**EXPLORING THE PHYLOGENETIC EVOLUTION AND GEOGRAPHIC TRANSMISSION PATTERNS OF THE RICE BLAST *MAGNAPORTHE ORYZAE* IN AFRICA**

**Margaret Wanjiku**

**SG30/PU/36151/18**

A research proposal submitted to the department of biochemistry and biotechnology in the school of pure and applied sciences, in partial fulfillment of the requirements for the award of the degree of Master of Science in Bioinformatics at Pwani University.

December, 2019

# 

# **DECLARATION**

This proposal is my original work and has not been presented for a degree in any other university

Signature……………………………………. Date………………………………………….

**Margaret Wanjiku**

**Pwani University**

This proposal has been submitted for examination with my approval as the university supervisor

Signature……………………………………. Date………………………………………….

**Dr Wilton Mbinda**

**Department of Biochemistry and Biotechnology**

**Pwani University**

This proposal has been submitted for examination with my approval as the host institution supervisor

Signature……………………………………. Date………………………………………….

**Dr Jean Baka Domelevo Entfellner**

**Head of Bioinformatics**

**BecA-ILRI Hub**

This proposal has been submitted for examination with my approval as the host institution supervisor

Signature……………………………………. Date………………………………………….

**Dr Geoffrey Onaga**

**National Crops Resources Research Institute (NaCRRI)**

**Table of Contents**

[**DECLARATION** 2](#_Toc29203129)

[**ABSTRACT** 4](#_Toc29203130)

[**CHAPTER 1** 5](#_Toc29203131)

[**1.0 INTRODUCTION** 5](#_Toc29203132)

[**1.1 Background** 5](#_Toc29203133)

[**1.2 Statement of the problem** 6](#_Toc29203134)

[**1.3 Objectives** 7](#_Toc29203135)

[**1.3.1 Main objective** 7](#_Toc29203136)

[**1.3.2 Specific objectives** 7](#_Toc29203137)

[**1.4 Justification** 7](#_Toc29203138)

[**1.5 Hypothesis** 7](#_Toc29203139)

[**CHAPTER 2** 7](#_Toc29203140)

[**2.0 LITREATURE REVIEW** 7](#_Toc29203141)

[**2.1 History of the disease** 8](#_Toc29203142)

[**2.3 Mating systems and mode of reproduction** 8](#_Toc29203143)

[**2.4 Pathogenicity and virulence** 9](#_Toc29203144)

[**CHAPTER 3** 9](#_Toc29203145)

[**3.0 METHODOLOGY** 9](#_Toc29203146)

[**3.1 *M. oryzae* isolates** 9](#_Toc29203147)

[**3.2 Single Nucleotide Polymorphisms (SNPs) discovery and variant calling** 11](#_Toc29203148)

[**3.3 Phylogenetic analysis** 12](#_Toc29203149)

[**3.4 Population structure analysis** 12](#_Toc29203150)

[**REFERENCES** 14](#_Toc29203151)

[**APPENDICES** 17](#_Toc29203152)

[**Appendix 1: Work plan** 17](#_Toc29203153)

[**Appendix 2: Budget** 18](#_Toc29203154)

# **ABSTRACT**

Rice blast disease is caused by a fungal pathogen, *Magnaporthe oryzae.* Outbreaks of the disease reoccur in rice growing areas resulting in major losses to rice production, consequently threatening the global food security. Previous studies were able to map the origin and lines of descent, the population structure and the virulence dynamics of the pathogen in various regions of the world. There was a study that placed isolates from South East Asia in the basal position of the phylogeny. This supports the postulation that *M. oryzae* originated from Asia, from where it dispersed and continues to spread to the rest of the world.

In Africa, studies conducted using molecular analysis based on non-sequence based markers and a few conserved genes have shown minor differences between strains from major rice growing areas. There has been no study focusing on the genetic diversity and the phylogeography of M. oryzae within Africa using whole genome data. Moreover, previous global phylogenetic studies using whole genome data involved only a few isolates from a limited number of African countries. As a result, this has provided insufficient clues about patterns of introduction and migration routes of the pathogen in and within Africa.

