IDENTIFICATION AND CHARACTERIZATION OF SELECTED TARO VIRUSES IN KENYA USING NEXT-GENERATION SEQUENCING

DAVID KARIUKI MURUU

MASTER OF SCIENCE

(Molecular Biology and Bioinformatics)

JOMO KENYATTA UNIVERSITY

OF

AGRICULTURE AND TECHNOLOGY

2024

Identification and Characterization of Selected Taro Viruses in Kenya Using Next-Generation Sequencing

David Kariuki Muruu

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Molecular Biology and Bioinformatics of the Jomo Kenyatta University of Agriculture and Technology

2024

DECLARATION

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

Signature…… Date……..

David Kariuki Muruu

This thesis has been submitted for examination with our approval as University Supervisors

Signature…… Date……..

Prof. Johnson Kinyua, PhD

JKUAT, Kenya

Signature…… Date……..

Dr. Benard Mware, PhD

International Livestock Research Institute, Kenya

Signature…… Date……..

Dr. Sita Ghimire, PhD

International Livestock Research Institute, Kenya

DEDICATION

This Thesis is dedicated to my father Mr. Muruu for his passion for education as evidenced by his commitment to educating his children in pursuit of knowledge despite the limited resources coupled with competing needs. Secondly, I dedicate this piece of work to my lovely wife Judy, our son Jayson, and our daughter Lilian for their continuous support, especially during the times I had to be away from home for studies. I could not have done it without you! Your care and support mean the world to me. Thanks, and I wish you all God's blessings.

To all people who care about our food and nutritional security, may this thesis shed light and be an inspiration to overcome the many challenges we face with respect to coming up with solutions to achieving sustainable development goal number 2 "ZERO HUNGER"

ACKNOWLEDGMENT

First, all glory goes to the Almighty God who has given me the strength and grace to reach this far. It is surely by His mercies that this process has been a success.

Secondly, I would like to greatly thank my supervisors, Drs. Sita Ghimire, Bernard Mware and Johnson Kinyua for their scientific guidance throughout my master's study and for the time they devoted to the organization of this work, their valuable insights, comments, intelligence, and continuous patience have been instrumental in the success of my research and inspired me during my personal development. I wish you God's blessings.

Dr. Isaac Njaci, and research associates, Leah, Dorothy, and Eunice am greatly indebted to you for your invaluable advice, assistance, and support during my research.

I am profoundly grateful to the BecA-ILRI Hub management and donors who gave me an opportunity to do my research in a world-class Laboratory. To all the staff and ABCF fellows who took the time to support me in my work, I thank you. We can contribute to the attainment of a zero-hunger country and continent together by offering scientific solutions to emerging scientific problems.

Last and not least, I would like to thank my family for their continuous support, ever-lasting understanding, love, and sacrifices. You have made a difference with your support and presence in my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABBREVIATIONS AND ACRONYMS

ABCF	Africa Biosciences Challenge Fund
AGO	Argonaute
BecA	Biosciences Eastern and Central Africa
Bp	Base Pairs
CBDV	*Colocasia Bobone Disease Virus*
DCI	Dicer-Like Enzymes
DNA	Deoxyribonucleic Acid
DsDNA	Double-Stranded DNA
DsMV	*Dasheen Mosaic Virus*
ELISA	Enzyme-Linked Immune-Sorbent Assay
FAO	Food and Agriculture Organization
hpRNAs	Hairpin RNAs
ILRI	International Livestock Research Institute
MEGA	Molecular Evolutionary Genetics Analysis
MiRNA	MicroRNA
NGS	Next Generation Sequencing
Nt	Nucleotide
ORF	Open Reading Frame

PAMP	Pathogen-Associated Molecular Pattern
PCR	Polymerase Chain Reaction
RDP	Recombination Detection Program
RdRp	RNA-Dependent RNA Polymerases
RI	Replication Intermediates
RNA	Ribonucleic Acid
RT-PCR	Reverse Transcriptase - Polymerase Chain Reaction
SiRNA	Short Interfering RNAs
sRNA	Small RNAs
TaBCHV	*Taro Bacilliform China Virus*
TaBV	*Taro Bacilliform Virus*
TaRV	*Taro Reovirus*
TaVCV	*Taro Vein Chlorotic Virus*
TLB	*Taro Leaf Blight*
VAMP	Virus-Associated Molecular Pattern
VsiRNAs	Virus-Derived Small Interfering RNAs

DEFINITION OF OPERATIONAL TERMS

Library Preparation	It is the process of converting a genomic DNA sample (or cDNA sample) into a pool of DNA fragments with adapters which can then be sequenced on an NGS instrument.
Next Generation Sequencing	It is a high-throughput method that employs DNA sequencing technologies capable of processing multiple DNA sequences in parallel to determine a portion of the nucleotide sequence of an individual's genome.
Polymerase Chain Reaction	It is a laboratory technique used to amplify a specific piece of DNA into many copies typically from a sample that contains very tiny amounts of that DNA thus allowing the specific DNA piece to be detected.
Sequencing	It is a general laboratory technique for determining the exact order of nucleotides, or bases, often referred to by the first letters of their chemical names: A for Adenine, T for Thymine, C for Cytosine, and G for Guanine in a DNA molecule which encodes the biological information of an organism.
Small RNAs	They are short, typically ~ 18 to 30 nucleotides, non-coding RNA molecules that regulate gene expression by targeting diverse categories of genes from different biological and metabolic processes.

Virus isolate	It is a very basic term that implies that a specific virus was isolated from an infected single host and is given a name so that its origin is known.

ABSTRACT

Plant viruses are one among the biggest limiting factors of crop production globally with a potential of causing up to 100% yield loss. Identification and characterization of viruses is a critical step in formulating effective disease management strategies and programs. Several diagnostic tools, including physical, biological, serological, and molecular approaches are available for detection and analysis of viral pathogens. These methods require sequence information for the target viruses and thus making it difficult to target and study emerging viruses or even identify rare pathogens. Taro is an important food security crop affected by several viral pathogens which have not been characterized in detail because of lack of adequate sequence information of these pathogens. Insufficient sequence information on a pathogen impedes the ability to understand and develop a control strategy for such pathogens. Next generation sequencing tools, especially small RNA sequencing, is a powerful technology for robust detection and identification of both known and unknown viruses due to its ability to multiplex detections with no prior knowledge of the virus sequence. NGS offers advantages over all existing methods of identifying a viral pathogen, including removing the need for targeting a specific pathogen or requiring sequence information for that pathogen, identifying multiple pathogens in a single sample, and eliminating the need for costly and often ineffective culturing or antibody laboratory tests. Herein, PCR and RT-PCR assays using degenerate primers were employed to screen for previously known viral pathogens infecting taro germplasm in Kenya. In-depth sequencing of small RNAs isolated from both virus symptomatic and asymptomatic Taro germplasm using next-generation sequencing technology was subsequently used to detect the unknown viral pathogens. Subsequent bioinformatics analyses revealed the presence of both DNA and RNA viruses. Detected DNA viruses included the *Taro Bacilliform Virus* (TaBV) and *Taro Bacilliform CH Virus* (TaBCHV), which are badnaviruses specific to Taro, the *sweet potato Badnavirus B, sugarcane bacilliform virus,* and *sweet potato leaf curl virus*. The RNA viruses included the Colocasia Bobone Disease Virus, a rhabdovirus specific to Taro, and the *East African Cassava mosaic virus, sweet potato feathery mottle,* and *Phaseolus vulgaris alphaendornavirus*. A Citrus exocortis viroid was also detected. Interestingly, the wild Taro relatives, including tannia, had no viral sequence hits affirming that wild species possess some level of tolerance to viral infections, possibly because of having a rich reservoir of resistance genes useful in breeding cultivars with a genetically controlled resistance against numerous diseases. Moreover, the viral severity in the fields was also positively correlated with the number of viruses identified, as evidenced by the positive correlation between the viral prevalence based on the visual identification of symptoms and the TaBV PCR results. These results collectively demonstrate the reliability of the sRNA deep sequencing data in determining virus and viroid diversity. This study reports the Taro viruses and viroids circulating in Kenya and comprehensively describes the prevalence, distribution, and sequence variability of TaBV in Kenya. The study forms a basis for developing effective management strategies to support the prevention and control of Taro viruses.

CHAPTER ONE

INTRODUCTION

1.1 Background Information

In sub-Saharan Africa, 2 in 10 people are undernourished with over 709 million people being moderately to serverely food insecure (FAO statistics, 2022). Despite this situation, there is over-reliance on cereals for supply of dietary energy as compared to root and tuber crops (Global Nutrition Report, 2022). In Kenya, it is estimated that over 14 million people are food insecure, attributed to climate change and high rate of inflation, mostly relying on food relief (FAO statistics, 2022). There is over-reliance on maize as stapple food, yet it is plagued by perennial supply shortage and price fluctuations. Yet many households have limited alternatives food choices thus persistently suffer food insecurity. Food security is "a situation in which all people, always, have physical, social and economic access to sufficient, safe and nutritious food which meets their dietary needs and food preferences for an active and healthy life" (Kenya Food Security Steering Group, 2008). To achieve this, over reliance on maize must be prevented by providing alternative food choices especially neglected root and tuber crops. Taro holds the potential to provide alternative staple food and relief from over-reliance on maize in many communities in Kenya. It is an important source of fine highly digestible starch consisting of 17-28% amylose and amylopectin suitable for persons with digestive problems. It is also useful to people with an allergic reaction to cereals because it contains higher protein levels in comparison to other aroids. Moreover, it is rich in carotene, potassium, calcium, phosphorous, iron, riboflavin, thiamine, niacin, vitamin A, vitamin C, and dietary fibre (Temesgen & Retta, 2015). It is a basic source of starch in the diet of many communities in central, western, and southwest Kenya (Akwee *et al.,* 2015).

Despite Taro's potential to ameliorate malnutrition and food insecurity, it is a neglected orphaned, under-researched crop with narrow genetic diversity whose improvement and production intensification would require a robust clean seed delivery system as its production in East Africa is mainly affected by viral infections, *Taro leaf blight* (TLB) and lack of clean planting materials. One of the pre-requisites of an

efficient clean seed delivery system is the availability of effective rapid virus screening protocols. Studies in other countries and regions report *Dasheen Mosaic Virus* (DsMV), *Taro Bacilliform Virus* (TaBV), Colocasia *bobone disease virus* (CBDV), *Taro vein chlorosis virus* (TaVCV) and *Taro reovirus* (TaRV) as the main taro infecting viruses (Revill *et al.*, 2005) with *Dasheen Mosaic Virus* (DsMV) and *Taro Bacilliform Virus* (TaBV) being reported as the most predominant viruses (Revill *et al.*, 2005). These Taro viruses have varying adverse effects on taro production both individually or when in synergy with some infections causing up to a hundred percent loss (Revill *et al.*, 2005). Although Kidanemariam *et al*. (2022) reported the presence of different taro viruses in East Africa, the major taro infecting viruses in Kenya remains unknown hence rendering it impossible to develop high throughput diagnostic assays to determine the prevalence and distribution of the viruses across the country. Further, it complicates the development of an effective system for the production of clean (virus-free) planting materials as well as the development of innate antiviral defense strategies that are more effective in viral disease control. This study aimed to identify and characterize viruses infecting taro in Kenya as a basis for effective disease management strategy and control of the viruses using next-generation sequencing technology of the small RNAs and subsequent bioinformatics data analysis due to its robust detection and identification of both known and unknown viruses. This study also compared viral susceptibility of the edible and wild Taro genotypes in Kenya that will further form a basis of genotypes that could be used for taro improvement through various breeding programs. These outputs will contribute to the development of disease-resistance strategies for enhanced Taro productivity in the country and beyond.

1.2 Problem Statement

Taro production in Kenya is faced with a myriad of challenges despite its potential to ameliorate food and nutritional insecurity; it is difficult for farmers to get disease-free planting materials due to the lack of an effective virus-cleaning system (Akwee *et al*., 2015; Onyeka *et al.,* 2014). This has led farmers to recycle and borrow planting materials from each other leading to viral diseases build-up and spread. The identity of taro viruses circulating in Kenya remains unknown. This phenomenon has rendered it difficult to determine the incidence and distribution of these viral diseases in the

country which would further inform the best management strategies that could be used to control and/or manage the viruses. All these are due to a lack of specific high throughput diagnostic assays that can only be developed based on specific relevant information, including the sequence information of the viruses (Wanyama and Mardell, 2006). Several diagnostic tools, including physical, biological, serological, and molecular approaches are available for detection and analysis of viral pathogens. Notably, molecular approaches require sequence information for the target viruses, thus making it difficult to target and study emerging viruses or even identify rare pathogens. This study therefore employed small RNA sequencing, which does not require any prior knowledge of Taro viral diseases sequence information, to generate base line information of viruses infecting taro in Kenya as a basis for effective management and control of taro viruses.

1.3 Justification

The ability to detect viruses is basic to all areas of viral investigation. The symptoms appearing on the leaves of the plants infected by viruses are an obvious piece of evidence for the diagnosis of viral diseases. However, virus identification might be misleading if the diagnosis is based on the symptom alone due to multiple viruses manifesting themselves in the same way in terms of symptoms. Therefore, correct identification of a virus, or viruses, infecting a particular crop is essential in characterizing it as well as the development of effective control measures to be applied (Babu *et al*., 2011). Established molecular methods for virus identification that have been in use have two major shortcomings, i.e., sequence information for the target viruses must be known, making it difficult to target and study emerging viruses and an individual analytical test must be conducted to confirm or refute each pathogen, making identification of rare or unexpected pathogens difficult, and identification of previously unknown pathogens impossible (Wang *et al*., 2002).

This study, therefore, focused on identifying and characterizing Taro viruses in Kenya using next-generation sequencing (NGS) of small RNAs because of their robust detection and identification of both known and unknown viruses due to its ability to multiplex detections with no prior knowledge of the virus sequence. Notably, NGS

offers advantages over all existing methods of identifying a viral pathogen, including removing the need for targeting a specific pathogen or requiring sequence information for that pathogen, identifying multiple pathogens in a single sample, and eliminating the need for costly and often ineffective culturing or antibody laboratory tests (Wu *et al*., 2015). The findings of this study provide base line information regarding viruses infecting Taro in Kenya consequently forming a basis for developing effective management and control strategies of taro viruses through the development of an effective system for the production of clean (virus-free) planting materials, development of specific high throughput diagnostic assays to determine the incidence and distribution of the viruses across the country, and development of innate antiviral defense strategies that are more effective in viral disease control. Such strategies will contribute greatly to improving Taro productivity in the country and region (Revill *et al.*, 2005; Palapala *et al.*, 2009).

1.4 Hypothesis

The null hypothesis is degenerate primers specific to Taro viruses cannot effectively determine the prevalence and distribution of viruses infecting Taro in Kenya.

1.5 Research Questions and Objectives

This study sought to answer the following questions:

1. Can degenerate primers specific to Taro effectively determine the prevalence and distribution of viruses infecting Taro in Kenya? Would it require the development of specific high throughput diagnostic assays to determine their prevalence and distribution?
2. What are the molecular characteristics of the major Taro viruses circulating in Kenyan Taro fields?

1.5.1 General Objective

To screen and characterize Taro viruses in selected agro-ecologies in Kenya.

1.5.2 Specific Objectives

1. To screen viruses and viroids infecting taro in selected agro-ecologies in Kenya using PCR and RT-PCR assays and small RNA sequencing
2. To characterize the major taro viruses circulating in Kenyan Taro fields

1.6 Scope of the Study

Viral surveys and sample collection was done in nine major Taro growing counties in Kenya: Kiambu, Murang'a, Meru, Nyeri, Siaya, Busia, Kakamega, Kisii, and Machakos. These counties were chosen because they have numerous rivers and streams where most farmers plant taro on their banks. They were also considered due to their varying Agro-ecological and climatic classifications. Visual identification of previously reported viral symptoms in both edible and wild relatives of Taro (Revill *et al.*, 2005; Yang *et al*., 2003*a,* 2003*b*) was done and the prevalence scores subsequently determined from ten randomly selected plants in each field. Leaves from Taro plants showing viral symptoms were collected for PCR and RT-PCR screening of previously reported viruses. Leaves from asymptomatic plants were also collected to determine whether some viruses are latent. Suckers from both the edible and wild relatives of Taro were also collected and planted in pots for monitoring of viral symptoms over a period of two months to compare viral susceptibility and collection of samples for small RNA sequencing. Of note, most of the wild relatives of Taro showed no viral disease symptoms. This study comprehensively covers the identification and characterization of viruses infecting Taro germplasm in Kenya.

1.7 Limitations of the Study

The limitation of funds for field work made it impossible to extend the field work to all counties that grow Taro which include Kirinyaga, Embu, Makueni, and Nakuru counties. Moreover, we were only able to do sanger sequencing of 15 TaBV isolates out of the 115 TaBV isolates because of limited funds. However, the selection for the survey and isolates for sanger sequencing was done in such a way that it was representative of all the agroecological zones where taro is grown, and hence no loss of valuable information was expected.

CHAPTER TWO

LITERATURE REVIEW

2.1 Taro (*Colocasia Esculenta L.*)

Taro (also known as arrowroot, tannia, cocoyam, or nduma in Kenya) is among the most important root crops for food and economic security in sub-Saharan African and Asia-Pacific countries (Chaïr *et al.*, 2016; He *et al.*, 2015). It is a food crop that is cultivated for both its foliage and root which are utilized as a food source (Lee, 1999). It is a member of the family Araceae and is believed to be one of the earliest root crops cultivated in the world (Plucknett, 1976). Taro is vegetatively propagated from suckers or sets and can grow up to 2 meters high. Taro is a rich source of proteins, minerals, carbohydrates, and vitamins making it very useful to people with an allergic reaction to cereals because it contains higher protein, carotene, potassium, calcium, phosphorous, iron, riboflavin, thiamine, niacin, vitamin A, vitamin C, and dietary fibre levels in comparison to other aroids (Akwee *et al.*, 2015). It is a basic source of starch in the diet of many communities in central, western, and southwest Kenya.

