MORPHOLOGICAL AND TRANSCRIPTOME ASSESSMENT OF SELECTED AFRICAN TOMATO AND EGGPLANT ACCESSIONS FOR IMPROVED ADAPTABILITY, QUALITY AND QUANTITY

MUNGAI GRACE WACHEKE

DOCTOR OF PHILOSOPHY
(Biotechnology)

JOMO KENYATTA UNIVERSITY OF AGRICULTURE AND TECHNOLOGY

2019

Morphological and Transcriptome Analysis of Selected African Tomato and Eggplant Accesions for Improved Adaptability, Quality and Quantity

Mungai Grace Wacheke

A thesis submitted in fulfillment of the requirements for the award of the Degree of Doctor of Philosophy of Science in Biotechnology in the Jomo Kenyatta University of Agriculture and Technology

2019

DECLARATION

This thesis is my original work and has not been presented for any award in any university.

Signature……Date……...

Grace Wacheke Mungai

This thesis has been submitted for examination with our approval as university supervisors.

Signature…….....Date…….....

Prof. Willis Owino, PhD

JKUAT, Kenya

Signature.....Date.....

Dr. Jane Ambuko, PhD

JKUAT, Kenya

Signature……..Date……

Prof. Aggrey Bernard Nyende, PhD

JKUAT, Kenya

DEDICATION

I dedicate this work to my loving husband Paul, my daughters; Esther and Josphine and my son Parmenas. Thank you for your immeasurable support. I can't thank you enough.

ACKNOWLEDGEMENT

I thank God almighty for giving me the chance, human and financial resources and ability to carry out this project. I acknowledge the support of my supervisors; Prof. Willis Owino, Prof. Aggrey Nyende. Dr. Jane Ambuko and Prof. Jim Giovannoni for standing with me during the whole project process. Thank you for your guidance and patience.

I acknowledge the support provided by U.S. Agency for International Development through the Partnerships for Enhanced Engagement in Research (PEER) program Sub-Grant Number: PGA-2000003426 to Prof. Willis Owino and Prof Jim J. Giovannoni without which this work would not have been done. I thank the National Commission for Science Technology and Innovation (NACOSTI) for the partial financial support to enable me finish the project.

I aknowledge the support of Dr. Tsvetelina Stoilova of AVRDC- World Vegetable Center in Arusha, Tanzania, seed repository for the provision of the African tomato and African eggplant seeds that were used in this project.

I also thank members of Prof. Jim Giovannoni lab at the Boyce Thomson Institute for Plant Research (BTI), Cornell University, USA ;- Prof Jim Giovannoni, Dr. Julia Vrebalov, Yimin Xu, Betsy Ampofo and Dr. Zhangjun Fei, for technical assistance during the study. Many thanks to Bryan Bell (BTI) and Mary Wanjiru (JKUAT) for their support in taking care of the African tomato and African eggplants in the respective greenhouses.

I acknowledge the help of Dr. George Michuki and Dr. Cecilia Waruhiu of The Africa Genomics Centre and Consultancy for their support during the data filtering and downstream analysis. I also thank the support of my collegues at the Institute of

Biotechnology Research;- Dr. Cecilia, Dr. Neondo, Dr. Amanuel, Dr. Mibei, Mercy, Rotich, Catherine and Jane for their support. I appreciate your continued support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

APPENDICES

LIST OF ABBREVIATIONS

DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
PCR	Polymerase chain reaction
NGS	Next generation sequencing
RNA-seq	RNA-sequencing
QTL	Quantitative Trait locus
MAS	Marker assisted selection
RFLP	Restriction Fragment Length Polymorphism
SNPs	Single Nucleotide Polymorphisms
GWAS	Genome-Wide Association Study

ABSTRACT

RNA sequence (RNASeq), a high – throughput sequencing technology that can generate information for characterizing RNA content and the composition of a given sample**.** African Solanaceae crops survive in very harsh conditions and have been used in Africa for decades as a source of food and for income. Despite the economic and nutrition importance of the African tomato and African eggplants have been underutilized because of lack of adequate information on their morphological and transcriptomic analysis. The main objectives of this study were to analyze the African tomato and eggplant transcriptome using morphological and RNASeq techniques to unveil genes associated with resistance to biotic and abiotic stress as well as genes associated with able to improve productivity and quality of the currently cultivated exotic varieties. Complete randomized block design and balanced block chain design were used to set up morphological and transcriptomic experiments respectively. A total of 67 African tomato and 72 African eggplant accessions morphological traits were analyzed using the standard phenotype descriptors. Out of the initial accessions, ten African eggplants and seventeen African tomato accessions were selected and planted, each with three biological replicates for transcriptomic analysis. Genstat and Darwins softwares were used to analyze data generated from vegetative and reproduction morphology characterization experiments. Vegetative and reproduction morphological traits were sampled for mophological characterization. For transcriptome analyses, leaf samples were collected 3 weeks after planting while fruit samples at mature green, mature breaker and mature red stage. Raw Sequences were cleaned and filtered using the Next Generation Sequencing Tool kit while TopHat software was used to identify differential gene expression, Single nucleotide polymorphism mining, gene ontology and gene of interest comparison. The filtered sequences of African tomato were aligned to the reference genome using TopHat while in African eggplants, *de novo* assembly was done using Trinity software. An in house reference genome was constructed using all the 80 African eggplant sample sequence from all the four stages. The generated reference genome was used to align the African eggplant accessions using TopHat. The findings of the morphological study reveal significant variation within African tomato and eggplant contributed by their plant growth habit and fruit morphology. A total of 329,018,858 and 303,754,051 sequences for African eggplant and tomato respectively, were obtained after filtering. A total of 18,129 and 173,194 genes were differentially expressed from African tomato and eggplant respectively. Significant differential gene expression was observed between the various fruiting stages in both African tomato and African eggplant at α 0.001. African tomato, African eggplant accessions expressed unknown genes which could be characterized more to unveil novel genes among them for breeding. The generated African eggplant reference genome is of great use in improving the current eggplant database. The SNPs in African tomato and African eggplant accessions revealed that variation among the accessions was more dependant on epigenetic factors since they grouped according to geographical locations. Variation in the African tomato and African eggplant accession transcriptomes was mainly dependent on the fruit development stage rather than the accession. This study revealed that environmental variables have an impact on gene flow patterns, which may influence spatial and progressive dispersal of genetic variation. For the gene comparative

analysis, gene expression profiles indicated that African tomato and African eggplant are closely related. There were many similar genes in the African tomato and African eggplant that confer resistance to abiotic, biotic stresses, that can be used to improve shelf life and yield to cultivated commercial cultivars. Transcriptome analysis is able to reveal genes that are being actively expressed in specific tissue and species of interest.

CHAPTER ONE

INTRODUCTION

1.1 African tomato and African eggplants

African tomato originated from South America, and were possibly brought to African continent by the Portuguese or Spanish from the 17th Century (Hirakawa *et al.*, 2014). However, having been in Africa for that long period they have acquired unique traits to enable them adapt to African climatic conditions. African Tomato and African eggplant are important crops in terms of their economic and nutritional value. They are associated with resistance genes for biotic, abiotic stress better shelf life and higher yielding (Mibei *et al*,. 2017).

African eggplants are widely distributed across sub-Saharan Africa and in many places throughout non arid part of Africa and India because of their considerable improvements in the fruit sizes and shapes as compared to the cultivated eggplant species. The most cultivated species of the African eggplants, include *Solanum aethiopicum, S. macrocarpon* and *S. anguivi* species (Adeniji and Agatha, 2012). These crops are commonly consumed in Tanzania, Uganda and Rwanda and the fruity forms are important component of vegetable diet, sold in grocery stores and retail outlets (Adeniji and Agatha, 2012).

1.2 Transcriptome analysis

Several interspecific genetic linkage maps and comparative gene analysis have been constructed between cultivated *Solanaceae* species and their wild relatives with the

aim of identifying genes that can be incoperated to obtain mproved varieties to counter both biotic and abiotic challenges that reduce their yield (Shirasawa *et al.,* 2013). These maps allow identification of the genes responsible for interspecific phenotypic variations, including disease resistance, fruit size and shape, and plant architecture (Shirasawa *et al.,* 2013). However, few genetic studies have reported intraspecific variations due to their narrow genetic diversity (Shirasawa *et al,* 2013; Hamilton *et al*., 2012). Massive parallel sequencing and genotyping methods have contributed to progress in genetics and genomics today, these include; Next-generation sequencers (NGSs), such as HiSeq2500 (Illumina), the GS FLX system analyses of genomes of several organisms (Hamilton and Buell, 2012; Schatz *et al.,* 2012).

These techniques allow a huge number of the nucleotide variations to be identified cheaply within a relatively short period of time. Whole-genome or whole-transcriptome analyses have become an option for genetic non-model organisms and will soon be standard practice for molecular ecological studies (Ellegren *et al.,* 2012; Barrio *et al,*. 2016).

Materials from wild relatives, ancestors and accessions held in germplasm collections of crop species contains an under exploited wealth of genetic variation, and will therefore offer a useful gene pool to cope with existing and new breeding challenges. Exploiting wild and early domesticated resources has the potential to genetically enrich crops with alleles that can improve traits. These dynamic challenges regarding climate change has led to requirement for sustainable production and a growing demand for more and better food.

In this study, African tomato and African eggplant lines were morphologically characterized and sequenced to find genes responsible for biotic and abiotic stresses which can be used to reduce the crop yield loss.

1.3 Economic Importance of African tomato and eggplant

African It is cultivated on every continent except Antarctica. Global tomato production (tonnes) has grown by 47% from 2001 to 2011, (FAOSTAT, 2014). According to small starter entrepreneurship magazine, tomatoes are Africa's most consumed fruit (or vegetable); eaten by millions of people across our continent's diverse religious, ethnic and social groups. Both in a raw and processed forms, tomatoes are central to most African diets and remain a regular ingredient in many soups, stews, sauces and dishes across the continent (FAOSTAT, 2014). Unfortunately, Africa does not produce enough tomatoes to meet its own needs. Almost every country in Africa consumes more tomatoes than it produces. Africa already spends nearly $1 billion on importing tomato products.

African eggplant is among the oldest vegetables. It is an indigenous tropical African crop grown in most African countries for its nutritional, medicinal and economic values of the leaves and fruits (Chadha and Oluocha, 2003). Mibei et al., (2017) stated that African eggplant contains a lot of minerals, vitamins, carbohydrate and water which are important and highly beneficial for the maintenance of health and prevention of diseases. Chadha and Oluocha (2003), reported that garden egg as a vegetable, has been recommended to handle malnutrition problem in Africa, especially among women of childbearing age and children under 5 years old.

In 2011, 46.8 million tons of eggplant (*Solanum melongena*) was produced in the top four producing countries, namely China (27.7 million tons), India (11.8 million tons), Egypt (1.1 million tons) and Turkey (8.2 million tons), according to the Food and Agriculture Organization (FAO) of the United Nations (http://faostat.fao.org) (Yang *et al*., 2014a).

According to data from the FAOSTAT, 2014 annual global yield of eggplant was 51.3 megatons in 2014. Over 90% total eggplant production was from Asia during the period from 2010 to 2014. The global production of eggplant has been growing and reached 48.4 million tons in 2012, which was roughly one-third of the total production of tomato (FAOSTAT, 2014).

Nevertheless, eggplant has been much less recognized as a target for molecular genetics research than tomato and potato, probably because it is produced and consumed less widely, especially in Western countries. As described by Chen *et al.,* (2017), many of the agronomically important traits in eggplant are shared by other solanaceous crops, and in most cases, the genetics of these traits have been investigated in detail in the tomato and potato. From the botanical and agronomical points of view, compared with the two *Solanum* model species, eggplant has many unique aspects including extra-large fruit size, high tolerance to biotic and abiotic stresses, and parthenocarpy (Chen *et al*., 2017).

The major obstacle to wider cultivation of African eggplant is low seed germination rates and flower abscission, and limited knowledge about the African eggplant and African tomato genetically. In this project, phenotypic characteristics and transcriptome analysis using the RNASeq data from illumina platform was done to

observe the African tomato and African eggplants morphological and molecular diversity.

1.4 Problem statement

Cultivated *Solanaceae* crops production has been hampered by diseases and pests causing low yields which lead to low income and losses. African *Solanaceae* especially the African tomato and African eggplant have been thriving comfortably while the cultivated varieties are being overcome by the biotic and abiotic stresses. In Africa, the *Solanaceae* species have not been well characterized to determine their morphological and genes of intrest. Their potential to improve the already cultivated varieties has not been investigated. Cultivated *Solanaceae* crops are also improved hybrids with a narrow genetic base making them vulnerable to diseases and pests.

Eggplant is one of the most popular vegetable since it is delicious in taste, and is an excellent source of fibers, vitamins, minerals as well as certain polyphenols that exhibit antioxidant activities (Zouine *et al.*, 2012; Knapp *et al.*, 2013). Despite the widespread cultivation of eggplant, many factors continue to cause extensive losses from planting to harvest of eggplant, such as pests, diseases and weeds (Daunay *et al.*,2012). Low temperature, which affects pollination and fertilization, is also a serious constraint to the yield and quality of early-maturing eggplant varieties (Chen *et al* ., 2017). However, undesirable consequences such as malformed fruits, loss of flavour and drug residual, which are induced by inappropriate use of the plant growth regulators, would reduce the fruit value and have a potential risk to consumers (Chen *et al* ., 2017).

The cultivated tomato and eggplant is susceptible to many bacterial and fungal pathogens such as the *Verticillium dahlia* fungus as well as insects and nematodes (Collonnier *et al*., 2001), which cause significant yield losses. Therefore, improving resistance to biotic and abiotic stresses is one of the main objectives of eggplant and tomato breeding programs (Yang *et al*., 2014a).

Worldwide loss of biodiversity, low prices for major plant production commodities, interest in more sustainable and diversified agriculture, and increasing demand for new foods and plant products Scherr *et al.*, (2007) has triggered the understanding of the African *Solanaceae* species. Commercially, underexploited plant derived compounds also attract innovators and entrepreneurs as a source for novel materials.

1.5 Justification

Solanaceae species crosses with their wild relatives with varying degrees of difficulty; their African relatives can be used as sources of genes for crop improvement. African Solanaceae species are interesting resources of genetic variation for introgression breeding and comprise exclusive sources of many resistance genes for cultivated varieties (Kaushik *et al*., 2016).

With the availability of the complete tomato genome (Sol Genomics network, 2012), it has become possible to perform genome-wide transcriptome analysis to study gene expression patterns across different plant tissues and under different conditions without *de novo* assembly. Next-generation high-throughput RNA sequencing technology (RNA-seq) using massively parallel sequencing has revolutionized transcriptome analysis. RNA-seq can detect all expressed genes without the

generation of an array or probes, with reduced background noise and large dynamic range. This is particularly important in species such as tomato and eggplant.

Characterization of the African tomato and African eggplant which harbors a lot of resistant genes against many diseases will help develop both biotic and abiotic stress-resistant varieties as well as improved shelf life. There is a need to identify useful genes that can be exploited in breeding programs.

1.6 Objectives

1.6.1. General objective

Determination of the morphology and transcriptome of the African eggplants and AfricanTomato accessions for improved adaptability, yield and quality.

1.6.2 Specific objectives

1. To characterize the African eggplant and tomato using Morphological descriptors
2. To assess the transcriptome of the African tomato and eggplant using High Throughput Next Generation Sequencing
3. To determine transcriptome diversity of the African tomato and eggplant using Higher Throughput next generation sequencing.
4. To assess genes associated with pest resistance, disease resistance, drought resistance, effective yield and shelf-life quality.
5. To determine comparative analysis of the African eggplant and African tomato genes.

1.6.3 Null Hypotheses

1. Morphological characterization of African tomato and eggplant is not possible.

2. Determination of African tomato and African eggplant transcriptome is not possible.

3. The African tomato and African eggplant plants do not transcriptome diversity

4. African tomato and African eggplant does not have genes that can be used for improvement

5. There similarity between the African tomato and African eggplant accession genes is not significant.

CHAPTER TWO

LITERATURE REVIEW.

2.1 African Solanaceae crops

The Solanaceae comprises a number of economically important food crops. These crops are important to agriculture, food security, human nutrition and health (Ray-Yu and Ojiewo, 2011). They include globally-consumed peppers (*Capsicum* sp), potato (*Solanum tuberosum*), cultivated tomato (*S. lycopersicum*), wild tomato (*S. hirsutum, S. peruvianum and S. pennellii*), cultivated eggplant (*S. melongena*), African eggplants (*S. aethiopicum, S. macropcarpon* and *S. anguivi*) and African nightshades (*S. scabrum* and *S. villosum*) (Knapp *et al.*, 2004). Many members of the family contain potent alkaloids, and some are highly toxic, but many, including tomatoes, potatoes, eggplant, bell/chili peppers, and tobacco are widely consumed. Although their fruits and vegetables are widely used, the leaves of these plants can be toxic to humans due to the presence of alkaloids, such as α-tomatine in tomato leaves. The rich source of alkaloids and other secondary metabolites makes Solanaceae plant species have a high potential for drug discovery (Lee, 2007).

2.1.1 The African eggplant

African eggplants are used as ethnomedicinal herbs. They are the wild relatives of the cultivated eggplants (Sękara *et al.*, 2007; Stàgel *et al.*, 2008). They are important fruit and leaf vegetables in Africa (Schippers, 2000) because both leaves and fruits are edible. The African eggplants are possibly native to Africa (Bukenya and Carasco, 1999). They are widely dispersed across sub-Saharan Africa and in many

places throughout non arid part of Africa as compared to the cultivated eggplant species. The comonly cultivated species of the African eggplants in Africa and India, includes the *Solanum aethiopicum, S. macrocarpon* and *S. anguivi* species because of their diverse fruits sizes and shapes Adeniji and Agatha, 2012).

The African eggplants are highly polymorphic and variable in plant structure, fruits and leaf characters. The leaves are large, hairy on the underside and alternate on the stems. Leaf prickles and hairiness are more pronounced in wild types (Jagatheeswari, 2014). The fruit of an eggplant is a fleshy berry with diverse colours ranging from black, white, green, shiny purple and yellow and the skin has stripes and patches. The shape of fruits varies from round to oblong, cylindrical, long and oval in shape. (Frary *et al.*, 2007).

The leaves and fruits of African eggplants are bitter taste, this could be accredited to the presence of alkaloids (mainly glycoalkaloids and phenolic compounds) as described by Abukutsa-Onyango (2003). The presence of alkaloids, phenolic acids and anthocyanins has led to the eggplant being used in traditional medicine (Frary *et al.*, 2007). The edibility and use of the African eggplant in traditional systems of medicine is mainly determined by bitterness for a long time (Chadha and Mndiga, 2007). There is increasing evidence that intake of their leaves and fruits reduce the incidence of chronic diseases including diabetes and artherosclerosis (Kwon *et al.*, 2008; Elekofehinti *et al.*, 2012, Mibei et al,. 2017). African eggplants are associated with resistance to drought, floods, molds mldews and certain soil- borne plant pathogens. (Sękara *et al.*, 2007). They can be intercropped or grown in small pots providing a high yield of fruit from a small area.

Domestication of the African eggplants, human selection, mutation, hybridization and natural inter-crossing have resulted in expansion in fruit size, colour and shape while decreasing fruit bitterness and leaf prickliness (Frary *et al.*, 2007). like many other crops indigenous to Africa, the African eggplants are easy to grow making it a good plant for research (Abukutsa-Onyango, 2003; Mibei et al 2017).

2.1.2 African tomato

African tomato plants are members of the Solanaceae family, native to South America, mainly across the regions of Peru, Chile, Ecuador, Colombia and the Galapagos Islands (Kole, 2007). In Africa, they are used as food, fruits salads and for medicinal purposes (Grandillo and Chetelat, 2011). They are a rich source of genes, which harbor genetic diversity that yields heritable variation in fruit chemistry. This could be exploited to identify genes regulating their synthesis and accumulation (Lee *et al.*, 2014).

They have diverse morphological characteristics based on their different geographical distribution which is reflected in their genetic diversity. According to Grandillo and Chetelat, (2011), physical barriers such as deserts and mountains have kept the African tomato species genetically distinct. African tomatoes are important for breeding, as sources of desirable traits, and for evolutionary studies (Bolger *et al.*, 2014).

African tomatoes are rich sources of folate, vitamin C, and potassium, with carotenoids being the most abundant phytonutrients (Shalom *et al.,* 2011). Lycopene is the most prominent carotenoid followed by beta-carotene, gamma-carotene and

phytoene as well as several minor carotenoids (Mibei *et al*., 2017). The antioxidant activity of lycopene as well as several other carotenoids and their abundance in African tomato make it a rich source of antioxidant activity. African Tomato also contains several other components that are beneficial to health, including vitamin E, trace elements, flavonoids, phytosterols, and several water-soluble vitamins (Mibei *et al*., 2017).

Like all known species of the genus *Lycopersicon*, tomato is a diploid; has 2n=24 chromosomes, and a genome size of 2.0 pg/2c 9.5×105 Kb/1c (950 Mbp), which is composed of 77% heterochromatin and 23% euchromatin (Lee *et al*,. 2012). The genus *Lycopersicon* includes both self-incompatible and self-compatible species, with the latter varying in their degree of out crossing.

Tomato is self-pollinating, but can easily hybridize within the species or cross with wild relatives under appropriate conditions, thus permitting gene introgression from wild relatives (Saito *et al*., 2011). Tomato improvement has increased by the exploitation of exotic resources and the introgression of new valuable genes into the tomato gene pool. The replacement of inbred lines by hybrids has remarkably increased yield, while the genetic gain rate has been reduced due to low genetic diversity.

Some studies have been done using the wild tomato species to genetically improve the domesticated tomato which has a poor genetic diversity from inbreeding during the domestication of the tomato (Zhang *et al.*, 2006).

Comparative genomics within this various genera and species and their wild relatives have greatly accelerated understanding of their genome evolution and the genetic mechanisms that confer phenotypic diversity to these species diversity within cultivated tomatoes (Wu and Tanksley, 2010).

2.2 Origin and distribution of African tomato and eggplant.

Solanacea crops originated from South America, and were possibly brought to Africa continent by the Portuguese or Spanish from the 17th Century (Hirakawa *et al.,* 2014). However, having been in Africa for that long period they have acquired unique traits to enable them adapt to them to African climatic conditions.

2.3 Constraints of African tomato and eggplant production

Biotic and abiotic factors have been attributed to low yields and the high cost of production. For instance, farmers used excessive pesticides to control pests and diseases. According to Osei *et al.,* (2014), low diversity among commercial tomato varieties has been identified as one of the major factors that predispose the crop to biotic and abiotic constrains. Crop accessions that have been used widely in breeding work and are always thought to harbour valuable traits lost among cultivated varieties and the exploitation of such traits increases research findings and knowledge of the genetic variability which facilitates breeding for wider geographic adaptability (Hanson *et al*., 2007).

In Africa, there are large numbers of tomato accessions stored in gene banks whose phenotypic and genotypic traits are largely undocumented. Knowledge of this

diversity is important to broaden the genetic resource base for future tomato crop improvement programs (Tembe *et al.*, 2017).

2.4 Genetic resources

2.4.1 Tomato and Eggplant genetic resources:

Tomato and eggplant genetic resources include old and new cultivars, primitive cultivars, breeding lines, accessions or heirloom tomatoes, and wild related species (Shallom *et al* ., 2011). Within cultivated tomato, genetic variation is very low; thus, there has long been an interest in searching for genes in exotic and primitive germplasm and closely related species. New breeding strategies now permit an in depth study and effective exploitation of the genetic diversity of wild relatives and accessions.

2.5 Applications of molecular markers

Various molecular markers have been used in *Solanaceae* species studies, for germplasm characterization Rao *et al.,* (2016), the evaluation of genetic diversity and species relationships within the genus *Lycopersicon* (Kochieva *et al.,* 2002), the genetic structure and diversity of wild *Lycopersicon* species populations determination of relationships between *Solanaceae* species and fingerprinting (He *et al.,* 2012; Rao *et al.,* 2016), the purity control of cultivars and variety identification (Bredemeijer *et al.,* 2002; Cook *et al.,* 2003), the identification of markers linked to important genes, map-based gene cloning, and genetic mapping (Saliba-Colombani *et al.,* 2000; Areshchenkova and Ganal, 2002).

Based on molecular markers, levels of intraspecific species polymorphism have been estimated to be very low. This is attributed to the self-pollination and self-fertilization of exotic tomato cultivars, combined with their narrow genetic base (He *et al*., 2012; Tam *et al*., 2005; Sabatini *et al*., 2006; Kochieva *et al.,* 2002). The molecular characterization of accessions has been carried out with many kinds of markers i.e. AFLP, next generation sequencing resources among others (Rao *et al*., 2016). The completion of the tomato genome, (Sol Genomics Networks, 2012), has enabled a lot of research in tomato and other related *Solanaceae* crops.

2.6 Transcriptome analysis

Due to technological limitations at present, sequence information from transcripts cannot be retrieved as a whole, but is randomly decomposed into short reads of up to several hundred base pairs. The assembly of this short reads can either be de novo assembly in absence of a reference genome or transcriptome information or mapping assembly where reads are aligned reference genome in the database. According to Xu *et al*., (2015), *De novo* transcriptome assembler Trinity, efficiently reconstructs full-length transcripts across a broad range of expression levels and sequencing depths to assemble Clean reads from the two Solanum into contigs and clustered into transcripts.

Transcriptome analysis of very large genomes have been done using the *de novo* sequencing of the RNA- seq data especially for species of considerable biological interest for reasons that relate to factors such as their evolutionary significance or economic importance. Examples include, pea (Franssen *et al*., 2011), chestnut (Barakat *et al*., 2009), and Japanese knotweed (Hao *et al*., 2011).

