016	-	ML-G184	A		This study
14	D1/s11	*Eleusine coracana*	Pallisa, Uganda	2000	+	ML-G33	A		This study
15	D8/s15	*Eleusine coracana*	Busia, Uganda	2000	-	ML-G44	A		This study
16	D15/s47	*Eleusine coracana*	Soroti, Uganda	2002	-	ML-G54	A		This study
17	U13	*Eleusine coracana*	Katakwi, Uganda	2016	-	ML-G67	A		This study
18	U47	*Eleusine coracana*	Hoima, Uganda	2016	-	ML-G85	A		This study
19	MG03	*Eleusine coracana*	Bangalore, India	2012	+	NK	NK	LNTJ000000000	Shirke *et al.*, 2016
20	MG04	*Eleusine coracana*	Bangalore, India	2012	+	NK	NK	LNTK000000000	Shirke *et al.*, 2016
21	MG12	*Eleusine coracana*	Bangalore, India	2012	+	NK	NK	LNTO000000000	Shirke *et al.*, 2016
22	CD156	*Eleusine indica*	NK	NK	NK	NK	NK	UELY00000000	NK

23	US71	*Setaria italica*	NK	NK	NK	NK	NK	UCOF00000000	NK
24	Guy11	*Oryza sativa*	Guyana	1978	NK	NK	NK	MQOP00000000	NK
25	FJ81278	*Oryza sativa*	Fujian, China	1981	NK	NK	NK	MQOO00000000	NK
26	Mo70-15	Lab strain	NK	NK	NK	NK	NK	AACU00000000	Dean *et al*., 2005
27	BTJP4.1	*Triticum aestivum*	NK	NK	NK	NK	NK	UEMB00000000	NK
28	G22	*Eleusine coracana*	Japan	1976	+	na	na	M77662	Dobinson *et al*., 1993

Serial numbers 1-18 are isolates for which the genome sequencing was done in this study and 19-27 are genome assemblies downloaded via NCBI and used for comparative analysis

Grasshopper (*grh*) element: *M. oryzae* isolates identified as *grh* positive (+) or *grh* negative (-) based on the presence or absence of the two amplicons in PCR screening with the PKE and PES primer pairs

[a]The genotypes identified based on the multi-locus sequence data identity among the 300 isolates including historical and contemporary populations

*The *grh* element DNA sequence of the isolate G22 was downloaded and used for the identification and comparative analysis

**The major groups identified based on the phylogenetic analysis of the multi-locus sequence data in Figure 3.12 and Table 3.8C

NA, Isolate T11 was distinctive and was not included in either Group A or Group B; NK, Not Known

Mo70-15 is a lab strain developed by crossing the Guy11 isolate from rice and a weeping love grass isolate from which the first genome sequence data was generated (Dean *et al*., 2005).

4.3 Results

4.3.1 Genome sequencing, *de novo* assembly characteristics and gene prediction

Genome level data generation and analysis of the finger millet blast pathogen *M. oryzae* in Eastern Africa was carried out utilising a set of 18 isolates representing the genetic and biogeographic diversity of the 300 isolates characterised based on the analyses of the HIS4, ITS, HyP2 and HyP1 markers (Chapter 3). The two approaches pursued in this work were: i) generation of reference genome data utilising a combination of Illumina and PacBio platforms for 2 isolates representing the major phylogenetic groups A and B identified, and ii) genome resequencing using the Illumina platform of 16 isolates representing the two groups.

In the first approach, isolate K23/123 representing Group A and the *grh* negative isolates, and E34 representing Group B and the *grh* positive isolates were used to develop the reference genomes. The total number of Illumina (MiSeq600) and PacBio sequence reads generated for the two reference genomes were 13.74 and 16.91 million for isolates K23/123 and E34, respectively (Table 4.2). The Illumina reads length varied from 35 to 301 bases and the PacBio reads length varied from 1000 to 92,752 bases. Using the combination of raw data set from Illumina paired end and mate pair sequencing and the PacBio SMRT sequencing platforms resulted in a total sequence of approximately 8.21 billion bases (Bb) for K23/123 and 12.16 Bb for E34. Curated reads from the two platforms were assembled using SPAdes and CANU programmes along with other programmes such as Quickmerge to polish the assemblies. This led to the generation of reference genome assemblies comprising 21 and 31 scaffolds for isolates K23/123 and E31, respectively. The assembly resulted in a total genome size of approximately 42.58 Mb for isolate K23/123 and 44.15 Mb

for isolate E34. Other key statistical parameters generated and evaluated are suggestive of the high quality of the two *de novo* assemblies. Genome assembly quality assessment using BUSCO version 3.0.2 (Simao *et al*., 2015) with the Sordariomycetes ortholog gene data set (3725 genes) as reference showed the presence of the orthologs ranging from 3568 (95.7 % gene content) in isolate E34 to 3641 (97.8 % gene content) in isolate K23/123. To predict the total gene set for each isolate, the nuclear genomes were analysed using the Augustus pipeline with *Magnaporthe grisea* as the model and the total protein coding genes predicted were 10,692 in isolate K23/123 and 11,271 in isolate E34 (Table 4.2).

Table 4.2. Genome assembly's statistics and gene prediction information of the two reference *M. oryzae* isolates representing the two major phylogenetic group A and B

Features*	*M. oryzae* isolates**	
	K23/123	E34
No. of reads from PE library (M)	6.45	3.55
No. of reads from MP library (M)	10.11	9.26
No. of reads from SMRT library (M)	0.35	0.94
Total reads (M)	16.91	13.75
Total sequence generated (Bb)	8.21	12.16
Coverage (X)	192.69	275.45
Assembly size (Mb)	42.58	44.15
No. of Scaffolds	21	31
Largest Scaffolds (Mb)	7.30	12.35
N50 (Mb)	5.36	4.73
N75 (Mb)	4.14	3.28
L50	4	3
L75	6	6
BUSCO completeness (%)	97.8	95.7
No. of BUSCO genes	3641	3568
GC content (%)	49.92	50.17
Predicted protein coding genes	10692	11271

* PE – Paired end; MP – Mate pair; M – Million, Mb – Million bases, Bb – Billion bases
N50 – the length for which the collection of all contigs covers at least half (50%) an assembly size,
N75 - the length for which the collection of all contigs covers at least 75% an assembly size,
L50 & L75 – the number of contigs equal to or longer than N50 and N75, respectively.
BUSCO – Benchmarking Universal Single-Copy Orthologs
** Isolate details are presented in Table 2.1 and Table 4.1.

In the second approach, resequencing of 16 isolates representing the molecular and geographic diversity of the FMB pathogen populations in Eastern Africa was carried out using the lllumina HiSeq4000 system and PE libraries. The total number of reads generated for each isolate varied from approximately 15.18 (isolate K5) to 26.06 million (isolate K48/115n) (Table 4.3). A uniform read length of 100 bases was achieved and the total sequence data generated for each isolate varied from 3.04 billion bases (Bb) for isolate K5 to 5.21 Bb for isolate K48/115n. The *de novo* assembly of the curated reads was performed using SPAdes assembler at default setting and scaffolds less than 1 kb were filtered. The

overall coverage of the assemblies for the 16 isolates varied from ≥70X to 128X and the estimated size of the assembled genomes ranged from 40.66 Mb (isolate T12) to 43.75 Mb (isolate E7) as shown in Table 4.3. The number of scaffolds of the assembled genomes varied from 1367 (isolate E7) to 2188 (isolate K48/115n) as presented in Table 4.3 BUSCO analysis showed the presence of the gene orthologs ranging from 3603 (96.7 % gene content) in isolate K38 to 3618 (97.2 % gene content) in isolate U13 (Table 4.3). The predicted total protein coding gene set for each of the 16 isolates varied from 10,566 in isolate U47 to 10,778 in isolate K5 (Table 4.3).

Table 4.3: Genome assembly's statistics and gene prediction information of 16 resequencing *M. oryzae* isolates representing diverse geographical locations in Eastern Africa

FEATURES*	*M. oryzae* isolates**								
	K5	**K38**	**K13/67**	**K48/115n**	**D1/s11**	**D8/s15**	**D15/s47**	**U13**	**U47**
No. of reads from PE library (M)	15.18	21.63	24.53	26.06	20.94	21.02	23.05	25.47	23.23
Total reads (M)	15.18	21.63	24.53	26.06	20.94	21.02	23.05	25.47	23.23
Total sequence generated (Bb)	3.04	4.33	4.91	5.21	4.19	4.20	4.61	5.09	4.65
Coverage (X)	72.03	105.22	118.55	128.04	101.10	101.71	111.36	121.63	114.00
Assembly size (Mb)	42.15	41.11	41.39	40.70	41.42	41.33	41.39	41.88	40.75
No. of Scaffolds	2127	1869	1876	2188	1942	1886	1894	2037	1646
Largest Scaffolds (Mb)	0.248	0.225	0.193	0.159	0.225	0.171	0.237	0.221	0.520
N50 (Mb)	0.035	0.039	0.039	0.034	0.036	0.039	0.038	0.036	0.111
N75 (Mb)	0.020	0.021	0.021	0.019	0.020	0.021	0.021	0.020	0.051
L50	357	319	330	359	335	325	323	343	113
L75	758	679	692	763	713	693	687	733	243
BUSCO completeness (%)	96.8	96.7	96.9	96.8	97.1	97.1	97.0	97.2	97.1
No. of BUSCO genes	3604	3603	3608	3605	3616	3614	3612	3618	3616
GC content (%)	50.07	50.22	50.12	50.52	50.09	50.12	50.10	50.10	51.30
Predicted protein coding genes	10778	10657	10647	10740	10646	10639	10647	10706	10566

(Table 4.3 Continued)

FEATURES*	*M. oryzae* isolates**						
	E1	**E7**	**E17**	**T5**	**T11**	**T12**	**T17**
No. of reads from PE library (M)	21.39	21.85	24.68	23.08	20.13	21.35	21.63
Total reads (M)	21.39	21.85	24.68	23.08	20.13	21.35	21.63
Total sequence generated (Bb)	4.28	4.37	4.94	4.61	4.03	4.27	4.32
Coverage (X)	103.73	99.80	120.11	110.35	96.66	105.03	104.82
Assembly size (Mb)	41.23	43.75	41.11	41.83	41.64	40.66	41.26
No. of Scaffolds	1832	1367	2117	1932	2145	2163	1839
Largest Scaffolds (Mb)	0.224	0.671	0.170	0.206	0.167	0.320	0.180
N50 (Mb)	0.039	0.120	0.036	0.037	0.036	0.033	0.039
N75 (Mb)	0.021	0.053	0.019	0.021	0.020	0.018	0.021
L50	317	107	348	341	339	360	324
L75	674	242	750	720	731	774	681
BUSCO completeness (%)	97.0	97.0	96.9	97.0	96.8	96.8	96.9
No. of BUSCO genes	3614	3613	3609	3613	3607	3607	3610
GC content (%)	50.19	50.57	50.35	50.08	50.17	50.39	50.10
Predicted protein coding genes	10629	10594	10731	10724	10709	10698	10628

* PE – Paired end; M – Million; Mb – Million bases; Bb – Billion bases;
N50 – the length for which the collection of all contigs covers at least half (50%) an assembly size;
N75 - the length for which the collection of all contigs covers at least 75% an assembly size;
L50 & L75 – the number of contigs equal to or longer than N50 and N75, respectively;
BUSCO – Benchmarking Universal Single-Copy Orthologs
** Isolate details are presented in Table 2.1 and Table 4.1.

4.3.2 Genome synteny between finger millet blast pathogen isolates

Synteny plot analysis of the genome assemblies of K23/123 and E34 was carried out using the SyMap programme (Figure 4.1). Large scaffolds 1 to 8 of isolate K23/123 mapped to large scaffolds 1 to 7 of isolate E34 showing major syntenic blocks. Some of the smaller scaffolds of K23/123 (e.g. scaffold 9) also mapped to some of the large scaffolds in E34 (e.g. scaffold 2). Furthermore, some of the mapping patterns either with K23/123 (e.g. scaffold 4) or with E34 (e.g. scaffold 2 and 4) are suggestive of putative genomic rearrangements such as duplication and/or translocation. A number of scaffolds (11) in K23/123 did not map to any of the scaffolds in E34. Likewise, a number of scaffolds (20) in E34 did not map to any of the scaffolds in K23/123. These scaffolds were designated as putative isolate specific regions (Figure 4.1). The size of the regions specific to K23/123 and E34 is approximately 0.5 Mb and 1.6 Mb, respectively.

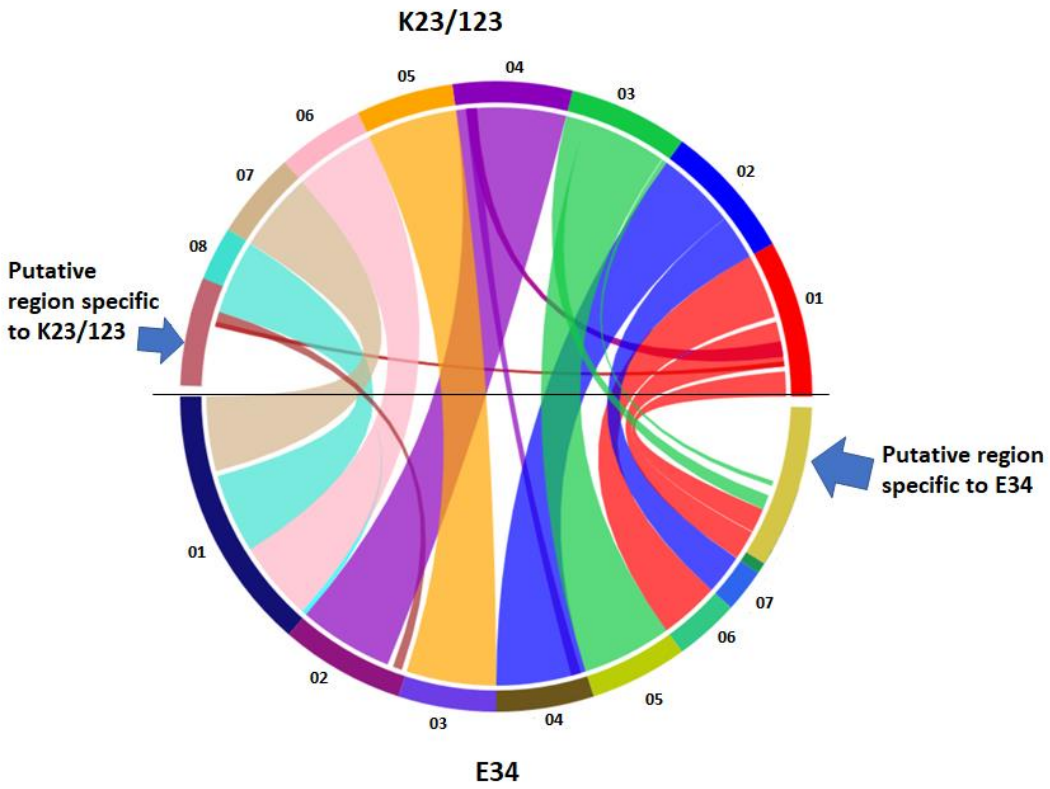

Figure 4.1. The synteny plot between K23/123 and E34 genomes obtained using SyMap

The circular plot shows mapping of some of the syntenic blocks between the scaffolds of K23/123 in the upper half with the scaffolds in E34 in the lower half of the circle. Syntenic mapping of the scaffolds is indicated by the corresponding colour blocks between K23/123 and E34.

4.3.3 Comparative analysis of the predicted protein coding genes among the finger millet blast pathogen *M. oryzae* isolates in Eastern Africa

The predicted protein coding genes identified in each isolate using the Augustus pipeline were further analysed with the OrthoVenn program to identify the distribution pattern of the genes among the 18 FMB isolates representing the pathogen populations in Eastern Africa. This enabled the identification of genes common to two or more isolates, and genes specific to an isolate. Based on the OrthoVenn analysis, 10216 protein coding genes were identified as common between *M. oryzae* isolates K23/123 and E34, whilst 19 and 37 genes were unique to isolates K23/123 and E34, respectively as shown in the Venn diagram (Figure 4.2).

OrthoVenn analysis of the 16 FMB pathogen isolates that were resequenced with isolates K23/123 and E34 used to develop the reference genomes enabled the identification of genes common among each resequenced isolate and the reference isolates as well as genes that are specific to each isolate (Table 4.4). For example, genes common among any one of the resequenced isolate and the two reference isolates ranged from 10,139 (isolate U47) to 10,173 (isolate D15/s47). The OrthoVenn programme initially identifies the genes common to each resequenced isolate and the two reference isolates. Subsequently, the programmes analyse the left-over gene set to identify genes common to the resequenced isolate and K23/123 or the resequenced isolate and E34 or K23/123 and E34 (Table 4.4). Finally, the number of genes specific to an isolate is identified, which varied from 0 to 38 among the 18 isolates including the 2 reference isolates (Table 4.4).

OrthoVenn analysis was also carried out amongst isolates from each of the 4 countries surveyed in Eastern Africa to assess the gene distribution pattern (Figures 4.3 – 4.6). For

example, 10,164 protein coding genes were common among the 5 FMB pathogen isolates from Kenya including the reference isolate K23/123, whilst the number of genes unique to each isolate varied from 0 to 7 (Figure 4.3). In addition to the 10,164 genes identified as common among the five isolates, some genes were identified as common among 4 isolates only (e.g. 14 to 109), some genes were identified as common among 3 isolates only (e.g. 11 to 28) and some genes were identified as common among two isolates only (e.g. 8 to 40) as shown in Figure 4.3. Among the 5 isolates from Uganda, 10,247 protein coding genes were common and the number of genes unique to each isolate varied from 0 to 2 (Figure 4.4). A further 11 to 95 genes were common among 4 isolates, 6 to 25 were common among 3 isolates and 6 to 29 were common among 2 isolates (Figure 4.4). In the case of the 4 FMB pathogen isolates from Tanzania, 10,212 genes common and the number of specific genes to each isolate varied from 0 to 5 (Figure 4.5). In addition, 14 to 182 genes were common among three isolates and 5 to 98 genes were common among two isolates (Figure 4.5). Similarly, among the 4 FMB pathogen isolates from Ethiopia including the reference isolate E34, 10,112 genes were common and the number of genes specific to each isolate varied from 0 to 22 (Figure 4.6). A further 26 to 124 genes were common among 3 isolates and 17 to 79 genes were common among 2 isolates (Figure 4.6).

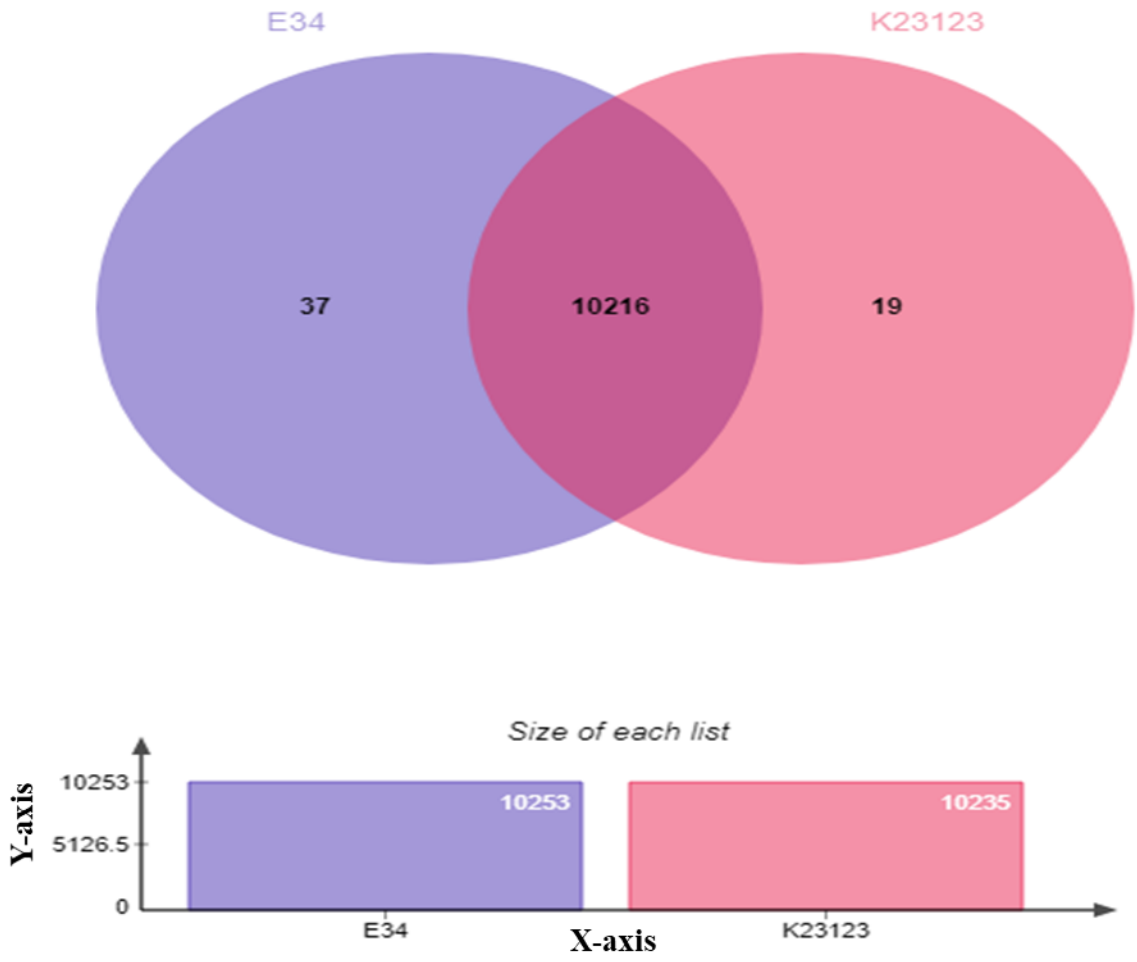

Figure 4.2. Distribution pattern of the predicted protein coding genes between isolates E34 and K23/123 used to develop the reference genomes as representatives of the major phylogenetic Groups A and B as well as the *grh* positive and *grh* negative isolates

Gene distribution pattern between the isolate E34 (purple) and isolate K23/123 (red) based on OrthoVenn analysis. The Venn diagram shows the clustered genes common to the two isolates (e.g. 10,216) and the number of genes specific to each isolate. The X-axis shows the isolates analysed and the Y-axis shows the number of clustered genes present.

The output was generated using OrthoVenn - http://www.bioinfogenome.net/OrthoVenn/start.php

Table 4.4 Distribution pattern of the predicted protein coding genes among the 16 resequenced finger millet blast pathogen isolates compared to the isolates K23/123 and E34 developed as reference genomes

Isolate	Distribution pattern of predicted protein coding genes						
	*Common to the isolate, K23/123 and E34	**Common to the isolate & K23/123 only	**Common to the isolate & E34 only	**Common to K23/123 & E34 only	Isolate specific	K23/123 specific	E34 specific
K5	10151	217	78	70	5	7	33
K38	10161	197	41	62	0	8	30
K13/67	10156	217	39	63	0	8	37
K48/115	10144	120	162	74	0	12	20
D1/s11	10159	226	40	59	0	5	36
D8/s15	10168	226	36	47	0	5	37
D15/s47	10173	222	39	50	0	6	35
U13	10169	213	57	53	0	7	35
U47	10139	184	50	82	1	14	38
E1	10159	210	41	59	0	5	32
E7	10147	91	164	70	0	18	27
E17	10165	115	174	53	3	11	21
T5	10144	210	65	72	1	8	23
T11	10143	204	68	75	0	9	30
T12	10151	108	173	67	0	12	27
T17	10148	207	34	74	0	6	35

Predicted protein coding genes of each isolate were clustered with those of K23/123 and E34 using OrthoVenn

*Column shows the number of genes common to each resequenced isolate and the reference isolates K23/123 and E34

** The left-over gene set is analysed by OrthoVenn progamme to identify genes common to each resequenced isolate and K23/123 only; each resequenced isolate and E34 only; K23/123 and E34 ony. The process enables the programme to identify genes specific to each resequenced isolate, K23/123, and E34.

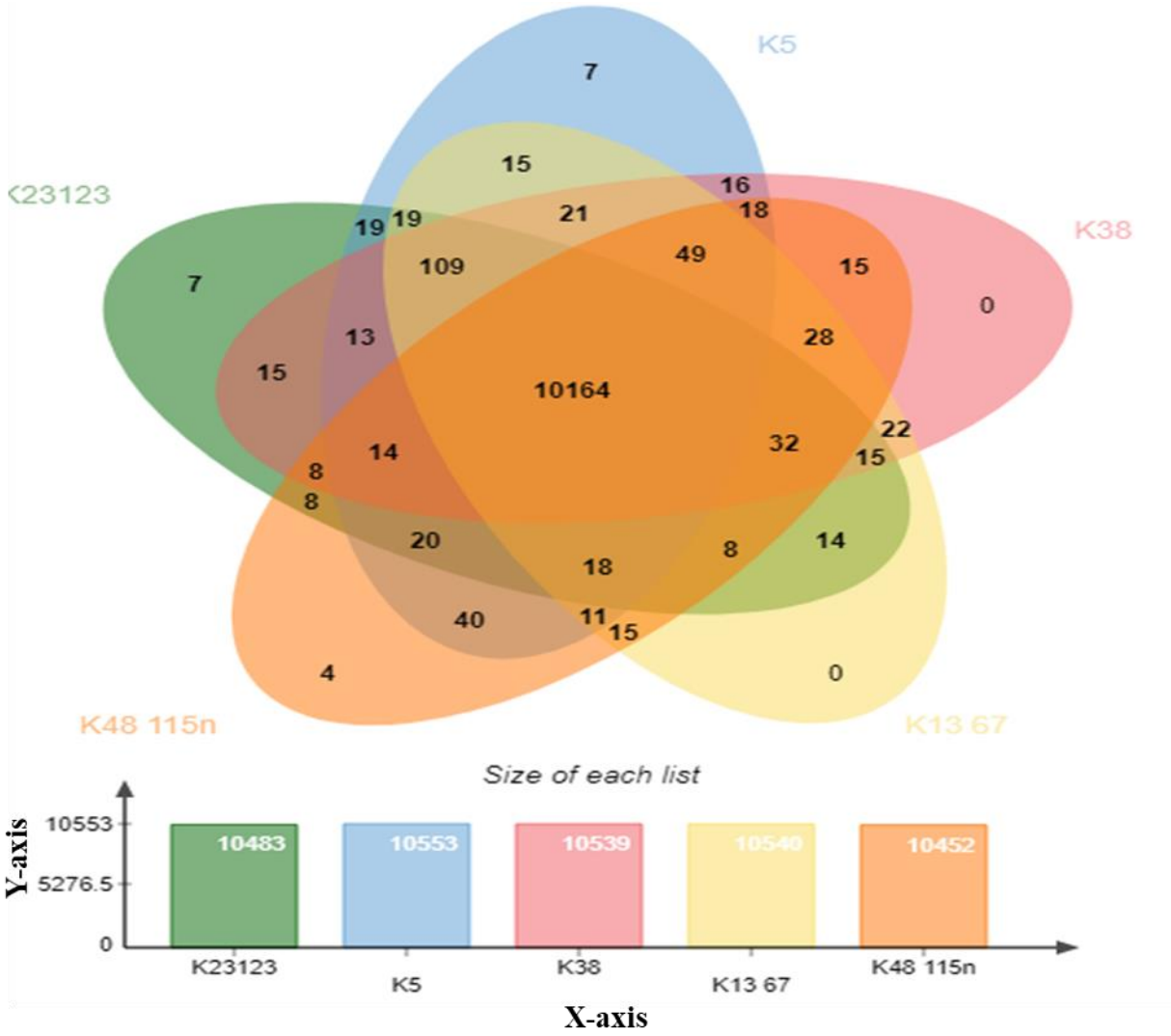

Figure 4.3. Distribution pattern of predicted protein coding genes among FMB pathogen isolates including reference isolate K23/123 from Kenya

Comparison of isolates K23/123 (green), K5 (blue), K38 (red), K13/67 (yellow) and K23/123 (orange). The Venn diagram shows the number of genes common to all 5 isolates (e.g. 10,164), genes common among 4 isolates, among 3 isolates and among 2 isolates, and the number of genes specific to each isolate. The X-axis shows the isolates analysed and the Y-axis shows the number of clustered genes present.

The output was generated using OrthoVenn - http://www.bioinfogenome.net/OrthoVenn/start.php

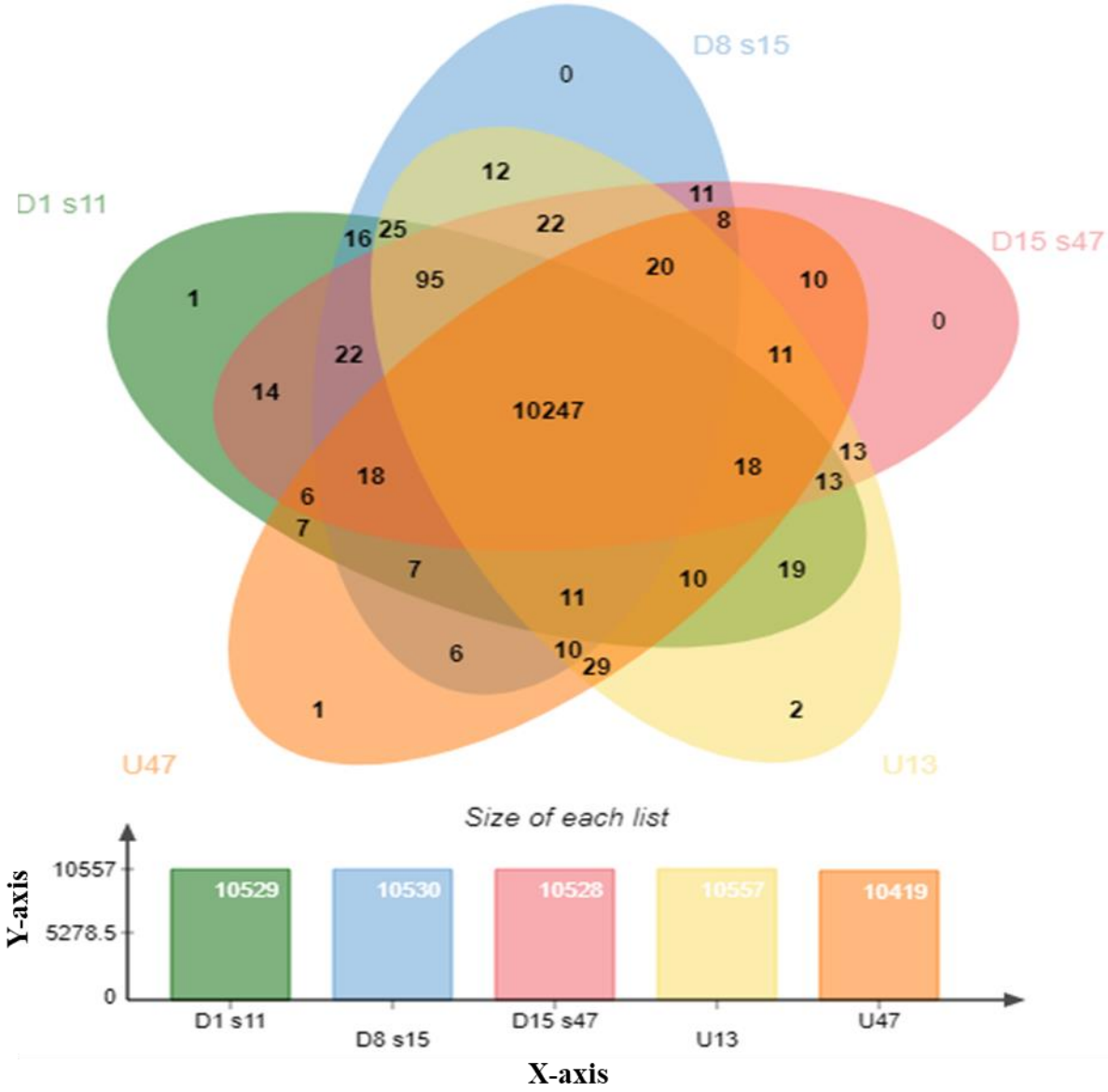

Figure 4.4. Distribution pattern of the predicted protein coding genes among FMB pathogen isolates from Uganda

Comparison of isolates D1/s11 (green), D8/s15 (blue), D15/s47 (red), U13 (yellow) and U47 (orange). The Venn diagram shows the number of genes common to all 5 isolates (e.g. 10,247), genes common among 4 isolates, among 3 isolates and among 2 isolates, and the number of genes specific to each isolate. The X-axis shows the isolates analysed and the Y-axis shows the number of clustered genes present.

The output was generated using OrthoVenn - http://www.bioinfogenome.net/OrthoVenn/start.php

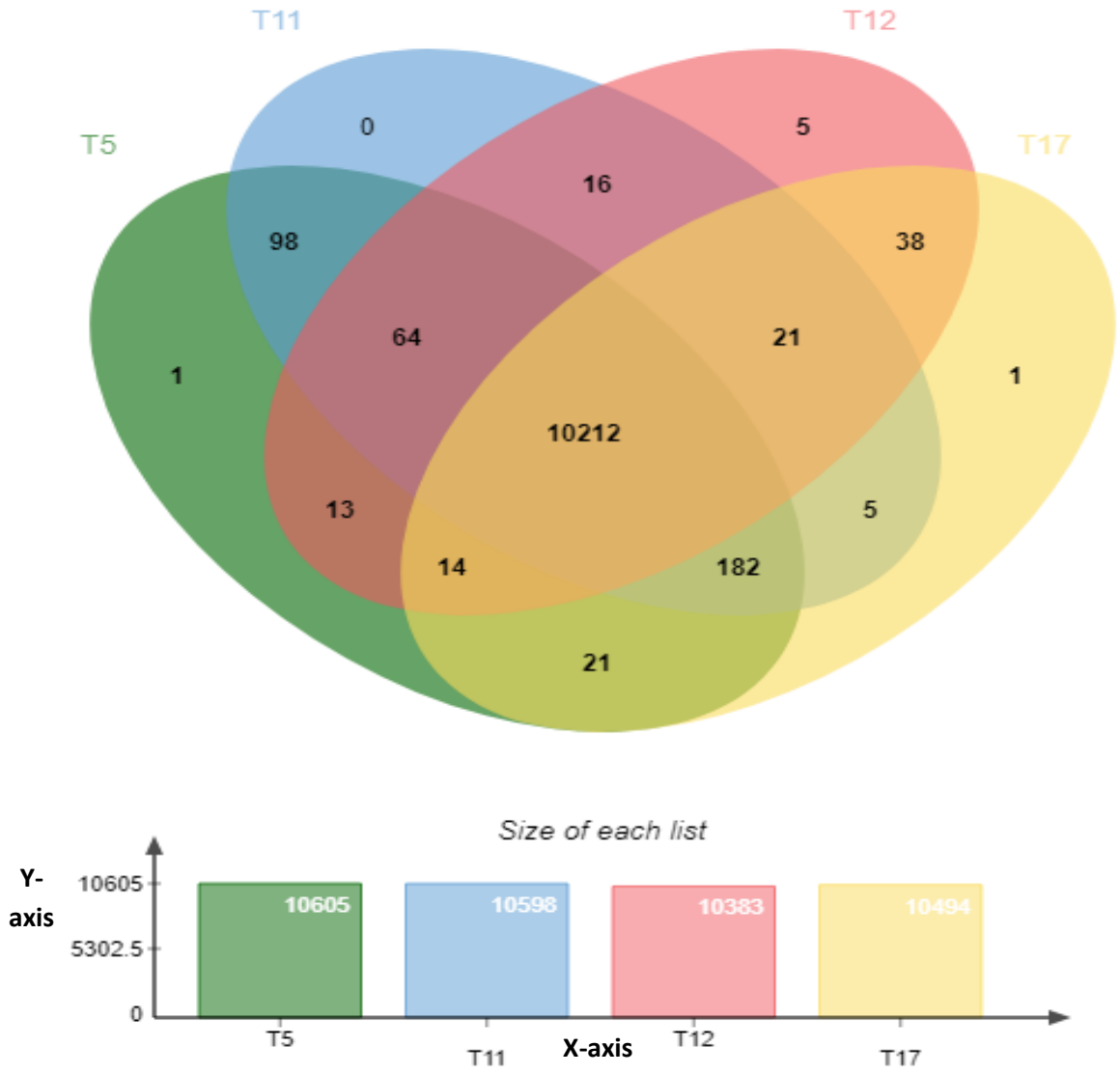

Figure 4.5. Distribution pattern of predicted protein coding genes among FMB pathogen isolates from Tanzania

Comparison of isolates T5 (green), T11 (blue), T12 (red) and T17 (yellow). The Venn diagram shows the number of genes common to all 4 isolates (e.g. 10,212), genes common among 3 isolates and among 2 isolates, and the number of genes specific to each isolate. The X-axis shows the isolates analysed and the Y-axis shows the number of clustered genes present.

The output was generated using OrthoVenn - http://www.bioinfogenome.net/OrthoVenn/start.php

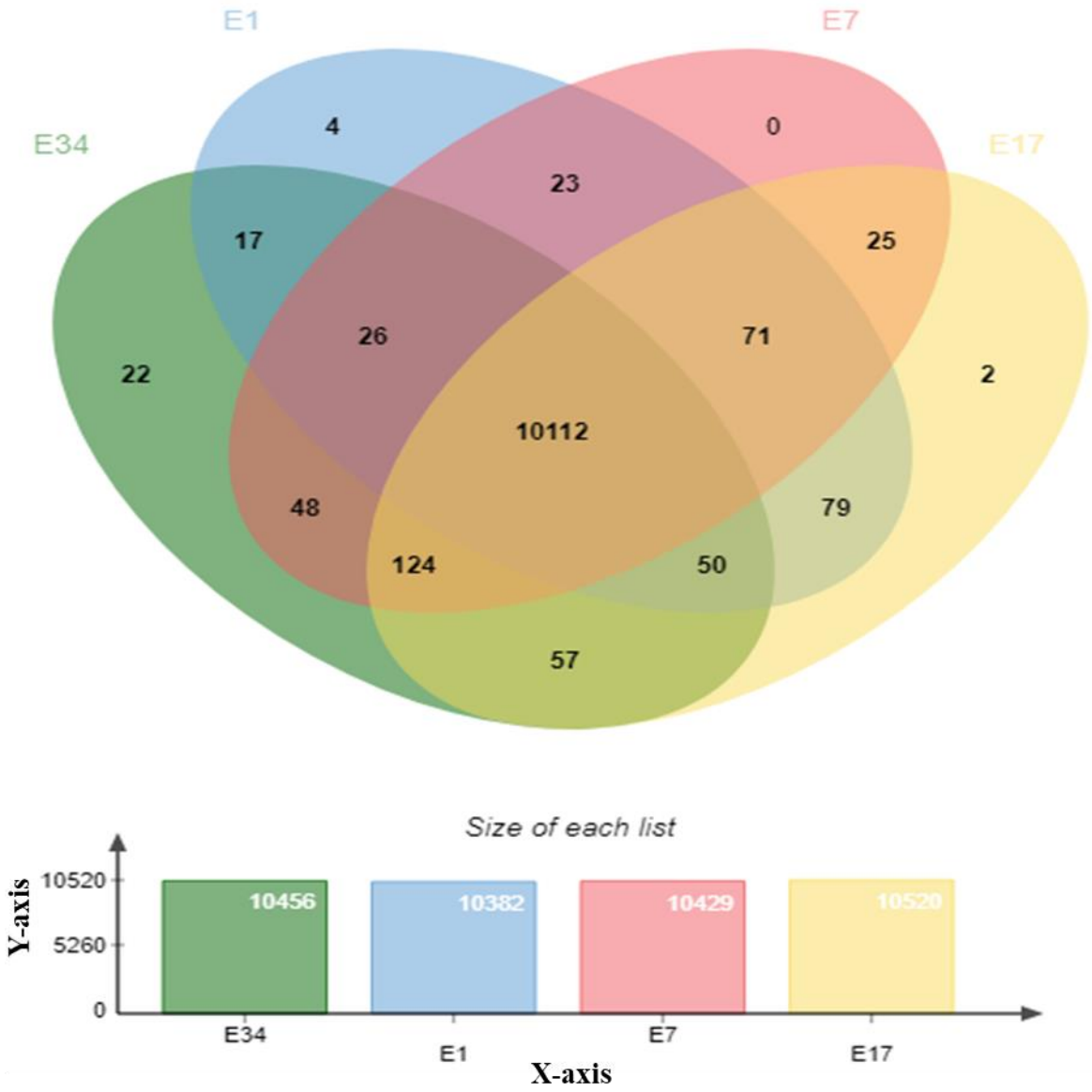

Figure 4.6. Distribution pattern of predicted protein coding genes among FMB pathogen isolates including the reference isolate E34 Ethiopia

Comparison of isolates E34 (green), E1 (blue), E7 (red) and E17 (yellow). The Venn diagram shows the number of genes common to all four isolates (e.g. 10,112), genes common among 3 isolates and among 2 isolates, and the number of genes specific to each isolate. The X-axis shows the isolates analysed and the Y-axis shows the number of clustered genes present.

The output was generated using OrthoVenn - http://www.bioinfogenome.net/OrthoVenn/start.php

4.3.4 Evaluation of the concept of genes common to all isolates, and genes specific to certain isolates using selected avirulence genes as a model

4.3.4.1 Distribution of the *AVR-Piz-T* gene in FMB *M. oryzae* isolates from Eastern Africa

The nucleotide sequence of the *AVR-Piz-T* deposited in the GenBank for Guy11 (isolate from rice) with the accession number LC175909.1 was extracted and used to interrogate the genomes of the 18 FMB pathogen isolates from Eastern Africa (Table 4.1) using BLASTN analysis. The BLASTN analysis showed that all 18 isolates harboured the *AVR-Piz-T* gene (327 bases) in their genomes reflecting a typical example of a gene common to all isolates as identified in the earlier parts of this Chapter (section 4.3.3).

The multiple sequence alignment of the *AVR-Piz-T* gene for the 18 isolates with the Guy 11 as reference shows the nucleotide substitutions (Figure 4.7). The nucleotide sequence variability resulted in changes in the amino acid sequence due to non-synonymous substitutions (Figure 4.8).

```
            30                                                        90
Guy 11      CCTCTTCACCGGGCTCGCCTCCGCCAGCTTCGTACAATGCAATCATCATCTCCTGTACAAT
Group FMB1  CCTCTTCACCGGGCTCGCCTCCGCCAACTTCGTACAATGCAATCATCATCTCCTGTACAAT
Group FMB2  CCTCTTCACCGGGCTCGCCTCCGCCAACTTCGTACAATGCAATCATCATCTCCTGTACAAT

            120                                                       180
Guy 11      GGCGGGTTGGGCCGTTAGATTTTACGAAGAAAAACCAGGGCAGCCAAAGAGGCTGGTCGCG
Group FMB1  CGCGGGTTGGGCCGTTAGATTTTACGAAGAAAAACCCAACCAGCCAAAGAGGCTGGTCGCG
Group FMB2  CGCGGGTTGGGCCGTTAGATTTTACGAAGAAAAACCAGACCAGCCAAAGAGGCTGGTCGCG

            181                                                       240
Guy 11      ATTTGCAAAAACGCGTCACCCGTACACTGCAACTATCTGAAATGCACCAATTTGGCAGCA
Group FMB1  ATTTGCAAAAACGCGTCACCCGTACACTGCAACTATCTGAAATGCACCAATTTGGCACCA
Group FMB2  ATTTGCAAAAACGCGTCACCCGTACACTGCAACTATCTGAAATGCACCAATTTGGCACCA
```

Figure 4.7. Multiple sequence alignment of the *AVR-PIZ-T* gene among the 18 *M. oryzae* isolates from finger millet in Eastern Africa with Guy 11 as reference

Nucleotide substitutions in the test isolates are shown in red with the corresponding bases in Guy11 shown in green. Only the regions that showed nucleotide substitutions are presented and the numbers above the alignment refer to base positions in the gene sequence.

Group FMB1 contained 17 isolates: K5, K38, K23/123, K13/67, D1/s11, D8/s15, D15/s47, E1, E7, E17, E34, T5, T11, T12, T17, U13, U47.

Group FMB2 contained 1 isolate (K48/115n).

Guy11 from rice blast in South America was used as the reference.

Further detail of the isolates is available in Table 4.1.