This study will focus on whole genome data obtained by illumina sequencing. 45 isolates of *M. oryzae* from various African countries, 3 from the Philippines and 2 from Argentina will be primarily characterized based on the occurrence of SNPs throughout the genomes. We will then look at the evolution of particular genes such as *Pi9, Pita2, Pil2* known to be important in virulence and pathogenicity according to previous studies. Signatures of adaptation to different environmental conditions in these strains will be studied. Having a better understanding of the pathogen’s phylogeograhy will help in informing future measures taken towards combating its spread and in managing the disease.

Key words: phylogeography, whole genome data, genetic diversity, population structure, virulence

# **CHAPTER 1**

## **1.0 INTRODUCTION**

### **1.1 Background**

Rice blast is a disease caused by a fungal pathogen, *Magnaporthe oryzae.* Its outbreaks keep reoccurring in areas where rice is grown resulting in major losses to rice production globally and in turn threatens global food security. The amount of rice destroyed by rice blast disease annually can feed up to 60 million people. The disease has been found to be difficult to control but knowing the structure of the current population and its diversity will help in managing the disease (Dean et al., 2005; Zhong et al., 2018). The disease affects rice plants at all stages of growth. It infects all parts of the plant except the roots. The leaves and panicles being the most seriously affected (Chuwa et al., 2015). The spread of rice blast disease indicates the evolution of new pathotypes and enhancement of population diversity is unknown (Onaga et al., 2015). South East Asia was found to be the centre of origin of *M. oryzae*. Therefore, the pathogen originated from Asia and then spread to the rest of the world. M. oryzae reproduces both sexually and asexually. However, sexual reproduction was found in just a few areas of South East Asia. This was demonstrated using the two opposite mating types of *M. oryzae*. The mating types are Mat1-1 and Mat1-2 which correlate to the reproduction capabilities of *M. oryzae* and is an important aspect to consider when studying the divergence of their populations. The crosses were done *in vitro* and they produced viable offspring. The sexual reproduction capability was found to be lost outside the centre of origin. Few strains from Africa were include in this study and so insufficient clues were provided regarding the point of entry of the pathogen to Africa and the possible routes of migration. Inferring these routes and identifying the genetic diversity within strains will likely aid in making future predictions regarding epidemics, spread and management of the disease and in the proposition of new control strategies (Gladieux et al., 2018; Saleh et al., 2014; Zhong et al., 2018). *M. oryzae* is predominantly asexual and thus results in a clonal population structure with limited genetic variability characterized by a few clonal lineages distributed across large geographical areas. The centre of origin of a pathogen is likely to consist of populations with more genetic variability than recently founded populations. By analysing the distribution of genetic diversity within and among populations, it is possible to identify centres of diversity and patterns of migration. Long distance dispersal of plant pathogens may occur naturally by air-dispersed spores and human-mediated movement of infected plant material and seeds (Saleh et al., 2012). Some studies in Africa used molecular analysis techniques. They used non sequence based markers and some conserved genes. They showed very small differences between strains from major rice growing areas within Africa (Chuwa et al., 2015; Onaga et al., 2015). By using genotyping by sequencing method, evidence of differentiation was found as well as possible adaptation along geographic lines, though at a lower representation to allow confident phylogenetic positioning among African isolates (Mutiga et al., 2017).

Therefore, there have been no studies on the phylogenetic evolution and geographic transmission patterns of rice blast *M. oryzae* in Africa, from the point of first entry and spread within the region. The spread of the pathogen throughout the world through human mediated or natural mechanisms does not give any insight as to the genetic diversity and the patterns of migration of the pathogen.

This study will focus on whole genome data obtained through an Illumina sequencing platform (Illumina HiSeq 4000 sequencer). 50 isolates of *M. oryzae* from various African countries will be primarily characterized based on the occurrence of SNPs throughout the genome. The study will also look at the evolution of particular genes such as *Pi9, Pita2, Pil2* known to be important in virulence and pathogenicity according to previous studies (Mutiga et al., 2017). Signatures of adaptation to different environmental conditions in these strains will be studied. Having a better understanding of the pathogen’s phylogeograhy will help in informing future measures taken towards combating its spread and in managing the disease. Consequently, that should have far reaching beneficial effects on consumer livelihoods globally.