De novo RNA-seq has been used to identify genetic polymorphisms which has a great potential as a platform for molecular breeding, wherein multiple cultivars or close-related species with variations in traits of interest are sequenced and genetic variation is identified (Zhang *et al.*, 2012; Haseneyer *et al*., 2011). The most commonly used sequencing platforms are the sequencing by synthesis based GA / HiSeq / MiSeq machines from Illumina and the sequencing by ligation SOLiD system, Semiconductor chip based IonTorrent system, Helicos' solid phase based Genetic Analysis Platform and the single-molecule real-time sequencing-based approach from Pacific Biosciences or Oxford Nanopore (Mardis 2008; Eid *et al.*, 2009; Raz *et al*., 2011; Chung et al., (2014)

Transcriptome analysis by next-generation sequencing (RNA-seq) allows analysis of a transcriptome at matchless resolution. RNA-seq has an upper hand over other techniques because it does not rely on a prior information on the sequence under investigation, thereby allowing analysis of a poorly characterized species (Hoeijmakers *et al*,. 2012).

RNA-seq technique has been widely used for species transcriptome characterization (Hoeijmakers *et al*,. 2012). Transcriptome analysis is able to reveal genes that are being actively expressed in specific tissue and species of interest, and also facilitate the discovery of potential molecular markers especially for *de novo* assembly of non-model organisms (He et al., 2012, Li et al., 2014, Sun et al., 2013). The availability of sufficient genome or transcriptome data are potentially useful for studies on differential gene expressions, gene regulatory mechanisms, and molecular marker application (Waiho *et al*., 2017).

The first RNA-seq studies in non-model organisms used transcriptome information obtained by sequencing from a single individual or a pool of individuals to construct microarrays for quantifying individual gene expression (Vera et al., 2008). RNA-seq has brought about a decrease in costs, increasing yields and improved bioinformatics data processing making it possible to obtain both sequence information and a measure of gene expression for several individuals directly by sequencing (Wolf et al., 2010).

RNA-seq captures a wider range of expression values, scales linearly even at extreme values (count tables) whereas microarrays show saturation of analogue type fluorescent signals and will soon be the standard even for large experiments. (Marioni *et al*., 2008, Nookaew *et al*., 2012). RNA-seq further provides information on RNA splice events; these are not readily detected by standard microarrays (Mortazavi *et al*., 2008). Microarrays also has a disadvantage of introducing biases in gene expression measurements due to its propensity for cross-hybridization.

Another advantage of RNA-seq over other next-generation approaches that it provides information on RNA splice events and reduces the genome to a more manageable size like restriction-site-associated DNA tags (Mortazavi *et al*., 2008; Elshire *et al*., 2011). RNA-seq data is directly derived from functional genomic elements, usually protein-coding genes. It allows users to investigate differential gene expression patterns between populations, for example in the context of speciation (Wolf *et al.*, 2010) or eco-type-specific adaptation (Lenz *et al*. 2013).

2.7 RNA-seq and Single nucleotide polymorphisms (SNPs)

High-throughput sequencing of mRNA which was primarily developed to analyse global gene expression in different tissues is an efficient way to discover coding SNPs. Next-generation sequencing (NGS) technology has produced immense biological data (Liu *et al.,* 2013). When multiple individuals with different genetic backgrounds are used, RNA-Seq is very effective for the identification of SNPs (Sathya *et al.*, 2014).

RNA-seq reveals information on sequence variation at individuals' genomes and transcriptomes allowing inferring patterns of allele-specific expression that can be relevant to environmental response and adaptation before even being examined in the wild (Guo et al. 2004; Tirosh et al., 2009). Transcriptome analysis thus constitutes a meaningful resource to develop a large number of popular molecular markers such as SNPs and microsatellites (Sathya et al., 2014).

Among all classes of molecular markers, SNPs are the most attractive markers for various reasons. They are abundant in number, co-dominant in nature, and amenable for high throughput genotyping (Hamid *et al*., 2015). SNPs are useful for characterizing allelic variation, for genome-wide mapping, and as a tool for marker-assisted selection (Wencai *et al*., 2003).

SNP markers covering the whole genome of cultivated tomato have been developed and genome-wide association studies (GWAS) performed with the aim of understanding relationship between genetic and phenotypic variations in cultivated tomato, (Gopalakrishnan *et al*., 2014). SNPs can be mined from sequence data to characterize allelic variation, genome-wide mapping, to identify the biogeography

and origins of different species, and for marker-assisted selection (Ding *et al*., 2016; Yang *et al*., 2004).

2.7.1 Transctiptome diversity of the tomato

SNPs have been useful for characterizing allelic variation, for genome-wide mapping, and as a tool for marker-assisted selection (Wencai *et al.*, 2003). Ttranscriptome and phenotypic variations in cultivated tomato have been done successfully and SNP markers covering the whole genome of cultivated tomato were developed and genome-wide association studies (GWAS) were performed (Hirakawa et al., 2013).

Tomato has been used as a model plant in classical and molecular genetics, due to its autogamous diploidy and a relatively compact genome (950 Mb). In 2012, the tomato genome consortium, published the whole-genome sequence and Hirakawa *et al,*. (2013) inferred the functions of 200 SNPs among the transcribed sequences of cultivated tomato lines by determining their positions in predicted genes on the tomato genome. These results have accelerated the understanding of genetic mechanisms that confer phenotypic variations among tomato cultivars (Hirakawa *et al*., 2013).

2.7.2 Transcriptome diversity of the eggplant

The eggplant (*Solanum melongena L*.) is a vegetable crop species belonging to the *Solanaceae* family, which includes economically important species such as tomato (*S. lycopersicum L*.), potato (*S. tuberosum L*.), pepper (*Capsicum annuum L*.), and tobacco (*Nicotiana tabacum L*.), (Mishra *et al*., 2013).

Genetic diversity of any vegetable crop including brinjal arises either due to geographical separation or due to genetic barriers to gene flow. Whether differences in geographic origin (source) imply genetic distance in parental selection for hybridization is still a matter of polemic. Assessment of the diversity and relationships of the cultivated species facilitates the establishment of conservation strategies, the use of genetic resources in breeding programmes, and the study of the crop evolution (Mishra *et al*., 2013).

Sequencing and genotyping methods such as HiSeq2500 (Illumina), the GS FLX system analyses of genomes of several organisms have contributed to progress in genetics and genomics (Hamilton and Buell, 2012; Schatz *et al.*, 2012). Allowing enormous number of the nucleotide polymorphisms to be identified cheaply and within a fairly short period of time.

Next –generation sequencing technology methods helps obtain accurate SNP and INDEL profile information. There are different types of SNPs. SNPs can either be Synonymous or Non synonymous. A synonymous SNP is a coding SNP that does not change the protein sequence while a non- synonymous SNP is a coding SNP that changes the protein sequence (Lee *et al*., 2014).

The identified SNPs can be used for polymorphic analysis of germplasm collections, which allows genetic analyses such as Quantitative Trait Analysis (QTL) mapping, GWAS, and genomic selection (Varshney *et al.*, 2012). Large-scale SNP genotyping is often performed with commercially available array-based platforms, such as Infinium (Illumina), Golden Gate (Illumina), and Axiom Genotyping Solution (Affymetrix).

2.7.3 Other areas where SNPs have been utilized

The general application of NGS include SNP and other methods for discovery of variations, whose downstream usefulness includes; linkage map construction, genetic diversity analyses, association mapping and marker assisted selection (Nielsen *et al*., 2011).

In human, SNPs variations could account for over 90% of all phenotype differences,the variations in SNPs could be responsible for individual differences in the way they respond to certain diseases, response to drugs and serve as popular biomarkers in pharmacogenomics studies and their response to various treatments (Gopalakrishnan *et al*., 2014).

In this study, 17 African tomatoes and 10 African eggplant were sequenced using illumina platform to discover novel SNPs that were be used to estimate transcriptome diversity among the African tomatoes and African eggplant, using the ratio of the SNPs contributing to the phenotypic and transcriptome variation.

The identified SNPs were analyzed in four different Categories of samples (before fruiting, mature green, mature breaker and mature red) of the African tomato and eggplant processed by RNA seq technology.

CHAPTER THREE

GENERAL MATERIALS AND METHODS

3.1. Sampling sites and sampling

Sixty African tomato accessions and sixty seven African eggplant accessions were collected from germplasm conservation stations at the Asian Vegetable Center, Regional Center for Africa (AVRDC-RCA), Arusha, Tanzania (Table 3:1 and Table 3:2) .

3.2 Viability check and Pre- germination

Ten seeds of each accessions were planted on petri-dishes with wet paper for 10 days ensuring the paper did not dry out during this time. Accessions with over 70% germination were used for pre-germination on trays containing peat moss media in the greenhouses at Jomo Kenyatta University (JKUAT), Institute of Biotechnology Research laboratory.

After four weeks, germinated seeds were transplanted into potting bags containing well mixed forest soil and manure in the ratio of 3:1 and placed in the IBR greenhouse.

Table 3:1. African tomato accessions collected from AVRDC-RCA.

Origin	Accession name	Origin	Accession Name	Origin	Accession Name
Morocco	V1005871	**Ethiopia**	V1006825	**Madagascar**	RV101896
	V1005872		V1006826		RV102098
	V1005873		V1006832		RV102104
	V1005874		V1006833	**Egypt**	V1005889
	V1005875		V1006837		V1005895
	V1005876		V1006838		V1006480
	VI005878		V1006840	**South Africa**	V1006892
	V1005905		V1006847		V1007108
	V1005986		V1006865		V1007539
	V1005987		V1006869		V1007540
	V1005988	**Madagascar**	RV101884		V1008098
	V1005989		RV101885		V1008099
	V1005990		RV101887		V1008916
Kenya	Tindi 050580		RV101888		V1030375
	Tindi 050589		RV102100		V1030852
Ethiopia	V1006827		RV102102		V1035028
	V1006828		RV102107	**Tanzania**	V1006972
	V1006841		RV102109		RV102114
	V1006842		RV102111	**Mauritius**	V1030379
	V1006848		RV102112		V1030380

Key: Table showing the various African tomato used in this study and their country where they were sampled

Table 3:2. African Eggplant accessions collected from AVRDC-RCA

Country of origin	Accession no	Country of origin	Accession no	Country of origin	Accession no
Mali	RV100215	**Mali**	RV100328	**Senegal**	RV100248
	RV100217		RV100330		RV100259
	RV100234		RV100333	**Ivory coast**	RV100386
	RV100239		RVI00334	**Malawi**	RV1002100
	RV100240		RV100447	**Ghana**	RV100380
	RV100241	**Uganda**	RV100359	**Rwanda**	RV100346
	RV100242		RV100360	**Gabon**	RV100185
	RV100243		RV100364	**Burkina Faso**	RV100332
	RV100245		RV100377	**Unknown**	RV100218
	RV100247				RV100246
	RV100249	**Tanzania**	RV100161		RV100265
	RV100250		RV100163		RV100432
	RV100252		RV100166		RV100438
	RV100260		RV100169		RV100453
	RV100261		RV100190		RV100445
	RV100263		RV100199		RV100449
	RV100264		RV100431		RV100452
	RV100268		RV100511		RV100456
	RV100270	**Cameroon**	RV100335		RV100455
	RV100271		RV100342		RV1001201
	RV100273		RV100343		
	RV100274	**Madagascar**	RV100382		
	RV100327		RV100385		

Table showing the various African Eggplant used in this study and their country where they were sampled

3.3 Experimental design and layout

A complete randomized block design (CRDB) was used to set up an evaluation plot in an open field at the JKUAT farm. The 60 African tomato and 67 African eggplant accessions were sown in three blocks each containing 3 x 2 m three plots. Eighteen plants per accession were grown with a total of 1, 080 and 1,206 plants from all 60 and 67 African Tomato and Eggplant accessions respectively.

Standard uniform crop management practices were applied to all entries; this included labelling, watering daily, weeding regularly, and pruning using secateurs to remove the old lower leaves, according to the "Guide Agricole 2004" a vegetable cultivation manual, three main pesticides were sprayed to the crops against the three main pests infesting the African tomato and eggplant (Table 3:3).

Table 3:3. Pesticides used to spray against the common Solanaceae crops

Pest	Pesticide	Dose (l)	Infestation rate	Application (2wks)
Powdery Mildew	Microthiol	3.0g	Mild	Once
Promopsis Fruit Rot	Mancozeb 80 NP	2.0 g	Mild	Once
Fruit and shoot borer	Lambda-cyhalothrine	0.5ml	Mild	Once

3.4 General methodology for RNA-seq

3.4.1 Development of accessions for African tomato and African eggplant

A total of 17 of the 60 African tomato and 10 of 67 African eggplants accessions were selected for molecular characterization based on their unique morphological

characteristics like fruit colour, size, shape and fruit surface structure (Table 3:4 and Table 3:5). The accessions were planted in a greenhouse at the Boyce Thomson Institute for Plant Research (BTI), Cornell University, USA from March to May, 2015 under controlled standard growth conditions (16 h light/8 h dark conditions; 26°C day, 20°C night in Giovannoni Cornell mix soil).

Young leaves from the 3 weeks' old plantlets were harvested for RNA extraction. The plantlets were later transplanted to pots containing the Giovannoni soil mixture. The seedlings (one per pot) were grown in 15 cm-diameter pots containing Giovannoni mix growth media using complete block design with three replications. The spacing of 30 cm between plants and 50 cm between the rows was maintained uniform treatments of fertilizers, weeding, watering, pesticides and pruning. Fruit samples were collected at three stages i.e. mature green, mature breaker and mature red for RNA extraction and further molecular work.

Table 3:4. Selected African eggplants used in this study

S/no	Code no.	Accession Id	Country of origin	Recommended Naming
1	1e	RV100343	Cameroon	1e_RV100343_Cam_tal_ro_red
2	3e	RV100201	Unknown	3e_RV100201_UNK_Tal_ obl _ gre
3	4e	RV100332	Burkina Faso	4e_RV100332_BurF_Tal_ro_red
4	6e	RV100445	Unknown	6e_RV100445_UNK_Pro_ obl _ gre
5	10e	RV100265	Unknown	10e_RV100265_UNK_Pro_ro_red
6	13e	RV100432	Unknown	13e_RV100432_UNK_VPRO_ro_red
7	14e	RV100246	Unknown	14e_RV100246_UNK_Int_obl_red
8	17e	RV100327	Mali	17e_RV100327_Mali_int_ro_red
9	23e	RV100330	Mali	23e_RV100330_Mali_pro_ro_gre
10	28e	RV1001201	Unknown	28e_RV1001201_UNK_int_ro_red

Key : **In recommended naming;** Cam – Cameroon, tal- Tall, ro - round, UNK- Unknown, obl- oblong, gre- green, BurF- Burkina Faso, Pro- Prostrate, Vpro – very prostrate, int- intermediate

Table 3:5. Selected African tomato accessions used in this study

NO	Code No.	Acc. No.	Country of origin	Suggested naming of the samples inclusive of 2 phenotypes
1	1at	V1005987	Morocco	1A_V1005987_Mor_kid_red
2	2at	V1006833	Ethiopia	2at_V1006833_Eth_obl_red
3	4at	V1005872	Morocco	4at_V1005872_Kid_red
4	5at	VI005878	Morocco	5at_V1005878_Mor_ro_red
5	6at	RV102114	Tanzania	6at_Rv102114_Tanz_ro_red
6	7at	V1007108	S. Africa	7at_V1007108_S.Afr_obl_red
7	8at	Tindi 050580	Kenya	8at_Tindi_050580_Ken_ro_yel
8	9at	RV102112	Madagascar	9at_Rv102112_Mad_ov_pin
9	10at	Tindi 050589	Kenya	10at_Tindi_050589_kenya_round_yellow
10	11at	V1006838	Ethiopia	11at_V1006838_Ethiopia_round_red
11	12at	V1006842	Ethiopia	12at_V1006842_Eth_ro_yel
12	13at	V1006826	Morocco	13at_v1006826_Mor_kid_red
13	15at	V1005874	Ethiopia	15at_v1005874_Eth_ro_red
14	16at	V1030380	Mauritius	16at_v1030380_Mau_ov_red
15	17at	V1006892	S. Africa	17at_v1006892_S.Afr_ov_pink
16	18at	V1035028	S. Africa	18at_v1035028_S.Afr_ro_red
17	19at	V1005875	Morocco	19at_v1005875_Mor_ro_red

Key: In recommended naming; S. Afr- South Africa, Mor- Morocco, Eth- Ethiopia, Tanz -Tanzania, Ken - Kenya, Mau- Mauritius, Kid-Kidney, Obl- Oblong, Ro - Round, Ov- Oval

3.4.2. Sample collection

Leaf samples were collected from 3 weeks old tagged seedlings with sterile forceps and immediately placed in labelled falcon tubes containing liquid nitrogen. The leaf samples were ground in liquid nitrogen using a mortar and pestle to a fine powder which was kept at - 86ºC for further molecular work.

In both African tomato and African eggplants, fruit samples were collected at three fruiting stages; mature green, mature breaker and mature red. The seeds were removed using a sterile scalpel blade, the pulp was chopped into cubes and frozen using liquid nitrogen. A grinder / thriller was used to mill the frozen cubes into a fine powder and kept in labelled falcons' tubes and stored in – 86 ºC for further molecular work.

3.4.3 RNA extraction, quality checking, quantification and library construction

RNA extraction from both leaf and fruit samples was done using modified Trizol extraction (Appendix I) method followed by single strand specific RNA-Seq library construction as described by Silin *et al*., 2014 with modifications (Appendix II).

RNA quality check was done by observing the bands on 0.8% agarose gel (0.8 g agarose gel) dissolved in 100ml of Lithium acetate (pH 5.8) to confirm presence and quality of nucleic acids. The purity and nucleic acid concentration was determined by measurement of the absorbance at 260 and 280nm in a nanadrop ND-1000 Spectrophotometer (Nanadrop Technologies, Inc., Wilmington, DE, USA). Equal amounts of total RNA (10ng) were used for library construction following the Silin *et al.* (2014) protocol (Appendix II).

3.4.4 PCR Enrichment of the library

PCR enrichment of the library was done to attain a sequencing threshold as follows initial denaturation at 94°C for 2 min, followed by 12-16 cycles of amplification (98°C for 30s, 65°C for 30s, 72°C for 20s). Final extensions of 72°C for 2min, followed by a 4°C hold the enriched libraries were purified using 1.4 volumes of (AMPure XP) Sera Mag beads, and eluted with 20µl of Tris EDTA (pH 6.4) (Appendix II).

3.4.5 Barcoding of the libraries for multiplexed sequencing

Concentration of each library was measured using Quan-IT DNA HS assay kit (single –tube), equal amounts (20ng) of each barcode library were combined. Then the libraries were concentrated using 1.4 volume of Sera Meg (AMPure XP) and eluted with 10µ l TE (Appendix II).

3.4.6 Sequencing

Multiplexed libraries using 41 barcodes (1 lane) were sequenced at the Biotechnology Resource Center (BRC) using the Illumina platform (Hiq 2500). The raw data obtained was filtered through the standard Illumina pipeline. The filtered Fastq files were further subjected to more stringent quality control process using the Next generation sequencing kit (NGS QC Toolkit (v2.3)) to remove the low-quality reads and reads with adaptor/primer contamination. The high-quality filtered reads were used for further downstream processing.

3.4.7 Sequence quality control and validation

Using NGSQC was used to set parameters ;- phred score of Q < 20, and a high quality read length 70 %. Ribosomal contamination was filtered out from the high quality RNA – seq reads using Ribopicker v 0.4.3 (Lee *et al*., 2014; Wang *et al*., 2010). NGSTool kit was used to check the sequence GC content.

3.4.8 Biological replicates validation

Dendogram, scatter plot and principal component analysis (PCA) at ∞ 0.001 were used to check whether the replicates used for sequencing were true to type.

CHAPTER FOUR

MORPHOLOGICAL CHARACTERIZATION OF THE AFRICAN TOMATO AND EGGPLANT.

Abstract

Africa's *Solanaceae* crops including African tomato and eggplant are used as food and for sustenance. However, their full potential has not been tapped due to lack of information regarding their diversity. This study evaluated morphological diversity of sixty African tomato and sixty seven eggplant accessions collected from Asian Vegetable Research and development center- regional center for Africa (AVRDC-RC) in Arusha Tanzania. Random Complete block design was used while planting. Data collected was subjected to GenStat's and Darwin 6 softwares. In African tomato, a dendogram grouped the accessions into three clusters. Cluster 1, 2 and 3 had 26, 31 and 3 accessions respectively. Singletons formed suggest divergent morphological background useful for out-crossing to other accessions. Clustering was attributed to leaf blade colour, leaf vein colour, fruit surface texture and fruit colour. Overlapping accessions in biplot analysis revealed close relationship among the accessions studied. Considerably, wide morphological diversity was observed. Accessions V1006825, V1006841 and RV102112 clustered far from other accessions showing high divergence. Tindi 050589 and Tindi 050590 were closely located and far from other accessions. Analysis of variance showed significant phenotypic variations among the accessions at $P<0.05$. In African eggplant, the dendogram grouped the accessions into two main clusters with majority falling in cluster 2 revealing a narrow genetic base in the cluster. Many singletons formed suggest

divergent genetic background hence useful for out crossing to other accessions. Clustering was contributed by the plant growth habit, fruit shape and fruit colour. Cluster 1 was constituted by accessions exhibiting prostrate and very prostrate growth habit while cluster 2 was composed of intermediate, tall and very tall. Overlapping accessions in the biplot revealed close relationship between many of the accessions studied as well as a considerably wide diversity for a few accessions. Accessions RV100328, RV100194 and RV100346 clustered far from the rest showing high variation based on morphological characters. Analysis of variance showed significant phenotypic variations in the landraes at $P<0.05$. Eight of the 14 Principal component (PC) analysis were significant accounting for 70.6% variation. PC-1 accounted for 16.02 % whereas PC-2 accounted for 12.29%. The findings of this study reveal significant variation among African tomato and eggplant contributed by plant growth habit and fruit morphology.overall, the morphological characterization used in this study show that phenotypic markers are important in charactering african tomato and eggplants.

4.1 Introduction

The African tomato is an important fruit and vegetable crop. It is extensively used in salads as well as for culinary purposes. It is also used in various processed forms including; pastes, sauces, pulps, juices, ketchup and as flavoring ingredients in soups, meat or fish dishes (Osei *et al* ., 2014).

The fruit contains significant amounts of vitamin A and C, lycopene, beta-carotene, magnesium, iron, phosphorus, potassium, riboflavin, niacin, sodium and thiamine with antioxidant properties and potential beneficial health effects (Zhang *et al.,*

2009). Its consumption is believed to reduce entire cholesterol, LDL cholesterol and triglyceride levels in white blood cells, reducing cardiovascular risk related with type 2 diabetes, decrease risk of breast cancer, neck cancers and strongly protect against neurodegenerative diseases (Freedman *et al.,* 2008).

African eggplant, is also referred to as the bitter tomato, Ethiopian eggplant or nakati, Ethiopian nightshades, garden eggs, and mock tomato, these names are a result of its varied morphology, with ripe fruit often looking like a cross between an eggplant and a tomato (Osei *et al*., 2014). The orange-red fruit is eaten boiled, steamed, pickled, or in stews with other vegetables or meats. Young leaves are cut and used in soups.

African eggplants leaves are known to have extremely high beta carotene, ascorbic acid, calcium, iron and proteins. Vitamin E, folic acid, ascorbic acid, calcium, iron, proteins and riboflavin are also found in high quantities in eggplant fruit. Leaves also contain oxalic acid and alkaloids, which possess anti-inflammatory and immunosuppressive properties. The bitter taste in leaves is attributed to furostanol glycosides especially saponins (Sarah and Maina, 2008).

Both African tomato and eggplant exhibits extensive diversity and can be produced in marginal areas. The determination of variability among accessions is essential to the maintenance and utilization of germplasm resources (Mwirigi *et al.* 2009). Systematic study and evaluation of germplasm is of great importance for current, future agronomic and genetic improvement of the crop (Reddy *et al*. 2004).

To identify and estimate the genetic diversity of plants, various methods can be used including morphological and molecular markers. Morphological markers are used mostly to study of genetic variation in plant species since they are important analytical features for distinguishing genotypes (Osei *et al*., 2014). Exotic tomato and eggplant production has been greatly affected by biotic and abiotic stresses. However, their African counterparts are adapted to harsh growing conditions and seem to possess superior genes for adaptation to these biotic and abiotic stresses.

Knowledge of the African tomato morphological diversity is important to expand the genetic resource base for future crop improvement programs.The aim of present study was to determine the morphological diversity within different selected African tomato and eggplant accessions using standard morphological descriptors (INIBAP, 2003).

4.2 Materials and methods

4.2.1 Phenotypic characterization

Data was obtained from six tagged plants selected from the 18 plants of each accessions. Phenotyping was carried out using nine quantitative and nine qualitative traits to estimate the levels of variation among the African Tomato and Eggplant accessions (Table 4:1). Leaf length, width, plant height and width, leaf colour, petiole colour and internode height, data was collected at flowering stage, while the fruit data was collected at mature green, mature breaker and mature red stages.The data was recorded in excel spread sheet (Table 4:1).