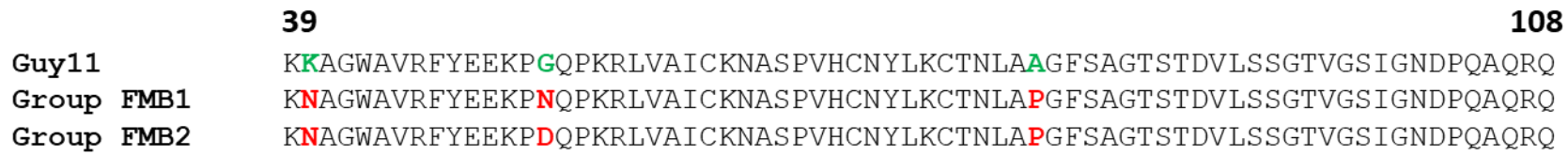

Figure 4.8. Multiple sequence alignment of amino acid encoded by AVR-Piz-T protein in *M. oryzae* isolates representing the 2 allelic groups identified among finger millet blast isolates in Eastern Africa with reference to data from isolate Guy11

Amino acid substitutions in the test isolates are shown in red and the corresponding amino acid in Guy 11 are shown in green. Only the regions showing amino acid substitution are presented. Numbers above the alignment refer to amino acid positions in the protein sequence.

Group FMB1 contained 17 isolates: K5, K38, K23/123, K13/67, D1/s11, D8/s15, D15/s47, E1, E7, E17, E34, T5, T11, T12, T17, U13, U47.

Group FMB2 contained 1 isolate (K48/115n).

Guy11 from rice blast in South America was used as the reference.

Further detail of the isolates is available in Table 4.1

4.3.4.2 Distribution of the *AVR-Pik* gene in FMB *M. oryzae* isolates from Eastern Africa

The nucleotide sequence of *AVR-Pik* deposited in the GenBank for Guy11 with the accession number AB498876.1 was extracted and used to interrogate the genome of 18 FMB isolates from Eastern Africa (Table 4.1) using BLASTN analysis. The BLASTN analysis showed that only 6 of the isolates contained the *AVR-Pik* gene (342 bases) in their genome reflecting a typical pattern of isolate specific genes.

The multiple sequence alignment of the *AVR-Pik* of the 6 isolates with Guy11 as the reference showed nucleotide substitution (Figure 4.9). The nucleotide sequence variability resulted in amino acid sequence changes due to non-synonymous substitutions (Figure 4.10).

```
            61                                                        120
Guy 11      GCCGAAACGGGCAACAAATATATAGAAAAACGCGCTATCGACCTAAGTCGAGAGCGAGAC
Group PiK1  GCTGAAACGGGCAACAAATATATAGAAAAACGCACTATCGACCTGAGTCGAGAGCGAGAC

            121                                                       180
Guy 11      CCTAACTTTTTCGACAACCCTGGTATTCCTGTACCCGAATGTTTTTGGTTTATGTTTAAA
Group PIK1  CCTAACTTTTTCGACCACCCTGGTATTCCTGTACCCGAATGTTTTTGGTTTATGTTTAAA

            181                                                       240
Guy 11      AACAACGTACGTCAAGATGATGGAACCTGTTACAGCTCTTGGAAAATGGACATGAAAGTT
Group PIK1  AACAACGTACGTCAAGATGCTGGAACCTGTTACAGCTCTTGGAAAATGGACATGCGAGTT
```

Figure 4.9. Multiple sequence alignment of the *AVR-Pik* gene in the 6 FMB *M. oryzae* isolates and the Guy 11 used as reference

Nucleotide substitutions in the test isolate are shown in red and the corresponding bases in Guy 11 are shown in green. Only the regions that showed nucleotide substitutions were extracted from the alignment and presented. Number above the alignment refer to base positions in the gene sequence.

Group PIK1 included the 6 FMB isolates from Eastern Africa: E7, E34, K5, K48/115n, U13, U47;

Guy 11 from rice blast in South America was used as the reference

Further details of the isolates can be seen in Table 4.1

```
            30                                                  80

Guy 11      KRAIDLSRERDPNFFDNPGIPVPECFWFMFKNNVRQDDGTCYSSWKMDMKV

Group PIK1  KRTIDLSRERDPNFFDHPGIPVPECFWFMFKNNVRQDAGTCYSSWKMDMRV
```

Figure 4.10. Multiple sequence alignment of the AVR-Pik protein in the 6 *M. oryzae* isolates from finger millet in Eastern Africa with corresponding data from Guy 11

Amino acid substitutions in the test isolates are shown in red and the corresponding amino acid in Guy 11 are shown in green. Only the regions of the amino acid sequence showing variation was extracted from the alignment and presented. Numbers above alignment refer to amino acid positions in the protein sequence.

Group PIK1 contained 6 isolates from *Eleusine coracana* in Eastern Africa: E7, E34, K5, K48/115n, U13, U47; Guy11, a rice blast isolate was used as reference.

4.3.5 Genome-wide variation of finger millet blast pathogen isolates in Eastern Africa based on single nucleotide polymorphisms

Genome-wide variation among the *M. oryzae* representing the blast pathogen populations in finger millet production systems in 4 Eastern African countries Kenya, Uganda, Ethiopia and Tanzania was evaluated based on the number of single nucleotide polymorphisms (SNPs). Curated Illumina PE sequence reads of seventeen isolates including K23/123 were mapped onto the ~44.15 Mb reference genome of the isolate E34. The homozygous SNPs between each isolate and the reference genome were identified utilising a suite of bioinformatics tools both for initial SNP calling and subsequent verification analyses (Chapter 2, Section 2.7.6). Indels and heterozygous SNPs were filtered to minimise differences due to misalignment or multiple regions alignment as the extracted genomic DNAs were from the haploid mycelium of *M. oryzae* isolates. The number of SNPs identified among the 17 isolates, compared to the reference genome of isolate E34, varied considerably ranging from 40,612 for isolate E7 to 200,949 for isolate K13/67. *M. oryzae* isolates within a country also displayed considerable differences in the number of SNPs, e.g. the range among the 4 isolates from Tanzania was 73,236 in isolate T12 to 197442 in isolate T17 (Table 4.5 and 4.6). The SNPs identified in this analysis reflect the genome-wide differences among *M. oryzae* isolates including the genic regions of both coding and non-coding as well as the intergenic regions.

Table 4.5. Single nucleotide polymorphisms (SNPs) identified among 17 finger millet blast pathogen isolates in comparison with isolate E34 used as the reference genome

S/N	Isolate code*	SNPs
1	E7	40,612
2	T12	73,236
3	K48/115n	79,837
4	E17	106,197
5	T11	173,656
6	U47	178,137
7	K23/123	185,409
8	U13	186,375
9	D1/s11	189,636
10	K5	190,144
11	D8/s15	190,823
12	K38	191,933
13	E1	196,140
14	T5	196,317
15	T17	197,442
16	D15/s47	199,587
17	K13/67	200,949

*Isolates code: K, Kenya; U & D, Uganda; T, Tanzania and E, Ethiopia

Further details of the isolates are included in Table 2.1 (Chapter 2)

Table 4.6. The range of single nucleotide polymorphisms (SNPs) identified among *M. oryzae* isolates from each country

Country	No. of isolates	Range of SNPs
Kenya	5	79,837 - 200,949
Uganda	5	178,137 - 199,587
Tanzania	4	73,236 - 197,442
Ethiopia	3	40,612 - 196,140

4.3.6 Phylogenomic analysis of *M. oryzae* isolates representing the finger millet blast pathogen populations in Eastern Africa based on ortholog genes in the genomes

Benchmarking Universal Single-Copy Orthologs (BUSCO) analysis was used to initially interrogate the genomes of 27 *M. oryzae* isolates from different hosts and geographical locations for the presence of the 3725 single copy gene set conserved in fungal species of the Sordariomycetes. The analysis showed that the BUSCO genes present in the genomes of the 27 *M. oryzae* isolates varied from 3568 to 3641. Of these, 944 genes were identified as phylogenetically informative among the *M. oryzae* isolates using the randomised accelerated maximum likelihood (RAxML) analysis (Stamatakis, 2014). A phylogenetic tree was constructed based on the concatenated alignment of the 944 single copy genes using the Accurate Species Tree Algorithm (ASTRAL, Mirarab *et al*., 2014). The phylogenetic tree branches with 80 % (= 0.8) or more bootstrap support value (BSV) are shown in black line and those with less than 80 % BSV are shown in dotted lines (Figure 4.11).

The tree topology distinguished the 27 *M. oryzae* isolates into two major clades. Among these, Clade 1 showed clear monophyletic relationship of finger millet blast (FMB) pathogen isolates from Eastern Africa and Asia (India) based on 80 % or more BSV. Clade 1 was further divided into '1A' and '1B', where Clade 1A comprised FMB pathogen isolates from Eastern Africa only and Clade 1B contained FMB pathogen isolates from India (Figure 4.11).

Clade 1A was further partitioned into two groups of Eastern Africa *M. oryzae* isolates from finger millet (1Ai and 1Aii). Clade 1Ai comprised 5 isolates from Uganda (U13, U47, D1/s11, D8/s15 and D15/s47), 4 isolates from Kenya (K5, K38, K13/67 and K23/123), 3 isolates from Tanzania (T5, T11, T17) and isolate E1 from Ethiopia. These included both *grh* positive (5) and *grh* negative (8) isolates. Clade 1Aii included isolates E7, E17 and E34 from Ethiopia,

isolate T12 from Tanzania and isolate K48/115n Kenya. These isolates were all *grh* positive. Within Clade 1Ai, some isolates showed a strong phylogenetic relationship based on their geographic location e.g. isolates T5, T11 and T17 from Tanzania with 80 % or more BSV (Figure 4.11).

Clade 2 contained isolates from various hosts with the 3 isolates (FJ81278, Guy11 and Mo70-15) from rice (*Oryza sativa*) forming a strong group (Clade 2Ai) distinct from the wheat blast isolate (BJTP4.1, *Triticum aestivum*) (Clade 2Aii). Clade 2B included the isolates from foxtail millet (US71, *Setaria italic*) and wild finger millet (CD156, *Eleusine indica*) as observed in Figure 4.11.

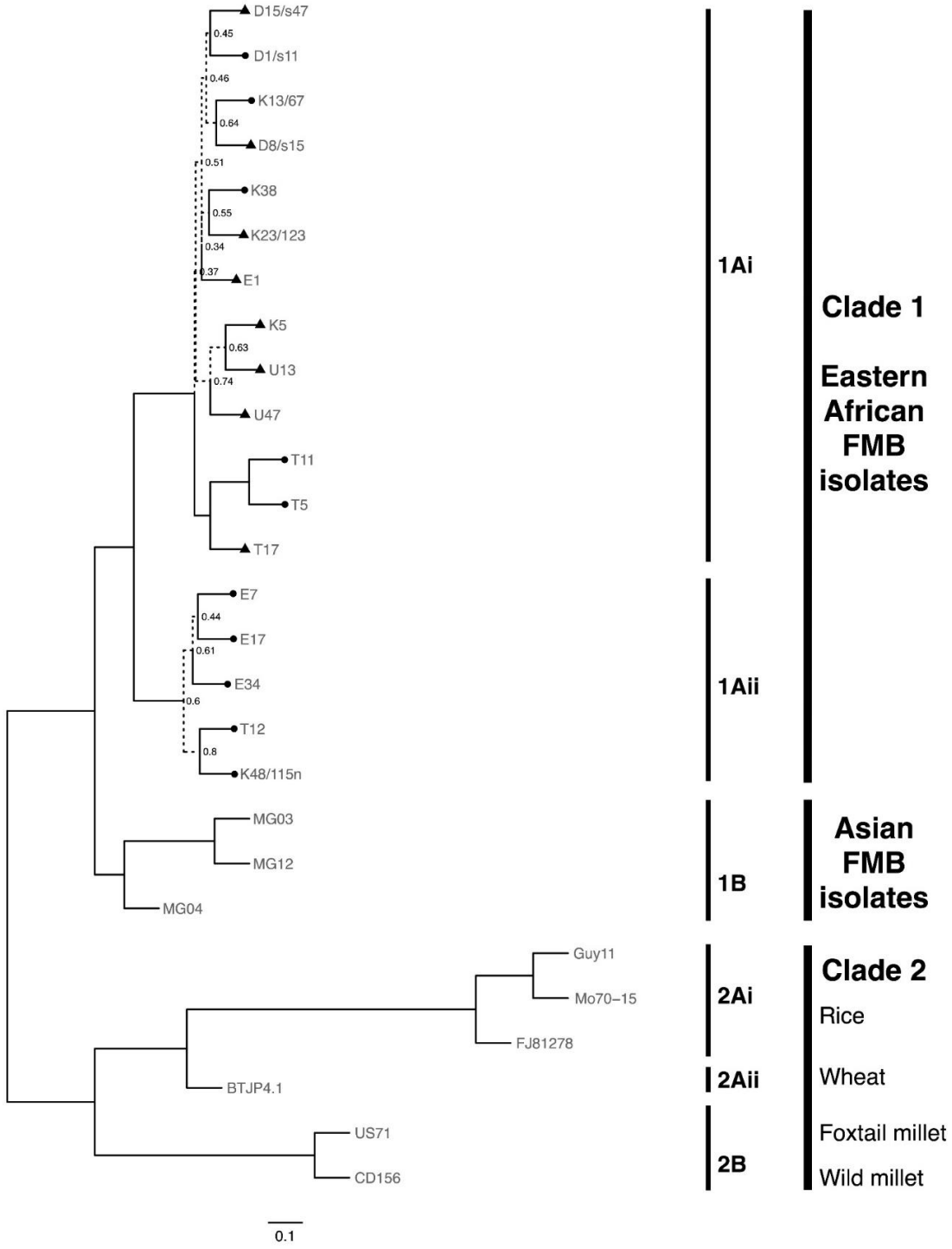

Figure 4.11. Maximum parsimony consensus phylogeny tree based on the 944-single copy ortholog genes

The tree reflects the diversity and phylogenetic relationships among 27 *M. oryzae* isolates from various hosts. Isolates from *Eleusine coracana* in Clade 1; isolates from *Oryza sativa, Triticum aestivum*, *Setaria italica* and *Eleusine indica* in Clade 2. Branches with 80 % (= 0.8) or more bootstrap support values (BSV) are shown in black lines, whilst branches with less than 80 % BSV are shown with dotted lines. Among the FMB isolates from Eastern Africa in Clade 1A, the circles and triangles at branch ends indicate *grh* positive and *grh* negative isolates, respectively.

4.3.7 Identification of the complete *Grasshopper* (*grh*) element in the Eastern African FMB pathogen and comparative analysis of the *grh* positive and *grh* negative isolates

The genome assemblies of isolates E34 and K23/123 were analysed for the presence of *grh* repeat element by BLASTN analysis using the previously reported 5233 bases of *grh* fragment sequence (Accession number M77662, Dobinson *et al*., 1993). In this study, for the first time, the full-length sequence of the *grh* element with 7610 bases and its typical structure was identified in the genome of isolate E34 isolate from Ethiopia (Figure 4.12A). The full *grh* element length was defined based on the identification of the nucleotide sequence of the 198-bp long terminal repeats (LTRs) at the 5' and 3' end of the *grh* element (Figure 4.12B). The LTRs were flanked by a short direct repeat (AAATA) at both ends in the genome of E34 (Figure 4.12B). The *grh* element included LTRs, gag protein encoded by 1047 bases, pol protein encoded by 4196 bases and an uncharacterised region of 2202 bases.

Comparative analysis of the genome assemblies of the isolates E34 (*grh* positive) and K23/123 (*grh* negative) provided very different patterns. In isolate E34, BLASTN analysis resulted in 6 hits corresponding to the number of copies of full-length sequence of *grh* element present in scaffolds 1, 5, and 7. Further, ~ 56 copies of the *grh* element varying in size ranging from 425 to 7000 bases were identified in E34 with nearly 28 copies in the 5000 to 7000 bases range (Table 4.9). On the contrary, genome assembly of K23/123 did not contain any copies of the full-length *grh* element. However, incomplete or partial sequence of the *grh* element was present and the maximum length identified was 5975 bases located in scaffold 5 of the K23/123 genome assembly. Other BLASTN hits in the genome of K23/123 ranged from 196 to 5000 bases (Table 4.7).

Comparative analysis of the *grh* element sequence from isolate E34 with the data from *M. oryzae* isolate G22 from finger millet in Japan from which the partial *grh* element sequence was originally reported (Dobinson *et al*., 1993) showed a high level of sequence identity (~ 99 %). On the contrary, comparative analysis of the ~5795 base sequence from the largest size fragment of the *grh* element in the genome of K23/123 with corresponding ~5975 bases of the *grh* sequence from E34 showed a high level of variation (Figure 4.13).

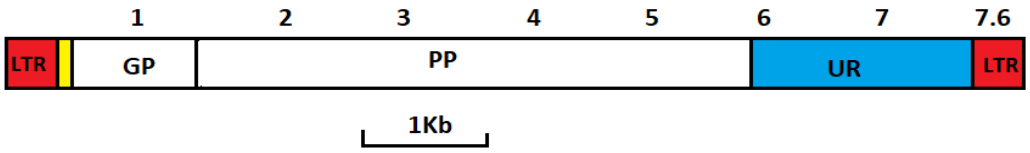

Figure 4.12A. Schematic representation of the structure of the full-length sequence of the *Grasshopper* (*grh*) element

Long terminal repeats (LTRs, red colour), the intergenic region (yellow colour) adjacent to the 5' LTR, the complete gag protein (GP), complete pol protein (PP), uncharacterised region (UR, blue colour). Each LTR is 198-nucleotide bases, intergenic region is 43 bases, GP is 1049 bases, PP is 4196 bases and the UR is 2202 bases. The *grh* element full sequence and structure was defined following the identification of the LTRs at both ends which were flanked by the short direct repeats as shown in figure 4.6B.

Figure 4.12B. Sequence of the Grasshopper (*grh*) element showing the flanking short repeats, long terminal repeats (LTRs) and some nucleotide bases internal to the LTRs

The flanking short repeats are in blue colour, the LTRs are in red colour, and nucleotide bases internal to the LTRs are in black colour.

Table 4.7. Distribution pattern of *Grasshopper* (*grh*) element copies and their size range in the genome of finger millet blast pathogen isolates E34 and K23/123 from Eastern Africa

	Size (bases) and number of copies of the *grh* element sequence in the genome			
Isolate	≥ 7000	6000 ≥ 6999	≥ 5000	< 5000
E34	6	5	23	28
K23/123	0	0	3	29

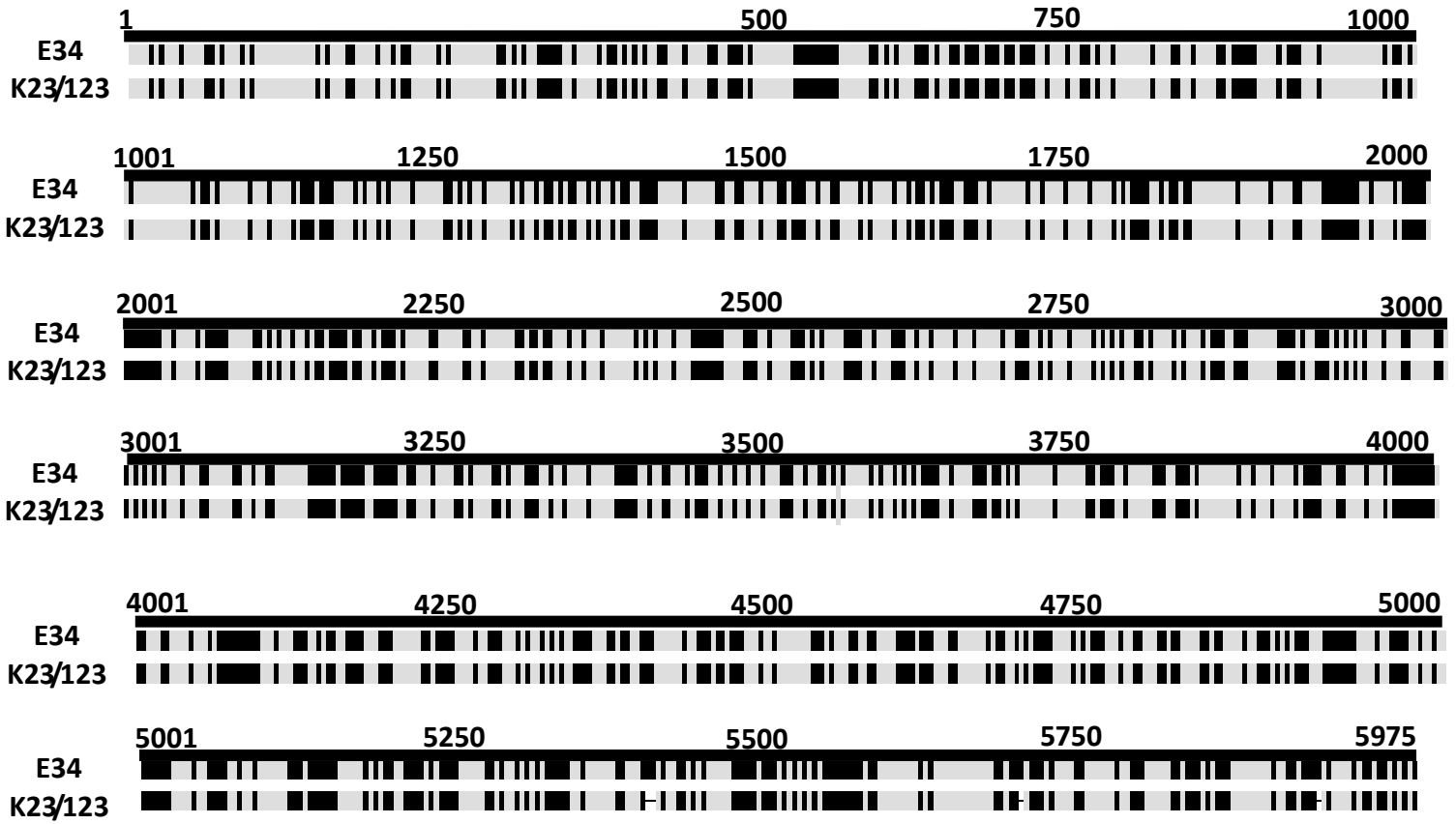

Figure 4.13. Schematic representation of the high sequence variation between the Grasshopper (*grh*) element in the genomes of isolates E34 (*grh* positive) and K23/123 (*grh* negative)

The black lines and bars reflect the level of nucleotide differences between the isolates; as noted in the results description, *grh* sequence in E34 showed high level of identity to the *grh* element originally described from FMB isolate G22 (Dobinson *et al*., 1993).

4.4 Discussion

Work reported in this chapter focused on the development and comparative analysis of the genome data from a distinctive set of 18 *M. oryzae* isolates (Table 4.1) representing the biogeographic diversity of the finger millet blast pathogen populations in Eastern Africa. The different DNA fragment libraries and the NGS platforms used proved effective in the genome level characterisation of the representative isolates. For instance, use of a combination of the Illumina MiSeq (PE and MP libraries) and PacBio Sequel (SMRT library) platforms was effective to develop the reference genome data of isolates E34 and K23/123 representing the two sub-populations identified by the MLS analysis. Likewise, the Illumina HiSeq platform using PE libraries was effective in resequencing the genomes of 16 isolates. This is reflected by the genome assembly parameters including the high levels of representation of the BUSCO genes ranging from 95.7 % to 97.8 % (Tables 4.2 and 4.3). The assembled genome size of the 18 isolates ranged from 40.66 Mb to 44.15 Mb, which is within the range of the assembled genomes size (35.8 Mb to 46.4 Mb) of a wide range of *M. oryzae* isolates (206, available at the NCBI database). Considerable differences in the genome size of *M. oryzae* isolates from diverse hosts and geographic locations have been reported (e.g., Xue *et al*., 2012; Shirke *et al*., 2016; Yoshida *et al*., 2016; Gladieux *et al*., 2018). Based on in-depth comparative genomic analysis of well-characterised isolates, the genome size variation reflects the differences in the genome evolution and architecture including the nature and distribution of repetitive regions and consequent genomic rearrangements as well as the presence and absence of specific components such as AVR genes (Yoshida *et al*., 2016; Bao *et al*., 2017). In this context, the NGS methodologies used

as well as the sequence coverage depth are important caveats before drawing firm conclusions (Dean *et al*., 2005; Bao *et al*., 2017; Gladieux *et al*., 2018).

This study is the first effort in establishing the reference genomes of the FMB pathogen isolates and the assembly scaffold levels achieved (21 to 31) is comparable to the reference genome of the rice blast isolate 70-15 (53 scaffolds). The genome assembly data for K23/123 has already been deposited at NCBI following successful completion of their validation tests (Accession Number PHFK00000000; Shittu *et al*., 2017, unpublished). Despite the major syntenic blocks identified between the E34 and K23/123 genome assemblies, the SyMap analysis revealed small regions suggestive of translocation in the genome of E34 and duplication in the genome of K23/123 (Figure 4.1). Similar patterns of genomic rearrangements (chromosomal translocations) have recently been reported between two rice blast isolates in India (Gowda *et al*., 2015) as well as in Japan (Bao *et al*., 2017). Presence of retroelements and transposable elements in these genomic regions has been considered to play a key role in mediating the translocations in *M. oryzae* isolates (Shnyreva, 2002; Gowda *et al*., 2015; Bao *et al*., 2017). The putative translocated regions identified in the *M. oryzae* isolate(s) in this study provide the basis to test whether similar mechanisms operate in the FMB pathogen.

Furthermore, the genomic regions identified as putatively specific to E34 (~1.6 Mb) and K23/123 (0.5 Mb) in the synteny analysis are equally interesting and important (Figure 4.1). This type of isolate-specific genomic regions has been strongly related to genes driving the host interaction patterns in the rice blast pathogen (e.g., Yoshida *et al*., 2009; Xue *et al*., 2012; Dong *et al*., 2015; Gowda *et al*., 2015). For instance, compared to the isolate 70-15 reference genome, isolate Ina168 contained a 1.68 Mb unique region containing 316

candidate effector genes (Yoshida *et al*., 2009) and isolate 98-06 had a similar 1.4 Mb unique region with 134 candidate effector genes (Dong *et al*., 2015). The putative isolate-specific genomic regions identified in this study presents an excellent platform for further comparative genomic and functional analysis of the effector/AVR genes, which is completely unexplored in the FMB pathogen at this time.

Prediction of the total gene set for each of the 18 isolates using Augustus programme (Stanke *et al*., 2006) and their comparative analysis provided an insight into potential gene content differences in the FMB pathogen (Table 4.2). The total gene content in these *M. oryzae* isolates ranged from 10,566 (isolate U47) to 11,271 (isolate E34), which is comparable to the total gene set predicted in the genomes of blast pathogen isolates from various hosts including rice (Gowda *et al*., 2015; Shirke *et al*., 2016; Yoshida *et al*., 2016). However, OrthoVenn analysis of the protein-coding genes in E34 and K23/123, predicted 19 genes and 37 genes putatively unique to each isolate, respectively in addition to 10,216 genes that are common (Figure 4.2). Similar comparative analysis of the protein-coding genes among the FMB pathogen isolates from the same country predicted up to 38 genes as potentially unique to an isolate (Table 4.3). The putative unique genes identified are likely to play key roles in determining the outcome of the host-pathogen interactions, which can be further analysed by functional annotation and analysis. Initial tests of this hypothesis carried out by comparative analysis of the 18 FMB pathogen isolates using two AVR genes well characterised in rice blast as models provided results reflecting differential gene content in various isolates. For instance, the *AVR-Piz-T* gene was present in all isolates, albeit in two allelic forms, which is an example of the genes common in the FMB pathogen (Figure 4.7). Conversely, the *AVR-Pik* gene was present only in 6 isolates, which is an example of genes

unique to certain isolates (Figure 4.9). The differences in the gene content of the FMB pathogen isolates recorded in this study aligns well with the emerging concept of the gene expansion and contraction (also referred as gene gain and loss), which has been strongly linked with the host range and pathogenic processes in major systems such as rice blast and anthracnose diseases caused by *Colletotrichum* species (Powell *et al*., 2008; Cantarel *et al*., 2008; Zhao *et al*., 2013; Morales-Cruz *et al*., 2015; Baroncelli *et al*., 2016). Availability of a whole range of computational algorithms such as InterProScan, RunIprscan and HMMER, and databases such as dbCAN and MEROPs is facilitating further investigations of the gene expansion and contraction in crop fungal pathogens (e.g. Morales-Cruz *et al*., 2015; Baroncelli *et al*., 2016).

Comparative analysis based on the homozygous SNPs identified for 17 *M. oryzae* isolates using E34 as the reference assembly revealed considerable variation in genome wide differences (SNPs ranged from 40,612 to 200,949) in the FMB pathogen in Eastern Africa (Table 4.5 and 4.6). *M. oryzae* isolates within a country also showed considerable differences based on SNPs as in the case Ethiopia (e.g. 40,612 - 196,140 SNPs; Table 4.6). In India, where finger millet is widely grown in some parts, 3 FMB pathogen isolates from the same district were found to be genetically distinct based on genome wide SNPs (Shirke *et al*., 2016). Similar patterns of genome wide differences based on SNPs have been reported with rice blast isolates from the same field in China (Xue *et al*., 2012), different districts in India (Gowda *et al*., 2015) and different regions in Kenya (Mwongera, 2018). Among the 17 *M. oryzae* isolates representing the FMB pathogen in Eastern Africa, 4 isolates (E7, T12, K48/115n and E17) showed an SNP range of 40,612 - 106,197. These isolates belonged to the one of the sub-populations (Group B) identified based on MLS analysis (Table 3.8C). In the

13 other isolates, the SNPs ranged from 173,656 - 200,949 and the majority of these isolates belonged to the second sub-population (Group A). The patterns based on genome wide SNPs further highlight the occurrence and spread of two sub-populations of the FMB pathogen within Eastern Africa meriting further investigations.

Phylogenomic analysis of the 18 FMB pathogen isolates from Eastern Africa along with 9 other *M. oryzae* isolates based on 944 genes (from the Sordariomycetes - BUSCO) provided insights into the evolutionary relationships of the blast pathogen associated with various hosts (Figure 4.11). Existence of host specific or host adapted forms of the blast pathogen *M. oryzae* related to key crops such as rice and wheat and their phylogenetic relationships are well established (Couch *et al*., 2005; Chiapello *et al*., 2015; Yoshida *et al*., 2016; Gladieux *et al*., 2018). Clades 1 and 2 defined in this study based on the phylogenomic analysis reflected the varying degrees of association of the host-related forms of the blast pathogen. It is important to emphasise the monophyletic nature of the finger millet blast isolates inclusive of the 18 from Eastern Africa and 3 from India suggesting a common origin of the pathogen (Clade 1). This model fits well with the widely recognised view of Eastern Africa as the primary centre of origin and diversification of finger millet and India as the secondary centre of diversity following the introduction into Asia ~3000 years ago from Africa (Chiapello *et al*., 2015; Gimode *et al*., 2016). This work is the first report to establish the phylogenomic relationship between FMB pathogen isolates from Eastern Africa and Asia. Previous reports had suggested some level of reintroduction of FMB pathogen from India into Eastern Africa (Dobinson *et al*., 1993; Takan *et al*., 2012). However, within Clade 1, FMB pathogen isolates from Eastern Africa and India formed distinctive genetic groups (Sub-Clades 1A and 1B). Further, the FMB pathogen isolates from Eastern Africa were also divided

into two genetic groups (Sub-Clades 1Ai and Aii), which broadly relate to the two sub-populations identified by the MLS analysis in Chapter 3. As discussed earlier, diversity of the planting material used in finger millet production systems in Eastern Africa along with seed movement is likely to have led to the pathogen population patterns identified (De Villiers *et al*., 2015; Howlett *et al*., 2015; Gimode *et al*., 2016; Tesfaye and Mengistu, 2017; Lule *et al*., 2018). Some recent phylogenomic studies mainly focused on rice blast and wheat blast have included a limited number of isolates from finger millet (*Eleusine coracana*) and/or wild millet (*E. indica*) for wider comparison (Yoshida *et al*., 2016; Shirke *et al*., 2016; Gladieux *et al*., 2018). In one such study, based on 3 isolates from *E. coracana* (in Japan and Philippines) and 5 isolates from *E. indica* (in China, Bolivia, Brazil and Ivory coast), 2 lineages of *M. oryzae* have been reported (Gladieux *et al*., 2018). An earlier study based on *grh* fingerprinting and ITS sequencing also suggested this pattern (Tanaka *et al*., 2009). In this context, it is pertinent to mention that the CD156 isolate from *E. indica* in Ivory coast included in the present study was part of Clade 2B and very distinctive to the FMB pathogen isolates in Clade 1, reflecting the two lineages discussed above.

Grasshopper (*grh*) is a key repeat element identified in the *M. oryzae* from *Eleusine species* including finger millet. The size of *grh* element was reported as ~8 Kb defined by the flanking short direct repeat sequence and long terminal repeats (LTRs). However, only the partial sequence of 5233 bases without the flanking short direct repeat sequence and long terminal repeats (LTRs) at the 3' end was deposited in the GenBank with accession number M77662 from isolate G22 (Dobinson *et al*., 1993). Utilising the reference genome assembly of E34, this study has identified the complete DNA sequence of *grh* element including the 3'

flanking short direct repeat sequence and long terminal repeats (LTRs) for the first time to define 7610 bases including a 2202 base region with no homology to any known fungal sequences (Figure 4.6 A and B). The DNA sequence of the *grh* element identified has high homology to the previously submitted *grh* element sequence of isolate G22 in the GenBank (Dobinson *et al*., 1993). Comparative genomic analysis of E34 and K23/123 assemblies provided a novel insight into presence and amplification pattern of the *grh* element among the finger millet blast pathogen *M. oryzae* populations in Eastern Africa.

Previous research based on Southern hybridisation and/or PCR had reported that only some FMB pathogen isolates contained the *grh* element (Dobinson *et al*., 1993; Tanaka *et al*., 2009). This type of isolates was mainly thought to be present in Asia (Japan and India) with very limited presence in Eastern Africa (Viji *et al*., 2000; Tanaka *et al*., 2009; Takan *et al*., 2012). Using the previously developed diagnostic PCR (Takan *et al*., 2012), this study showed for the first-time wide occurrence of *grh* positive *M. oryzae* isolates from finger millet in Eastern Africa (110 isolates mainly from Ethiopia and Tanzania; Chapter 3, Table 3.13 and Figure 3.18A). This prompted further in-depth analysis of the genome-level data from 18 isolates representing the *grh* positive and *grh* negative *M. oryzae* isolates, in this Chapter. Comparative analysis revealed that both types of isolates contained the *grh* element but with major differences in the genomic architecture. For instance, in isolate E34, which represents the *grh* positive type isolates, there were ~34 copies of the *grh* element with the size of the copies ranging from ≥ 5000 bases to ≥7000 bases. Conversely, in K23/123, which represents the *grh* negative type isolates, there were only 3 copies of ≥ 5000 bases (Table 4.7) with the largest ~5975 bases. Comparative analysis of the corresponding ~5975 bases in

isolates E34 and K23/123 revealed a significant level of nucleotide differences of ~844 bases (Figure 4.13). Further analysis using the ~5233 bases of *grh* sequence originally reported from isolate G22 from finger millet blast in Japan (Accession number M77662, Dobinson *et al*., 1993) with corresponding sequence in E34 revealed high homology with ~10 nucleotides differences (Figure 4.14). Conversely, comparison of *grh* fragment sequences from K23/123 with G22 revealed a striking difference of ~729 bases in K23/123 (Figure 4.15) spanning the regions used for designing the *grh* diagnostic PCR primers from the original sequence in G22 (Figure 4.17). The high level of nucleotide differences recorded in isolates such as K23/123 is reflective of the *grh* negative results reported previously by Southern hybridisation (Dobinson *et al*., 1993; Tanaka *et al*., 2009) as well as diagnostic PCR (Takan *et al*., 2012). Recent work with genome data from 3 isolates of the FMB pathogen in India also showed the presence of *grh* elements (Shirke *et al*., 2016). Overall, based on present results, there appear to be at least two types of *grh* elements in the FMB pathogen isolates with differences in their genomic architecture particularly in Eastern Africa, which requires further investigations.

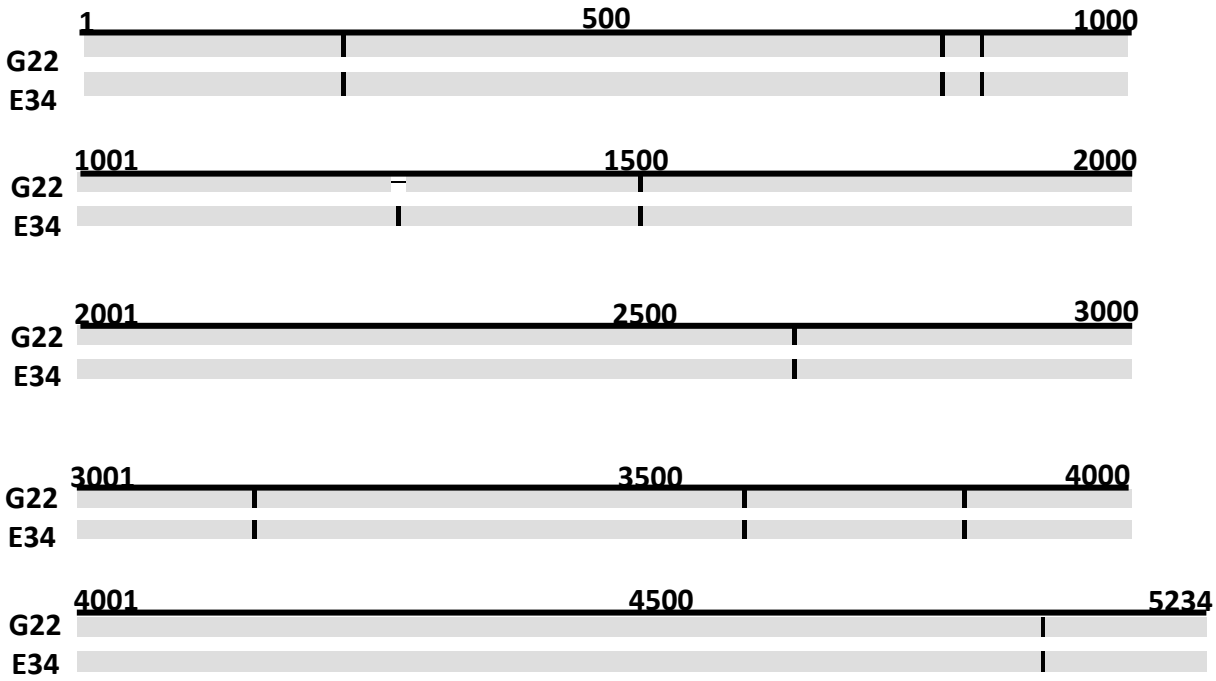

Figure 4.14. Schematic representation of the variation between the *grh* element in the genome of isolates E34 and G22

The black lines reflect the level differences between the isolates.

The *grh* sequence of G22 from which the element was originally described (Dobinson *et al*.,1993) was used as reference.

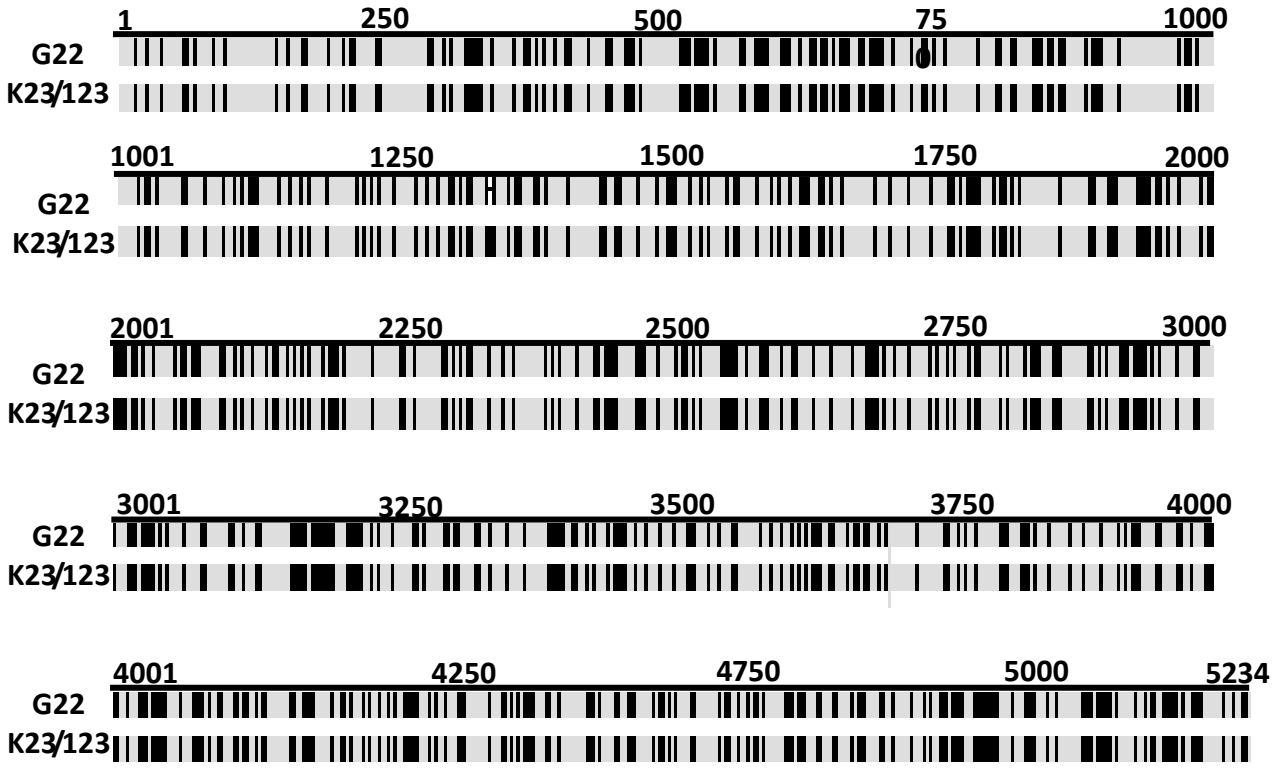

Figure 4.15. Schematic representation of the high degree of variation between the *Grasshopper* (*grh*) element in the genomes of isolates K23/123 and G22

The black lines and bars reflect the level of nucleotide differences between the isolates.

The *grh* sequence of G22 from which the element was originally described (Dobinson *et al*.,1993) was used as reference.

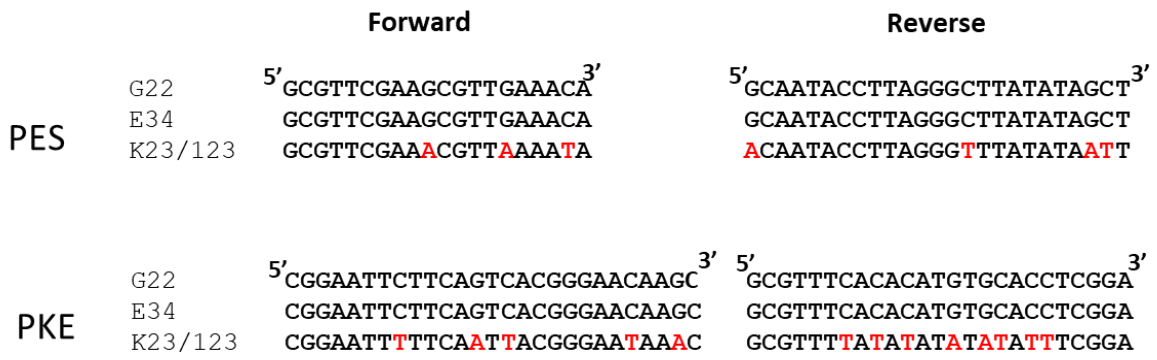

Figure 4.16. Grasshopper diagnostic PCR primer regions showing variation in the isolate K23/123 compared to the isolates G22 and E34

The bases that are different in K23/123 are shown in red.

In conclusion, the work reported in this chapter led to the development of a significant volume of genomic resources for the FMB pathogen. Two reference genomes have been developed for isolates E34 and K23/123 representing the two sub-populations of the FMB pathogen in Eastern Africa along with the 16 resequenced genomes. Comparative genomic analysis has evidenced putative isolate specific regions and/or genes likely to be pivotal in driving host interaction patterns, providing a platform for further functional annotation and analysis. Analysis of two model AVR genes from rice blast revealed presence and absence polymorphisms as well as different allelic versions in the FMB pathogen opening-up a vast unexplored area of research. SNP data as well as the phylogenomic analysis results have enhanced our understanding of the FMB pathogen origin and evolution with the identification a monophyletic group from Eastern Africa and India, potentially related to movement of the host material and the pathogen within and across continents. Finally, the present work has revealed the occurrence of at least two types of the *Grasshopper* element in the FMB pathogen within Eastern Africa with significant differences in their genomic architecture, with the potential to further clarify this aspect in the FMB pathogen isolates representing other geographic locations such as India. Overall, the genomic resources developed and the genome level knowledge gained will enable rapid advances in our understanding of the finger millet blast pathogen genetics, pathology and biology including the sexual reproductive potential, which is the main focus of investigation in the next chapter.