### **1.2 Statement of the problem**

Millions of people rely on rice as a source of livelihood in the world. However, global rice production is constrained or threatened by rice blast. With the expected increase in cultivation of rice, the disease will inevitably continue to spread and cause havoc in all the rice growing fields. The consequences are significant reductions in rice production as well as a strain on economies. If the disease causing pathogen is not contained, the global food security will continue being threatened as rice will continue dying in the fields.

### **1.3 Objectives**

#### **1.3.1 Main objective**

1. To determine the possible point of entry of the *M. oryzae* pathogen to Africa and find any possible transmission patterns within Africa.

#### **1.3.2 Specific objectives**

1. To investigate the evolution of genes that are known to be important for virulence and pathogenicity.
2. To examine the signatures of adaptation within the African strains.

### **1.4 Justification**

In order to contain the challenges posed by rice blast, it is important to investigate the phylogenetic transmission of the pathogen within Africa. This study will focus on whole genome data of 50 isolates from different parts of Africa. Results of the study will seek to fill in the knowledge gap regarding the genetic diversity of the pathogen within Africa as well as feed into the efforts being put in towards management and combating the spread of the disease.

### **1.5 Hypothesis**

Null hypothesis; there is no genetic diversity of the pathogen within the African strains

Alternate hypothesis; there is a genetic diversity of the pathogen within the African strains

# **CHAPTER 2**

## **2.0 LITREATURE REVIEW**

While there has been much research on the transmission patterns of rice blast disease globally, few researchers have taken into consideration how the disease causing pathogen *M. oryzae* entered Africa and spread to other countries. There is need to find out the structure of the current African *M. oryzae* population as well as how the diversity is generated, to help in managing the disease.

### **2.1 History of the disease**

Rice blast disease was first reported in Asia over three centuries ago. It has since spread to 85 countries and counting. Its adaptation to different environmental conditions is said to be very high (Odjo et al., 2011). A global phylogenetic study carried that used whole genome data of 50 different isolates of the disease causing pathogen from various times and places in the world. However, the amount of African strains used in this study were very few. They found 6 lineages of the disease causing pathogen. They also found that the pathogen did separate a millennium ago. Their findings were in line with the findings of Saleh et al., 2012. (Gladieux et al., 2018).

### **2.3 Mating systems and mode of reproduction**

A global phylogenetic study found that there are 2 mating types in M. oryzae populations namely Mat1-1 and Mat1-2 which correlate to the reproduction capabilities of *M. oryzae*. Reproduction capability is an important aspect to consider when studying the divergence of populations of *M. oryzae*. They found 3 clades from the phylogenetic, PCA and STRUCTURE analysis. Two of the clades had descended from the same ancestral population whereas the third clade was more diversified to signify that that population had come from many different lineages. The first two clades only consisted of either one of the two mating types entirely while the third one had a mixture of both mating types in an almost equal proportion. They also estimated the time of divergence to be about 700-1000 years ago. The estimated divergence times were based on the time the strain was collected and the SNPs. Only a few African strains were involved in the study. (Zhong et al., 2018). Another global study found 3 clades with Asia being the origin where the pathogen was found to have been reproducing sexually, from where it spread to the rest of the world. The conclusion on this region being the origin was because some of the strains from this region were found in all the clades and also the clade containing the South East Asian isolates was the most diverse clade. However sexual reproduction was found in just a few areas of South East Asia. Together with the equal distribution of the two mating types Mat1-1 and Mat1-2 in that population, it proved that South East Asia is the centre of origin of the pathogen. The strains were genotyped using microsatellite markers. Sexual reproduction was demonstrated using the two opposite mating types, with one being required to be female-fertile in order to produce viable offspring. They also showed sexual recombination capabilities. The sexual reproduction capability was found to be lost outside this known centre of origin (Saleh et al., 2014, 2012).