Table 4:1. Qualitative and quantitative traits measured in African Tomato

	Quantitative Traits Measured		Qualitative traits observed
1.	Flower Heads (FH)	1.	Plant growth habit - (very tall, tall, intermediate, prostrate, very prostrate.)
2.	Fruit Length (FL)	2.	Fruit colour – (red yellow, purple, green, yellow with strips, white, orange)
3.	Fruit Width (FW)	3.	Fruit Shape – (Round, oval, oblong)
4.	Internode Height (IH)	4.	Fruit texture - (smooth, ridged)
5.	Plant Breadth (PB)	5.	Leaf base shape - (sessile, asymmetrical, symmetrical)
6.	Plant Height (PH)	6.	Leaf blade colour – (light green, green, dark green)
7.	Leaf Blade Length (LBL)	7.	Petiole colour- (light green, green, purple with green stripes, purple)
8.	Leaf Blade Width (LBW)	8.	Stem colour – (light green, green, dark green, purple with green stripes, purple, dark purple)
9.	Number of Leaf Lobules (LLL)	9.	Vein colour- (light green, green, dark green, purple with green stripes, purple, dark purple)

Source: INIBAP, (2003). Both quantitative and qualitative traits were measured. Vegetative traits were measured when the plants were 50% had flowered while reproductive data was taken at three fruiting stages: - mature green, mature breaker and mature red

4.2.2 Data analyses

Analysis of variance (ANOVA) was carried out to determine sources of genetic diversity of the measured nine quantitative traits. Means for each trait were separated by the least significant difference (LSD) at ($p<0.05$) to show the significant differences.

Phenotypic correlation coefficients were computed to examine the degree of association among the quantitative traits. Multivariate analysis of variance (MANOVA) was conducted to reveal the patterns of phenotypic diversity of quantitative traits studied. Means of each quantitative character were standardized before subjecting to principal component analysis (PCA) as was suggested by Reddy *et al.* (2004).

The standardized data of 11 quantitative traits were then used as an input for the PCA biplot loading and cluster analysis. Measurements of similarity and dissimilarity were derived by calculating the Euclidean distance between pairs of parameters (Mead *et al.,* 2002). Statistical analyses was done using the Darwin 6 and GenStat Discovery software, version number 12.1.3338, 12th, Edition.

4.3 Results

4.3.1 Morphological characterization stage of the African tomato accessions

Differences were observed in the various morphological characteristics Iin both vegetative and reproductive stages (Table 4:2). Data was collected at both vegetative and reproduction stages. Vegetative data included stem colour, petiole colour, leaf base shape, leaf colour, plant growth habit, height of the plant at 50% fruiting (Plate1, 2, 3 and 4). While at reproductive stage data included fruit shape, colour and texture (Plate 5).

Table 4:2. Qualitative Morphological variations in African Tomato accessions

	Accession	Growth Habit	Stem colour	Leaf Vein Colour	Leaf Blade Colour	Petiole Colour	Leaf Base Shape	Fruit Shape	Fruit Colour	Fruit Texture
1	RV102111	Determinate	Green + Purple	Light Green	Green	Green + Purple	Asymmetrical	Round	Red	Rough
2	V10050580	Determinate	Light purple	Purple	Green	Purple	Asymmetrical	Round	Yellow	Smooth
3	V1005990	Determinate	Green	Light Green	Green	Green + Purple	Asymmetrical	Round	Rough	Rough
4	TINDI/050589	Determinate	Purple	Purple	Green	Green + Purple	Asymmetrical	Round	Yellow	Smooth
5	V1008098	Determinate	Green	Green	Green	Green + Purple	Symmetrical	Round	Red	Rough
6	V1006826	Determinate	Green	Light Green	Green	Green + Purple	Asymmetrical	Kidney	Red	Rough
7	RV101887	Determinate	Green	Light Green	Green	Green + Purple	Symmetrical	Oval	Red	Smooth
8	RVI02I00	Determinate	Green	Green	Green	Green	Asymmetrical	Red	Red	Rough
9	RV102112	Determinate	Purple + Green	Light purple	Green	Green + Purple	Asymmetrical	Oval	Pink	Smooth
10	RV102098	Indeterminate	Green	Light Green	Light Green	Light Green	Asymmetrical	Oval	Red	Rough
11	RV102107	Indeterminate	Green	Green + Purple	Dark Green	Green + Purple	Asymmetrical	Red	Red	Rough
12	V1006833	Indeterminate	Green	Green	Green	Green + Purple	Symmetrical	Oblong	Red	Rough
13	V1006837	Indeterminate	Green + Purple	Light Green	Green	Green + Purple	Symmetrical	Oval	Red	Smooth
14	RV102109	Indeterminate	Light Green	Light Green	Dark Green	Green + Purple	Asymmetrical	Red	Red	Smooth
15	V1030379	Indeterminate	Green + Purple	Purple	Green	Green + Purple	Asymmetrical	Round	Red	Rough
16	V1005876	Indeterminate	Green + Purple	Green	Green	Purple	Symmetrical	Round	Red	Rough
17	RV102102	Indeterminate	Green	Green	Green	Green	Symmetrical	Round	Red	Rough
18	V1006827	Indeterminate	Green + Purple	Purple	Green	Purple	Asymmetrical	Round	Red	Smooth

	Accession	Growth habit	Stem colour	Leaf Vein Colour	Leaf Blade Colour	Petiole Colour	Leaf Base Shape	Fruit Shape	Fruit Colour	Fruit Texture
19	RVI02I04	Indeterminate	Green	Green	Green	Green	Symmetrical	Round	Red	Smooth
20	V1005879	Indeterminate	Green	Green + Purple	Dark Green	Green + Purple	Asymmetrical	Round	Red	Rough
21	V1005905	Indeterminate	Green	Green + Purple	Green	Green + Purple	Symmetrical	Round	Red	Smooth
22	V10035028	Indeterminate	Green	Green	Green	Green + Purple	Asymmetrical	Round	Red	Rough
23	V1005871	Indeterminate	Green	Light Green	Green	Green + Purple	Asymmetrical	Kidney	Red	Rough
24	V1006480	Indeterminate	Green	Light Green	Green	Green + Purple	Asymmetrical	Round	Red	Smooth
25	RV101885	Indeterminate	Green + Purple	Light Green	Green	Green + Purple	Asymmetrical	Round	Red	Smooth
26	V100578	Indeterminate	Green	Green	Green	Green	Symmetrical	Round	Red	Rough
27	V1006848	Indeterminate	Green	Light Green	Green	Green + Purple	Asymmetrical	Round	Red	Rough
28	RV101896	Indeterminate	Purple + Green	Light Green	Green	Green	Symmetrical	Round	Red	Rough
29	V1006869	Indeterminate	Green	Green + Purple	Green	Green + Purple	Asymmetrical	Round	Red	Smooth
30	V1005878	Indeterminate	Green	Green	Green	Green + Purple	Asymmetrical	Round	Red	Rough
31	V1005895	Indeterminate	Green	Green	Green	Green + Purple	Asymmetrical	Round	Red	Round
32	V1008099	Indeterminate	Green	Light Green	Green	Green + Purple	Asymmetrical	Round	Red	Smooth
33	RVI006832	Indeterminate	Green	Green	Green	Purple	Symmetrical	Round	Red	Smooth
34	V1030375	Indeterminate	Green	Green	Green	Purple	Symmetrical	Round	Red	Smooth
35	V1006865	Indeterminate	Green	Light Green	Green	Green + Purple	Asymmetrical	Round	Red	Smooth
36	V1005986	Indeterminate	Green	Green + Purple	Dark Green	Green + Purple	Asymmetrical	Round	Red	Rough

	Accession	Growth habit	Stem colour	Leaf Vein Colour	Leaf Blade Colour	Petiole Colour	Leaf Base Shape	Fruit Shape	Fruit Colour	
37	V1005872	Indeterminate	Green	Green + Purple	Dark Green	Green + Purple	Symmetrical	Kidney	Red	Rough
38	V1005989	Indeterminate	Green + Purple	Green	Green	Green + Purple	Asymmetrical	Oval	Red	Rough
39	V1008916	Indeterminate	Green	Light Green	Green	Green + Purple	Asymmetrical	Round	Red	Smooth
40	V1006892	Indeterminate	Green	Light Green	Light Green	Green + Purple	Asymmetrical	Oval	Pink	Smooth
41	LO5942	Indeterminate	Green	Light Green	Green	Green + Purple	Asymmetrical	Oval	Red	Smooth
42	RV101888	Indeterminate	Green + Purple	Green	Green	Green + Purple	Asymmetrical	Round	Red	Smooth
43	V1006838	Indeterminate	Green	Green	Light Green	Green + Purple	Symmetrical	Round	Red	Rough
44	V1005987	Indeterminate	Green	Light Green	Green	Green + Purple	Asymmetrical	Kidney	Red	Rough
45	V1006840	Indeterminate	Green	Light Green	Green	Green + Purple	Asymmetrical	Kidney	Red	Rough
46	RV02114	Indeterminate	Green + Purple	Light Green	Green	Green + Purple	Asymmetrical	Round	Red	Smooth
47	V1006841	Indeterminate	Green + Purple	Light Green	Green	Green + Purple	Asymmetrical	Round	Red	Rough
48	V1009873	Indeterminate	Green	Light Green	Green	Green + Purple	Asymmetrical	Round	Red	Rough
49	V100852	Indeterminate	Green	Light Green	Green	Green + Purple	Symmetrical	Round	Red	Rough
50	V1005875	Indeterminate	Green + Purple	Light Green	Green	Green + Purple	Asymmetrical	Round	Red	Rough
51	V1007540	Indeterminate	Green + Purple	Light Green	Green	Green + Purple	Asymmetrical	Round	Red	Rough
52	RVIOI884	Indeterminate	Green	Light Green	Green	Green	Asymmetrical	Round	Red	Rough
53	V1006842	Indeterminate	Green	Light Green	Green	Green	Asymmetrical	Round	Yellow	Smooth
54	V1006847	Indeterminate	Green	Green + Purple	Green	Green	Asymmetrical	Round	Red	Smooth
55	V1006972	Indeterminate	Green	Green + Purple	Green	Green	Symmetrical	Oval	Red	Rough
56	V1006828	Indeterminate	Green	Green	Green	Green + Purple	Asymmetrical	Round	Red	Smooth
57	V1006864	Indeterminate	Green + Purple	Light Green	Green	Green + Purple	Asymmetrical	Round	Red	Rough
58	V1007539	Indeterminate	Green + Purple	Light Green	Green	Green + Purple	Asymmetrical	Round	Red	Rough
59	V1006825	Indeterminate	Green	Light Green	Green	Green + Purple	Symmetrical	Kidney	Red	Rough
60	V1007108	Indeterminate	Green	Light Green	Green	Green + Purple	Asymmetrical	Oblong	Red	Rough

Key: Qualitative descriptors used in morphological characterization of the African tomato**: Growth habits** - 51 accessions had indeterminate growth habit while only 9 had determinate growth habit**. Stem colour**- most accessions were green (39), few light purple (1), green + purple (18) and purple (1). **Leaf vein colour** - was either light green, green, green + purple or purple. **Leaf vein colour**: - light green, green, green + purple or purple. Leaf blade colour: - light green, green or dark green. **Petiole colour**: - Light green, green, green + purple or purple. **Leaf base shape**- leaf base was either asymmetrical (39) or symmetrical (21). **Fruit shape**- round, oblong, oval, kidney. Fruit colour- red, yellow, and pink. Fruit surface texture- either rough or smooth

Stem colour; 39 accessions had green stems while 18, 1 and 1 accessions had purple and green colour, purple and light purple colours respectively. The light purple on the stem indicated the presence of anthocyanins in this accessions (Plate 4:1). Peter *et al*. (2008) observed that anthocyanin in tomato can be observed on stem, petioles leaf veins and on fruit

Plant Growth habit and leaf colour

Variation was observed on the plant growth habit with 51 accessions having the indeterminate growth habit and only 9 accessions having determinate growth habit. Leaf colour varied from green to light green to dark green (Plate 4:2 and Table 4:1).

Petiole colour

Variation was observed in the African tomato accessions with most accessions (49) showing presence of anthocyanin.

Leaf margins, leaf vein, leaf lobbing margins and Leaf base shape: The African tomato showed either assymmetrical 39 accessions (c) or heart shape 21 accessions (d), at the point where the stalk meets the leaf. (Plate 4:1 and Table 4:2)

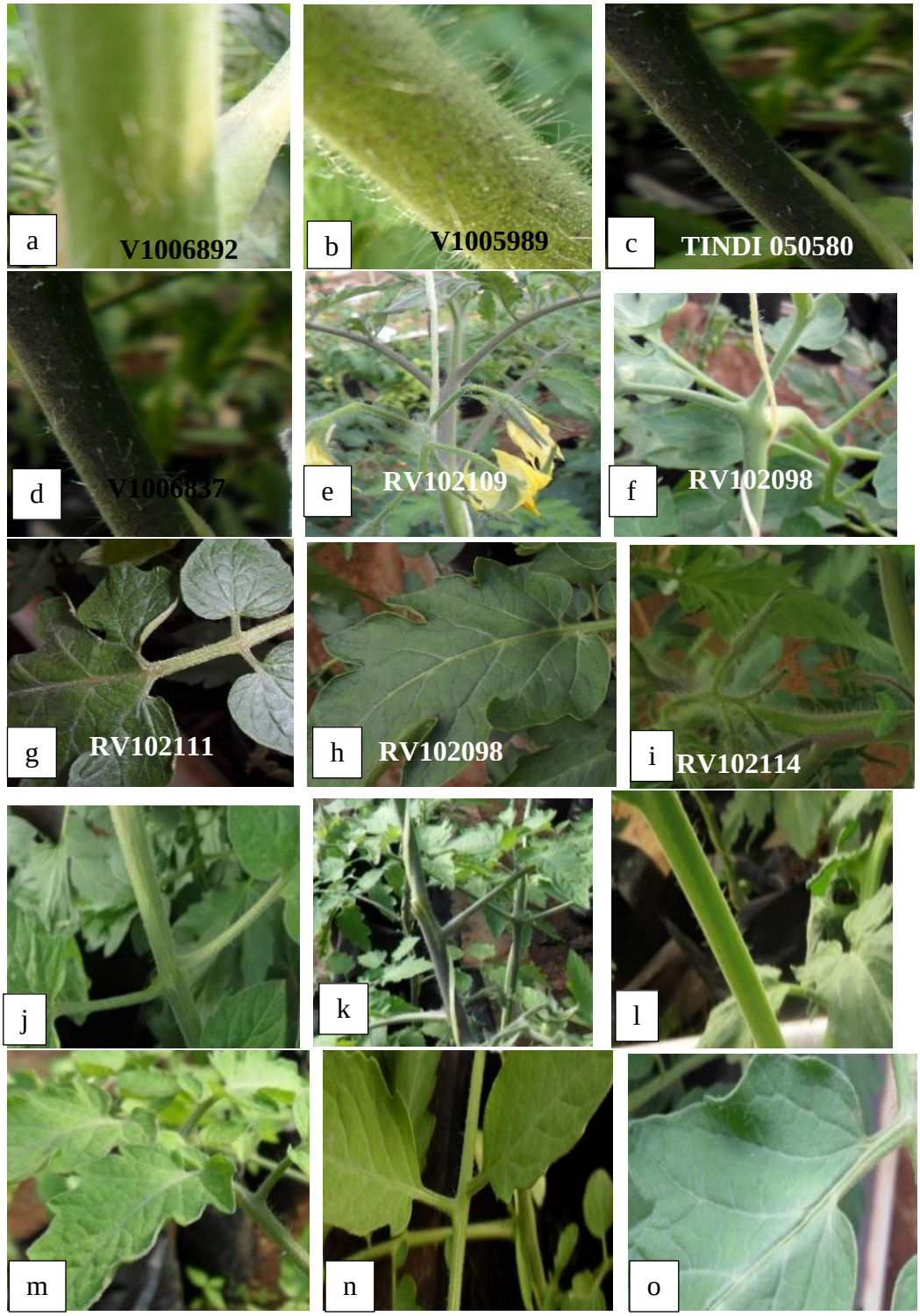


Plate 4:1. African tomato Vegetative morphological variation

Stem colour. The stem colour ranged from green; 39 accessions (a), green and purple; 18 accessions (b) purple;1 accession (c) light purple; 1 accession (d). **Plant growth habit.** The growth habit was either indeterminate (e) or determinate (f); Leaf colour ranged from dark green green (LG) 5 accessions (g), green (G) 53 accessions (h) and light green (LG) 2 accessions(f) **Petiole colour**. Petioles were either i-purple (4 accessions), j- Green (6 accession), k- green purple (49 accessions), and l- light green (1 accession) **Leaf morphology**. All accessions had lobbed edged leaves though with varying numbers and the vein colour varied from one accession to another; (i), symmetrical (n), or assymetrical (o)

Fruit shape: The African tomato exhibited diversity in shapes including; - Seven accessions showed kidney shape, two accessions showed oblong shape, eight had oval shape while most accessions (Forty-three) had a round shape (Plate 4:2 and Table 4:2). **Fruit colour**: African tomato had red colour (55). Two accessions had Pink while three accessions were yellow in colour (Plate 4:2 and Table 4:2). **Fruit surface texture**African tomato had either a rough (36 accessions) or smooth (24 accessions) fruit surface (Plate 4:2 and Table 4:2).

Plate 4:2. African tomato Fruit morphology. Shape: African tomato were either round (43- b, c, d, f), oval (8-g) oblong (2- e) or kidney (7-a) in shape. **Fruit surface**: The surface were either smooth (24- b, c, d, f) or ridged (36-a). **Fruit colour**: In colour, they were either red (55-a, b, e), pink (2-f), or yellow (3-c, d) in colour when ripe. **Shoulder colour**: The shoulder colours were either red (55- a, e, g), green (1- b), pink (2- f) or yellow (3- c, d).

4.3.1.1 Phenotypic diversity using qualitative traits

Phenotypic diversity for individual qualitative traits revealed a high degree of variation among the studied populations (Table 4:3) using the Shannon - Weaver diversity index to estimate phenotypic diversity of eight qualitative traits studied. The highest phenotypic diversity index (H') for traits studied recorded was 0.99 in petiole colour, stem colour and vein colour with a total mean phenotypic diversity index of 7.89. Substantial variation was observed in stem colour and vein colour (Table 4:3).

The existence of high variability as shown by diversity values recorded indicates that the diversity among the accessions is due to variation in traits. Overall, a high value of (H') represent a diverse and equally distributed classes for an individual trait. On the contrary, a lower value indicates less diversity since Shannon - Weaver diversity index accounts for abundance and evenness of a population present in a community.

Qualitative morphological parameters showed a close relationship between the 60 accessions. Morphological features used in the delimitation of the accessions were the presence or absence of ridge on the fruit, fruit habit, leaf orientation and general fruit morphology.

Table 4:3. Diversity index (H´) values explaining the genetic diversity of the African tomato accessions based on qualitative traits.

Qualitative Traits	Genetic Index (H')
Fruit colour	0.99
Fruit shape	0.99
Fruit texture	0.99
Leaf base shape	0.99
Leaf blade colour	0.99
Petiole colour	0.99
Stem colour	0.98
Vein colour	0.97
Total diversity	7.89
Average genetic index (h')	**0.99**

Key: The higher the genetic index (H') reflects high variability in terms of qualitative traits. Genetic index of the African tomato qualitative traits ranged from 0.97-0.99

4.3.1.2. Morphological variations in the quantitative traits of African tomato

The African tomato showed significant diversity in the quantitative traits that were selected for their characterization (Table 4:4). Plant growth habit was either determinate or indeterminate with plant height range of 40 cm (RV102111) to 120.2 cm (V1007108) . Fruit weigth ranged from 4.2g to 166.9g. Fruit length ranged from 2.4 cm (TINDI050580) to 18.8cm (V1005871) and fruit weight ranged from 6.3cm (TINDI 050580) to 25.9cm (VI005871) (Table 4:4).

Table 4:4 . Quantitative Morphological variations in African Tomato accessions based on (IPGRI, 2003)

	Accession	Plant Height	Plant growth habit	Leaf Blade Length	Leaf Blade Width	Leaf Lobules	Plant Breadth	Petal Number	Internode No.	Fruit Height	Fruit Length	Fruit Width	Fruit Weight
1	RV102111	40.0	Determinate	11.8	6.4	11.0	4.1	6.0	9.4	7.0	7.2	15.2	50.2
2	TINDI 50580	40.0	Determinate	6.3	3.6	10.0	2.9	5.0	3.9	6.0	2.4	6.3	4.2
3	V1005990	49.0	Determinate	11.9	6.2	10.0	4.1	6.0	7.6	9.0	9.1	17.8	77.5
4	TINDI/050589	52.0	Determinate	7.0	2.9	11.0	2.9	5.0	5.0	7.0	3.2	6.3	4.3
5	V1008098	64.6	Determinate	10.6	6.2	10.0	4.1	5.0	9.4	6.0	8.1	16.0	61.4
6	V1006826	66.2	Determinate	16.7	6.4	11.0	3.5	7.0	5.8	10.0	9.5	22.7	69.7
7	RV101887	67.2	Determinate	13.9	7.5	12.0	3.9	6.0	7.0	8.0	8.7	15.3	93.0
8	RVI02I00	68.4	Determinate	10.9	5.8	10.0	4.1	6.0	7.8	9.0	9.0	17.8	81.4
9	RV102112	68.4	Determinate	13.4	7.3	9.4	2.9	5.0	5.3	7.0	6.2	11.1	28.2
10	RV102098	71.8	Indeterminate	12.0	5.2	9.0	3.4	5.0	9.0	6.0	8.7	16.1	77.6
11	RV102107	74.0	Indeterminate	13.5	7.4	12.0	4.3	6.0	8.3	6.0	7.8	15.9	60.0
12	V1006833	74.2	Indeterminate	12.6	4.9	13.0	3.8	6.0	5.7	8.0	8.2	11.8	47.8
13	V1006837	74.2	Indeterminate	11.7	6.1	13.0	3.5	5.0	6.0	6.0	8.3	12.2	47.3
14	RV102109	75.6	Indeterminate	8.5	4.3	8.0	3.5	6.0	6.2	7.0	8.2	16.0	62.8
15	V1030379	76.8	Indeterminate	12.1	7.1	110.	3.7	6.0	7.0	7.0	9.7	21.3	118.8
16	V1005876	78.6	Indeterminate	7.8	4.5	8.0	3.9	7.4	5.6	12.0	9.7	21.3	117.6
17	RV102102	78.8	Indeterminate	13.3	8.6	10.0	4.9	5.2	5.8	10.0	8.7	15.4	66.4
18	V1006827	79.4	Indeterminate	9.9	6.5	11.0	4.5	4.0	4.8	8.0	4.1	8.2	8.3
19	RVI02I04	81.4	Indeterminate	14.4	8.7	11.0	4.8	5.0	5.4	6.0	9.2	19.1	97.3
20	V1005879	83.0	Indeterminate	14.4	7.7	9.0	4.7	6.0	10.6	9.0	8.2	16.5	128.5
21	V1005905	84.4	Indeterminate	11.8	8.4	7.8	5.1	8.0	7.6	8.0	7.3	13.4	39.6
22	V10035028	85.4	Indeterminate	9.7	5.4	8.6	3.4	6.0	8.8	7.0	7.6	15.2	53.8
23	V1005871	85.6	Indeterminate	9.4	5.2	8.0	3.8	8.0	6.3	8.0	18.8	25.9	166.9
24	V1006480	87.0	Indeterminate	8.9	5.1	10.0	3.4	6.0	6.4	8.0	6.4	13.4	34.8
25	RV101885	88.8	Indeterminate	11.3	5.6	13.0	4.7	6.0	6.6	6.0	8.8	17.7	80.8
26	V100578	89.6	Indeterminate	12.4	6.9	11.0	4.3	7.0	5.0	11.0	8.8	18.1	78.2
27	V1006848	90.6	Indeterminate	11.5	6.8	11.0	4.0	6.0	6.3	6.0	6.6	13.1	32.1
28	RV101896	91.8	Indeterminate	12.2	7.3	12.0	5.3	6.0	7.2	7.0	10.5	21.7	154.0
29	V1006869	92.6	Indeterminate	13.5	8.1	11.0	4.6	7.0	8.9	7.0	9.7	20.0	108.0
30	V1005878	93.0	Indeterminate	9.1	5.7	10.0	3.9	7.0	11.0	10.0	8.8	18.1	38.0