Chapter 5

5.0 Sexual reproductive capability of the finger millet blast pathogen *M. oryzae* populations in Eastern Africa

5.1 Introduction

Pathogen diversity is principally driven by a combination of processes including mutation, reproduction, recombination and natural selection (Saleh *et al*., 2012; Seidl and Thomma, 2014, Milgroom, 2015). These processes aid crop pathogens' adaptation to diverse environmental conditions as well as in enhancing their ability to overcome host resistance (Saleh *et al*., 2012; Seidl and Thomma, 2014; Milgroom, 2015). Among these processes however, reproduction is recognised to play a critical role in pathogen evolution and genetic diversity (Billiard *et al*., 2012). Fungi display diverse modes of reproductive behaviour including asexual, sexual and parasexual reproduction (Seidl and Thomma, 2014). In some species, the sexual reproductive forms may be cryptic or facultative, thus making it more difficult to detect the structures in nature even under specific conditions (Saleh *et al*., 2012). Understanding the key factors influencing an organism's reproductive mode is fundamentally important to decipher the evolutionary process (Billiard *et al*., 2012). Like various fungal species, reproductive behaviour of the cereal blast pathogen *M. oryzae* is complex as it predominantly reproduces asexually in nature during crop infection cycles, but its sexual reproductive capability under *in vitro* conditions in the lab is well known (Hebert, 1971, Itoi *et al*., 1983). The key question regarding whether the pathogen reproduces sexually in nature in any of the crop production systems remains unresolved. However, the

haploid asexual state of the pathogen can be readily recognised when isolated from their host and grown in pure culture normally (Ajello and Cheng, 1967).

M. oryzae is one of the earliest and well-studied pathogenic fungi and its sexual or perfect state was first demonstrated by crossing isolates from crabgrass (*Digitaria sanguinalis* L.), demonstrating the heterothallic nature of the fungus (Herbert, 1971). Ever since, efforts have been made to produce the perfect state of the *M. oryzae* under controlled conditions on artificial media using hermaphroditic isolates from *Eleusine coracana* and *Oryza sativa* as testers to cross new isolates from the field (Kato and Yamaguchi, 1982; Itoi *et al*., 1983; Viji and Gnanmanickam, 1998; Metwatanakarn *et al*., 1999; Takan *et al*., 2012). Compatible strains in the crosses produce perithecia, which is the fruiting body of the fungus reflecting the sexual state (Metwatanakarn *et al*., 1999; Takan *et al*., 2012). Besides the artificial media, perithecia production has been observed on dead tissue of leaf sheaths and nodes of standing rice plants in a humid chamber under control conditions (Nottegham and Silue, 1992; Hayashi *et al*., 1997). The perithecia produced on the dead tissue contained asci and viable ascospores similar to those observed on the artificial media (Notteghem and Silue, 1992; Hayashi *et al*., 1997; Takan *et al*., 2012). However, to date, there are no reports of perithecia production by *M. oryzae* isolates in the field (Zeigler, 1998; Saleh *et al*., 2012).

Consistent efforts have been made to study the sexual reproductive potential of *M. oryzae* isolates from different hosts in various parts of the world using mating assays under controlled conditions, and successful production of the perfect stage of the fungus had been reported in Japan (Kato and Yamaguchi, 1982), China (Zeng *et al*., 2009), Thailand (Mekwatanakam *et al*., 1999), India (Viji and Gnanamanickam, 1998; Adarisini *et al*., 1999; Karthikeyan and Gnanamanickam, 2008b; Dayakar *et al*., 2012), Argentina (Consolo *et al*.,

2005) and Sub-Saharan African countries such as Kenya and Uganda (Takan *et al*., 2012). As recognised in other ascomycetes, the perfect state of *M. oryzae* is governed by genes found as a single locus designated as *MAT1*, with individual isolates containing either a *MAT1-1* or *MAT1-2* gene (Turgeon and Yoder, 2000; Paoletti *et al*., 2005; Kanamori *et al*., 2007; Amorim *et al*., 2017). These mating type genes are idiomorphs (Glass *et al*., 1988; Debuchy and Coppin, 1992; Kang *et al*., 1994). The *MAT1-1* locus (3.5 kb) contains a characteristic alpha (α) box gene, whereas the *MAT1-2* locus (2.5kb) contains a single open reading frame (ORF) encoding a high mobility group (HMG) gene (Paoletti *et al*., 2005; Kanamori *et al*., 2007; Tsui *et al*., 2013; Bihon *et al*., 2014) as shown in Figure 5.1. Kanamori *et al*. (2007) further identified novel ORF(s) designated as *MAT1-1-3* and *MAT1-2-2* within the *MAT1-1* and *MAT1-2* idiomorphs, respectively. The sequences of these genes encode pheromone precursors designated as *MF1-1* and *MF2-1* related to the corresponding genes and their structure in *Saccharomyces cerevisiae* (Shen *et al*., 1999; Bobrowicz *et al*., 2002). The mating type genes *MAT1-1* and *MAT1-2* from *M. oryzae* isolates have been cloned and sequenced using a genomic subtraction method (Kang *et al*., 1994).

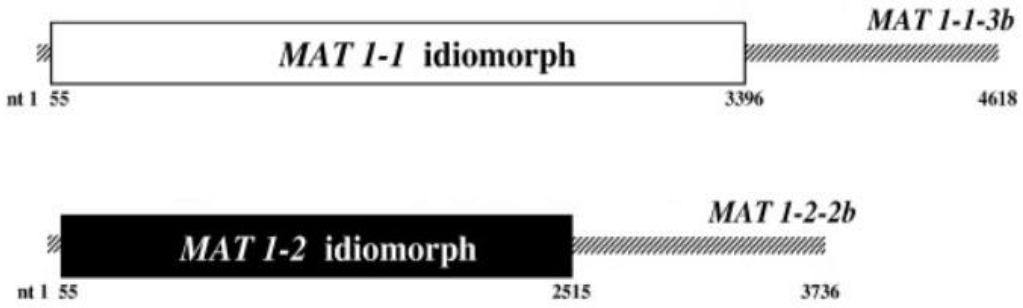

Figure 5.1. Schematic showing the organisation of the *MAT1* locus in *M. oryzae*

The white and dark boxes are *MAT1* idiomorphs presented as *MAT 1-1* and *MAT 1-2*, and striped lines are the flanking regions. The images were modified from Kanamori *et al*., 2007.

Availability of the sequence of these mating type genes has facilitated characterisation of the *M. oryzae* populations into MAT 1-1 or MAT 1-2 using molecular probes or PCR primers instead of performing time consuming and labour-intensive crossing assays with tester strains without knowing the mating type characteristic of a new isolate. These approaches have been used to assess the mating type distribution pattern among the blast pathogen populations from different hosts in various parts of the world. Subsequently, clonal or skewed distribution of the two mating types among the *M. oryzae* populations associated with rice from different parts of the world (Adarisini *et al*., 1999; Consolo *et al*., 2005; Takan *et al*., 2012; Samanta *et al*., 2014; Imam *et al*., 2015; Onaga *et al*., 2015; Sirisathaworn *et al*., 2017) and other hosts such as wheat (D'Avila *et al*., 2016) has been reported commonly. However, with the finger millet blast pathogen populations in Kenya and Uganda, an overall near equal distribution of the two mating types was observed (Takan *et al*., 2012).

The sexual phase of *M. oryzae* requires two strains of opposite mating types and at least one of which must be female-fertile, to come in contact with the other and interact to produce perithecia (Saleh *et al.*, 2012). The pattern of fertility in *M. oryzae* isolates ranges from sterile – the inability to mate with any other isolates, male fertile – the ability to mate with another isolate leading to the production of perithecia by that isolate; female fertile – the ability to mate with another isolate and produce perithecia and hermaphrodite - the ability to mate with another isolate leading to the production of perithecia by both the isolates wherein the isolates are able to act as both male and female (Itoi *et al*., 1983). However, the pattern of fertility observed so far largely correlates with the host-specific forms of *M. oryzae*. For example, many *M. oryzae* isolates from plant hosts such as finger millet, goose grass, weeping love grass are hermaphrodites producing mature perithecia generating numerous viable ascospores (Valent *et al*., 1991; Notteghem and Silue, 1992). In contrast, *M. oryzae* isolates from rice, in crosses with hermaphrodite tester isolates, are mainly either male or female sterile with only a very limited number producing mature perithecia or producing barren perithecia or perithecia with poorly viable ascospores (Valent *et al*., 1991; Notteghem and Silue, 1992). It has been suggested that *M. oryzae* isolates infecting rice might have lost their female fertility during the dispersal of the fungus from its centre of origin to the other parts of the world (Saleh *et al*., 2012). Even though the sexual phase of *M. oryzae* has not been reported in nature, characterisation of the finger millet blast pathogen in Kenya and Uganda revealed the presence of hermaphrodite and female fertile isolates in the region (Takan *et al*., 2012). This pattern is suggestive of the potential for sexual reproduction and recombination (Crawford *et al*., 1986; Saleh *et al*., 2012). The ability of *M. oryzae* isolates to produce perithecia is thought to be controlled by genes at

several regions, segregating independent of the mating type and pathogenicity on different hosts (Kolmer and Ellingboe, 1987). Understanding the occurrence of recombination in crop fungal pathogen populations is imperative in predicting the potential frequency of particular genetic groups that might represent distinctive pathotypes compatible with resistant cultivars or specific strains resistant to fungicides (Petes, 2001; Bernstein and Bernstein, 2010; Milgroom, 2015). Some of the criteria employed to predict the potential for recombination include the presence of sexual structures, genotype diversity levels, possible recombinant genotypes, and genome wide diversity. This can be evaluated using the mating type ratios, genotypic diversity index and gametic disequilibrium (Milgroom, 2015).

5.1.2 Aim and Objectives

The main aim of the work reported in this chapter is to assess the sexual reproductive capability underlying the possibility for recombination of the blast pathogen *M. oryzae* populations prevailing in finger millet production systems in Eastern Africa specifically in Kenya, Uganda, Tanzania and Ethiopia. The Objectives are:

1. To determine the mating type of the *M. oryzae* isolates using the mating type-specific PCR and map the distribution pattern of the two mating types MAT1-1 and MAT 1-2.
2. To assess the fertility status of the *M. oryzae* isolates by performing the mating crosses with designated testers of the opposite mating type.
3. To integrate the mating type distribution and the fertility status data as a basis for evaluating the sexual reproduction and recombination potential in shaping the finger

millet blast pathogen populations in the context of effective management of host resistance.

5.2 Experimental approach

Mating type-specific PCR assay with MAT 1-1 forward and reverse, and MAT 1-2 forward and reverse primers was carried out (section 2.8.1) to determine the mating type of the 224 *M. oryzae* isolates representing the contemporary populations used in this study. Each batch of PCR assays included positive control isolates TH3 (MAT 1-1) and K23/123 (MAT 1-2), respectively and a tube with sterile water instead of fungal genomic DNA as negative control. Mating crosses of new isolates representing the contemporary populations were carried out using previously characterised tester isolates of known mating type. The testers used were MAT 1-1 isolates 4136-4-3, I-R-22 and TH3, and MAT 1-2 isolates Guy11, JP 15, BR 62, K23/123 and D15/s47. Mating crosses were set up by co-culturing a new isolate of known mating type with a tester isolate of the opposite mating type. In these assays, the inoculum discs of the two isolates were placed 4 cm apart on OMA plates, which were then incubated under a specific temperature and light regime for approximately 4 weeks (section 2.8.2 and 2.8.3). The mating crosses were monitored periodically over this period for production of perithecia and where present, the perithecia were counted under a stereomicroscope. Each new isolate from each cross was scored as fertile or infertile based on the presence and absence of perithecia either on the side of the new isolate or the tester or both. The fertile isolates were classified as male or female or hermaphrodite according to the nomenclature of Itoi *et al*. (1983) based on the pattern of production of perithecia by the new isolate and the tester (section 2.8.2 and 2.8.3).

5.3 Results

5.3.1 Mating type identification and their distribution pattern

Initially, the efficacy of the mating type-specific primers designated MAT 1-1F and R, and MAT 1-2F and R and the PCR conditions (Takan *et al*., 2012) was validated in distinguishing the previously characterised *M. oryzae* tester isolates into mating types MAT 1-1 and MAT 1-2 based on the production of an amplicon of the expected size (data not shown). Subsequently, the mating type-specific PCR was used routinely to screen the 224 *M. oryzae* isolates representing the contemporary populations in this study. Each batch of PCR included appropriate positive and negative controls as described previously. Each *M. oryzae* isolate was identified as either MAT 1-1 or MAT 1-2 based on the PCR product size corresponding to ~960 bp or 802 bp, respectively, with reference to the product size of the positive control isolates, TH3 and K23/123 (example, Figure 5.2). Based on this assay the mating type of 224 *M. oryzae* isolates from Eastern Africa was determined as MAT 1-1 or MAT 1-2 (Table 5.1). Among these isolates, 55.8 % (125 isolates) were MAT 1-1 and 44.2 % (99 isolates) were MAT 1-2 (Figure 5.3). Both mating types were present in all 4 countries surveyed in Eastern Africa. However, based on the characterisation of 63 isolates from Ethiopia, 45 from Kenya, 58 from Tanzania, and 58 from Uganda collected from various locations (districts), the overall proportion of the two mating types varied in these countries (Figure 5.4). In Ethiopia, Kenya and Uganda, the proportions of MAT 1-1 isolates are higher than MAT 1-2 isolates, whilst in Tanzania opposite pattern was recorded with a clearly higher proportion of MAT 1-2 isolates compared to the MAT 1-1 type (Figure 5.4). For instance, in Kenya, Uganda and Ethiopia, the proportion of MAT 1-1 isolates ranged from 56.9 % to 68.3 % and the proportion of MAT1-2 isolates ranged from 31.7 % to 43.1 %.

Among the 58 *M. oryzae* isolates from Tanzania, 39.7 % were MAT 1-1 and 60.3 % were MAT 1-2 (Table 5.1). Further, the distribution the MAT 1-1 and MAT 1-2 isolates at different locations (districts) within each of the countries revealed complex patterns. For example, in Tanzania, MAT 1-1 was clearly dominant in Mbozi (9 out of 10 isolates) and MAT 1-2 was dominant (16 out of 21 isolates) in Njombe. In Nkasi (10 isolates), MAT 1-1 and MAT 1-2 levels were comparable, where as in Madaba, MAT 1-1 was not recorded as all 6 isolates tested were MAT 1-2. Similar complexity was observed in the distribution pattern of the *M. oryzae* isolates belonging to the two mating types in various districts in the other 3 countries surveyed in Eastern Africa (Table 5.1).

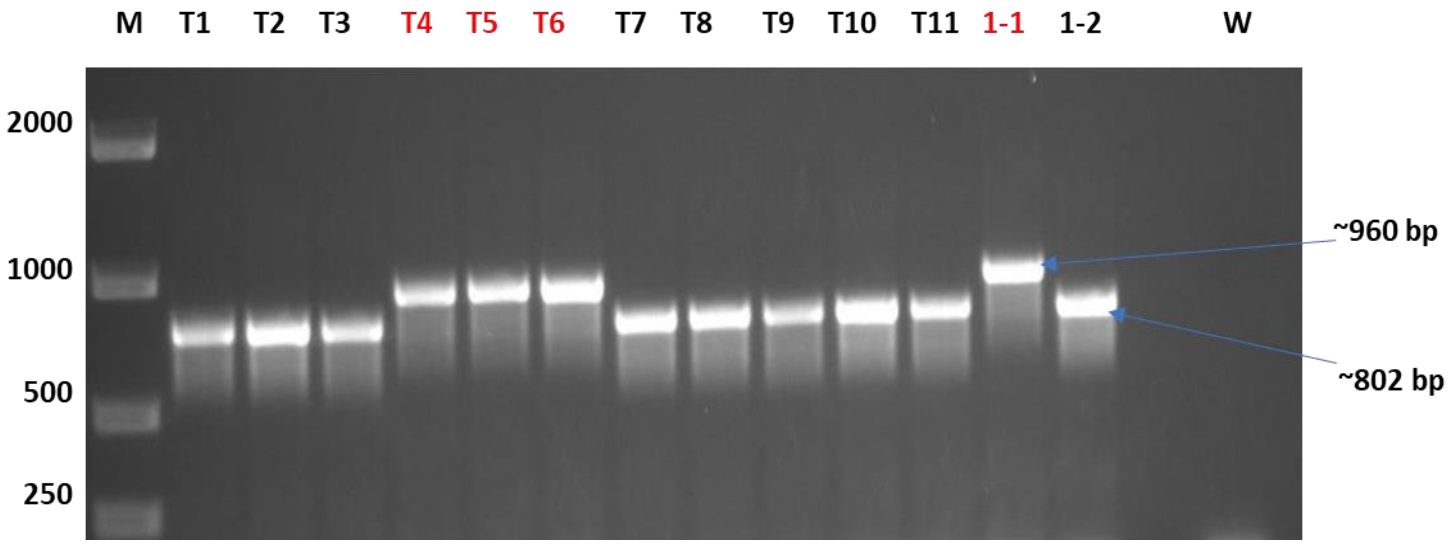

Figure 5.2. Agarose gel showing the *MAT 1-1* (960 bp) and *MAT 1-2* (802 bp) specific amplicons in a set of *M. oryzae* isolates
Isolates with the 960 bp PCR fragment were identified as MAT 1-1 and isolates with 802 bp PCR fragment were identified MAT 1-2, with reference to positive controls.

Lane M is the molecular known size DNA marker (Easy ladder1, Bioline-UK); Lane W – water (negative control).

Lanes 1 to 11 are *M. oryzae* isolates T1 to T11 from Tanzania; Lanes 1-1 and 1-2 are isolates TH3 (MAT 1-1 positive control) and K23/123 (MAT 1-2 positive control), respectively.

Further details of the isolates are available in Table 2.1A.

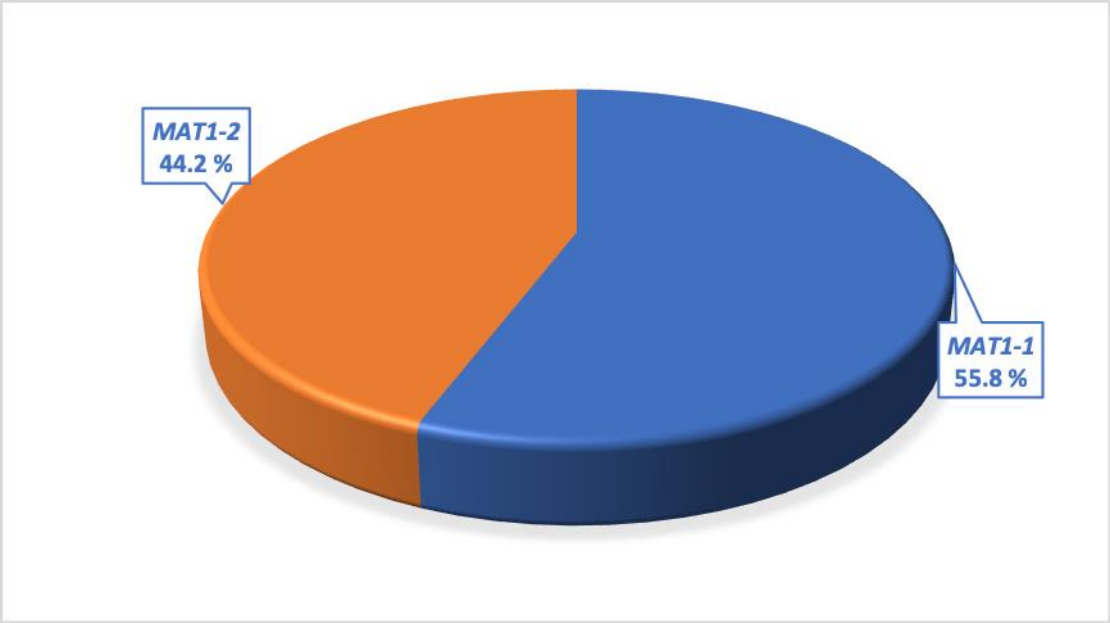

Figure 5.3. Proportion of the *M. oryzae* isolates of the two mating types MAT 1-1 and MAT 1-2 in finger millet production systems in Eastern Africa

Percentage values were calculated based on the number of isolates containing either the *MAT1-1* or the *MAT1-2* mating type gene among the 224 *M. oryzae* isolates representing the contemporary population in this study.

The blue and orange colours relate to MAT 1-1 and MAT 1-2 mating type isolates, respectively.

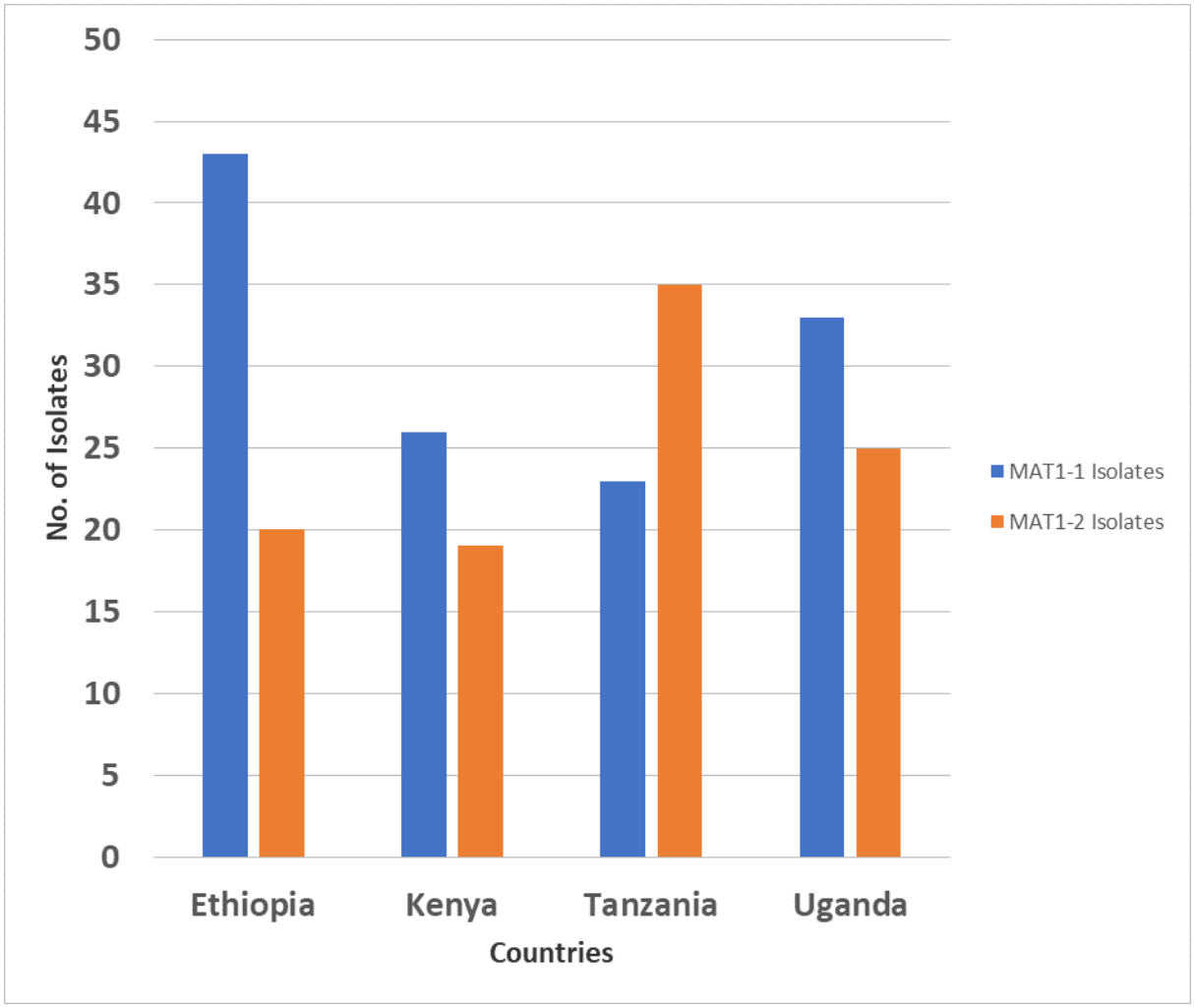

Figure 5.4. Pattern of occurrence of the MAT 1-1 and MAT 1-2 type *M. oryzae* isolates in the four countries

The number of MAT 1-1 isolates is represented by the blue colour bars and number of MAT 1-2 isolates is represented by the orange colour bars.

Table 5.1. Distribution pattern of the two mating types MAT 1-1 and MAT 1-2 among *M. oryzae* isolates in some of the locations (districts) in Ethiopia, Kenya, Tanzania and Uganda

Region/ Country	District*	Number of isolates[x]	MAT 1-1[Y]	MAT 1-2[Z]
Eastern Africa	**Total**	**224**	**125 (55.8 %)**	**99 (44.2 %)**
Ethiopia	Waju Tuka	4	4	0
	Diga	10	9	1
	Nedjo	5	5	0
	Banja	5	4	1
	Bahir Dar Zuria	6	0	6
	Dure Bete	4	2	2
	Mecha	4	2	2
	Total	**63**	**43 (68.3 %)**	**20 (31.7 %)**
Kenya	Siaya	5	1	4
	Kisumu	6	5	1
	Busia	13	8	5
	Kisii	14	9	5
	Total	**45**	**26 (57.8 %)**	**19 (42.2 %)**
Tanzania	Nkasi	10	4	6
	Sumbawanga	5	3	2
	Mbozi	10	9	1
	Njombe	21	5	16
	Madaba	6	0	6
	Momba	6	2	4
	Total	**58**	**23 (39.7 %)**	**35 (60.3 %)**
Uganda	Arua	4	1	3
	Serere	4	2	2
	Kumi	6	5	1
	Ngora	4	4	0
	Alebtong	4	2	2
	Manafwa	4	0	4
	Tororo	4	1	3
	Total	**58**	**33 (56.9 %)**	**25 (43.1 %)**

* District level data shown is only the locations from which at least 4 or more isolates were collected; whereas, totals shown for Eastern Africa and each Country are based on the overall number of isolates collected and characterised from various locations

[x]Number of *M. oryzae* isolates representing the contemporary population (2015-2017) [Y]Number of MAT 1-1 isolates present in Eastern Africa or in a country or in certain districts

[Z]Number of MAT 1-2 isolates present in Eastern Africa or in a country or in certain districts

Numbers in parentheses are percentage values of the proportion of MAT 1-1 and MAT 1-2 isolates.

5.3.2 Fertility status of the *M. oryzae* contemporary populations

Initial validation experiments were performed to determine the sexual fitness of the mating testers (I-R-22, TH3 and 4136-4-3 as MAT 1-1) and BR62, Guy11, JP15 and K23/123 as MAT 1-2). This was carried out by crossing the opposite mating type isolates among the testers to mainly check the production of perithecia. The assay is temperature, light and the growth medium sensitive. The crossing assays were performed on OMA plates under regulated conditions. Initial screening showed variation in the number of perithecia produced per crossing. Mating testers I-R-22, TH3, 4136-4-3 and K23/123 retained their sexual fitness as evidenced by the production of perithecia with asci and viable ascospores. However, Guy11 was extremely slow growing, which affected the interaction zone (crossing point) and the mating period (data not shown). Further, isolates BR62 and JP15 were not efficient testers in mating with all the isolates crossed. Subsequently, 6 hermaphrodite isolates (D9/s76, D9/s9, D3/s3, D15/s47, K9/47 and K30/90w) from previous work (Takan *et al*., 2012), were selected and screened as potential testers as described above. This led to the identification of isolate D15/s47 (MAT 1-2) as a suitable mating tester. D15/s47 was isolated from finger millet in Uganda and showed similar levels of mating capability as the previously established tester isolates (Data not shown). Following the two standardisation experiments, isolates I-R-22, TH3 and 4136-4-3 (MAT 1-1) and isolates K23/123 and D15/s47 (MAT 1-2) were selected as the mating testers to screen the 224 *M. oryzae* isolates representing the contemporary population.

With each new MAT 1-1 isolate, crossing assays were performed with two testers of the opposite mating type, and with each new MAT 1-2 isolate, crossing assays were performed with three testers of the opposite mating type (Table 5.2). From a total of 547 crosses, 226

crosses were compatible and resulted in the production of perithecia (Table 5.2). With some isolates, perithecia were observed as early as 10 to 12 days after crossing, but with most isolates, perithecia were formed after 3 to 4 weeks after crossing. Further, taking the compatible crosses into account, it was clear that some of new the isolates were not successfully compatible will all the testers of the opposite mating type. Consequently, new isolates that produced perithecia with at least one tester isolate were considered fertile (137 isolates) and new isolates that were not compatible with any of the mating testers were considered infertile (87 isolates). However, among the fertile isolates, their ability to mate with opposite mating type testers varied considerably. For example, among the MAT 1-2 isolates (99), 14 produced perithecia with all three MAT 1-1 testers used, 33 isolates produced perithecia with two MAT 1-1 testers and 21 isolates produced perithecia with only one MAT 1-1 tester used. Among the MAT 1-1 isolates (125), 33 isolates produced perithecia with two MAT 1-2 testers and 36 isolates produced perithecia with one MAT 1-2 tester. The new mating tester isolate D15/s47 (MAT 1-2) identified in the study proved to be efficient in the crossing assays (Table 5.2). This also suggests that a higher proportion of the MAT 1-2 isolates are fertile (68 out of 99 isolates) compared to the MAT 1-1 isolates (69 out of 125 isolates).

Table 5.2. Proportion of compatible crosses of new *M. oryzae* isolates with the mating testers based on the production of perithecia

M. oryzae tester isolates	Tester mating type*	Number of crosses	**Number of crosses with perithecia				Total no. of crosses with perithecia
			Ethiopia	Kenya	Tanzania	Uganda	
I-R-22	1-1	99	13	15	15	16	59
TH3	1-1	99	7	14	7	16	44
4136-4-3	1-1	99	8	8	4	3	23
D15/s47	1-2	125	23	6	8	14	51
K23/123	1-2	125	18	6	9	16	49
Total	na	**547**	**69**	**49**	**43**	**65**	**226**

*MAT 1-1 testers were used in crossing assays with new isolates established as MAT 1-2 by the PCR method.
*MAT 1-2 testers were used in crossing assays with new isolates established as MAT 1-1 using the PCR method.
NA, Not applicable

** Numbers for each country are related to the number of successful crosses with the testers (and don't directly relate to the number of isolates collected/tested from each country)

Overall, among the *M. oryzae* isolates characterised across the 4 countries (Ethiopia, Kenya, Tanzania and Uganda) in Eastern Africa, 61% were fertile (137) and 39 % were infertile (87) as presented in Figure 5.5. However, among the 4 countries, the proportion of fertile isolates in Ethiopia (47 out of 63 isolates) and Uganda (38 out of 58 isolates) was clearly higher compared to Kenya (23 out of 45 isolates) and Tanzania (29 out of 58 isolates), which both showed comparable proportions of fertile and infertile isolates (Figure 5.6). The distribution pattern of the fertile *M. oryzae* isolates at various locations (districts) across the 4 countries in Eastern Africa is detailed in Table 5.3.

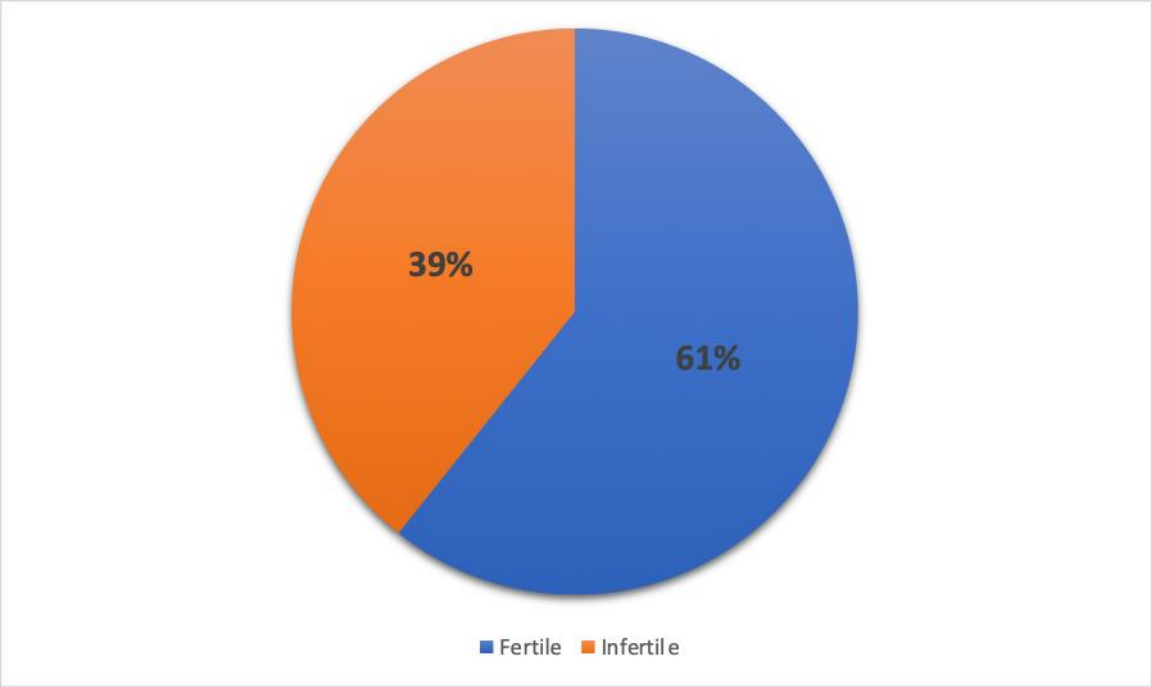

Figure 5.5. Proportion of the fertile and infertile *M. oryzae* isolates overall in Eastern Africa

Percentage values are based on the number of fertile and infertile isolates identified among the 224 isolates representing the contemporary population (collection period was 2015 to 2017).

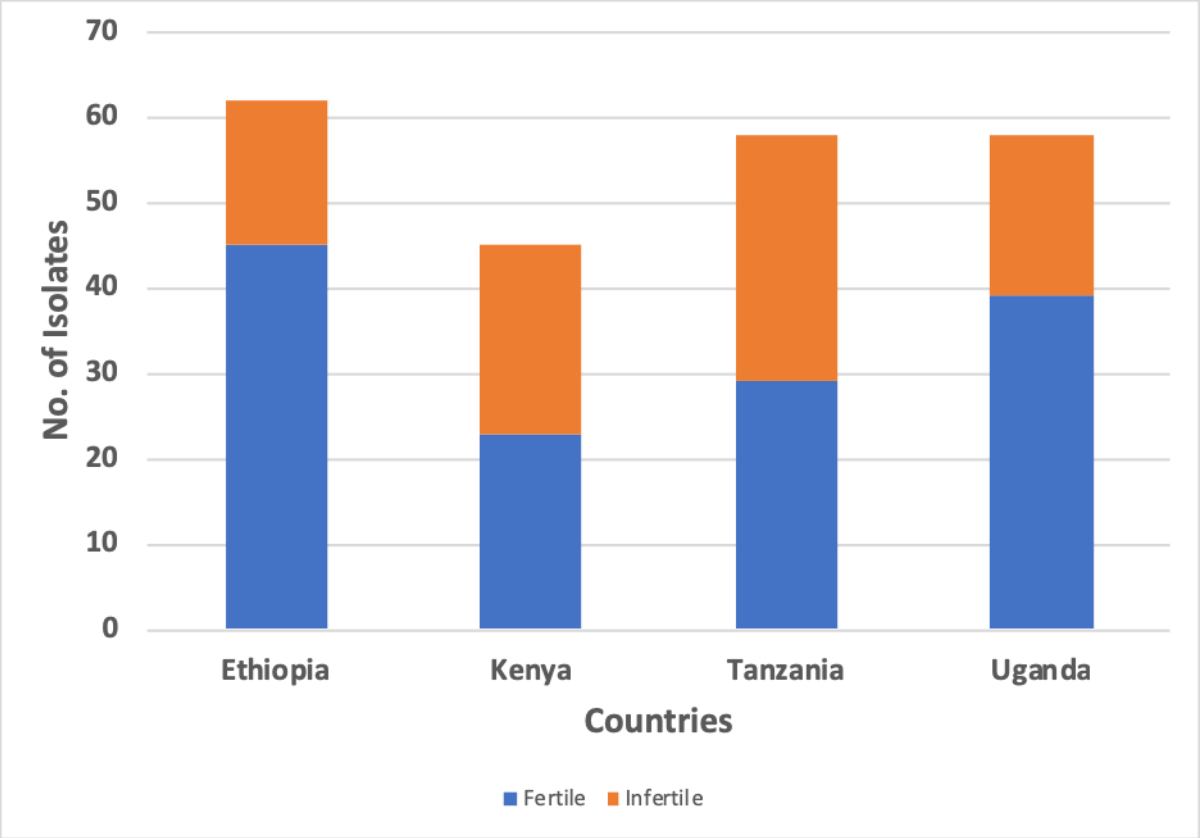

Figure 5.6. Proportion of fertile and infertile *M. oryzae* isolates identified in Ethiopia, Kenya, Tanzania and Uganda

Identification is based on the sexual compatibility and incompatibility of the new isolates to the standard mating testers

Identification as fertile isolates is based on the production of perithecia in compatible crosses with at least one tester isolate of the opposite mating type

Blue and Orange bars represent the fertile and infertile isolates, respectively among the total number of isolates characterised in Ethiopia (63), Kenya (45), Tanzania (58) and Uganda (58).

Table 5.3. Distribution of the fertile and infertile finger millet blast pathogen isolates across the four countries

Isolates status*	Country	No. of isolates in country	Isolates	Districts**
Fertile (137)	Ethiopia	47	E3, E4, E5, E7, E8, E9, E12, E13, E14, E15, E16, E17, E18, E20, E21, E23, E25, E27, E28, E30, E31, E33, E36, E37, E38, E39, E34, E40, E41, E42, E44, E45, E47, E48, E49, E50, E51, E52, E53, E54, E56, E57, E58, E59, E60, E61, E62	20[a]
	Kenya	23	K1, K6, K7, K8, K10, K13, K14, K17, K19, K20, K21, K24, K25, K26, K27, K30, K31, K32, K36, K37, K39, K40, K44, K45	7[b]
	Tanzania	29	T1, T2, T5, T6, T9, T10, T11, T12, T14, T15, T16, T17, T18, T21, T22, T24, T26, T29, T31, T32, T33, T35, T36, T42, T47, T50, T51, T53, T58	6[c]
	Uganda	38	U1, U2, U3, U5, U6, U7, U8, U9, U10, U11, U13, U15, U19, U20, U21, U22, U23, U25, U27, U29, U30, U31, U35, U36, U37, U38, U39, U40, U42, U43, U44, U48, U49, U52, U53, U55, U56, U57,	17[d]
Infertile (87)	Ethiopia	16	E1, E2, E6, E10, E19, E22, E24, E26, E29, E32, E35, E41, E43, E46, E55, E63	10[e]
	Kenya	22	K2, K3, K4, K5, K9, K11, K12, K15, K16, K18, K22, K23, K28, K29, K33, K34, K35, K38, K41, K42, K43	5[f]
	Tanzania	29	T3, T4, T7, T8, T13, T19, T20, T23, T25, T27, T28, T30, T34, T37, T38, T39, T40, T41, T43, T44, T45, T46, T48, T49, T52, T54, T55, T56, T57	5[g]
	Uganda	20	U4, 12, U14, U16, U17, U18, U24, U26, U28, U32, U33, U34, U41, U45, U46, U47, U50, U51, U54, U58	15[h]

*Fertile isolates are compatible with at least one mating tester and produced perithecia.

* Infertile isolates are incompatible with all the testers and no perithecia were produced in crosses.

** Some of the districts appear in both categories as both fertile and infertile isolates are present in these locations.

[a]The 20 districts in Ethiopia include Adet, Diga, Lallo Assabi, Banja, Bahir Dar Zuria, Bure, Demecha, Angebo, Wayu Tuka, Nedjo, Bila, Boji Bermeji, Gimbi, Leka dulecha, Dangla, Dure Bete, Guangau, Mecha, Ankussa-Abdo Gor, Jabi-Tana.

[b]The 7 districts in Kenya include Busia, Bungoma, Eldret, Marakwet, Kakamega, Kisii, Siaya.

[c]The 6 districts in Tanzania include Nkasi, Njombe, Mbozi, Madaba, Momba, Sumbawanga.

[d]The 17 districts in Uganda include Arua, Serere, Kumi, Pallisa, Katakwi, Amuria, Alebtong, Lira, Apac, Gulu, Amuru, Agago, Manafwa, Tororo, Masindi, Ngora, Lamwo.

[e]The 10 districts in Ethiopia include Sire, Wayu Tuka, Diga, Nedjo, Leta Sibu, Leka dulecha, Dure Bete, Mandura, Bahir Dar Zuria, Qilxxu Kara.

[f]The 5 districts in Kenya include Kisumu, Siaya, Busia, Kisii, Eldoret.

[g]The 5 districts in Tanzania include Nkasi, Sumbawanga, Njombe, Momba, Mbozi.

[h]The 15 districts in Uganda include Moyo, Pallisa, Katakwi, Kumi, Ngora, Apac, Lira, Kitgum, Agago, Tororo, Serere, Hoima, Amuria, Amuru, Mbale.

Sexual behaviour of fertile *M. oryzae* isolates is further evaluated using the Itoi *et al*. (1983) nomenclature based on perithecia production by the new isolate and/or the tester in crosses as described in previous sections and the patterns shown here (Figure 5.9A and B). Sexual behaviour of the fertile *M. oryzae* isolates identified among the contemporary population from Eastern Africa is detailed in Table 5.4. In vast majority of the isolates and crosses mature perithecia with asci and ascospores were observed. However, with some isolates and crosses, the perithecia produced were barren lacking asci and ascospores (e.g. E20, E25, E53, U25, U31, U42, and U56) as indicated in Table 5.4.

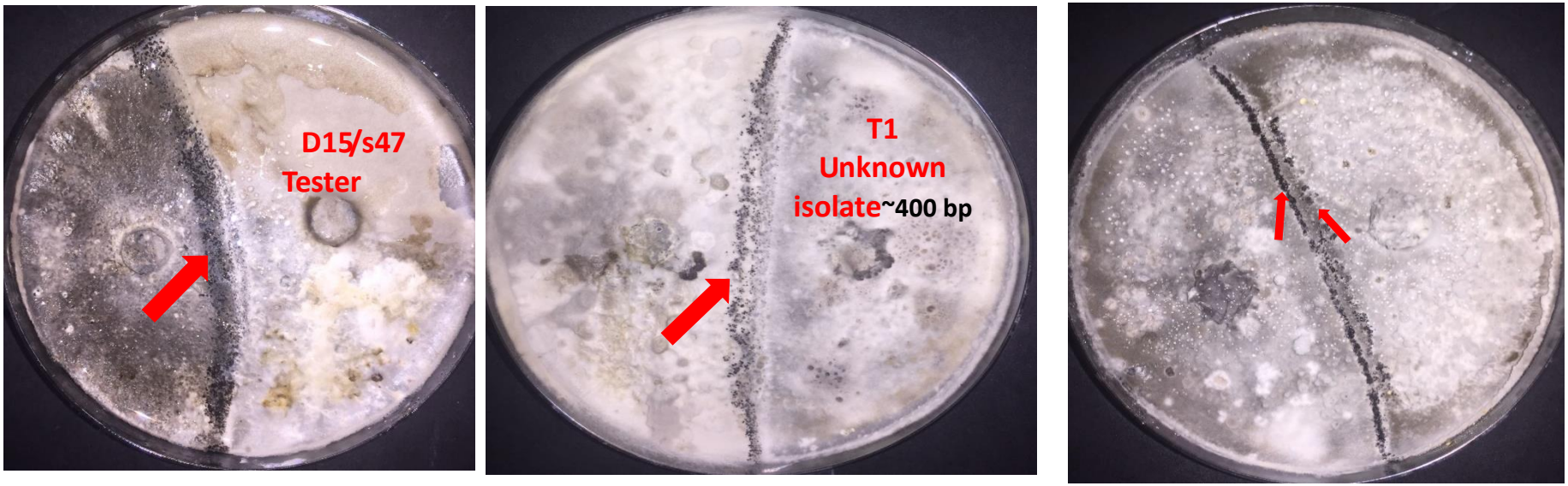

Figure 5.7A. Sexual status of *M. oryzae* isolates collected from finger millet from Ethiopia and Tanzania

The images show the pattern of perithecia production by female, male and hermaphrodite isolates.