2.2 Taro Production Status and Potential

Taro farming in the world is mostly characterized by small holder production systems relying on minimum external resource input (Singh *et al*, 2012). The world production of Taro is estimated at around 12 million tons per annum produced from about 1.8 million hectares with an average yield of 6.9t/ha (FAO statistics, 2022). Most of the global production comes from developing countries. Global data on taro production indicates that Nigeria is the highest world's taro producer having produced 3.22 million tonnes that accounts for 25.86% of the world's Taro (cocoyam) as of 2021. Ethiopia, China, Cameroon, and Ghana account for another 55.77% totaling to 81.63% of global taro production (FAO statistics, 2022). In Africa, the average yield of taro is around 4.9t/ha with high production of global taro production; about seventy-four percent (74%) comes from West and Central African countries (FAO statistics, 2022). In East Africa, high production of taro comes from Rwanda and Burundi with both countries having an average yield of 4.5t/ha (FAO statistics, 2022). The Food and

Agriculture Organization (FAO) database report on the production of major crops uses the label taro to represent the total production of all Colocasia and Xanthosoma spp. However, there are no FAO reports for Kenya, Uganda, and Tanzania. This phenomenon attests to the fact there are no concerted efforts by the various stakeholders to improve the taro production system in East Africa. While Pacific Islands countries have been able to achieve significant levels of exports in Taro, with 10,000-12,000 tons exported annually, African countries like Kenya have not given much attention to the underutilized and orphan crop even though Taro is well adapted to different agro-ecological conditions (Akwee *et al*., 2015). In Kenya, the production of Taro is extremely low compared to the neighboring countries like Rwanda and Burundi which are exporters of Taro. It is even extremely low in comparison to other root and tuber crops like cassava, sweet potato, and yams. Hindering factors such as lack of planting materials, improved Taro varieties, pests and diseases, limited research activities, and information research on taro germplasm varieties are the main causes of underutilization of the crop compared to Pacific-Islands communities (Akwee *et al*., 2015). In some parts of Kenya like the Mount Kenya region, parts of the Rift valley, western and Nyanza, Taro is grown by small-scale farmers near the streams and riverbanks because most of them lack modern irrigation facilities for upland Taro cultivation (Akwee *et al*., 2015).

Taro production systems are regarded as an informal production activity and are managed outside the convectional market and economic channels by both researchers and policy makers (Onyeka *et al.*, 2014). It suffers low productivity, probably due to low-quality planting materials and low levels of value-addition and processing (Wanyama and Mardell, 2006). Yet, in a region that is characterized by increasing poverty levels and food insecurity because of the ever-increasing population, Taro could contribute substantially to the food and income security of many households. Unlike in Kenya, the Pacific Island countries cultivate taro under greater intense systems and has the highest percentage contribution to the diet. As such, Taro has the potential to ameliorate household food hunger and malnutrition for people who live below the poverty line such as rural families in Kenya. In addition to contributing to sustained food security in the domestic market, it can also bring in export earnings to the country (Revill *et al.*, 2005; Palapala *et al.*, 2009).

2.3 Pests and Diseases of Taro

Taro production is largely affected by pests and diseases (Ivancic, 1992). Amongst the pests, the taro beetle is of great concern. It burrows into the corms of Taro, making smooth-sided tunnels that run together to form large cavities causing secondary rots to develop (Masamdu, R. and Simbiken, N., 2000). Taro diseases are classified as viral, bacterial, and fungal diseases including oomycetes. Due to its vegetative nature of propagation, young suckers and sets from the previous crop are mostly used as planting material. Some of these materials are already infected by viruses while others get infected through transmission by herbivorous insect vectors, such as the aphids and the leaf hoppers causing massive diseases to build up and spread. This has resulted in severe yield reductions in recent years (Babu *et al*., 2011). The viral diseases reported in other countries, especially in Asia include *Dasheen mosaic virus* (DsMV), *Colocasia bobone disease virus* (CBDV), *Taro bacilliform virus* (TaBV), *Taro vein chlorosis virus* (TaVCV) and *Taro reovirus* (Revill *et al.*, 2005). Of the five viruses, TaBV is the only DNA virus classified as a badnavirus (Yang *et al*., 2003*a*, 2003*b*) while the other four viruses are RNA viruses; DsMV is a potyvirus while CBDV, TaVCV, and TaRV are rhabdoviruses. Studies in Asian countries report that the *Dasheen Mosaic Virus* (DsMV) and *Taro Bacilliform Virus* (TaBV) are the most predominant Taro viruses (Revill *et al.*, 2005).

DsMV is reported to be found wherever Taro is grown infecting both the edible and ornamental aroids (Zettler and Hartman, 1986, 1987; Jackson, 1980; Shaw *et al*., 1979). It is transmitted by aphids and is caused by a stylet-borne, flexuous, rod-shaped virus that is spread by aphids. It is characterized by chlorotic and feathery mosaic patterns on the leaf, distortion of leaves, and stunted plant growth although cultivars vary considerably in symptom expression. Taro bacilliform virus (TaBV) is transmitted by the plant hopper. It is thought to occur in combination with CBDV to cause “alomae” disease (James *et al*., 1973). Alomae disease is considered the most destructive virus disease of taro (Jackson & Gollifer, 1975; Rodoni *et al.*, 1994). Symptoms include crinkling of young leaves that fail to develop normally, the presence of thickened veins and lamina, shortening of the petioles, and the presence of irregularly shaped outgrowths on the petioles. Infected plants ultimately die due to the

development of systemic necrosis (Rodoni *et al*., 1994). Infection with TaBV alone results in a range of mild symptoms including stunting, mosaic, and down-curling of the leaf blades while infection of taro with CBDV alone results in a disease known as bobone which is characterized by stunting, leaf distortion, and presence of galls on the petioles (Jackson 1978).

2.4 Identification and Characterization of Plant Viruses

The ability to detect viruses is basic to all areas of viral investigation. The symptoms appearing on the leaves of the plants infected by viruses are an obvious piece of evidence for the diagnosis of viral diseases. However, virus identification might be misleading if the diagnosis is based on the symptom alone due to multiple viruses manifesting themselves in the same way in terms of symptoms. Therefore, correct identification of a virus, or viruses, infecting a particular crop is essential in characterizing it as well as the development of effective control measures to be applied. Detection methods depend on different properties of the virus. In the past three decades, several methods have been reported for the detection and identification of viruses. To date, there are several diagnostic tools, including physical, biological, serological, and molecular approaches that have been employed to identify and characterize viral pathogens. Traditionally, viral pathogens are detected on cultured cell monolayers that exhibit cytopathic effects or plaques, or by antibody neutralization tests. However, many viral types are not cultivable in the laboratory, and antibody neutralization tests depend on the availability of quality antiserum (Wang *et al*., 2002), hindering the identification, discovery, and research of these pathogens. Several methods have been adopted for routine use in diagnosis, pathology, and virology laboratories. These include the two most successfully established virus detection methods: the protein-based enzyme-linked immune-sorbent assay (ELISA) for both DNA and RNA viruses and the nucleotide sequence-based reverse transcription polymerase chain reaction (RT-PCR) for RNA viruses or polymerase chain reaction (PCR) for DNA viruses (Boonham *et al*., 2014). However, established molecular methods have two major shortcomings, i.e., sequence information for the target viruses must be known, making it difficult to target and study emerging viruses and an individual analytical test must be conducted to confirm or refute each pathogen,

making identification of rare or unexpected pathogens difficult, and identification of previously unknown pathogens impossible.

More recently, next-generation sequencing (NGS) technology which allows for sequencing the total nucleic acid content in disease samples and, subsequently, identifying the pathogens by analysis of the NGS data using bioinformatics tools (Wu *et al*., 2015) has become a very powerful technology for robust detection and identification of both known and unknown viruses due to its ability to multiplex detections with no prior knowledge of the virus sequence. NGS offers advantages over all existing methods of identifying a viral pathogen, including removing the need for targeting a specific pathogen or requiring sequence information for that pathogen, identifying multiple pathogens in a single sample, and eliminating the need for costly and often ineffective culturing or antibody laboratory tests. One of the powerful, robust NGS approaches for the identification of viral pathogens is small RNA sequencing (Cox-Foster *et al*., 2007; Quan *et al*., 2008; Li *et al*., 2012). Using sRNA assembly and analysis, Wu *et al*., (2010) were able to identify several known and unknown viruses infecting invertebrate cell lines. In plant systems, NGS technologies have also been applied for virus and viroid identification (Adams *et al*., 2009; Al Rwahnih *et al*., 2009; Kreuze *et al*., 2009; Navarro *et al*., 2009; Coetzee *et al*., 2010; Hagen *et al*., 2011). Deep sequencing of sRNAs is the most promising detection and identification technique with abilities to identify RNA and DNA viruses in sweet potatoes (Kreuze *et al*., 2009) and pospiviruses in grapevine (Navarro *et al*., 2009). Hagen *et al*., (2011) extended NGS applications to identify RNA and DNA viruses in tomatoes and squash.

Regulatory small RNAs including small interfering RNAs (siRNAs), hairpin RNAs (hpRNAs), and Piwi-interacting RNAs (piwi RNAs), ranging in size from 20 to 24 nucleotides (nt), are ubiquitous components of endogenous plant transcriptomes and are a common response to exogenous viral infections (Wu *et al*., 2012). Most viruses are RNA viruses whose genomes contain imperfect regulatory stem-loops and produce double-stranded RNA (dsRNA) replication intermediates (RIs) by viral RNA-dependent RNA polymerases (RdRp). As a virus-associated molecular pattern (VAMP, a form of PAMP), these dsRNAs are siRNA precursors and represent the hallmark of RNA virus infection. VAMPs are associated with Dicer-like enzymes

(DCL), which can cleave dsRNA into viral-derived small interfering RNAs (vsiRNAs) that, upon loading into Argonaute (AGO) proteins, improve antiviral defenses through RNA silencing. In principle, 'dicing' of dsRNA viral RIs would be enough to mediate antiviral silencing. It is well known that plants are infected by either DNA or RNA viruses that produce vsiRNAs because of virus-induced innate resistance. Despite being short sequences (21 to 24 nt in length), vsiRNAs overlap each other and cover the entire viral genome (Kreuze *et al*., 2009; Wu *et al*., 2012). The accumulated level of vsiRNAs corresponding to infection by different viruses has been reported to be variable in different host plants (Donaire *et al*., 2009, Lin *et al*., 2010) and hence can be used to determine the most predominant virus in a sample. Therefore, deep sequencing and assembly of total small RNAs for virus-infected samples offer a specific and sensitive approach to virus discovery and in regulation of a numerous biological processes in plants, including plant developmental processes and metabolism, maintenance of the genome integrity, immunity against pathogens, and abiotic stress responses.

2.5 Control of Plant Viruses

Plant viruses are known to cause considerable losses in production. In Taro, virus infection causes a reduction in corm size and quality, with yield losses of up to 20% having been reported (Zettler & Hartman, 1986). In the past, an efficient management strategy was to control insect vectors using pesticides. However, with the growing resistance of insect vectors to pesticides and increased health concerns regarding the continued use of pesticides on food crops, the demand for a critical reassessment of strategies used in viral disease management has revolutionized the fight against plant viruses. One of the most promising approaches is the use of clean planting materials to boost productivity. Reyes *et al.,* (2006) demonstrated that the virus-free plants produced a 25% higher yield than the virus-infected plants in an experiment comparing the field performance between dasheen mosaic virus-free and virus-infected *in vitro* plants of cocoyam (*Xanthosoma* spp.) in Nicaragua. This was so even though most of the virus-free plants were re-infected during the experimental period. Similarly, studies on *Xanthosoma Sagittifolia* have shown that yield gains of more than seven times are possible with virus-free plants (Valverde *et al.*, 1997). These reports

demonstrate the benefits of using virus-free planting materials as it leads to increased yields. This demonstrates the need for virus-free planting materials in cropping systems. Elimination of plant viruses has been successfully achieved through one or combinations of two or more in vitro techniques (Awan *et al.*, 2007). Five methods are currently used for virus elimination: meristem culture, thermotherapy, chemotherapy, cryotherapy, and electrotherapy. Elimination of viruses using these techniques alone or in combination has been attempted for different crops such as banana, yam, potatoes, plums, and sweet potatoes among others with different degrees of success. Asami *et al*. (2013) reported the successful elimination of badnavirus from taro collected from Burundi using meristem culture followed by thermotherapy treatment of all positive *in vitro* plantlets for 20 days at a daily temperature of 38°C, photoperiod of 16 hours and 8 hours of the dark period at 28°C, intensity continued light of 5000 lux, and 70% relative humidity. However, the success of these virus elimination techniques is dependent on the correct identification of the viruses present and the development of highly specific high through put assays for their detection. This is critical for virus indexing systems used in virus cleaning programs as well as breeding for resistance against these viruses. As such, proper identification and characterization of taro viruses are vital in unlocking other efficient virus management and control strategies.

CHAPTER THREE

METHODOLOGY

3.1 Study Design

This study was cross-sectional based research that aimed to generate baseline information on viruses infecting taro in Kenya. The data generated provide information necessary for formulating effective management and control strategies for taro viruses through the development of an effective system to produce clean (virus-free) planting materials, development of specific high throughput diagnostic assays to determine the incidence and distribution of the viruses across the country and development of innate antiviral defense strategies that are more effective in viral disease control.

3.2 Viral Survey and Sample Collection

Viral surveys and sampling were conducted between August and October 2017 in Kiambu, Murang'a, Meru, Nyeri, Siaya, Busia, Kakamega, Kisii, and Machakos counties of Kenya (Figure 3.1) to determine the prevalence and distribution of viruses affecting Taro.

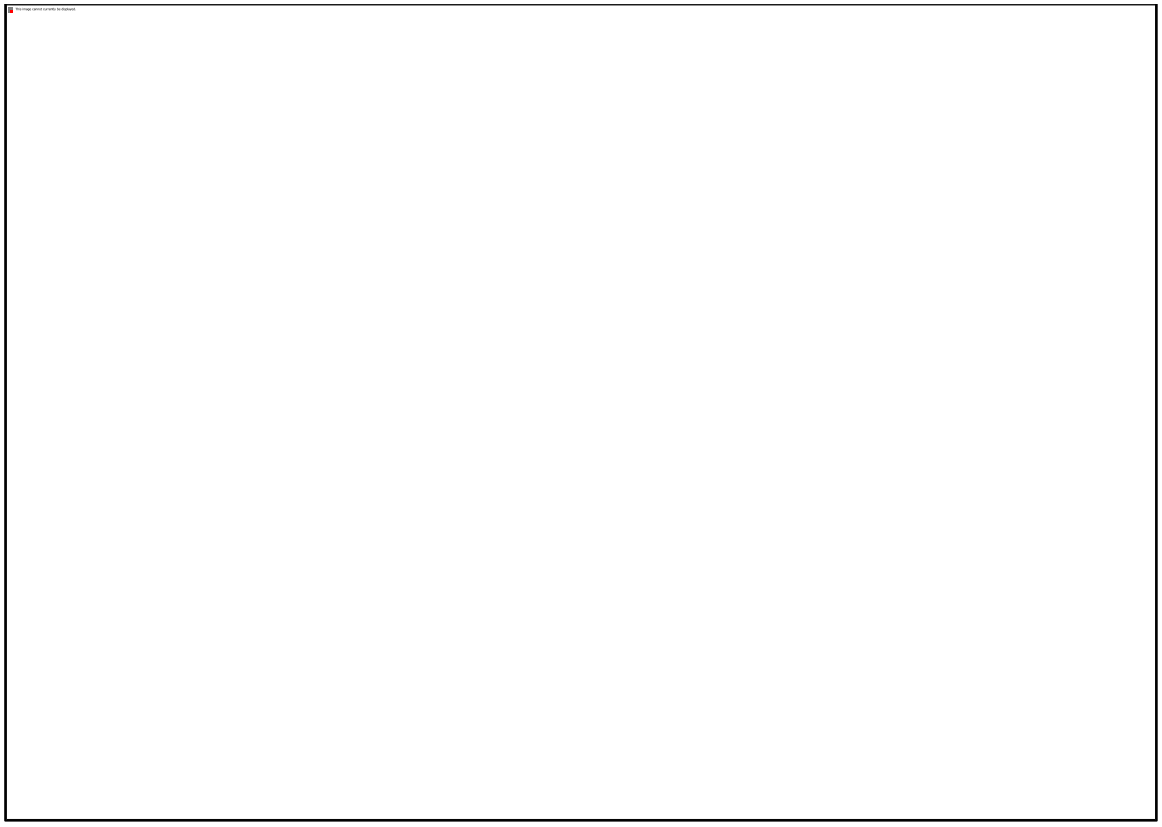


Figure 3.1: Kenyan Map Showing the Nine Counties (Highlighted in Green and Labeled) where Viral Survey and Sampling were done

In each county, viral surveys were done in ten fields across various administrative units (wards) and the viral disease prevalence was recorded. Virus symptomatic leaf samples and young suckers of edible and wild Taro were also collected for virus screening. Table 3.1 highlights the wards where the viral survey was done in each county. The geographical locations of the farms where the viral survey and sample collection was done are outlined in Appendix 12. The leaf samples were preserved in tubes containing silica gel upon collection, while the suckers were placed in plastic bags to prevent dying off. The leaf samples were then transported to the BecA-ILRI hub laboratory in Nairobi, Kenya, for *in vitro* laboratory analysis. The suckers collected from the various farms were planted in potted autoclaved soil containing compost and were subsequently monitored for viral symptoms for two months in the screen house. These plants were used for Small RNA work.

Table 3.1: Wards where the Viral Survey was done

County	Wards where the viral survey was done
Kiambu	Ngewa, Kikuyu, Karai, Gitaru, Nyathuna
Murang'a	Ngenda, Kigumo
Meru	Mikinduri, Nyaki, Kiguchwa, Kiamurio, Nkomo, Mitunguu, Nyagene, Maraa
Nyeri	Ruguru, Iriaini, Karatina, Kirimukuyu
Siaya	Central Alego, Yala township
Busia	Bunyala Central, Hajula
Kakamega	Mahiakalo, Butsotso East, Ingotse Matia
Kisii	Obaracho, Nyakoe
Machakos	Kathiani
Total	**30 wards**

3.3 Data Collection in the Field

The prevalence of Taro viruses-like diseases was determined through visual aid by counting plants showing viral symptoms, including stunting, leaf rolling, mosaic, and down curling of the leaf blades in every ten plants picked randomly across each field.