### **2.4 Pathogenicity and virulence**

There was a study that examined how diverse the virulence is and how genetically related 122 isolates of the pathogen are from 9 countries in both East and West Africa. However, some samples collected from some countries were very few. For future studies, the recommendation is to get more representative samples from all the countries involved in the study. They looked at SNPs from data generated by genotyping by sequencing. They found 7 clades of the pathogen differing in virulence. The isolates from West Africa were found to have more evolved forms of virulence compared to those from East Africa. They also found an association between virulence and the clades. This they inferred was because West Africa is known to have grown rice longer than East Africa. They looked at known resistance genes. These R-genes can be used in marker-assisted breeding programs aimed at developing African rice cultivars that are resistant to *M. oryzae* (Mutiga et al., 2017). Another study looked at 88 isolates of the disease causing pathogen from East Africa, West Africa and the Philippines (Asia). They used amplified fragment length polymorphism (AFLP) markers. They found that there was no population structure and a significant flow of genes was evident. They also analysed the pathogenicity of the isolates which showed that aggression was variable amongst the East African isolates, indicating possible racial diversity. They found that there is a possibility that the East African population either consists of only one genetic population or the flow of genes is quite significant (Onaga et al., 2015).

# **CHAPTER 3**

## **3.0 METHODOLOGY**

### **3.1 *M. oryzae* isolates**

This study will focus on 50 isolates of *M. oryzae.* 45from different geographical locations in Africa, 3 from the Philippines and 2 from Argentina. All the isolates were collected during different time points and extracted from different parts of infected rice plants. They were sequenced under the Illumina HiSeq 4000 sequencer at Beijing Genomics Institute (BGI). The following table shows each of the isolate by collection site and the country of origin.

|  |  |  |
| --- | --- | --- |
| **Sample Name** | **collection site** | **Country** |
| 1- | Akagoma | Burundi |
| 4- | Nyazatare | Rwanda |
| 5 | Asome 4 | Togo |
| 5- | Asome 4 | Togo |
| 6 | Alavayo 1 | Togo |
| 7 | Akuse | Ghana |
| 7- | Akuse | Ghana |
| 9 | Savalou | Benin |
| 11 | Beterou2 | Benin |
| 12 | Malanville | Benin |
| 12- | Malanville | Benin |
| 13 | Ibadan | Nigeria |
| 14 | Abeokuta | Nigeria |
| 15 | Banfora | Burkina Faso |
| 15- | Banfora | Burkina Faso |
| 16 | Niena Dionkele | Burkina Faso |
| 16- | Niena Dionkele | Burkina Faso |
| 17 | Mai | Mali |
| 18 | Mali | Mali |
| 21- | Kyela | Tanzania |
| 22 | Kikusya | Tanzania |
| 22- | Kikusya | Tanzania |
| 23- | Sokoine | Tanzania |
| 24 | Kahama | Tanzania |
| 25 | Ifakara | Tanzania |
| 27 | Dakawa | Tanzania |
| 29 | Nguka | Kenya |
| 32 | Bugiri | Uganda |
| 34 | Mbale | Uganda |
| 36 | Doho | Uganda |
| 38 | Galiraya | Uganda |
| 39- | IRRI | Philippines |
| 41 | Cyayabaga | Rwanda |
| 43 | Nyazatare | Rwanda |
| 47 | Namulonge | Uganda |
| 49 | Lira | Uganda |
| 52 | Cankuzo-Mbaraga | Burundi |
| 54 | Akagoma | Burundi |
| 55 | Cankuzo-Mbaraga | Burundi |
| 56 | Mat 1 | Guyana |
| 58 | Ibadan | Nigeria |
| 59 | Edozhigi | Nigeria |
| 60 | Mg 6.1 | Argentina |
| 61 | Mg 6.1 | Argentina |
| 64 | Caliraya, Laguna | Philippines |
| 65 | KE 216 | Kenya |
| 69 | Tz0078 | Tanzania |
| 70 | tz0014 | Tanzania |
| 71 | 961/M36-1-3-10-1 | Philippines |