Female isolate formed perithecia on the side of the new isolate E34; **Male** isolate formed perithecia on the side of the tester isolate (TH3); and **Hermaphrodite** isolate formed perithecia by the new isolate (T5) as well as the tester isolate (K23/123).

The 137 fertile new isolates were designated according to the Itoi *et al*. (1983).

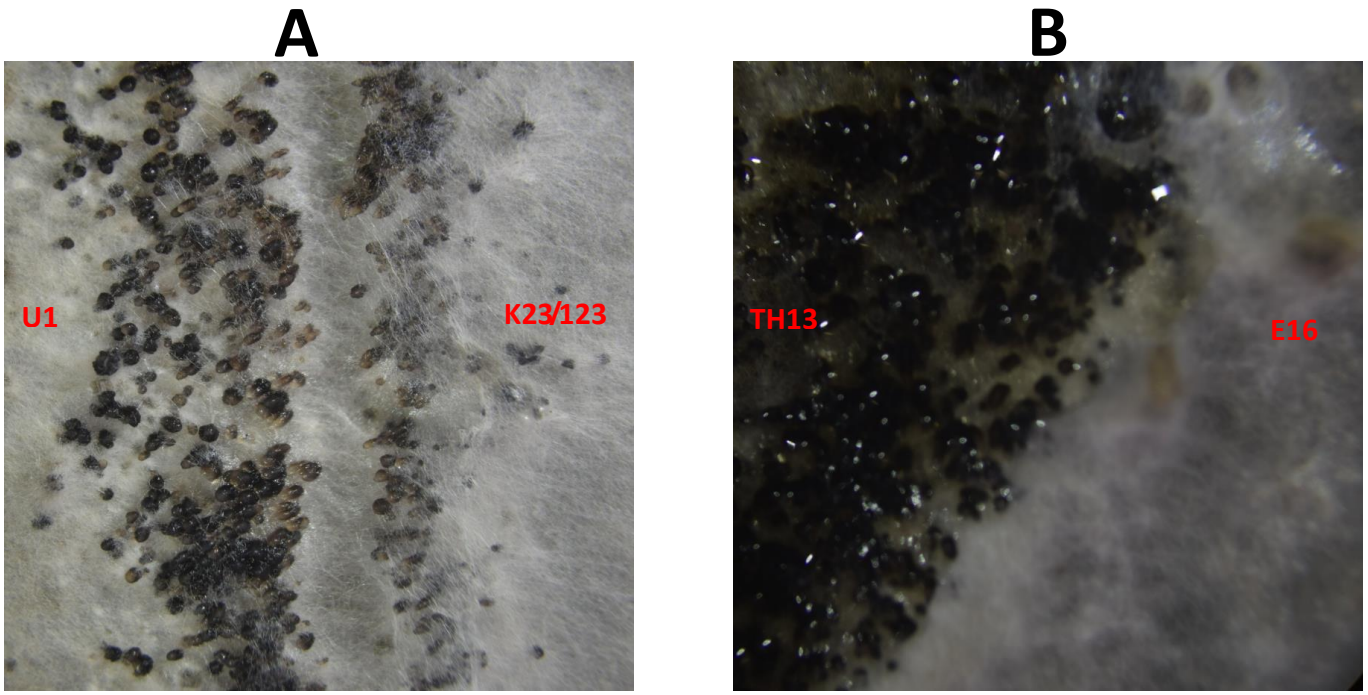

Figure 5.7B. Perithecia production by new *M. oryzae* isolates from finger millet and the tester isolates

A) Perithecia produced on both sides of the compatibility zone by hermaphrodite isolates U1 and K23/123

B) Perithecia produced by a compatible cross between isolates TH3 and E16 acting as female and male fertile, respectively

Table 5.4. Sexual behaviour of the fertile FMB pathogen isolates identified among the contemporary population in Eastern Africa based on crosses with mating testers

Isolate*	Mating type[a]	Tester[b]	No. of Perithecia[c]		Sexual behaviour[d]	Sexual structures[e]
			Isolate	Tester		
E3	MAT 1-2	4136-4-3	600	1200	Hermaphrodite	Perithecia, asci, ascospores
E3	MAT1-2	TH3	650	110	Hermaphrodite	Perithecia, asci, ascospores
E3	MAT1-2	I-R-22	720	1300	Hermaphrodite	Perithecia, asci, ascospores
E4	MAT1-1	K23/123	0	1320	Female	Perithecia, asci, ascospores
E5	MAT1-1	K23/123	586	700	Hermaphrodite	Perithecia, asci, ascospores
E5	MAT1-1	D15/s47	500	720	Hermaphrodite	Perithecia, asci, ascospores
E7	MAT1-1	K23/123	0	500	Male	Perithecia, asci, ascospores
E7	MAT1-1	D15/S47	80	100	Hermaphrodite	Perithecia, asci, ascospores
E8	MAT1-1	D15/s47	800	920	Hermaphrodite	Perithecia, asci, ascospores
E9	MAT1-1	D15/S47	0	750	Male	Perithecia, asci, ascospores
E11	MAT1-1	K23/123	0	800	Male	Perithecia, asci, ascospores
E12	MAT1-2	TH3	0	400	Male	Perithecia, asci, ascospores
E12	MAT1-2	I-R-22	0	400	Male	Perithecia, asci, ascospores
E13	MAT1-2	4136-4-3	0	1000	Male	Perithecia, asci, ascospores
E13	MAT1-2	TH3	0	1100	Male	Perithecia, asci, ascospores
E13	MAT1-2	I-R-22	0	959	Male	Perithecia, asci, ascospores
E14	MAT1-2	TH3	80	988	Hermaphrodite	Perithecia, asci, ascospores
E14	MAT1-2	I-R-22	0	1105	Male	Perithecia, asci, ascospores
E15	MAT1-1	K23/123	1105	1450	Hermaphrodite	Perithecia, asci, ascospores
E15	MAT1-1	D15/s47	1200	1500	Hermaphrodite	Perithecia, asci, ascospores
E16	MAT1-2	TH3	0	1000	Male	Perithecia, asci, ascospores
E16	MAT1-2	I-R-22	0	270	Male	Perithecia, asci,

						ascospores
E17	MAT1-2	K23/123	0	900	Male	Perithecia, asci, ascospores
E18	MAT1-1	D15/s47	500	200	Hermaphrodite	Perithecia, asci, ascospores
E20	MAT1-1	D15/s47	0	400	Male	Perithecia
E20	MAT1-1	K23/123	0	500	Male	Perithecia
E21	MAT1-2	K23/123	0	800	Male	Perithecia, asci, ascospores
E23	MAT1-1	D15/s47	0	400	Male	Perithecia, asci, ascospores
E25	MAT1-1	K23/123	0	100	Male	Perithecia
E27	MAT1-1	D15/s47	600	400	Hermaphrodite	Perithecia, asci, ascospores
E28	MAT1-1	D15/s47	0	650	Male	Perithecia, asci, ascospores
E30	MAT1-1	D15/s47	0	500	Male	Perithecia, asci, ascospores
E31	MAT1-1	D15/s47	70	100	Hermaphrodite	Perithecia, asci, ascospores
E33	MAT1-1	D15/s47	200	400	Hermaphrodite	Perithecia, asci, ascospores
E34	MAT1-1	D15/s47	500	0	Female	Perithecia, asci, ascospores
E36	MAT1-1	K23/123	150	170	Hermaphrodite	Perithecia, asci, ascospores
E37	MAT1-1	K23/123	0	500	Male	Perithecia, asci, ascospores
E38	MAT1-1	K23/123	0	1000	Male	Perithecia, asci, ascospores
E38	MAT1-1	D15/s47	0	1004	Male	Perithecia, asci, ascospores
E39	MAT1-1	D15/s47	0	500	Male	Perithecia, asci, ascospores
E40	MAT1-1	K23/123	250	400	Hermaphrodite	Perithecia, asci, ascospores
E40	MAT1-1	D15/s47	900	1205	Hermaphrodite	Perithecia, asci, ascospores
E43	MAT1-1	D15/s47	1020	1000	Hermaphrodite	Perithecia, asci, ascospores
E44	MAT1-1	K23/123	1000	1020	Hermaphrodite	Perithecia, asci, ascospores
E44	MAT1-1	D15/s47	0	500	Male	Perithecia, asci, ascospores
E45	MAT1-2	I-R-22	0	600	Male	Perithecia, asci, ascospores
E47	MAT1-2	4136-4-3	0	700	Male	Perithecia, asci, ascospores
E47	MAT1-2	I-R-22	0	100	Male	Perithecia, asci,

						ascospores
E48	MAT1-1	K23/123	0	800	Male	Perithecia, asci, ascospores
E48	MAT1-1	D15/s47	500	540	Hermaphrodite	Perithecia, asci, ascospores
E49	MAT1-1	D15/s47	0	960	Male	Perithecia, asci, ascospores
E50	MAT1-1	K23/123	0	200	Male	Perithecia, asci, ascospores
E50	MAT1-1	D15/s47	0	700	Male	Perithecia, asci, ascospores
E51	MAT1-1	K23/123	0	700	Male	Perithecia, asci, ascospores
E52	MAT1-2	TH3	0	150	Male	Perithecia, asci, ascospores
E52	MAT1-2	I-R-22	0	700	Male	Perithecia, asci, ascospores
E53	MAT1-1	D15/s47	0	200	Male	Perithecia
E54	MAT1-2	I-R-22	0	100	Male	Perithecia, asci, ascospores
E54	MAT1-2	4136-4-3	0	120	Male	Perithecia, asci, ascospores
E56	MAT1-2	TH3	0	70	Male	Perithecia, asci, ascospores
E57	MAT1-2	I-R-22	0	1800	Male	Perithecia, asci, ascospores
E58	MAT1-2	4136-4-3	0	100	Male	Perithecia, asci, ascospores
E58	MAT1-2	I-R-22	0	560	Male	Perithecia, asci, ascospores
E59	MAT1-2	4136-4-3	804	0	Female	Perithecia, asci, ascospores
E59	MAT1-2	I-R-22	0	890	Male	Perithecia, asci, ascospores
E60	MAT1-1	K23/123	0	250	Male	Perithecia, asci, ascospores
E61	MAT1-2	4136-4-3	0	1050	Male	Perithecia, asci, ascospores
E61	MAT1-2	TH3	1025	0	Female	Perithecia, asci, ascospores
E62	MAT1-2	4136-4-3	0	200	Male	Perithecia, asci, ascospores
E62	MAT1-2	TH3	0	250	Male	Perithecia, asci, ascospores
E62	MAT1-2	I-R-22	858	0	Female	Perithecia, asci, ascospores
K1	MAT1-2	4136-4-3	700	800	Hermaphrodite	Perithecia, asci, ascospores
K1	MAT1-2	TH3	851	900	Hermaphrodite	Perithecia, asci,

						ascospores
K1	MAT1-2	I-R-22	450	700	Hermaphrodite	Perithecia, asci, ascospores
K6	MAT1-2	4136-4-3	1500	0	Female	Perithecia, asci, ascospores
K6	MAT1-2	TH3	1500	1000	Hermaphrodite	Perithecia, asci, ascospores
K6	MAT1-2	I-R-22	1500	1200	Hermaphrodite	Perithecia, asci, ascospores
K7	MAT1-2	TH3	0	800	Male	Perithecia, asci, ascospores
K7	MAT1-2	I-R-22	0	900	Male	Perithecia, asci, ascospores
K8	MAT1-2	4136-4-3	0	1000	Male	Perithecia, asci, ascospores
K8	MAT1-2	TH3	0	700	Male	Perithecia, asci, ascospores
K8	MAT1-2	I-R-22	0	800	Male	Perithecia, asci, ascospores
K10	MAT1-1	K23/123	0	2500	Male	Perithecia, asci, ascospores
K10	MAT1-1	D15/s47	0	1500	Male	Perithecia, asci, ascospores
K13	MAT1-2	TH3	0	1056	Male	Perithecia, asci, ascospores
K13	MAT1-2	I-R-22	0	1000	Male	Perithecia, asci, ascospores
K14	MAT1-2	I-R-22	0	1000	Male	Perithecia, asci, ascospores
K17	MAT1-2	TH3	0	506	Male	Perithecia, asci, ascospores
K17	MAT1-2	I-R-22	0	1000	Male	Perithecia, asci, ascospores
K19	MAT1-1	K23/123	0	2000	Male	Perithecia, asci, ascospores
K19	MAT1-1	D15/s47	0	2000	Male	Perithecia, asci, ascospores
K20	MAT1-2	TH3	0	700	Male	Perithecia, asci, ascospores
K20	MAT1-2	I-R-22	0	1000	Male	Perithecia, asci, ascospores
K21	MAT1-1	D15/s47	1500	1500	Hermaphrodite	Perithecia, asci, ascospores
K24	MAT1-2	TH3	0	1150	Male	Perithecia, asci, ascospores
K24	MAT1-2	I-R-22	0	2005	Male	Perithecia, asci, ascospores
K25	MAT1-1	K23/123	1790	2000	Hermaphrodite	Perithecia, asci, ascospores

K25	MAT1-1	D15s47	1300	1795	Hermaphrodite	Perithecia, asci, ascospores
K26	MAT1-2	4136-4-3	0	600	Male	Perithecia, asci, ascospores
K26	MAT1-2	TH3	0	890	Male	Perithecia, asci, ascospores
K26	MAT1-2	I-R-22	0	500	Male	Perithecia, asci, ascospores
K27	MAT1-2	TH3	0	800	Male	Perithecia, asci, ascospores
K27	MAT1-2	I-R-22	0	700	Male	Perithecia, asci, ascospores
K30	MAT1-2	4136-4-3	0	200	Male	Perithecia, asci, ascospores
K30	MAT1-2	TH3	0	100	Male	Perithecia, asci, ascospores
K30	MAT1-2	I-R-22	0	200	Male	Perithecia, asci, ascospores
K31	MAT1-1	K23/123	0	1200	Male	Perithecia, asci, ascospores
K32	MAT1-2	TH3	2500	2550	Hermaphrodite	Perithecia, asci, ascospores
K32	MAT1-2	I-R-22	0	1680	Male	Perithecia, asci, ascospores
K36	MAT1-1	D15/s47	300	1200	Hermaphrodite	Perithecia, asci, ascospores
K37	MAT1-2	4136-4-3	0	450	Male	Perithecia, asci, ascospores
K37	MAT1-2	TH3	0	800	Male	Perithecia, asci, ascospores
K37	MAT1-2	I-R-22	0	800	Male	Perithecia, asci, ascospores
K39	MAT1-1	K23/123	0	950	Male	Perithecia, asci, ascospores
K40	MAT1-1	K23/123	0	2750	Male	Perithecia, asci, ascospores
K40	MAT1-1	D15/s47	0	2500	Male	Perithecia, asci, ascospores
K44	MAT1-2	4136-4-3	0	1200	Male	Perithecia, asci, ascospores
K44	MAT1-2	TH3	0	1700	Male	Perithecia, asci, ascospores
K44	MAT1-2	I-R-22	0	2000	Male	Perithecia, asci, ascospores
K45	MAT1-2	4136-4-3	0	800	Male	Perithecia, asci, ascospores
K45	MAT1-2	I-R-22	800	2560	Male	Perithecia, asci, ascospores
T1	MAT1-2	TH3	0	500	Male	Perithecia, asci,

						ascospores
T1	MAT1-2	I-R-22	0	1200	Male	Perithecia, asci, ascospores
T2	MAT1-2	TH3	0	650	Male	Perithecia, asci, ascospores
T2	MAT1-2	I-R-22	0	1000	Male	Perithecia, asci, ascospores
T5	MAT1-1	K23/123	400	800	Hermaphrodite	Perithecia, asci, ascospores
T5	MAT1-1	D15/s47	400	2100	Hermaphrodite	Perithecia, asci, ascospores
T6	MAT1-1	K23/123	0	600	Male	Perithecia, asci, ascospores
T6	MAT1-1	D15/s47	0	550	Male	Perithecia, asci, ascospores
T9	MAT1-2	4136-4-3	0	100	Male	Perithecia, asci, ascospores
T9	MAT1-2	I-R-22	0	800	Male	Perithecia, asci, ascospores
T10	MAT1-2	I-R-22	0	800	Male	Perithecia, asci, ascospores
T11	MAT1-2	4136-4-3	0	150	Male	Perithecia, asci, ascospores
T11	MAT1-2	TH3	0	1400	Male	Perithecia, asci, ascospores
T12	MAT1-2	TH3	0	100	Male	Perithecia, asci, ascospores
T12	MAT1-2	I-R-22	0	800	Male	Perithecia, asci, ascospores
T14	MAT1-2	I-R-22	0	200	Male	Perithecia, asci, ascospores
T15	MAT1-2	I-R-22	0	920	Male	Perithecia, asci, ascospores
T16	MAT1-2	I-R-22	0	800	Male	Perithecia, asci, ascospores
T17	MAT1-2	4136-4-3	800	0	Female	Perithecia, asci, ascospores
T17	MAT1-2	TH3	500	1200	Hermaphrodite	Perithecia, asci, ascospores
T17	MAT1-2	I-R-22	1150	1200	Hermaphrodite	Perithecia, asci, ascospores
T18	MAT1-1	K23/123	0	150	Male	Perithecia, asci, ascospores
T18	MAT1-1	D15/s47	0	200	Male	Perithecia, asci, ascospores
T21	MAT1-1	K23/123	0	200	Male	Perithecia, asci, ascospores
T21	MAT1-1	D15/s47	0	500	Male	Perithecia, asci, ascospores

T22	MAT1-2	4136-4-3	400	0	Female	Perithecia, asci, ascospores
T22	MAT1-2	I-R-22	0	600	Male	Perithecia, asci, ascospores
T24	MAT1-1	K23/123	0	2600	Male	Perithecia, asci, ascospores
T26	MAT1-1	K23/123	0	100	Male	Perithecia, asci, ascospores
T26	MAT1-1	D15/s47	0	100	Male	Perithecia, asci, ascospores
T29	MAT1-1	K23/123	0	500	Male	Perithecia, asci, ascospores
T31	MAT1-1	D15/s47	0	100	Male	Perithecia, asci, ascospores
T32	MAT1-1	D15/s47	0	100	Male	Perithecia, asci, ascospores
T35	MAT1-2	I-R-22	0	500	Male	Perithecia, asci, ascospores
T36	MAT1-1	K23/123	500	100	Hermaphrodite	Perithecia, asci, ascospores
T36	MAT1-1	D15/s47	100	0	Female	Perithecia, asci, ascospores
T42	MAT1-1	D15/s47	0	400	Male	Perithecia, asci, ascospores
T47	MAT1-2	TH3	900	0	Female	Perithecia, asci, ascospores
T50	MAT1-2	I-R-22	0	900	Male	Perithecia, asci, ascospores
T51	MAT1-2	I-R-22	0	300	Male	Perithecia, asci, ascospores
T53	MAT1-1	K23/123	0	100	Male	Perithecia, asci, ascospores
T58	MAT1-2	4136-4-3	0	100	Male	Perithecia, asci, ascospores
T58	MAT1-2	TH3	0	150	Male	Perithecia, asci, ascospores
T58	MAT1-2	I-R-22	0	100	Male	Perithecia, asci, ascospores
U1	MAT1-1	K23/123	1006	650	Hermaphrodite	Perithecia, asci, ascospores
U2	MAT1-2	TH3	0	1200	Male	Perithecia, asci, ascospores
U2	MAT1-2	I-R22	0	700	Male	Perithecia, asci, ascospores
U3	MAT1-2	4136-4-3	0	100	Male	Perithecia, asci, ascospores
U3	MAT1-2	TH3	0	900	Male	Perithecia, asci, ascospores
U3	MAT1-2	I-R-22	0	400	Male	Perithecia, asci,

						ascospores
U5	MAT1-2	4136-4-3	500	600	Hermaphrodite	Perithecia, asci, ascospores
U5	MAT1-2	TH3	490	550	Hermaphrodite	Perithecia, asci, ascospores
U5	MAT1-2	I-R-22	500	570	Hermaphrodite	Perithecia, asci, ascospores
U6	MAT1-1	K23/123	500	1115	Hermaphrodite	Perithecia, asci, ascospores
U6	MAT1-1	D15/s47	590	1300	Hermaphrodite	Perithecia, asci, ascospores
U7	MAT1-1	K23/123	0	1300	Male	Perithecia, asci, ascospores
U7	MAT1-1	D15/s47	0	1150	Male	Perithecia, asci, ascospores
U8	MAT1-1	K23/123	0	1054	Male	Perithecia, asci, ascospores
U8	MAT1-1	D15/s47	0	1204	Male	Perithecia, asci, ascospores
U9	MAT1-1	K23/123	0	1500	Male	Perithecia, asci, ascospores
U9	MAT1-1	D15/s47	0	1250	Male	Perithecia, asci, ascospores
U10	MAT1-2	TH3	0	600	Male	Perithecia, asci, ascospores
U10	MAT1-2	I-R-22	0	510	Male	Perithecia, asci, ascospores
U11	MAT1-1	K23/123	0	1057	Male	Perithecia, asci, ascospores
U11	MAT1-1	D15/s47	0	968	Male	Perithecia, asci, ascospores
U13	MAT1-1	K23/123	500	1205	Hermaphrodite	Perithecia, asci, ascospores
U13	MAT1-1	D15/s47	0	456	Male	Perithecia, asci, ascospores
U15	MAT1-1	K23/123	0	500	Male	Perithecia, asci, ascospores
U19	MAT1-1	D15/s47	0	250	Male	Perithecia, asci, ascospores
U20	MAT1-2	TH3	0	400	Male	Perithecia, asci, ascospores
U20	MAT1-2	I-R-22	0	800	Male	Perithecia, asci, ascospores
U21	MAT1-2	4136-4-3	700	0	Male	Perithecia, asci, ascospores
U21	MAT1-2	TH3	600	800	Hermaphrodite	Perithecia, asci, ascospores
U21	MAT1-2	I-R-22	900	1251	Hermaphrodite	Perithecia, asci, ascospores

U22	MAT1-1	K23/123	0	400	Male	Perithecia, asci, ascospores
U22	MAT1-1	D15/s47	0	500	Male	Perithecia, asci, ascospores
U23	MAT1-1	K23/123	0	400	Male	Perithecia, asci, ascospores
U25	MAT1-2	TH3	0	100	Male	Perithecia
U25	MAT1-2	I-R-22	0	250	Male	Perithecia
U27	MAT1-1	K23/123	0	1200	Male	Perithecia
U27	MAT1-1	D15/s47	0	400	Male	Perithecia
U29	MAT1-2	TH3	0	450	Male	Perithecia
U30	MAT1-1	K23/123	0	250	Male	Perithecia, asci, ascospores
U30	MAT1-1	D15/s47	0	200	Male	Perithecia, asci, ascospores
U31	MAT1-2	TH3	0	500	Male	Perithecia
U31	MAT1-2	I-R-22	0	320	Male	Perithecia
U35	MAT1-2	I-R-22	0	250	Male	Perithecia
U36	MAT1-2	TH3	0	400	Male	Perithecia, asci, ascospores
U36	MAT1-2	I-R-22	0	450	Male	Perithecia, asci, ascospores
U37	MAT1-2	4136-4-3	0	350	Male	Perithecia, asci, ascospores
U37	MAT1-2	I-R-22	0	350	Male	Perithecia, asci, ascospores
U38	MAT1-2	TH3	0	100	Male	Perithecia
U39	MAT1-2	I-R-22	0	900	Male	Perithecia
U40	MAT1-1	K23/123	0	800	Male	Perithecia, asci, ascospores
U40	MAT1-1	D15/s47	0	900	Male	Perithecia, asci, ascospores
U42	MAT1-2	TH3	0	1000	Male	Perithecia, asci, ascospores
U42	MAT1-2	I-R-22	0	400	Male	Perithecia
U43	MAT1-2	TH3	0	450	Male	Perithecia, asci, ascospores
U43	MAT1-2	I-R-22	0	250	Male	Perithecia, asci, ascospores
U44	MAT1-2	TH3	0	350	Male	Perithecia, asci, ascospores
U44	MAT1-2	I-R-22	0	100	Male	Perithecia
U48	MAT1-1	K23/123	0	720	Male	Perithecia
U49	MAT1-1	D15/s47	0	1230	Male	Perithecia
U52	MAT1-1	K23/123	0	200	Male	Perithecia
U53	MAT1-1	K23/123	0	1200	Male	Perithecia, asci, ascospores
U53	MAT1-1	D15/s47	400	1400	Hermaphrodite	Perithecia, asci, ascospores

U54	MAT1-1	D15/s47	0	700	Male	Perithecia, asci, ascospores
U56	MAT1-2	TH3	0	950	Male	Perithecia, asci, ascospores
U56	MAT1-2	I-R-22	0	100	Male	Perithecia
U57	MAT1-2	TH3	700	1000	Hermaphrodite	Perithecia, asci, ascospores

*Isolates were collected from: E-Ethiopia, K- Kenya, T- Tanzania and U- Uganda.

[a]Isolate mating type determined by the PCR assay.

[b]Tester isolates of opposite mating type including isolate D15/s47 identified as a mating tester in this study have the capability to cross with new isolates and produce perithecia.

[c]Perithecia is the fruiting body of *M. oryzae* resulting from the cross of two compatible isolates of opposite mating type.

[d]Sexual behaviour was determined according to the nomenclature by Itoi *et al*. (1983) where, Hermaphrodites – Perithecia produced on the side of the mating tester as well as the new isolate; Female – Perithecia produced by the new isolate only; Male – Perithecia produced by the mating tester only.

[e]Includes examples where isolates and crosses produced barren perithecia without asci and ascospores, these are denoted as Perithecia in the corresponding rows.

Among the fertile *M. oryzae* isolates identified overall in Ethiopia, Kenya, Uganda and Tanzania, in terms of sexual behaviour a major proportion of the crosses led to their identification in terms of sexual behaviour as males (175), followed by hermaphrodite (46) and female (10) as summarised below (Table 5.5). However, there were some variations in these trends among the 4 countries. For example, the data for Kenya reflects a higher proportion of hermaphrodites, where as in Uganda no female isolates were identified (Table 5.5).

Table 5.5. Summary of sexual behaviour of the fertile *M. oryzae* isolates identified among the contemporary population from Ethiopia, Kenya, Tanzania and Uganda

Country	**Sexual behaviour of the FMB pathogen *M. oryzae* isolates based on the number of crosses**			**Total***
	Male[a]	Female[b]	Hermaphrodite[c]	
Ethiopia	46	5	20	71
Kenya	39	1	10	50
Tanzania	35	4	5	44
Uganda	55	0	11	66
*Total	175	10	46	

[a]Male isolate that supported the production of perithecia only by the tester isolate.

[b]Female isolates that formed perithecia only by the new isolates.

[c]Hermaphrodite isolates that formed perithecia both by the new isolate and the tester isolate.

[*]Numbers reflect the sexual behaviour of each isolate with one or more testers (and not related to the total number of isolates collected and characterised from each country or the four countries together).

5.4 Discussion

Work in this chapter focused on assessing the sexual reproductive potential of the finger millet blast (FMB) pathogen, utilising 224 *M. oryzae* isolates representing the contemporary populations (2015-2017) from Eastern Africa. These isolates were collected in four countries including Kenya (45), Tanzania (58), Uganda (58) and Ethiopia (63). Integrated use of the mating type specific PCR and mating culture assays was effective in addressing the objectives of the study. A number of investigations have been carried out to assess the reproductive behaviour in *M. oryzae* isolates associated with crop and grass hosts in diverse geographic locations (Karthikeyan and Gnanamanickam, 2008b; Zeng *et al*., 2009; Tredway *et al*., 2003; Consolo *et al*., 2005; Takan *et al*., 2012; Onaga *et al*., 2015; Kema *et al*., 2018). However, only some studies have used the combination of the two methodological approaches required to develop a fuller assessment of the blast pathogen sexual reproductive potential (e.g. Takan *et al*., 2012).

The PCR assay clearly showed the presence of both mating types at a high proportion (MAT 1-1, 56 % and MAT 1-2, 44 %) among the 224 FMB pathogen isolates from Kenya, Tanzania, Uganda and Ethiopia in Eastern Africa (Figures 5.2 and 5.3). The overall pattern is similar to previous findings where the presence of both mating types at a high proportion (MAT 1-1, 47 % and MAT 1-2, 53 %) was recorded in Kenya and Uganda during 2000-2004 (Takan *et al*., 2012). Despite the similarity in overall pattern, there is a noticeable difference in the proportion of the two mating types between the two studies reflective of potential differences associated with space and time. This is highlighted by the clear differences in the proportion of the mating types MAT 1-1 and MAT 1-2 across the four countries in Eastern Africa. MAT 1-1 isolates were dominant in 3 countries, ranging between ~57 % in Uganda

and ~68 % in Ethiopia with the corresponding range of MAT 1-2 isolates from 38 % to 43%. Conversely, in Tanzania, MAT 1-2 isolates were dominant at ~60 % and MAT 1-1 isolates were 40 % (Table 5.1). Mating type distribution pattern among the FMB pathogen populations varied considerably in different countries in Asia. For example, in India, where finger millet has a long cultivation history (~3000 years) as in Eastern Africa (~5000 years) presence of both mating types at a high proportion MAT 1-1, 54 % and MAT1-2, 46 % among 28 isolates) was recorded (Viji and Gnanamanickam, 1998). Limited evidence available in Japan and Nepal, where the cultivation of finger millet is considered more recent, suggests that the distribution pattern of the two mating types MAT 1-1 and MAT 1-2 is highly skewed with the domination of MAT 1-1 isolates (Yaegashi and Udagawa, 1978).

In the case of blast pathogen populations from other crops and grass hosts, the mating type distribution pattern again varies according to the host as well as the geographic location. Occurrence of both mating types at variable proportions has been reported among the blast pathogen isolates infecting rice and wheat at different geographical locations such as East Africa, West Africa and Brazil (Takan *et al*., 2012; Onaga *et al*., 2015; Urashina *et al*., 1993; Maciel *et al*., 2014, Mwongera, 2018). Further, in rice, wheat, perennial ryegrass and St. Augustine turf grass blast pathogen populations, where both mating types were recorded, the common pattern is highly skewed distribution either dominated by MAT 1-1 or MAT 1-2 (Nottenghem and Silue, 1992; Kumar *et al*., 1999; Dayakar *et al*., 2000; Tredway *et al*., 2003; Takan *et al*., 2012; Maciel *et al*., 2014; Onaga *et al*., 2015, Mwongera, 2018). For example, In East Africa, among 80 isolates of the rice blast pathogen, only 17 were MAT 1-2 mating type (Onaga *et al*., 2015). In addition, the presence of only one mating type was recorded among the blast pathogen isolates infecting rice and perennial ryegrass in Argentina, Thailand and

USA (Viji and Uddin 2002; Consolo *et al*., 2005; Peixoto, 2014; D'Avilla *et al*., 2016; Sirisathaworn *et al*., 2017).

It is important to emphasise that, despite the overall high proportion of occurrence of both mating types at a country level or regional level in Eastern Africa, their distribution pattern varies considerably at a local or district level within a country (Table 5.1). For instance, high proportion of both mating types was recorded in locations such as Busia and Kisii in Kenya and Nkasi in Tanzania. Conversely, dominance of one of the mating types was recorded in districts such as Diga and Banja in Ethiopia and Siaya and Kusumu in Kenya. Finally, presence of only one mating type was recorded in some areas and the examples include Waju Tuka and Bahir Dar Zuria in Ethiopia, and Ngora and Manafwa in Uganda. The number of isolates characterised from each location is a caveat in the wider interpretation of these distinctive patterns at a local level. Nonetheless, there is some commonality to the patterns recorded from contemporary populations (2015-2017) of the FMB pathogen in Eastern Africa and the data from historical populations (2000-2004) within the region. For example, occurrence of both mating types at a high proportion in the Busia district, Kenya as recorded in the present study as well as previously (Takan *et al*., 2012). A similar pattern was observed among the rice blast pathogen isolates from the same location in Asia with near equal distribution of both mating types among the contemporary and historical populations. On this basis, the occurrence of sexual reproduction has been suggested (Saleh *et al*., 2012). Likewise, based on the wide occurrence of both mating types in wheat blast pathogen populations in Brazil, scope for the episodic occurrence of sexual reproduction among regular cycles of asexual reproduction has been proposed (Maciel *et al*., 2014).

Presence of both mating types within a location is reflective of the opportunity for the occurrence of sexual reproduction (Saleh *et al*., 2012; Maciel *et al*., 2014; Saleh *et al*., 2014). However, it is well recognised that, this pattern alone will not clearly reveal the sexual reproductive capability of fungal populations (e.g. Couch *et al*., 2005; Maciel *et al*., 2014). Thus, in the present study, the fertility status of the contemporary population was determined in mating assays (Table 5.2). The results highlighted the presence of both fertile and infertile *M. oryzae* isolates in Eastern Africa across the four countries with the fertile isolates occurring at a higher frequency (~ 60 %, Figure 5.5 and Table 5.3). This pattern is similar to the previous findings, where a high proportion of fertile isolates (84 to 89 %) was found in the FMB pathogen in Kenya and Uganda (Takan *et al*., 2012). Similarly, in India, field isolates of *M. oryzae* infecting finger millet and wild millet have generally been reported to be very fertile (Viji and Gnanmannickam, 1998, Karthikeyan and Gnanamanickam, 2008). Likewise, high fertility status of *M. oryzae* isolates has also been reported from grass hosts such as weeping love grass and goosegrass (Crawford *et al*., 1986; Tredway *et al*., 2003). However, the pattern differs markedly among the rice blast pathogen populations in diverse locations, where the level of fertile isolates is low. For instance, only 33 % of the rice blast pathogen isolates were fertile in Thailand (Mekwatanakan *et al*., 1999), 39.6 % of the *M. oryzae* isolates from rice were fertile in India (Kumar *et al*., 1999) and only 7.2 % of the isolates were fertile in Argentina (Consolo *et al*., 2005).

Among the FMB pathogen isolates characterised in this study, ~40 % were infertile. *M. oryzae* isolates that are infertile have been commonly reported with the rice blast system in diverse locations (Mekwatanakan *et al*., 1999; Kumar *et al*., 1999; Consolo *et al*., 2005) and also with other hosts such as finger millet and foxtail millet in India and East Africa (Viji and

Gnanamanickam, 1998; Karthikeyan and Gnanamanickam, 2008; Takan *et al*., 2012). It has been suggested that the infertility observed might have been due to the incompatibility of the isolates with the mating testers used in these assays (Viji and Gnanamanickam, 1998). However, the influence of genetic changes leading to infertility is an equally important consideration, particularly based on the use of well-characterised testers (Milgroom, 2015).

Within Eastern Africa, based on the present sampling during 2015-2017, the range of fertile *M. oryzae* isolates of the FMB pathogen varied from 50 % in Tanzania to 74 % in Ethiopia (Table 5.3), which is lower than the range reported in Kenya and Uganda (84% - 89%) based on field sampling during 2000 – 2004. Further, the high fertility status of *M. oryzae* isolates from weeping love grass and goosegrass has been attributed mainly to the presence of hermaphrodites (Crawford *et al*., 1986; Tredway *et al*., 2003). Similarly, a high proportion of hermaphroditic *M. oryzae* isolates (64% of the fertile isolates) from finger millet has been recorded in Kenya and Uganda (Takan, 2007; Takan *et al*., 2012). However, in this study, among the fertile isolates, the highest proportion behaved as males, followed by hermaphrodites, and females (Table 5.4 and 5.5). In general, with the rice blast system in various locations, occurrence of hermaphroditic isolates is very rare and normally *M. oryzae* isolates behaving as the males is common (e.g. Consolo *et al*., 2005). Some exceptions have been reported as in the case of specific rice blast pathogen populations in a part of China and India with the presence of a very high frequency of female fertile isolates (Saleh *et al*., 2012).

Increase in the proportion of *M. oryzae* isolates associated with finger millet in Eastern Africa behaving as fertile males suggests a loss of hermaphroditic and female reproductive behaviour. It is important to point out that it was not possible to establish a direct

correlation between the mating type distribution pattern and the fertility status of the *M. oryzae* isolates within a location. The effect of the space and time as well as underlying genetic processes are key considerations in this context. Further, in this study, a proportion of fertile isolates produced barren perithecia (without asci and/or ascospores, Table 5.4). The perithecial barrenness has been ascribed to genetic differences in compatibility and genetic abnormalities including chromosomal duplications and mutations (Saleh *et al*., 2012). The loss of fertility over space and time in crop fungal pathogens such as *M. oryzae* has been attributed to genetic processes as well as agro-ecological contexts such as a selective advantage of female sterile strains and host selection pressure (Saleh *et al*., 2012). Thus, the reproductive cycle including the occurrence of sexual reproduction in *M. oryzae* populations depend on a whole range of factors including the isolate genotype, compatibility, genetic regulation, and agro-ecological contexts including gene and genotype flow (Viji and Gnanamanickam, 1998; Milgroom, 2015; Saleh *et al*., 2012; Bazin *et al*., 2014). The overall pattern of mating type distribution and the fertility status of the contemporary pathogen populations in the FMB pathogen in Eastern Africa reflects a strong potential for random mating as reported in other plant and animal pathogens (e.g. Kema *et al*., 2018; Persinoti *et al*., 2018).

In conclusion, the present study confirmed the wide occurrence of a high proportion of both mating types MAT 1-1 and MAT 1-2 among the *M. oryzae* isolates associated with finger production in Eastern Africa. The frequency of occurrence of the two mating types vaired across the 4 countries and in various locations/districts within the countries. Presence and wide distribution of a high proportion of fertile *M. oryzae* isolates was recorded among the contemporay populaitons of the FMB pathgoen. Among the fertile isolates, a higher

proportion of isolates showed sexual behaviour as males, followed by hermaphrodites and females. A comparison of the data with the contemporary populaitons (2015-17) with previous data from historic populaitons (2000-04) suggest a decrease in fertility level as well as the proportion of hermaphrodites and females. The overall pattern of the mating type distribution and the fertility status of the *M. oryzae* isolates associated with finger millet blast in Easteran Africa shows strong potential for a mixed reproductive system including asexual reproduction and episodic sexual reproduction and recombination.

Chapter 6

6.0 General Discussion, Conclusions and Future Perspectives

It is now widely recognized that finger millet originated in the highland stretches across Ethiopia and Uganda and was first domesticated in Eastern Africa, where landraces and traditional varietal mixtures have been cultivated for ~5000 years. Finger millet was subsequently introduced into Asia, particularly India ~ 3000 years ago and the crop is widely cultivated in certain parts (Hilu *et al*., 1979; de Wet *et al*., 1984; Gimode *et al*., 2016). Thus, Eastern Africa and India are recognised as the primary and secondary centres of finger millet diversity, respectively (Mathur *et al*., 2012).

There is a growing recognition of the food and nutritional security potential of finger millet both in Africa and India (Onyango, 2016; Wafula *et al*., 2018). However, in Eastern Africa as well as in India, blast disease caused by *M. oryzae* constitutes one of the major constraints to finger millet production and causes substantial loss of grain, which if saved is enough to feed millions of people annually (Lenne *et al*., 2007; Takan, 2007; Takan *et al*., 2012). Compared to the considerable progress made in deciphering the population biology and infection biology of major systems such as rice blast and wheat blast, research into these aspects with the finger millet systems lags behind significantly.

Integrating the application of molecular genetic, genomic and computational analsyses as well as biological assays, the present study characterised a distinctive set of 300 isolates of the finger millet blast (FMB) pathogen *M. oryzae* from Eastern Africa. This set includes 76 isolates representing historical populations from 2000-2004 and 224 isolates representing contemporary populations from 2015-2017, covering four key finger millet producing

countries Ethiopia, Kenya, Tanzania, and Uganda in Eastern Africa. Furthermore, it is important to emphasise that, based on the information available, the FMB pathogen isolates characterised in this study have originated from a real mixture of landraces, local varieties and some improved varieties across the locations surveyed in Eastern Africa. For example, in Tanzania and Ethiopia, the blast pathogen isolates almost exclusively represent landraces, at least where known. Whilst in Kenya and Uganda, the blast isolates represent a wider spectrum of improved varieties, local varieties and landraces. In Eastern Africa, movement of finger millet seeds and/or plant material via cultural exchanges, research for development programmes and trade is well known (Lenne *et al*., 2007; Gimode *et al*., 2016). Population diversity and structure of fungal pathogens of crops is known to be influenced by the historical background of the host as well as the agro-environmental conditions of the geographical locations (McDonald, 1997). The high genotype diversity, distribution of some genotypes across country boundaries and the overall continuous genetic variation pattern allied to the high proportion of the two mating types and the high fertility status of the FMB pathogen in Eastern Africa is typical of the patterns recorded with the blast pathogen in the centres of diversity of rice in Asia and wheat in Brazil (Kumar *et al*., 1999; Mekwatanakarn, *et al*., 1999; Chen *et al*., 2006; Maciel *et al*., 2014). Further, the genetic diversity and population structure recorded in Eastern Africa is also reflective of genotype and/or gene flow due to pathogen movement linked to anthropogenic activities (Nangoti, *et al*., 2004; Takan *et al.,* 2012; Gimode *et al*., 2016).

In this context, some of the key patterns from the population diversity analysis in Chapter 3, the genome-wide analysis in Chapter 4 and the assessment of sexual reproductive capacity in Chapter 5 provide a holistic view of the finger millet blast pathogen biology, genetics and

evolution particularly in Eastern Africa (e.g. Table 6.1). The MLS analysis used has been previously utilised to analyse populations of major fungal pathogens including *M. oryzae* (Couch and Kohn, 2002; Couch *et al*., 2005; Armitage *et al*., 2015; Baroncelli *et al*., 2015; Baroncelli *et al*., 2017; Luo *et al*., 2017; Almiman, 2018; Jagadeesh *et al*., 2018a; Mannai *et al*., 2018). MLS phylogenies distinguished the FMB pathogen associated with finger millet in Eastern Africa into two sub-populations (Groups A and B) that are geographically clustered. Group A contained ~90 % of *M. oryzae* isolates from finger millet in Kenya and Uganda. Both mating types were recorded at a high proportion in both countries. In the PCR assay, the vast majority of the isolates from this group tested *grh* negative (Table 6.1). Group B mostly contained *M. oryzae* isolates from Ethiopia and Tanzania. Nearly ~96 % of these were *grh* positive in the PCR assay. Both mating types were present at high proportion in Tanzania and in Ethiopia MAT 1-1 type was predominant (Table 6.1).

Phylogenomic analysis based on 944 BUSCO genes also divided the 18 representative isolates into two groups (Clade 1Ai and 1Aii; Figure 4.11) further substantiating the pattern of two sub-populations identified by the MLS analysis discussed above. In addition, the two ranges of genome-wide SNPs identified also divided the 18 representative isolates into two groups broadly corresponding to the two sub-populations/groups discussed above. Further, Ethiopia and Tanzania have a high proportion of *grh* element containing *M. oryzae* isolates and at least two types of *grh* element with significant differences in their sequence and genomic architecture are present in different *M. oryzae* isolates within Eastern Africa within the context of the two sub-populations/groups considered above. This emerging pattern of the two sub-populations overall merits further phylogenomic investigations of the FMB pathogen movement within and across the African and Asian continents.

Table 6.1. Overview of the characteristics of the blast pathogen *M. oryzae* isolates associated with finger millet production in Eastern Africa

Group*	Total isolates per group	Nucleotides differences (%)**	Country	Total isolates per country	*Grasshopper* (*grh*) element PCR***		Mating type (PCR)****		Population type
					Positive	Negative	MAT 1-1	MAT 1-2	
A	181	0-3	Uganda	97	2	95	50	47	H & C
			Kenya	66	8	58	36	30	H & C
			Tanzania	9	0	9	3	6	C
			Ethiopia	9	2	7	3	6	C
B	111	0-11	Ethiopia	54	53	1	39	15	C
			Tanzania	48	48	0	21	27	C
			Kenya	6	6	0	3	3	H
			Uganda	3	0	3	1	2	H

H, historical population (2000-2004) and C, contemporary population (2015-2017);

Eight isolates were not represented in either Group A or B, and these include five isolates from Kenya (H & C), two from Uganda (H), and one from Tanzania;

*The two sub-populations identified in the multi-locus sequence analysis (Figure 3.12 and Table 3.8C); Further details of the isolates available in Table 2.1A.

**Percentage of nucleotide sequence variation within a group

****Grasshopper* (*grh*) element: *M. oryzae* isolates identified as *grh* positive (*grh*+) or *grh* negative (*grh*-) based on the presence or absence of the two amplicons in the PCR screening with the PKE and PES primer pairs

****Mating types: *M. oryzae* isolates identified as MAT 1-1 or MAT 1-2 using the mating type specific PCR (Figure 5.3 and Table 5.1).