Viral prevalence in each farm was determined using the formulae:

P = (n/N) * 100% where; P = prevalence; n = number of plants with viruses-like symptoms; N (10) = total number of plants assessed.

3.4 Nucleic Acids Extraction, PCR, and RT-PCR

Total plant DNA and RNA were extracted from 290 leaf samples using the Qiagen DNeasy Plant Mini Kit and the Qiagen RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA), respectively. The quality and quantity of the DNA and RNA were checked using 1.0% Agarose gel electrophoresis and the NanoDrop® ND-1000

Spectrophotometer (Thermo Scientific, USA). The polymerase chain reaction (PCR) technique was used to detect the DNA virus *Taro bacilliform virus* (TaBV), while the reverse transcription-polymerase chain reaction (RT-PCR) technique was used to detect the RNA viruses, i.e., *Dasheen mosaic virus* (DsMV), *Colocasia bobone disease virus* (CBDV), *Taro vein chlorosis virus* (TaVCV) and *Taro reovirus*, using degenerate primers. The extracted DNA and RNA were standardized to 50ng/µl and then subjected to PCR and RT-PCR, respectively. The total volume for the PCR reaction was 20µL, composed of 10µL 2× Master Mix, 0.6µL forward primer (10µM), 0.6µL reverse primer (10µM), 5µL DNA template, and 3.8µL PCR grade water. The primer pairs used to amplify TaBV, DsMV, CBDV, and TaVCV and *Taro reovirus* were TaBV1/4 (Yang et al., 2003b), DsMV 3F/3R (Maino, 2003), CBDVF1/R2 (Revill et al., 2005), and TaVCV1/2 (Revill et al., 2005), respectively. Amplified products were stained with 2.5µl of gel red per 100ml of Tris-borate-EDTA buffer, electrophoresed through a 1.5% agarose gel, and visualized using a UV trans-illuminator (Syngene™ Ingenius 3 Manual Gel Documentation System, Thermo Scientific, USA) to detect the expected band. PCR amplicons were submitted for sanger sequencing, and the resultant sequences were analyzed/cleaned, and assembled using CLC Genomics Workbench v 8.0.3 (https://digitalinsights.qiagen.com/).

3.5 TaBV Sequencing and Phylogenetic Analysis

Direct sequencing of TaBV-positive samples was done in both directions using the forward and reverse primers; TaBV-1 (5'-CKSTGYAARSAACATGGTCTTG-3') and TaBV-4 (5'-TAATCAAGYGGWGGGAGYTTCTC-3') using the Big Dye Terminator v. 3.1 Cycle Sequencing kit (Thermo Fisher Scientific) at the Biosciences Eastern and Central Africa Hub (BecA-ILRI hub) laboratory in Nairobi, Kenya. Sequencing data were processed and analyzed using CLC Genomics Workbench v 8.0.3 (https://digitalinsights.qiagen.com/). Sequences were compared to other TaBV sequences in the NCBI database using BLAST algorithms available on the NCBI website (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Virus sequences were further aligned and analyzed with the Classical sequence alignment tool in CLC. Phylogenetic trees were constructed from CLC-aligned sequences on create tree tool on CLC using the maximum-likelihood method and a Jukes-Cantor parameter model with 1000

bootstrap replications. Pair wise sequence comparison (PASC) was carried out on aligned sequences using the create pairwise comparison tool on CLC Genomics Workbench v 8.0.3 (https://digitalinsights.qiagen.com/).

3.6 Small RNA Extraction, Library Preparation, and Sequencing

A representative sample of 48 leaf samples comprising 27 edible symptomatic Taro, 9 edible asymptomatic Taro, and 12 wild Taro relatives was picked from the potted plants is the screen house for virus detection and identification upon Small RNAseq. Small RNA extraction was performed using the Qiagen MiRNeasy mini-Kit (Qiagen, Valencia, CA, USA) following the manufacturer's instructions. The Small RNA quality was checked using the Bioanalyzer nanochip (Agilent, Santa Clara, CA, USA), followed by small RNA library preparation using the NebNext Small RNA Library Prep Set for Illumina (New England Biolabs, Ipswich, MA, USA). Small RNA fragments were ligated with single-stranded adapters, first at the 3′-end, then at the 5′ end, followed by reverse transcription and PCR amplification to generate the DNA template colonies. The library was then sequenced using the Illumina Miseq sequencing platform.

3.7 Small RNA Sequence Processing and Virus Screening

The raw small RNA reads were pre-processed for adapter removal and quality filtering using Trimmomatic (version 0.39) (Bolger *et al.,* 2014). The clean fastq sequences were then converted to fasta format using the seqkit toolkit (Shen *et al.,* 2016). The reads were then collapsed into single unique sequences to remove duplicates using the FASTX short reads processing toolkit (FASTX-Toolkit: FASTQ/a short-reads pre-processing tools; http://hannonlab.cshl.edu/fastx_toolkit). The small RNA reads were then searched against the NCBI non-redundant nucleotides (nt) database using BLASTn algorithm set at default parameters on the NCBI website (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to identify possible viruses in the samples.

CHAPTER FOUR

RESULTS

4.1 Symptoms and Field Prevalence Data of Taro Viral Diseases

Data collection in the field was mainly through visual observation. Plants showing viral symptoms, including stunting, leaf rolling, shrinkage, deformed leaves with mosaic and yellow veins, down curling of the leaf blades, and dwarfism (Figure 4.1) in every ten plants picked randomly across each field, were counted, and recorded.

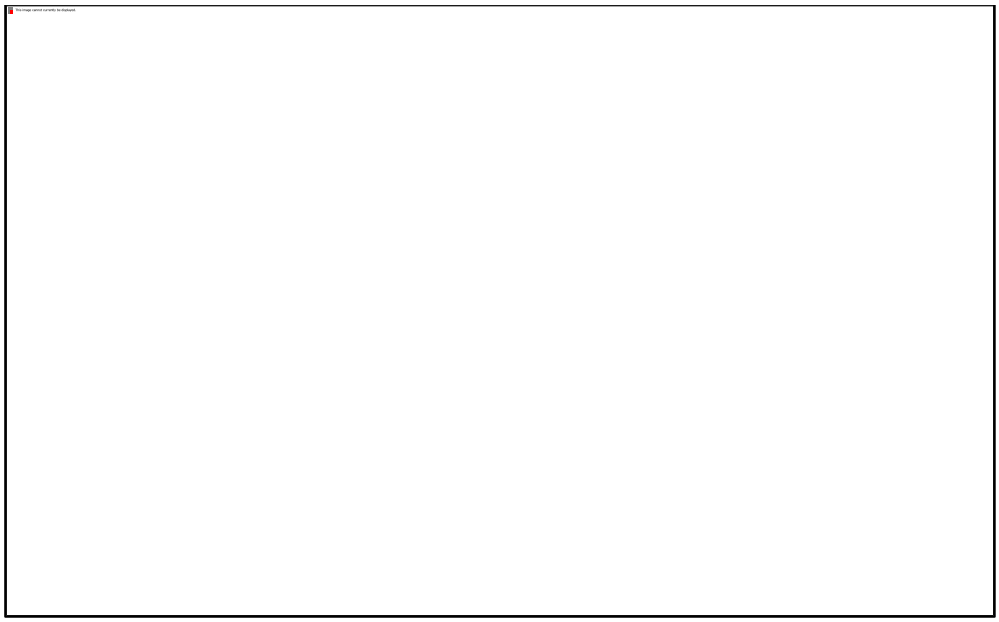


Figure 4.1: Symptoms of Taro Viral Diseases in the Field. (A & B) Deformed Leaves with Mosaic Formations; (C) Yellowing of Veins; (D) Conspicuous Yellowing of the Leaf with Brown Spots; (E) Leathery Leaf; (F) Browning and Drying of the Leaf Margin and Blades

Viral symptoms were common in all fields surveyed, affecting 32-60% of the surveyed plants. Nyeri county had the highest average prevalence at 60%, with a prevalence range of 40% to 90% in the fields surveyed. In contrast, Murang'a and Meru counties had the least average prevalence at 32%, with an incidence range of 20% to 60% (Figure 4.2). The other seven counties had an average prevalence of between 38% and 57%, with field prevalences ranging between 30% and 70% (Appendix I-IX). Out of

the 90 taro fields surveyed, 43 fields had a prevalence of more than 50%, while the rest had prevalences ranging between 20 and 49%. Notably, all farms surveyed, except one, had no history of chemical use and all farmers used manure to fertilize their farms (Appendix XII).

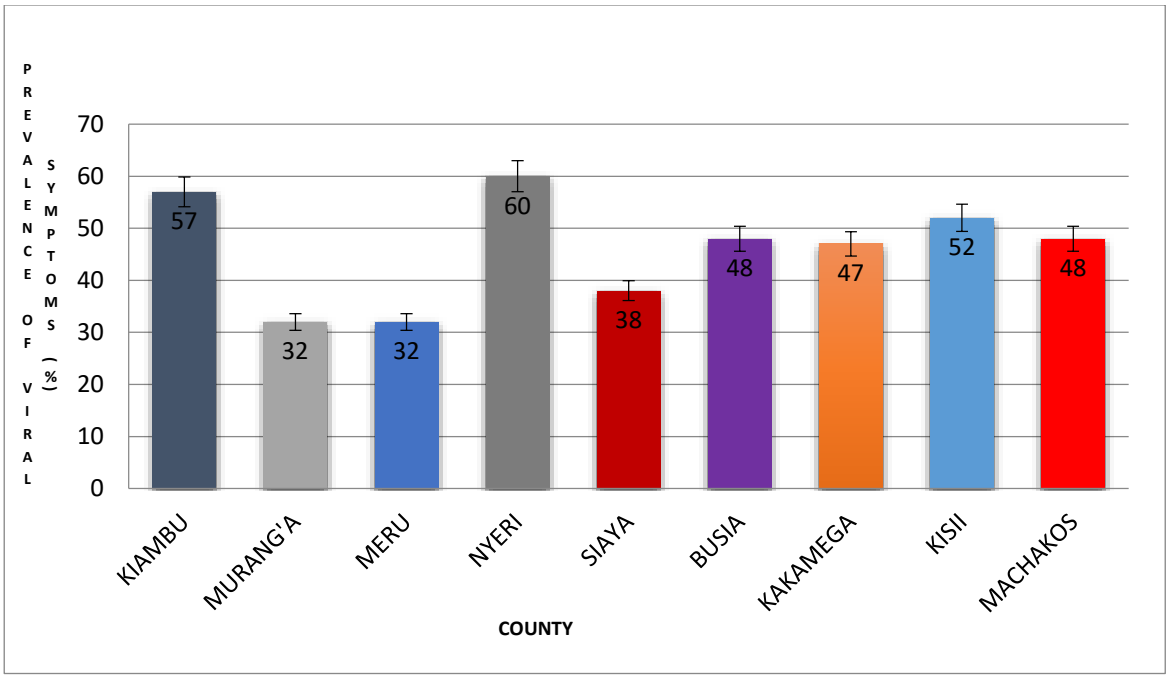


Figure 4.2: Graph Showing the Viral Prevalence among the Nine Counties Surveyed. The Error Bars were Calculated Based on the Average Prevalence Per County

4.2 Detection of Viruses by PCR and RT-PCR

DNA and RNA samples were subjected to PCR and RT-PCR to screen for five viruses, including *Dasheen Mosaic Virus* (DsMV), *Taro Bacilliform Virus* (TaBV), *Colocasia bobone disease virus* (CBDV), *Taro vein chlorosis virus* (TaVCV), and *Taro reovirus* (TaRV), reported to infect taro in other countries and regions. Among the five viruses, only *Taro Bacilliform Virus* (TaBV) was detected through PCR. An amplicon of 320bps, which was the expected band size for TaBV, was obtained for the positive samples (Figure 4.3).

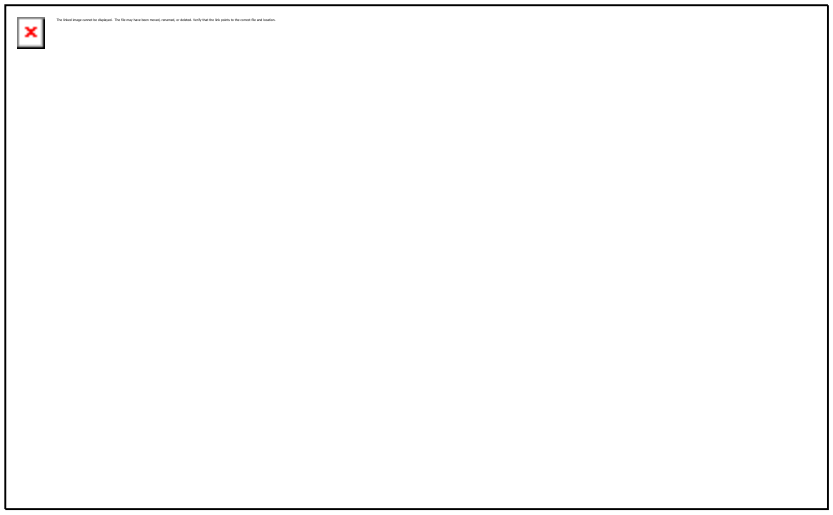


Figure 4.3: Gel Image Showing Amplicons of the Expected Size, 320bps of TaBV Positive Samples. The Ladder is the Thermo Scientific Generuler 1 Kb Plus DNA Ladder

Detections were made in samples collected across the nine counties at varying incidences. Nyeri County had the highest incidence at 63%, while Machakos County had the least at 23%. Of the 270 samples from edible taro, 115 tested positive, with positive samples identified in all the 30 wards surveyed across the nine counties (Table 4.1). Notably, all 20 samples collected from wild relatives of taro, including tannia, tested negative for TaBV.

Table 4.1: Summary of TaBV-Positive Samples Based on the PCR Assay in the Various Sampling Locations

County	Number of Samples collected	Samples that tested positive for TaBV	% Positive
Kiambu	30	14	47
Murang'a	30	15	50
Meru	30	14	47
Nyeri	30	19	63
Siaya	30	14	47
Busia	30	8	27
Kakamega	30	16	53
Kisii	30	8	27
Machakos	30	7	23
Total	**270**	**115**	**43**

4.3 Sanger Sequencing, Phylogenetic Analysis, and Pairwise Sequence Comparison of TaBV

Fifteen TaBV amplicons spread across the nine counties, i.e., Kiambu - 1, Murang'a – 2, Meru – 2, Nyeri – 3, Siaya - 2, Busia - 1, Kakamega – 1, Kisii, and Machakos – 1, were randomly selected for sanger sequencing of the 320 bp conserved RT/RNase H-coding sequence of TaBV ORF 3 for comparison with other TaBV viral sequence data, especially from East Africa. BLAST analysis of the consensus sequences revealed that all 15 isolates showed the highest nucleotide identity (94-99%) to four TaBV isolates from East Africa, i.e., Tz24 (MG833013), Ug75 (MG017323), Tz17 (MG017322), and Ke52 (MG017321) (Appendix X). Of note, the Tz24 isolate had been isolated from tannia, a wild taro relative in Tanzania. The parameters of blastn program were: Expect value = 0.05, Match = 1, Mismatch scores = -2, and Gapcosts = 0,2.5, with the low complexity filter set to Yes.

All the 15 TaBV isolates had the highest nucleotide identity to Tz24 (MG833013) but at varying percentages (Table 4.2).

Table 4.2: The percentage Identities of the 15 TaBV Isolates to Tz24 (MG833013)

Isolate	Accession	Description	Percentage Identity	E-value
Kiambu1	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	98.792	1.23E-163
Siaya10	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	96.108	1.16E-148
Busia12	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	98.799	9.39E-165
Kakamega17	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	93.75	1.58E-137
Murang'a2	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	94.828	1.22E-148
Kisii22	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	93.983	1.40E-143
Machakos24	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	96.875	6.80E-147
Murang'a3	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	94.828	1.25E-148
Machakos32	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	95.415	5.10E-153
Meru4	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	97.329	4.27E-158
Meru5	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	97.399	1.23E-163
Nyeri6	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	97.289	7.43E-156
Nyeri7	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	94.062	1.99E-131
Nyeri8	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	97.289	7.56E-156
Siaya9	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	96.073	4.10E-148

Phylogenetic analysis of the 15 conserved RT/RNase H-coding sequences of ORF 3 revealed that the isolates formed two major clades. Ug75, Tz24, Ke52, and Tz17 isolated from east Africa clustered with those from Kakamega, Siaya, and Kiambu and one from Nyeri away from isolates from the other counties (Figure 4.4). Notably, no isolates from any county formed a distinct group in this tree but the two isolates from Murang’a county showed a close relationship.