	Accession	Plant Height	Plant growth habit	Leaf Blade Length	Leaf Blade Width	Leaf Lobules	Plant Breadth	Petal Number	Internode No.	Fruit Height	Fruit Length	Fruit Width	Fruit Weight
31	V1005895	93.0	Indeterminate	10.9	5.7	11.0	4.7	11.0	6.0	12.0	9.1	18.2	86.2
32	V1008099	93.2	Indeterminate	11.4	6.8	9.0	4.0	6.0	6.8	12.0	7.8	17.7	53.0
33	RVI006832	93.6	Indeterminate	10.4	6.2	10.0	4.8	6.0	6.0	12.0	6.2	11.9	25.8
34	V1030375	93.6	Indeterminate	10.4	6.2	10.0	4.8	6.0	6.0	12.0	11.2	21.1	133.2
35	V1006865	94.0	Indeterminate	9.3	5.1	9.0	3.8	5.0	6.4	10.0	7.3	15.2	41.9
36	V1005986	94.2	Indeterminate	16.3	6.4	8.0	4.0	6.0	8.2	7.0	10.2	21.2	127.7
37	V1005872	94.8	Indeterminate	13.8	7.8	10.0	4.2	7.0	10.6	6.0	10.0	19.9	111.6
38	V1005989	95.4	Indeterminate	10.9	6.4	9.0	4.0	5.0	6.8	12.0	7.4	15.0	46.7
39	V1008916	95.6	Indeterminate	10.6	6.4	12.0	4.0	5.0	6.8	8.0	6.0	12.9	34.4
40	V1006892	95.6	Indeterminate	10.6	6.4	12.0	4.0	5.0	6.8	8.0	4.6	8.3	10.0
41	LO5942	96.2	Indeterminate	14.2	6.9	8.0	3.7	5.0	5.2	7.0	6.7	12.7	38.9
42	RV101888	96.4	Indeterminate	9.6	6.5	10.0	5.3	5.0	7.2	8.0	8.9	15.4	69.1
43	V1006838	96.4	Indeterminate	11.9	6.8	10.2	4.8	6.0	7.9	9.0	5.3	12.0	21.2
44	V1005987	96.8	Indeterminate	9.8	6.0	10.0	4.1	7.0	7.6	10.0	10.5	21.1	107.4
45	V1006840	96.8	Indeterminate	9.8	6.0	10.0	4.1	7.0	7.6	10.0	8.8	20.2	109.7
46	RV02114	100.6	Indeterminate	10.6	5.7	11.0	3.5	6.0	7.4	9.0	5.0	10.2	16.5
47	V1006841	100.6	Indeterminate	10.6	5.7	11	3.5	6	7.4	9	5.9	11.5	23.4
48	V1009873	106.2	Indeterminate	11.0	5.0	12.0	3.6	7.0	8.5	8.0	8.2	16.5	64.1
49	V100852	106.2	Indeterminate	11.0	5.0	12.0	3.6	7.0	8.5	8.0	10.1	20.2	125.2
50	V1005875	106.4	Indeterminate	12.3	6.5	8.0	3.6	8.0	7.8	10.0	8.1	18.3	73.8
51	V1007540	106.4	Indeterminate	12.3	6.5	8.0	3.6	8.0	7.8	10.0	11.0	21.6	147.6
52	RVIOI884	106.6	Indeterminate	10.5	6.6	11.0	4.8	5.0	9.0	14.0	5.4	10.0	17.4
53	V1006842	106.6	Indeterminate	10.5	6.6	11.0	4.8	5.0	9.0	14.0	9.1	19.7	101.8
54	V1006847	108.4	Indeterminate	10.3	2.1	11.0	5.0	6.0	4.9	3.0	6.4	13.0	34.6
55	V1006972	108.4	Indeterminate	10.3	2.1	11.0	5.0	6.0	4.9	3.0	5.3	9.5	26.5
56	V1006828	109.0	Indeterminate	10.6	6.5	9.0	4.1	6.0	7.0	7.0	5.7	10.5	22.0
57	V1006864	114.0	Indeterminate	11.0	4.8	9.0	3.9	6.0	6.6	8.0	5.5	10.7	19.8
58	V1007539	114.0	Indeterminate	11.0	4.8	9.0	3.9	6.0	6.6	8.0	7.8	15.1	52.2
59	V1006825	120.2	Indeterminate	9.0	4.5	8.0	3.4	10.0	7.2	11.0	8.0	19.9	85.8
60	V1007108	120.2	Indeterminate	9.0	4.5	8.0	3.4	10.0	7.2	11.0	8.8	12.3	44.0

Key: Quantitative descriptors used in morphological characterization of African Tomato. Plant height ranged from 40 cm to 120.2cm. **Leaf blade length** (6.3- 16.3cm), Leaf blade width (2.1-8.7cm), leaf lobules (7-13), Plant breadth (2.9-5.3cm), petal number (4.0-11.0), internode no. (3.9-11.0), fruit length (3.0- 14cm), Fruit width (6.3-25.9cm) fruit weight (4.2- 166.9g)

4.3.1.3 Variation in quantitative traits African tomato

The means of the quantitative traits, their minimum and maximum values were significant at P≤0.05 as shown in the Table 4:5. Plant height of African tomato ranged from 40 cm to 151 cm, Fruit height ranged from 4cm to 12cm, the stem circumference ranged from 2.5cm to 10cm, the fruit length ranged from 2.5 cm to 13 cm, while the fruit diameter range was 5.4 cm to 26 cm. Combined analysis of variance revealed significant difference (P ≤ 0.05) among accessions for all the experimental characters. Mean data showed high range for all the studied traits. Coefficient of variation (CV %) varied from 2.2 % for number of leaf lobule to 14% for fruit height (Table 4:5).

Table 4:5. Variation of the African tomato quantitative traits by variance percentage, least significance difference.

Trait	Unit	min	Max	mean	Mean square	LSD (0.05)	F	CV%
FH	cm	4.00	12.00	6.67±1.24	37.99	0.18	0.23	14.0
FL	cm	2.50	13.40	7.94 ± 0.25	27.40	0.40	0.51	3.2
FW	cm	5.40	26.00	15.72±0.70	88.97	1.13	2.02	4.5
IH	cm	3.00	12.00	6.67±0.37	11.10	0.60	0.84	5.6
PB	cm	2.50	10.00	4.28±0.50	1.75	0.81	0.22	11.7
PH	cm	40.00	151.00	92.7±2.70	3403.72	0.24	12.94	2.9
LBL	cm	6.20	17.50	11.71±0.78	134.43	1.26	2.26	6.7
LBW	cm	3.20	9.50	6.14±0.59	72.37	0.96	2.87	9.7
LL	No.	7.00	13.00	9.88±0.21	13.13	0.34	0.15	2.2
PN	No	5.00	11.00	6.04±0.15	12.24	0.24	0.16	2.5
Y	kg	4.30	180.4	67.33±2.12	2356.12	3.41	0.36	3.2

(P≤0.05 significance level).

Key: FH- fruit height, FL- Fruit height , FW- fruit width, IH- Internode height , PB- Plant bredth, PH- Plant height, LBL- Leaf blade length, LBW- Leaf blade width, LL- Leaf lobules, PN- Petal numbers ,Y-Yield.Min- minimum, Max- maximum

4.3.1.4 Principal component analysis of the quantitative traits.

The first seven principal components (PC1, PC 2, PC3, PC4, PC5, PC6, and PC7 analyzed covered 68.51% variation within the 14 dimensions generated (Table 4:6). The quantitative traits that contributes more to PC 1 were fruit length, width internode height and leaf length PC1 accounted for 20.89 % of the total variation. The PC 2 accounted for 11.63 % of the total variation due to fruit length and fruit mass PC 3, PC 4, PC 5 and PC 6 accounted for 9.54, 7.92, 6.75, 6.16 % of total variation contributed by petal no., plant height, petal no. and plant width respectively PC 7 accounted for 5.52 % contributed by leaf length, leaf width, plant height and plant width (Table 4:6).

Table 4:6. Principal component analysis (PCA) of the quantitative traits for African tomato

Quantitative parameters	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Flower heads	0.07258	-0.40814	0.09085	0.11223	0.0303	0.04267	0.14039
Fruit length	**0.41411**	**0.23125**	0.11139	-0.10733	0.05275	0.02169	0.03877
Fruit mass	-0.38372	**0.24584**	0.19038	-0.18398	0.08386	-0.06275	0.04988
Fruit texture	-0.29795	0.12817	-0.03516	0.05423	0.19827	0.11512	0.07797
Fruit width	**0.41417**	-0.21053	0.17539	-0.14587	0.06099	0.02006	-0.02095
Internode height	**0.21518**	-0.03836	-0.0615	0.05328	-0.0867	-0.17874	-0.59779
Leaf length	**0.71981**	0.11673	-0.44241	0.15117	**0.18627**	-0.27005	**0.21874**
Leaf width	0.12114	0.12054	-0.45903	0.14075	-0.0819	-0.39829	**0.23909**
Plant height	0.10945	-0.40855	0.0698	**0.20491**	-0.0913	-0.01389	**0.22363**
Plant width	0.18552	-0.24753	0.04562	0.08003	-0.2033	**0.33455**	**0.22363**
% Variation	**20.89**	**11.63**	**9.54**	**7.92**	**6.75**	**6.16**	**5.52**
Cumulative % variation	**20.89**	**32.62**	**42.16**	**50.08**	**56.83**	**62.99**	**68.51**

Principal component showing traits contributing to variation in African Tomato

Key: **PC**- Principle component, **PC1 - PC7** gave significant variation in African tomato. **PC1** – Main variation contributors were Fruit length, fruit width and leaf length. **PC2** – fruit length and fruit mass. **PC3** – fruit mass, fruit length and fruit width. **PC4**- Plant height was the main variant contributor. **PC5**- Fruit Texture and leaf length. **PC6**- Internode height and plant width. **PC7**- Plant width, plant height, leaf width and leaf length

4.3.1.4 Simple matrix correlation of the phenotypic traits in African Tomato

Significant positive correlation (r= 0. 446, 0,264, 0.271, 0.3539, 0.65, and 0.93, <0.01) was observed between fruit yield and leaf width, petal number, internode height fruit shape fruit length and fruit width respectively. Significant positive correlation (r= 0.72, < 0.01) was observed between leaf width and leaf blade length. Yield per plant showed significant positive correlation with fruit weight (r= 0.65, <0.01) (Table 4:7).

Low positive correlation was observed between stem colour (r= 0.055, 0.259,0.104, 0.192, 0.106, 0.109, < 0.01) and leaf lobule, leaf vein colour, petiole colour, fruit colour, shoulder colour and fruit texture respectively. Low positive correlations were also observed between leaf blade length (r= 0.224, 0.224,0.258,0.252, 0.175, < 0.01) and leaf lobule, internode height, fruit length,fruit width and fruit yield respectively. Other parameters which showed low positive correlations were leaf width (r= 0.199, 0.149, 0.105, 0.182, <0.01) and leaf lobule, plant width, internode height, fruit lenth respectively). Low positive correlation was also observed between leaf lobule (r= 0.112 and 0.219, < 0.01) and petiole colour and fruit texture respectively. Leaf blade colour (r= 0.22, 0.16, 0.14, 0.1 <0.01) and leaf blade colour shoulder colour and fruit texture respectively. Leaf blade colour had low positive correlation (r= 0.25, 0.17,0.2, 0.1,<0.01) with internode height, fruit length, fruit width and fruit yield respectively. Petiole colour had low positive correlation (r= 0.215, 0.714, 0.153, 0.107, 0.159, <0.01) with petal number, internode height, shoulder colour and fruit texture respectively. Plant height had low positive correlation with (r= 0.187, 0.117, 0.112, <0.01) with petal number,internode height and fruit width respectively. Plant width had low positive correlation (r= 0.269, 0.1445, 0.187, 0.27, 0.22, 0.23, 0.229,

< 0.01) with petal number, internode height, flower heads, fruit shape, fruit length, fruit width and fruit yield.

Petal number had low positive correlation (r= 0.0158, 0.161, 0.157, 0.104, 0.264, < 0.01) with internode height, fruit shape, fruit length and fruit yield respectively. Internode height had low positive correlation of (r= 0.175, 0.271,< 0.01) with fruit length and fruit yield respectively). Shoulder colour had a low positive correlation (r= 0.21, < 0.01) with fruit texture. Fruit shape had a low positive correlation (r= 0.193, <0.01) with fruit length.

High negative correlation was observed between stem colour (r= -0.27, -0.34, -0.31,-0.28, < 0.01) with leaf blade length, internode height, fruit width and fruit yield respectively. Leaf blade length had a high negative correlation (r= -0.23, -0.2, -0.26, < 0.01) with petiole colour, flower heads and shoulder colour respectively.

Leaf width had a high negative correlation (r = -0.23, -0.24,<0.01) with fruit colour and shoulder colour respectively. Leaf lobules had a high negative correlation (r = -0.27, -0.32, -0.2,<0.01) with plant width, flower heads and fruit width respectively. Leaf vein colour had a high negative correlation (r= -0.22, -0.2, < 0.01) with leaf blade colour and internode height respectively. Plant height had a high negative correlation (-0.241, -0.217 <0.01) with fruit colour and fruit texture respectively. Plant width had a high negative correlation (-0.299, -0.269, -0.293 <0.01) with fruit colour, fruit shape and fruit texture respectively. Internode height had a high negative correlation (-0.29, and -0.29 < 0.01) for fruit colour and fruit texture. The number of flower heads had a high negative correlation (-0.21, -0.2, <0.01) with fruit colour and fruit texture respectively (4:7).

Fruit colour also had a high negative correlation (-0.39, -0.32,<0.01) with fruit length nad fruit width respectively. Fruit shoulder colour had a high negative correlation (-0.3, -0.22, .21,<0.01) with fruit length, fruit width and fruit texture respectively. The fruit shape, fruit length and fruit width had a high negative correlation (-0.2724,-0.23, -0.39 <0.01) with fruit texture respectively (4:7).

Table 4:7. Simple correlation matrix of the phenotypic traits in African Tomato

	SC	LBL	LW	LL	LVC	LBC	PC	PH	PW	PN	IH	FH	FC	SHC	FSH	FL	FW	FT	FY
SC	1																		
LBL	-0.27	1																	
LW	-0.1	**0.72***	1																
LL	0.055	0.224	0.199	1															
LVC	0.259	-0.1	0.079	-0.05	1														
LBC	-0.11	0.11	0.075	0.065	0.22	1													
PC	0.104	-0.23	-0.16	0.112	**0.39***	**0.48***	1												
PH	-0.03	-0.14	-0.14	-0.2	-0	-0.05	-0.14	1											
PW	-0.09	0.01	0.147	-0.27	-0	-0.06	-0.14	**0.446***	1										
PN	-0.15	-0.08	-0.16	-0.11	-0.1	0.02	0.215	0.187	0.2619	1									
IH	-0.34	0.226	0.105	0.005	-0.2	0.25	0.174	0.117	0.1445	0.158	1								
FH	-0.01	-0.2	-0.16	-0.32	0.08	-0.25	-0.1	**0.377***	0.1867	0.161	-0.12	1							
FC	0.192	-0.18	-0.23	0.055	0.16	-0.06	0.056	-0.241	-0.299	-0.13	-0.29	-0.21	1						
SHC	0.106	-0.26	-0.24	-0.01	0.14	-0.07	0.153	-0.172	-0.169	-0.02	-0.1	-0.16	**0.761***	1					
FSH	-0.02	0.061	-0.06	-0.05	-0.1	0.03	0.107	-0.064	0.2694	0.157	0.351	-0.04	-0.08	-0.11	1				
FL	-0.19	0.258	0.182	-0.04	0.07	0.17	0.044	-0.02	0.2209	0.104	0.175	-0.01	-0.39	-0.3	0.1928	1			
FW	-0.31	0.252	0.091	-0.2	-0.1	0.2	-0.07	0.112	0.2302	**0.342***	**0.361***	0.05	-0.32	-0.22	**0.3533***	**0.64***	1		
FT	0.109	-0.06	-0.06	0.219	0.1	0.04	0.159	-0.217	-0.293	-0.16	-0.29	-0.2	**0.352***	0.21	-0.2724	-0.23	**-0.39**	1	
FY	-0.28	0.175	**0.446***	-0.17	-0.1	0.1	-0.09	-0.003	0.229	0.264	0.271	0.044	-0.23	-0.04	**0.3539***	**0.65***	**0.93***	**-0.37**	1

Key: Key: Bolded signify positive and negative correlation in African tomato morphological traits. Highest positive correlation was observed between fruit yield and fruit width (r= 0.93, <0.01) highest negative correlation (r=-0.39 <0.01) between fruit texture nd fruit width. **SC**- Stem colour, **LBL**- leaf blade colour, **LW**- leaf width, **LL**- leaf lobule, **LVC**- leaf vein colour, **LBC**- leaf blade colour, **PC**- petiole colour, **PH**- Plant height, **PW**- plant width, **PN**- petal number, **IH**- Internode height, **FH**- flower heads, **FC**- fruit colour, **SHC**- shoulder colour, **FSH**- fruit shape, **FL**- fruit length, **FW**- fruit width, **FT-** fruit texture, **FY**-fruit yield

4.3.1.5 Biplot presentation of the African tomato

The biplot shows that morphologically, the African tomato accessions are diverse with some overlapping of the accessions. Accessions V1006825, V1006841 and RV102112 clustered far from other accessions showing high divergence while Tindi 050589 and Tindi 050590 are closely clustered and far from other accessions (Figure 4:1).

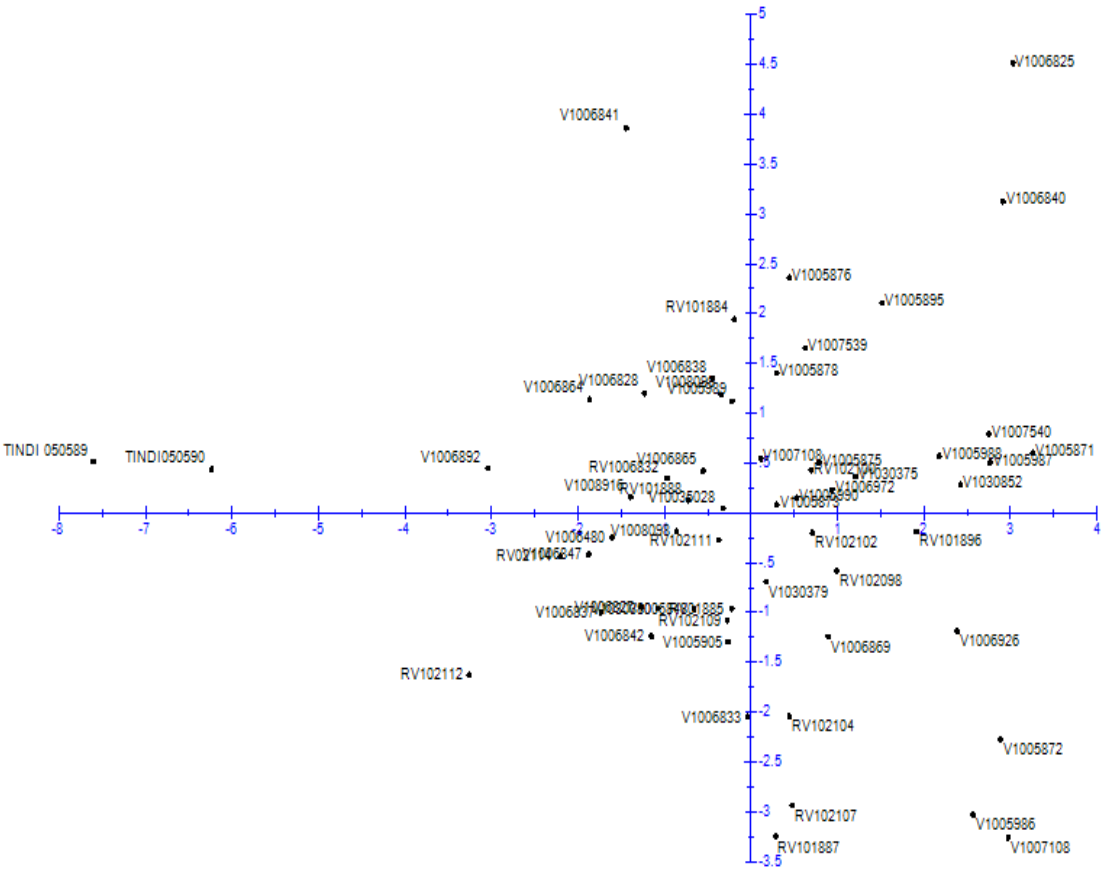


Figure 4:1. Biplot presentation of the African tomato accessions
Key: **A, B, C and D** represent four quadrants of a biplot. **A**- Some accessions grouped together but far from others, accessions in this quadrant had small sized fruits, low leaf length, and low fruit yield. **B**- Represents accessions with indeterminate plant growth habit, **C**- all accessions in this Quadrant had asymmetrical leaf base and **D**- Had medium and big fruits

4.3.1.6 Cluster analysis of 60 accessions of African tomato.

The 60 accessions grouped into three major clusters with many sub clusters using morphological data (Figure 4:2). Cluster 1 had 26 accessions, cluster 2 had 31 accessions while cluster 3 had only 3 accessions.

Main contributing factor in cluster 1 was the leaf vein colour; with all the 26 accessions having green colour. All accessions in cluster 2 had red fruits while cluster 3 had same vein colour (light green), same leaf blade colour (green) same fruit colour (red) and the same fruit texture (ridged fruit surface).

Accessions with similar phenotypes i.e. Tindi 050589 and Tindi 050590 which were of cherry type grouped closely using morphological data. This had same fruit size (cherry), same fruit colour (yellow), same fruit shape (round) same fruit surface texture (smooth) and were both indeterminate examples of other accessions that grouped together were V1006480 and V1005905, V1005989 and V1006865, V1008916 and V1006848, V1007108 and V1007539, V1006864 and V10035028.

Other phenotypes that dictated the clustering together were colour of the fruit, shape, and texture, leaf blade colour. The results of the clustering analysis using the Darwin's 6 software showed that branching occurred at a very low phenon line.

Clustering of the African tomato accessions according to morphological traits

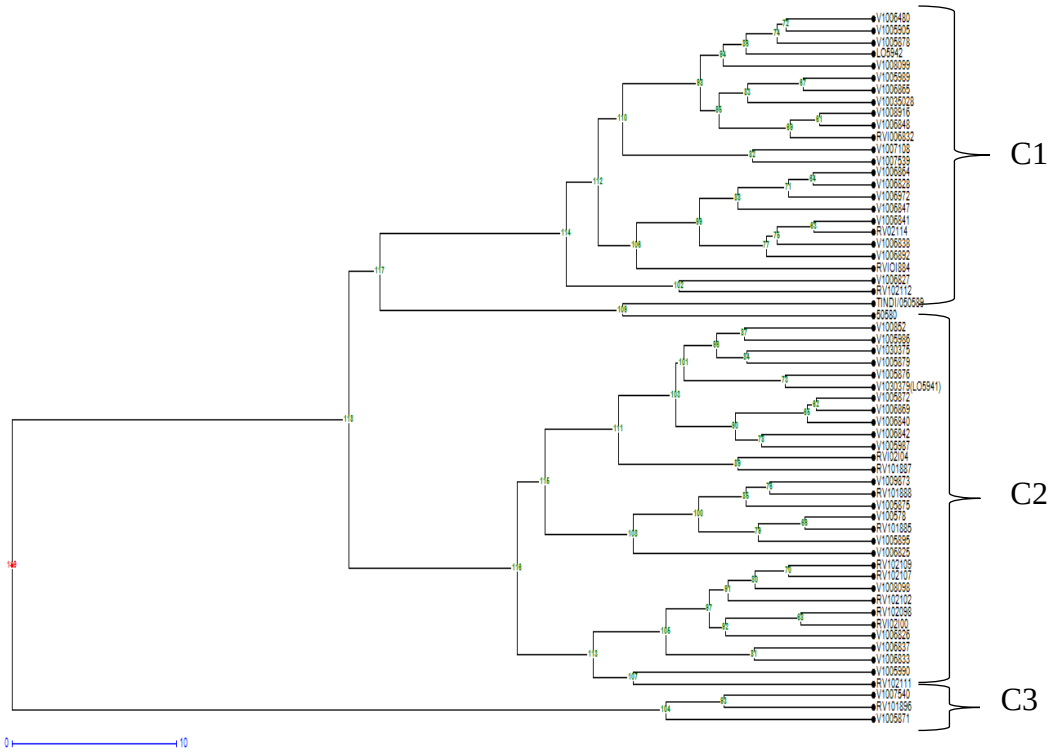


Figure 4.2. Clustering of 60 African tomato using the Darwin's 6 software. Cluster 1 - accessions with same leaf vein colour**. Cluster 2-** had red fruit**. Cluster 3-** had light green colour, green leaf blade colour, red fruits and same fruit texture

4.3.2 Morphological characterization of selected African eggplant accessions.

Significant differences were observed in the various morphological characteristics evaluated at the vegetative stage and reproductive stages (Table 4:8). A range of variation was observed at reproductive stage (fruit colour, fruit shape, fruit size and texture) of the African eggplant.

4.3.2.1 Characterization using Qualitative traits

The African eggplant exhibited significant variation when the qualitative traits were considered.

Plant growth habit (PGH): The sixty - seven accessions clustered into five groups i.e. very tall, tall, intermediate, prostrate, and very prostrate accessions according to their height (Table 4:8).

Number of prickles present (PN): Sixty-four percent of all accessions did not have prickles on the leaf surface, while 3% (RV100185 and RV100247) had more than 20 prickles, others ranged from 16.4, 1.5,7.5 and 6% for 4, 8 and 15 prickles on the leaf surface respectively (Table 4:8).

Leaf lobules (LLL): Leaf blade length ranged between 7.7cm (RV100380) - 32.5cm (RV100364 and RV100383) while leaf blade width ranged 5.5cm (RV100328) to 21cm (RV100447). The number of leaf lobules was between 5 to 16 (Table 4:8).

Stem colour (SC): Stem colour formed six groups; light green (10), green (23), dark green (1), green with purple stripes (25) purple (2) and dark purple (6) (Table 4:8).

Leaf base shape (LB): Only three groups were formed when the shape of the leaf base was considered, with one accession (RV100199) showing the sessile behavior, 40 had asymmetrical and 26 had the symmetrical leaf base shape (Table 4:8).

Fruit colour (FC): Differences in fruit shape and colour was observed as the most important trait in this study (Plate 4:3 and Table 4:8). The fruit shape ranged from round, oval, and oblong shapes, the diverse colour included the red (Plate 4:3 a, b and c), yellow (Plate 4:3d), orange, purple (Plate 4:3h), yellow with green stripes (Plate 4:3g) and white (Plate 4:3f).

Fruit shape (FS) and Texture: Round fruit shape dominated the fruit shape category with 57 accessions (Plate 4:3a, c and f); oval had 7 accessions (Plate 4:3b,

4:3d, 4:3g and 4:3h) while oblong shape had 3 accessions (Plate 4:3e and h). Only two types of fruit surface texture were observed; ridged surface (51 accessions - 3a) and smooth surface (16 accessions) (Plate 4:3b, 4:3c, 4:3d,4:3e, 4:3f, 4:3g, 4:3h) and Table 4:8).