The *grh* element and mating type data for the historical isolates were adopted from previous work (Takan, 2007).

The overall genetic variation pattern of the *M. oryzae* pathogen populations associated with finger millet production in Eastern Africa linked to sexual reproduction potential aligns well with the model of ancient and recombining populations as previously suggested with the rice and finger millet blast system in the Indian Himalayas and East Africa (Kumar *et al*., 2009; Takan *et al*., 2012). The high proportion of MAT 1-1 and MAT 1-2 and the presence of a high proportion of fertile isolates among the *M. oryzae* isolates associated with finger millet production in Eastern Africa are typical of the patterns observed in other plant and animal fungal pathogen where strong potential for random mating has been suggested (Kumar *et al*., 1999; Couch *et al*., 2005; Paoletti *et al*., 2005; Fraser *et al*., 2007; O'Gorman *et al*., 2009; Amorium *et al*., 2017; Kema *et al*., 2018; Persinoti *et al*., 2018). The presence of both mating type isolates in the two sub-populations identified (Table 6.1) suggests the possiblity for increase in genetic diversity over time through sexual reproduction potential linked to enhanced pathogenic diversity leading to the break-down of host resistance. Further, formation of perithecia on dead plant tissue has been shown by experimental co-inoculation of opposite mating type isolates of *M. oryzae* on rice plants (Silue and Notteghem, 1990; Hayashi *et al*., 1997). These reports highlight the need for effective clearance of the crop debris in locations such as Eastern Africa for the effective control of finger millet blast disease as well as to reduce the chances of episodic sexual reproductive cycles. With wheat blast system in Brazil, a mixed reproductive system including asexual cycles and episodic sexual cycle has been reported following the wide occurrence of both mating types allied to high pathogen genetic diversity (Maciel *et al*., 2014).

The two *M. oryzae* isolates E34 and K23/123, used to develop the reference genomes, showed putative translocations that could lead to genomic rearrangements and similar patterns have been reported extensively in fungi, e.g. a large fraction of translocation occurring among or between *Colletotrichum* species isolates (Rao and Nandineni, 2017; Buiate *et al*., 2017). Identification of the genes present in these locations, as well as in the putative unique genomic fragments identified in E34 and K23/123 would necessitate further refinement of gene prediction using algorithms such as MAKER and BRAKER and comparative analysis (Cantarel *et al*., 2008; Campbell *et al*., 2014; Hoff *et al*., 2016). Likewise, the identification of allelic versions of the *AVR-Piz-T* gene among the 18 representative isolates suggests the scope for further analysis in the finger millet blast system as the existence of allelic versions of AVR genes and their role in the rice blast system is well recognized (e.g. Khang *et al*., 2008; Chuma *et al*., 2011; Huang *et al*., 2014; Bialas *et al*., 2017).

This research work has achieved significant advances in the knowledge of the finger millet blast pathogen genetic diversity, population structure and phylogenetic relationships as well as the sexual reproductive capability across four countries in Eastern Africa countries. Further, a considerable level of genomic resources and key opportunities for further computational and functional analysis as well as phylogenomic of the pathogen origin and evolution has been generated. These knowledge and resources will feed into ongoing and future efforts for effective finger millet blast management by various stakeholders such as the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) and National Agricultural Research and Extension Services (NARES). The progress achieved in this study with the FMB pathogen genetics and biology would provide a critically important framework

for further research in the identification, development and deployment of host resistance in Eastern Africa. The new resources developed in this work will also provide a baseline for future monitoring of the FMB pathogen populations in Sub-Saharan Africa, where the disease is viewed as a threat to the food and nutrition security potential of finger millet.

6.2 Conclusions and Future Perspectives

In conclusion, this study has led to the development of new knowledge and resources presented below:

1. Two novel genetic markers HyP1 and HyP2 were identified and developed for the finger millet blast (FMB) pathogen *M. oryzae*. These markers were phylogenetically informative with a high content of SNPs and offered high resolution applicable in deciphering the pathogen genotype diversity and distribution pattern across different countries.

2. The HyP1 and HyP2 markers were used in conjunction with the ITS and HIS4 markers selected from known phylogenetic markers.

3. Single- and multi-locus analysis provided a clear assessment of the genotype diversity and distribution pattern both at a country level and a regional overview in Eastern Africa.

4. ITS and HIS4 identified 7-9 genotypes among the 300 isolates analysed, whilst HyP1 and Hyp2 identified 80-85 genotypes. This pattern clearly shows that the ITS and

HIS4 markers reflects the conserved areas and the HyP1 and HyP2 markers reflect the highly variable regions in the genomes of the FMB pathogen.

5. The genotype diversity and their distribution pattern have been deciphered in each of the four countries surveyed in Eastern Africa based on the Multi-locus sequence (MLS) analysis. Certain genotypes included *M. oryzae* isolates from two or more countries. Further analysis of the data enabled the clustering of these genotypes into genetic groups in each country.

6. Multi-locus sequence (MLS) analysis revealed high genetic diversity of the FMB pathogen in Eastern Africa with 207 genotypes, which showed a continuous genetic variation pattern.

7. Vast majority of the genotypes identified in Eastern Africa were distinguished into two sub-populations designated as Groups A and B with geographic clustering. Group A predominantly included *M. oryzae* isolates from Kenya and Uganda, whilst the Group B predominantly included pathogen isolates from Ethiopia and Tanzania.

8. PCR-based screening revealed for the first time, the presence of the *Grasshopper* (*grh*) element in a high proportion of the FMB pathogen isolates in Eastern Africa specifically in Ethiopia and Tanzania (in 85 % of the isolates).

9. Recent advances in the next generation sequencing technologies have been effectively utilised to establish high quality reference genome assemblies of two isolates (E34 and K23/123) representing the two sub-populations identified in the FMB pathogen in Eastern Africa.

10. Genome resequence data has been developed and assembled for sixteen *M. oryzae* isolates representing the genetic and geographic diversity of the FMB pathogen in Eastern Africa.

11. Comparative analysis based on Synteny mapping and OrthoVenn analysis enabled the identification of genomic regions and genes putatively unique to specific *M. oryzae* isolates representing the FMB pathogen in Eastern Africa. These unique regions/genes are likely to influence host interaction patterns/pathogenicity. This hypothesis was tested initially by selecting two well-characterised avr gene models from rice blast. The results showed that *AVR-Piz-T* gene was present in all 18 isolates tested revealing two allelic forms, whilst *AVR-Pik* gene was only present in 6 of these isolates.

12. Phylogenomic analysis using a core set of Sordariomycete-BUSCO genes showed for the first time, the monophyletic nature of the finger millet blast pathogen from Eastern Africa and Asia (Clade A). Within Clade A, two sub-clades are evident broadly corresponding to the two sub-populations identified in the MLS analysis. Likewise, genome-wide SNPs of the set of 18 isolates also revealed two ranges, which relate the two sets of isolates to the above discussed sub-populations/Clades.

13. Based on in-depth analysis of the reference assemblies of isolates E34 and K23/123, for the first time, the full DNA sequence of *Grasshopper* (*grh*) element was identified and the structure defined from E34.

14. New insights of *grh* presence and its genomic architecture gained by identifying at least two versions of *grh* element that are present among the *M. oryzae* isolates

associated with finger millet in Eastern Africa. Isolates E34 and K23/123 typify the two versions of the *grh* with E34 showing a high level of sequence identity to the original sequence from G22, whilst K23/123 showed a high level of nucleotide differences.

15. Mating type specific PCR revealed clearly that both mating types MAT 1-1 (56 %) and MAT 1-2 (44 %) are present at a high proportion among the *M. oryzae* isolates representing the contemporary population in Eastern Africa as well as in the four countries surveyed, albeit at variable levels.

16. Mating culture assays revealed that a high proportion of *M. oryzae* isolates representing the contemporary population Identification are fertile (60 %). Among the fertile isolates, male sexual behaviour was dominant, followed by hermaphrodites and females. These emerging patterns are indicative of a gradual decrease in fertility and loss of hermaphroditic and female sexual behaviour.

17. Integration of the mating type distribution data with the fertility status information strongly suggets the possibility of a mixed reproduction system including asexual cycles and episodic sexual cycles of the FMB pathogen in Eastern Africa.

18. Overall, the new knowledge and resources developed provide an improved understanding of the biology and genetics of the finger millet blast pathogen *M. oryzae*. This would serve as a framework to take a structured approach for the identification, development and deployment of finger millet resistance to blast in Eastern Africa as well as future monitoring of pathogen population dynamics. Furthermore, the new knowledge and resources developed would serve as platform

for advancing the comparative genomic and functional analyses in this key pathosystem.

Based on the current results, suggested further work includes:

1. Investigations of the evidence for sexual recombination in field populations by allelic distribution and linkage disequilibrium analysis.
2. Comparative analysis and annotation of transposable elements and repeat elements in selected isolates including the genomic architecture of the *grh* element building on the genome sequence data.
3. Phylogenomic analysis of a wider set of *M. oryzae* isolates associated with finger millet production in Eastern Africa, India and other geographic locations to decipher the pathogen origin, evolution and transcontinental movement.
4. In-depth computational analysis and annotation of the genomic regions and genes unique to specific isolates from the Synteny mapping and OrthoVenn analyses.
5. Interrogation of the genomic architecture of key genes such as candidate effector genes including avr genes to identify genomic islands reflective of adaptive divergence and to decipher the role of transposable elements and repeat elements.
6. Explore the utility of the established rice blast pathogen mutants as a model to test the functional complementation of allelic forms of candidate effector/avirulence genes from the finger millet blast isolates to initiate the identification of avr gene mediated host interactions.

7.0 References

Abed-Ashtiani, F., Kadir, J., Nasehi, A., Hashemian-Rahaghi, S., Vadamalai, G., Rambe, S., 2016. Characterisation of *Magnaporthe oryzae* isolates from rice in peninsular Malaysia. Czech Journal of Genetics & Plant Breeding 52, 145-156.

Adarisini, V.B.P., Viji, G., Gnanamanickam, S., 1999. Mating type distribution, fertility and pathogenicity of rice isolates of *Magnaporthe grisea* in four rice growing regions of India. Indian Phytopathology 52, 28-34.

Adipala, E., Okoboi, C., Ssekabembe, C., Ogenga-Latigo, M., 1994. Foliar diseases and yield of finger millet [*Eleusine coracana* (L.)] under monocropping and intercropping systems. Discovery and Innovation 6, 301-305.

Afouda, L., Godjo, A., Sere, Y., Nouatin, G., Akossou, A., 2010. Transmission of *Pyricularia grisea* (Herbert) Barr by seeds to rice seedlings in Benin. Annals of Agricultural Science 14, 8-14.

Aguileta, G., Marthey, S., Chiapello, H., Lebrun, M.-., Rodolphe, F., Fournier, E., Gendrault-Jacquemard, A., Giraud, T., 2008. Assessing the Performance of Single-Copy Genes for Recovering Robust Phylogenies. Systematic Biology 57, 613-627.

Ajello, L., Cheng, S., 1967. The perfect state of *Trichophyton mentagrophytes*. Sabouraudia 5, 230-234.

Almiman, B.F., 2018. Molecular genetics and genomic characterization of an emerging mycotoxigenic pathogen *Fusarium proliferatum*. PhD Thesis, University of Bedfordshire, UK.

Amorim, R., Savi, D.C., Ferreira-Maba, L., Aluizio, R., Goulin, E.H., Takita, M.A., Machado, M.A., Glienke, C., 2017. MAT gene idiomorphs suggest a heterothallic sexual cycle in the citrus pathogen *Phyllosticta citricarpa*. European Journal of Plant Pathology 147, 325-337.

Amruta, N., Kumar, M.P., Puneeth, M., Sarika, G., Kandikattu, H.K., Vishwanath, K., Narayanaswamy, S., 2018. Exploring the potentiality of novel rhizospheric bacterial Strains against the rice blast fungus *Magnaporthe oryzae*. Plant Pathology Journal 34, 126-138.

Andersen, A.L., Henry, B., Tullis, E., 1947. Factors affecting infectivity, spread, and persistence of *Pyricularia oryzae* cav. Phytopathology 37, 94-110.

Anderson, P.K., Cunningham, A.A., Patel, N.G., Morales, F.J., Epstein, P.R., Daszak, P., 2004. Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. Trends in Ecology & Evolution 19, 535-544.

Armitage, A.D., Barbara, D.J., Harrison, R.J., Lane, C.R., Sreenivasaprasad, S., Woodhall, J.W., Clarkson, J.P., 2015. Discrete lineages within *Alternaria alternata* species group: Identification using new highly variable loci and support from morphological characters. Fungal Biology 119, 994-1006.

Aru, J., Wanyera, N., Okori, P., 2015. Inheritance of resistance to (NGR1) pathogen isolates of *Pyricularia grisea* in GULU-E finger millet last resistant variety of Uganda. Uganda Journal of Agricultural Sciences 16, 187-194.

Babu, T.K., Sharma, R., Upadhyaya, H., Reddy, P., Deshpande, S., Senthilvel, S., Sarma, N., Thakur, R., 2013. Evaluation of genetic diversity in *Magnaporthe grisea* populations adapted

to finger millet using simple sequence repeats (SSRs) markers. Physiological and Molecular Plant Pathology 84, 10-18.

Babu, T.K., Thakur, R., Upadhyaya, H., Reddy, P.N., Sharma, R., Girish, A., Sarma, N., 2013. Resistance to blast (*Magnaporthe grisea*) in a mini-core collection of finger millet germplasm. European Journal of Plant Pathology 135, 299-311.

Babu, T.K., Sharma, R., Thakur, R.P., Upadhyaya, H.D., Reddy, P.N., Girish, A.G., 2015. Selection of host differentials for elucidating pathogenic variation in *Magnaporthe grisea* populations adapted to finger millet (*Eleusine coracana*). Plant Disease. 99, 1784-1789.

Babujee, L., Gnanamanickam, S., 2000. Molecular tools for characterization of rice blast pathogen (*Magnaporthe grisea*) population and molecular marker-assisted breeding for disease resistance. Current Science-Bangalore 78, 248-257.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology 19, 455-477.

Bao, J., Chen, M., Zhong, Z., Tang, W., Lin, L., Zhang, X., Jiang, H., Zhang, D., Miao, C., Tang, H., Zhang, J., Lu, G., Ming, R., Norvienyeku, J., Wang, B., Wang, Z., 2017. PacBio sequencing reveals transposable elements as a key contributor to genomic plasticity and virulence variation in *Magnaporthe oryzae*. Molecular Plant 10, 1465-1468.

Barbeau, W., Hilu, K., 1993. Protein, calcium, iron, and amino acid content of selected wild and domesticated cultivars of finger millet. Plant Foods for Human Nutrition 43, 97-104.

Baroncelli, R., 2012. *Colletotrichum acutatum* sensu lato: from diversity study to genome analysis. PhD Thesis, University of Warwick, UK.

Baroncelli, R., Amby, D.B., Zapparata, A., Sarrocco, S., Vannacci, G., Le Floch, G., Harrison, R.J., Holub, E., Sukno, S.A., Sreenivasaprasad, S., 2016. Gene family expansions and contractions are associated with host range in plant pathogens of the genus *Colletotrichum*. BMC Genomics 17, 1-17.

Baroncelli, R., Talhinhas, P., Pensec, F., Sukno, S.A., Le Floch, G., Thon, M.R., 2017. The *Colletotrichum acutatum* species complex as a model system to study evolution and host specialization in plant pathogens. Frontiers in Microbiology 8, 2001-2017.

Baroncelli, R., Zapparata, A., Sarrocco, S., Sukno, S.A., Lane, C.R., Thon, M.R., Vannacci, G., Holub, E., Sreenivasaprasad, S., 2015. Molecular diversity of Anthracnose pathogen populations associated with UK strawberry production suggests multiple introductions of three different *Colletotrichum* species. PloS One 10, e0129140.

Bazin, E., Mathé-Hubert, H., Facon, B., Carlier, J., Ravigné, V., 2014. The effect of mating system on invasiveness: some genetic load may be advantageous when invading new environments. Biological Invasions 16, 875-886.

Beddington, J., 2010. Food security: contributions from science to a new and greener revolution. Philosophical Transactions of the Royal Society London. Series B Biological Society 365, 61-71.

Bernstein, H., Bernstein, C., 2010. Evolutionary origin of recombination during meiosis. Bioscience 60, 498-505.

Białas, A., Zess, E.K., De la Concepcion, Juan Carlos, Franceschetti, M., Pennington, H.G., Yoshida, K., Upson, J.L., Chanclud, E., Wu, C., Langner, T., 2017. Lessons in effector and NLR biology of plant-microbe systems. Molecular Plant-Microbe Interactions 31, 34-45.

Bihon, W., Wingfield, M.J., Slippers, B., Duong, T.A., Wingfield, B.D., 2014. MAT gene idiomorphs suggest a heterothallic sexual cycle in a predominantly asexual and important pine pathogen. Fungal Genetics and Biology 62, 55-61.

Billiard, S., López-Villavicencio, M., Hood, M., Giraud, T., 2012. Sex, outcrossing and mating types: unsolved questions in fungi and beyond. Journal of Evolutionary Biology 25, 1020-1038.

Bobrowicz, P., Pawlak, R., Correa, A., Bell-Pedersen, D., Ebbole, D.J., 2002. The *Neurospora crassa* pheromone precursor genes are regulated by the mating type locus and the circadian clock. Molecular Microbiology 45, 795-804.

Bohnert, H.U., Fudal, I., Dioh, W., Tharreau, D., Notteghem, J.L., Lebrun, M.H., 2004. A putative polyketide synthase/peptide synthetase from *Magnaporthe grisea* signals pathogen attack to resistant rice. Plant Cell 16, 2499-2513.

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

Bonman, J., 1992. Durable resistance to rice blast disease environmental influences. In: Anonymous breeding for disease resistance. Springer, pp. 115-123.

Boukaew, S., Prasertsan, P., 2014. Suppression of rice sheath blight disease using a heat stable culture filtrate from *Streptomyces philanthi* RM-1-138. Crop Protection 61, 1-10.

Bourett, T.M., Howard, R.J., 1990. In vitro development of penetration structures in the rice blast fungus *Magnaporthe grisea*. Canadian Journal of Botany 68, 329-342.

Bruns, T.D., Vilgalys, R., Barns, S.M., Gonzalez, D., Hibbett, D.S., Lane, D.J., Simon, L., Stickel, S., Szaro, T.M., Weisburg, W.G., 1992. Evolutionary relationships within the fungi: analyses of nuclear small subunit rRNA sequences. Molecular Phylogenetics and Evolution 1, 231-241.

Buiate, E.A., Xavier, K.V., Moore, N., Torres, M.F., Farman, M.L., Schardl, C.L., Vaillancourt, L.J., 2017. A comparative genomic analysis of putative pathogenicity genes in the host-specific sibling species *Colletotrichum graminicola* and *Colletotrichum sublineola*. BMC Genomics 18, 67-75.

Calvero, S.B., 1994. Developing models to predict favourable environments for rice blast. Masters Thesis, Oregon State University, USA.

Campbell, M.S., Holt, C., Moore, B., Yandell, M., 2014. Genome annotation and curation using MAKER and MAKER-P. Current Protocols in Bioinformatics 48, 4111-41139.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., Henrissat, B., 2008. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Research 37, D233-D238.

Cantarel, B.L., Korf, I., Robb, S.M., Parra, G., Ross, E., Moore, B., Holt, C., Sanchez Alvarado, A., Yandell, M., 2008. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Research 18, 188-196.

Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972-1973.

Carbone, I., Kohn, L.M., 1999. A method for designing primer sets for speciation studies in filamentous ascomycetes. Mycologia 91, 553-556.

Castroagudin, V., Danelli, A., Moreira, S.I., de Assis Reges, Juliana Teodora, Carvalho, G., Maciel, J.L.N., Bonato, A.L.V., Forcelini, C.A., Alves, E., McDonald, B., 2017. The wheat blast pathogen *Pyricularia graminis-tritici* has complex origins and a disease cycle spanning multiple grass hosts. Biorxiv 10, 203455-203520.

Castroagudin, V.L., Ceresini, P.C., de Oliveira, S.C., Reges, J.T.A., Maciel, J.L.N., Bonato, A.L.V., Dorigan, A.F., McDonald, B.A., 2015. Resistance to QoI Fungicides is widespread in Brazilian populations of the wheat blast pathogen *Magnaporthe oryzae*. Phytopathology 105, 284-294.

Ceresini, P.C., Castroagudín, V.L., Rodrigues, F.Á., Rios, J.A., Aucique-Pérez, C.E., Moreira, S.I., Croll, D., Alves, E., de Carvalho, G., Maciel, J.L.N., 2018. Wheat blast: from its origins in

South America to its emergence as a global threat. Molecular Plant Pathology 10, 12747-12765.

Chadha, S., Sharma, M., 2014. Transposable elements as stress adaptive capacitors induce genomic instability in fungal pathogen *Magnaporthe oryzae*. PloS One 9, e94415.

Chakraborty, M., Baldwin-Brown, J.G., Long, A.D., Emerson, J., 2016. Contiguous and accurate de novo assembly of metazoan genomes with modest long read coverage. Nucleic Acids Research 44, E147-E159.

Chakraborty, S., Newton, A.C., 2011. Climate change, plant diseases and food security: an overview. Plant Pathology 60, 2-14.

Chao, C.T., Ellingboe, A.H., 1991. Selection for mating competence in *Magnaporthe grisea* pathogenic to rice. Canadian Journal of Botany 69, 2130-2134.

Chen, C., Lian, B., Hu, J., Zhai, H., Wang, X., Venu, R., Liu, E., Wang, Z., Chen, M., Wang, B., 2013. Genome comparison of two *Magnaporthe oryzae* field isolates reveals genome variations and potential virulence effectors. BMC Genomics 14, 887-895.

Chen, D., Zeigler, R.S., Leung, H., Nelson, R.J., 1995. Population structure of *Pyricularia grisea* at two screening sites in the Philippines. Phytopathology 85, 1011-1020.

Chen, Q., Wang, Y., Zheng, X., 2006. Genetic diversity of *Magnaporthe grisea* in China as revealed by DNA fingerprint haplotypes and pathotypes. Journal of Phytopathology 154, 361-369.

Chen, C., Lian, B., Hu, J., Zhai, H., Wang, X., Venu, R.C., Liu, E., Wang, Z., Chen, M., Wang, B., Wang, G.L., Wang, Z., Mitchell, T.K., 2013a. Genome comparison of two *Magnaporthe oryzae* field isolates reveals genome variations and potential virulence effectors. BMC Genomics 14, 8872164-8872185.

Chen, Y., Yang, X., Yuan, S.-., Li, Y.-., Zhang, A.-., Yao, J., Gao, T.-., 2015. Effect of azoxystrobin and kresoxim-methyl on rice blast and rice grain yield in China. Annals of Applied Biology 166, 434-443.

Chiapello, H., Mallet, L., Guerin, C., Aguileta, G., Amselem, J., Kroj, T., Ortega-Abboud, E., Lebrun, M., Henrissat, B., Gendrault, A., 2015. Deciphering genome content and evolutionary relationships of isolates from the fungus *Magnaporthe oryzae* attacking different host plants. Genome Biology and Evolution 7, 2896-2912.

Chisholm, S.T., Coaker, G., Day, B., Staskawicz, B.J., 2006. Host-microbe interactions: shaping the evolution of the plant immune response. Cell 124, 803-814.

Chuma, I., Isobe, C., Hotta, Y., Ibaragi, K., Futamata, N., Kusaba, M., Yoshida, K., Terauchi, R., Fujita, Y., Nakayashiki, H., 2011. Multiple translocation of the *AVR-Pita* effector gene among chromosomes of the rice blast fungus *Magnaporthe oryzae* and related species. PLoS Pathogens 7, e1002147.

Cochrane, L., 2014. Agricultural intensification in Ethiopia: Review of recent research. African Journal of Agricultural Research 9, 2377-2390.

Consolo, V.F., Cordo, C.A., Salerno, G.L., 2008. DNA fingerprint and pathotype diversity of *Pyricularia oryzae* populations from Argentina. Australasian Plant Pathology 37, 357-364.

Consolo, V.F., Cordo, C.A., Salerno, G.L., 2005. Mating-type distribution and fertility status in *Magnaporthe grisea* populations from Argentina. Mycopathologia 160, 285-290.

Coppin, E., Debuchy, R., Arnaise, S., Picard, M., 1997. Mating types and sexual development in filamentous ascomycetes. Microbiology Molecular Biology Reviews 61, 411-428.

Cook, D.E., Mesarich, C.H., Thomma, B.P.H.J., 2015. Understanding plant immunity as a surveillance system to detect invasion. Annual Review of Phytopathology 53, 541-563.

Correa-Victoria, F.J., Zeigler, R.S., 1993. Pathogenic variability in *Pyricularia grisea* at a rice blast hot spot breeding site in eastern Colombia. Plant Disease 77, 1029-1035.

Correll, J., Harp, T., Guerber, J., Zeigler, R., Liu, B., Cartwright, R., Lee, F., 2000. Characterization of *Pyricularia grisea* in the United States using independent genetic and molecular markers. Phytopathology 90, 1396-1404.

Couch, B.C., Fudal, I., Lebrun, M.H., Tharreau, D., Valent, B., van Kim, P., Notteghem, J.L., Kohn, L.M., 2005. Origins of host-specific populations of the blast pathogen *Magnaporthe oryzae* in crop domestication with subsequent expansion of pandemic clones on rice and weeds of rice. Genetics 170, 613-630.

Couch, B.C., Kohn, L.M., 2002. A multilocus gene genealogy concordant with host preference indicates segregation of a new species, *Magnaporthe oryzae* from *Magnaporthe grisea*. Mycologia 94, 683-693.

Crawford, M.S., Chumley, F.G., Weaver, C.G., Valent, B., 1986. Characterization of the heterokaryotic and vegetative diploid phases of *Magnaporthe grisea*. Genetics 114, 1111-1129.

Croll, D., Laine, A., 2016. What the population genetic structures of host and pathogen tell us about disease evolution. New Phytologist 212, 537-539.

Croll, D., McDonald, B.A., 2012. The accessory genome as a cradle for adaptive evolution in pathogens. PLoS Pathogens 8, e1002608.

Croll, D., Lendenmann, M.H., Stewart, E., McDonald, B.A., 2015. The impact of recombination hotspots on genome evolution of a fungal plant pathogen. Genetics 201, 1213-1228.

Cruz, C.D., Valent, B., 2017. Wheat blast disease: danger on the move. Tropical Plant Pathology 42, 210-222.

D'Ávila, L.S., Lehner, M.S., Filippi, M.C.C., Scheuermann, K.K., Del Ponte, E.M., 2016. Genetic structure and mating type analysis of the *Pyricularia oryzae* population causing widespread epidemics in southern Brazil. Tropical Plant Pathology 41, 297-305.

Dai, Y., Jia, Y., Correll, J., Wang, X., Wang, Y., 2010. Diversification and evolution of the avirulence gene *AVR-Pita1* in field isolates of *Magnaporthe oryzae*. Fungal Genetics and Biology 47, 973-980.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., 2011. The variant call format and VCFtools. Bioinformatics 27, 2156-2158.

Das, I., 2017. Millet Diseases: Current Status and Their Management. Millets and Sorghum: Biology and Genetic Improvement 10, 291-322.

Dass, A., Sudhishri, S., Lenka, N., 2013. Integrated nutrient management to improve finger millet productivity and soil conditions in hilly region of Eastern India. Journal of Crop Improvement 27, 528-546.

De Datta, S.K., 1981. Principles and Practices of Rice Production. International Rice Research Institute, New York: John Wileys & Sons, pp. 618.

De Villiers, S.M., Michael, V.N., Manyasa, E.O., Saiyiorri, A.N., Deshpande, S., 2015. Compilation of an informative microsatellite set for genetic characterisation of East African finger millet (*Eleusine coracana*). Electronic Journal of Biotechnology 18, 77-82.

De Wet, J., Rao, K.P., Brink, D., Mengesha, M., 1984. Systematics and evolution of *Eleusine coracana* (Gramineae). American Journal of Botany 71, 550-557.

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

Dean, R.A., Talbot, N.J., Ebbole, D.J., Farman, M.L., Mitchell, T.K., Orbach, M.J., Thon, M., Kulkarni, R., Xu, J., Pan, H., Read, N.D., Lee, Y., Carbone, I., Brown, D., Oh, Y.Y., Donofrio, N.,

Jeong, J.S., Soanes, D.M., Djonovic, S., Kolomiets, E., Rehmeyer, C., Li, W., Harding, M., Kim, S., Lebrun, M., Bohnert, H., Coughlan, S., Butler, J., Calvo, S., Ma, L., Nicol, R., Purcell, S., Nusbaum, C., Galagan, J.E., Birren, B.W., 2005. The genome sequence of the rice blast fungus *Magnaporthe grisea*. Nature 434, 980-986.

Debnath, A., Bhutia, K.L., Sumpi, H., 2018. Mapping and Deployment of Blast Resistance Gene in RiceA Work in Progress. International Journal of Current Microbiology and Applied Science 7, 2073-2094.

Debuchy, R., Coppin, E., 1992. The mating types of *Podospora anserina*: functional analysis and sequence of the fertilization domains. Molecular and General Genetics 233, 113-121.

Deininger, P.L., Batzer, M.A., Hutchison III, C.A., Edgell, M.H., 1992. Master genes in mammalian repetitive DNA amplification. Trends in Genetics 8, 307-311.

Depotter, J.R., Seidl, M.F., van den Berg, Grardy CM, Thomma, B.P., Wood, T.A., 2017. A distinct and genetically diverse lineage of the hybrid fungal pathogen *Verticillium longisporum* population causes stem striping in British oilseed rape. Environmental Microbiology 19, 3997-4009.

Devanna, N.B., Vijayan, J., Sharma, T.R., 2014. The blast resistance gene *Pi54* of cloned from *Oryza officinalis* interacts with *AVR-Pi54* through its novel non-LRR domains. PLoS One 9, e104840.

Devulapalle, K.S., Suryanarayanan, S., 1995. Cross-incompatibility among Indian isolates of *Pyricularia grisea*. Plant Disease 79, 779-781.

DeZwaan, T.M., Carroll, A.M., Valent, B., Sweigard, J.A., 1999. *Magnaporthe grisea* pth11p is a novel plasma membrane protein that mediates appressorium differentiation in response to inductive substrate cues. Plant Cell 11, 2013-2030.

Dida, M.M., Devos, K.M., 2006. Finger millet. In: Anonymous Cereals and Millets. Springer, pp. 333-343.

Dida, M.M., Wanyera, N., Dun, M.L.H., Bennetzen, J.L., Devos, K.M., (2008). Population structure and diversity in finger millet (*Eleusine coracana*) germplam. Tropical Plant Biology 1, 131-141.

Dioh, W., Tharreau, D., Notteghem, J.L., Orbach, M., Lebrun, M., 2000. Mapping of avirulence genes in the rice blast fungus, *Magnaporthe grisea*, with RFLP and RAPD markers. Molecular Plant-Microbe Interactions 13, 217-227.

Disthaporn, S., 1994. Current rice blast epidemics and their management in Thailand. Rice Blast Disease, 321-332.

Dobinson, K.F., Harris, R.E., Hamer, J.E., 1993. *Grasshopper*, a long terminal repeat (LTR) retroelement in the phytopathogenic fungus *Magnaporthe grisea*. Molecular Plant-Microbe Interactions 6, 114-120.

Dobrindt, U., Hochhut, B., Hentschel, U., Hacker, J., 2004. Genomic islands in pathogenic and environmental microorganisms. Nature Reviews Microbiology 2, 414-420.

Dodds, P.N., Rathjen, J.P., 2010. Plant immunity: towards an integrated view of plant–pathogen interactions. Nature Reviews Genetics 11, 539-545.

Don, L.D., Chapter, K.B., Urashima, A.S., Yukio, T., Akira, N., Shimayama, S., 1999. DNA fingerprint analysis of population structure of rice blast fungus in Japan. Japan Plant Pathology Newletter 65, 15-24.

Donaldson, G.C., Ball, L.A., Axelrood, P.E., Glass, N.L., 1995. Primer sets developed to amplify conserved genes from filamentous ascomycetes are useful in differentiating *Fusarium species* associated with conifers. Applied and Environmental Microbiology 61, 1331-1340.

Dong, Y., Li, Y., Zhao, M., Jing, M., Liu, X., Liu, M., Guo, X., Zhang, X., Chen, Y., Liu, Y., 2015. Global genome and transcriptome analyses of *Magnaporthe oryzae* epidemic isolate 98-06 uncover novel effectors and pathogenicity-related genes, revealing gene gain and lose dynamics in genome evolution. PLoS Pathogens 11, e1004801.

Dorjey, S., Dolkar, D., Sharma, R., 2017. Plant growth promoting Rhizobacteria *Pseudomonas*: A Review. International Journal of Current Microbiology and Applied Sciences 6, 1335-1344.

Duncan, J., Cooke, D., Birch, P., Toth, R., Bridge, P., Couteaudier, Y., Clarkson, J., 1998. Molecular variability in sexually reproducing fungal plant pathogens. Molecular Variability of Fungal Pathogens 17, 19-39.

Esele, J.P., Odelle, S., 1995. Progress in breeding for resistance to finger millet blast disease at Serere research station. In: Eighth EARSAM Regional Workshop on Sorghum and Millets, pp.100-110.

Esele, J., Odelle, S., 1992. Progress in breeding for resistance to finger millet blast at Serere research station. In: Proceedings of the 6th EARSAM Regional Workshop on Sorghum and Millets, 30th Oct.–5th Nov, pp. 93-99.

Eto, Y., Ikeda, K., Chuma, I., Kataoka, T., Kuroda, S., Kikuchi, N., Don, L., Kusaba, M., Nakayashiki, H., Tosa, Y., 2001. Comparative analyses of the distribution of various transposable elements in *Pyricularia* and their activity during and after the sexual cycle. Molecular and General Genetics 264, 565-577.

Faino, L., Seidl, M.F., Shi-Kunne, X., Pauper, M., van den Berg, G.C., Wittenberg, A.H., Thomma, B.P., 2016. Transposons passively and actively contribute to evolution of the two-speed genome of a fungal pathogen. Genome Research 26, 1091-1100.

Farman, M.L., 2007. Telomeres in the rice blast fungus *Magnaporthe oryzae*: the world of the end as we know it. FEMS Microbiology Letters 273, 125-132.

Farman, M.L., 2002. *Pyricularia grisea* isolates causing gray leaf spot on perennial ryegrass (*Lolium perenne*) in the United States: Relationship to *P. grisea* isolates from other host plants. Phytopathology 92, 245-254.

Farman, M.L., Leong, S.A., 1998. Chromosome walking to the *AVR1-CO39* avirulence gene of *Magnaporthe grisea*: discrepancy between the physical and genetic maps. Genetics 150, 1049-1058.

Farman, M.L., Taura, S., Leong, S.A., 1996b. The *Magnaporthe grisea* DNA fingerprinting probe MGR586 contains the 3' end of an inverted repeat transposon. Molecular and General Genetics 251, 675-681.

Farman, M.L., Tosa, Y., Nitta, N., Leong, S.A., 1996a. MAGGY, a retrotransposon in the genome of the rice blast fungus *Magnaporthe grisea*. Molecular and General Genetics 251, 665-674.

Farman, M., Peterson, G., Chen, L., Starnes, J., Valent, B., Bachi, P., Murdock, L., Hershman, D., Pedley, K., Fernandes, J.M., 2017. The Lolium pathotype of *Magnaporthe oryzae* recovered from a single blasted wheat plant in the United States. Plant Disease 101, 684-692.

Feau, N., Decourcelle, T., Husson, C., Desprez-Loustau, M., Dutech, C., 2011. Finding single copy genes out of sequenced genomes for multilocus phylogenetics in non-model fungi. PLoS One 6, e18803.

Fernandez, J., Orth, K., 2018. Rise of a Cereal Killer: The biology of *Magnaporthe oryzae* biotrophic growth. Trends in Microbiology 26, 582-597.

Finckh, M.R., 2008. Integration of breeding and technology into diversification strategies for disease control in modern agriculture. In: Anonymous Sustainable Disease Management in a European Context. Springer, pp. 399-409.

Fisher, M.C., Henk, D.A., Briggs, C.J., Brownstein, J.S., Madoff, L.C., McCraw, S.L., Gurr, S.J., 2012. Emerging fungal threats to animal, plant and ecosystem health. Nature 484, 186-195.

Fisher, M.C., Gow, N.A., Gurr, S.J., 2016. Tackling emerging fungal threats to animal health, food security and ecosystem resilience. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 371, 206-210.

FRAC 2018. List plant pathogenic organisms resistant to disease control agents. FRAC Code numbering, 11-48.

Fraser, J.A., Stajich, J.E., Tarcha, E.J., Cole, G.T., Inglis, D.O., Sil, A., Heitman, J., 2007. Evolution of the mating type locus: insights gained from the dimorphic primary fungal pathogens *Histoplasma capsulatum*, *Coccidioides immitis*, and *Coccidioides posadasii*. Eukaryotic Cell 6, 622-629.

Galagan, J.E., Calvo, S.E., Borkovich, K.A., Selker, E.U., Read, N.D., Jaffe, D., FitzHugh, W., Ma, L., Smirnov, S., Purcell, S., 2003. The genome sequence of the filamentous fungus *Neurospora crassa.* Nature 422, 859-864.

Geiser, D.M., Gueidan, C., Miadlikowska, J., Lutzoni, F., Kauff, F., Hofstetter, V., Fraker, E., Schoch, C.L., Tibell, L., Untereiner, W.A., 2006. *Eurotiomycetes*: *Eurotiomycetidae* and *Chaetothyriomycetidae*. Mycologia 98, 1053-1064.

George, M., Nelson, R., Zeigler, R., Leung, H., 1998. Rapid population analysis of *Magnaporthe grisea* by using rep-PCR and endogenous repetitive DNA sequences. Phytopathology 88, 223-229.

Ghazanfar, M.U., Habib, A., Sahi, S., 2009. Screening of rice germplasm against *Pyricularia oryzae* the cause of rice blast disease. Pakistan Journal of Phytopathology 21, 41-51.

Gimode, D., Odeny, D.A., de Villiers, E.P., Wanyonyi, S., Dida, M.M., Mneney, E.E., Muchugi, A., Machuka, J., de Villiers, S.M., 2016. Identification of SNP and SSR markers in finger millet using next generation sequencing technologies. PLoS One 11, e0159437.

Gladieux, P., Condon, B., Ravel, S., Soanes, D., Maciel, J.L.N., Nhani, A., Jr, Chen, L., Terauchi, R., Lebrun, M.H., Tharreau, D., Mitchell, T., Pedley, K.F., Valent, B., Talbot, N.J., Farman, M., Fournier, E., 2018. Gene flow between divergent cereal and grass specific lineages of the rice blast fungus *Magnaporthe oryzae*. mBio 9, e01219-17.

Glass, N.L., Donaldson, G.C., 1995. Development of primer sets designed for use with the PCR to amplify conserved genes from filamentous ascomycetes. Applied and Environmental Microbiology 61, 1323-1330.

Glass, N.L., Vollmer, S.J., Staben, C., Grotelueschen, J., Metzenberg, R.L., Yanofsky, C., 1988. DNAs of the two mating-type alleles of *Neurospora crassa* are highly dissimilar. Science 241, 570-573.

Godfray, H.C., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty, J., Robinson, S., Thomas, S.M., Toulmin, C., 2010. Food security: the challenge of feeding 9 billion people. Science 327, 812-818.

Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M., Louis, E.J., Mewes, H.W., Murakami, Y., Philippsen, P., Tettelin, H., Oliver, S.G., 1996. Life with 6000 genes. Science 274, 563-567.

Goodwin, S.B., M'barek, S.B., Dhillon, B., Wittenberg, A.H., Crane, C.F., Hane, J.K., Foster, A.J., Van der Lee, Theo AJ, Grimwood, J., Aerts, A., 2011. Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. PLoS Genetics 7, e1002070.

Gopalakrishnan, S., Vadlamudi, S., Bandikinda, P., Sathya, A., Vijayabharathi, R., Rupela, O., Kudapa, H., Katta, K., Varshney, R.K., 2014. Evaluation of *Streptomyces* strains isolated from herbal vermicompost for their plant growth-promotion traits in rice. Microbiological Research 169, 40-48.

Gowda, M., Shirke, M.D., Mahesh, H.B., Chandarana, P., Rajamani, A., Chattoo, B.B., 2015. Genome analysis of rice-blast fungus *Magnaporthe oryzae* field isolates from southern India. Genomics Data 5, 284-291.

Grandaubert, J., Lowe, R.G., Soyer, J.L., Schoch, C.L., Van de Wouw, Angela P, Fudal, I., Robbertse, B., Lapalu, N., Links, M.G., Ollivier, B., 2014. Transposable element-assisted evolution and adaptation to host plant within the *Leptosphaeria maculans-Leptosphaeria biglobosa* species complex of fungal pathogens. BMC Genomics 15, 891-899.

Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29, 1072-1075.

Hajime, K., Yamamoto, M., Yamaguchi-Ozaki, T., Kadouchi, H., Iwamoto, Y., Nakayashiki, H., Yukio, T., Mayama, S., Naoki, M., 2000. Pathogenicity, mating ability and DNA restriction fragment length polymorphisms of *Pyricularia* populations isolated from Gramineae, Bambusideae and Zingiberaceae plants. Journal of General Plant Pathology 66, 30-47.

Hall, T., Biosciences, I., Carlsbad, C., 2011. BioEdit: an important software for molecular biology. GERF Bulletin of Biosciences 2, 60-61.

Hamada, T., Asanagi, M., Satozawa, T., Araki, N., Banba, S., Higashimura, N., Akase, T., Hirase, K., 2014. Action mechanism of the novel rice blast fungicide tolprocarb distinct from that of conventional melanin biosynthesis inhibitors. Journal of Pesticide Science 39, 152-158.

Hamer, J.E., Farrall, L., Orbach, M.J., Valent, B., Chumley, F.G., 1989. Host species-specific conservation of a family of repeated DNA sequences in the genome of a fungal plant pathogen. Proceedings of the National Academy of Sciences of the United States of America 86, 9981-9985.

Han, J., Kim, S.H., Lee, Y., Lee, S., Hwang, J., Kim, J.Y., Yoon, S.J., Lee, G.K., 2014. Comparison of targeted next-generation sequencing with conventional sequencing for predicting the responsiveness to epidermal growth factor receptor-tyrosine kinase inhibitor (EGFR-TKI) therapy in never-smokers with lung adenocarcinoma. Lung Cancer 85, 161-167.

Hatakeyama, M., Aluri, S., Balachadran, M.T., Sivarajan, S.R., Patrignani, A., Grüter, S., Poveda, L., Shimizu-Inatsugi, R., Baeten, J., Francoijs, K., 2017. Multiple hybrid de novo genome assembly of finger millet, an orphan allotetraploid crop. DNA Research 25, 39-47.

Hayashi, N., Li, C.Y., Li, J.L., Naito, H., 1997. In vitro production on rice plants of perithecia of *Magnaporthe grisea* from Yunnan, China. Mycological Research 101, 1308-1310.

Head, S.R., Komori, H.K., LaMere, S.A., Whisenant, T., Van Nieuwerburgh, F., Salomon, D.R., Ordoukhanian, P., 2014. Library construction for next-generation sequencing: overviews and challenges. BioTechniques 56, 61-77.

Hebert, T., 1971. The perfect stage of *Pyricularia grisea*. Phytopathology 61, 83-87.

Hibbett, D., Abarenkov, K., Kõljalg, U., Öpik, M., Chai, B., Cole, J., Wang, Q., Crous, P., Robert, V., Helgason, T., 2016. Sequence-based classification and identification of Fungi. Mycologia 108, 1049-1068.

Hilu, K.W., Wet, J.M.J., Harlan, J.R., 1979. Archaeobotanical Studies of *Eleusine coracana* ssp. *coracana* (Finger Millet). American Journal of Botany 66, 330-331.

Hittalmani, S., Mahesh, H., Shirke, M.D., Biradar, H., Uday, G., Aruna, Y., Lohithaswa, H., Mohanrao, A., 2017. Genome and Transcriptome sequence of Finger millet (*Eleusine coracana* (L.) Gaertn.) provides insights into drought tolerance and nutraceutical properties. BMC Genomics 18, 465-474.

Hoff, K.J., Lange, S., Lomsadze, A., Borodovsky, M., Stanke, M., 2016. BRAKER1: unsupervised RNA-Seq-based genome annotation with GeneMark-ET and AUGUSTUS. Bioinformatics 32, 767-769.