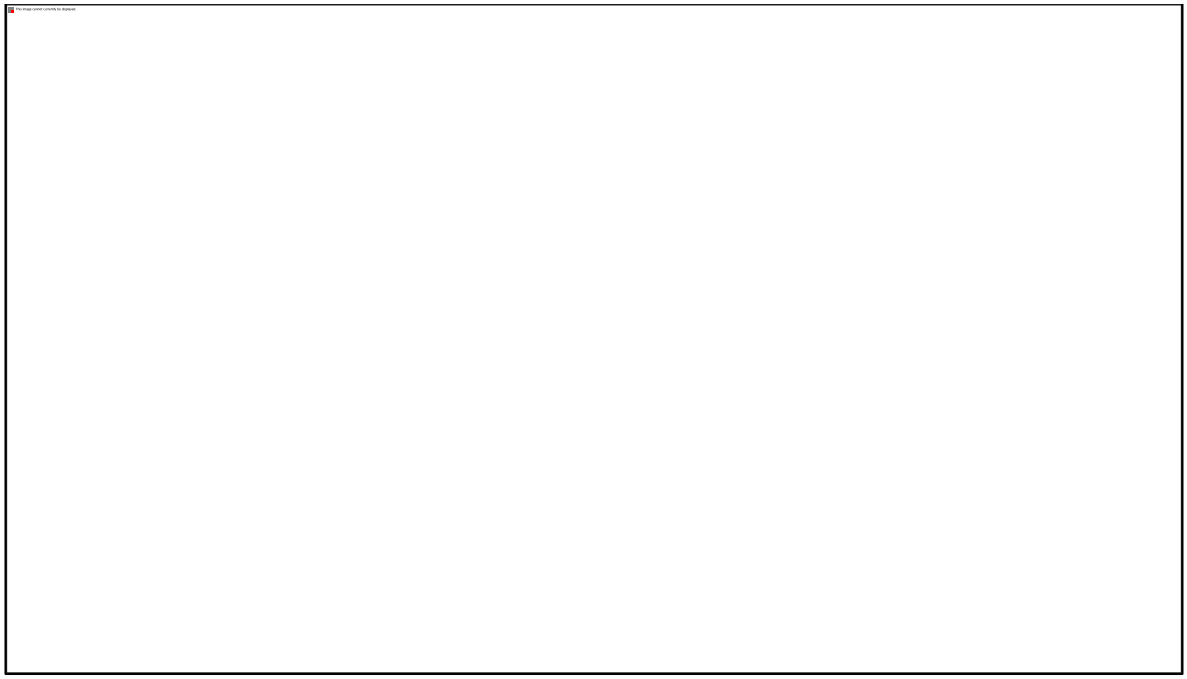


Figure 4.4: Phylogram of the 15 Isolates and 4 TaBV Isolates from East Africa (all with Bootstrap Numbers) Based on the 320bp RT/Rnase H-Coding Sequence ORF 3 of TaBV. The Maximum-Likelihood Method and a Jukes-Cantor Parameter Model with 1000 Bootstrap Replications were Employed.

Pairwise sequence comparison of the 15 sequences revealed a 33-94% sequence variability of the putative reverse transcriptase (RT)/ribonuclease H (RNase H) coding region of the TaBV isolates. Isolates from Nyeri county had the highest nucleotide sequence variability of up to 58% while those from Murang'a had the least variability of up to 8%. Isolates from Siaya, Kakamega, Kiambu, and one Nyeri isolate had the greatest nucleotide sequence identity ranging from 84% to 92% while those in the rest of the counties had identities ranging between 71% and 94% (Figure 4.5). Sequences of the 15 isolates have been deposited in the National Center for Biotechnology Information (NCBI) database accession numbers ON853594 – ON853608; Kiambu 1 (ON853594), Siaya 10 (ON853595), Busia 12 (ON853596), Kakamega 17 (ON853597), Murang'a 2 (ON853598), Kisii 22 (ON853599), Machakos 24 (ON853600), Murang'a 3 (ON853601), Machakos 32 (ON853602), Meru 4 (ON853603), Meru 5 (ON853604), Nyeri 6 (ON853605), Nyeri 7 (ON853606), Nyeri 8 (ON853607), and Siaya 9 (ON853608).

Figure 4.5: Pairwise Sequence Comparison of the 15 TaBV Sequences

4.4 Small RNA Sequencing of Symptomatic and Asymptomatic Taro Plants

Agarose gel electrophoresis of the small RNA samples revealed that they were of good quality and quantity (Figure 4.6).

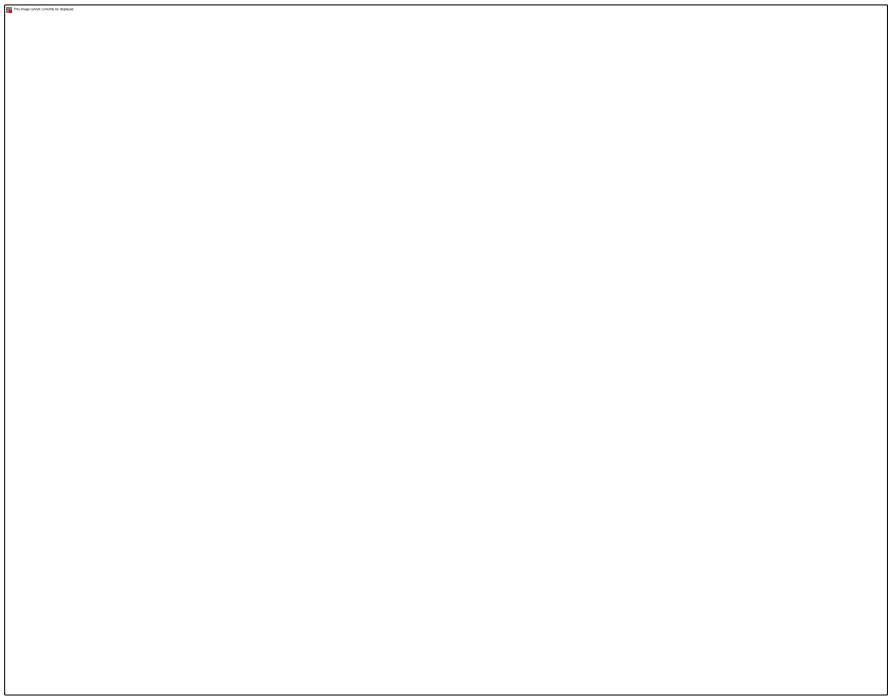


Figure 4.6: Gel Image Showing Small RNAs of Representative Samples with a Size of Around 100 Bps. The Ladder is the Thermo Scientific Generuler 1 Kb Plus DNA Ladder

Subsequent library preparation using the NebNext Small RNA Library Prep Set kit yielded good quality and quantity of the small RNA libraries embodied by having a conspicuous band of 150bp in a 2% agarose gel (Figure 4.7) and Agilent 2100 bioanalyzer electropherogram peaks ranging between 141 and 153 (Figure 4.8), highlighting an enrichment of small RNAs in the libraries.

Figure 4.7: Gel Image Showing Small RNA Libraries of Representative Samples with a Conspicuous Band at 150 Bp. The Ladder is the Thermo Scientific Generuler 1 Kb Plus DNA Ladder

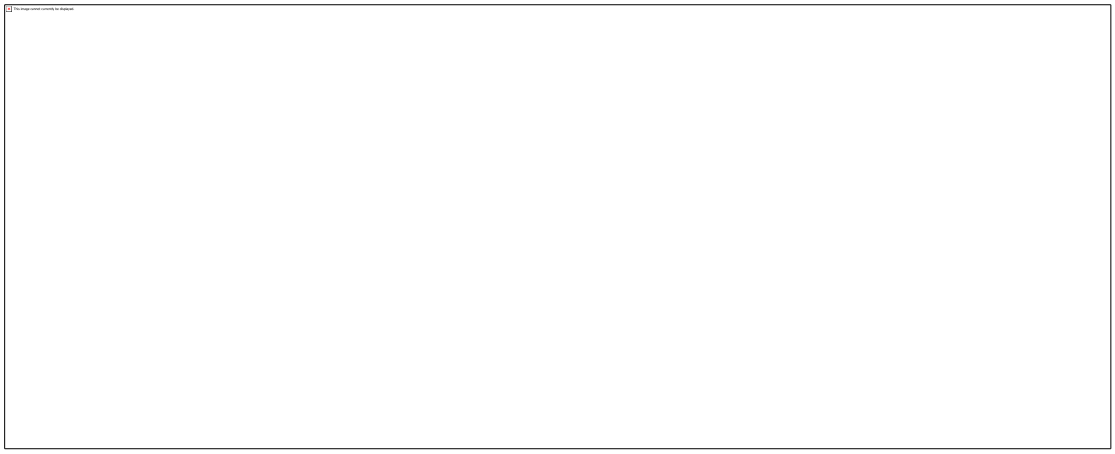


Figure 4.8: Aligent 2100 Bioanalyzer Electropherogram Trace of Representative Small RNA Libraries (A) Sample T9 and (B) Sample T3 Showing Desired Peaks of 141 and 143, Respectively.

High throughput sequencing of libraries from edible and wild Taro conducted on an Illumina Miseq platform yielded a total of 25,707,591 small RNA reads. Subsequent trimming of adapters and low-quality reads left 24,224,023 clean small RNA reads comprising 17,642,010, 6,552,208, and 29,805 reads from symptomatic,

asymptomatic, and wildtype samples, respectively. The length of the clean reads ranged between 35 and 70 base pairs (Figure 4.9).

Figure 4.9: Graphical View of the Percentage Ratios of the Lengths (Base Pairs) of the Clean Small RNA Reads

The clean reads were analyzed independently by blasting them against the NCBI non-redundant nucleotides (nt) database. Notably, there were detections of the *Taro Bacilliform Virus* and *Taro Bacilliform CH Virus*, which are badnaviruses specific to Taro, the *Colocasia Bobone Disease Virus*, a rhabdovirus specific to Taro in both the symptomatic and asymptomatic samples. Other badnaviruses identified were: *Sweet potato Badnavirus B* and *Sugarcane bacilliform virus*. The S*weet potato leaf curl virus* and *East African Cassava mosaic virus* were the only begomoviruses, while the *Sweet potato feathery mottle virus* was the only potyvirus detected. *Phaseolus vulgaris alphaendornavirus*, a rhabdovirus, was predominant in all samples except in the wild Taro relatives, while the *Citrus exocortis viroid* was the only viroid detected (Appendix XI). The viral hits were predominantly on the coding regions of the viruses and the viroid, highlighting the actual occurrence of the viruses and viroid in Taro (Table 4.4). Of note, there were no viral nor viroid hits amongst the clean reads obtained from the wild Taro relatives.

Table 4.4: Representative Viral Hits Showing the Predominance of the Query Sequences in the Coding Regions of Viruses

Querry ID	Accession	Percentage Identity	E-value	Organism	Description
20123-79	gi|1343311831|gb|MF375892.1|	100	0.001	Viruses	Phaseolus vulgaris alphaendornavirus 1 polyprotein gene
20521-76	gi|1343311831|gb|MF375892.1|	100	0.001	Viruses	Phaseolus vulgaris alphaendornavirus 1 polyprotein gene
21054-73	gi|1343311831|gb|MF375892.1|	100	0.001	Viruses	Phaseolus vulgaris alphaendornavirus 1 polyprotein gene
1924571-1	gi|1587104515|gb|MH142485.1|	100	0.000109	Viruses	Sugarcane bacilliform virus isolate HNSb4 polyprotein gene
1924571-1	gi|1587104523|gb|MH142489.1|	100	0.000109	Viruses	Sugarcane bacilliform virus isolate HNSb8 polyprotein gene
1213887-1	gi|770114323|gb|KM009090.1|	100	0.001	Viruses	Sweet potato badnavirus A isolate KSR675 NORA-II(2) polyprotein gene
1809189-1	gi|769469209|gb|KM000051.1|	100	0.0000311	Viruses	Sweet potato badnavirus A isolate Tanzania Carrot C-1(2) polyprotein gene
1809189-1	gi|769469211|gb|KM000052.1|	100	0.0000311	Viruses	Sweet potato badnavirus A isolate Tanzania Carrot C-2(2) polyprotein gene
1809189-1	gi|770114343|gb|KM009100.1|	100	0.0000311	Viruses	Sweet potato badnavirus B isolate 28-Kawogo Uganda (2) polyprotein gene
1213887-1	gi|770114341|gb|KM009099.1|	100	0.001	Viruses	Sweet potato badnavirus B isolate 28-Kawogo Uganda polyprotein gene
1623560-1	gi|156857638|gb|EF591125.1|	100	8.53E-16	Viruses	Sweet potato begomovirus isolate BG9 movement protein (V2)
1623560-1	gi|270268864|gb|GQ268223.1|	100	5.17E-08	Viruses	Sweet potato begomovirus strain SPBV-D-DNA-[1]
1327058-1	gi|73852992|emb|AJ811971.1|	100	2.34E-26	Viruses	Sweet potato chlorotic stunt virus CP gene for coat protein
1327058-1	gi|73852994|emb|AJ811972.1|	100	1.09E-24	Viruses	Sweet potato chlorotic stunt virus CP gene for coat protein

Querry ID	Accession	Percentage Identity	E-value	Organism	Description
1327058-1	gi|115342812|gb|DQ864344.1|	100	1.09E-24	Viruses	Sweet potato chlorotic stunt virus isolate 115-2S coat protein (CP) gene
3582729-1	gi|2394166|gb|AF015541.1|AF015541	100	0.000403	Viruses	Sweet potato feathery mottle virus coat protein (CP) gene
399778-2	gi|295841641|dbj|AB509459.1|	100	1.78E-22	Viruses	Sweet potato feathery mottle virus gene for polyprotein
934417-1	gi|295841649|dbj|AB509463.1|	100	1.78E-22	Viruses	Sweet potato feathery mottle virus gene for polyprotein
2495520-1	gi|1399734333|gb|MG017325.1|	100	0.000000669	Viruses	Taro bacilliform CH virus isolate Ke43
3020826-1	gi|1399734333|gb|MG017325.1|	100	3.02E-10	Viruses	Taro bacilliform CH virus isolate Ke43
1639466-1	gi|1389516188|gb|MG017359.1|	100	0.001	Viruses	Taro bacilliform CH virus isolate Ug52 ORF3 gene
1658564-1	gi|1389516188|gb|MG017359.1|	100	0.000000182	Viruses	Taro bacilliform CH virus isolate Ug52 ORF3 gene
3770473-1	gi|1389516188|gb|MG017359.1|	100	0.001	Viruses	Taro bacilliform CH virus isolate Ug52 ORF3 gene

4.5 The Relationship between Virus Detection and Disease and Co-infection of Taro in Kenya

The symptoms observed on infected leaves varied significantly and ranged from stunting, leaf rolling, shrinkage, deformed leaves with mosaic and yellow veins, down curling of the leaf blades, and dwarfism (Figure 4.1). There was a clear association between virus presence and symptoms in Taro plants as evidenced by a correlation of viral disease prevalence and TaBV detection in the various counties. There were, however, cases in which plants were asymptomatic but viruses were detected through NGS. Both symptomatic and asymptomatic samples had DNA and RNA virus hits in the same sample. *Taro Bacilliform Virus* and *Taro Bacilliform CH Virus* were the most predominant viruses in both symptomatic and asymptomatic plants, occurring in nearly all samples, and were mostly coupled with the *Colocasia Bobone Disease Virus*. Other combinations, including the TaBV/TaBCHV and the other rhabdovirus,

potyvirus, begomoviruses, or all were also observed mostly in the symptomatic samples. This finding strongly suggested co-infection of taro with multiple viruses.

CHAPTER FIVE

DISCUSSION, CONCLUSIONS, AND RECOMMENDATIONS

5.1 Discussion

Viral diseases cause severe yield losses and quality decline in Taro worldwide, thus having a huge economic impact on farmers (Babu *et al.,* 2011). In Kenya, the occurrence and distribution of the major viruses and viroids infecting taro remain poor, limiting the development of robust disease management strategies to mitigate their spread. Herein, viral symptoms, including stunting, leaf rolling, shrinkage, deformed leaves with mosaic and yellow veins, leaf rolling, and dwarfism (Yosop *et al.*, 2019), were common in all taro fields surveyed, affecting 32-60% of the surveyed plants. Notably, all farms surveyed, except one, had no history of chemical use and all farmers used manure to fertilize their farms, further highlighting the low efforts put towards Taro production. The prevalence and distribution data obtained in this study demonstrated that taro viral infection are both highly diverse and prevalent in Kenya, necessitating a need for in-depth studies on the virus diversity in Kenyan Taro fields. This study employed small RNA sequencing and PCR techniques to identify viruses infecting taro in Kenya and is the first to report a comprehensive occurrence and distribution of the major viruses infecting taro in Kenya. Of note, TaBV was distributed in all 30 wards across the nine counties surveyed. To date, TaBV appears to be host specific mainly transmitted through vegetative propagation, especially through sharing of diseased planting materials and mealybugs in a semipersistent manner (Gollifer *et al.,* 1977; Macanawai *et al.,* 2005). TaBV infection has been associated with symptoms such as downward curling of the leaf blades, vein clearing, and stunting in some cultivars (Yang *et al.,* 2003a; Revill *et al.,* 2005; Kidanemariam *et al.,* 2018). Though Yang *et al.* (2003b) and Revill *et al.* (2005) reported that there is no correlation between symptoms and the presence of TaBV, this study shows that there is indeed a direct correlation between the two because the high prevalence of TaBV symptoms in all the fields surveyed was consistent with the findings of this study that revealed extensive detection of TaBV in both symptomatic and asymptomatic plants. However, proper documentation of these symptoms should be

made to avoid inconsistencies that may be caused by mixed infections because it is a common occurrence in Taro (Revill *et al.,* 2005). The sequence variability in the putative reverse transcriptase (RT)/ribonuclease H (RNase H) coding region from *Taro bacilliform virus* (TaBV) isolates from different counties ranged between 33% and 94%. This finding was inconsistent with that of Yang *et al.* (2003a) who reported a maximum nucleotide variability of 19.7% within the RT/RNase H-coding region of TaBV isolates from throughout the Pacific. The high variability was attributed to a second Badnavirus, *Taro Bacilliform CH Virus*, that was recently identified to infect taro (Kazmi *et al.,* 2015), and potentially other uncharacterized *Taro Bacilliform Virus* strains. The major difference between TaBV and TaBCHV genomes is in the number of ORFs. TaBV possesses four ORFs, all encoded on the plus-strand of the viral DNA, with the size and organization of ORFs 1-3 consistent with most badnaviruses (Yang *et al.*, 2003a). In contrast, TaBCHV encodes six putative ORFs, with ORFs 1-4 analogous to TaBV and an additional two small ORFs at the 3' end of ORF 3 (Kazmi *et al.*, 2015). Given that the RT/RNase H-coding region of ORF 3 is the most conserved region of the genome and a nucleotide (nt) difference of greater than 20% in this part of the genome is used for species demarcation in badnaviruses according to the International Committee on Taxonomy of Viruses (ICTV) (King *et al.,* 2012), the high sequence variability in the putative RT/RNase H-coding region from TaBV isolates from different counties alludes to the fact that there may be a third TaBV strain circulating in Kenyan Taro fields. This phenomenon is further supported by the detection of other badnaviruses, including the *sweet potato Badnavirus B* and the *sugarcane bacilliform virus* through small RNA sequencing.