Plate 4:3. African eggplant Fruit morphology.

Variation in fruit shape, colour, size and texture; a) round ridged red, b) oval smooth red, c) round small smooth red, d) oblong ridged yellow, e) oblong smooth green, f) round smooth white, g) oval smooth orange with stripes, h) oblong smooth purple fruit.

Table 4:8. Qualitative morphological variations in African eggplant Accessions.

	Accessions	PG	LVC	LBC	SC	LBS	FC	FS	FT	SN	Accession	PG	LVC	LBC	SC	LBS	FC	FS	FT
1	RV100383	VT	L. G	L. G	L. G	Sym	Y+G	O	R	**34**	RV1002100	Int	D. G	G	P	Asym	R	R	R
2	RV100511	VT	L. G	G	L. G	Asym	O	O	S	**35**	RV100333	Int	L. G	G	L. G	Asym	R	O	S
3	RV100431	VT	P	D.G	D.P	Asym	R	R	R	**36**	RV100266	Int	L. G	G	L. G	Asym	R	R	R
4	RV100364	VT	P	D.G	D.P	Sym	R	R	R	**37**	RV100327	Int	G	G	G	Asym	R	R	R
5	RV100352	VT	L. G	L. G	G	Sym	R	R	R	**38**	RV100264	Int	P+G	D.G	P+G	Sym	R	O	R
6	RV100260	VT	G	D.G	P+G	Asym	R	R	R	**39**	RV100243	Int	L. G	G	G	Sym	R	R	R
7	RV100169	VT	G	G	P+G	Asym	Y	O	R	**40**	RV100250	Int	G	G	G	Asym	R	R	R
8	RV100359	VT	L. G	G	L. G	Asym	R	O	R	**41**	RV100455	Int	G	G	G	Asym	P	R	S
9	RV100458	VT	P+G	D.G	P+G	Sym	R	O	R	**42**	RV100261	Int	G	G	G	Asym	P	Ob	S
10	RV100190	VT	G	D.G	G	Sym	R	R	S	**43**	RV100453	Int	P+G	G	P+G	Asym	R	R	S
11	RV100335	T	L. G	G	G	Sym	R	R	R	**44**	RV100241	Int	L. G	G	Gr	Asym	R	R	R
12	RV100249	T	L. G	G	G	Sym	R	R	R	**45**	RV100242	Int	P+G	G	P+G	Asym	R	R	S
13	RV100194	T	D. P	D.G	D.P	Sym	R	R	R	**46**	RV100328	Int	P+G	D.G	P+G	Sym	G	R	R
14	RV100356	T	G	G	G	Asym	R	R	S	**47**	RV100342	Int	L. G	D.G	P+G	Asym	R	R	R
15	RV100382	T	L. G	G	G	Asym	R	R	R	**48**	RV100386	P	L. G	G	G	Asym	R	R	R
16	RV100332	T	G	D.G	P+G	Asym	R	R	R	**49**	RV100380	P	G	G	P+G	Sym	Y	O	R
17	RV100239	T	G	G	G	Asym	R	R	R	**50**	RV100240	P	L.G	D.G	G	Asym	R	R	R
18	RV100245	T	L. G	G	P+G	Sym	R	R	R	**51**	RV100325	P	LG	G	G	Asym	R	R	R
19	RV100343	T	G	D.G	G	Sym	R	R	S	**52**	RV100445	P	G	G	G	Asym	G	Ob	S
20	RV100201	T	L. G	G	L. G	Asym	G	Ob	S	**53**	RV100330	P	P	D.G	D.P	Sym	G	R	R
21	RV100447	T	P	D.G	P	Sym	R	R	R	**54**	RV100271	P	L. G	G	D.G	Sym	R	R	R
22	RV100161	T	G	G	L. G	Asym	R	R	R	**55**	RV100234	P	G	G	P+G	Asym	Red	R	R
23	RV100456	T	G	D.G	P+G	Asym	R	R	R	**56**	RV100273	P	G	G	P+G	Sym	Red	R	R
24	RV100377	T	P+G	D.G	P+G	Asym	R	R	R	**57**	RV100265	P	G	G	G	Asym	Red	R	R
25	RV100247	T	G	G	P+G	Sym	R	R	R	**58**	RV100215	P	G	D.G	G	Asym	G	R	R
26	RV100343-C	T	G	D.G	G	Sym	R	R	S	**59**	RV100334	P	G	D.G	G	Asym	O	R	R
27	RV100217	T	G	G	L. G	Asym	R	R	S	**60**	RV100270	P	G	G	G	Asym	R	R	R
28	RV100268	Int	D. P	DP	D.P	Asym	R	R	R	**61**	RV100 199	V. P	P+G	G	P+G	Sess	W	R	S
29	RV100218	Int	L. G	G	L. G	Asym	R	R	S	**62**	RV100274	V. P	G	D.G	P+G	Asym	R	R	R
30	RV100346	Int	L. G	G	P+G	Asym	R	R	R	**63**	RV100360	V. P	L.G	G	L. G	Sym	O	R	S
31	RV100259	Int	G	D.G	P+G	Sym	O	R	R	**64**	RV100432	V. P	G	G	P+G	Sym	Red	R	R
32	RV100248	Int	D.G	G	P+G	Sym	R	R	R	**65**	RV100236	V. P	G	G	P+G	Sym	Y	R	R
33	RV100452	Int	G	D.G	G	Asy	G	R	R										

Key : GH (Growth Habit) :- VT- very tall, T- Tall, Int- intermediade, P- Prostate, VP- very prostate. LVC (Leaf Vein Colour):- G- Green, D.P- Dark Purple,L.G- Light Green, D.G- Dark Green, P+G- purple and green stripes. LBC (Leaf Blade Colour)- D.G- Dark Green, G- green, DP- dark purple, L.G- Light Green. SC (Stem Colour) G- Green, L.G- Light green, P+G- Purple and green stripes, P- purple, D.P- Dark purple, LBS (Leaf base shape)- Sym- Symmetrical , Asym- Asymetrical , Sess- Sessile., FC (Fruit Colour)- R- red , O- orange, G- Green, Y- Yellow, Y+G- yellow and green stripes, FS (Fruit shape)- R- round, O- Oval, OB- Oblong, FT (Fruit texture) S- Smooth, R- ridge

4.3.2.2 Phenotypic diversity of the African eggplants index using qualitative traits

Phenotypic diversity for individual qualitative traits revealed a high degree of variation among the studied populations (Table 4:9) using the Shannon - Weaver diversity index to estimate phenotypic diversity of nine qualitative traits studied.

The highest phenotypic diversity index (H') for traits studied recorded was 1.0 in plant growth habit, petiole colour, stem colour and vein colour with a total mean phenotypic diversity index of 6.0. Substantial variation was observed in fruit colour, fruit shape, fruit texture, leaf base and leaf lade colour (Table 4:9).

Table 4:9. Diversity index (H´) values explaining the genetic diversity of the accessions based on qualitative traits.

Qualitative traits	Genetic index(H')
Plant growth habit	1.000
Fruit colour	0.951
Fruit shape	0.988
Fruit texture	0.992
Leaf base shape	0.987
Leaf blade colour	0.995
Petiole colour	1.000
Stem colour	1.000
Vein colour	1.000
Total	**6.040**

Key: Genetic index (H') in African eggplant ranged from 0.951-1.000. The higher the genetic index (H') reflects high variability in terms of qualitative traits.

The characterization data and the multivariate analysis performed may be useful to select a subset of accessions that represent most of the morphological diversity of the African eggplants.

4.3.2.3 Characterization using Quantitative traits

Plant growth habit (PGH): The sixty - seven accessions clustered into five groups i.e. very tall (10), tall (17), intermediate (21), prostrate (13) and very prostrate (5) accessions according to their height (Table 4:10). Accession RV100511 was the tallest (104 cm) while RV100360 (23cm) was the shortest among the 67 African eggplants accessions.

Number of prickles present (PN): Sixty-four percent of all accessions did not have prickles on the leaf surface, while 3% (RV100185 and RV100247) had more than 20 prickles, others ranged from 16.4, 1.5,7.5 and 6% for 4, 8 and 15 prickles on the leaf surface respectively (Table 4:10).

Leaf blade length ranged between 7.7cm (RV100380) - 32.5cm (RV100364 and RV100383) while leaf blade width ranged 5.5cm (RV100328) to 21cm (RV100447).

Plant breadth (PB): Plant breath ranged between 3.5 to 6.3 cm while internode length ranged between 2.9 to 7.7 cm (Table 4:10).

Fruit size: Fruits varied in length having between 2 to 12 cm and the fruit width having between 4 to 22.6 cm (Table 4:10).

Table 4:10. Quantitative morphological variations in African eggplant Accessions.

	Accessions	PH	LBL	LBW	PN	FL	FW	S/No	Accessions	PH	LBH	LBW	PN	FL	FW
1	RV100383	101	31.0	26.0	0	6.6	13.8	**34**	RV1002100	78	16.8	11.0	0	7.3	17.3
2	RV100511,	104	24.7	20.4	0	5.0	7.3	**35**	RV100333	75	19.2	17.7	<3	5.4	13.0
3	RV100431	102	24.0	18.0	0	5.0	13.0	**36**	RV100266	78	14.0	10.5	0	5.4	13.0
4	RV100364	102	31.0	21.0	0	6.0	17.0	**37**	RV100327	76	13.3	10.7	0	4.8	13.0
5	RV100352	100	17.2	16.2	0	6.0	15.0	**38**	RV100264	77	18.0	18.0	0	6.7	16.4
6	RV100260	101	18.5	15.4	< 3	6.3	17.6	**39**	RV100243	72	13.5	10.0	<3	5.4	15.0
7	RV100169	100	21.3	19.5	0	4.6	8.6	**40**	RV100250	70	10.3	8.0	<3	5.0	12.8
8	RV100359	102	23.7	19.7	0	12.6	15.5	**41**	RV100455	75	24.0	21.2	0	5.0	13.0
9	RV100458	103	12.0	8.5	0	5.2	10.0	**42**	RV100261	67	18.6	15.4	<3	6.8	7.5
10	RV100190	101	13.0	9.9	0	2.0	4.0	**43**	RV100453	87	13.7	11.0	0	2.2	6.3
11	RV100335	94	16.0	12.2	0	10.2	22.5	**44**	RV100241	66	18.0	12.0	8	8.5	16.0
12	RV100249	94	22.2	17.5	0	3.5	8.0	**45**	RV100242	65	20.0	15.0	0	1.0	5.0
13	RV100194	93	22.3	15.0	15	4.0	12.0	**46**	RV100328	76	10.0	5.5	0	9.0	21.0
14	RV100356	95	22.5	18.7	0	2.0	5.0	**47**	RV100342	72	12.5	10.5	15	6.4	15.6
15	RV100382	94	12.9	10.1	0	.0	10.0	**48**	RV100386	45	14.8	11.1	0	9.8	21.5
16	RV100332	95	23.0	15.0	15	5.0	11.0	**49**	RV100380	42	7.7	5.6	0	5.3	9.8
17	RV100239	96	11.2	10.0	0	5.0	12.0	**50**	RV100240	44	23.5	17.0	0	5.5	13.3
18	RV100245	96	22.0	17.0	< 3	4.0	8.0	**51**	RV100325	43	12.8	19.0	0	7.0	17.0
19	RV100343	94	21.1	20.0	0	2.5	5.8	**52**	RV100445	45	16.0	13.0	0	5.0	13.0
20	RV100201	95	21.2	15.0	8	11.0	11.5	**53**	RV100330	45	23.0	17.0	<3	4.8	12.3
21	RV100447	99	29.0	22.0	8	4.5	11.0	**54**	RV100271	45	17.2	16.2	0	6.0	15.0
22	RV100161	94	25.9	12.5	0	5.0	13.3	**55**	RV100234	42	23.0	15.0	0	6.0	15.0
23	RV100456	98	23.0	17.8	4	6.5	17.4	**56**	RV100273	43	22.0	15.4	0	6.1	16.0
24	RV100377	96	28.1	19.0	0	5.7	12.4	**57**	RV100265	44	13.4	11.5	0	5.0	13.0
25	RV100247	92	21.0	19.0	>20	4.0	10.5	**58**	RV100215	44	20.0	17.0	0	5.5	13.0
26	RV100343-C	90	21.1	20.0	0	2.5	5.8	**59**	RV100334	40	20.8	13.4	0	7.0	15.0
27	RV100217	93	21.0	15.0	< 3	3.2	6.0	**60**	RV100270	45	17.3	15.0	4	5.5	13.0
28	RV100268	75	25.4	19.5	0	5.6	12.4	**61**	RV100 199	30	29.0	19.0	0	12.0	15.5
29	RV100218	77	16.8	11.0	0	2.0	7.0	**2**	RV100274	39	20.9	17.7	0	5.0	12.0
30	RV100346	75	14.0	10.1	0	2.2	5.0	**63**	RV100360	23	23.7	19.5	0	12.7	15.6
31	RV100259	75	20.4	17.5	0	7.0	16.0	**64**	RV100432	30	15.0	14.0	0	4.0	9.6
32	RV100248	76	22.0	17.0	<3	8.0	14.0	**65**	RV100236	37	20.5	18.0	0	7.0	17.0
33	RV100452	77	16.7	15.7	<3	7.0	15.0								

Key: **PH**- plant height (cm), **LBL**- Leaf blade length (cm), **LBW**- leaf blade width (cm), **PN**- Petal number (no), **FL**- Fruit length (cm), **FW**- Fruit width (cm). **Plant height** ranged from 23 cm to 104cm. **Number of princkles** 64%- accessions without prickles on the leaf surface, 3% > 20 prickles, and 6% had 15 or less prickles. **Leaf blade length** (7.7-32.5cm), **Leaf blade width** (5.5-21cm), **Plant breadth** (3.5-6.3cm), **internode no**. (2.9-7.7cm), **fruit length** (2.0-12.0cm), **Fruit width** (4.0-22.6cm).

4.3.2.4 Morphological variation in African eggplant accessions using quantitative traits

The African eggplant accessions showed significant diversity in the quantitative traits that were selected for their characterization. The means of the quantitative traits, their minimum and maximum values were significant at P≤0.01. Plant height, leaf blade length, leaf blade width, and fruit width had a significant contribution to the variation among the selected African eggplant with a mean of 75.17±3.59, 19.44±1.50, 14.76±1.69 and 12.37±1.42 respectively (Table 4:11).

Table 4:11. Coefficient variance percentage, least significant difference, means and mean squares for nine quantitative traits (P≤0.01 significance level). In African eggplant

Trait	Units	Minimum	Maximum	Mean	Mean square	LSD (0.05)	F	CV%
Flower Heads	No	1.00	13.00	4.309±0.48	37.99	0.54	0.23	11.1
Fruit Length	Cm	1.50	13.00	5.56 ± 0.72	27.40	0.81	0.51	12.9
Fruit Width	Cm	4.00	23.00	12.37±1.42	88.97	1.60	2.02	11.5
Internode Height	Cm	1.00	9.30	4.63±0.92	11.10	1.03	0.84	8.5
Plant Breadth	Cm	3.00	7.00	4.90±0.46	1.75	0.52	0.22	9.5
Plant Height	Cm	22.00	119.00	75.17±3.59	3403.72	4.05	12.94	4.8
Leaf Blade Length	Cm	9.00	32.50	19.44±1.50	134.43	1.69	2.26	7.7
Leaf Blade Width	Cm	6.30	26.20	14.76±1.69	72.37	1.91	2.87	11.5
Leaf Lobules	No	5.00	16.00	8.75±0.38	13.13	0.44	0.15	4.4

Key: No- Numbers, Cm- centimeters. **Plant height** – 75.17±3.59, **Leaf blade length**- 19.44±1.15, **Leaf blade width**- 14.76±1.69, **Fruit width**-12.37±1.42

4.3.2.5 Principal component analysis of the quantitative traits of the African eggplant

The first eight principal components (PC1, PC 2, PC3, PC4, PC5, PC6, PC7 and PC 8 accounted for 70.6% variation for eight PC within the 14 dimensions generated (Table 4:12). The Eigenvectors decreased significantly from principal component 8 from 5.34 % to 5.06 % (Table 4:12). This suggests that after principal component 8 more principal components did not describe much variation. PC-1 accounted for 16.02 % mainly contributed by Plant growth habit, fruit colour, fruit width and length, whereas PC-2 accounted for 12.29% mainly contributed by fruit texture, leaf blade width and length and number flower heads. PC-3 was contributed mainly by fruit texture, number of flower heads, plant breath. PC-4 was contributed by plant growth habit, number of leaf lobules. PC-5 was dictated mainly by leaf blade width, number of prickles and leaf blade length.

PC-6 was contributed by number of flower heads, fruit length, and number of prickles PC-7 was accounted by fruit colour, internode height; whereas PC-8 was contributed mainly by fruit shape, number of prickles plant breadth (Table 4:12).

The first, second and third PCs with a cumulative of 41.03 % revealed the most variation among the populations, showing a high degree of correlation among the traits studied (Table 4:12).

Table 4:12. Principal component analysis showing the main contributors to variation in African eggplant

Traits	PC-1	PC-2	PC-3	PC-4	PC-5	PC-6	PC-7	PC-8
Number of Flower Heads	-0.15897	0.27744*	0.11538	-0.21196	0.02384	0.38472*	0.10237	-0.18399
Fruit Length	0.28841*	-0.20113	-0.39739	0.18032	-0.07939	0.33111*	-0.04052	0.08979
Fruit Width	0.18062*	-0.43483*	-0.35880	0.05968	0.01560	0.15866	0.02196	0.04421
Internode Height	-0.08371	0.12666	-0.22202	-0.20757	-0.51360*	-0.05897	0.20670*	0.07094
Leaf Blade Length	-0.21312	0.24640*	-0.46125*	-0.02428	0.16944	-0.16278	-0.17892	-0.02010
Leaf blade width	-0.15968	0.23779*	-0.49039*	0.00986	0.21654*	-0.10560	-0.08740	-0.02063
Number of leaf lobules	0.14892	0.14362	-0.12556	0.24815	-0.28188	-0.00530	-0.58571*	0.02506
Plant Breadth	0.03963	-0.20629	0.09199	0.16662	-0.22362	-0.61760*	0.04144	0.10840
Plant Height	-0.02030	-0.00653	-0.07088	-0.16291	-0.27069	-0.00309	-0.16444	-0.61130*
Eigenvalue	2.802	2.439	2.175	1.495	1.239	1.169	1.086	1.015
% Variation	14.75	12.83	11.45	7.87	6.52	6.15	5.75	5.34
%Cumulative Variation	14.75	27.58	39.03	46.9	53.42	59.57	65.32	70.66

Key: PC- Principle component. Out of 14 principle components 8 were significant. Main variation contributors included; PC1- plant growth habit, fruit colour, fruit width and fruit length. PC2-fruit texture, leaf blade width, leaf length, number of flower heads. PC3-Fruit texture, number of flower heads, plant breath. PC4- Plant growth habit, number of leaf lobules. PC5- Leaf blade width, number of prickles, leaf blade length. PC6- Number of flower heads, fruit length, number of prickles. PC7- Fruit colour, internode height. PC8- Fruit shape, number of prickles, plant breadth

4.3.2.6 Relationships among the African eggplant accessions using Factorial analysis.

Overlapping among the accessions revealed a close relationship between many of the accession studied. Similar to the dendogram output, closely related accessions overlapped when the factorial analysis was considered at (axes ½) example of overlapping accessions were RV100445 and RV100265, RV100215 and RV100271, RV100236 and RV100274, RV100273 and RV100234. Accessions RV100328, RV100194 and RV100346 grouped far from the rest showing high variation based on morphological traits (Figure 4:3).

In Figure 3, Accessions which plot in different parts of the biplot would be the most informative for distinguishing accessions with narrow and wide diversity. For

example, accessions RV100199, RV100611, RV100261, RV100446, RV100194, and RV100247 were plotted far from the other accessions in Figure 4:3.

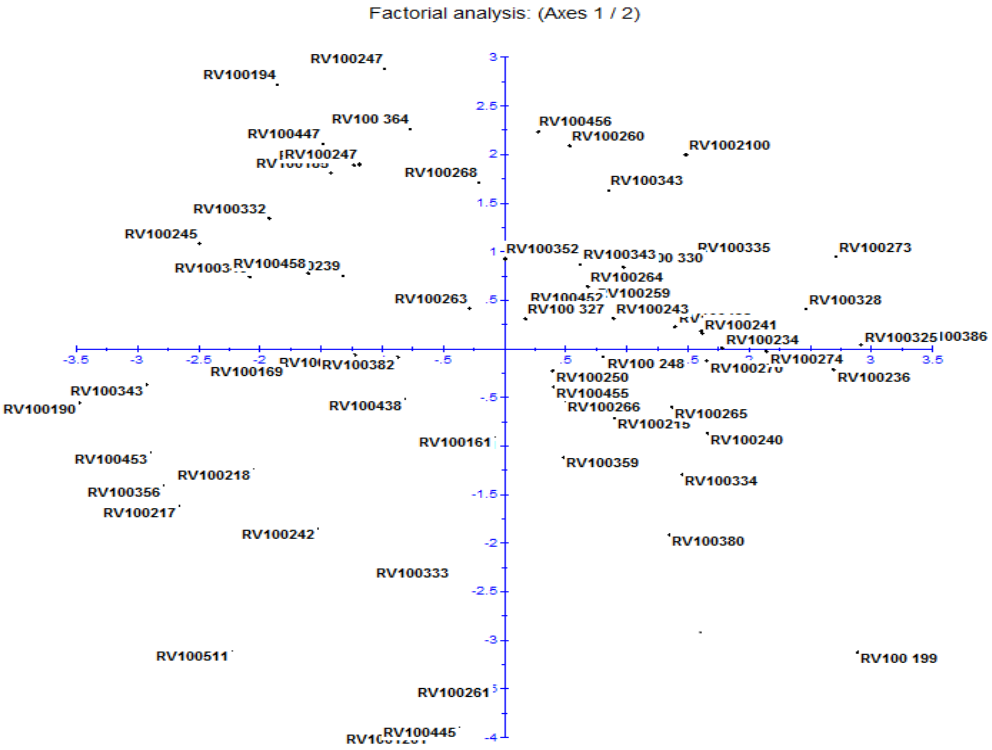


Figure 4.3. Interactions between the African eggplant accessions using the factorial analysis. **Quadrant A-** Most accessions were either tall or very tall (75cm to 102cm), **B-** Accessions without prickles, **C-** most accessions were round and red, the fruit texture was smooth. **D-** Most accessions were ridged

4.3.2.7 Cluster analysis of the African eggplants using Darwin 6 software

Clustering procedure using the Darwin 6 software grouped the 67 accessions into two main groups and many sub-clusters. Cluster 1 had 18 accessions, all with round fruit shape (RV100455, RV100199, RV100380, RV100456, RV100325, RV100432, RV100386, RV100265, RV100270, RV100215, RV100271, RV100334, RV100236, RV100274, RV100330, RV100240, RV100273 and RV100234).

Cluster 1 had a main sub cluster in which all eight accessions had a ridged fruit surface structure except accession RV100199 which had a white smooth fruit surface, (RV100334, RV100236, RV10027, RV100330, RV100240 RV100273 and RV100234) (Figure 4:4).

Accessions in cluster 1 were either prostrate or very prostrate. Most of the accessions clustered very closely showing a narrow diversity amongst the accessions, that is accessions RV100445 and RV100265 which had similar plant growth habit they did not have prickles, had same vein leaf stem colour, same leaf base shape same fruit height and width, RV100215 and RV100271 had same plant growth habit, no prickles had same fruit shape and texture, RV100236 and RV100274, shared same plant growth habit, leaf length and width, no prickles, same fruit shape and texture RV100273 and RV100234 with all parameters similar except slight difference in plant height, leaf width and height, fruit length and width therefore they grouped very closely (Figure 4:4).

Cluster 2 had the most accessions (49). The clustering was mainly dictated by the growth habit. This cluster had two main sub clusters namely, 2a and 2b. Cluster 2a constituted of 27 accessions with intermediate growth habit and they included (RV100190, RV100352, RV100458, RV100239 RV100382 RV100161, RV100438,

RV100364, RV100377, RV100431, RV100447, RV100335, RV1001201, RV100359, RV100456 RV100260, RV100343, RV100247, RV100194, RV100332, RV100511, RV100245, RV100169, RV100249, RV100217, RV100343-CN012, and RV100356) (Figure 4:4).

Sub cluster 2b comprised of 22 accessions all with tall and very tall growth habit and they included RV100328, RV100241, RV1002100, RV100250, RV100342, RV100327, RV100266, RV100243, RV100185, RV100268, RV100455, RV100259, RV100264, RV100452, RV100248, RV100333, RV100261, RV100242, RV100346, RV100218, RV100453 and RV100263) (Figure 4:4).

Clustering of the selected African eggplants was not dependent on the country of origin, but mainly on the plant growth habit, fruit colour and texture.