Hoff, K.J., Stanke, M., 2013. Web AUGUSTUS—a web service for training AUGUSTUS and predicting genes in eukaryotes. Nucleic Acids Research 41, W123-W128.

Hofstetter, V., Miadlikowska, J., Kauff, F., Lutzoni, F., 2007. Phylogenetic comparison of protein-coding versus ribosomal RNA-coding sequence data: a case study of the Lecanoromycetes (Ascomycota). Molecular Phylogenetics and Evolution 44, 412-426.

Howlett, B.J., Lowe, R.G., Marcroft, S.J., van de Wouw, Angela P, 2015. Evolution of virulence in fungal plant pathogens: exploiting fungal genomics to control plant disease. Mycologia 107, 441-451.

Huang, J., Si, W., Deng, Q., Li, P., Yang, S., 2014. Rapid evolution of avirulence genes in rice blast fungus *Magnaporthe oryzae*. BMC Genetics 15, 45-50.

Hunkapiller, T., Kaiser, R.J., Koop, B.F., Hood, L., 1991. Large-scale and automated DNA sequence determination. Science 254, 59-67.

Hyon, G., Nga, N.T.T., Chuma, I., Inoue, Y., Asano, H., Murata, N., Kusaba, M., Tosa, Y., 2012. Characterization of interactions between barley and various host-specific subgroups of *Magnaporthe oryzae* and *Magnaporthe grisea*. Journal of General Plant Pathology 78, 237-246.

Ignacimuthu, S., Ceasar, S.A., 2012. Development of transgenic finger millet (*Eleusine coracan*a (L.) Gaertn.) resistant to leaf blast disease. Journal of Biosciences 37, 135-147.

Imam, J., Alam, S., Mandal, N.P., Maiti, D., Variar, M., Shukla, P., 2015. Molecular diversity and mating type distribution of the rice blast pathogen *Magnaporthe oryzae* in North-East and Eastern India. Indian Journal of Microbiology 55, 108-113.

Inoue, K., Kitaoka, H., Park, P., Ikeda, K., 2016. Novel aspects of hydrophobins in wheat isolate of *Magnaporthe oryzae*: Mpg1, but not Mhp1, is essential for adhesion and pathogenicity. Journal of General Plant Pathology 82, 18-28.

Islam, M.T., Croll, D., Gladieux, P., Soanes, D.M., Persoons, A., Bhattacharjee, P., Hossain, M.S., Gupta, D.R., Rahman, M.M., Mahboob, M.G., 2016. Emergence of wheat blast in Bangladesh was caused by a South American lineage of *Magnaporthe oryzae*. BMC Biology 14, 84-90.

Itoi, S., Mishima, T., Arase, S., Nozu, M., 1983. Mating behaviour of Japanese isolates of *Pyricularia oryzae*. Phytopathology 73, 155-158.

Jagadeesh, D., Kumar, M.P., Chandrakanth, R., Devaki, N., 2018a. Molecular diversity of internal transcribed spacer among the monoconidial isolates of *Magnaporthe oryzae* isolated from rice in Southern Karnataka, India. Journal of Genetic Engineering and Biotechnology 8, 118-123.

Jagadeesh, D., Kumar, M.P., Devaki, N., 2018b. Population analysis of *Magnaporthe oryzae* by using endogenous repetitive DNA sequences and mating-type alleles in different districts of Karnataka, India. Journal of Applied Genetics 59, 365-375.

Jardim, S.S., Schuch, A.P., Pereira, C.M., Loreto, E.L.S., 2015. Effects of heat and UV radiation on the mobilization of transposon mariner-Mos1. Cell Stress and Chaperones 6, 1-9.

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

Kachroo, P., Leong, S., Chattoo, B., 1994. Pot2, an inverted repeat transposon from the rice blast fungus *Magnaporthe grisea*. Molecular and General Genetics 245, 339-348.

Kachroo, P., Leong, S.A., Chattoo, B.B., 1995. Mg-SINE: a short interspersed nuclear element from the rice blast fungus, *Magnaporthe grisea*. Proceedings of the National Academy of Sciences of the United States of America 92, 11125-11129.

Kanamori, M., Kato, H., Yasuda, N., Koizumi, S., Peever, T.L., Kamakura, T., Teraoka, T., Arie, T., 2007. Novel mating type-dependent transcripts at the mating type locus in *Magnaporthe oryzae*. Gene 403, 6-17.

Kang, S., Chumley, F.G., Valent, B., 1994. Isolation of the mating-type genes of the phytopathogenic fungus *Magnaporthe grisea* using genomic subtraction. Genetics 138, 289-296.

Kang, S., Lebrun, M.H., Farrall, L., Valent, B., 2001. Gain of virulence caused by insertion of a Pot3 transposon in a *Magnaporthe grisea* avirulence gene. Molecular Plant-Microbe Interactions 14, 671-674.

Kang, S., Sweigard, J.A., Valent, B., 1995. The PWL host specificity gene family in the blast fungus *Magnaporthe grisea*. Molecular Plant-Microbe Interactions 8, 939-948.

Karthikeyan, V., Gnanamanickam, S., 2008a. Biological control of Setaria blast (*Magnaporthe grisea*) with bacterial strains. Crop Protection 27, 263-267.

Karthikeyan, V., Gnanamanickam, S., 2008b. Determining the fertility status of Setaria infecting *Magnaporthe grisea* isolates with standard testers and identification of tolerant cultivar of *Setaria italica*. Mycopathologia 166, 227-323.

Kato, H., Yamaguchi, T., 1982. The perfect state of *Pyricularia oryzae* Cav. from rice plants in culture. Japanese Journal of Phytopathology 48, 607-612.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution 30, 772-780.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647-1649.

Keller, H., Boyer, L., Abad, P., 2016. Disease susceptibility in the zig-zag model of host-microbe interactions: only a consequence of immune suppression?. Molecular Plant Pathology 17, 475-479.

Kema, G.H., Gohari, A.M., Aouini, L., Gibriel, H.A., Ware, S.B., van Den Bosch, F., Manning-Smith, R., Alonso-Chavez, V., Helps, J., M'Barek, S.B., 2018. Stress and sexual reproduction affect the dynamics of the wheat pathogen effector AVRStb6 and strobilurin resistance. Nature Genetics 50, 375-380.

Khalil, M., Moubasher, H., Hasan, F., 2014. Biological control of rice blast disease by *Streptomyces flavotricini*. Research Journal of Pharmaceutical, Biological and Chemical Sciences 5, 1453-1461.

Khang, C.H., Park, S., Lee, Y., Valent, B., Kang, S., 2008. Genome organization and evolution of the *AVR-Pita* avirulence gene family in the *Magnaporthe grisea* species complex. Molecular Plant-Microbe Interactions 21, 658-670.

Khanum, S.A., Shashikanth, S., Sathyanarayana, S.G., Lokesh, S., Deepak, S.A., 2009. Synthesis and antifungal activity of 2-azetidinonyl-5-(2-benzoylphenoxy) methyl-1, 3, 4-oxadiazoles against seed-borne pathogens of *Eleusine coracana* (L.) Gaertn. Pest Management Science 65, 776-780.

Kihoro, J., Bosco, N.J., Murage, H., Ateka, E., Makihara, D., 2013. Investigating the impact of rice blast disease on the livelihood of the local farmers in greater Mwea region of Kenya. SpringerPlus 2, 308-310.

Kinfe, H., Yiergalem, T., Alem, R., Redae, W., Desalegn, Y., Welegerima, G., 2017. Yield performance and adaptability of finger millet landrace in North Western Tigray, Ethiopia. World News of Natural Sciences 15, 98-111.

Kisandu, D., Ntundu, W., Marandu, W., Mgonja, M., 2005. Germplasm Collection and Evaluation of Finger Millet in Tanzania: Challenges and Opportunities for Improved Production. Finger Millet Blast Management in East Africa. Creating opportunities for improving production and utilization of finger millet, 23-32.

Kito, H., Takahashi, Y., Sato, J., Fukiya, S., Sone, T., Tomita, F., 2003. Occan, a novel transposon in the Fot1 family, is ubiquitously found in several *Magnaporthe grisea* isolates. Current Genetics 42, 322-331.

Klaubauf, S., Tharreau, D., Fournier, E., Groenewald, J., Crous, P.W., De Vries, R.P., Lebrun, M., 2014. Resolving the polyphyletic nature of *Pyricularia* (Pyriculariaceae). Studies in Mycology 79, 85-120.

Koeck, M., Hardham, A.R., Dodds, P.N., 2011. The role of effectors of biotrophic and hemibiotrophic fungi in infection. Cellular Microbiology 13, 1849-1857.

Kohli, M., Mehta, Y., Guzman, E., Viedma, L., Cubilla, L., 2011. *Pyricularia* blast-a threat to wheat cultivation. Czech Journal of Genetics and Plant Breeding 47, S130-S134.

Koide, Y., Kobayashi, N., Xu, D., Fukuta, Y., 2009. Resistance genes and selection DNA markers for blast disease in rice (*Oryza sativa* L.). Japan Agricultural Research Quarterly: JARQ 43, 255-280.

Kõljalg, U., Nilsson, R.H., Abarenkov, K., Tedersoo, L., Taylor, A.F., Bahram, M., Bates, S.T., Bruns, T.D., Bengtsson-Palme, J., Callaghan, T.M., 2013. Towards a unified paradigm for sequence-based identification of fungi. Molecular Ecology 22, 5271-5277.

Kolmer, J.A., Ellingboe, A.H., 1988. Genetic relationships between fertility and pathogenicity and virulence to rice in *Magnaporthe grisea*. Canadian Journal of Botany 66, 891-897.

Kusaba, M., Hirata, K., Sumida, Y., Yamagashira, A., Konagai-Urata, H., Yaegashi, H., 2006. Molecule genetic characterization and host specificity of *Pyricularia* isolates from annual ryegrass in Japan. Plant Pathology Journal 5, 72-79.

Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H., Phillippy, A.M., 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Research 27, 722-736.

Korinsak, S., Tangphatsornruang, S., Pootakham, W., Wanchana, S., Plabpla, A., Jantasuriyarat, C., Patarapuwadol, S., Vanavichit, A., Toojinda, T., 2018. Genome-wide association mapping of virulence gene in rice blast fungus *Magnaporthe oryzae* using a genotyping by sequencing approach. Genomics 011, 245-260.

Kumar, G., Maharshi, A., Patel, J., Mukherjee, A., Singh, H., Sarma, B., 2016. Trichoderma: a potential fungal antagonist to control plant diseases. SATSA Mukhapatra-Annual Technical Issue 21, 206-218.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular Biology and Evolution 33, 1870-1874.

Kumar, J., Nelson, R.J., Zeigler, R.S., 1999. Population structure and dynamics of *Magnaporthe grisea* in the Indian Himalayas. Genetics 152, 971-984.

Kunova, A., Pizzatti, C., Bonaldi, M., Cortesi, P., 2014. Sensitivity of nonexposed and exposed populations of *Magnaporthe oryzae* from rice to Tricyclazole and Azoxystrobin. Plant Disease 98, 512-518.

Kunova, A., Pizzatti, C., Cortesi, P., 2013. Impact of tricyclazole and azoxystrobin on growth, sporulation and secondary infection of the rice blast fungus, *Magnaporthe oryzae*. Pest Management Science 69, 278-284.

Kurahashi, Y., 2001. Melanin biosynthesis inhibitors (MBIs) for control of rice blast. Pesticide Outlook 12, 32-35.

Kutcher, H., Balesdent, M., Rimmer, S., Rouxel, T., Chevre, A., Delourme, R., Brun, H., 2010. Frequency of avirulence genes in *Leptosphaeria maculans* in western Canada. Canadian Journal of Plant Pathology 32, 77-85.

Lara-Á lvarez, I., Tharreau, D., Aguilar-Portero, M., Castejón-Muñoz, M., 2010. Evidence for rapid changes in the population genetic structure of *Magnaporthe oryzae* in Southern Spain. Journal of Phytopathology 158, 785-791.

Latha, A.M., Rao, K.V., Reddy, V.D., 2005. Production of transgenic plants resistant to leaf blast disease in finger millet (*Eleusine coracana* (L.) Gaertn.). Plant Science 169, 657-667.

Law, J.W., Ser, H., Khan, T.M., Chuah, L., Pusparajah, P., Chan, K., Goh, B., Lee, L., 2017. The potential of *Streptomyces* as biocontrol agents against the rice blast fungus, *Magnaporthe oryzae* (*Pyricularia oryzae*). Frontiers in Microbiology 8, 3-10.

Le, M.T., Arie, T., Teraoka, T., 2010. Population dynamics and pathogenic races of rice blast fungus, *Magnaporthe oryzae* in the Mekong Delta in Vietnam. Journal of General Plant Pathology 76, 177-182.

Lear, G., Dickie, I., Banks, J., Boyer, S., Buckley, H.L., Buckley, T.R., Cruickshank, R., Dopheide, A., Handley, K.M., Hermans, S., Kamke, J., Lee, C.K., MacDiarmid, R., Morales, S.E. Orlovich, D.A., Smissen, R., Wood, J., Holdaways, R., 2018. Methods for the extraction, storage, amplification and sequencing of DNA from environmental samples. New Zealand Journal of Ecology 42, 10-50.

Leigh, J., Bryant, D., Steel, M., 2015. PopART. Full-feature software for haplotype network construction. Methods in Ecology and Evolution 6, 1110-1116.

Lenné, J., Thomas, D., 2006. Integrating Crop—Livestock Research and Development in Sub-Saharan Africa: Option, Imperative or Impossible? Outlook on Agriculture 35, 167-175.

Lenne, J.M., Takan, J.P., Mgonja, M.A., Manyasa, E.O., Kaloki, P., Wanyera, N., Okwadi, J., Muthumeenakshi, S., Brown, A.E., Tamale, M., Sreenivasaprasad, S., 2007. Finger millet blast disease management: a key entry point for fighting malnutrition and poverty in East Africa. Outlook on Agriculture 36, 78-82.

Leslie, J.F., Klein, K.K., 1996. Female fertility and mating type effects on effective population size and evolution in filamentous fungi. Genetics 144, 557-567.

Leung, H., Nelson, R., Leach, J., 1993. Population structure of plant pathogenic fungi and bacteria. Advances in Plant Pathology 36, 240-245.

Levy, M., Correa-Victoria, F.J., Zeigler, R.S., Xu, S., Hamer, J.E., 1993. Genetic diversity of the rice blast fungus in a disease nursery in Colombia. Phytopathology 83, 1427-3143.

Levy, M., Romao, J., Marchetti, M.A., Hamer, J.E., 1991. DNA Fingerprinting with a dispersed repeated sequence resolves pathotype diversity in the rice blast fungus. Plant Cell 3, 95-102.

Lewis, K.A., Tzilivakis, J., Warner, D.J., Green, A., 2016. An international database for pesticide risk assessments and management. human and ecological risk assessment: An International Journal 22, 1050-1064.

Li, H., 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27, 2987-2993.

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25, 1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078-2090.

Li, L.M., Grassly, N.C., Fraser, C., 2014. Genomic analysis of emerging pathogens: methods, application and future trends. Genome Biology 15, 541-550.

Li, Q., Jiang, Y., Ning, P., Zheng, L., Huang, J., Li, G., Jiang, D., Hsiang, T., 2011. Suppression of *Magnaporthe oryzae* by culture filtrates of *Streptomyces globisporus* JK-1. Biological Control 58, 139-148.

Li, W., Wang, B., Wu, J., Lu, G., Hu, Y., Zhang, X., Zhang, Z., Zhao, Q., Feng, Q., Zhang, H., 2009. The *Magnaporthe oryzae* avirulence gene *AVRPiz-t* encodes a predicted secreted

protein that triggers the immunity in rice mediated by the blast resistance gene *Piz-t*. Molecular Plant-Microbe Interactions 22, 411-420.

Linnarsson, S., 2010. Recent advances in DNA sequencing methods–general principles of sample preparation. Experimental Cell Research 316, 1339-1343.

Lis-Balchin, M., Deans, S., 1997. Bioactivity of selected plant essential oils against *Listeria monocytogenes*. Journal of Applied Microbiology 82, 759-762.

Liu, S., Dean, R.A., 1997. G protein α subunit genes control growth, development, and pathogenicity of *Magnaporthe grisea*. Molecular Plant-Microbe Interactions 10, 1075-1086.

Long, D., Lee, F., TeBeest, D., 2000. Effect of nitrogen fertilization on disease progress of rice blast on susceptible and resistant cultivars. Plant Disease 84, 403-409.

Lopez, A.L.C., Cumagun, C.J.R., 2018. Genetic structure of *Magnaporthe oryzae* populations in three island groups in the Philippines. European Journal of Plant Pathology 56, 1-18.

Louwaars, N. P., de Boef, W.S., 2012. Integrated seed sector development in Africa: A conceptual framework for creating coherence between practices, programs and policies. Journal of Crop Improvement 26, 34 -45.

Lule, D., de Villiers, S., Fetene, M., Odeny, D.A., Rathore, A., Das, R.R., Tesfaye, K., 2018. Genetic diversity and association mapping of Ethiopian and exotic finger millet accessions. Crop and Pasture Science 69, 879-891.

Luo, Y., Hou, L., Förster, H., Pryor, B., Adaskaveg, J., 2017. Identification of *Alternaria species* Causing Heart Rot of Pomegranates in California. Plant Disease 101, 421-427.

Luo, J., Qiu, H., Cai, G., Wagner, N.E., Bhattacharya, D., Zhang, N., 2015. Phylogenomic analysis uncovers the evolutionary history of nutrition and infection mode in rice blast fungus and other Magnaporthales. Scientific Reports 5, 9448-9455.

Machida, M., Asai, K., Sano, M., Tanaka, T., Kumagai, T., Terai, G., Kusumoto, K., Arima, T., Akita, O., Kashiwagi, Y., 2005. Genome sequencing and analysis of *Aspergillus oryzae*. Nature 438, 1157-1163.

Maciel, J.L.N., Ceresini, P.C., Castroagudin, V.L., Zala, M., Kema, G.H., McDonald, B.A., 2014. Population structure and pathotype diversity of the wheat blast pathogen *Magnaporthe oryzae* 25 years after its emergence in Brazil. Phytopathology 104, 95-107.

Mahesh, H., Meghana, S., Shailaja, H., Mk, P., Mahadevu, P., Mv, C., Malali, G., 2016. Acquisition of the Grasshopper Retro Transposon by Rice *Magnaporthe* isolates indicates a Dynamic gene flow between rice and non-rice *Magnaporthe* population. Journal of Pathology and Microbiology 1, 1011-1022.

Malaker, P.K., Barma, N.C., Tiwari, T., Collis, W.J., Duveiller, E., Singh, P.K., Joshi, A.K., Singh, R.P., Braun, H., Peterson, G.L., 2016. First report of wheat blast caused by *Magnaporthe oryzae* pathotype *Triticum* in Bangladesh. Plant Disease 100, 2330-2333.

Manidipa, R., Dutta, S.G., Venkata, R.C., 2013. *Pseudomonads*: Potential biocontrol agents of rice diseases. Research Journal of Agriculture and Forestry Sciences 1, 19-25.

Mannai, S., Horrigue-Raouani, N., Boughalleb-M'Hamdi, N., 2018. Characterization of *Fusarium species* associated with apple decline in Tunisian nurseries. Journal of Biological Studies 1, 14-34.

Manyasa, E., Tongoona, P., Shanahan, P., Mgonja, M., De Villiers, S., 2015. Genetic diversity in East African finger millet (*Eleusine coracana* (L.) Gaertn) landraces based on SSR markers and some qualitative traits. Plant Genetic Resources 13, 45-55.

Marangoni, M.S., Nunes, M.P., Fonseca Jr, N., Mehta, Y.R., 2013. *Pyricularia* blast on white oats: a new threat to wheat cultivation. Tropical Plant Pathology 38, 198-202.

Marone, D., Russo, M.A., Laidò, G., De Vita, P., Papa, R., Blanco, A., Gadaleta, A., Rubiales, D., Mastrangelo, A.M., 2013. Genetic basis of qualitative and quantitative resistance to powdery mildew in wheat: from consensus regions to candidate genes. BMC Genomics 14, 562-565.

Marthey, S., Aguileta, G., Rodolphe, F., Gendrault, A., Giraud, T., Fournier, E., Lopez-Villavicencio, M., Gautier, A., Lebrun, M.H., Chiapello, H., 2008. FUNYBASE: a FUNgal phYlogenomic dataBASE. BMC Bioinformatics 9, 456-458.

Martinez, D., Larrondo, L.F., Putnam, N., Gelpke, M.D.S., Huang, K., Chapman, J., Helfenbein, K.G., Ramaiya, P., Detter, J.C., Larimer, F., 2004. Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. National Biotechnology 22, 695-699.

Mathur, P. N., 2012. Global strategy for the Ex Situ conservation of finger millet and its wild relatives. Global Crop Diversity Trust & ICRISAT, pp. 7-57.

McCormack, J.E., Hird, S.M., Zellmer, A.J., Carstens, B.C., Brumfield, R.T., 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. Molecular Phylogenetics and Evolution 66, 526-38.

McDonald, B.A., McDermott, J.M., 1993. Population genetics of plant pathogenic fungi. Bioscience 4, 311-319.

McDonald, B.A., 1997. The population genetics of fungi: tools and techniques. Phytopathology 87, 448-453.

McMullan, M., Rafiqi, M., Kaithakottil, G., Clavijo, B.J., Bilham, L., Orton, E., Percival-Alwyn, L., Ward, B.J., Edwards, A., Saunders, D.G., 2018. The ash dieback invasion of Europe was founded by two genetically divergent individuals. Nature Ecology & Evolution 2, 1000-1005.

Mekwatanakarn, P., Kositratana, W., Phromraksa, T., Zeigler, R., 1999. Sexually fertile *Magnaporthe grisea* rice pathogens in Thailand. Plant Disease 83, 939-943.

Mew, T.W., Leung, H., Savary, S., Vera Cruz, C.M., Leach, J.E., 2004. Looking ahead in rice disease research and management. Critical Reviews in Plant Sciences 23, 103-127.

Mgonja, M., Audi, P., Mgonja, A., Manyasa, E., Ojulong, O., 2013. Integrated Blast and Weed Management and Microdosing in Finger Millet: A HOPE Project Manual for Increasing Finger Millet Productivity. Project report. International Crops Research Institute for the Semi-Arid Tropics, Patancheru, Hyderabad, Andhra-Pradesh, pp. 760-1000.

Miah, G., Rafii, M., Ismail, M., Sahebi, M., Hashemi, F., Yusuff, O., Usman, M., 2017. Blast disease intimidation towards rice cultivation: a review of pathogen and strategies to control. Journal of Animal and Plant Sciences 27, 1058-1066.

Miki, S., Matsui, K., Kito, H., Otsuka, K., Ashizawa, T., Yasuda, N., Fukiya, S., Sato, J., Hirayae, K., Fujita, Y., 2009. Molecular cloning and characterization of the *AVR-Pia* locus from a Japanese field isolate of *Magnaporthe oryzae*. Molecular Plant Pathology 10, 361-374.

Milgroom, M.G., 2015. Population Biology of Plant Pathogens: Genetics, Ecology, and Evolution. The American Phytopathological Society (APS), 120-200.

Mirarab, S., Reaz, R., Bayzid, M.S., Zimmermann, T., Swenson, M.S., Warnow, T., 2014. ASTRAL: genome-scale coalescent-based species tree estimation. Bioinformatics 30, i541-i548.

Morales-Cruz, A., Amrine, K.C., Blanco-Ulate, B., Lawrence, D.P., Travadon, R., Rolshausen, P.E., Baumgartner, K., Cantu, D., 2015. Distinctive expansion of gene families associated with plant cell wall degradation, secondary metabolism, and nutrient uptake in the genomes of grapevine trunk pathogens. BMC Genomics 16, 469-475.

Motlagh, M.R.S., Hbibi, F., Ebadi, A.A., 2015. Genetic diversity of *Pyricularia grisea*, the causal agent of rice blast. Acta Scientiarum Polonorum-Hortorum Cultus 14, 43-50.

Mudita, I., Chiduza, C., Richardson-Kageler, S., Murungu, F., 2008. Evaluation of different strategies of intercropping maize (*Zea mays* L.) and soya bean [*Glycine max* (L.) Merrill] under small-holder production in sub-humid Zimbabwe. Journal of Agronomy 7, 237-243.

Mulè, G., Susca, A., Stea, G., Moretti, A., 2004. A species-specific PCR assay based on the calmodulin partial gene for identification of *Fusarium verticillioides*, *F. proliferatum* and *F. subglutinans*. European Journal of Plant Pathology 110, 495-502.

Mulualem, T., Melak, A., 2013. A survey on the status and constraints of finger millet (*Eleusine coracana* L.) production in Metekel Zone, North Western Ethiopia. Direct Research Journal of Agriculture and Food Science 1, 67-72.

Mutiga, S., Rotich, F., Ganeshan, V.D., Mwongera, D., Mgonja, E., Were, V., Harvey, J., Zhou, B., Wasilwa, L., Feng, C., 2017. Assessment of the virulence spectrum and its association with genetic diversity in *Magnaporthe oryzae* populations from Sub-Saharan Africa. Phytopathology 107, 852-863.

Mwongera, D.T., 2018. Determining the population structure and avirulence gene repertoire of the rice blast fungus *Magnaporthe oryzae* in Kenya by comparative genome analysis. PhD Thesis, University of Exeter, UK.

Nakayashiki, H., Matsuo, H., Chuma, I., Ikeda, K., Betsuyaku, S., Kusaba, M., Tosa, Y., Mayama, S., 2001. Pyret, a Ty3/Gypsy retrotransposon in *Magnaporthe grisea* contains an extra domain between the nucleocapsid and protease domains. Nucleic Acids Research 29, 4106-4113.

Narasimhan, V., Danecek, P., Scally, A., Xue, Y., Tyler-Smith, C., Durbin, R., 2016. BCFtools/RoH: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data. Bioinformatics 32, 1749-1751.

National Research Council (US) 1996. Finger millet. In: Lost crops of Africa: Board on Science, Technology for International Development. Volume I: Grains. National Academy Press, pp.39-58.

Negi, Y., Prabha, D., Garg, S., Kumar, J., 2017. Biological control of ragi blast disease by chitinase producing fluorescent *Pseudomonas* isolates. Organic Agriculture 7, 63-70.

Nierman, W.C., Pain, A., Anderson, M.J., Wortman, J.R., Kim, H.S., Arroyo, J., Berriman, M., Abe, K., Archer, D.B., Bermejo, C., 2005. Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. Nature 438, 1151-1171.

Nilsson, R.H., Kristiansson, E., Ryberg, M., Hallenberg, N., Larsson, K., 2008. Intraspecific ITS variability in the kingdom Fungi as expressed in the international sequence databases and its implications for molecular species identification. Evolutionary Bioinformatics 4, S653-S660.

Nilsson, R.H., Taylor, A., Adams, R.I., Baschien, C., Bengtsson-Palme, J., Cangren, P., Coleine, C., Daniel, H., Glassman, S.I., Hirooka, Y., Irinyi, L., Irsenaite, R., Martin-Sanchez, P., Meyer, W., Oh, S., Sampaio, J.P., Seifert, K.A., Sklenar, F., Stubbe, D., Suh, S., Summerbell, R., Svantesson, S., Unterseher, M., Visagie, C.M., Weiss, M., Woudenberg, J., Wurzbacher, C., Van, d.W., Yilmaz, N., Yurkov, A., Koljalg, U., Abarenkov, K., 2018. Taxonomic annotation of public fungal ITS sequences from the built environment - a report from an April 10-11, 2017 workshop (Aberdeen, UK).

Nitta, N., Farman, M., Leong, S., 1997. Genome organization of *Magnaporthe grisea*: integration of genetic maps, clustering of transposable elements and identification of genome duplications and rearrangements. Theoretical and Applied Genetics 95, 20-32.

Nottéghem, J., Silue, D., 1992. Distribution of the mating type alleles *in Magnaporthe grisea* populations pathogenic on rice. Phytopathology 82, 421-424.

Nurnberger, T., Brunner, F., 2002. Innate immunity in plants and animals: emerging parallels between the recognition of general elicitors and pathogen-associated molecular patterns. Current Opinion in Plant Biology 5, 456-465.

O'Connell, J., Schulz-Trieglaff, O., Carlson, E., Hims, M.M., Gormley, N.A., Cox, A.J., 2015. NxTrim: optimized trimming of Illumina mate pair reads. Bioinformatics 31, 2035-2037.

O'Donnell, K., Ward, T.J., Robert, V.A., Crous, P.W., Geiser, D.M., Kang, S., 2015. DNA sequence-based identification of *Fusarium*: current status and future directions. Phytoparasitica 43, 583-595.

O'Gorman, C.M., Fuller, H.T., Dyer, P.S., 2009. Discovery of a sexual cycle in the opportunistic fungal pathogen *Aspergillus fumigatus*. Nature 457, 471-480.

Obilo, O., Daniel, A., Ihejirika, G., Ofor, M., Adikuru, N., 2012. Control of rice blast (*Magnaporthe grisea*) disease using various organic manures. Journal of Agricultural Extension and Rural Development 15, 1198-1205.

Oduori, C.O., 2008. Breeding investigations of finger millet characteristics including blast disease and striga resistance in western Kenya. PhD Thesis, Mississippi State University, USA.

Oduori, C., Kanyenji, B., 2005. Finger Millet in Kenya: Importance, Advances in R&D, Challenges and Opportunities for Improved Production and Profitability. Finger Millet Blast Management in East Africa. Creating opportunities for improving production and utilization of finger millet. Mgonja, M.A., Lenne, J.M., Manyasa, E., Sreenivasaprasad, S. (eds), pp.10-20.

Oduori, C., 1993. Small millets production and research in Kenya. Advances in small millets (Riley KW, Gupta SC, Seetharam A and Mushonga JN, eds.). New Delhi, India: Oxford and IBH Publishing, 67-73.

Ogier, J., Calteau, A., Forst, S., Goodrich-Blair, H., Roche, D., Rouy, Z., Suen, G., Zumbihl, R., Givaudan, A., Tailliez, P., 2010. Units of plasticity in bacterial genomes: new insight from the comparative genomics of two bacteria interacting with invertebrates, *Photorhabdus* and *Xenorhabdus*. BMC Genomics 11, 568-570.

Okagaki, L.H., Sailsbery, J.K., Eyre, A.W., Dean, R.A., 2016. Comparative genome analysis and genome evolution of members of the Magnaporthaceae family of fungi. BMC Genomics 17, 135-140.

Onaga, G., Wydra, K.D., Koopmann, B., Sere, Y., von Tiedemann, A., 2015. Population structure, pathogenicity and mating type distribution of *Magnaporthe oryzae* isolates from East Africa. Phytopathology 678, 68-75.

Onyango, A.O., 2016. Finger millet: food security crop in the arid and semi-arid lands (ASALs) of Kenya. World Environment 6, 62-70.

Orbach, M.J., Farrall, L., Sweigard, J.A., Chumley, F.G., Valent, B., 2000. A telomeric avirulence gene determines efficacy for the rice blast resistance gene *Pi-ta*. Plant Cell 12, 2019-2032.

Oryokot, J., 2001. Finger millet (*Eleusine coracana*) (L.) Gaertn) in Agriculture in Uganda. Vol. II. Crops (Mukiibi JK, ed.). Fountain publishers/CTA/NARO Uganda pp. 29-41.

Oses-Ruiz, M., Sakulkoo, W., Littlejohn, G.R., Martin-Urdiroz, M., Talbot, N.J., 2017. Two independent S-phase checkpoints regulate appressorium-mediated plant infection by the rice blast fungus *Magnaporthe oryzae*. Proceedings of the National Academy of Sciences of the United States of America 114, E237-E244.

Osuna-Canizalez, F., De Datta, S., Bonman, J., 1991. Nitrogen form and silicon nutrition effects on resistance to blast disease of rice. Plant Soil 135, 223-231.

Ou, S., 1985. Rice diseases second Edition. Commonwealth Mycological Institute Kew, Surrey, England 380-390.

Owere, L., Tongoona, P., Derera, J., Wanyera, N., 2014. Farmers' perceptions of finger millet production constraints, varietal preferences and their implications to finger millet breeding in Uganda. Journal of Agricultural Science 6, p126-130.

Pagliaccia, D., Urak, R.Z., Wong, F., Douhan, L.I., Greer, C.A., Vidalakis, G., Douhan, G.W., 2018. Genetic Structure of the Rice Blast Pathogen (*Magnaporthe oryzae*) over a Decade in North Central California rice fields. Microbial Ecology 75, 310-317.

Pais, M., Win, J., Yoshida, K., Etherington, G.J., Cano, L.M., Raffaele, S., Banfield, M.J., Jones, A., Kamoun, S., Saunders, D.G., 2013. From pathogen genomes to host plant processes: the power of plant parasitic oomycetes. Genome Biology 14, 211-217.

Panwar, P., Jha, A.K., Pandey, P., Gupta, A.K., Kumar, A., 2011. Functional markers based molecular characterization and cloning of resistance gene analogs encoding NBS-LRR disease resistance proteins in finger millet (*Eleusine coracana*). Molecular Biology Reports 38, 3427-3436.

Paoletti, M., Rydholm, C., Schwier, E.U., Anderson, M.J., Szakacs, G., Lutzoni, F., Debeaupuis, J., Latgé, J., Denning, D.W., Dyer, P.S., 2005. Evidence for sexuality in the opportunistic fungal pathogen *Aspergillus fumigatus*. Current Biology 15, 1242-1248.

Park, S., Chi, M., Milgroom, M.G., Kim, H., Han, S., Kang, S., Lee, Y., 2010. Genetic stability of *Magnaporthe oryzae* during successive passages through rice plants and on artificial medium. Plant Pathology Journal 26, 313-320.

Park, S., Milgroom, M.G., Han, S., Kang, S., Lee, Y., 2003. Diversity of pathotypes and DNA fingerprint haplotypes in populations of *Magnaporthe grisea* in Korea over two decades. Phytopathology 93, 1378-1385.

Park, S., Milgroom, M., Han, S., Kang, S., Lee, Y., 2008. Genetic differentiation of *Magnaporthe oryzae* populations from scouting plots and commercial rice fields in Korea. Phytopathology 98, 436-442.

Park, J.Y., Jin, J., Lee, Y.W., Kang, S., Lee, Y.H., 2009. Rice blast fungus (*Magnaporthe oryzae*) infects Arabidopsis via a mechanism distinct from that required for the infection of rice. Plant Physiology 149, 474-486.

Patel, R., Tripathi, S., 1998. Epidemiology of Blast of Finger-Millet Caused by *Pyricularia grisea* (CKE) Sacc. Advances in Plant Sciences 11, 73-76.

Peixoto, L.F., 2014. Identification of the compatibility, sexuality, fertility and avirulence in populations of *Magnaporthe oryzae*, of Brazilian rice plantations. PhD Thesis, Goias Federal University, Brazil.

Perelló, A., Martinez, I., Sanabria, A., Altamirano, R., Sibole, J., 2017. Pathogenicity of isolates of *Magnaporthe* spp. from wheat and grasses infecting seedlings and mature wheat plants in Argentina. Plant Pathology 66, 1149-1161.

Persinoti, G.F., Martinez, D.A., Li, W., Dogen, A., Billmyre, R.B., Averette, A., Goldberg, J.M., Shea, T., Young, S., Zeng, Q., Oliver, B.G., Barton, R., Metin, B., Hilmioglu-Polat, S., Ilkit, M., Graser, Y., Martinez-Rossi, N.M., White, T.C., Heitman, J., Cuomo, C.A., 2018. Whole-genome analysis illustrates global clonal population structure of the ubiquitous Dermatophyte Pathogen *Trichophyton rubrum*. Genetics 208, 1657-1669.

Petes, T.D., 2001. Meiotic recombination hot spots and cold spots. Nature Reviews Genetics 2, 360-362.

Pilet-Nayel, M., Moury, B., Caffier, V., Montarry, J., Kerlan, M., Fournet, S., Durel, C., Delourme, R., 2017. Quantitative resistance to plant pathogens in pyramiding strategies for durable crop protection. Frontiers in Plant Science 8, 1831-1835.

Planet, P.J., Narechania, A., Chen, L., Mathema, B., Boundy, S., Archer, G., Kreiswirth, B., 2017. Architecture of a species: phylogenomics of *Staphylococcus aureus*. Trends in Microbiology. 25, 153-166.

Poczai, P., Hyvönen, J., 2010. Nuclear ribosomal spacer regions in plant phylogenetics: problems and prospects. Molecular Biology Reports 37, 1897-1912.

Pooja, K., Katoch, A., 2014. Past, present and future of rice blast management. Plant Science Today 1, 165-173.

Postel, S., Kemmerling, B., 2009. Plant systems for recognition of pathogen-associated molecular patterns. Seminars in Cell and Development Biology 20, 1025-1031.

Powell, A.J., Conant, G.C., Brown, D.E., Carbone, I., Dean, R.A., 2008. Altered patterns of gene duplication and differential gene gain and loss in fungal pathogens. BMC Genomics 9, 147-154.

Prasanna Kumar, M., Amruta, N., Manjula, C., Puneeth, M., Teli, K., 2017. Characterisation, screening and selection of *Bacillus subtilis* isolates for its biocontrol efficiency against major rice diseases. Biocontrol Science and Technology 27, 581-599.

Purseglove, J.W., 1982. Tropical Crops. Monocotyledons. Vol. 1. Longman Group Limited, pp.46-156.

Raffaele, S., Kamoun, S., 2012. Genome evolution in filamentous plant pathogens: why bigger can be better. Nature Reviews Microbiology 10, 417-420.

Raffaele, S., Farrer, R.A., Cano, L.M., Studholme, D.J., MacLean, D., Thines, M., Jiang, R.H., Zody, M.C., Kunjeti, S.G., Donofrio, N.M., Meyers, B.C., Nusbaum, C., Kamoun, S., 2010. Genome evolution following host jumps in the Irish potato famine pathogen lineage. Science 330, 1540-1543.

Rajasekaran, N., Nithya, M., Rose, C., Chandra, T., 2004. The effect of finger millet feeding on the early responses during the process of wound healing in diabetic rats. Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease 1689, 190-201.

Ramashia, S., Gwata, E., Meddows-Taylor, S., Anyasi, T., Jideani, A., 2018. Some physical and functional properties of finger millet (*Eleusine coracana*) obtained in sub-Saharan Africa. Food Research International 104, 110-118.

Ramdial, H., Latchoo, R., Hosein, F., Rampersad, S., 2016. Phylogeny and haplotype analysis of fungi within the *Fusarium incarnatum-equiseti* species Complex. Phytopathology 107, 109-120.

Ramirez-Prado, J., Moore, G.G., Horn, B.W., Carbone, I., 2008. Characterization and population analysis of the mating-type genes in *Aspergillus flavus* and *Aspergillus parasiticus*. Fungal Genetics and Biology 45, 1292-1299.

Rathour, R., Singh, B., Sharma, T., Chauhan, R., 2004. Population Structure of *Magnaporthe grisea* from North-western Himalayas and its Implications for Blast Resistance Breeding of Rice. Journal of Phytopathology 152, 304-312.

Ray, S., Singh, P.K., Gupta, D.K., Mahato, A.K., Sarkar, C., Rathour, R., Singh, N.K., Sharma, T.R., 2016. Analysis of *Magnaporthe oryzae* genome reveals a fungal effector, which is able to induce resistance response in transgenic rice line containing resistance gene, *Pi54*. Frontiers in Plant Science 7, 1140-1144.

Rehmeyer, C., Li, W., Kusaba, M., Kim, Y.S., Brown, D., Staben, C., Dean, R., Farman, M., 2006. Organization of chromosome ends in the rice blast fungus, *Magnaporthe oryzae*. Nucleic Acids Research 34, 4685-4701.

Ribot, C., Cesari, S., Abidi, I., Chalvon, V., Bournaud, C., Vallet, J., Lebrun, M., Morel, J., Kroj, T., 2013. The *Magnaporthe oryzae* effector *AVR1–CO39* is translocated into rice cells independently of a fungal-derived machinery. Plant Journal 74, 1-12.

Ribot, C., Hirsch, J., Balzergue, S., Tharreau, D., Nottéghem, J., Lebrun, M., Morel, J., 2008. Susceptibility of rice to the blast fungus, *Magnaporthe grisea*. Journal of Plant Physiology. 165, 114-124.

Robert, Vincent Albert Romain Ghislain, 2016. Sequence-based classification and identification of fungi. Mycologia 108, 16-130.

Ronald, P., 2011. Plant genetics, sustainable agriculture and global food security. Genetics 188, 11-20.

Ross, M.G., Russ, C., Costello, M., Hollinger, A., Lennon, N.J., Hegarty, R., Nusbaum, C., Jaffe, D.B., 2013. Characterizing and measuring bias in sequence data. Genome Biology 14, R51-R55.

Roumen, E., Levy, M., Nottéghem, J., 1997. Characterisation of the European pathogen population of *Magnaporthe grisea* by DNA fingerprinting and pathotype analysis. European Journal of Plant Pathology 103, 363-371.

Rouxel, T., Balesdent, M., 2017. Life, death and rebirth of avirulence effectors in a fungal pathogen of Brassica crops, *Leptosphaeria maculans*. New Phytologist 214, 526-532.

Ryan, J.R., 2016. Chapter 8 - Biological Threat to Agriculture. In: Ryan, J.R. (Ed.), Biosecurity and Bioterrorism (Second Edition). Butterworth-Heinemann, pp. 185-216.

Sakulkoo, W., Osés-Ruiz, M., Garcia, E.O., Soanes, D.M., Littlejohn, G.R., Hacker, C., Correia, A., Valent, B., Talbot, N.J., 2018. A single fungal MAP kinase controls plant cell-to-cell invasion by the rice blast fungus. Science 359, 1399-1403.

Saleh, A.S., Zhang, Q., Chen, J., Shen, Q., 2013. Millet grains: nutritional quality, processing, and potential health benefits. Comprehensive Reviews in Food Science and Food Safety 12, 281-295.

Saleh, D., Milazzo, J., Adreit, H., Fournier, E., Tharreau, D., 2014. South-East Asia is the center of origin, diversity and dispersion of the rice blast fungus, *Magnaporthe oryzae*. New Phytologist. 201, 1440-1456.

Saleh, D., Xu, P., Shen, Y., Li, C., Adreit, H., Milazzo, J., Ravigne, V., Bazin, E., Notteghem, J., Fournier, E., 2012. Sex at the origin: an Asian population of the rice blast fungus *Magnaporthe oryzae* reproduces sexually. Molecular Ecology 21, 1330-1344.

Saleh, D., Milazzo, J., Adreit, H., Tharreau, D., Fournier, E., 2012. Asexual reproduction induces a rapid and permanent loss of sexual reproduction capacity in the rice fungal pathogen *Magnaporthe oryzae*: results of in vitro experimental evolution assays. BMC Evolution Biology 12, 422148-422160.

Samanta, S., Dhua, U., Nayak, S., Behera, L., Mukherjee, A., 2014. Mating types analysis of *Magnaporthe oryzae* populations by molecular methods. Open Biotechnology Journal 8, 6-12.

Sanchez Jr, E., Asano, K., Sone, T., 2011. Characterization of Inago1 and Inago2 retrotransposons in *Magnaporthe oryzae*. Journal of General Plant Pathology 77, 239-242.

Sánchez-Vallet, A., Fouché, S., Fudal, I., Hartmann, F.E., Soyer, J.L., Tellier, A., Croll, D., 2018. The genome biology of effector gene evolution in filamentous plant pathogens. Annual Review of Phytopathology 56, 21-40.

Sanger, F., Air, G.M., Barrell, B.G., Brown, N.L., Coulson, A.R., Fiddes, J.C., Hutchison III, C.A., Slocombe, P.M., Smith, M., 1977. Nucleotide sequence of bacteriophage ϕX174 DNA. Nature 265, 687-690.

Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences of the United States of America

74, 5463-5467.

Santana, M.F., Silva, J.C., Mizubuti, E.S., Araújo, E.F., Condon, B.J., Turgeon, B.G., Queiroz, M.V., 2014. Characterization and potential evolutionary impact of transposable elements in the genome of *Cochliobolus heterostrophus*. BMC Genomics 15, 536-540.

Savary, S., Ficke, A., Aubertot, J., Hollier, C., 2012. Crop losses due to diseases and their implications for global food production losses and food security. Food Security 4, 519-537.

Sawada, H., Sugihara, M., Takagaki, M., Nagayama, K., 2004. Monitoring and characterization of *Magnaporthe grisea* isolates with decreased sensitivity to scytalone dehydratase inhibitors. Pest Management Science: formerly Pesticide Science 60, 777-785.