Degenerate primers are limited in providing a true reflection of viral diversity in the field, especially when viruses have not been previously described or are remarkably divergent from those previously characterized (Wu *et al.,* 2015). Herein, the degenerate primers of already reported Taro viruses only detected TaBV and were thus not adequate to conclude viral diversity in the tested samples.

This study thus further employed Next-generation sequencing (NGS) of small RNAs to broaden the possibility of Taro virus diagnostics because of its superiority in detecting multiple viruses present even in the absence of disease symptoms (Adams

et al., 2009; Wu *et al.,* 2015). NGS technologies have significantly advanced our ability to comprehensively study plant viruses, especially in vegetatively propagated crops, such as Taro that are prone to virus accumulation and co-infections (Kreuze *et al.,* 2009). NGS has been applied for viral diversity studies, detection, and virus genome assembly and reconstruction. For instance, NGS has been successfully used to detect viruses in different crop plants, such as grapevines (Coetzee *et al.*, 2010; Jo *et al.,* 2015), sweet potatoes (Kashif *et al.,* 2012), tomatoes (Li *et al.,* 2012), garlic (Wylie *et al.,* 2014), pear (Jo *et al.,* 2016), pepper (Jo *et al.,* 2017), and orange (Matsumura *et al.,* 2017). NGS has also identified new viruses infecting sweet potatoes (Gu *et al.,* 2014).

In this study, small RNA sequencing revealed badnaviruses, begomoviruses, potyviruses, and rhabdoviruses, including the *Taro Bacilliform Virus*, *Taro Bacilliform CH Virus*, which are badnaviruses specific to taro, and *Colocasia Bobone Disease Virus*, a rhabdovirus specific to taro to be the main viruses infecting taro in Kenya. Similar studies in the Pacific Islands report *Dasheen Mosaic Virus* (DsMV), *Taro Bacilliform Virus* (TaBV), Colocasia bobone disease virus (CBDV), *Taro vein chlorosis virus* (TaVCV), and *Taro reovirus* (TaRV) to be the main viruses infecting taro (Revill *et al.,* 2005; Yusop *et al.,* 2019).

Studies in Asian countries report that the *Dasheen Mosaic Virus* (DsMV), a potyvirus, and *Taro Bacilliform Virus* (TaBV), a badnavirus, are the most predominant Taro viruses (Revill *et al.*, 2005; Yusop *et al.,* 2019). Interestingly, Zettler and Hartman, (1986 & 1987), Jackson, (1980), and Shaw *et al.* (1979) reported that DsMV is found wherever Taro is grown, infecting both the edible and ornamental aroids, and is characterized by chlorotic and feathery mosaic patterns on the leaf, distortion of leaves, and stunted plant growth. In this study, DsMV was not detected using degenerate primers and small RNA sequencing, but the associated symptoms were common in some fields surveyed. However, the *sweet potato feathery mottle virus*, a potyvirus, was predominantly detected in both symptomatic and asymptomatic plants, strongly suggesting a potyvirus(es) was in circulation in taro growing areas in Kenya. This finding supported the speculations regarding the extent of potyviruses present in taro

due to cross-species transmission via mechanical transmission or aphid vectors that feed on different host plants (Yusop *et al.*, 2019).

TaBV is thought to occur in combination with CBDV to cause “alomae” disease (James *et al.*, 1973), which is considered taro's most destructive virus disease (Jackson and Gollifer, 1975; Rodoni *et al.*, 1994). Herein, TaBV and TaBCHV, which are badnaviruses specific to taro, and CBDV, a rhabdovirus specific to taro, were detected through small RNA sequencing. They mostly occurred as co-infections in the analyzed samples, plus two to three other RNA and DNA viruses. Alomae symptoms including crinkling of young leaves that fail to develop normally, thickening of veins and lamina, shortening of the petioles, and irregularly shaped outgrowths on the petioles, which cause plants to ultimately die because of the development of systemic necrosis (Rodoni *et al.*, 1994), were not common in the fields surveyed. However, there were numerous cases of stunting, mosaic, and down-curling of the leaf blades, which are associated with infection with TaBV alone, and leaf distortion and the presence of galls on the petioles (Jackson, 1978), which are associated with infection with CBDV alone to cause bobone disease. Combined inferencing of the field symptoms and the sequencing results suggest the potential development of alomae disease in Kenyan taro fields in the future or its presence in other regions that were not surveyed.

The PCR-based detection of TaBV in all the nine counties surveyed and subsequent detection of *Taro Bacilliform CH Virus* and other badnaviruses, including the *sweet potato Badnavirus B* and the *Sugarcane bacilliform virus*, confirmed the wide host range of badnaviruses, especially in perennial hosts that are propagated vegetatively (Bhat *et al.*, 2016). These findings were consistent with other studies which postulate that TaBV is widespread in almost all taro-growing regions, including Kenya (Kidanemariam *et al.*, 2018), and its occurrence alongside a putative rhabdovirus, CBDV, which leads to the lethal Alomae disease (Yang *et al.*, 2003; Higgins *et al.*, 2016).

To date, there is no report of begomoviruses infecting Taro worldwide. However, small RNA sequencing and subsequent blasting of the sequence reads strongly suggested that begomoviruses(es) could be circulating in taro fields in Kenya after

several begomoviruses, including the *East African Cassava mosaic virus* and *sweet potato leaf curl virus*, were predominantly detected in both symptomatic and asymptomatic plants. Begomoviruses are single-stranded DNA plant viruses transmitted by whiteflies of the *Bemisia tabaci* complex. They are important groups of emerging plant viruses infecting numerous vegetables, root, and fiber crops in subtropical and tropical regions (Navas-Castillo *et al.*, 2011). Their symptoms include leaf curling, mosaic, vein yellowing, and generalized leaf yellowing, often accompanied by stunting (Leke *et al.*, 2015). These symptoms were common amongst the taro fields surveyed, further suggesting the possibility of having a begomovirus(es) circulating in Kenyan taro. Begomoviruses, such as the *East African cassava mosaic Kenya virus* (EACMKV), have already been reported in Kenya (Bull *et al.*, 2006). The mixed cropping systems that are common in the small holder farms where most of the samples were collected could also have caused the transfer of some virus particles from the host plants, such as cassava and sweet potatoes, to taro as an alternate host through insect vectors. Notably, begomoviruses have been reported to cause significant yield losses in other root crops. For example, cassava mosaic diseases in Sub-Saharan Africa cause yield losses exceeding $2 billion annually (Thresh *et al.*, 1997). This discovery is the first report of a begomovirus(es) potentially infecting taro. Nonetheless, further studies should be done to identify whether there are specific begomoviruses infecting taro and their influence on yield. Viral detections in the asymptomatic plants suggested latency in some of these viruses or the lack of manifestation of the symptoms, possibly because of low viral titer.

Viroids are small, single-stranded, circular RNAs that induce specific diseases in higher plants despite lacking protein-coding capacity (Kovalskaya & Hammond, 2014). Their small size and distinct molecular structure make them potent molecular features for inducing resistance to viral pathogens through RNA silencing (Sano *et al.*, 2010). The presence of viroids affects disease severity and symptom manifestation in plants and are molecular vehicles for the introduction of diseases (Natalia *et al.*, 2014). Detection of a viroid in a host is thus useful because of its considerable economic importance. To date, there are no reports of viroids infecting taro. However, this study detected the Citrus exocortis viroid to infect taro. Viroids have been reported to cause diseases in many species, including *Solanum*, causing the tubers to be small, elongated,

distorted, and cracked. They are mainly transmitted through vegetative propagation and aphids (Owens *et al.*, 2009). Their symptoms are like those of many plant virus infections, including stunting, vein discoloration and clearing, chlorotic or necrotic spots, leaf distortion and mottling cankers, and tuber malformations (Natalia *et al.*, 2014). Of note, tuber malformation was a common decry from most farmers during our sample collection in the farmers' fields across the nine counties that were surveyed. These findings allude that Taro is also affected by viroids. However, whether this was a case of cross-species transmission without any significant impact or points to the existence of Taro viroids should be investigated using homology-independent approaches that combine deep sequencing of small RNAs with a computational algorithm.

Interestingly, the wild Taro relatives had no viral or viroid hits. These findings affirmed that wild species possess some level of tolerance to viral infections, possibly because of having a rich reservoir of resistance genes useful in breeding cultivars with a genetically controlled resistance against numerous diseases (Okoń *et al.*, 2021).

Generally, the viral symptoms observed during the survey were consistent with the viruses detected. Moreover, the viral severity in the fields was also positively correlated with the number of viruses identified, as evidenced by the positive correlation between the viral incidences based on the visual identification of symptoms and the PCR results. These results ascertained the major virus families infecting taro in Kenya and demonstrated the reliability of the sRNA deep sequencing data in determining virus and viroid diversity. This discovery is the first report of a viroid infecting Taro.

5.2 Conclusions

This study aimed to identify and characterize Taro viruses in Kenya. Specifically, the study aimed to:

1. To identify viruses and viroids infecting Taro in Kenya using PCR and RT-PCR assays and small RNA sequencing. This objective was largely met through small RNA sequencing and subsequent analyses, which revealed the

major viruses circulating in Kenyan Taro fields. *Taro Bacilliform Virus* and *Taro Bacilliform CH Virus* were the most predominant viruses. This finding is supported by the PCR and RT-PCR detection of Taro viruses reported in other countries and regions using degenerate primers which also revealed the presence of *Taro Bacilliform Virus* in all farms surveyed.

2. To characterize the major taro viruses circulating in Kenyan Taro fields. This objective was also met though with only one virus, i.e., the *Taro Bacilliform Virus* (TaBV). Sanger sequencing of 15 amplicons drawn from samples from nine counties and subsequent bioinformatics analysis using CLC Genomics Workbench v 8.0.3 revealed a 33-94% sequence variability in the putative reverse transcriptase (RT)/ribonuclease H (RNase H) coding region of the TaBV isolates. The abundance of badnaviruses further confirmed the wide host range of badnaviruses, especially in perennial hosts that are propagated vegetatively.

The successful achievement of these objectives constitutes the first comprehensive report of Taro viruses and viroids in Kenya and forms a basis for further studies, including genetic resources for virus-taro interactions, and insight for developing robust management strategies to mitigate their spread.

In conclusion, the results of this study support the rejection of the alternative hypothesis which stated that degenerate primers specific to Taro viruses can effectively determine the prevalence and distribution of viruses infecting Taro in Kenya. The results lead to acceptance of the null hypothesis that degenerate primers specific to Taro viruses cannot effectively determine the prevalence and distribution of viruses infecting Taro in Kenya and thus require the development of specific high throughput diagnostic assays to determine their prevalence and distribution.

5.3 Recommendations

5.3.1 Policy Recommendations

Taro production in Kenya is faced with a myriad of challenges despite its potential to ameliorate food and nutritional insecurity. Of great interest is the lack of disease-free

planting materials attributed to the lack of an effective virus cleaning system brought by the poor attention and neglection of taro and the lack of research funds allocated to Taro research. Consequently, farmers recycle and borrow planting materials from each other leading to viral diseases build-up and spread. In this regard, I recommend the following:

1. The government through the ministry of Agriculture and Livestock Development to develop a framework that will enable research on Taro to ensure up-to-date knowledge of Taro production statistics and agronomic requirements.
2. The Kenya Agricultural and Livestock Research Organization (KALRO) establishes a Taro seed production facility to ensure the availability of clean seed as well as improved Taro cultivars, which would promote Taro production in the country and propel Kenya to be a Taro exporter.

5.3.2 Recommendations for Further Research

The high incidence and co-existence of viruses previously not known to infect Taro require further exploration to determine their association, which potentially end up with super strains capable of limiting taro productivity. Future studies should focus on:

1. Developing effective management strategies to support the prevention and control of Taro viruses, including removing infected crops, controlling insect vectors, and developing virus-free planting materials.
2. Taro viroids, especially in viroid-induced RNA silencing as a transcriptional machinery for inducing resistance to viral pathogens, to decipher their role in plant-virus interactions and their potential use as effective modulators of taro defense mechanisms.

REFERENCES

Adams, I.P., Glover, R.H., Monger, W.A., Mumford, R., Jackeviciene, E., Navalinskiene, M., Samuitiene, M., & Boonham, N. (2009) Next-generation sequencing and metagenomic analysis: a universal diagnostic tool in plant virology. *Mol Plant Pathol 10*, 537–545.

Akwee, P.E., Netondo, G., Kataka, J.A., & Va, P. (2015) A critical review of the role of Taro Colocasia esculenta L. (Schott) to food security: A comparative analysis of Kenya and Pacific Island taro germplasm. *Scientia Agriculturae*, *9*, 101-8.

Anthony, M.B., Marc, L., & Bjoern, U. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data, *Bioinformatics*, *30*(15), 2114–2120.

Babu, B., Hegde, V., Makeshkumar, T., & Jeeva, M. L. (2011) Detection and identification of Dasheen mosaic virus infecting Colocasia esculenta in India. *Indian Journal of Virology*, 22(1), 59–62.

Bhat, A.I., Hohn, T., & Selvarajan, R. (2016). Badnaviruses: The Current Global Scenario. *Viruses*, *8*(6), 177.

Bull, S.E., Briddon, R.W., Sserubombwe, W.S., Ngugi, K., Markham, P.G., & Stanley, J. (2006) Genetic diversity and phylogeography of cassava mosaic viruses in Kenya. J. Gen. *Virol. 87*, 3053-3065.

Coetzee, B., Freeborough, M.J., Maree, H.J., Celton, J.M., Rees, D.J.G., & Burger, J.T. (2010) Deep sequencing analysis of viruses infecting grapevines: virome of a vineyard. *Virology*, *400*, 157–163.

Gollifer, D.E., Jackson, G.V.H., & Dabek, A.J. (1977) The occurrence and transmission of viruses of edible aroids in the Solomon Islands and the Southwest Pacific. *International Journal of Pest Management, 23,* 171–7.

Gu, Y.H., Tao, X., Lai, X.J., Wang, H.Y., & Zhang, Y.Z. (2014) Exploring the polyadenylated RNA virome of sweet potato through high-throughput sequencing. *PLoS ONE, 9,* e98884.

Gollifer, D.E. & Jackson, G.V.H. (1975) Disease and pest problems of Taro (*Colocasia esculenta* L. Schott) in the British Solomon Islands. *Pesticide Articles and News Summarie, 21*, 45-53.

Hadidi A., Sun, L., & Randles, J.W. (2022) Modes of Viroid Transmission. *Cells. Feb 18*;11(4):719.

Hammond, R.W. & Owens, R.A. (2006). Viroids. New and Continuing Risks for Horticultural and Agricultural Crops. *APSnet Feature article*.

Higgins, C.M., Bejerman, N., Li, M., James, A.P., Dietzgen, R.G., Pearson, M.N., Revill, P.A., & Harding, R.M. (2016) Complete genome sequence of Colocasia bobone disease-associated virus, a putative cytorhabdovirus infecting Taro. *Arch. Virol. 161,* 745–748.

Ivancic A. (1992) Breeding and genetics of Taro (*Colocasia esculenta (L)*) Schott Ministry of Agriculture and Lands, Solomon Islands UNDP, Food and Agriculture Organizations of the United Nations, 1-97.

Jackson, G.V.H. (1978). Alomae and bobone diseases of Taro. South Pacific Commission Advisory Leaflet No. 8.

Jackson, G.V.H. (1980) 'Diseases and pests of taro.' (SPC: Noumea).

James M., Kenten R.H., Woods, R.D. (1973). Virus-like particles associated with two diseases of *Colocasia esculenta* (L.) Schott in the British Solomon Islands. *Journal of General Virology*, *21,* 145–153.

Jo, Y., Choi, H., Kim, S.-M., Kim, S.-L., Lee, B. C., & Cho, W. K. (2017). The pepper virome: natural co-infection of diverse viruses and their quasispecies. *BMC Genomics 18,* 453.

Jo, Y., Choi, H., Kim, S.-M., Kim, S.-L., Lee, B. C., & Cho, W. K. (2016). Integrated analyses using RNA-Seq data reveal viral genomes, single nucleotide variations, the phylogenetic relationship, and recombination for Apple stem grooving virus. *BMC Genomics, 17*, 579.

Jo, Y., Choi, H., Kyong Cho, J., Yoon, J.-Y., Choi, S.-K., & Kyong Cho, W. (2015). In silico approach to reveal viral populations in grapevine cultivar Tannat using transcriptome data. *Sci. Rep. 5,* 15841.

Kazmi, S. A., Yang, Z., Hong, N., Wang, G., & Wang, Y. (2015). Characterization by small RNA sequencing of Taro bacilliform CH Virus (TaBCHV), a novel Badnavirus. *PLoS One*.

Kashif, M., Pietilä, S., Artola, K., Jones, R. A. C., Tugume, A. K., Makinen, V., et al. (2012). Detection of viruses in sweet potato from Honduras and Guatemala augmented by deep-sequencing of small-RNAs. *Plant Dis*. *96*, 1430–1437.

Kidanemariam, D.B., Sukal, A.C., Abraham, A.D., Stomeo, F., Dale, J.L., James, A.P. & Harding, R.M. (2018) Identification and molecular characterization of Taro bacilliform virus and Taro bacilliform CH virus from East Africa. *Plant Pathol. 67,* 1977–1986.

Kidanemariam, D.B., Sukal, A.C., Abraham, A.D., Njuguna, J.N., Stomeo, F., Dale, J.L., James, A.P., Harding, R.M. (2022) Incidence of RNA viruses infecting taro and tannia in East Africa and molecular characterisation of dasheen mosaic virus isolates. *Ann Appl Biol*. *180*(2), 211-223.