Clustering og the African eggplant

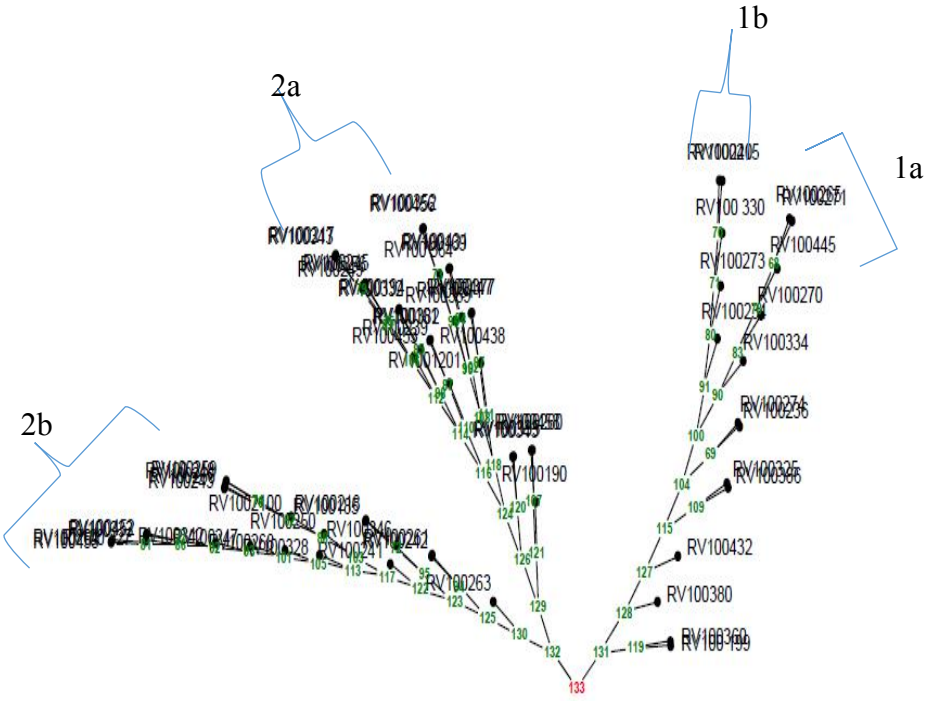


Figure 4.4. Cluster analysis of Sixty-Seven African eggplants accessions using Darwin's 6 software. African eggplant clustered into two main clusters with sub clusters. **Cluster 1a** – had round fruits , **Cluster 1b**- ridged fruit surface, prostrate and very prostrate, no prickles , same plant growth habits, same fruit shape and texture, same leaf width and leaf length. **Cluster 2** had 49 accessions. **Sub cluster 2a**- (27 accessions) - all with intermediate growth habit. **Sub cluster 2b** (22 accessions) - were tall and very tall growth habit

4.3.6 Comparison of the African tomato and eggplant using morphological traits

a) Shape.

There was significant similarity between the two crops in terms of shape. With some accessions looking very similar. Some of the shapes that was shared among the two crops included ;- round, oval, kidney shape (Table 4:13).

b) **Size**.

Both African tomato and eggplant exhibited different sizes with some showing the cherry types, while others were big sized. For instance in African tomato, Tindi 050580, Tindi 050589, RV102111, African eggplant also had some accessions with cherry type fruits i.e. RV100511, RV100169, RV100190, RV100249, RV100356,

RV100245, RV100343, RV100217, RV100218, RV100346 and RV100242 (Table 4:13).

c) Colour.

African tomato was not as diverse in colour as African eggplant. However they shared the red and yellow colour, with most of them looking same. Most the the African tomato and eggplants were red in colour , making it difficult to differentiate the two crops (Table 4:13).

d) Fruit surface texture.

Both African tomato and eggplant had either smooth or ridged fruit surface (Table 4:13).

Table 4:13. Comparison of the African tomato and eggplant accessions morphological traits at fruiting stage.

a) African tomato			b) African eggplant			Remarks
Tindi	**050589**		RV100343	RV100246	RV100327	
V1006826						Some african tomato and eggplant were cherry/plum type, medium and large sized. Colour among the cherry/plum types ranged between red and round.
						
V1006842			RV100246		RV100330	
	V1006838					Fruit surface texture for both african tomato and eggplant was either smooth or ridged,
						
						The shape in African tomato and eggplant ranged from round, oval, oblong, kidney,

Key: a) African tomato, b) African eggplant. Similarities were observed in the morphological traits of African tomato and African eggplant in; Shape- both had similarities in round, oval and kidney shapes, **Size-** Both had cheery sized, medium sized and big size**. Colour-** they both shared red and yellow colours**. Fruit surface texture-** the two crops fruit surface was either smooth or ridged.

4.4 Discussion

Significant morphological diversity among the African tomato and eggplant using the morphological descriptors was observed. This variance was manifested by both vegetative traits and reproductive parameters.

In African tomato, 28 accessions showed presence of anthocyanins on the stem and 53 accessions had anthocyanins on the petiole (Table 4:2). According to Peter *et al* ., (2008), presence of purple colourlation on the stems, petioles and leaf veins show high levels of anthocyanins.

Dendogram representation showed that African tomato accessions clustered with branching occurring at a very low phenon line, this suggest that a broad to overall similarity among all the accessions, this can be caused by the ability of the African tomato to self pollinate s also observed by Lawal *et al*., (2007).

Morphological variation caused by the quantitative traits of the African tomato showed a coefficient variation of 2.2% to 14% (Table 4:5). Accornding to Bernousi *et al* ., (2011), any coefficient variation below 20% is considered good.

Principal component analysis (PCA) showed that of the 14 dimensions tested, only seven had an eigenvalue of 1 (Table 4:6). Accornding to Chatfield and Collins (1980) any eigenvalue less than 1 should be eliminated while eigenvalues of 1 and above are considered significant.

In a simple correlation matrix of the African tomato, both strong positive and high negative correlation was observed among the phenotypic traits. Strong positive correlation on yield were r=0.446, 0.264, 0.271, 0.3539, 0.65 and 0.95 for fruit yield , leaf width, internode height,fruit shape, fruit length and fruit width respectively

(Table 4:7) suggesting that these traits can be used for selection. However, phenotypic traits with a high negative correlation cannot be used during selection program. This agrees to findings of Slewinski, (2012) and Kisua *et al.,* (2015) who stated that parameters with stron correlation are important in selection programs .

Other strong positive correlation on yield included leaf width, plant diameter would contribute to good quantity of food synthesized by plant during photosynthesis as plant width could serve well in water and translocated food from aerial plant as also observed by Shafiei, (2000).

The qualitative morphology analysis showed that the fruit morphology had the highest genetic index of 0.99 (Table 4:3). Similar results were observed by Lawal *et al.,* (2007). In African eggplant, significant morphological diversity was observed at vegetative and reproductive stages. Fruit morphology was diverse in terms of shape, shape , size, colour and fruit texture , while vegetative morphology included presence of prickles on the stem and on leaf surfaces, other sources of variation included the plant growth habit ranging from very tall, tall, intermediate, postrate and very prostrate (Table 4:8).

Qualitative morphology showed that plant growth habit, petiole, stem and vein colour were among the main source of variation with a genetic index of 1 (Table 4:9). Eight of the fourteen dimension tested in principal component analysis had eigenvalues equal to 1. This indicated that only eight were significant, while the others with eigenvalues less than 1 were not significant . Main variance contributed by plant growth habit, fruit colour, fruit width, fruit length, fruit texture, number of prinkles, fruit shape and number of flower heads (Table 4:8).

At the selection and breeding level, considerable phenotypic differences among and within the phenotypic traits with a high positive correlation may be used for selecting best accessions or selection of parents in obtaining F1 hybrid herotic for yield or with intermediate or new characteristics (Adeniji and Aloyse 2012).

From the comparison done morphologically, African tomato and eggplant seem to be notably related (Table 4:13). This is in agreement with Romano *et al.,* 2014 and Zhou *et al,* 2009 on effects of grafting on tomato and eggplant that tomato and eggplants are closely related being from the same family and can be easily grafted.

Overall, the morphological characteization analysis in this study shows that phenotypic markers were useful in characterizing the African tomato and African eggplant

CHAPTER FIVE

TRANSCRIPTOME ANALYSIS OF AFRICAN TOMATO AND EGGPLANT CROP USING HIGH THROUGHPUT NEXT GENERATION SEQUENCING FOR IMPROVED RESISTANCE TO BIOTIC AND ABIOTIC STRESS, FOR IMPROVED SHELF LIFE AND HIGHER YIELD.

Abstract

RNA-seq is the use of high – throughput sequencing technologies for characterizing the RNA content and the composition of a given sample**.** The African Solanaceae crops transcriptome has not been done to determine or reveal the genes responsible for their adaptability to harsh climatic conditions. The main objective of this study was to analyze the African tomato and African eggplant transcriptome using RNA-seq technique which is able to reveal genes of importance to improved adaptability to harsh conditions, improved yield and quality. Balanced block chain design was used to set up the experiment. Ten African eggplant and seventeen African tomato accessions were planted each with 3 biological replicates. Leaf samples were collected 3 weeks after planting and fruit samples were sampled at mature green, mature breaker and mature red stage each two replicates. Raw sequences were cleaned and filtered using the NGS tool kit while TopHat software was used for identification of differential gene expression. The African tomato filtered sequences were aligned to the reference genome using TopHat while in African eggplants, *de novo* assembly was done using Trinity software. A reference genome for African eggplant was constructed using all the 80 samples from the ten samples collected in duplicates from four stages. This reference genome was used to align the African

eggplant accessions using TopHat. A total of 329,018,858 and 303,754,051 sequences from African eggplant and tomato respectively, were obtained after filtering. A total of 18,129 and 173,194 genes were differentially expressed from African tomato and eggplant respectively. Significant differential gene expression was observed between the various fruiting stages in both African tomato and eggplant at α 0.001. African tomato and African eggplant accessions expressed unknown genes which could be studied for their importance in tomato and eggplant crop improvement . The generated African eggplant reference genome can be of great use in improving the current eggplant database. Results from this study showed that African tomato and eggplant harbours resistance genes to abiotic and biotic stresses which can be utilized to improve the cultivated cultivar.

5.1 Introduction

Transcriptome analysis is the study of gene expression. Previously, gene expression studies were restricted to small- scale quantitative PCR analyses of candidate genes or were dependent on cross-species hybridization on microarrays (Naurin et al., 2008).

RNA-seq, (whole-transcriptome shotgun sequencing), refers to the use of high-throughput sequencing technologies to characterize the RNA content and composition of a given sample (Nielsen et al,. 2011). With the rapid development of next-generation sequencing technology, RNA deep-sequencing (RNA-seq) has become more efficient, less expensive and highly reproducible (Cloonan et al 2008; Nagalakshmi et al,. 2008).

In eggplant research, RNA-sequencing has been used to characterize the transcriptomes of resistant (R) and susceptible (S) strains before (R0, S0) and after (R1, S1) *R. solanacearum* inoculation. First, RNA-Seq technique was used to classify a mass of transcripts from two eggplant lines inoculated with *R. solanacearum*. Secondly, it was used to study gene expression changes where gene functions following *R. Solanacearum*-inoculation were identified and annotated.

These genes were involved in Salicyclic Acid and Jasmonic Acid signaling pathways which are involved in the defense response of eggplant against *R. Solanacearum* infection (Chen *et al.*, 2017).

With the availability of a high-quality tomato genome sequence and next-generation sequencing (NGS), RNA-seq technology has rapidly become a popular tool for genome- wide expression profiling, providing the potential to better understand the comprehensive host-pathogen interactions (Chen *et al* ., 2017).

RNA-seq has been used to study patterns of plant hormone gene expression under normal sunlight conditions to elucidate relationships between circadian rhythms, plant hormones, and prediction of phenome using gene expression analysis. RNA-seq analysis has also been used to study tomato plant hormone mechanisms, for instance Salicylic Acid (SA) has been shown to inhibit Abscisic acid (ABA) and the downstream responses reflected cellular SA/ABA ratios (Nielsen et al., 2011). In contrast, Jasmonic Acid (JA) production is promoted by ET signalling, which suppresses SA signals. These hormone pathways lead to the expression of genes that suppress other pathways, suggesting that each hormone is intricately regulated by interactions and cross relationships between environmental and growth response variables (Wang et al., 2013).

According to Tanigaki *et al*., (2015), RNA-seq has revealed that plant phenotypes are highly dependent on cultivation conditions, and corresponding changes in gene expression are strongly correlated with multiple plant hormone-related processes, including gating stomata and flower initiation. Chen *et al*., (2017) analyzed the differentially expressed genes between parthenocarpic and non- parthenocarpic eggplants, for instance, transcriptome profiles of flower buds of a parthenocarpic eggplant line PP05, and two non-parthenocarpic eggplant lines PnP05 and GnP05, were analyzed using the next generation sequencing (RNA-seq) technology. In his study, Chen *et al*., (2017) successfully screened differentially expressed genes between the parthenocarpic and non-parthenocarpic eggplants, so as to comprehensively exploit parthenocarpic genes and bring insight into the mechanisms of parthenocarpy in eggplant using the RNA-seq technique.

Numerous transcriptome data sets have been produced, made publicly available, and reanalysed by other researchers (Wang *et al*., 2013; Ahmad *et al*,. 2015). However, so far, RNA-seq analyses of African tomato and eggplant have not been reported. This study was done to analyse the transcriptome of the African tomato and African eggplant to identify genes of intrest in terms yield, quality , biotic and abiotic stresses.

5.2 Materials and methods

5.2.1. Transcripts assembly and differential gene expression in African tomato.

The African tomato non-rRNA Fastq sequences were aligned to the tomato genome from the Sol genomic network (Solanum Lycopersicum GCF-000188115.3_SL2.5.0) for differential gene expression analysis.

TopHat software was used to align the non-rRNA to *Solanum lycopersicum* SL2.5.,0 genome from ensembl-genomes, for initial assembly, this is because TopHat identifies splice junctions and handles assembly of reads to reference genome even where big gaps (introns) are present. This is important for gene expression in coding regions only. Splice junctions occur between an intron and exon, it was also used to convert the non rRNA sequence to a BAM file, then Cufflink was used to assemble the transcripts followed by the Cuff Compare which compared two or more transcripts. the compared transcripts were merged using the Cuff Merge and Cuffdiff was used for differential gene expressions. The Cummerbund was used to plot abundance of the differential genes expressed while the R Studio was used to visualize graphs charts and tables (Figure 5).

5.2.2. Transcript assembly and de novo assembly in African eggplant

The Next generation sequencing tool kit (NGSTool kit) was used to remove the adapters used during the sequencing process, low quality sequences with a phred quality score Q < 20 and a percentage high quality cut off read length for high quality was 70% and ambiguous sequences with N. The High quality clean reads of all the 80 samples (10 samples each with 4 fruiting stages and 2 biological replicates each) were assembled into 1 long transcript using the Trinity. This consecious sequence served as the reference genome for the eggplant.

5.2.3. Differential gene expression, data visualization and presentation

TopHat and Genome Analysis Tool kit (GATK) was used to align each of the samples to the generated reference genome (Figure 5:1). Data was visualized in excel spread sheets, in R-studio, PCA and scatter plots analysis.

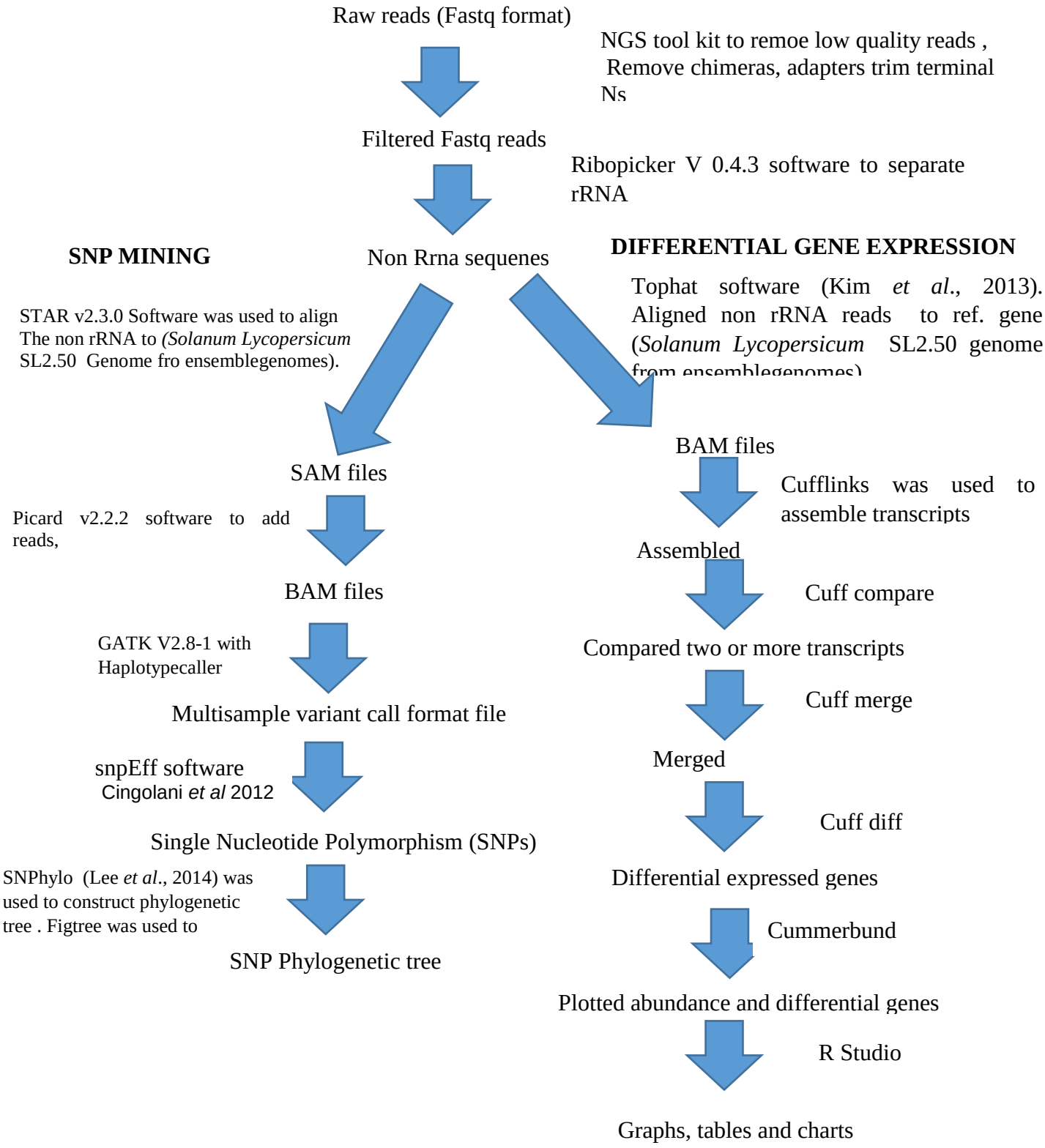


Figure 5:1. Flowchart showing the various steps and bioinformatics tools used in differential gene expression and SNP mining in African tomato

5.3 Results for transcriptome analysis of African tomato accessions

5.3.1. Fruit Samples collected

Photos of fruits sampled at three development stages included mature green (MG), mature breaker (MB), and mature red (MR) as shown in Plate 5:1 **.**

MG MB MR MG MB MR

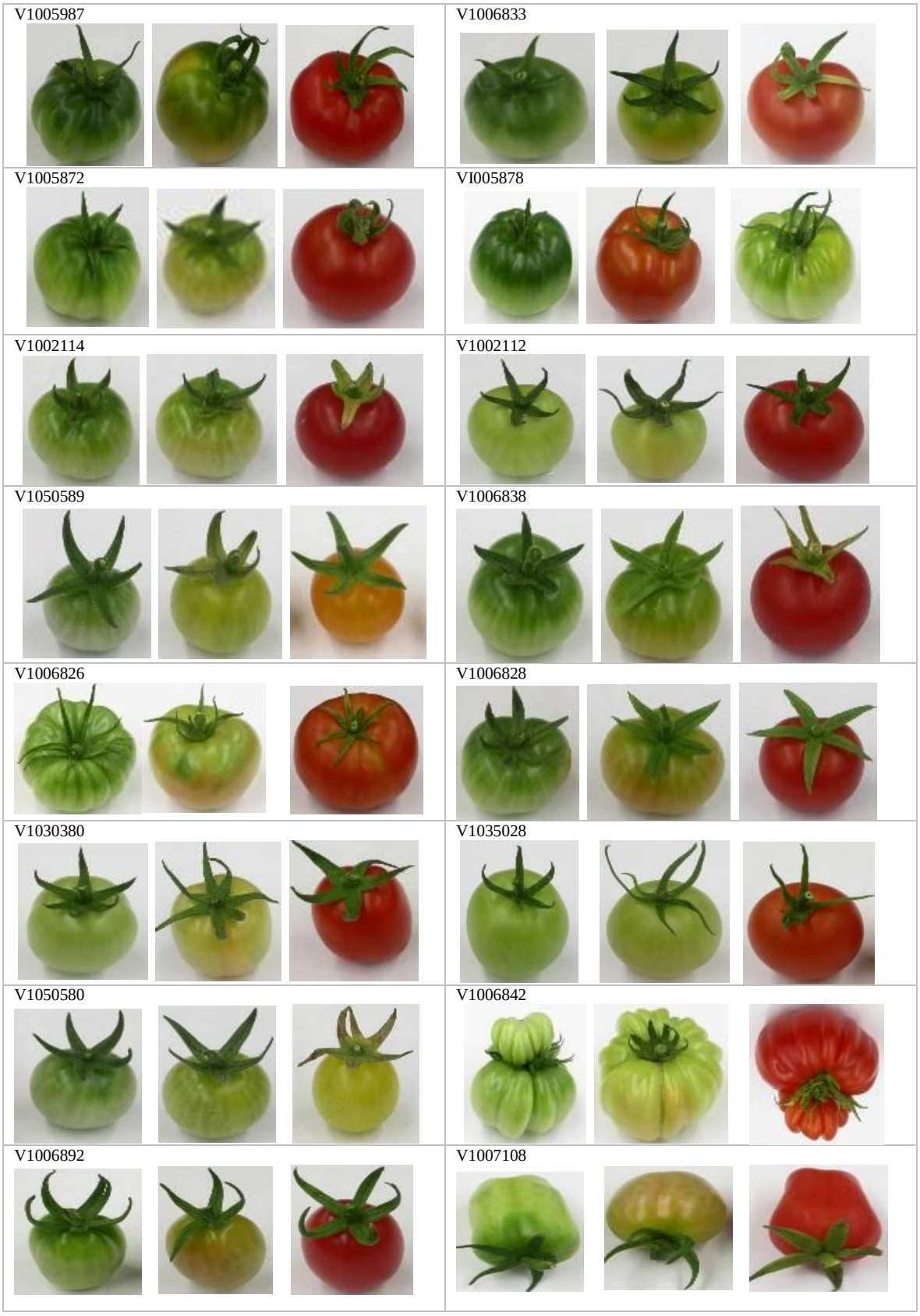


Plate 5:1. African tomato collected at the various sampling stages ie (Mature green (MG), mature breaker (MB) and mature red (MR).

5.3.2 Quality check and sequence validation of the African tomato accessions

After filtering, 90.8 million reads before fruiting, 91.6 million reads at mature green stage, and 84.2 million reads at mature breaker and 82.4million reads mature red were retained and used for further analysis. (Table 5:1). The percentage of the high quality filtered reads (non rRNA filtered sequences) ranged between 66.66% to 95.08 % (Table 5:1 and Figure 5:2). The GC content had its peak between 40-45% in all the African tomato accessions (Figure 5:3).