### **3.2 Single Nucleotide Polymorphisms (SNPs) discovery and variant calling**

Fastqc (v0.11.7) will be used to assess the quality of all the reads. The sequenced raw reads will be indexed and then aligned to a reference genome, MG8, using bowtie2 (v2.3.4.1) that outputs SAM files which are text-based files (Langmead & Salzberg, 2012). Samtools (v1.8) will then be used to convert the SAM files to BAM files (a binary counterpart of the SAM file). This is because it is easier for computer programs to work with binary files. Samtools will also be used to sort and index the resulting BAM files. After aligning reads to a genome, the alignments resulting from that are usually in a random order with respect to their positions in the reference genome. Any meaningful downstream analysis to be carried out later would require a further manipulation of the BAM file. The BAM files will be sorted using Samtools, such that the reads are ordered on the basis of their alignment coordinates on each chromosome. After that, the genome sorted BAM files will be indexed, to allow quick extraction of alignments overlapping particular genomic regions. IGV will be used to view the genome and it requires that the genomes are indexed beforehand (Li et al., 2009). Bcftools (v1.8) will be used to perform variant calling and generate VCF files which will then be filtered to produce a list of positions where at least one of the VCF files has a polymorphism (SNP) (Li, 2011a, 2011b). IGV (v2.4.5) and igvtools (v2.3.98) will be used for whole genome viewing (Robinson et al., 2017; Thorvaldsdóttir et al., 2013). Vcftools will be used to filter out all the indels and microsatellites from the VCF files to remain with only positions with SNPs (Danecek et al., 2011). We will analyse the rate of non-synonymous mutations to the rate of synonymous mutations (dN/dS analysis). This will help in identifying selection signatures by looking at all the genomic regions under selection pressure.

### **3.3 Phylogenetic analysis**

Phylogenetic trees of the SNP data will be constructed using BIONJ, adistance based phylogeny reconstruction algorithm (Gascuel, 1997). The trees will be pseudo alignments made of SNP positions. Molecular dating of phylogenetic trees will then be done to estimate the time when the evolutionary events occurred as well as the rates of molecular evolution. This will be done using the Bayesian Evolutionary Analysis by Sampling Trees (BEAST) software, (v1.8.3) (Rieux & Balloux, 2016; Suchard et al., 2018).

### **3.4 Population structure analysis**

The population structure will be inferred by clustering all the 50 isolates of *M. oryzae* into subpopulations based on the SNP data. The gene flow amongst these subpopulations will also be evaluated to help in inferring the patterns of migration of the pathogen and the genetic diversity amongst them. This will be done using a program called STRUCTURE (v3.2.2). Among the outputs of this program is a matrix (the Q matrix), that assigns an individual a probability depending on the subpopulation it belongs to (Pritchard et al., 2000). This matrix is the incorporated as a fixed effect in the “Unified-Mixed Model” for association analysis (Yu et al., 2006). In addition, principal component analysis (PCA), a method that summarizes data will be used to characterize the genetic variation patterns within and among the populations (Ma et al., 2012). Eventually, we will perform a genome wide association studies (GWAS) analysis. It will be carried out in combination with the metadata, to determine potential links between genotype and phenotype, providing adaptation to certain environments. However, a shortcoming may arise due to the limited number of samples available to carry this out. Most GWAS analyses require large numbers of samples to be effective.

# **REFERENCES**

Andrews S. (2010). FastQC: a quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc> ## for citing fastqc tool kit

Chuwa, C. J., Mabagala, R. B., & Reuben, M. S. O. W. (2015). *Assessment of Grain Yield Losses Caused by Rice Blast Disease in Major Rice Growing Areas in Tanzania*. *4*(10), 2211–2218.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., … Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Dean, R. A., Talbot, N. J., Ebbole, D. J., Farman, M. L., Mitchell, T. K., Orbach, M. J., … Dirren, B. W. (2005). The genome sequence of the rice blast fungus Magnaporthe grisea. *Nature*, *434*(7036), 980–986. https://doi.org/10.1038/nature03449

Gascuel, O. (1997). BIONJ: An improved version of the NJ algorithm based on a simple model of sequence data. *Molecular Biology and Evolution*, *14*(7), 685–695. https://doi.org/10.1093/oxfordjournals.molbev.a025808

Gladieux, P., Ravel, S., Rieux, A., Cros-Arteil, S., Adreit, H., Milazzo, J., … Tharreau, D. (2018). Coexistence of multiple endemic and pandemic lineages of the rice blast pathogen. *MBio*, *9*(2), 1–18. https://doi.org/10.1128/mBio.01806-17