King, A.M.Q., Adams, M.J., Lefkowitz, E.J., Carstens, E.B. (2012). *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses*. San Diego: Academic Press.

Kreuze, J. F., Perez, A., Untiveros, M., Quispe, D., Fuentes, S., Barker, I., & Simon, R. (2009) Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: a generic method for

diagnosis, discovery, and sequencing of viruses. *Virology, 388*(1), 1–7.

Leke, W.N., Mignouna, D.B., & Brown, J.K. (2015). Begomovirus disease complex: emerging threat to vegetable production systems of West and Central Africa. *Agric & Food Secur, 4,* 1.

Li, R., Gao, S., Hernandez, A. G., Wechter, W. P., Fei, Z., & Ling, K. S. (2012) Deep sequencing of small RNAs in tomato for virus and viroid identification and strain differentiation. *PLoS ONE, 7,* e37127.

Macanawai, A.R, Ebenebe, A.A, & Hunter, D. (2005) Investigations into the seed and mealybug transmission of Taro bacilliform virus. *Australasian Plant Pathology, 34,* 73–6.

Maino, M.K. (2003) *The development of a serological-based diagnostic test for Dasheen mosaic potyvirus (DsMV).* Unpublished M Sc Thesis, Queensland: Queensland University of Technology.

Matsumura, E., Coletta-Filho, H., Nouri, S., Falk, B., Nerva, L. & Oliveira, T. (2017) Deep sequencing analysis of RNAs from citrus plants grown in a citrus sudden death-affected area reveals diverse known and putative novel viruses. *Viruses, 9,* 92.

Natalia, K. & Rosemarie, W.H. (2014) Molecular biology of viroid–host interactions and disease control strategies, *Plant Science, 228,* 48-60.

Navas-Castillo, J., Fiallo-Olivé, E., & Sánchez-Campos, S. (2011) Emerging virus diseases transmitted by whiteflies. *Annu. Rev. Phytopathol. 49,* 219–248.

Ndabikunze, B.K., Talwana, H.A.L., Mongi, R.J. et al. (2011) Proximate and mineral composition of cocoyam (*Colocasia esculenta L*. and *Xanthosoma sagittifolium L*.) grown along the Lake Victoria Basin in Tanzania and Uganda. *African Journal of Food Science, 5,* 248– 54.

Okoń, S., Ociepa, T., Nucia, A., Cieplak, M., & Kowalczyk, K. (2021) Is Every Wild Species a Rich Source of Disease Resistance? *Avena fatua* L.-Potential Donor of Resistance to Powdery Mildew. *Plants (Basel, Switzerland)*, *10*(3), 560.

Onyeka, J. (2014) Status of Cocoyam (Colocasia esculenta and Xanthosoma spp) in West and Central Africa: Production, Household Importance and the Threat from Leaf Blight. Lima (Peru). *CGIAR Research Program on Roots, Tubers and Bananas (RTB)*.

Owens, R.A. & Verhoeven, J.Th.J. (2009) Potato spindle tuber. *The Plant Health Instructor*. DOI: 10.1094/PHI-I-2009-0804-01. Reviewed 2015.

Revill, P., Jackson, G., Hafner, G. (2005) Incidence and distribution of viruses of Taro (*Colocasia esculenta*) in Pacific Island countries. *Australasian Plant Pathology, 35*, 327–31.

Rodoni, B.C, Dale, J.L., & Harding, R.M. (1994) Review of alomae disease of Taro. *Papua New Guinea Journal of Agriculture, Forestry and Fisheries*, *37*, 14-18.

Sano, T., Barba, M., Li, S.F., & Hadidi, A. (2010) Viroids and RNA silencing: mechanism, role in viroid pathogenicity and development of viroid-resistant plants. *GM Crops. 1*(2), 80-6.

Shaw, D.E, Plumb, R.T., &Jackson G.V.H. (1979) Virus diseases of Taro (*Colocasia esculenta*) and *Xanthosoma* spp. in Papua New Guinea. *Papua New Guinea Agriculture Journal*, *30*, 71-97.

Shen, W., Le, S., Li, Y., & Hu, F. (2016) SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. *PLoS One. 11*(10), e0163962.

Singh, D., Jackson, D., Hunter, D., Fullerton R., Lebot, V., Tailor, M., Josef, T., et al. (2012). Taro Leaf Blight - A threat to food security. *Open access Agriculture*, *2*, 182-203.

Talwana, H.A.L., Serem, A,K., & Ndabikunze, B.K. (2009) Production Status and Prospects of Cocoyam (*Colocasia esculenta (L.) Schott*.) in East Africa. *Journal of Root Crops*, *35*, 98–107.

Temesgen., M. & Retta., N. (2015) Nutritional potential, Health and Food Security Benefits of Taro Colocasia esculenta (L.): A Review. *The Open Food Science Journal, 26,* 23 – 30.

Thresh, J.M., Otim-Nape, G.W., Legg, J.P., & Fargette, D. (1997) African cassava mosaic virus disease: the magnitude of the problem. *African Journal of Root and Tuber Crops*. *2* (1/2), 13-19.

Wanyama, D. & Mardell, G. (2006) Community of taro producers. Retrieved from www. Sustainable kenya.info.

Wylie, S. J., Li, H., Saqib, M., & Jones, M.G.K. (2014). The global trade in fresh produce and the vagility of plant viruses: a case study in garlic. *PLoS ONE 9,* e105044.

Wu, Q.F., Ding, S.W., Zhang, Y, J., & Zhu, S.F. (2015) Identification of viruses and viroids by next-generation sequencing and homology-dependent and homology-independent algorithms. *Annu Rev Phytopathol*, *53*, 425–444.

Yang, I.C., Hafner, G.J., Revill, P.A., Dale, J.L., & Harding, R.M. (2003b) Sequence diversity of South Pacific isolates of Taro bacilliform virus and the development of a PCR-based diagnostic test. *Archives of Virology*, *148,* 1957– 68.

Yusop, M.S.M., Saad, M.F.M., Talip, N., Baharum, S.N., & Bunawan, H. (2019) A Review on Viruses Infecting Taro (Colocasia esculenta (L.) Schott). *Pathogens*. *8*(2), 56.

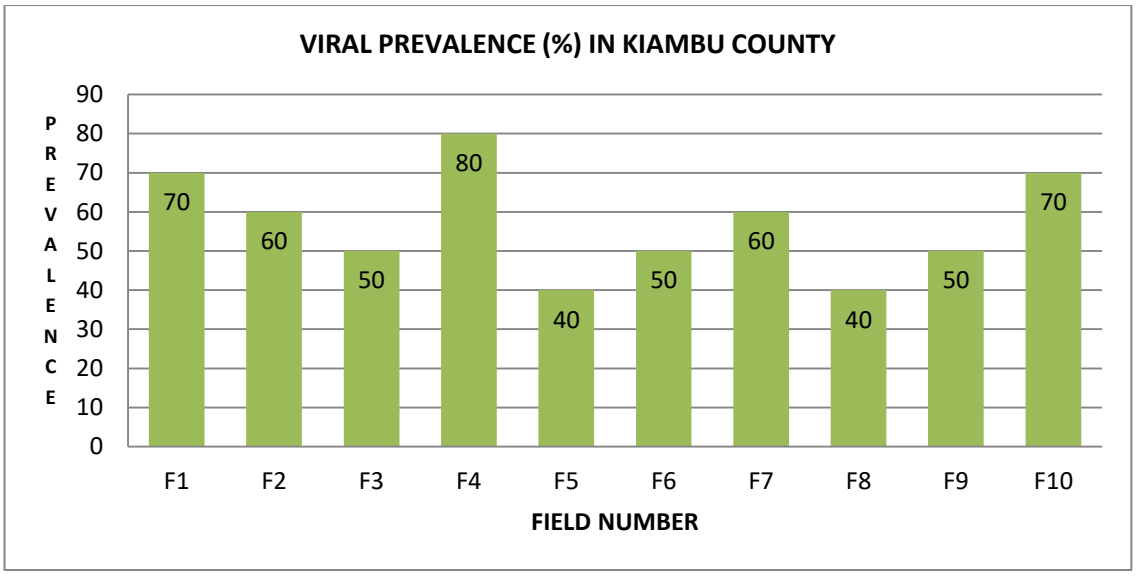
Zettler, F.W. & Hartman, R.D. (1986) Dasheen mosaic virus and its' control in cultivated aroids. Extension Bulletin, ASPAC Food and Fertilizer

Technology Centre for the Asian and Pacific region, *Taipei. No. 233,* 13.

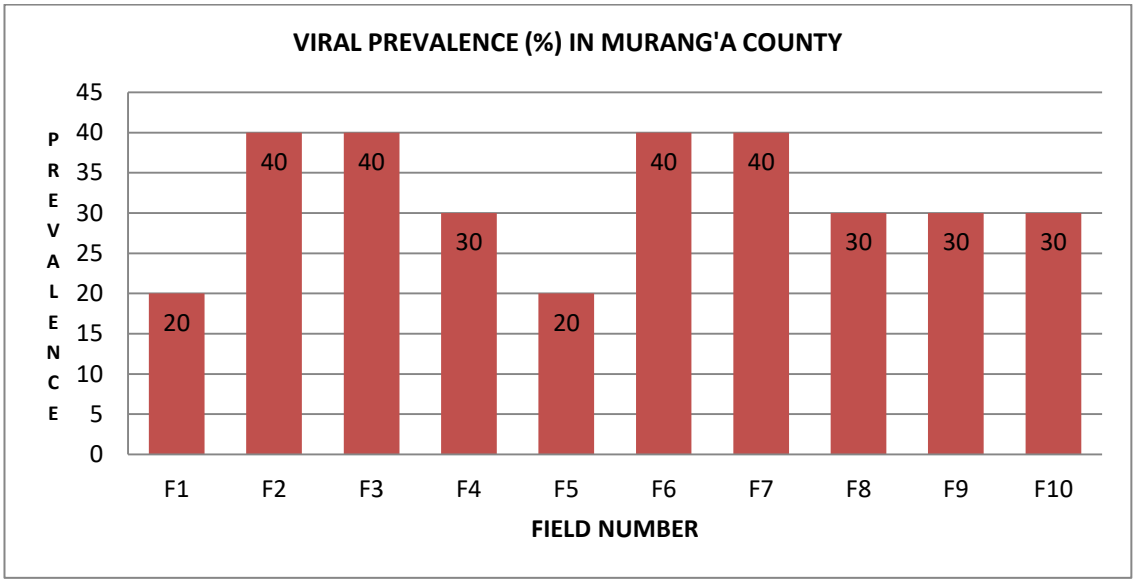
Zettler, F.W. & Hartman, R.D. (1987) Dasheen mosaic virus as a pathogen of cultivated aroids and control of the virus by Tissue culture. *Plant Disease, 71*, 958–963.

APPENDICES

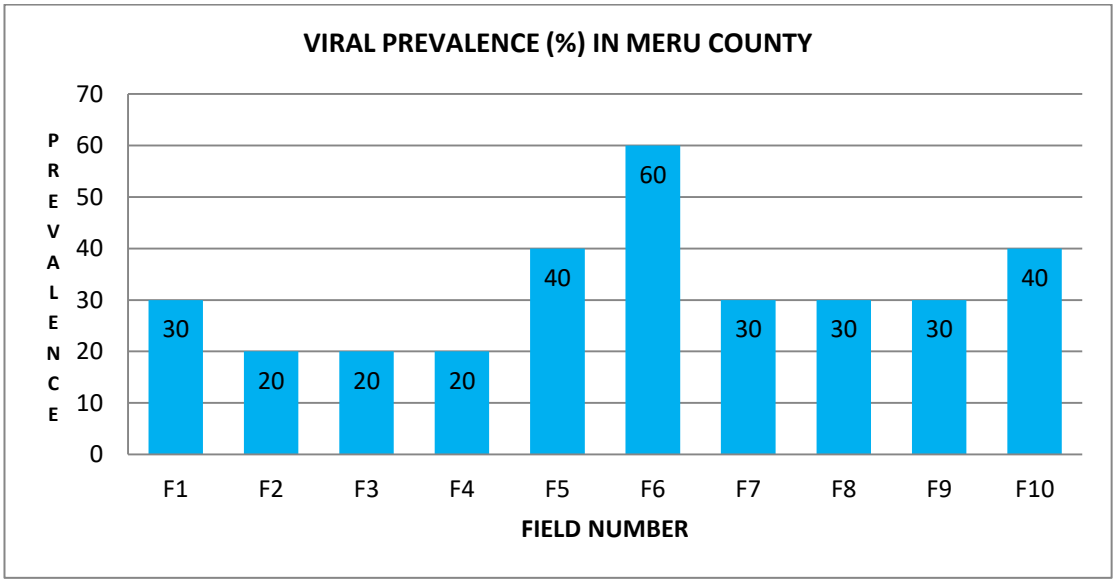
Appendix I: Graph Showing the Viral Prevalence in Kiambu County among the Ten Fields Surveyed



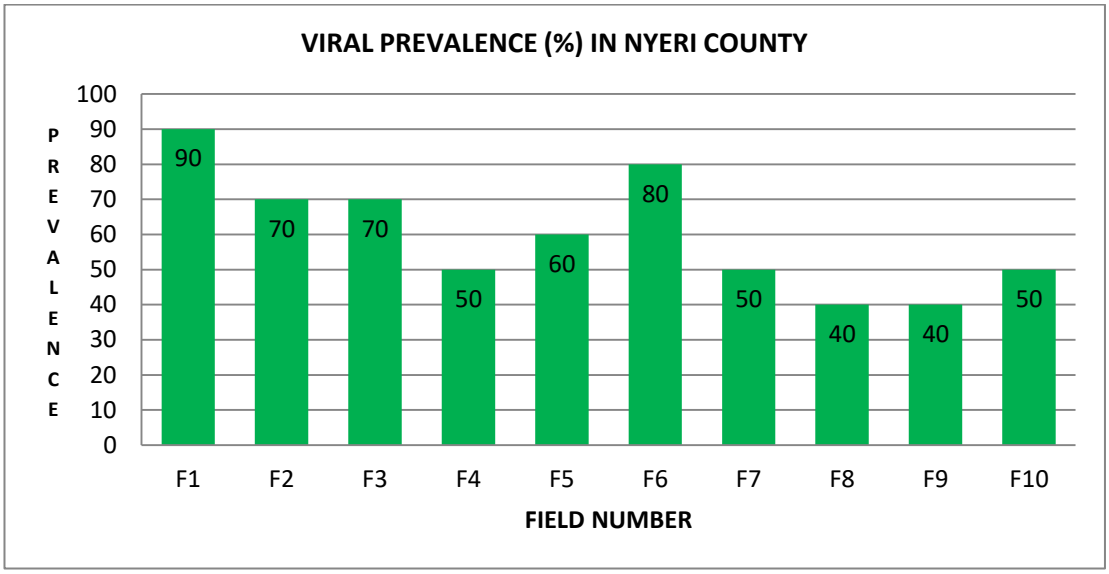
Appendix II: Graph Showing the Viral Prevalence in Murang'a County among the Ten Fields Surveyed



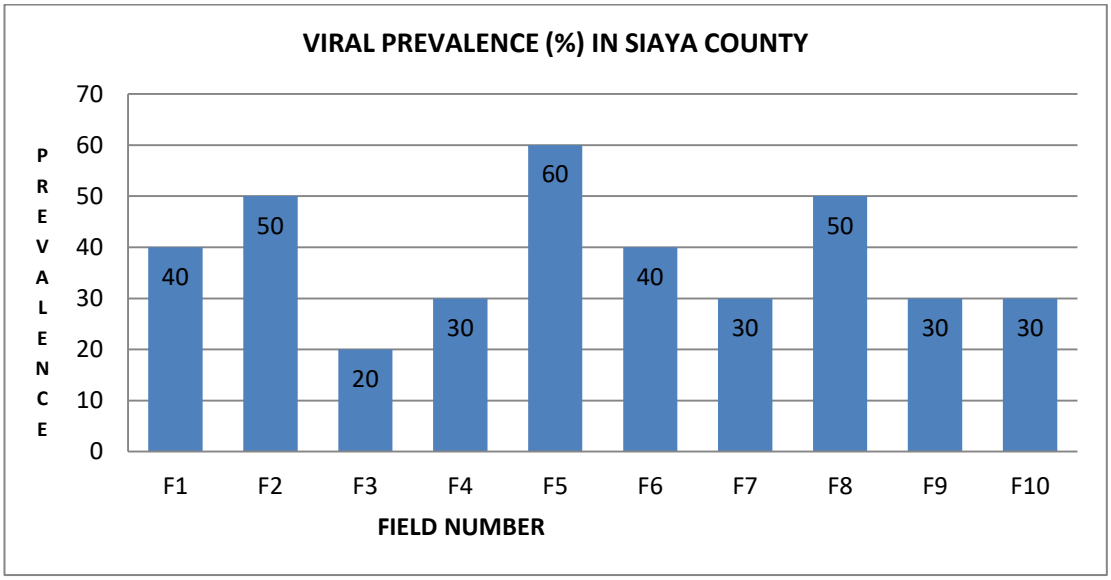
Appendix III: Graph Showing the Viral Prevalence in Meru County among the Ten Fields Surveyed