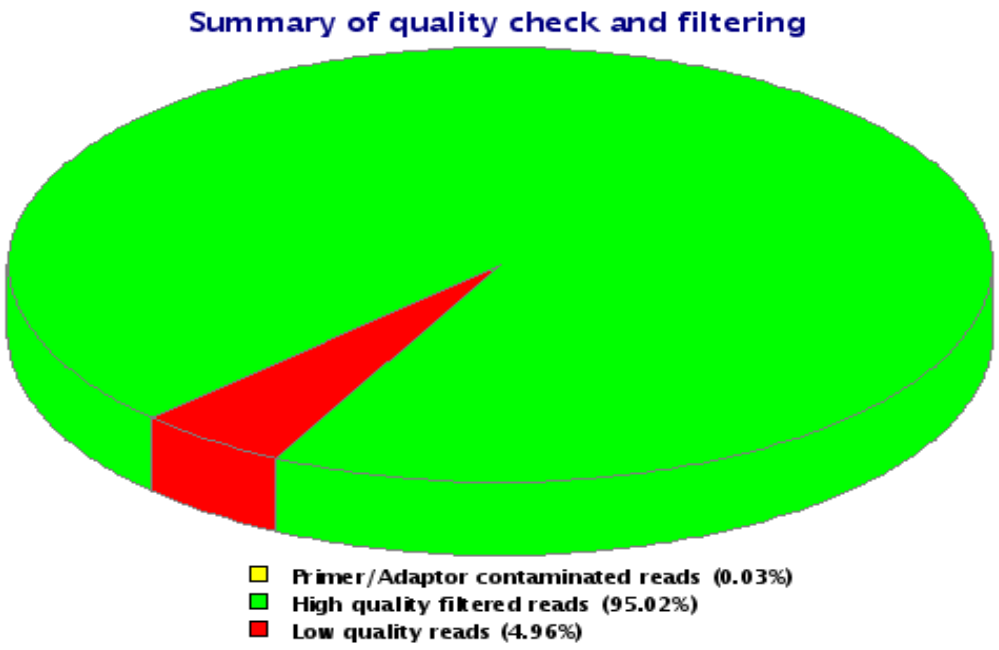


Figure 5.2: Quality of raw data and the % age contaminant (primer/ adaptor, low quality reads). Composition of contaminants included; Primer/ adaptor contaminated reads (0.03%), low quality reads (4.96%) and high quality filtered reads (95.02%)

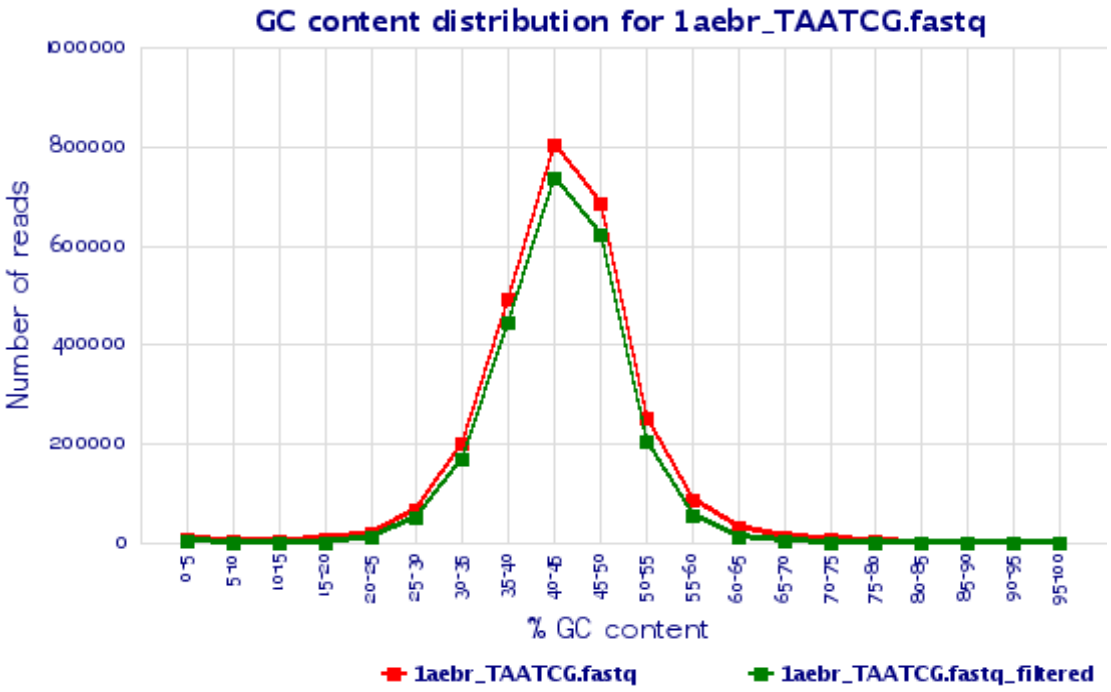


Figure 5.3: GC content distribution in Sample 1aebr. GC content in all Africa tomato accessions had its peak at 40-45%

Table 5:1. Quality of sequencing reads; raw reads, clean reads and % high quality reads for the African tomato accessions

Accession stage	Total raw reads	Total HQ reads	TotalHQ filtered reads	HQ Filtered reads (%)	Accession stage	Total raw reads	Total HQ reads	TotalHQ filtered reads	HQ Filtered reads (%)
1atbf	5,426,387	4,754,125	4,705,716	86.72	**2atmr**	3,404,500	2,532,197	2,505,396	73.59
1btbf	11,444,763	10,715,630	10,697,037	93.47	**2btmr**	4,151,364	3,011,654	2,951,412	71.09
1atmg	4,840,643	4,395,734	4,364,506	90.16	**4atbf**	5,556,033	5,321,002	5,271,002	94.87
1btmg	5,146,137	4,879,802	4,855,432	94.35	**4btbf**	4,006,465	3,065,291	3,046,299	76.03
1atbrk	3,256,318	2,417,229	2,397,862	73.64	**4atmg**	3,850,260	2,860,542	2,852,488	74.09
1btbrk	4,232,002	3,118597	3,056,448	72.22	**4btmg**	3,455,500	2,593,072	2,588,559	74.91
1atmr	4,027,609	3,489,020	3,485,959	86.55	**4atbrk**	3,410,143	3,096,847	3,096,175	90.79
1btmr	4,793,925	3,509,169	3,500,179	73.01	**4btbrk**	2,937,664	2,626,788	2,626,393	89.40
2atbf	3,476,247	2,606,953	2,581,477	74.26	**4atmr**	3,799,612	2,874,459	2,873,518	75.63
2btbf	3,754,879	3,487,095	3,449,994	91.88	**4btmr**	2,699,987	2,010,681	2,009,993	74.44
2atmg	4,928,620	3,679,754	3,677,229	74.61	**5atbf**	6,808,919	6,200,596	6,195,281	90.99
2btmg	4,896,483	3,666,904	3,666,486	74.88	**5btbf**	2,417,042	2,067,295	2,040,001	84.40
2atbrk	5,321,594	4,819,138	4,757,855	89.41	**5atmg**	3,795,424	2,761,787	2,753,912	72.56
2btbrk	5,076,526	4,590,025	4,501,686	88.68	**5btmg**	5,326,499	4,649,958	4,642,580	87.16
5atbrk	4,604,418	4,180,294	4,165,814	90.47	**7atmg**	6,096,329	4,590,423	4,584,734	75.20
5btbrk	4,161,739	3,794,981	3,794,088	91.17	**7btmg**	5,699,003	4,284,053	4,282,924	75.15
5atmr	4,572,146	3,414,244	3,369,733	73.70	**7atbrk**	6,048,966	5,438,853	5,335,248	88.20
5btmr	4,018,622	2,991,984	2,940,623	73.17	**7btbrk**	6,113,974	5,585,183	5,578,846	91.25
6atbf	3,964,481	3,476,932	3,475,104	87.66	**7atmr**	4,930,261	3,689,395	3,673,377	74.51
6btbf	307,531	269,416	269,306	87.57	**7btmr**	4,321,111	3,196,211	3,145,391	72.79
6atmg	4,030,213	3,665,710	3,664,901	90.94	**8atbf**	4,055,841	3,755,340	3,668,221	90.44
6btmg	5,691,892	5,167,750	5,150,574	90.49	**8btbf**	3,706,391	3,388,374	3,378,648	91.16
6atbrk	4,219,020	3,127,919	3,073,458	72.85	**8atmg**	4,636,502	3,491,778	3,488,742	75.25

Accession stage	Total raw reads	Total HQ reads	Total HQ filtered reads	HQ Filtered reads (%)	Accession stage	Total raw reads	Total HQ reads	Total HQ filtered reads	HQ Filtered reads (%)
6btbrk	3,512,894	2,605,012	2,555,841	72.76	**8btmg**	4,821,505	3,614,767	3,612,454	74.92
6atmr	4,226,269	3,629,880	3,624,243	85.76	**8atbrk**	3,937,841	2,959,916	2,944,055	74.76
6btmr	4,845,874	3,951,372	3,944,640	81.40	**8btbrk**	3,775,554	2,827,170	2,810,240	74.3
7atbf	3,879,000	3,551,140	3,545,057	91.39	**8atmr**	4,323,044	3,904,047	3,897,422	90.15
7btbf	4,963,224	4,448,602	4,423,086	89.12	**8btmr**	4,546,538	4,247,318	4,246,034	93.39
9atbf	4,027,892	2,760,045	2,684,967	66.66	**10btmr**	3,144,715	2,356,045	2,354,901	74.88
9btbf	6,939,085	6,523,695	6,462,867	93.14	**11atbf**	5,574,678	4,763,321	4,761,699	85.42
9atmg	5,075,566	4,608,914	4,600,680	90.64	**11btbf**	4,704,093	3,953,002	3,950,808	83.99
9btmg	5,755,411	4,989,877	4,974,197	86.43	**11atmg**	4,396,313	3,988,711	3,987,373	90.70
9atbrk	4,366,952	3,978,979	3,970,585	90.92	**11btmg**	7,307,687	6,626,574	6,623,910	90.64
9btbrk	3,005,425	2,263,851	2,258,867	75.16	**11atbrk**	4,462,911	4,070,525	4,059,462	90.96
9atmr	3,303,346	2,312,239	2,236,103	67.69	**11btbrk**	5,192,306	4,721,743	4,711,447	90.74
9btmr	3,949,370	2,955,840	2,954,988	74.82	**11atmr**	4,730,146	4,488,326	4,484,875	94.81
10atbf	37,789,450	35,421,804	35,203,789	93.16	**11btmr**	4,539,622	4,321,319	4,316,119	95.08
10btbf	4,456,758	4,075,011	4,020,153	90.20	**12atbf**	5,092,825	3,429,421	3,425,324	67.26
10atmg	4,868,097	4,412,257	4,401,120	90.41	**12btbf**	4,931,112	3,990,971	3,988,764	89.00
10btmg	7,963,885	7,214,717	7,177,941	90.13	**12atmg**	5,047,975	4,572,875	4,557,244	90.28
10atbrk	3,536,194	3,326,434	3,324,584	94.02	**12btmg**	2,926,491	2,184,067	2,179,808	74.49
10btbrk	4,657,201	4,207,796	4,192,494	90.02	**12atbrk**	3,934,164	2,934,264	2,911,800	74.01
12atbrk	3,934,164	2,934,264	2,911,800	74.01	**15atmg**	5,139,275	4,656,425	4,641,162	90.31
12btbrk	3,797,220	2,779,989	2,716,576	71.54	**15btmg**	6,669,550	5,944,820	5,779,864	86.66
12atmr	5,968,672	5,533,617	5,529,271	92.64	**15atbrk**	3,737,327	2,790,984	2,766,702	74.03
12btmr	7,904,236	7,185,639	7,176,873	90.80	**15btbrk**	3,842,000	2,866,393	2,828,031	73.61
13atbf	4,372,434	3,990,916	3,987,744	91.20	**15atmr**	5,454,620	5,109,415	5,095,349	93.41
13btbf	4,747,599	4,241,897	4,232,807	89.16	**15btmr**	5,711,235	5,353,223	5,328,583	93.30

Accession and stage of fruiting	Total raw reads	Total HQ reads	Total HQ filtered reads	HQ Filtered reads (%)	Accession and stage of fruiting	Total raw reads	Total HQ reads	Total HQ filtered reads	HQ Filtered reads (%)
13atmg	3,400,426	2,541,553	2,538,455	74.65	16atbf	3,521,221	3,351,448	3,337,737	94.79
13btmg	4,534,758	3,401,198	3,395,244	74.87	16btbf	3,896,896	3,697,392	3,663,685	94.02
13atbrk	4,006,635	2,983,727	2,968,043	74.08	16atmg	4,352,028	3,251,068	3,230,451	74.23
13btbrk	4,053,612	3,034,923	3,014,344	74.36	16btmg	4,155,906	3,110,248	3,102,150	74.64
13atmr	5,897,895	5,594,842	5,582,000	94.64	16atbrk	4,208,768	3,140,519	3,138,193	74.56
13btmr	4,480,806	4,234,491	4,222,938	94.25	16btbrk	4,017,317	3,010,594	3,006,918	74.85
15btbf	4,087,239	3,836,656	3,832,332	93.76	16atmr	4,138,592	3,087,139	3,084,802	74.54
15btbf	1,974,654	1,728,686	1,723,887	87.30	16btmr	4,327,058	3,236,995	3,233,440	74.73
17atbf	3,515,651	3,325,598	3,323,122	94.52	18atmr	5,288,544	4,546,978	4,543,319	85.91
17btbf	2,779,319	2,628,228	2,617,343	94.17	18btmr	4,404,279	4,181,537	4,171,473	94.71
17atmg	4,457,894	4,053,098	4,046,468	90.77	19atbf	6,215,297	5,836,401	5,830,733	93.81
17btmg	4,817,901	4,355,628	4,339,713	90.07	19btbf	6,925,038	6,518,737	6,508,407	93.98
17atbrk	3,471,952	2,617,344	2,608,056	75.12	19atmg	4,937,177	4,485,682	4,483,732	90.82
17btbrk	3,355,114	2,502,850	2,478,478	73.87	19btmg	3,467,849	3,143,079	3,141,629	90.59
17atmr	4,379,732	3,925,906	3,925,040	89.62	19atbrk	10,801,928	10,191,760	10,187,598	94.31
17btmr	4,287,988	3,878,098	3,877,445	90.43	19btbrk	8,760,562	8,090,672	8,080,603	92.24
18atbf	10,164,193	8,035,722	7,992,867	78.64	19atmr	3,661,310	2,950,431	2,938,956	80.27
18btbf	6,866,644	6,534,544	6,496,518	94.61	19btmr	4,056,345	3,272,385	3,264,108	80.47
18atmg	5,284,349	4,793,116	4,788,764	90.62					
18btmg	6,189,042	5,567,502	5,561,031	89.85					
18atbrk	3,784,357	2,838,498	2,826,898	74.70					
18btbrk	7,144,935	6,146,854	6,106,081	85.46					

Key: Raw reads means – sequences with all contaminants - adaptors, chimeras, primers, ribosomal RNA, clean reads means – sequences whose without the contaminants but has rRNA, while high quality reads means, all contaminants have been removed and they have met the threshold set. Mg- mature green, MB- Mature breaker, MR- Mature red, HQ- High quality

5.3.3. Dendogram representation for biological replicates validation

Dendogram representation between the biological replicates showed that the replicates grouped close to each other. Each replicate of the different fruit development stage was true to type, since there were no mixups between the replicates and stages. The stage before fruiting however grouped far from the other three stages (mature green, mature breaker and mature red fruit stages) (Figure 5:4)

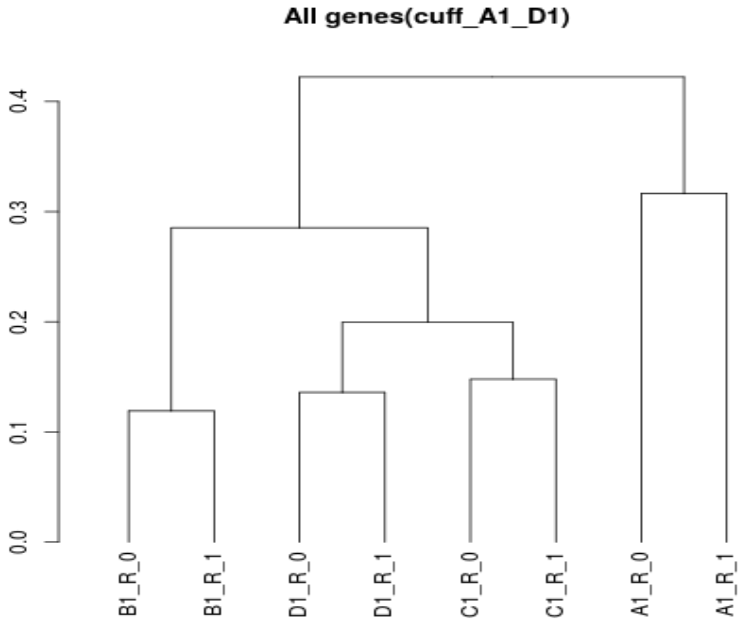


Figure 5:4. Dendogram representation of replicates (A- Before fruiting, B- mature green, C- mature breaker, D- mature red).

5.3.4. Alignment to the reference genome and gene expression

Short sequencing reads were mapped to the annotated tomato reference genome (*Solanum Lycopersicum* GCF-000188115.3_SL2.5.0) using TopHat with the default parameters. Properly mapped reads were separated from the unmapped reads using Filter SAM by setting the flag values in SAMtool.

Among the 90.8 M reads before fruiting, about 80.89- 94.87% reads were properly mapped to the tomato genome, in the 91.6 M reads in mature green stage, 74.9-94.3% mapped to the tomato genome. In mature breaker, 73.64-94.64 % of the 84.2 M reads mapped to the tomato genome while in mature red, 73.59- 94.64% of the 82.4 M mapped to the tomato genome. After performing alignment, SAMtools was used to remove duplicate reads (Table 5:2).

A total of 18,129 genes were differentially expressed at α=0.02 in the 17 African tomato accessions. Among these genes, 700 genes were novel/unknown, 16,226 were putative/predicted genes while 1,137 genes were known genes (Table 5:2).

Table 5:2. Total number of genes expressed in all the 17 African tomato accessions

Type of gene	No of expressed gene
Total genes expressed	18, 129
Known genes	1,137
Predicted / putative genes	16,226
Unknown genes	700

5.3.5 Differential gene expression

Differential gene expression was observed at α= 0.001. Scatter plot matrix Analysis (Figure 5:5), and SignGene plot analysis (Figure 5:6) showed that there was significant differential gene expression at the different fruit developmental stages. The differential gene expression increased as the fruit progressed and vice versa. For instance, in accession V1005987, only 2, 610 were differentially expressed between BF and MG stages, 4,230 genes were differentially expressed between BF and MB stages and 4,689 between BF with MR stage (Figure 5:5 and 5:6)

Significant differential gene expression was also observed between the specific stages for instance, the number of differentially expressed genes between MB and MR was 889; between MB and MG was 1,348; and 2,143 differential expressed genes between MG and MR (Figure 5:6).

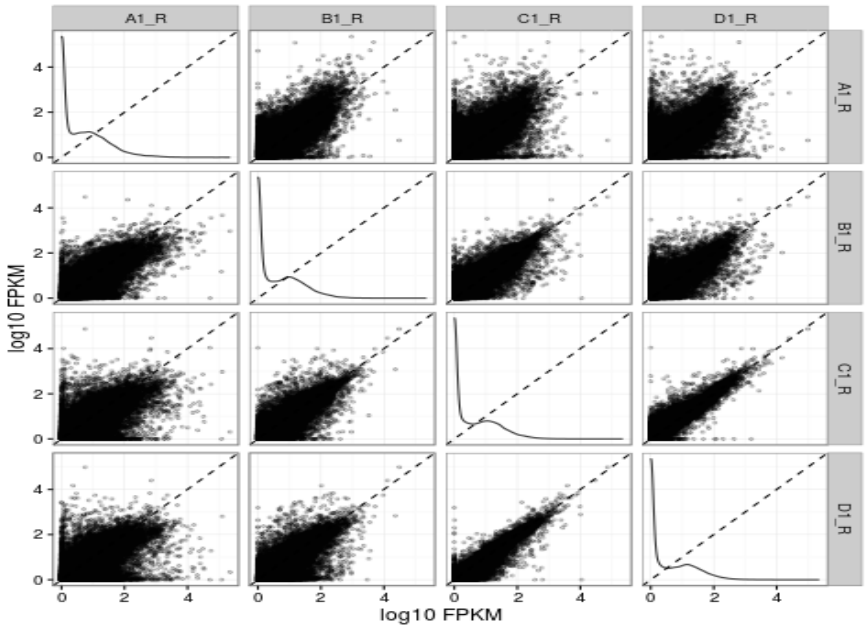
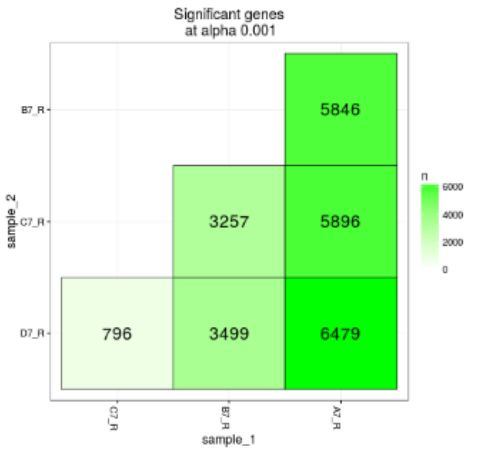
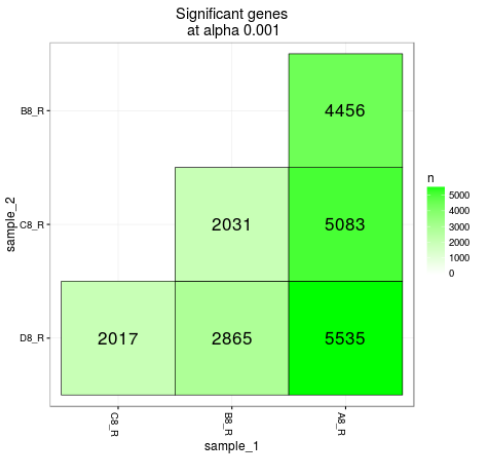
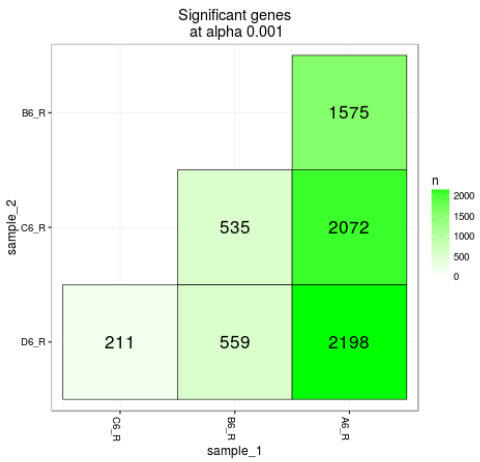


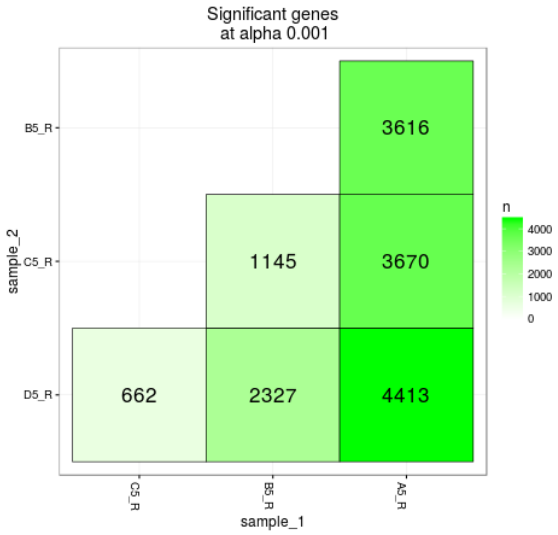
Figure. 5:5. Scater matrix showing differential gene expression among and within the four different African tomato fruit development stages at ∞ 0.001 (A- before fruiting, B- mature green, C- mature breaker, D- mature red).

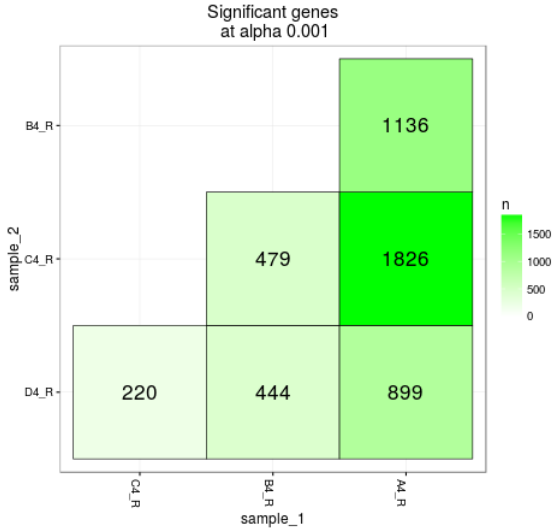
There was significant differential gene expression between all the different fruiting stages. The stages BF and MR stages had the highest differential genes expression in all the accessions; followed by between BF and MB, BF and maturMG, MG and MR, then MG and MB while there was no much diferential gene expression between MB and MR (Figure 5:5 and 5:6).