Sayyari, E., Mirarab, S., 2016. Fast coalescent-based computation of local branch support from quartet frequencies. Molecular Biology and Evolution 33, 1654-1668.

Schafer, W., 1994. Molecular mechanisms of fungal pathogenicity to plants. Annual Review of Phytopathology 32, 461-477.

Schoch, C.L., Sung, G., López-Giráldez, F., Townsend, J.P., Miadlikowska, J., Hofstetter, V., Robbertse, B., Matheny, P.B., Kauff, F., Wang, Z., 2009. The Ascomycota tree of life: a phylum-wide phylogeny clarifies the origin and evolution of fundamental reproductive and ecological traits. Systematic Biology 58, 224-239.

Schoch, C.L., Seifert, K.A., Huhndorf, S., Robert, V., Spouge, J.L., Levesque, C.A., Chen, W., Fungal Barcoding Consortium, Fungal Barcoding Consortium Author List, 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for

Fungi. Proceedings of the National Academy of Sciences of the United States of America 109, 6241-6246.

Scholthof, K.B., 2007. The disease triangle: pathogens, the environment and society. Nature Reviews Microbiology, 5, 152-156.

Schuster, S.C., 2007. Next-generation sequencing transforms today's biology. Nature Methods 5, 16-24.

Seehausen, O., Butlin, R.K., Keller, I., Wagner, C.E., Boughman, J.W., Hohenlohe, P.A., Peichel, C.L., Saetre, G., Bank, C., Brännström, Å., 2014. Genomics and the origin of species. Nature Reviews Genetics 15, 176-180.

Seidl, M.F., Thomma, B.P., 2014. Sex or no sex: evolutionary adaptation occurs regardless. Bioessays 36, 335-345.

Séré, Y., Onasanya, A., Afolabi, A., Mignouna, H., Akator, K., 2007. Genetic diversity of the blast fungus, *Magnaporthe grisea* (Hebert) Barr, in Burkina Faso. African Journal of Biotechnology 6, 71-81.

Séré, Y., Sy, A., Sie, M., Akator, S., Onasanya, A., Kabore, B., Conde, C., Traore, M., Kiepe, P., 2011. Importance of varietal improvement for blast disease control in Africa. Japan International Research Center for Agriculture Sciences, JIRCAS Newletter 70, 77-90.

Shekhar, S., Roch, S., Mirarab, S., 2017. Species tree estimation using ASTRAL: how many genes are enough? IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB) 15, 1738-1747.

Shen, W., Bobrowicz, P., Ebbole, D.J., 1999. Isolation of pheromone precursor genes of *Magnaporthe grisea*. Fungal Genetics and Biology 27, 253-63.

Shendure, J., Ji, H., 2008. Next-generation DNA sequencing. Nature Biotechnology 26, 1135-1138.

Sheppard, S.K., Guttman, D.S., Fitzgerald, J.R., 2018. Population genomics of bacterial host adaptation. Nature Reviews Genetics 19, 549-565.

Shirke, M.D., Mahesh, H., Gowda, M., 2016. Genome-wide comparison of *Magnaporthe species* reveals a host-specific pattern of secretory proteins and transposable elements. PloS One 11, e0162458.

Shnyreva, A., 2003. Transposable elements are the factors involved in various rearrangements and modifications of the fungal genomes. Russian Journal of Genetics 39, 505-518.

Shobana, S., Krishnaswamy, K., Sudha, V., Malleshi, N., Anjana, R., Palaniappan, L., Mohan, V., 2013. Finger millet (Ragi, *Eleusine coracana* L.): a review of its nutritional properties, processing, and plausible health benefits. Advances in Food and Nutrition Research 69, 1-39.

Silva, G.B., Prabhu, A.S., Filippi, M.C., Trindade, M.G., Araújo, L.G., Zambolim, L., 2009. Genetic and phenotypic diversity of *Magnaporthe oryzae* from leaves and panicles of rice in commercial fields in the State of Goiás, Brazil. Tropical Plant Pathology 34, 77-86.

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M., 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210-3222.

Singh, U.B., Malviya, D., Wasiullah, Singh, S., Pradhan, J.K., Singh, B.P., Roy, M., Imram, M., Pathak, N., Baisyal, B.M., Rai, J.P., Sarma, B.K., Singh, R.K., Sharma, P.K., Kaur, S.D., Manna, M.C., Sharma, S.K., Sharma, A.K., 2016. Bio-protective microbial agents from rhizosphere eco-systems trigger plant defense responses provide protection against sheath blight disease in rice (*Oryza sativa* L.). Microbiological Research 192, 300-312.

Sirisathaworn, T., Srirat, T., Longya, A., Jantasuriyarat, C., 2017. Evaluation of mating type distribution and genetic diversity of three *Magnaporthe oryzae* avirulence genes, *PWL-2*, *AVR-Pii* and *AVR-Piz-t*, in Thailand rice blast isolates. Agriculture and Natural Resources 51, 7-14.

Skamnioti, P., Gurr, S.J., 2009. Against the grain: safeguarding rice from rice blast disease. Trends in Biotechnology 27, 141-150.

Soderlund, C., Nelson, W., Shoemaker, A., Paterson, A., 2006. SyMAP: A system for discovering and viewing syntenic regions of FPC maps. Genome Research 16, 1159-1168.

Sone, T., Suto, M., Tomita, F., 1993. Host species-specific repetitive DNA sequence in the genome of *Magnaporthe grisea*, the rice blast fungus. Bioscience Biotechnology and Biochemistry 57, 1228-1230.

Sone, T., Takeuchi, S., Miki, S., Satoh, Y., Ohtsuka, K., Abe, A., Asano, K., 2013. Homologous recombination causes the spontaneous deletion of *AVR-Pia* in *Magnaporthe oryzae*. FEMS Microbiology Letters 339, 102-109.

Sood, S., Babu, B.K., 2016. Finger Millet. In: Anonymous Broadening the Genetic Base of Grain Cereals. Springer, pp. 225-256.

Soreng, R.J., Peterson, P.M., Romaschenko, K., Davidse, G., Zuloaga, F.O., Judziewicz, E.J. Filgueiras, T.S., Davis, J.I., Morrone, O., 2015. A worldwide phylogenetic classification of the Poaceae (Gramineae). Journal of Systematics & Evolution 53, 117-120.

Sreenivasaprasad, S., Meehan, B., Mills, P., Brown, A., 1996. Phylogeny and systematics of 18 *Colletotrichum* species based on ribosomal DNA spacer sequences. Genome 39, 499-512.

Sreenivasaprasad, S., Takan, J., Muthumeenakshi, S., Obilana, A., Manyasa, E., Audi, P., Oduori, C., Coll, R., Brown, A., Talbot, N., Bandyopadhyay, R., 2005. Finger millet blast in East Africa: Pathogen diversity and disease management. Finger millet blast management in East Africa. Creating opportunities for improving production and utilization of finger millet, 34-45.

Srivastava, D., Shamim, M., Kumar, D., Pandey, P., Khan, N., Singh, K., 2014. Morphological and molecular characterization of *Pyricularia oryzae* causing blast disease in Rice (*Oryza sativa*) from North India. International Journal of Scientific and Research publication 4, 2250-3153.

Srivastava, D., Shamim, M., Kumar, M., Mishra, A., Pandey, P., Kumar, D., Yadav, P., Siddiqui, M.H., Singh, K.N., 2017. Current status of conventional and molecular interventions for blast resistance in rice. Rice Science 24, 299-321.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312-1313.

Stanke, M., Keller, O., Gunduz, I., Hayes, A., Waack, S., Morgenstern, B., 2006. AUGUSTUS: ab initio prediction of alternative transcripts. Nucleic Acids Research 34, W435-W439.

Starnes, J.H., 2013. Characterization and distribution of novel non-LTR retroelements driving high telomere RFLP diversity in clonal lines of *Magnaporthe oryzae*. PhD Thesis, University of Kentucky, USA.

Starnes, J.H., Thornbury, D.W., Novikova, O.S., Rehmeyer, C.J., Farman, M.L., 2012. Telomere-targeted retrotransposons in the rice blast fungus *Magnaporthe oryzae*: agents of telomere instability. Genetics 191, 389-406.

Strange, R.N., Scott, P.R., 2005. Plant disease: a threat to global food security. Annual Review of Phytopathology 43, 83-116.

Stukenbrock, E.H., Jørgensen, F.G., Zala, M., Hansen, T.T., McDonald, B.A., Schierup, M.H., 2010. Whole-genome and chromosome evolution associated with host adaptation and speciation of the wheat pathogen *Mycosphaerella graminicola*. PLoS Genetics 6, e1001189.

Sucher, N.J., Hennell, J.R., Carles, M.C., 2012. DNA fingerprinting, DNA barcoding, and next generation sequencing technology in plants. In: Anonymous Plant DNA Fingerprinting and Barcoding. Springer, pp. 13-22.

Sun, X., Kang, S., Zhang, Y., Tan, X., Yu, Y., He, H., Zhang, X., Liu, Y., Wang, S., Sun, W., Cai, L., Li, S., 2013. Genetic Diversity and Population Structure of Rice Pathogen *Ustilaginoidea virens* in China. PloS One 8, e76879.

Suprapta, D.N., 2012. Potential of microbial antagonists as biocontrol agents against plant fungal pathogens. Journal of ISSAAS 18, 1-8.

Suzuki, F., Arai, M., Yamaguchi, J., 2007. Genetic analysis of *Pyricularia grisea* population by rep-PCR during development of resistance to scytalone dehydratase inhibitors of melanin biosynthesis. Plant Disease 91, 176-184.

Suzuki, H., 1975. Meteorological factors in the epidemiology of rice blast. Annual Review of Phytopathology 13, 239-256.

Swami, S.B., Thakor, N.J., Gurav, H.S., 2013. Effect of soaking and malting on finger millet (*Eleusine coracana*) grain. Agricultural Engineering International: CIGR Journal 15, 194-200.

Sweigard, J.A., Carroll, A.M., Kang, S., Farrall, L., Chumley, F.G., Valent, B., 1995. Identification, cloning, and characterization of *PWL2,* a gene for host species specificity in the rice blast fungus. Plant Cell 7, 1221-1233.

Swerdlow, H., Wu, S., Harke, H., Dovichi, N.J., 1990. Capillary gel electrophoresis for DNA sequencing: laser-induced fluorescence detection with the sheath flow cuvette. Journal of Chromatography A 516, 61-67.

Taheri, P., Irannejad, A., 2014. Genetic structure of various *Magnaporthe oryzae* populations in Iran and Uruguay. Australasian Plant Pathology 43, 287-297.

Takan, J.P., 2007. A Study of the Genetic diversity, sexuality and pathogenicity of the finger millet blast pathogen *Magnaporthe grisea* populations in two East African countries. PhD Thesis, University of Exeter, UK.

Takan, J., 2004. Finger millet blast pathogen diversity and management in East Africa: A summary of project activities and outputs. International sorghum and millets newsletter 45, 66-69.

Takan, J., Chipili, J., Muthumeenakshi, S., Talbot, N., Manyasa, E., Bandyopadhyay, R., Sere, Y., Nutsugah, S., Talhinhas, P., Hossain, M., Brown, A.E., Sreenivasaprasad, S., 2012. *Magnaporthe oryzae* populations adapted to finger millet and rice exhibit distinctive patterns of genetic diversity, sexuality and host interaction. Molecular Biotechnology 50, 145-158.

Talbot, N.J., 2003. On the trail of a cereal killer: exploring the biology of *Magnaporthe grisea*. Annual Reviews in Microbiology 57, 177-202.

Talbot, N.J., 1995. Having a blast: exploring the pathogenicity of *Magnaporthe grisea*. Trends in Microbiology 3, 9-16.

Talbot, N.J., Salch, Y.P., Ma, M., Hamer, J.E., 1993. Karyotypic variation within clonal lineages of the rice blast fungus, *Magnaporthe grisea*. Applied and Environmental Microbiology 59, 585-593.

Talbot, N.J., Kershaw, M.J., Wakley, G.E., De Vries, O., Wessels, J., Hamer, J.E., 1996. MPG1 Encodes a Fungal Hydrophobin Involved in Surface Interactions during Infection-Related Development of *Magnaporthe grisea*. Plant Cell 8, 985-999.

Talhinhas, P., Sreeinvasaprasad, S., Neves-Marthis, J., Oliverira, H., 20002. Genetic and morphological characterization of *Colletotrichum acutatum* causing anthracnose of lupins. Phytopathology 92, 986-996.

Tamreihao, K., Ningthoujam, D.S., Nimaichand, S., Singh, E.S., Reena, P., Singh, S.H., Nongthomba, U., 2016. Biocontrol and plant growth promoting activities of a *Streptomyces corchorusii* strain UCR3-16 and preparation of powder formulation for application as biofertilizer agents for rice plant. Microbiological Research 192, 260-270.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24, 1596-1599.

Tanabe, Y., Saikawa, M., Watanabe, M.M., Sugiyama, J., 2004. Molecular phylogeny of Zygomycota based on EF-1α and RPB1 sequences: limitations and utility of alternative markers to rDNA. Molecular Phylogenetics and Evolution 30, 438-449.

Tanaka, M., Nakayashiki, H., Tosa, Y., 2009. Population structure of *Eleusine* isolates of *Pyricularia oryzae* and its evolutionary implications. Journal of General Plant Pathology 75, 173-180.

Taylor, J.W., 2015. Evolutionary Perspectives on Human Fungal Pathogens. Cold Spring Harbour Perspectives in Medicine 5, a019588.

Teixeira, M.d.M., Patané, J.S., Taylor, M.L., Gómez, B.L., Theodoro, R.C., de Hoog, S., Engelthaler, D.M., Zancopé-Oliveira, R.M., Felipe, M.S., Barker, B.M., 2016. Worldwide phylogenetic distributions and population dynamics of the *genus Histoplasma*. PLoS Neglected Tropical diseases 10, e0004732.

Terauchi, R., Kanzaki, H., Fujisaki, K., Takagi, H., Abe, A., Yoshida, K., Okuyama, Y., Tamiru, M., Saitoh, H., 2016. Whole genome sequencing approaches to understand *Magnaporthe*-rice interactions. Physiological and Molecular Plant Pathology 95, 4-7.

Tesfaye, K., Mengistu, S., 2017. Phenotypic characterization of Ethiopian finger millet accessions (*Eleusine coracana* (L.) Gaertn), for their agronomically important traits. Acta Universitatis Sapientiae, Agriculture and Environment 9, 107-118.

Thon, M.R., Pan, H., Diener, S., Papalas, J., Taro, A., Mitchell, T.K., Dean, R.A., 2006. The role of transposable element clusters in genome evolution and loss of synteny in the rice blast fungus *Magnaporthe oryzae*. Genome Biology 7, R16-R20.

Tosa, Y., Hirata, K., Tamba, H., Nakagawa, S., Chuma, I., Isobe, C., Osue, J., Urashima, A., Don, L., Kusaba, M., 2004. Genetic constitution and pathogenicity of Lolium isolates of

Magnaporthe oryzae in comparison with host species-specific pathotypes of the blast fungus. Phytopathology 94, 454-462.

Tredway, L., Stevenson, K., Burpee, L., 2005. Genetic structure of *Magnaporthe grisea* populations associated with St. Augustinegrass and tall fescue in Georgia. Phytopathology 95, 463-471.

Tredway, L., Stevenson, K., Burpee, L., 2003. Mating type distribution and fertility status in *Magnaporthe grisea* populations from turfgrasses in Georgia. Plant Disease 87, 435-441.

Trevathan, L.E., Moss, M.A., Blasingame, D., 1994. Ryegrass blast. Plant Disease (USA), pp. 180-300.

Tsui, C.K., DiGuistini, S., Wang, Y., Feau, N., Dhillon, B., Bohlmann, J., Hamelin, R.C., 2013. Unequal recombination and evolution of the mating-type (*MAT*) loci in the pathogenic fungus *Grosmannia clavigera* and relatives. G3 (Bethesda) 3, 465-480.

Turgeon, B.G., Yoder, O., 2000. Proposed nomenclature for mating type genes of filamentous ascomycetes. Fungal Genetics and Biology 31, 1-5.

Uddin, B.W., Viji, G., Vincelli, P., 2003. Gray leaf spot of perennial ryegrass turf: an emerging problem for the turfgrass industry. USGA Green Section Record 41, 9-13.

Underhill, D.M., Iliev, I.D., 2014. The mycobiota: interactions between commensal fungi and the host immune system. Nature Reviews Immunology 40, 405-415.

Upadhyaya, H., Gowda, C., Reddy, V.G., 2007. Morphological diversity in finger millet germplasm introduced from Southern and Eastern Africa. Journal of SAT Agricultural Research 3, 1-3.

Upson, J.L., Zess, E.K., Białas, A., Wu, C., Kamoun, S., 2018. The coming of age of EvoMPMI: evolutionary molecular plant–microbe interactions across multiple timescales. Current Opinion in Plant Biology 44, 108-116.

Urashima, A.S., Hashimoto, Y., Don, L. D., Kusaba, M., Tosa, Y., Nakayashiki, H., Mayama, S., 1999. Molecular analysis of the wheat blast population in Brazil with a homolog of retrotransposon MGR583. Japanese Journal of Phytopathology 65, 429-436.

Urashima, A.S., Silva, C.P., 2011. Characterization of *Magnaporthe grisea* (*Pyricularia grisea*) from black oat in Brazil. Journal of Phytopathology 159, 789-795.

Urashima, A., Igarashi, S., Kato, H., 1993. Host range, mating type, and fertility of *Pyricularia grisea* from wheat in Brazil. Plant Disease 77, 1211-1216.

Valent, B., Farrall, L., Chumley, F.G., 1991. *Magnaporthe grisea* genes for pathogenicity and virulence identified through a series of backcrosses. Genetics 127, 87-101.

Van Dijk, E.L., Auger, H., Jaszczyszyn, Y., Thermes, C., 2014a. Ten years of next-generation sequencing technology. Trends in Genetics 30, 418-426.

Van Dijk, E.L., Jaszczyszyn, Y., Thermes, C., 2014b. Library preparation methods for next-generation sequencing: tone down the bias. Experimental Cell Research 322, 12-20.

Vasudevan, P., Kavitha, S., Priyadarisini, V.B., Babujee, L., Gnanamanickam, S.S., 2002. Biological control of rice diseases. Biological control of crop diseases. Gnanamanickam ed. Marcel Dekker Inc, New York, 11-32.

Viji, G., Gnanamanickam, S., 1998. Mating type distribution and fertility status of *Magnaporthe grisea* populations from various hosts in India. Plant Disease 82, 36-40.

Viji, G., Gnanamanickam, S., Levy, M., 2000. DNA polymorphisms of isolates of *Magnaporthe grisea* from India that are pathogenic to finger millet and rice. Mycological Research 104, 161-167.

Viji, G., Uddin, W., 2002. Distribution of mating type alleles and fertility status of *Magnaporthe grisea* causing gray leaf spot of perennial ryegrass and St. Augustinegrass turf. Plant Disease 86, 827-832.

Viji, G., Wu, B., Kang, S., Uddin, W., Huff, D., 2001. *Pyricularia grisea* causing gray leaf spot of perennial ryegrass turf: population structure and host specificity. Plant Disease 85, 817-826.

Visentin, I., Tamietti, G., Valentino, D., Portis, E., Karlovsky, P., Moretti, A., Cardinale, F., 2009. The ITS region as a taxonomic discriminator between *Fusarium verticillioides* and *Fusarium proliferatum*. Mycological Research 113, 1137-1145.

Wafula, W.N., Korir, N.K., Ojulong, H.F., Siambi, M., Gweyi-Onyango, J.P., 2018. Protein, Calcium, Zinc, and Iron contents of finger millet grain response to varietal differences and phosphorus application in Kenya. Agronomy 8, 24-30.

Waghunde, R.R., Shelake, R.M., Sabalpara, A.N., 2016. Trichoderma: A significant fungus for agriculture and environment. African Journal of Agricultural Research 11, 1952-1965.

Wakelin, S.A., Gomez-Gallego, M., Jones, E., Smaill, S., Lear, G., Lambie, S., 2018. Climate change induced drought impacts on plant diseases in New Zealand. Australasian Plant Pathology 34, 1-14.

Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A., Zeng, Q., Wortman, J., Young, S.K., 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PloS One 9, e112963.

Walker, D.M., Castlebury, L.A., Rossman, A.Y., White, J., James F., 2012. New molecular markers for fungal phylogenetics: Two genes for species-level systematics in the Sordariomycetes (Ascomycota). Molecular Phylogenetics and Evolution 64, 500-512.

Wang, B., Ebbole, D.J., Wang, Z., 2017a. The arms race between *Magnaporthe oryzae* and rice: Diversity and interaction of AVR and R genes. Journal of Integrative Agriculture 16, 2746-2760.

Wang, Q., Jiang, C., Wang, C., Chen, C., Xu, J., Liu, H., 2017b. Characterization of the two-speed subgenomes of *Fusarium graminearum* reveals the fast-speed subgenome specialized for adaption and infection. Frontiers in Plant Science 8, 140-150.

Wang, X., Jia, Y., Wamishe, Y., Jia, M.H., Valent, B., 2017c. Dynamic changes in the rice blast population in the united states over six decades. Molecular Plant-Microbe Interactions 30, 803-812.

Wang, Y., Coleman-Derr, D., Chen, G., Gu, Y.Q., 2015. OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species. Nucleic Acids Research 43, W78-W84.

Wanyera, N., 2005b. Facilitating the promotion of improved and blast resistant finger millet varieties to enhance production. Finger Millet Blast Management in East Africa. Creating opportunities for improving production and utilization of finger millet, 66-70.

Wanyera, N.M.W., 2005a. Finger Millet (*Eleusine coracana*) (L.) Gaertn) in Uganda. Finger Millet Blast Management in East Africa. Creating opportunities for improving production and utilization of finger millet, 1-9.

Weiner, A.M., Deininger, P.L., Efstratiadis, A., 1986. Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. Annual Review of Biochemistry 55, 631-661.

Werth, C.R., Hilu, K.W., Langner, C.A., 1994. Isozymes of *Eleusine* (Gramineae) and the origin of finger millet. American Journal of Botany 81, 1186-1197.

White, T.J., Bruns, T., Lee, S., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. PCR protocols: A Guide to Methods and Applications 18, 315-322.

Wilson, R.A., Talbot, N.J., 2009. Under pressure: investigating the biology of plant infection by *Magnaporthe oryzae*. Nature Reviews Microbiology 7, 185-195.

Wolf, J.B., Ellegren, H., 2017. Making sense of genomic islands of differentiation in light of speciation. Nature Reviews Genetics 18, 87-90.

Wood, V., Gwilliam, R., Rajandream, M., Lyne, M., Lyne, R., Stewart, A., Sgouros, J., Peat, N., Hayles, J., Baker, S., 2002. The genome sequence of *Schizosaccharomyces pombe*. Nature 415, 871-880.

Wöstemeyer, J., Kreibich, A., 2002. Repetitive DNA elements in fungi (Mycota): impact on genomic architecture and evolution. Current Genetics 41, 189-198.

Wu, J., Kou, Y., Bao, J., Li, Y., Tang, M., Zhu, X., Ponaya, A., Xiao, G., Li, J., Li, C., 2015. Comparative genomics identifies the *Magnaporthe oryzae* avirulence effector *AVRPi9* that triggers *Pi9*-mediated blast resistance in rice. New Phytologist 206, 1463-1475.

Wu, W., Wang, L., Zhang, S., Li, Z., Zhang, Y., Lin, F., Pan, Q., 2014. Stepwise arms race between *AVRPik* and *Pik* alleles in the rice blast pathosystem. Molecular Plant-Microbe Interactions 27, 759-769.

Xu, J., 2016. Fungal DNA barcoding. Genome 59, 913-392.

Xu, J., 2006. Fundamentals of fungal molecular population genetic analyses. Current Issues in Molecular Biology 8, 75-77.

Xue, M., Yang, J., Li, Z., Hu, S., Yao, N., Dean, R.A., Zhao, W., Shen, M., Zhang, H., Li, C., 2012. Comparative analysis of the genomes of two field isolates of the rice blast fungus *Magnaporthe oryzae*. PLoS Genetics 8, e1002869.

Yaegashi, H., Udagawa, S., 1978. The taxonomical identity of the perfect state of *Pyricularia grisea* and its allies. Canadian Journal of Botany 56, 180-183.

Yandell, M., Ence, D., 2012. A beginner's guide to eukaryotic genome annotation. Nature Reviews Genetics 13, 329-330.

Yashoda, H., Palakshappa, M., Kalappanavar, I., Mohankumar, H., Angadi, V., 2000. Chemical control of blast of rice caused by *Magnaporthe grisea* under rainfed direct seeded condition in Karnataka. Karnataka Journal of Agricultural Sciences 13, 372-375.

Yoshida, K., Saunders, D.G., Mitsuoka, C., Natsume, S., Kosugi, S., Saitoh, H., Inoue, Y., Chuma, I., Tosa, Y., Cano, L.M., 2016. Host specialization of the blast fungus *Magnaporthe oryzae* is associated with dynamic gain and loss of genes linked to transposable elements. BMC Genomics 17, 370-371.

Yoshida, K., Saitoh, H., Fujisawa, S., Kanzaki, H., Matsumura, H., Yoshida, K., Tosa, Y., Chuma, I., Takano, Y., Win, J., Kamoun, S., Terauchi, R., 2009. Association genetics reveals three novel avirulence genes from the rice blast fungal pathogen *Magnaporthe oryzae*. Plant Cell 21, 1573-1591.

You, M., Lanoiselet, V., Wang, C., Shivas, R., Li, Y., Barbetti, M., 2012. First report of rice blast (*Magnaporthe oryzae*) on rice (*Oryza sativa*) in Western Australia. Plant Disease 96, 1228-1233.

Yu, G., Smith, D.K., Zhu, H., Guan, Y., Lam, T.T., 2017. GGtree: An R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution 8, 28-36.

Zarandi, M.E., Bonjar, G.S., Dehkaei, F.P., Moosavi, S.A., Farokhi, P.R., Aghighi, S., 2009. Biological control of rice blast (*Magnaporthe oryzae*) by use of *Streptomyces sindeneusis* isolate 263 in greenhouse. American Journal of Applied Sciences 6, 194-199.

Zeigler, R.S., 1998. Recombination in *Magnaporthe grisea*. Annual Review of Phytopathology 36, 249-275.

Zeng, J., Feng, S., Cai, J., Wang, L., Lin, F., Pan, Q., 2009. Distribution of mating type and sexual status in Chinese rice blast populations. Plant Disease 93, 238-242.

Zhang, H., Zheng, X., Zhang, Z., 2016. The *Magnaporthe grisea* species complex and plant pathogenesis. Molecular Plant Pathology 17, 796-804.

Zhang, H., He, D., Kasetsomboon, T., Zhou, H., Li, P., Li, X., Jantasuriyarat, C., Zhou, B., 2013. Analysis of selected singleton transposable elements (SSTEs) and their application for the development of land PATE markers in *Magnaporthe oryzae*. Journal of General Plant Pathology 79, 96-104.

Zhang, S., Wang, L., Wu, W., He, L., Yang, X., Pan, Q., 2015. Function and evolution of *Magnaporthe oryzae* avirulence gene *AVRPib* responding to the rice blast resistance gene *Pib*. Scientific reports 5, 11642-11645.

Zhang, Z., Zhang, X., Zhou, Z., Hu, H., Liu, M., Zhou, B., Zhou, J., 2013. Solution structure of the *Magnaporthe oryzae* avirulence protein *AVRPiz-t*. Journal of Biomolecular NMR 55, 219-223.

Zhang, C., Huang, X., Wang, J., Zhou, M., 2009. Resistance development in rice blast disease caused by *Magnaporthe grisea* to tricyclazole. Pesticide Biochemistry and Physiology 94, 43-47.

Zhao, K., Wright, M., Kimball, J., Eizenga, G., McClung, A., Kovach, M., Tyagi, W., Ali, M.L., Tung, C., Reynolds, A., 2010. Genomic diversity and introgression in *O. sativa* reveal the impact of domestication and breeding on the rice genome. PloS One 5, e10780.

Zhao, Z., Liu, H., Wang, C., Xu, J., 2013. Comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. BMC Genomics 14, 274-276.

Zhong, Z., Chen, M., Lin, L., Han, Y., Bao, J., Tang, W., Lin, L., Lin, Y., Somai, R., Lu, L., 2018. Population genomic analysis of the rice blast fungus reveals specific events associated with expansion of three main clades. Multidisciplinary Journal of Microbial Ecology 12, 1867-1878.

8.0 Appendix

Appendix 1. Diversity and distribution pattern of the *M. oryzae* genotypes identified among the finger millet blast pathogen populations in Kenya, Uganda, Tanzania and Ethiopia based on the HyP1 marker sequences

Genotype	No. of isolates in genotype	Representative isolate	No. of countries present	No. of isolates per country	Isolate code	No. of districts	Population type
HyP1-G1	73	T56	4	25 (E)	**E2, E15, E19, E22, E26, E34, E35, E37, E38, E39, E41, E42, E44, E45, E47, E48, E49, E53, E54, E56, E57, E58, E59, E60, E63**	14	C
				10 (K)	**K14, K21, K41,** K24/127, K33/184, K33/189, K44/111p, K48/115n, K58/128p, K60/131p	6	H & C
				36 (T)	**T1, T2, T3, T6, T7, T8, T9, T10, T13, T18, T19, T22, T24, T26, T27, T28, T30, T31, T32, T33, T34, T36, T38, T41, T42, T43, T44, T45, T46, T47, T48, T49, T50, T51, T52, T56**	6	C
				2 (U)	D3/s9, D10/s73	2	H
HyP1-G2	72	U57	3	7 (E)	**E18, E27, E43, E46, E55, E61, E62**	6	C
				25 (K)	**K1, K4, K5, K7, K8, K13, K15, K22, K27, K42, K43**, K4/21p, K5/24w, K12/62, K21/68n, K22/118, K23/123, K26/76p, K28/82w, K36/98n, K55/124p, K5/23, K15/53n, K9/46, K65/159w	7	H & C
				40 (U)	**U1, U2, U4, U5, U7, U8, U20, U21, U24, U26, U32, U35, U36, U37, U39, U44, U55, U56, U57**, D2/s14, D1/s11, D3/s3, D3/s24, D5/s1, D7/s6, D9/s76, D14/s27, Pen-2-2, D1/s72, D4/s12, D9/s50, D9/s56, D9/s70, D10/s63, D10/s77, D13/s5, D14/s30, D15/s12, D15/s41, D4/s41	20	H & C
HyP1-G3	38	U58	3	9 (K)	**K2, K3, K10, K11, K12, K24, K26, K33, K35**	4	C
				1 (T)	**T35**	1	C
				28 (U)	**U3, U6, U9, U10, U11, U12, U13, U14, U17, U18, U19, U22, U23, U25, U29, U30, U31, U33, U38, U42, U46, U49, U50, U51, U53, U54, U58,** D15/s6	17	H & C

HyP1-G4	8	E33	1	8 (E)	**E21, E23, E24, E25, E28, E29, E31, E33**	5	C
HyP1-G5	5	K57/126p	1	5 (K)	**K34, K36, K39,** K45/112n, K57/126p	2	H & C
HyP1-G6	4	Secn-2-2	2	2 (K)	K13/67, **K28**	2	H & C
				2 (U)	Secn-2-2*, D1/s19	2	H
HyP1-G7	4	P665n-2-1	2	1 (K)	K47/114p	1	H
				3 (U)	E11p-1-1*, Odyp-2-1*, P665n-2-1*	1	H
HyP1-G8	3	T54	1	3 (T)	**T20, T53, T54**	1	C
HyP1-G9	3	U48	2	1 (K)	**K23**	1	C
				2 (U)	**U27, U48**	2	C
HyP1-G10	3	K40	1	3 (K)	**K37, K38, K40**	1	C
HyP1-G11	2	E11	1	2 (E)	**E7, E11**	2	C
HyP1-G12	2	E10	1	2 (E)	**E8, E10**	2	C
HyP1-G13	2	E32	1	2 (E)	**E9, E32**	2	C
HyP1-G14	2	K32	1	2 (K)	**K31, K32**	2	C
HyP1-G15	2	E3	1	2 (E)	**E1, E3**	2	C
HyP1-G16	2	U45	2	1 (K)	**K30**	1	C
				1 (U)	**U45**	1	C
HyP1-G17	2	T40	1	2 (T)	**T39, T40**	1	C
HyP1-G18	2	U52	1	2 (U)	**U41, U52**	2	C
HyP1-G19	2	K65/140n	2	1 (K)	K65/140n	1	H
				1 (U)	D2/s26	1	H
HyP1-G20	2	E51	1	2 (E)	**E50, E51**	2	C
HyP1-G21	2	D15/s37	1	2 (U)	D1/s50, D15/s37	2	H
HyP1-G22	2	T12	1	2 (T)	**T12, T16**	2	C
HyP1-G23	1	E17	1	1	**E17**	1	C
HyP1-G24	1	E40	1	1	**E40**	1	C
HyP1-G25	1	T11	1	1	**T11**	1	C
HyP1-G26	1	U34	1	1	**U34**	1	C
HyP1-G27	1	D8/S15	1	1	D8/S15	1	H
HyP1-G28	1	E12	1	1	**E12**	1	C
HyP1-G29	1	E6	1	1	**E6**	1	C
HyP1-G30	1	E20	1	1	**E20**	1	C
HyP1-G31	1	E36	1	1	**E36**	1	C

HyP1-G32	1	T14	1	1	**T14**	1	C
HyP1-G33	1	T58	1	1	**T58**	1	C
HyP1-G34	1	D6/s1	1	1	**D6/s1**	1	H
HyP1-G35	1	K17	1	1	**K17**	1	C
HyP1-G36	1	D12/s2	1	1	D12/s2	1	H
HyP1-G37	1	T17	1	1	**T17**	1	C
HyP1-G38	1	Gup-2-1	1	1	Gup-2-1	1	H
HyP1-G39	1	K14/74	1	1	K14/74	1	H
HyP1-G40	1	K65/142n	1	1	K65/142n	1	H
HyP1-G41	1	K1/15	1	1	K1/15	1	H
HyP1-G42	1	K18	1	1	**K18**	1	C
HyP1-G43	1	K25	1	1	**K25**	1	C
HyP1-G44	1	U28	1	1	**U28**	1	C
HyP1-G45	1	D1/s53b	1	1	D1/s53b	1	H
HyP1-G46	1	K8/40	1	1	K8/40	1	H
HyP1-G47	1	U15	1	1	**U15**	1	C
HyP1-G48	1	U16	1	1	**U16**	1	C
HyP1-G49	1	U43	1	1	**U43**	1	C
HyP1-G50	1	U47	1	1	**U47**	1	C
HyP1-G51	1	D11/s16	1	1	D11/s16	1	H
HyP1-G52	1	K20	1	1	**K20**	1	C
HyP1-G53	1	K45	1	1	**K45**	1	C
HyP1-G54	1	D4/s26	1	1	D4/s26	1	H
HyP1-G55	1	K9	1	1	**K9**	1	C
HyP1-G56	1	S1p-1-1	1	1	S1p-1-1*****	1	H
HyP1-G57	1	D15/s47	1	1	D15/s47	1	H
HyP1-G58	1	D9/s54	1	1	D9/s54	1	H
HyP1-G59	1	D1/s44	1	1	D1/s44	1	H
HyP1-G60	1	K19	1	1	**K19**	1	C
HyP1-G61	1	K29/164	1	1	K29/164	1	H
HyP1-G62	1	U40	1	1	**U40**	1	C

HyP1-G63	1	K29	1	1	**K29**	1	C
HyP1-G64	1	K44	1	1	**K44**	1	C
HyP1-G65	1	K16	1	1	**K16**	1	C
HyP1-G66	1	K64/137p	1	1	K64/137p	1	H
HyP1-G67	1	T21	1	1	**T21**	1	C
HyP1-G68	1	T37	1	1	**T37**	1	C
HyP1-G69	1	T4	1	1	**T4**	1	C
HyP1-G70	1	T15	1	1	**T15**	1	C
HyP1-G71	1	T23	1	1	**T23**	1	C
HyP1-G72	1	E13	1	1	**E13**	1	C
HyP1-G73	1	E4	1	1	**E4**	1	C
HyP1-G74	1	E30	1	1	**E30**	1	C
HyP1-G75	1	E14	1	1	**E14**	1	C
HyP1-G76	1	E16	1	1	**E16**	1	C
HyP1-G77	1	E52	1	1	**E52**	1	C
HyP1-G78	1	T25	1	1	**T25**	1	C
HyP1-G79	1	T57	1	1	**T57**	1	C
HyP1-G80	1	T5	1	1	**T5**	1	C
HyP1-G81	1	D10/s71	1	1	D10/s71	1	H
HyP1-G82	1	K6	1	1	**K6**	1	C
HyP1-G83	1	T55	1	1	**T55**	1	C
HyP1-G84	1	T29	1	1	**T29**	1	C
HyP1-G85	1	E5	1	1	**E5**	1	C

H, historical population (2000-2004) and C, contemporary population (2015-2017);

In the Isolate code, K-isolates collected from Kenya E-isolates collected from Ethiopia, T-isolates collected from Tanzania and U and D-isolates collected from Uganda; *Isolates from Uganda;

Isolates in bold and normal font are contemporary and historic collections, respectively. Further details of the isolates are available in **Table 2.1A.**

Appendix 2. Diversity and distribution pattern of the *M. oryzae* genotypes identified among the finger millet blast pathogen populations in Kenya, Uganda, Tanzania and Ethiopia based on the HyP2 marker sequences

Genotype	No. of isolates in genotype	Representative isolate	No. of countries present	No. of isolates per country	Isolate code	No. of districts	Population type
HyP2-G1	48	T57	2	35 (E)	**E2, E4, E5, E7, E11, E12, E14, E15, E19, E20, E22, E29, E33, E34, E36, E37, E38, E39, E41, E44, E45, E47, E48, E49, E50, E51, E52, E53, E54, E56, E57, E58, E59, E60, E63**	13	C
				13 (T)	**T1, T2, T3, T5, T8, T10, T22, T25, T47, T49, T51, T55, T57**	5	C
HyP2-G2	36	U36	3	2 (E)	**E46, E62**	2	C
				13 (K)	**K7, K9, K15, K18, K24, K27, K35, K42,** K5/23, K14/74, K47/114p, K60/131p, K65/140n	8	H & C
				21 (U)	**U4, U24, U25, U34, U36,** D6/s1, D2/s14, D2/s26, D3/s3, D5/s1, D9/s50, D9/s54, D9/s56, D9/s76, D10/s71, D13/s5, D14/s30, E11p-1-1*, Odyp-2-1*, Secn-2-2*	12	H & C
HyP2-G3	24	E9	3	3 (E)	**E9, E17, E27**	3	C
				7 (K)	**K1, K2, K3, K10, K11, K12, K33**	4	C
				14 (U)	**U6, U11, U12, U16, U17, U22, U23, U29, U42, U46, U49, U50, U52, U58**	10	C
HyP2-G4	20	T52	2	3 (K)	K33/184, K33/189, K48/115n	2	H
				17 (T)	**T6, T7, T9, T18, T19, T21, T24, T28, T32, T34, T36, T38, T41, T43, T44, T48, T52**	5	C
HyP2-G5	20	U54	2	11 (K)	**K5, K19, K23, K26, K39, K41, K43,** K9/46, K28/82w, K21/68n, K29/164	6	H & C
				9 (U)	**U10, U14, U27, U30, U38, U48, U54,** D1/s44, D14/s27	7	H & C
HyP2-G6	14	U31	3	4 (E)	**E3, E32, E43, E55**	4	C
				3 (K)	**K4, K45,** K1/15	3	H & C
				7 (U)	**U2, U31,** D1/s72, D15/s12, D15/s47, Gup-2-1*, P665n-2-1*	4	H & C

HyP2-G7	12	U45	2	3 (K)	**K30,** K5/24w, K65/159w	2	H & C
				9 (U)	**U15, U28, U37, U43, U44, U45**, D4/s41, D11/s16, D1/s53b,	8	H & C
HyP2-G8	12	U55	2	1 (K)	K26/76p	1	H
				11 (U)	**U3, U7, U8, U18, U26, U33, U39, U53, U55, U20**, D15/s6	9	H & C
HyP2-G9	12	U1	2	10 (E)	**E8, E10, E21, E23, E24, E25, E26, E30, E31, E42**	6	H & C
				2 (U)	**U1**, D9/s70	2	H & C
HyP2-G10	11	D7/s6	2	6 (K)	**K14**, **K25, K29**, K23/123, K36/98n, K65/142n	3	H & C
				5 (U)	**U56, U57**, D7/s6, D8/s15, D15/s41	5	H &C
HyP2-G11	5	T58	3	1 (K)	**K22**	1	C
				3 (T)	**T35, T53, T58**	2	C
				1 (U)	Pen2-2	1	H
HyP2-G12	4	T42	1	4 (T)	**T30, T33, T37, T42**	3	C
HyP2-G13	4	T16	2	1 (K)	K58/128p	1	H
				3 (T)	**T4, T12, T16**	2	C
HyP2-G14	4	U5	2	2 (K)	**K6, K40**	2	C
				2 (U)	**U5**, S1p-1-1*	2	H & C
HyP2-G15	3	K4/21p	2	1 (K)	K4/21p	1	H
				2 (U)	D3/s9, D4/s12	2	H
HyP2-G16	2	E13	1	2 (E)	**E6, E13**	2	C
HyP2-G17	2	E35	1	2 (E)	**E16, E35**	2	C
HyP2-G18	2	K13/37	1	2 (K)	**K13,** K13/67	2	H & C
HyP2-G19	2	T40	1	2 (T)	**T39, T40**	1	C
HyP2-G20	2	U13	1	2 (U)	**U9, U13**	2	C
HyP2-G21	2	D10/s63	1	2 (U)	D1/s50, D10/s63	2	H
HyP2-G22	2	K57/126p	1	2 (K)	K55/124p, K57/126p	1	H
HyP2-G23	1	D10/s73	1	1	D10/s73	1	H
HyP2-G24	1	T29	1	1	**T29**	1	C
HyP2-G25	1	K24/127	1	1	K24/127	1	H
HyP2-G26	1	K44/111p	1	1	K44/111p	1	H
HyP2-G27	1	T31	1	1	**T31**	1	C
HyP2-G28	1	T23	1	1	**T23**	1	C

HyP2-G29	1	T26	1	1	**T26**	1	C
HyP2-G30	1	T15	1	1	**T15**	1	C
HyP2-G31	1	T27	1	1	**T27**	1	C
HyP2-G32	1	T13	1	1	**T13**	1	C
HyP2-G33	1	T46	1	1	**T46**	1	C
HyP2-G34	1	U40	1	1	**U40**	1	C
HyP2-G35	1	T45	1	1	**T45**	1	C
HyP2-G36	1	D10/s77	1	1	D10/s77	1	H
HyP2-G37	1	T11	1	1	**T11**	1	C
HyP2-G38	1	D1/s11	1	1	D1/s11	1	H
HyP2-G39	1	E1	1	1	**E1**	1	C
HyP2-G40	1	K21	1	1	**K21**	1	C
HyP2-G41	1	K31	1	1	**K31**	1	C
HyP2-G42	1	K64/137p	1	1	K64/137p	1	H
HyP2-G43	1	K17	1	1	**K17**	1	C
HyP2-G44	1	K22/118	1	1	K22/118	1	H
HyP2-G45	1	E61	1	1	**E61**	1	C
HyP2-G46	1	D15/s37	1	1	D15/s37	1	H
HyP2-G47	1	D3/s24	1	1	D3/s24	1	H
HyP2-G48	1	K8/40	1	1	K8/40	1	H
HyP2-G49	1	U35	1	1	**U35**	1	C
HyP2-G50	1	D1/s19	1	1	D1/s19	1	H
HyP2-G51	1	E18	1	1	**E18**	1	C
HyP2-G52	1	K36	1	1	**K36**	1	C
HyP2-G53	1	K32	1	1	**K32**	1	C
HyP2-G54	1	K34	1	1	**K34**	1	C
HyP2-G55	1	D4/s26	1	1	D4/s26	1	H
HyP2-G56	1	K8	1	1	**K8**	1	C
HyP2-G57	1	K20	1	1	**K20**	1	C
HyP2-G58	1	U47	1	1	**U47**	1	C
HyP2-G59	1	T20	1	1	**T20**	1	C