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Li, H. (2011a). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, *27*(21), 2987–2993. https://doi.org/10.1093/bioinformatics/btr509

Li, H. (2011b). Improving SNP discovery by base alignment quality. *Bioinformatics*, *27*(8), 1157–1158. https://doi.org/10.1093/bioinformatics/btr076

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Ma, J., & Amos, C. I. (2012). Principal components analysis of population admixture. *PLoS ONE*, *7*(7). https://doi.org/10.1371/journal.pone.0040115

Mutiga, S. K., Rotich, F., Ganeshan, V. D., Mwongera, D. T., Mgonja, E. M., Were, V. M., … Correll, J. C. (2017). Assessment of the virulence spectrum and its association with genetic diversity in magnaporthe oryzae populations from sub-Saharan Africa. *Phytopathology*, *107*(7), 852–863. https://doi.org/10.1094/PHYTO-08-16-0319-R

Odjo, T., Ahohuendo, B. C., Onasanya, A., Akator, K., & Séré, Y. (2011). Analysis of magnaporthe oryzae population structure in Benin. *African Journal of Agricultural Research*, *6*(28), 6183–6188. https://doi.org/10.5897/AJAR11.1329

Onaga, G., Wydra, K., Koopmann, B., Séré, Y., & Von Tiedemann, A. (2015). Population structure, pathogenicity, and mating type distribution of Magnaporthe oryzae isolates from East Africa. *Phytopathology*, *105*(8), 1137–1145. https://doi.org/10.1094/PHYTO-10-14-0281-R

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959.

Rieux, A., & Balloux, F. (2016). Inferences from tip-calibrated phylogenies: A review and a practical guide. *Molecular Ecology*, *25*(9), 1911–1924. https://doi.org/10.1111/mec.13586

Robinson, J. T., Thorvaldsdóttir, H., Wenger, A. M., Zehir, A., & Mesirov, J. P. (2017). Variant review with the integrative genomics viewer. *Cancer Research*, *77*(21), e31–e34. https://doi.org/10.1158/0008-5472.CAN-17-0337

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*(4), 1440–1456. https://doi.org/10.1111/nph.12627

Saleh, D., Xu, P., Shen, Y., Li, C., Adreit, H., Milazzo, J., … Tharreau, D. (2012). Sex at the origin: An Asian population of the rice blast fungus Magnaporthe oryzae reproduces sexually. *Molecular Ecology*, *21*(6), 1330–1344. https://doi.org/10.1111/j.1365-294X.2012.05469.x

Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., & Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution*, *4*(1), 1–5. https://doi.org/10.1093/ve/vey016

Thorvaldsdóttir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, *14*(2), 178–192. https://doi.org/10.1093/bib/bbs017

Yu, J., Pressoir, G., Briggs, W. H., Bi, I. V., Yamasaki, M., Doebley, J. F., … Buckler, E. S. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics*, *38*(2), 203–208. https://doi.org/10.1038/ng1702

Zhong, Z., Chen, M., Lin, L., Han, Y., Bao, J., Tang, W., … Wang, Z. (2018). Population genomic analysis of the rice blast fungus reveals specific events associated with expansion of three main clades. *ISME Journal*, *12*(8), 1867–1878. https://doi.org/10.1038/s41396-018-0100-6

# **APPENDICES**

## **Appendix 1: Work plan**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Activity | Timeline | | | | | | | | | | | | |
|  | Sept | Oct | Nov | Dec | Jan | Feb | Mar | Apr | May | June | July | Aug | Sept |
| Developing project concept |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Proposal writing & corrections |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Proposal presentation at university |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Data analysis & project work |  |  |  |  | | | |  |  |  |  |  |  |
| Dissertation writing & corrections |  |  |  |  |  |  |  |  | | |  |  |  |
| Dissertation presentation |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Writing of manuscript |  |  |  |  |  |  |  |  |  |  |  | | |

## **Appendix 2: Budget**

|  |  |
| --- | --- |
| Description | Cost (US$) |
| Office space at BecA | 600.00 |
| Internet services | 100.00 |
| Access cards | 300.00 |
| Total | 1000.00 |