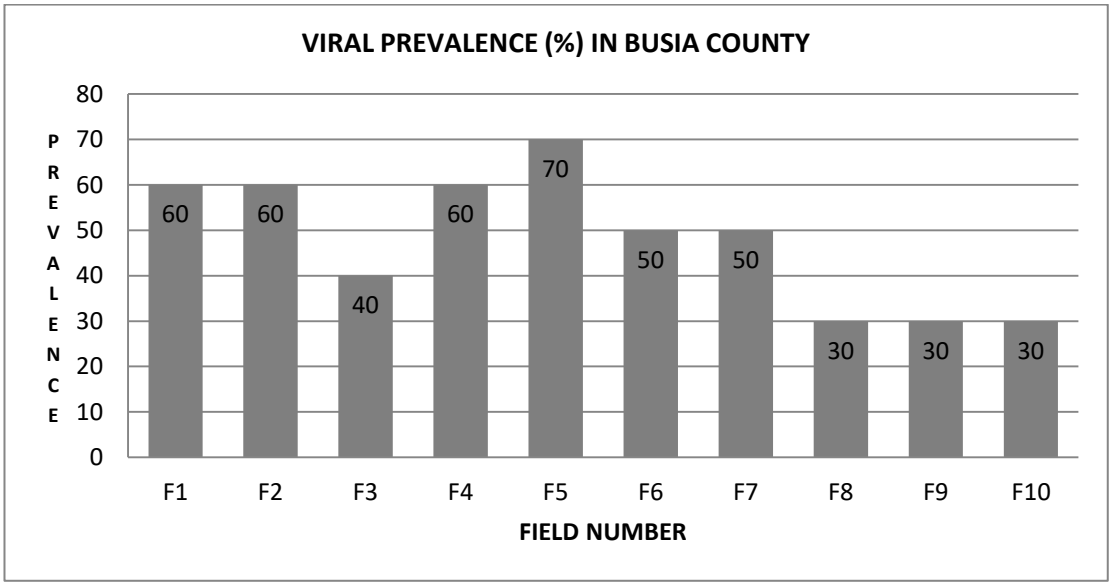
Appendix IV: Graph Showing the Viral Prevalence in Nyeri County among the Ten Fields Surveyed



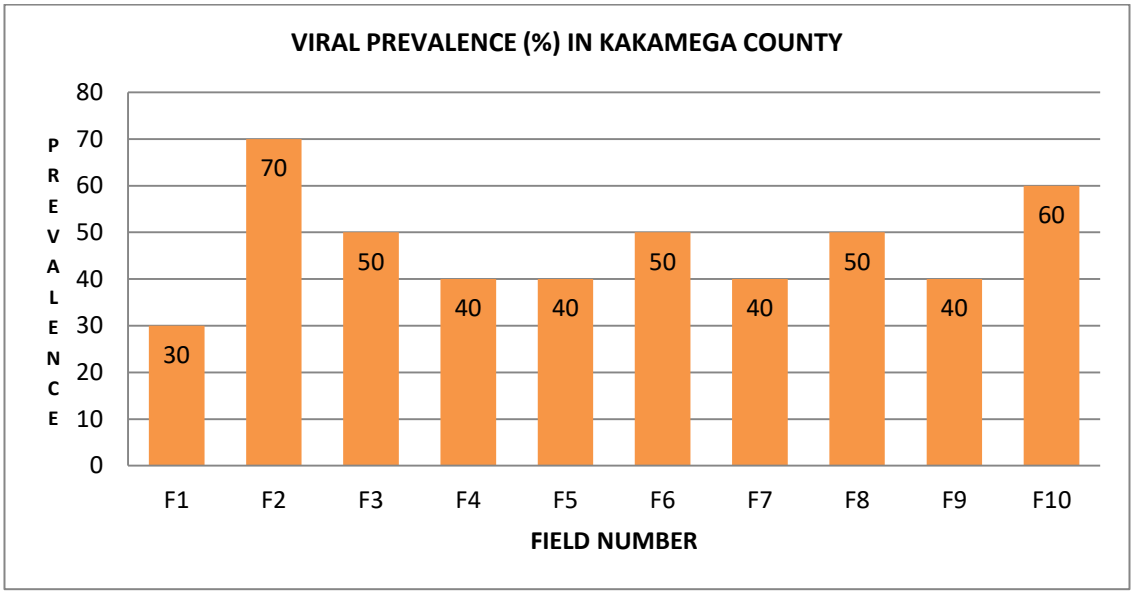
Appendix V: Graph Showing the Viral Prevalence in Siaya County among the Ten Fields Surveyed



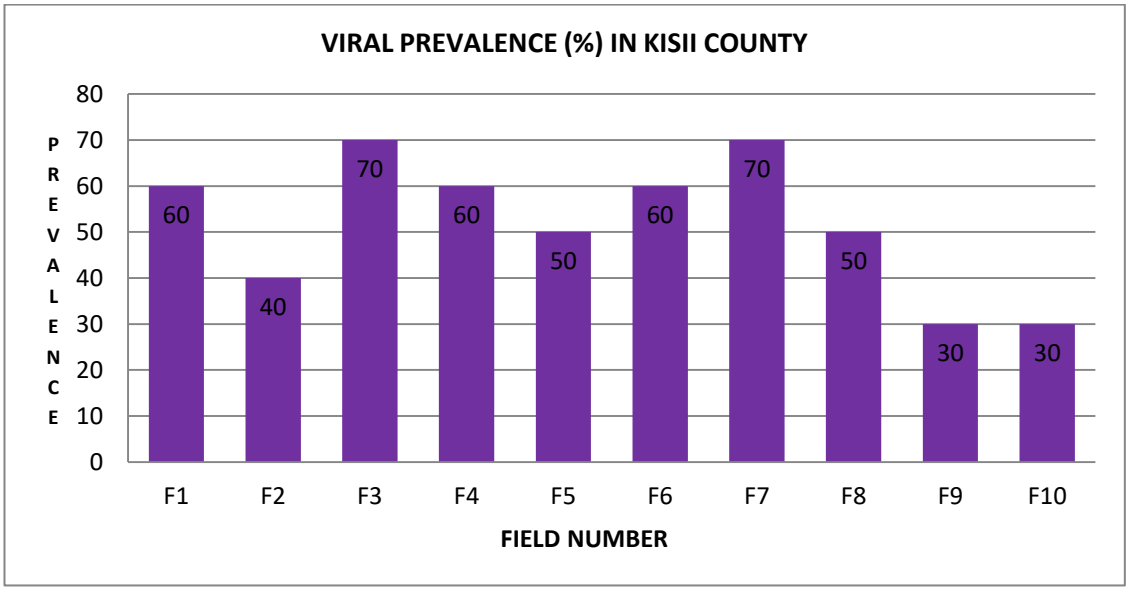
Appendix VI: Graph Showing the Viral Prevalence in Busia County among the Ten Fields Surveyed



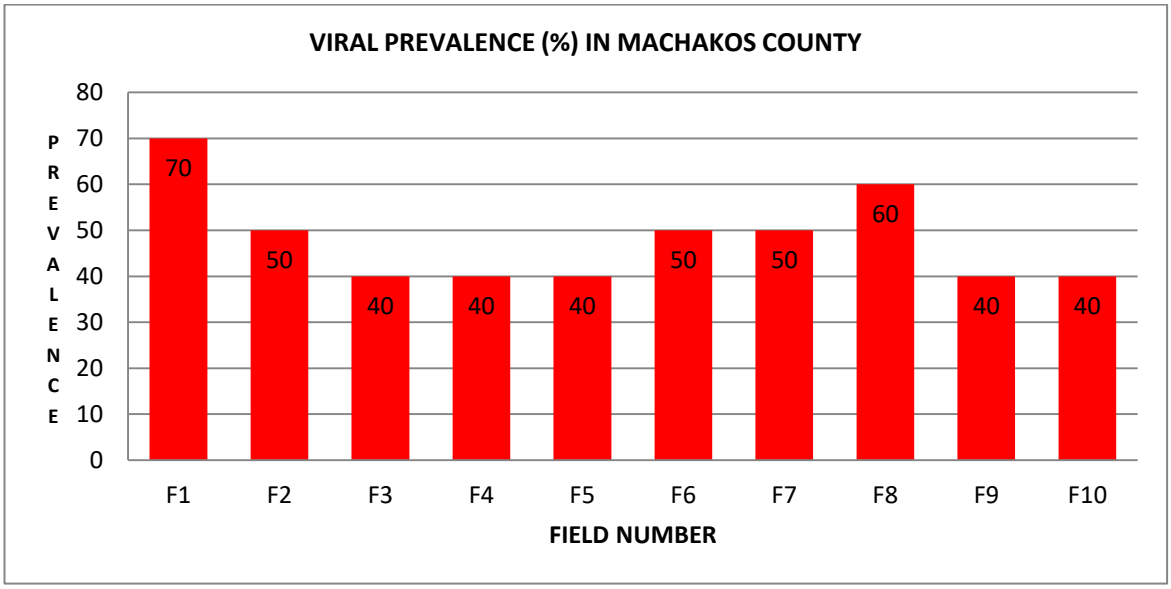
Appendix VII: Graph Showing the Viral Prevalence in Kakamega County among the Ten Fields Surveyed



Appendix VIII: Graph Showing the Viral Prevalence in Kisii County among the Ten Fields Surveyed



Appendix IX: Graph Showing the Viral Prevalence in Machakos County among the Ten Fields Surveyed



Appendix X: Blastn Results of all the 15 Isolates Showing the Top Five Hits Based on the Percentage Identity of the Query Sequence to the Database Accessions

Isolate	Accession	Description	Percentage Identity	E-value
Kiambu1	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	98.792	1.23E-163
Kiambu1	MG017323.1	Taro bacilliform virus isolate Ug75, complete genome	98.792	1.23E-163
Kiambu1	MG017322.1	Taro bacilliform virus isolate Tz17, complete genome	98.792	1.23E-163
Kiambu1	MG017321.1	Taro bacilliform virus isolate Ke52, complete genome	98.792	1.23E-163
Kiambu1	MG017336.1	Taro bacilliform virus isolate Ug45 ORF3 gene, partial cds	98.792	1.23E-163
Siaya10	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	96.108	1.16E-148
Siaya10	MG017323.1	Taro bacilliform virus isolate Ug75, complete genome	96.108	1.16E-148
Siaya10	MG017322.1	Taro bacilliform virus isolate Tz17, complete genome	96.108	1.16E-148
Siaya10	MG017321.1	Taro bacilliform virus isolate Ke52, complete genome	96.108	1.16E-148
Siaya10	MG017336.1	Taro bacilliform virus isolate Ug45 ORF3 gene, partial cds	96.108	1.16E-148
Busia12	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	98.799	9.39E-165
Busia12	MG017323.1	Taro bacilliform virus isolate Ug75, complete genome	98.799	9.39E-165
Busia12	MG017322.1	Taro bacilliform virus isolate Tz17, complete genome	98.799	9.39E-165
Busia12	MG017321.1	Taro bacilliform virus isolate Ke52, complete genome	98.799	9.39E-165
Busia12	MG017336.1	Taro bacilliform virus isolate Ug45 ORF3 gene, partial cds	98.799	9.39E-165
Kakamega17	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	93.75	1.58E-137

Kakamega17	MG017323.1	Taro bacilliform virus isolate Ug75, complete genome	93.75	1.58E-137
Kakamega17	MG017322.1	Taro bacilliform virus isolate Tz17, complete genome	93.75	1.58E-137
Kakamega17	MG017321.1	Taro bacilliform virus isolate Ke52, complete genome	93.75	1.58E-137
Kakamega17	MG017336.1	Taro bacilliform virus isolate Ug45 ORF3 gene, partial cds	93.75	1.58E-137
Murang'a2	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	94.828	1.22E-148
Murang'a2	MG017323.1	Taro bacilliform virus isolate Ug75, complete genome	96.084	1.22E-148
Murang'a2	MG017322.1	Taro bacilliform virus isolate Tz17, complete genome	94.828	1.22E-148
Murang'a2	MG017321.1	Taro bacilliform virus isolate Ke52, complete genome	94.828	1.22E-148
Murang'a2	MG017336.1	Taro bacilliform virus isolate Ug45 ORF3 gene, partial cds	94.828	1.22E-148
Kisii22	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	93.983	1.40E-143
Kisii22	MG017323.1	Taro bacilliform virus isolate Ug75, complete genome	93.948	5.05E-143
Kisii22	MG017322.1	Taro bacilliform virus isolate Tz17, complete genome	93.983	1.40E-143
Kisii22	MG017321.1	Taro bacilliform virus isolate Ke52, complete genome	93.983	1.40E-143
Kisii22	MG017336.1	Taro bacilliform virus isolate Ug45 ORF3 gene, partial cds	93.983	1.40E-143
Machakos24	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	96.875	6.80E-147
Machakos24	MG017323.1	Taro bacilliform virus isolate Ug75, complete genome	96.875	6.80E-147
Machakos24	MG017322.1	Taro bacilliform virus isolate Tz17, complete genome	96.875	6.80E-147
Machakos24	MG017321.1	Taro bacilliform virus isolate Ke52, complete genome	96.875	6.80E-147
Machakos24	MG017336.1	Taro bacilliform virus isolate Ug45 ORF3 gene, partial cds	96.875	6.80E-147

Murang'a3	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	94.828	1.25E-148
Murang'a3	MG017323.1	Taro bacilliform virus isolate Ug75, complete genome	94.54	5.83E-147
Murang'a3	MG017322.1	Taro bacilliform virus isolate Tz17, complete genome	94.828	1.25E-148
Murang'a3	MG017321.1	Taro bacilliform virus isolate Ke52, complete genome	94.828	1.25E-148
Murang'a3	MG017336.1	Taro bacilliform virus isolate Ug45 ORF3 gene, partial cds	94.828	1.25E-148
Machakos32	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	95.415	5.10E-153
Machakos32	MG017323.1	Taro bacilliform virus isolate Ug75, complete genome	95.677	1.42E-153
Machakos32	MG017322.1	Taro bacilliform virus isolate Tz17, complete genome	96.697	1.83E-152
Machakos32	MG017321.1	Taro bacilliform virus isolate Ke52, complete genome	95.415	5.10E-153
Machakos32	MG017336.1	Taro bacilliform virus isolate Ug45 ORF3 gene, partial cds	95.415	5.10E-153
Meru4	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	97.329	4.27E-158
Meru4	MG017323.1	Taro bacilliform virus isolate Ug75, complete genome	97.033	1.99E-156
Meru4	MG017322.1	Taro bacilliform virus isolate Tz17, complete genome	97.329	4.27E-158
Meru4	MG017321.1	Taro bacilliform virus isolate Ke52, complete genome	97.329	4.27E-158
Meru4	MG017336.1	Taro bacilliform virus isolate Ug45 ORF3 gene, partial cds	97.329	4.27E-158
Meru5	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	97.399	1.23E-163
Meru5	MG017323.1	Taro bacilliform virus isolate Ug75, complete genome	97.399	1.23E-163
Meru5	MG017322.1	Taro bacilliform virus isolate Tz17, complete genome	97.399	1.23E-163
Meru5	MG017321.1	Taro bacilliform virus isolate Ke52, complete genome	97.399	1.23E-163

Meru5	MG017336.1	Taro bacilliform virus isolate Ug45 ORF3 gene, partial cds	97.399	1.23E-163
Nyeri6	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	97.289	7.43E-156
Nyeri6	MG017323.1	Taro bacilliform virus isolate Ug75, complete genome	97.289	7.43E-156
Nyeri6	MG017322.1	Taro bacilliform virus isolate Tz17, complete genome	97.289	7.43E-156
Nyeri6	MG017321.1	Taro bacilliform virus isolate Ke52, complete genome	97.289	7.43E-156
Nyeri6	MG017336.1	Taro bacilliform virus isolate Ug45 ORF3 gene, partial cds	97.289	7.43E-156
Nyeri7	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	94.062	1.99E-131
Nyeri7	MG017323.1	Taro bacilliform virus isolate Ug75, complete genome	94.062	1.99E-131
Nyeri7	MG017322.1	Taro bacilliform virus isolate Tz17, complete genome	94.062	1.99E-131
Nyeri7	MG017321.1	Taro bacilliform virus isolate Ke52, complete genome	94.062	1.99E-131
Nyeri7	MG017336.1	Taro bacilliform virus isolate Ug45 ORF3 gene, partial cds	94.062	1.99E-131
Nyeri8	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	97.289	7.56E-156
Nyeri8	MG017323.1	Taro bacilliform virus isolate Ug75, complete genome	96.232	2.10E-156
Nyeri8	MG017322.1	Taro bacilliform virus isolate Tz17, complete genome	97.289	7.56E-156
Nyeri8	MG017321.1	Taro bacilliform virus isolate Ke52, complete genome	97.289	7.56E-156
Nyeri8	MG017336.1	Taro bacilliform virus isolate Ug45 ORF3 gene, partial cds	97.289	7.56E-156
Siaya9	MG833013.1	Taro bacilliform virus isolate Tz24, complete genome	96.073	4.10E-148
Siaya9	MG017323.1	Taro bacilliform virus isolate Ug75, complete genome	96.073	4.10E-148
Siaya9	MG017322.1	Taro bacilliform virus isolate Tz17, complete genome	96.073	4.10E-148

Siaya9	MG017321.1	Taro bacilliform virus isolate Ke52, complete genome	96.073	4.10E-148
Siaya9	MG017336.1	Taro bacilliform virus isolate Ug45 ORF3 gene, partial cds	96.073	4.10E-148

Appendix XI: Representative Blastn Results of Small RNA Sequences Revealing Hits to Various DNA and RNA Viruses