HyP2-G60	1	T14	1	1	**T14**	1	C
HyP2-G61	1	T17	1	1	**T17**	1	C
HyP2-G62	1	U51	1	1	**U51**	1	C
HyP2-G63	1	U32	1	1	**U32**	1	C
HyP2-G64	1	K44	1	1	**K44**	1	C
HyP2-G65	1	U41	1	1	**U41**	1	C
HyP2-G66	1	K38	1	1	**K38**	1	C
HyP2-G67	1	K16	1	1	**K16**	1	C
HyP2-G68	1	K28	1	1	**K28**	1	C
HyP2-G69	1	K45/112n	1	1	K45/112n	1	H
HyP2-G70	1	E28	1	1	**E28**	1	C
HyP2-G71	1	T54	1	1	**T54**	1	C
HyP2-G72	1	U19	1	1	**U19**	1	C
HyP2-G73	1	U21	1	1	**U21**	1	C
HyP2-G74	1	E40	1	1	**E40**	1	C
HyP2-G75	1	K37	1	1	**K37**	1	C
HyP2-G76	1	K12/62	1	1	K12/62	1	H
HyP2-G77	1	K15/53n	1	1	K15/53n	1	H
HyP2-G78	1	T50	1	1	**T50**	1	C
HyP2-G79	1	T56	1	1	**T56**	1	C
HyP2-G80	1	D12/s2	1	1	D12/s2	1	H

H, historical population (2000-2004) and C, contemporary population (2015-2017);

In the Isolate code, K-isolates collected from Kenya E-isolates collected from Ethiopia, T-isolates collected from Tanzania and U and D-isolates collected from Uganda; *Isolates from Uganda;

Isolates in bold and normal font are contemporary and historic collections, respectively. Further details of the isolates are available in **Table 2.1A.**

Appendix 3. Diversity and distribution pattern of the *M. oryzae* genotypes identified among the finger millet blast pathogen populations in Kenya, Uganda, Tanzania and Ethiopia based on the multi-locus sequence analysis

Genotype	No. of isolates in genotype	Representative isolate	No. of countries present	No. of isolates per country	Isolate code	No. of districts	Population type
ML-G1	26	T51	2	17 (E)	**E2, E15, E19, E22, E34, E37, E38, E39, E44, E48, E49, E53, E54, E58, E59, E60, E63**	12	C
				9 (T)	**T1, T2, T3, T8, T10, T22, T47, T49, T51**	4	C
ML-G2	15	T52	2	13 (T)	**T6, T7, T9, T18, T19, T24, T28, T32, T41, T43, T44, T48, T52**	4	C
				2 (K)	K33/184, K33/189	1	H
ML-G3	7	U58	2	4 (U)	**U6, U17, U49, U5**8	4	C
				3 (K)	**K10, K12, K33**	2	C
ML-G4	6	U23	2	2 (K)	**K2, K11**	2	C
				4 (U)	**U11, U12, U22, U23**	2	C
ML-G5	5	K15	3	2 (E)	**E46, E62**	2	C
				2 (U)	D5/s1, D9/s76	2	H
				1 (K)	**K15**	1	C
ML-G6	5	U55	1	5 (U)	**U8, U20, U26, U39, U55**	5	C
ML-G7	5	E31	1	5 (E)	**E21, E23, E24, E25, E31**	3	C
ML-G8	4	U24	2	3 (U)	D9/s50, D13/s5, **U24**	3	H & C
ML-G9	4	U56	2	2 (K)	K23/123, K36/98n	2	H
ML-G10	3	K4	2	1 (K)	**K4**	1	C
ML-G11	3	U48	2	1 (K)	**K23**	1	C
				2 (U)	**U27, U48**	2	C
ML-G12	3	U36	2	1 (K)	K5/23	1	H
				2 (U)	D14/s30, **U36**	2	H & C
ML-G13	3	Odyp-2-1	2	1 (K)	K47/114p	1	H
				2 (U)	E11p-1-1*, Odyp-2-1*	1	H
ML-G14	3	U54	1	3 (U)	**U14, U30, U54**	2	C
ML-G15	3	U33	1	3 (U)	D15/s6, **U18, U33**	3	H & C
ML-G16	2	T16	1	2 (T)	**T12, T16**	2	C
ML-G17	2	T33	1	2 (T)	**T30, T33**	1	C

ML-G18	2	T38	1	2 (T)	**T36, T38**	1	C
ML-G19	2	E33	1	2 (E)	**E29, E33**	1	C
ML-G20	2	E51	1	2 (E)	**E50, E51**	2	C
ML-G21	2	E57	1	2 (E)	**E56, E57**	1	C
ML-G22	2	K28/28w	1	2 (K)	K21/68n, K28/82w	1	H
ML-G23	2	K65/140n	2	1 (U)	D2/s26	1	H
				1 (K)	K65/140n	1	H
ML-G24	2	U25	2	1 (U)	**U25**	1	C
				1 (K)	**K24**	1	C
ML-G25	2	K42	2	1 (U)	D3/s3	1	H
				1 (K)	**K42**	1	C
ML-G26	2	U45	2	1 (U)	**U45**	1	C
				1 (K)	**K30**	1	C
ML-G27	2	U37	2	1 (U)	**U37**	1	C
				1 (K)	K65/159w	1	H
ML-G28	2	U10	2	1 (U)	**U10**	1	C
				1 (K)	**K26**	1	C
ML-G29	1	D1/s50	1	1	D1/s50	1	H
ML-G30	1	D2/s14	1	1	D2/s14	1	H
ML-G31	1	D1/s44	1	1	D1/s44	1	H
ML-G32	1	D1/s53b	1	1	D1/s53b	1	H
ML-G33	1	D1/s11	1	1	D1/s11	1	H
ML-G34	1	D1/s19	1	1	D1/s19	1	H
ML-G35	1	D3/s9	1	1	D3/s9	1	H
ML-G36	1	D3/s24	1	1	D3/s24	1	H
ML-G37	1	D4/s26	1	1	D4/s26	1	H
ML-G38	1	D4/s41	1	1	D4/s41	1	H
ML-G39	1	D6/s1	1	1	D6/s1	1	H
ML-G40	1	D4/s12	1	1	D4/s12	1	H
ML-G41	1	D9/s70	1	1	D9/s70	1	H
ML-G42	1	D10/s77	1	1	D10/s77	1	H
ML-G43	1	D11/s16	1	1	D11/s16	1	H

ML-G44	1	D8/s15	1	1	D8/s15	1	H
ML-G45	1	D12/s2	1	1	D12/s2	1	H
ML-G46	1	D9/s56	1	1	D9/s56	1	H
ML-G47	1	D9/s54	1	1	D9/s54	1	H
ML-G48	1	D10/s71	1	1	D10/s71	1	H
ML-G49	1	D10/s63	1	1	D10/s63	1	H
ML-G50	1	D10/s73	1	1	D10/s73	1	H
ML-G51	1	D14/s27	1	1	D14/s27	1	H
ML-G52	1	D15/s41	1	1	D15/s41	1	H
ML-G53	1	D15/s37	1	1	D15/s37	1	H
ML-G54	1	D15/s47	1	1	D15/s47	1	H
ML-G55	1	Gup-2-1	1	1	Gup-2-1*****	1	H
ML-G56	1	P665n-2-1	1	1	P665n-2-1*****	1	H
ML-G57	1	Pen-2-2	1	1	Pen-2-2	1	H
ML-G58	1	S1p-1-1	1	1	S1p-1-1*****	1	H
ML -G59	1	Secn-2-2	1	1	Secn-2-2*****	1	H
ML -G60	1	U1	1	1	**U1**	1	C
ML -G61	1	U2	1	1	**U2**	1	C
ML -G62	1	U3	1	1	**U3**	1	C
ML -G63	1	U4	1	1	**U4**	1	C
ML -G64	1	U5	1	1	**U5**	1	C
ML -G65	1	U7	1	1	**U7**	1	C
ML -G66	1	U9	1	1	**U9**	1	C
ML -G67	1	U13	1	1	**U13**	1	C
ML -G68	1	U15	1	1	**U15**	1	C
ML -G69	1	U16	1	1	**U16**	1	C
ML -G70	1	U19	1	1	**U19**	1	C
ML -G71	1	U21	1	1	**U21**	1	C
ML -G72	1	U28	1	1	**U28**	1	C
ML -G73	1	U29	1	1	**U29**	1	C
ML -G74	1	U31	1	1	**U31**	1	C

ML -G75	1	U32	1	1	**U32**	1	C
ML -G76	1	U34	1	1	**U34**	1	C
ML -G77	1	U35	1	1	**U35**	1	C
ML -G78	1	U38	1	1	**U38**	1	C
ML -G79	1	U40	1	1	**U40**	1	C
ML -G80	1	U41	1	1	**U41**	1	C
ML -G81	1	U42	1	1	**U42**	1	C
ML -G82	1	U43	1	1	**U43**	1	C
ML -G83	1	U44	1	1	**U44**	1	C
ML -G84	1	U46	1	1	**U46**	1	C
ML -G85	1	U47	1	1	**U47**	1	C
ML -G86	1	U50	1	1	**U50**	1	C
ML -G87	1	U51	1	1	**U51**	1	C
ML -G88	1	U52	1	1	**U52**	1	C
ML-G89	1	U53	1	1	**U53**	1	C
ML-G90	1	U57	1	1	**U57**	1	C
ML-G91	1	E1	1	1	**E1**	1	C
ML-G92	1	E3	1	1	**E3**	1	C
ML-G93	1	E4	1	1	**E4**	1	C
ML-G94	1	E5	1	1	**E5**	1	C
ML-G95	1	E6	1	1	**E6**	1	C
ML-G96	1	E7	1	1	**E7**	1	C
ML-G97	1	E8	1	1	**E8**	1	C
ML-G98	1	E9	1	1	**E9**	1	C
ML-G99	1	E10	1	1	**E10**	1	C
ML-G100	1	E11	1	1	**E11**	1	C
ML-G101	1	E12	1	1	**E12**	1	C
ML-G102	1	E13	1	1	**E13**	1	C
ML-G103	1	E14	1	1	**E14**	1	C
ML-G104	1	E16	1	1	**E16**	1	C
ML-G105	1	E17	1	1	**E17**	1	C

ML-G106	1	E18	1	1	**E18**	1	C
ML-G107	1	E20	1	1	**E20**	1	C
ML-G108	1	E26	1	1	**E26**	1	C
ML-G109	1	E27	1	1	**E27**	1	C
ML-G110	1	E28	1	1	**E28**	1	C
ML-G111	1	E30	1	1	**E30**	1	C
ML-G112	1	E32	1	1	**E32**	1	C
ML-G113	1	E35	1	1	**E35**	1	C
ML-G114	1	E36	1	1	**E36**	1	C
ML-G115	1	E40	1	1	**E40**	1	C
ML-G116	1	E41	1	1	**E41**	1	C
ML-G117	1	E42	1	1	**E42**	1	C
ML-G118	1	E43	1	1	**E43**	1	C
ML-G119	1	E45	1	1	**E45**	1	C
ML-G120	1	E47	1	1	**E47**	1	C
ML-G121	1	E52	1	1	**E52**	1	C
ML-G122	1	E55	1	1	**E55**	1	C
ML-G123	1	E61	1	1	**E61**	1	C
ML-G124	1	K1	1	1	**K1**	1	C
ML-G125	1	K1/15	1	1	K1/15	1	H
ML-G126	1	K3	1	1	**K3**	1	C
ML-G127	1	K4/21p	1	1	K4/21p	1	H
ML-G128	1	K5	1	1	**K5**	1	C
ML-G129	1	K5/24w	1	1	K5/24w	1	H
ML-G130	1	K7	1	1	**K7**	1	C
ML-G131	1	K6	1	1	**K6**	1	C
ML-G132	1	K8	1	1	**K8**	1	C
ML-G133	1	K8/40	1	1	K8/40	1	H
ML-G134	1	K9	1	1	**K9**	1	C
ML-G135	1	K9/46	1	1	K9/46	1	H
ML-G136	1	K12/62	1	1	K12/62	1	H

ML-G137	1	K13	1	1	**K13**	1	C
ML-G138	1	K13/67	1	1	K13/67	1	H
ML-G139	1	K14	1	1	**K14**	1	C
ML-G140	1	K14/74	1	1	K14/74	1	H
ML-G141	1	K15/53n	1	1	K15/53n	1	H
ML-G142	1	K16	1	1	**K16**	1	C
ML-G143	1	K17	1	1	**K17**	1	C
ML-G144	1	K18	1	1	**K18**	1	C
ML-G145	1	K19	1	1	**K19**	1	C
ML-G146	1	K20	1	1	**K20**	1	C
ML-G147	1	K21	1	1	**K21**	1	C
ML-G148	1	K22	1	1	**K22**	1	C
ML-G149	1	K22/118	1	1	K22/118	1	H
ML-G150	1	K24/127	1	1	K24/127	1	H
ML-G151	1	K25	1	1	**K25**	1	C
ML-G152	1	K26/76p	1	1	K26/76p	1	H
ML-G153	1	K28	1	1	**K28**	1	C
ML-G154	1	K29	1	1	**K29**	1	C
ML-G155	1	K29/164	1	1	K29/164	1	H
ML-G156	1	K31	1	1	**K31**	1	C
ML-G157	1	K32	1	1	**K32**	1	C
ML-G158	1	K34	1	1	**K34**	1	C
ML-G159	1	K35	1	1	**K35**	1	C
ML-G160	1	K36	1	1	**K36**	1	C
ML-G161	1	K37	1	1	**K37**	1	C
ML-G162	1	K38	1	1	**K38**	1	C
ML-G163	1	K39	1	1	**K39**	1	C
ML-G164	1	K40	1	1	**K40**	1	C
ML-G165	1	K41	1	1	**K41**	1	C
ML-G166	1	K43	1	1	**K43**	1	C
ML-G167	1	K44	1	1	**K44**	1	C

ML-G168	1	K44/111p	1	1	K44/111p	1	H
ML-G169	1	K45	1	1	**K45**	1	C
ML-G170	1	K55/124p	1	1	K55/124p	1	H
ML-G171	1	K65/142n	1	1	K65/142n	1	H
ML-G172	1	K45/112n	1	1	K45/112n	1	H
ML-G173	1	K48/115n	1	1	K48/115n	1	H
ML-G174	1	K57/126p	1	1	K57/126p	1	H
ML-G175	1	K58/128p	1	1	K58/128p	1	H
ML-G176	1	K60/131p	1	1	K60/131p	1	H
ML-G177	1	K64/137p	1	1	K64/137p	1	H
ML-G178	1	T4	1	1	**T4**	1	C
ML-G179	1	T5	1	1	**T5**	1	C
ML-G180	1	T11	1	1	**T11**	1	C
ML-G181	1	T13	1	1	**T13**	1	C
ML-G182	1	T14	1	1	**T14**	1	C
ML-G183	1	T15	1	1	**T15**	1	C
ML-G184	1	T17	1	1	**T17**	1	C
ML-G185	1	T20	1	1	**T20**	1	C
ML-G186	1	T21	1	1	**T21**	1	C
ML-G187	1	T23	1	1	**T23**	1	C
ML-G188	1	T25	1	1	**T25**	1	C
ML-G189	1	T26	1	1	**T26**	1	C
ML-G190	1	T27	1	1	**T27**	1	C
ML-G191	1	T29	1	1	**T29**	1	C
ML-G192	1	T31	1	1	**T31**	1	C
ML-G193	1	T34	1	1	**T34**	1	C
ML-G194	1	T35	1	1	**T35**	1	C
ML-G195	1	T37	1	1	**T37**	1	C
ML-G196	1	T39	1	1	**T39**	1	C
ML-G197	1	T40	1	1	**T40**	1	C
ML-G198	1	T42	1	1	**T42**	1	C

ML-G199	1	T45	1	1	**T45**	1	C
ML-G200	1	T46	1	1	**T46**	1	C
ML-G201	1	T50	1	1	**T50**	1	C
ML-G202	1	T53	1	1	**T53**	1	C
ML-G203	1	T54	1	1	**T54**	1	C
ML-G204	1	T55	1	1	**T55**	1	C
ML-G205	1	T56	1	1	**T56**	1	C
ML-G206	1	T57	1	1	**T57**	1	C
ML-G207	1	T58	1	1	**T58**	1	C

H, historical population (2000-2004) and C, contemporary population (2015-2017);

In the Isolate code, K-isolates collected from Kenya E-isolates collected from Ethiopia, T-isolates collected from Tanzania and U and D-isolates collected from Uganda; *Isolates from Uganda;

Isolates in bold and normal font are contemporary and historic collections, respectively. Further details of the isolates are available in **Table 2.1A.**

Appendix 4: Diversity and distribution pattern of the *M. oryzae* genotypes (shared and single) identified among the isolates collected from finger millet in various districts of Ethiopia based on the multi-locus sequence analysis

S/N	Genotype[a]	Isolate	District	No. of District	Plant part	Finger millet variety	Year of collection	Grasshopper[b]	Mating type[c]
1	ET-G1 (17)	E2	Wayu Tuka	11	Neck	Landrace	2015	+	1-1
2		E19	Wayu Tuka		Neck	NK	2016	+	1-1
3		E15	Bure		Head	Landrace	2015	+	1-1
4		E59	Bure		Head	Landrace	2015	+	1-2
5		E22	Nedjo		Neck	NK	2016	+	1-1
6		E34	Diga		Neck	NK	2016	+	1-1
7		E37	Diga		Neck	NK	2016	+	1-1
8		E38	Banja		Head	Landrace	2015	+	1-1
9		E39	Banja		Neck	Landrace	2015	+	1-1
10		E49	Banja		Head	Landrace	2015	+	1-1
11		E53	Mecha		Head	Landrace	2015	+	1-1
12		E54	Mecha		Neck	Landrace	2015	+	1-2
13		E48	Guangau		Neck	Landrace	2015	+	1-1
14		E60	Ankussa-Abdo Gor		Head	Landrace	2015	+	1-1
15		E58	Bahir Dar Zuria		Neck	Landrace	2015	+	1-2
16		E44	Dure Bete		Neck	Landrace	2015	+	1-1
17		E63	Qilxxu Kara		Head	Landrace	2015	+	1-1
18	ET-G2 (5)	E21	Wayu Tuka	3	Neck	NK	2016	+	1-1
19		E23	Nedjo		Neck	NK	2016	+	1-1
20		E24	Nedjo		Neck	NK	2016	+	1-1
21		E25	Nedjo		Neck	NK	2016	+	1-1
22		E31	Gimbi		Head	NK	2016	+	1-1
23	ET-G3 (2)	E29	Diga	1	Neck	NK	2016	+	1-1
24		E33	Diga		Head	NK	2016	+	1-1
25	ETG-4	E46	Mandura	2	Neck	Landrace	2015	-	1-2

26	(2)	E62	Jabi Tana		Neck	Landrace	2015	-	1-2
27	ET-G5	E50	Banja	2	Neck	Landrace	2015	+	1-1
28	(2)	E51	Mecha		Head	Landrace	2015	+	1-1
29	ET-G6 (2)	E56	Bahir Dar Zuria	1	Head	Landrace	2015	+	1-2
30		E57	Bahir Dar Zuria		Neck	Landrace	2015	+	1-2
31	ET-G7	E1	Sire	1	Neck	Landrace	2015	-	1-1
32	ET-G8	E3	Adet	1	Neck	Landrace	2015	+	1-2
33	ET-G9	E4	Diga	1	Head	Landrace	2015	+	1-1
34	ET-G10	E5	Diga	1	Neck	Landrace	2015	+	1-1
35	ET-G11	E6	Diga	1	Head	Landrace	2015	+	1-1
36	ET-G12	E7	Diga	1	Neck	Landrace	2015	+	1-1
37	ET-G13	E8	Lallo Assabi	1	Head	Landrace	2015	+	1-1
38	ET-G14	E9	Lallo Assabi	1	Neck	Landrace	2015	+	1-1
39	ET-G15	E10	Nedjo	1	Head	Landrace	2015	+	1-1
40	ET-G16	E11	Guangau	1	Neck	Landrace	2015	+	1-1
41	ET-G17	E12	Banja	1	Head	Landrace	2015	+	1-2
42	ET-G18	E13	Bahir Dar Zuria	1	Head	Landrace	2015	+	1-2
43	ET-G19	E14	Bahir Dar Zuria	1	Neck	Landrace	2015	+	1-2
44	ET-G20	E16	Demecha	1	Neck	Landrace	2015	+	1-2
45	ET-G21	E17	Angebo	1	Neck	Landrace	2016	+	1-1
46	ET-G22	E18	Wayu Tuka	1	Neck	NK	2016	+	1-1
47	ET-G23	E20	Diga	1	Neck	NK	2016	+	1-1
48	ET-G24	E26	Leta Sibu	1	Head	NK	2016	+	1-1
49	ET-G25	E27	Bila	1	Neck	NK	2016	+	1-1
50	ET-G26	E28	Boji Bermeji	1	Neck	NK	2016	+	1-1
51	ET-G27	E30	Lallo Asabi	1	Neck	NK	2016	+	1-1
52	ET-G28	E32	Diga	1	Neck	NK	2016	-	1-2

53	ET-G29	E35	Leka dulecha	1	Neck	NK	2016	+	1-2
54	ET-G30	E36	Leka dulecha	1	Head	NK	2016	+	1-1
55	ET-G31	E40	Dangla	1	Neck	Landrace	2015	+	1-1
56	ET-G32	E41	Dure Bete	1	Head	Landrace	2015	+	1-2
57	ET-G33	E42	Dure Bete	1	Neck	Landrace	2015	+	1-1
58	ET-G34	E43	Dure Bete	1	Head	Landrace	2015	-	1-2
59	ET-G35	E45	Leka dulecha	1	Head	NK	2015	+	1-2
60	ET-G36	E47	Guangau	1	Head	Landrace	2015	+	1-2
61	ET-G37	E52	Mecha	1	Neck	Landrace	2015	+	1-2
62	ET-G38	E55	Bahir Dar Zuria	1	Neck	Landrace	2015	-	1-2
63	ET-G39	E61	Jabi Tana	1	Head	Landrace	2015	-	1-2

[a]Genotypes identified based on the multi-locus sequence data of HIS4, ITS, HyP2 and HyP1

[b]*M. oryzae* isolates identified as grh positive (+) or grh negative (-) based on the presence or absence of the two amplicons in PCR screening with the PKE and PES primer pairs

[c]Mating type was determined using PCR assay and the *M. oryzae* isolates were characterised as MAT 1-1 or MAT 1-2

Numbers in parentheses are the number of isolates represented by a genotype

NK: Information not available

Appendix 5: Diversity and distribution pattern of the *M. oryzae* genotypes (shared and single) identified among the isolates collected from finger millet in various districts of Tanzania based on the multi-locus sequence analysis

S/N	Genotype[a]	Isolate	District	No. of District	Plant part	Finger millet variety	Year of collection	Grasshopper[b]	Mating type[c]
1	TZ-G1 (13)	T6	Njombe	4	Neck	Landrace	2016	+	1-1
2		T7	Njombe		Neck	Landrace	2016	+	1-2
3		T9	Njombe		Neck	Landrace	2016	+	1-2
4		T18	Njombe		Neck	Landrace	2016	+	1-1
5		T48	Njombe		Neck	Landrace	2016	+	1-2
6		T52	Njombe		Head	Landrace	2016	+	1-2
7		T19	Mbozi		Neck	Landrace	2016	+	1-1
8		T41	Mbozi		Neck	Landrace	2016	+	1-1
9		T43	Mbozi		Neck	Landrace	2016	+	1-1
10		T44	Mbozi		Neck	Landrace	2016	+	1-1
11		T24	Nkasi		Neck	Landrace	2016	+	1-1
12		T28	Nkasi		Neck	Landrace	2016	+	1-1
13		T32	Sumbawanga		Neck	Landrace	2016	+	1-1
14	TZ-G2 (9)	T1	Nkasi	4	Neck	Landrace	2015	+	1-2
15		T2	Nkasi		Neck	Landrace	2016	+	1-2
16		T3	Nkasi		Neck	Landrace	2016	+	1-2
17		T8	Njombe		Neck	Landrace	2016	+	1-2
18		T47	Njombe		Neck	Landrace	2016	+	1-2
19		T49	Njombe		Neck	Landrace	2016	+	1-1
20		T51	Njombe		Neck	Landrace	2016	+	1-2
21		T10	Madaba		Neck	Landrace	2016	+	1-2
22		T22	Momba		Neck	Landrace	2016	+	1-2
23	TZ-G3 (2)	T30	Nkasi	1	Neck	Landrace	2016	+	1-2
24		T33	Nkasi		Neck	Landrace	2016	+	1-2
25	TZ-G4 (2)	T36	Momba	1	Neck	Landrace	2016	+	1-1
26		T38	Momba		Neck	Landrace	2016	+	1-2
27	TZ-G5 (2)	T12	Sumbawanga	2	Neck	Landrace	2016	+	1-2
28		T16	Njombe		Neck	Landrace	2016	+	1-2

29	TZ-G6	T4	Sumbawanga	1	Neck	Landrace	2016	+	1-1
30	TZ-G7	T5	Mbozi	1	Neck	Landrace	2016	+	1-1
31	TZ-G8	T11	Madaba	1	Neck	Landrace	2016	+	1-2
32	TZ-G9	T13	Momba	1	Neck	Landrace	2016	+	1-1
33	TZ-G10	T14	Madaba	1	Head	Landrace	2016	-	1-2
34	TZ-G11	T15	Madaba	1	Head	Landrace	2016	+	1-2
35	TZ-G12	T17	Madaba	1	Neck	Landrace	2016	-	1-2
36	TZ-G13	T20	Njombe	1	Neck	Landrace	2016	-	1-2
37	TZ-G14	T21	Mbozi	1	Neck	Landrace	2016	+	1-1
38	TZ-G15	T23	Momba	1	Neck	Landrace	2016	+	1-2
39	TZ-G16	T25	Mbozi	1	Neck	Landrace	2016	+	1-2
40	TZ-G17	T26	Nkasi	1	Neck	Landrace	2016	+	1-1
41	TZ-G18	T27	Nkasi	1	Neck	Landrace	2016	+	1-2
42	TZ-G19	T29	Nkasi	1	Neck	Landrace	2016	+	1-1
43	TZ-G20	T31	Sumbawanga	1	Neck	Landrace	2016	+	1-1
44	TZ-G21	T34	Sumbawanga	1	Neck	Landrace	2016	+	1-2
45	TZ-G22	T35	Madaba	1	Neck	Landrace	2016	-	1-2
46	TZ-G23	T37	Momba	1	Neck	Landrace	2016	+	1-2
47	TZ-G24	T39	Mbozi	1	Neck	Landrace	2016	-	1-1
48	TZ-G25	T40	Mbozi	1	Neck	Landrace	2016	-	1-1
49	TZ-G26	T42	Mbozi	1	Neck	Landrace	2016	+	1-1
50	TZ-G27	T45	Njombe	1	Neck	Landrace	2016	+	1-2
51	TZ-G28	T46	Njombe	1	Neck	Landrace	2016	+	1-2
52	TZ-G29	T50	Njombe	1	Neck	Landrace	2016	+	1-2
53	TZ-G30	T53	Njombe	1	Neck	Landrace	2016	-	1-1
54	TZ-G31	T54	Njombe	1	Head	Landrace	2016	-	1-2
55	TZ-G32	T55	Njombe	1	Head	Landrace	2016	+	1-2
56	TZ-G33	T56	Njombe	1	Neck	Landrace	2016	+	1-1
57	TZ-G34	T57	Njombe	1	Head	Landrace	2016	+	1-2
58	TZ-G35	T58	Njombe	1	Head	Landrace	2016	-	1-2

[a]Genotypes identified based on the multi-locus sequence data of HIS4, ITS, HyP2 and HyP1

[b] *M. oryzae* isolates identified as grh positive (+) or grh negative (-) based on the presence or absence of the two amplicons in PCR screening with the PKE and PES primer pairs

[c]Mating type was determined using PCR assay and the *M. oryzae* isolates were characterised as MAT 1-1 or MAT 1-2

Numbers in parentheses are the number of isolates represented by a genotype

Appendix 6: Diversity and distribution pattern of the *M. oryzae* genotypes (shared and single) identified among the isolates collected from finger millet in various districts of Kenya based on the multi-locus sequence analysis

S/N	Genotype[a]	Isolate	District	No. of Districts	Plant part	Finger millet variety	Year of collection	Grasshopper[b]	Mating type[c]
1	KN-G1 (3)	K10	Busia	2	Neck	Improved	2016	-	1-1
2		K12	Busia		Neck	Improved	2016	-	1-1
3		K33	Kisii		Neck	NK	2017	-	1-1
4	KN-G2 (2)	K2	Kisumu	2	Neck	Improved	2016	-	1-1
5		K11	Busia		Neck	Improved	2016	-	1-1
6	KN-G3 (2)	K23/123	Busia	2	Neck	NK	2000	NI	NI
7		K36/98n	Teso		Neck	NK	2002	NI	NI
8	KN-G4 (2)	K33/184	Kericho	1	Panicle	NK	2000	NI	NI
9		K33/189	Kericho		Leaf	NK	2000	NI	NI
10	KN-G5 (2)	K21/68	Teso	1	Neck	NK	2002	NI	NI
11		K28/82w	Teso		NK	NK	2002	NI	NI
12	KN-G6	K58/128p	Gucha	1	Panicle	NK	2002	NI	NI
13	KN-G7	K24/127	Kisii central	1	Panicle	NK	2000	NI	NI
14	KN-G8	K44/111p	Kisii	1	Panicle	NK	2002	NI	NI
15	KN-G9	K6	Siaya	1	Neck	Improved	2016	-	1-2
16	KN-G10	K48/115n	Kisii	1	Neck	NK	2002	NI	NI
17	KN-G11	K21	Marakwet	1	Head	Landrace	2016	-	1-1
18	KN-G12	K41	Kisii	1	Head	NK	2017	-	1-1
19	KN-G13	K60/131p	Gucha	1	Panicle	NK	2002	NI	NI
20	KN-G14	K14	Bungoma	1	Neck	Improved	2016	-	1-2
21	KN-G15	K34	Kisii	1	Neck	NK	2017	-	1-1
22	KN-G16	K39	Kisii	1	Neck	NK	2017	-	1-1
23	KN-G17	K36	Kisii	1	Neck	NK	2017	-	1-1
24	KN-G18	K45/112n	Kisii	1	Neck	NK	2002	NI	NI
25	KN-G19	K57/126p	Gucha	1	Panicle	NK	2002	NI	NI
26	KN-G20	K64/137p	Homabay	1	Panicle	NK	2002	NI	NI

27	KN-G21	K16	Kisii	1	Neck	Improved	2016	-	1-2
28	KN-G22	K44	Kisii	1	Neck	NK	2017	-	1-2
29	KN-G23	K38	Kisii	1	Neck	NK	2017	+	1-1
30	KN-G24	K37	Kisii	1	Neck	NK	2017	+	1-2
31	KN-G25	K40	Kisii	1	Neck	NK	2017	+	1-1
32	KN-G26	K29	Busia	1	Neck	NK	2017	-	1-1
33	KN-G27	K12/62	Teso	1	Neck	NK	2000	NI	NI
34	KN-G28	K31	Busia	1	Neck	NK	2017	-	1-1
35	KN-G29	K17	Kakamega	1	Neck	Weed	2016	-	1-2
36	KN-G30	K25	Busia	1	Neck	NK	2017	-	1-1
37	KN-G31	K14/74	Teso	1	Neck	NK	2000	NI	NI
38	KN-G32	K47/114p	Kisii	1	Panicle	NK	2002	NI	NI
39	KN-G33	K15	Kimusu	1	Neck	Improved	2016	-	1-1
40	KN-G34	K7	Siaya	1	Neck	Improved	2016	-	1-2
41	KN-G35	K18	Eldoret	1	Neck	Landrace	2016	-	1-1
42	KN-G36	K22/118	Busia	1	Panicle	NK	2000	NI	NI
43	KN-G37	K65/142n	Teso	1	Neck	NK	2002	NI	NI
44	KN-G38	K42	Kisii	1	Neck	NK	2017	-	1-2
45	KN-G39	K55/124p	Gucha	1	Panicle	NK	2002	NI	NI
46	KN-G40	K5/23	Teso	1	Neck	NK	2000	NI	NI
47	KN-G41	K27	Busia	1	Neck	NK	2017	-	1-2
48	KN-G42	K9	Kisumu	1	Neck	Improved	2016	-	1-2
49	KN-G43	K8/40	Busia	1	Panicle	NK	2000	NI	NI
50	KN-G44	K65/140n	Alupe/Teso	1	Neck	NK	2002	NI	NI
51	KN-G45	K24	Busia	1	Neck	NK	2017	-	1-2
52	KN-G46	K35	Kisii	1	Neck	NK	2017	-	1-2
53	KN-G47	K1/15	Teso	1	Panicle	NK	2000	NI	NI
54	KN-G48	K45	Kisii	1	Neck	NK	2017	-	1-2
55	KN-G49	K15/53n	Teso	1	Neck	NK	2002	NI	NI
56	KN-G50	K4	Kisumu	1	Head	Improved	2016	+	1-1
57	KN-G51	K22	Kisumu	1	Neck	Improved	2016	-	1-1

58	KN-G52	K4/21p	Busia	1	Panicle	NK	2002	NI	NI
59	KN-G53	K26/76p	Teso	1	Panicle	NK	2002	NI	NI
60	KN-G54	K65/159w	Alupe/Teso	1	NK	NK	2002	NI	NI
61	KN-G55	K20	Eldoret	1	Neck	Improved	2016	-	1-2
62	KN-G56	K30	Busia	1	Neck	NK	2017	-	1-2
63	KN-G57	K5/24w	Busia	1	NK	NK	2002	NI	NI
64	KN-G58	K8	Siaya	1	Neck	Improved	2016	+	1-2
65	KN-G59	K13	Busia	1	Neck	Weed	2016	-	1-2
66	KN-G60	K13/67	Teso	1	Neck	NK	2000	NI	NI
67	KN-G61	K1	Siaya	1	Neck	Improved	2016	-	1-2
68	KN-G62	K3	Siaya	1	Neck	Improved	2016	-	1-1
69	KN-G63	K26	Busia	1	Neck	NK	2017	-	1-2
70	KN-G64	K9/46	Busia	1	Panicle	NK	2000	NI	NI
71	KN-G65	K23	Busia	1	Neck	NK	2017	-	1-1
72	KN-G66	K5	Kisumu	1	Neck	Improved	2016	-	1-1
73	KN-G67	K43	Kisii	1	Head	NK	2017	-	1-1
74	KN-G68	K32	Bungoma	1	Neck	NK	2017	-	1-2
75	KN-G69	K19	Eldoret	1	Neck	Landrace	2016	-	1-1
76	KN-G70	K29/164	Suba	1	Panicle	NK	2000	NI	NI
77	KN-G71	K28	Busia	1	Neck	NK	2017	-	1-1

[a]Genotypes identified based on the multi-locus sequence data of HIS4, ITS, HyP2 and HyP1

[b]*M. oryzae* isolates identified as grh positive (+) or grh negative (-) based on the presence or absence of the two amplicons in PCR screening with the PKE and PES primer pairs

[c]Mating type was determined using PCR assay and the *M. oryzae* isolates were characterised as MAT 1-1 or MAT 1-2

Numbers in parentheses are the number of isolates represented by a genotype NK: Information not available

NI: Isolates not included in the Grasshopper and mating type PCR analyses as theses have previously been characterised (Takan *et al*., 2012)

Appendix 7: Diversity and distribution pattern of the *M. oryzae* genotypes (shared and single) identified among the isolates collected from finger millet in various districts of Uganda based on the multi-locus sequence data analysis

S/N	Genotype[a]	Isolate	District	No. of District	Plant part	Finger millet variety	Year of collection	Grasshopper[b]	Mating type[c]
1	UG-G1 (5)	U8	Kumi	5	Neck	Local	2016	-	1-1
2		U20	Alebtong		Neck	NK	2016	-	1-2
3		U26	Apac		Neck	Local	2016	-	1-2
4		U39	Tororo		Neck	Local	2016	-	1-2
5		U55	Lamwo		Neck	Improved	2016	-	1-1
6	UG-G2 (4)	U6	Serere	4	Neck	Local	2016	-	1-1
7		U17	Ngora		Neck	Local	2016	-	1-1
8		U49	Kumi		Head	NK	2016	-	1-1
9		U58	Mbale		Neck	NK	2016	-	1-1
10	UG-G3 (4)	U11	Pallisa	2	Neck	Improved	2016	-	1-1
11		U12	Pallisa		Neck	NK	2016	-	1-1
12		U22	Lira		Neck	NK	2016	-	1-1
13		U28	Lira		Neck	Local	2016	-	1-2
14	UG-G4 (3)	D9/s50	Lira	3	Panicle	NK	2002	NI	NI
15		D13/s5	Katakwi		Panicle	NK	2002	NI	NI
16		U24	Apac		Neck	Local	2016	-	1-1
17	UG-G5 (3)	U14	Katakwi	2	Neck	Local	2016	-	1-1
18		U30	Amuru		Neck	Local	2016	-	1-1
19		U54	Amuru		Neck	Local	2016	-	1-1
20	UG-G6 (3)	D15/s6	Soroti	3	Panicle	NK	2000	NI	NI
21		U13	Katakwi		Neck	Local	2016	-	1-1
22		U18	Ngora		Neck	Local	2016	-	1-1
23	UG-G7 (2)	D5/s1	Iganga	2	Panicle	NK	2000	NI	NI
24		D9/s76	Lira		Panicle	NK	2002	NI	NI
25	UG-G8 (2)	E11p-1-1	Soroti	1	Panicle	NK	2004	NI	NI
26		Odyp-2-1	Soroti		Panicle	NK	2004	NI	NI
27	UG-G9	D14/s30	Kabermaido	2	Leaf	NK	2002	NI	NI

28	(2)	U36	Manafwa		Neck	Local	2016	-	1-2
29	UG-G10	D7/s6	Kamuli	2	Panicle	NK	2000	NI	NI
30	UG-G10	U56	Manafwa		Neck	Local	2016	-	1-2
31	UG-G11	U27	Gulu	2	Neck	Local	2016	-	1-1
32		U48	Kumi		Neck	Local	2016	-	1-1
33	UG-G12	D1/s72	Pallisa	2	Leaf	NK	2002	NI	NI
34		D15/s12	Soroti		Panicle	NK	2000	NI	NI
35	UG-G13	D15/s47	Soroti	1	Neck	NK	2002	NI	NI
36	UG-G14	U31	Gulu	1	Head	Local	2016	-	1-2
37	UG-G15	Gup-2-1	Soroti	1	Panicle	NK	2004	NI	NI
38	UG-G16	U2	Arua	1	Neck	Local	2015	-	1-2
39	UG-G17	P665n-2-1	Soroti	1	Neck	NK	2004	NI	NI
40	UG-G18	U1	Arua	1	Neck	Local	2015	-	1-1
41	UG-G19	D9/s70	Lira	1	Neck	NK	2002	NI	NI
42	UG-G20	D1/s19	Pallisa	1	Neck	NK	2000	NI	NI
43	UG-G21	Pen-2-2	Soroti	1	Neck	NK	2004	NI	NI
44	UG-G22	U5	Serere	1	Neck	Local	2016	-	1-2
45	UG-G23	S1p-1-1	Soroti	1	Panicle	NK	2004	NI	NI
46	UG-G24	D1/s11	Pallisa	1	Neck	NK	2000	NI	NI
47	UG-G25	U7	Serere	1	Neck	Improved	2016	-	1-1
48	UG-G26	U32	Kitgum	1	Neck	Local	2016	-	1-2
49	UG-G27	U3	Arua	1	Neck	NK	2015	-	1-2
50	UG-G28	U51	Ngora	1	Neck	Local	2016	-	1-1
51	UG-G29	U53	Alebtong	1	Neck	NK	2016	-	1-1
52	UG-G30	U21	Alebtong	1	Neck	NK	2016	-	1-2
53	UG-G31	D10/s77	Apac	1	Panicle	NK	2002	NI	NI
54	UG-G32	D4/s12	Tororo	1	Panicle	NK	2000	NI	NI
55	UG-G33	U28	Lira	1	Neck	Local	2016	-	1-2
56	UG-G34	U45	Serere	1	Neck	Local	2016	-	1-2
57	UG-G35	D11/s16	Masindi	1	Panicle	NK	2000	NI	NI

58	UG-G36	U15	Amuria	1	Head	Local	2016	-	1-1
59	UG-G37	U43	Masindi	1	Neck	Local	2016	-	1-2
60	UG-G38	U47	Hoima	1	Neck	Local	2016	-	1-1
61	UG-G39	U37	Manafwa	1	Neck	Local	2016	-	1-2
62	UG-G40	U44	Hoima	1	Neck	Local	2016	-	1-2
63	UG-G41	D4/s41	Tororo	1	Neck	NK	2002	NI	NI
64	UG-G42	D1/s53b	Pallisa	1	Leaf	NK	2002	NI	NI
65	UG-G43	D4/s26	Tororo	1	Panicle	NK	2000	NI	NI
66	UG-G44	D14/s27	Kabermaido	1	Panicle	NK	2000	NI	NI
67	UG-G45	D1/s44	Pallisa	1	Panicle	NK	2002	NI	NI
68	UG-G46	D15/s41	Soroti	1	Leaf	NK	2002	NI	NI
69	UG-G47	U57	Tororo	1	Neck	Local	2016	-	1-2
70	UG-G48	D8/s15	Busia	1	Panicle	NK	2000	NI	NI
71	UG-G49	D3/s24	Mbale	1	Panicle	NK	2000	NI	NI
72	UG-G50	D6/s1	Bugiri	1	Panicle	NK	2000	NI	NI
73	UG-G51	Secn-2-2	Soroti	1	Neck	NK	2004	NI	NI
74	UG-G52	U40	Tororo	1	Neck	Local	2016	-	1-1
75	UG-G53	D2/s26	Kumi	1	Leaf	NK	2000	NI	NI
76	UG-G54	U25	Apac	1	Neck	NK	2016	-	1-2
77	UG-G55	D3/s3	Mbale	1	Panicle	NK	2000	NI	NI
78	UG-G56	U35	Agago	1	Neck	Local	2016	-	1-2
79	UG-G57	U4	Moyo	1	Neck	Landrace	2015	-	1-2
80	UG-G58	D9/s54	Lira	1	Panicle	NK	2002	NI	NI
81	UG-G59	D10/s63	Apac	1	Panicle	NK	2002	NI	NI
82	UG-G60	D9/s56	Lira	1	Panicle	NK	2002	NI	NI
83	UG-G61	D12/s2	Nakasongola	1	Panicle	NK	2000	NI	NI
84	UG-G62	U34	Agago	1	Neck	Local	2016	-	1-1
85	UG-G63	U46	Hoima	1	Neck	Local	2016	+	1-1
86	UG-G64	U29	Amuru	1	Neck	Local	2016	-	1-2
87	UG-G65	U42	Arua	1	Neck	Local	2016	-	1-2
88	UG-G66	U50	Amuria	1	Neck	Local	2016	+	1-1

89	UG-G67	U19	Alebtong	1	Neck	Local	2016	-	1-1
90	UG-G68	U13	Katakwi	1	Neck	Local	2016	-	1-1
91	UG-G69	U9	Kumi	1	Neck	Local	2016	-	1-1
92	UG-G70	U16	Kumi	1	Neck	NK	2016	-	1-1
93	UG-G71	U52	Ngora	1	Neck	Local	2016	-	1-1
94	UG-G72	U41	Tororo	1	Neck	Local	2016	-	1-1
95	UG-G73	D10/s71	Apac	1	Leaf	NK	2002	NI	NI
96	UG-G74	D15/s37	Soroti	1	Leaf	NK	2002	NI	NI
97	UG-G75	D1/s50	Pallisa	1	Panicle	NK	2002	NI	NI
98	UG-G76	D3/s9	Mbale	1	Panicle	NK	2000	NI	NI
99	UG-G77	D10/s73	Apac	1	Leaf	NK	2002	NI	NI
100	UG-G78	D2/s14	Kumi	1	Panicle	NK	2000	NI	NI
101	UG-G79	U10	Kumi	1	Neck	Local	2016	-	1-2
102	UG-G80	U38	Manafwa	1	Neck	NK	2016	-	1-2

[a]Genotypes identified based on the multi-locus sequence data of HIS4, ITS, HyP2 and HyP1

[b]*M. oryzae* isolates identified as grh positive (+) or grh negative (-) based on the presence or absence of the two amplicons in PCR screening with the PKE and PES primer pairs

[c]Mating type was determined using PCR assay and the *M. oryzae* isolates were characterised as MAT 1-1 or MAT 1-2

Numbers in parentheses are the number of isolates represented by a genotype NK: Information not available

NI: Isolates not included in the *Grasshopper* and Mating types PCR analyses as these have previously been characterised (Takan *et al*., 2012)