Querry ID	Accession	Percentage Identity	E-value	Description
255550-3	gi|86651756|gb|DQ318790.1|	100	5.03E-28	Citrus exocortis viroid clone 2/3
281422-3	gi|86651756|gb|DQ318790.1|	98.667	5.03E-28	Citrus exocortis viroid clone 2/3
366597-2	gi|86651756|gb|DQ318790.1|	100	2.39E-11	Citrus exocortis viroid clone 2/3
388113-2	gi|86651756|gb|DQ318790.1|	98.462	1.82E-22	Citrus exocortis viroid clone 2/3
8000-193	gi|1343311831|gb|MF375892.1|	100	0.001	Phaseolus vulgaris alphaendornavirus 1 polyprotein gene
11849-147	gi|1343311831|gb|MF375892.1|	100	0.001	Phaseolus vulgaris alphaendornavirus 1 polyprotein gene
15620-113	gi|1343311831|gb|MF375892.1|	100	0.001	Phaseolus vulgaris alphaendornavirus 1 polyprotein gene
18815-88	gi|1343311831|gb|MF375892.1|	100	0.001	Phaseolus vulgaris alphaendornavirus 1 polyprotein gene
1924571-1	gi|1587104515|gb|MH142485.1|	100	0.000109	Sugarcane bacilliform virus isolate HNSb4 polyprotein gene
1924571-1	gi|1587104523|gb|MH142489.1|	100	0.000109	Sugarcane bacilliform virus isolate HNSb8 polyprotein gene
1213887-1	gi|770114323|gb|KM009090.1|	100	0.001	Sweet potato badnavirus A isolate KSR675 NORA-II(2) polyprotein gene
1213887-1	gi|769469215|gb|KM000054.1|	100	0.001	Sweet potato badnavirus A isolate Marooko polyprotein gene
1809189-1	gi|769469209|gb|KM000051.1|	100	0.0000311	Sweet potato badnavirus A isolate Tanzania Carrot C-1(2) polyprotein gene
1213887-1	gi|770114321|gb|KM009089.1|	100	0.001	Sweet potato badnavirus B isolate Bitambi polyprotein gene
1809189-1	gi|770114335|gb|KM009096.1|	100	0.0000311	Sweet potato badnavirus B isolate Cameroon-RH polyprotein gene
1623560-1	gi|156857638|gb|EF591125.1|	94.915	8.53E-16	Sweet potato begomovirus isolate BG9 movement protein (V2)
1623560-1	gi|270268864|gb|GQ268223.1|	100	5.17E-08	Sweet potato begomovirus strain SPBV-D-DNA-[1]
1327058-1	gi|73852992|emb|AJ811971.1|	97.333	2.34E-26	Sweet potato chlorotic stunt virus CP gene for coat protein
1327058-1	gi|73852994|emb|AJ811972.1|	96	1.09E-24	Sweet potato chlorotic stunt virus CP gene for coat protein
1327058-1	gi|73852990|emb|AJ811970.1|	96	1.09E-24	Sweet potato chlorotic stunt virus CP gene for coat protein
1488095-1	gi|1089801|dbj|D38543.1|SPFCI	95.238	5.17E-08	Sweet potato feathery mottle virus CI
3582729-1	gi|2394166|gb|AF015541.1|AF015541	100	0.000403	Sweet potato feathery mottle virus coat protein (CP) gene
399778-2	gi|295841641|dbj|AB509459.1|	97.059	1.78E-22	Sweet potato feathery mottle virus gene for polyprotein
934417-1	gi|295841649|dbj|AB509463.1|	94.595	1.78E-22	Sweet potato feathery mottle virus gene for polyprotein
934417-1	gi|295841645|dbj|AB509461.1|	94.595	1.78E-22	Sweet potato feathery mottle virus gene for polyprotein

1049883-1	gi|29466290|emb|AJ539130.1|	98.667	5.03E-28	Sweet potato feathery mottle virus gene for polyprotein
1049883-1	gi|29466292|emb|AJ539131.1|	96	1.09E-24	Sweet potato feathery mottle virus gene for polyprotein
1623560-1	gi|195963278|dbj|AB433786.1|	98.305	3.94E-19	Sweet potato leaf curl virus DNA
1623560-1	gi|195963292|dbj|AB433788.1|	94.915	8.53E-16	Sweet potato leaf curl virus DNA
1623560-1	gi|5702158|gb|AF104036.1|	96.61	1.83E-17	Sweet potato leaf curl virus DNA A
2249835-1	gi|295042100|emb|FN806776.1|	97.297	0.000000669	Sweet potato leaf curl virus DNA-A
1623560-1	gi|1115464097|dbj|LC201923.1|	98.305	3.94E-19	Sweet potato leaf curl virus genes for V1 protein
1623560-1	gi|614706430|gb|KF697069.1|	98.148	2.37E-16	Sweet potato leaf curl virus isolate [Greece:Crete 638:2013]
1623560-1	gi|614706439|gb|KF697070.1|	98.148	2.37E-16	Sweet potato leaf curl virus isolate [Greece:Crete 638-1:2013]
2495520-1	gi|1399734333|gb|MG017325.1|	91.304	0.000000669	Taro bacilliform CH virus isolate Ke43
3020826-1	gi|1399734333|gb|MG017325.1|	93.878	3.02E-10	Taro bacilliform CH virus isolate Ke43
127181-7	gi|794005704|gb|KP710178.1|	100	0.001	Taro bacilliform CH virus isolate TaBCHV-1
520654-2	gi|794005704|gb|KP710178.1|	100	0.001	Taro bacilliform CH virus isolate TaBCHV-1
521202-2	gi|794005704|gb|KP710178.1|	100	0.001	Taro bacilliform CH virus isolate TaBCHV-1
579530-2	gi|794005704|gb|KP710178.1|	96.875	0.001	Taro bacilliform CH virus isolate TaBCHV-1
586036-2	gi|794005675|gb|KP710177.1|	100	0.001	Taro bacilliform CH virus isolate TaBCHV-2
767720-1	gi|794005675|gb|KP710177.1|	94.737	0.00000865	Taro bacilliform CH virus isolate TaBCHV-2
2495520-1	gi|1510256154|gb|MG833014.1|	91.304	0.000000669	Taro bacilliform CH virus isolate Tz27
3020826-1	gi|1510256154|gb|MG833014.1|	93.878	3.02E-10	Taro bacilliform CH virus isolate Tz27
2495520-1	gi|1389516217|gb|MG017326.1|	91.304	0.000000669	Taro bacilliform CH virus isolate Tz36
3020826-1	gi|1389516217|gb|MG017326.1|	93.878	3.02E-10	Taro bacilliform CH virus isolate Tz36
2495520-1	gi|1389516222|gb|MG017327.1|	91.304	0.000000669	Taro bacilliform CH virus isolate Ug10
3020826-1	gi|1389516222|gb|MG017327.1|	93.878	3.02E-10	Taro bacilliform CH virus isolate Ug10
1639466-1	gi|1389516188|gb|MG017359.1|	100	0.001	Taro bacilliform CH virus isolate Ug52 ORF3 gene
1658564-1	gi|1389516188|gb|MG017359.1|	100	0.000000182	Taro bacilliform CH virus isolate Ug52 ORF3 gene
3770473-1	gi|1389516188|gb|MG017359.1|	100	0.001	Taro bacilliform CH virus isolate Ug52 ORF3 gene
70390-14	gi|1389516192|gb|MG017321.1|	100	0.001	Taro bacilliform virus isolate Ke52
103146-9	gi|1389516192|gb|MG017321.1|	100	0.000403	Taro bacilliform virus isolate Ke52
110532-8	gi|1389516192|gb|MG017321.1|	100	0.00000865	Taro bacilliform virus isolate Ke52
172390-5	gi|1389516192|gb|MG017321.1|	100	0.001	Taro bacilliform virus isolate Ke52
175553-5	gi|1389516192|gb|MG017321.1|	100	0.001	Taro bacilliform virus isolate Ke52

456834-2	gi|1389516197|gb|MG017322.1|	97.619	0.000000004	Taro bacilliform virus isolate Tz17
460638-2	gi|1389516197|gb|MG017322.1|	100	0.000403	Taro bacilliform virus isolate Tz17
465328-2	gi|1389516197|gb|MG017322.1|	100	5.06E-23	Taro bacilliform virus isolate Tz17
467322-2	gi|1389516197|gb|MG017322.1|	100	0.000403	Taro bacilliform virus isolate Tz17
478463-2	gi|1389516197|gb|MG017322.1|	100	0.001	Taro bacilliform virus isolate Tz17
493021-2	gi|1389516197|gb|MG017322.1|	100	0.000403	Taro bacilliform virus isolate Tz17
796461-1	gi|1510256149|gb|MG833013.1|	100	0.001	Taro bacilliform virus isolate Tz24
801431-1	gi|1510256149|gb|MG833013.1|	100	0.000000014	Taro bacilliform virus isolate Tz24
806163-1	gi|1510256149|gb|MG833013.1|	100	0.00000235	Taro bacilliform virus isolate Tz24
437532-2	gi|1389516202|gb|MG017323.1|	100	0.001	Taro bacilliform virus isolate Ug75
438985-2	gi|1389516202|gb|MG017323.1|	100	0.00000241	Taro bacilliform virus isolate Ug75
441735-2	gi|1389516202|gb|MG017323.1|	100	0.000403	Taro bacilliform virus isolate Ug75
442029-2	gi|1389516202|gb|MG017323.1|	100	0.0000311	Taro bacilliform virus isolate Ug75
442548-2	gi|1389516202|gb|MG017323.1|	100	0.001	Taro bacilliform virus isolate Ug75

Appendix XII: Metadata of the Individual Fields Surveyed

FIELD NUMBER	FIELD LOCATION (WARD)	LONGITUDE	LATITUDE	ALTITUDE (M)	FERTILIZER/ MANURE USE	CHEMICAL USE (YES/NO)
KIAMBU COUNTY						
F1	Ngewa	S -1 5' 54.9"	E 36 51' 57.5"	1636	Manure	No
F2	Kikuyu	S -1 14' 47.2"	E 36 39' 18.3"	1986	Manure	No
F3	Kikuyu	S -1 16' 47.3"	E 36 40' 58.5"	1865	Manure	No
F4	Karai	S -1 18' 38.1"	E 36 39' 47.5"	1847	Manure	No
F5	Kikuyu	S -1 16' 21.5"	E 36 41' 9.7"	1849	Manure	No
F6	Gitaru	S -1 13' 32.5"	E 36 42' 8.2"	1882	Manure	No
F7	Gitaru	S -1 13' 5.5"	E 36 41' 33.9"	1917	Manure	No
F8	Kabete	S -1 14' 1.2"	E 36 43' 42.2"	1780	Manure	No
F9	Kabete	S -1 14' 1.3"	E 36 43' 42.2"	1777	Manure	No
F10	Nyathuna	S -1 11' 28.4"	E 36 41' 51.4"	1927	Manure	No
MURANG'A COUNTY						
F1	Ngenda	S 0 46' 15.5"	E 37 37' 42.3"	1314	Manure	No
F2	Ngenda	S 0 46' 15.3"	E 37 7' 48.6"	1310	Manure	No
F3	Ngenda	S 0 46' 13.9"	E 37 7' 52.6"	1309	Manure	No
F4	Ngenda	S 0 45' 58.6"	E 37 7' 44.6"	1291	Manure	No
F5	Ngenda	S 0 46' 28.0"	E 37 7' 36.1"	1348	Manure	No
F6	Kigumo	S 0 48' 32.7"	E 36 58' 44.3"	1690	Manure	No
F7	Kigumo	S 0 48' 34.9"	E 36 58' 55.0"	1682	Manure	No
F8	Kigumo	S 0 49' 10.2"	E 36 59' 24.3"	1684	Manure	No
F9	Kigumo	S 0 49' 15.5"	E 36 59' 17.1"	1686	Manure	No
F10	Kigumo	S 0 48' 37.0"	E 37 0' 33.5"	1612	Manure	No
MERU COUNTY						
F1	Mikinduri	N 0 6' 41.5"	E 37 48' 39.7"	1178	Manure	No
F2	Nkomo	N 0 5' 14.8"	E 37 46' 29.4"	1179	Manure	No
F3	Nyaki	N 0 2' 17.0"	E 37 44' 32.9"	1139	Manure	No
F4	Nyaki	N 0 2' 25.7"	E 37 44' 31.0"	1148	Manure	No
F5	Mikinduri	N 0 5' 10.0"	E 37 51' 46.7"	1125	Manure	No
F6	Kiguchwa	N 0 9' 9.0"	E 37 51' 23.3"	1453	Manure	No
F7	Kiamurio	S 0 6' 0.3"	E 37 44' 52.5"	1152	Manure	No
F8	Nyagene	S 0 6' 32.8"	E 37 45' 10.5"	1061	Manure	No
F9	Mitunguu	S 0 6' 41.4"	E 37 45' 16.7"	1049	Manure	No
F10	Maraa	S 0 7' 30.1"	E 37 45' 34.5"	1009	Manure	Yes
NYERI COUNTY						
F1	Ruguru	S 0 20' 26.5"	E 37 5' 57.5"	1937	Manure	No
F2	Ruguru	S 0 20' 0.8"	E 37 5' 58.7"	1957	Manure	No

F3	Ruguru	S 0 20' 31.5"	E 37 5' 42.9"	1934	Manure	No
F4	Ruguru	S 0 20' 32.7"	E 37 5' 32.7"	1903	Manure	No
F5	Ruguru	S 0 20' 51.5"	E 37 5' 45.5"	1915	Manure	No
F6	Ruguru	S 0 20' 53.4"	E 37 5' 17.9"	1873	Manure	No
F7	Karatina	S 0 28' 59.5"	E 37 8' 8.8"	1744	Manure	No
F8	Iriaini	S 0 28' 31.4"	E 37 8' 32.4"	1737	Manure	No
F9	Kirimukuyu	S 0 28' 42.1"	E 37 6' 2.3"	1718	Manure	No
F10	Kirimukuyu	S 0 26' 28.2	E 37 5' 34.3	1767	Manure	No
SIAYA COUNTY						
F1	Central Alego	N 0 2' 20.1"	E 34 8' 51.8"	1143	Manure	No
F2	Central Alego	N 0 2' 31.6"	E 34 9' 15.8"	1142	Manure	No
F3	Central Alego	N 0 2' 19.7"	E 34 9' 46.6"	1145	Manure	No
F4	Central Alego	N 0 2' 10.2"	E 34 9' 55.5"	1138	Manure	No
F5	Yala Township	N 0 6' 43.0"	E 34 33' 0.4"	1397	Manure	No
F6	Yala Township	N 0 6' 41.2"	E 34 32' 57.6"	1395	Manure	No
F7	Yala Township	N 0 7' 8.5"	E 34 32' 52.1"	1402	Manure	No
F8	Yala Township	N 0 7' 50.7"	E 34 32' 42.5"	1407	Manure	No
F9	Yala Township	N 0 8' 13.9"	E 34 32' 45.1"	1408	Manure	No
F10	Yala Township	N 0 8' 20.7"	E 34 32' 41.3"	1408	Manure	No
BUSIA COUNTY						
F1	Bunyala Central	N 0 3' 47.4"	E 34 0' 53.0"	1134	Manure	No
F2	Bunyala Central	N 0 3' 52.7"	E 34 0' 54.3"	1137	Manure	No
F3	Bunyala Central	N 0 3' 41.8"	E 34 0' 59.4"	1140	Manure	No
F4	Bunyala Central	N 0 3' 40.7"	E 34 1' 0.7"	1138	Manure	No
F5	Bunyala Central	N 0 3' 37.7"	E 34 1' 10.4"	1138	Manure	No
F6	Bunyala Central	N 0 4' 20.4"	E 34 1' 26.7"	1132	Manure	No

F7	Bunyala Central	N 0 5' 40.2"	E 34 1' 28.2"	1138	Manure	No
F8	Hajula	N 0 5' 32.8"	E 34 0' 25.4"	1141	Manure	No
F9	Hajula	N 0 3' 55.4"	E 33 59' 31.4"	1136	Manure	No
F10	Hajula	N 0 2' 48.1"	E 33 59' 35.2"	1135	Manure	No
KAKAMEGA COUNTY						
F1	Mahiakalo	N 0 17' 18.7"	E 34 46' 5.8"	1507	Manure	No
F2	Butsotso East	N 0 19' 44.3"	E 34 43' 36.2"	1462	Manure	No
F3	Butsotso East	N 0 18' 57.7"	E 34 44' 4.5"	1493	Manure	No
F4	Butsotso East	N 0 19' 6.8"	E 34 43' 58.2"	1490	Manure	No
F5	Mahiakalo	N 0 17' 27.5"	E 34 46' 32.7"	1548	Manure	No
F6	Ingotse Matia	N 0 20' 56.3	E 34 44' 4.3"	1467	Manure	No
F7	Ingotse Matia	N 0 21' 42.6	E 34 44' 32.5"	1484	Manure	No
F8	Ingotse Matia	N 0 21' 5.8"	E 34 45' 16.1"	1487	Manure	No
F9	Ingotse Matia	N 0 21' 52.3"	E 34 41' 89"	1423	Manure	No
F10	Ingotse Matia	N 0 21' 13.1"	E 34 41' 34.2"	1463	Manure	No
KISII COUNTY						
F1	Obaracho	S 0 40' 41.8"	E 34 47' 27.6"	1607	Manure	No
F2	Obaracho	S 0 40' 19.6"	E 34 46' 26.4"	1637	Manure	No
F3	Nyakoe	S 0 36' 34"	E 34 46' 41.8"	1607	Manure	No
F4	Nyakoe	S 0 37' 17.4"	E 34 43' 54.4"	1500	Manure	No
F5	Nyakoe	S 0 37'21.4"	E 34 44' 11.1"	1514	Manure	No
F6	Nyakoe	S 0 37' 13.7"	E 34 43' 56.0	1500	Manure	No
F7	Nyakoe	S 0 37' 09"	E 34 43' 57.4"	1496	Manure	No
F8	Nyakoe	S 0 37' 18.7"	E 34 43' 54.1"	1497	Manure	No
F9	Nyakoe	S 0 47' 20.2	E 34 43' 52.2"	1495	Manure	No
F10	Nyakoe	S 0 38' 5.4"	E 34 43' 49.5"	1545	Manure	No
MACHAKOS COUNTY						
F1	Kathiani	S 1 30' 2.1"	E 37 17' 58.3"	1958	Manure	No
F2	Kathiani	S 1 30' 4.3"	E 37 17' 18.2"	1922	Manure	No
F3	Kathiani	S 1 30' 34"	E 37 18' 2.8"	1919	Manure	No
F4	Kathiani	S 1 30' 3.9"	E 37 18' 6"	1913	Manure	No
F5	Kathiani	S 1 30' 22.4	E 37 17' 59"	1885	Manure	No
F6	Kathiani	S 1 30' 26.8"	E 37 17' 57.3"	1876	Manure	No

F7	Kathiani	S 1 30' 28.4"	E 37 17' 56.9"	1875	Manure	No
F8	Kathiani	S 1 30' 22.2"	E 37 17' 52.2"	1890	Manure	No
F9	Kathiani	S 1 30' 23.9"	E 37 17' 54"	1884	Manure	No
F10	Kathiani	S 1 30' 26.9"	E 37 17' 55.5"	1883	Manure	No