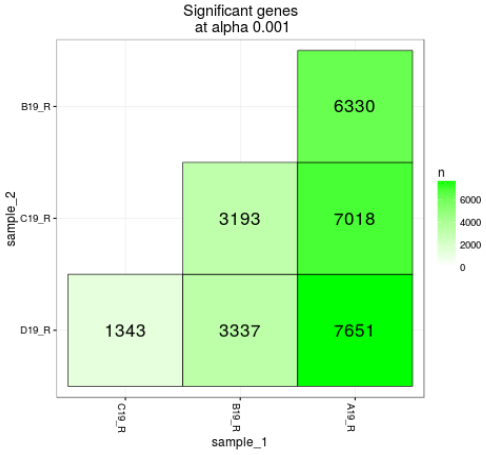

Significant genes
at alpha 0.001
B7_R
C7_R
D7_R
sample_2
5846
3257
5896
796
3499
6479
n
6000
4000
2000
0
C7_R
B7_R
A7_R
sample_1

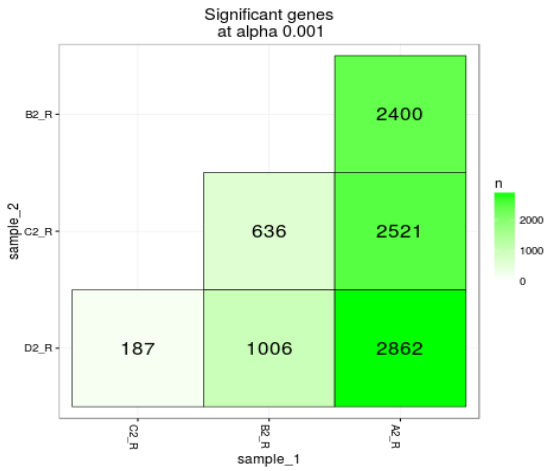

Significant genes
at alpha 0.001
B8_R
C8_R
D8_R
sample_2
4456
2031
5083
2017
2865
5535
n
5000
4000
3000
2000
1000
0
C8_R
B8_R
A8_R
sample_1

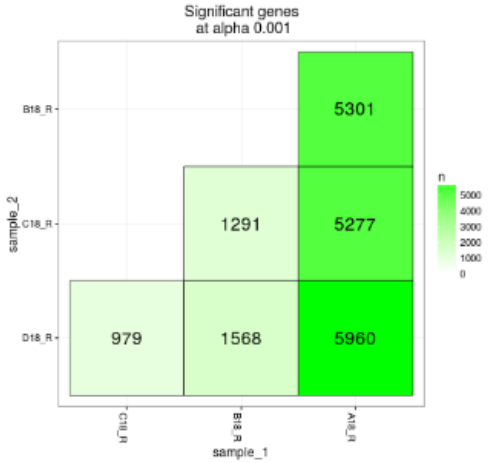

Significant genes
at alpha 0.001
B6_R
C6_R
D6_R
sample_2
1575
535
2072
211
559
2198
n
2000
1500
1000
500
0
C6_R
B6_R
A6_R
sample_1

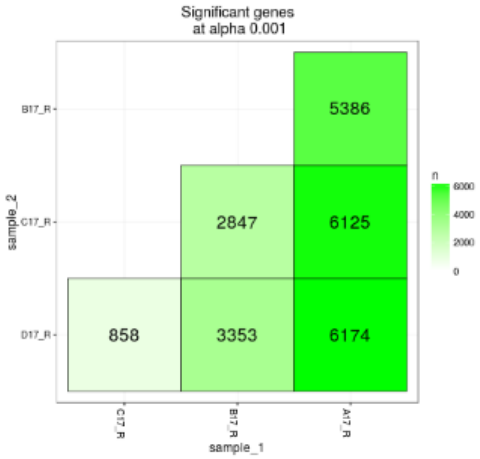

Significant genes
at alpha 0.001
B5_R
C5_R
D5_R
sample_2
3616
1145
3670
662
2327
4413
n
4000
3000
2000
1000
0
C5_R
B5_R
A5_R
sample_1

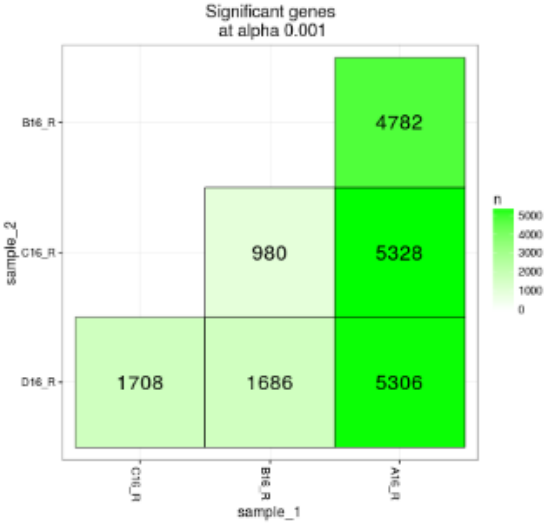

Significant genes
at alpha 0.001
B4_R
C4_R
D4_R
sample_2
1136
479
1826
220
444
899
n
1500
1000
500
0
C4_R
B4_R
A4_R
sample_1

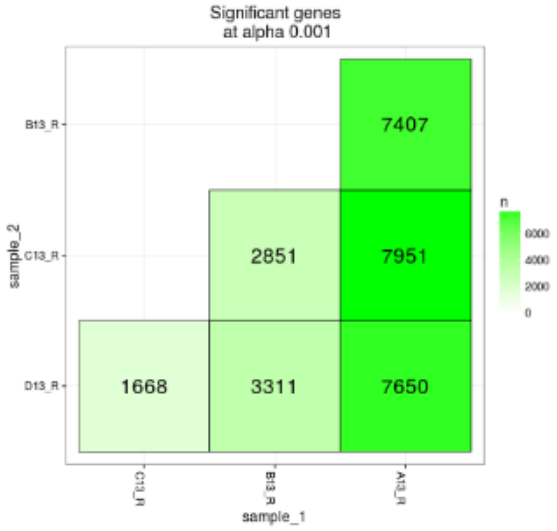

Significant genes
at alpha 0.001
B19_R
C19_R
D19_R
sample_2
6330
3193
7018
1343
3337
7651
n
6000
4000
2000
0
C19_R
B19_R
A19_R
sample_1

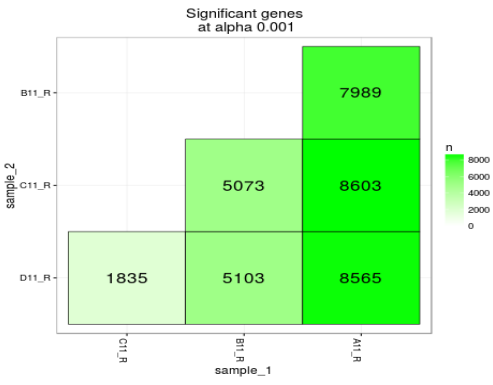

Significant genes
at alpha 0.001
B2_R
C2_R
D2_R
2400
636
2521
187
1006
2862
sample_2
C2_R
B2_R
A2_R
sample_1
n
2000
1000
0


Significant genes
at alpha 0.001
B18_R
C18_R
D18_R
5301
1291
5277
979
1568
5960
sample_2
C18_R
B18_R
A18_R
sample_1
n
5000
4000
3000
2000
1000
0


Significant genes
at alpha 0.001
B17_R
C17_R
D17_R
5386
2847
6125
858
3353
6174
sample_2
C17_R
B17_R
A17_R
sample_1
n
6000
4000
2000
0


Significant genes
at alpha 0.001
B16_R
C16_R
D16_R
4782
980
5328
1708
1686
5306
sample_2
C16_R
B16_R
A16_R
sample_1
n
5000
4000
3000
2000
1000
0


Significant genes
at alpha 0.001
B13_R
C13_R
D13_R
7407
2851
7951
1668
3311
7650
sample_2
C13_R
B13_R
A13_R
sample_1
n
6000
4000
2000
0


Significant genes
at alpha 0.001
B11_R
C11_R
D11_R
7989
5073
8603
1835
5103
8565
sample_2
C11_R
B11_R
A11_R
sample_1
n
8000
6000
4000
2000
0

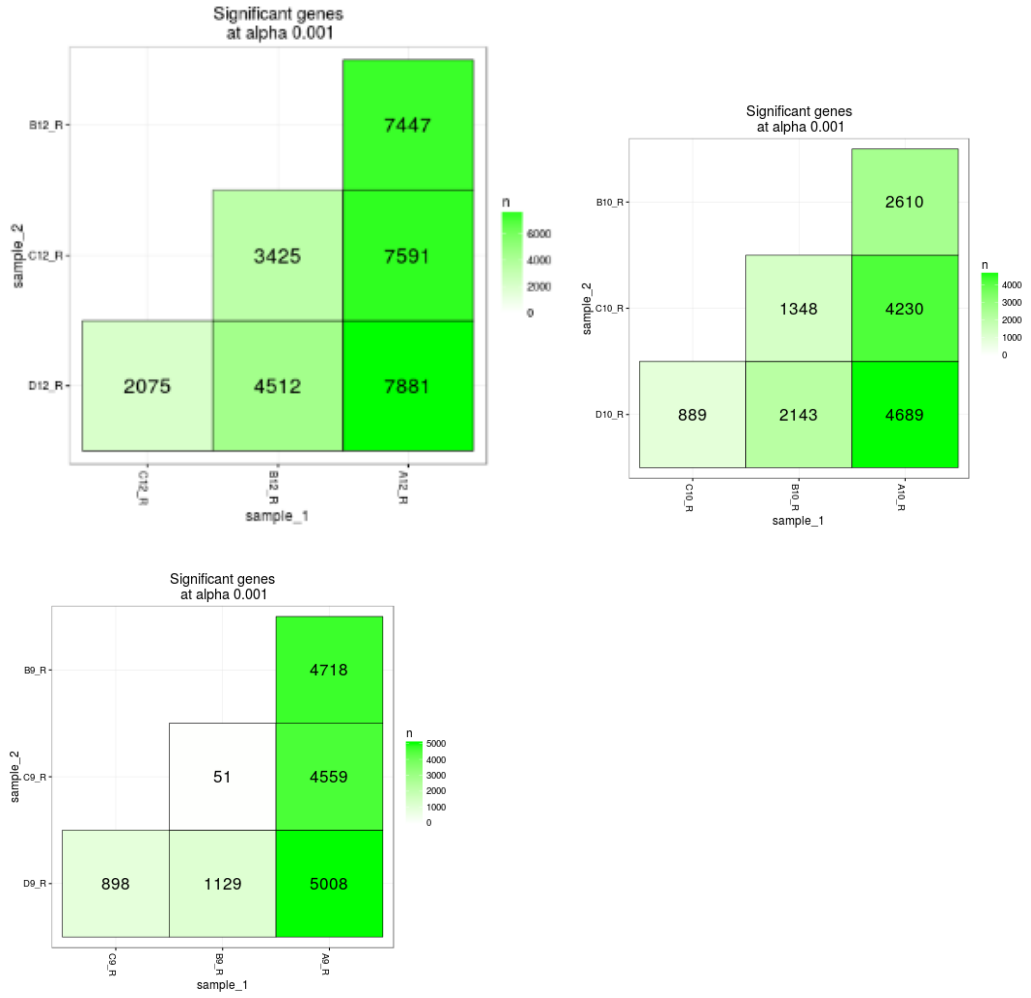


Figure 5:6. Significant Gene - plots showing differential gene expression among and within the four different African tomato fruit development stages at ∞ 0.001 (A - before fruiting, B - mature green, C- mature breaker, D- mature red).

In, V1005987 102, 5, 2 and 3 genes were differentially expressed before fruiting, mature green, mature breaker and mature red stages respectively. 19 genes that were 100% expressed in V1005987 before fruiting were also expressed in other accession at lower expressions. Only 3 genes were expressed in V1005987 at 100 % before fruiting and not in any other accession. These included: 1 uknown gene, LOC101260653 and LOC101262259. V1005987, V1007108, Tindi 050580 and V1005875 had a positive relationship in gene differential expression because they

shared similar genes . In mature green stage, 1 unknown gene was only expressed in V1005987 and V1005875 showing that these two accessionss had a genetic similarity.

In V1005987, Before fruiting stage, there were 7, 15 and 80 known genes, unknown genes and putative genes respectively. The known genes included;- 20ox-2, XET4, Dea1, TCP17, LOG8, TPX2, NDPS1.PHS1. In mature green, only 1 unknown gene was differentially expressed while four putative genes were expressed, these included: LOC101245521, LOC101262926, LOC101249128 and LOC101246223. In mature breaker, only 1 unknown gene was differentially expressed and 1 putative gene LOC101268579 was differentially expressed. In mature red , there were 2 uknown genes and 1 putative gene (LOC101249942) differentially expressed (Table 5:3 Appendix III).

Table 5:3. Differential gene expression of known, predicted and unkown genes at the different stages of fruit development in African tomato

	Accession	stage	Gene expressed			Remarks
			Known genes	**Predicted genes**	**Unknown genes**	
1	V1005987	BF	7	80	15	19 unique genes this accession (not found in the other accessions , this accession was closely related to V1005875
		MG	-	4	1	
		MB		1	1	
		MR		1	2	
2	V1006833	BF	9	248	72	Only one putative gene was specific to this accession.
		MG	-	1	-	
		MB	-	-	-	
		MR	-	-	-	
3	V1005872	BF	-	28	27	Only 3 genes were specific differntially expressed in V1005872
		MG	-	1	-	
		MB	-	-	-	
		MR	-	-	-	
4	V1005878	BF	6	296	28	5 genes were specific to V1005878. Closely related to RV102114 and V1006838
		MG		53	18	
		MB	-	-	1	
		MR	-	-	1	
5	V1007108	BF				PR-1a1 and SICYP735A2 genes were only expressed in this accession, 21 genes were specific to this accession, closely related to

						V1005878
		MG	1 (SICYP735A2)	69	12	
		MB	1 (PR-1a1)	6	1	
		MR	-	10	7	
6.	V1002114	BF				Only 4 genes were specific to V1002114. Accessions V1002114 is closely related to V1035028 and V1006838
		MG	-	2	1	
		MB	-	1	-	
		MR	-	-	1	
7.	Tindi 050580	BF	3 (fsm1, TCP23, TCP4)	199	43	4 Genes were specific to this accession closely related to V1007108
		MG	-	8	4	
		MB		2	-	
		MR	-	27	16	
8.	V1002112	BF	19	502	49	
		MG	-	-	-	
		MB	-	-	-	
		MR	-	10	19	
9.	Tindi 050589	BF	5	143	31	15 Genes were specifically expressed in this accession, closely related to Tindi 05080 and V1005878
		MG		25	11	
		MB	-	-	-	
		MR	-	12	9	
10	V1006838	BF	13	483	58	Cevi57 was expressed at the mature green stage, only 1 putative gene was specific to this accession.closely

						related to V1007108 and V1005878
		MG	1 (Cevi57)	16	10	
		MB	-	5	2	
		MR	-	-	-	
11	V1006842	BF	12	415	56	10 Genes were specifically expressed in this accession,
		MG	1 (Cevi57)	33	1	
		MB	-	-	1	
		MR	-	7	10	
12	V1006826	BF	8	356	45	
		MG	1 (GA2OX4)	25	10	
		MB	1	-	1	
		MR	-	1	4	
13	V1005874	BF	8	301	27	29 specific to this accession
		MG	-	38	11	
		MB	-	-	-	
		MR	1 (CHII4)	15	23	
14	V1030380	BF	15	408	48	12 specific to this accession
		MG	-	2	1	
		MB	-	3	3	
		MR	-	11	6	
15	V1006892	BF	7	307	24	
		MG	-	5	-	
		MB	-	-		
		MR	-	1	2	
16	V1035028	BF	3 (TCP 1, TCP 23, AGO 10)	89	16	21 specific to this accession
		MG	-	1	3	
		MB	-	-	7	
		MR	-	-	4	
17.	V1005875	BF	12	439	40	
		MG	-	13	6	

	MB	-	2	1
	MR	-	1	1

Key: Table showing the known, unknown and the putative genes as expressed in each African tomato accession at different fruiting stages .

In accession V1006833, 329 of the 18129 genes expressed were only expressed before fruiting. Of these 329, 72 genes were unknown, 248 putative and 9 known genes . Only 1 putative gene (LOC101250805) was specifically expressed in mature green stage of accession V1006833. However, there were no genes specific to mature breaker or mature red stages in accession V1006833 (Table 5:3).

Only 3 genes were specific In accession V1005872. These genes included;. LOC104645428, LOC101245027 and 1 unknown gene. 55 of the genes expressed in V1005872 were before fruiting stage. Of the 55 differentially expressed genes before fruiting, 27 were unknown while 28 were putative genes. Only 1 putative gene (LOC101245027) was specific to mature green stage in V1005872. However, there were no specific genes expressed at mature breaker and mature red stages (Table 5:3).

In accession V1005878, 330 genes were differentially expressed before fruiting. Of these, only 3 genes were specifically to V10005878. This included 2 unkown genes and 1 putative gene (LOC101258179). Of the 330, 6 were known genes, 28 were unknown while 296 genes were putative genes. The most closely related accession as far as gene expression is concerned is accession Tindi 050580 in mature green stage. At mature breaker, only 1 unknown gene was specific to accession V1005878. In mature red stage, only 1 gene was specific to this accession only and not in any other accession. Sample V1006838 seemed more close to accession V1005878 at mature red stage. 71 genes were differentially expressed at mature green stage only. Of these, 18 were unknown and 53 were putative gene.

In accession RV102114, 299 genes were differentially expressed before fuiting. Of these genes, only four were specific to accession RV102114. These included LOC104645734, LOC101254696, LOC101250916 and LOC104645704. 3 genes at mature green stage (2 putative genes and 1 unknown) however, only one gene was specific to mature green stage and the same gene was expressed in accession V1035028. One putative gene at mature breaker and one unkown mature red stages of RV102114. This same gene was differentially expressed also in s accession V1006838.

In accession V1007108, 357 genes of the 18,129 genes expressed in African tomato were differentially expressed, among therse were 49 unkown/ novel genes, 297 putative genes and 10 known genes. Of these, 63 genes were differentially expressed before fruiting. Only 2 genes were specific to accession V1007108 before fruiting, these included; LOC101249836 and 1 unknown/novel gene. In mature green stage, 82 genes were differentially expressed. SICYP735A2 gene was expressed in mature green stage in addition to 12 unknown genes and 69 putative genes. 12 genes were specific to accession V1007108 and not in any other accession. Accessions V1005878 and V1007108 had similar genes being differentilly expressed at mature green stage. At amature breaker stage, 6 putative genes, one unknown gene and one known gene (PR-1a1 gene) specifically expressed in accession V1007108 and not in any other accession. In mature red stage, there were 17 genes differentially expressed; 10 putative and 7 unknown. Of the 17, the7 genes were specific to accession V1007108 (Table 5:3).

In accession Tindi 050580, 302 genes were differentially expressed. Of these, 245 were differentially expressed before fruiting, (3 known - fsm1, TCP23 and TCP4, 43 Uknown genes and 199 putative genes). Accession V1007108 showed a lot similarities with accession Tindi 050580 before fruiting stage, mature green and mature breaker stage. In mature green stage, Tindi 050580 had 12 differentially expressed genes (4 unknown and 8 putative) with only four genesspecific to this accession and not in any other accession. In mature breaker stage, only 2 putative genes were differntially expressed. In mature red stage, there were 43 differentially expressed genes, of which 16 were unkown/novel and 27 were putative genes.

Accession RV102112 had 599 differential expressed genes of the total 18,129. Of these, 570 were differentially expressed before fruiting with 19 known, 49 unknown/novel and 502 putative. 29 genes were differntially expressed at mature red stage, with 19 unkown, and 10 putative genes. However, there was no differential gene expression at mature green and breaker stages (Table 5:3)

236 genes were differentially expresed in accession Tindi 050589. Of these, 179 were differentially expressed before fruiting, (5 known, 31 unknown and 143 putative genes) with 15 genes being specific to accession Tindi 050589. At mature green stage, 36 genes were differentially expressed, 11 unknown and 25 putative genes. In mature green only, 15 genes were specifically expressed in this accession and not in any other accession. Accession V1005878 and Tindi 050580 are the most closely related to accession Tindi 050589, At mature red stage, 21 genes were differntially expressed (9 unknown and 12 putative genes). Only 5 genes (all

unknown/novel) were specifically differentially expressed at mature red stage of accession Tindi 050589 and not in any other accession.

A total of 481 of 18,129 genes were differentially expressed in accession V1006838. Of these 441 were differentially expressed before fruiting (58 unknown/novel, 483 putative and 13 were known genes). In mature green stage, 27 genes were differentially expressed with 1 known gene (Cevi57) which is a defense related gene ectopically expressed in viriod infected plants Gadea *et al*., 1996. Ten genes differentialy expressed at mature green stage were unknown with 16 being putative. At mature breaker stage, there were 7 differentially expressed genes (2 unknown and 5 putative genes; only one putative gene (LOC104645380) was specific to sample V1006838. Accession V1005878 and V1007108 were the most closely related at this stage. In mature red stage, 6 genes were differentially expressed, all unknown with only one uknown/ novel gene which was specific to accession V1006838.

In accession V1006842, 529 of the 18,129 genes expressed in African tomato were differentially expressed. Of these, 483 genes were differentially expressed before fruiting (56 unknown, 12 known and 415 putative genes). Only 8 genes expressed were specific to accession V1006842. In mature green, 35 genes were differentially expressed (1 unknown, 1 known genes (Cevi57) and 33 putative genes). In mature breaker, only 1 unknown gene was differentially expressed. In mature red 10 genes were differentially expressed; 3 unknown and 7 putative genes. Of these, only 2 genes were specific to accession V1006842.

In accession V1006826, 452 genes were differentailly expressed. Of these, 409 were differentailly expressed before fruiting (45 unknown, 8 known and 356 putative

gene). In mature green stage, 36 genes were differentially expressed (1 known gene-GA2OX4 a gene involved in regulation of fruit set in tomato (Servani *et al*., 2007) , 10 unknown and 25 putative genes. Only 2 genes were differentially expressed in mature breaker (1 unkown and 1 known gene). At mature red stage, only 5 genes were differentially expressed; 4 unknown and 1 putative (Table 21 and Appendix 3)

Accession V1005874 had 425 differentially expressed genes. Of these, 336 were differentially expressed before fruiting (27 unknown, 8 known and 301 putative). In mature green stage, 49 genes were differentaily expressed (11 unknown and 38 putative), 5 of these genes were specific to mature green stage of accession V1005874 and not in other accessions. 39 genes were differentially expressed in mature red stage (23 unknown, 15 putative and 1 known gene - CH114). CH114 gene is usually expressed in *Cladosporium fluvum* infected leaves (Danhashet *et al.*, 1993). 24 of the genes expressed in mature red stage were specific to accession V1005874 and not in other accession (Table 5:3 and Appendix III).

Accession V1030380 had 500 genes differentially expressed. Of these, 471 genes were differentaily expressed before fruiting (48 unkown, 15 known, 408 putative). Only 7 genes were specific to accession V1030380 before fruiting. In mature green stage, only 3 genes were differentially expressed (1 unknown and 2 putative – LOC104645893 and LOC101264905). In matue green stage, only one putative gene was specific to accession 16, i.e. LOC101264905. In mature breaker stage, 6 genes were differentially expressed; 3 unknown and 3 putative (LOC101255311, LOC101267363 and LOC101255952). Four of the six genes expressed in mature

breaker stage were specific to accession V1030380. In mature red stage, 20 genes were differentially expressed (6 unknown and 11 putative).

346 genes were differentailly expressed in accession V1006892. Of these 338 genes were expressed before fruiting (24 unknown/novel, 7 known and 307 were putative) only 5 putative genes were differentially expressed at mature green stage and three genes at mature red stage (2 unkown and 1 putative gene) (Table 5:3).

123 of 18129 genes expressed in accession V1035028 was differentially expressed before fruiting (16 unknown, 3 known- TCP1, TCP23, AGO10 and 89 putative genes). 13 of the 123 genes were specific to accession 18 before fruiting, Only four genes (3 unkown and 1 putative) were expressed at mature green stage. At mature breaker, 7 unknown genes were differentially expressed and all were specific to accession 18 while only 4 unknown genes were differentially expressed at mature red stage. Only 1 gene was secific to accession V1035028 (Table 5:3 and Appendix III).

In accession V1005875, 525 genes were differentially expressed. Of these, 501 were expressed before fruiting (40 unknown/novel, 12 known and 439 putative) 41 genes of the 525, were specific to accession V1005875 and not in any other accession. In mature green stage, 19 genes were differentially expressed (6 unknown and 13 putative) of which 6 genes were specific to accession V1005875. Only 3 genes were differentially expressed at mature breaker stage (2 putative and 1 unknown) with the unknown/novel gene specific to accession V1005875. In mature red stage, 2 genes were differentially expressed (1 unknown and 1 putative- LOC101266137).

XET4 (xyloglucan endotransglycosylase) a GA3 regulated xyloglucan gene though expressed in all the other accessions was not expressed in accession Tindi 050580. Another interesting observation was in accession 7 (V1007108) which expressed PR1a1, a gene expressed upon viroid infection according to Aoki *et al*., (2010); expressed only in breaker stage. Another gene expressed in the same accession was SICYP735A2 (expressed at mature green stage) which is normally induced by ethylene gas and is involved in resisstance to Fusarium wilt (Catanzariti *et al* 2015).

5.4 Discussion

The African tomato transcripts mapped to the reference sequence within a range of 67-98% sequence identities; indicating that the African tomato though it belong to the same species it has some similar genes with the reference genome it also have adapted to the harsh African conditions and could be having other genes not present in the sequence available at the genebank.

The percentage GC content of the assembled transcripts for the African tomato accessions unigenes peaked at 40-45%. This is similar to the GC content of Arabidopsis transcripts (42.3%; TAIR version 10 cDNA) (Chen *et al*., 2017). Similar results were also observed by Chen and colleagues in his work on comparative transcriptome analysis of *Solanum melongena.*

A total of 329M raw reads were obtained from the 17 African tomato accessions and after filtering, a total of 303.8M were left which were aligned to the tomato reference genome available at the Sol genomic network and NCBI databases. (*Solanum Lycopersicum* GCF-000188115.3_SL2.5.0). After alignment, 18,129 genes were

expressed, from these, 1,137 genes are known genes, 16,226 were putative genes while 700 genes were unknown.

The dendogram representation, scatter matric and principal component analysis used to check the quality and validate the biological replicates used in this study showed that the various fruit development stages clustered together, however the stage before fruiting clustered far from mature green, breaker and red stages (Figure 5:4 and 5:5). This is because before fruiting, leaf sample was used unlike the other three stages where the fruit was used. Meaning that there are specific genes that are either upregulated or downregulated in fruit formation, maturation and ripening that could be expressed or not expressed before fruiting or possible presence of novel genes in this accessions. This also indicated that the biological replicates were sampled at the same time, this was important to ensure there were no errors in the type of genes expressed in a specific fruit development stage gene as also suggested by Wolf, (2013).

Using the scatter matric analysis, differential gene expression between mature green and mature breaker were less dispersed compared to mature red and before fruiting, this implies that the genes up or down regulated between mature breaker and mature red were more similar than before green and mature red and mature breaker (Figure 5:6).

There was the presence of differentially expressed ethylene regulating genes which according to Gapper *et al.,* (2014), controls the ripening processes in tomato. These genes could be studied further to see whether their delay role in the ripening processes in this accessions.

There were genes related to cell wall formation and degradation expressed in most African tomato accessions. These genes have been studied previously and can either fasten or slow ripening process Gapper *et al.,* (2014) hence can be used to improve the shelf life. Genes that were up regulated in mature breaker and mature red stages could be involved in cell wall formation or degradation and fruiting (Giovannoni, 2011). This is also in line with what Gapper *et al*., (2014) reported that most of the genes responsible for cell wall degradation are switched on or expressed at the onset of ripening.

There were specific accessions that expressed genes that are related to Abscisic acid, ethylene biosynthesis, anthocyanins, and heat shock proteins among others. These genes have been found to cause resistance to both abiotic and biotic stresses, increase shelf life and improve colour (Wang *et al*., 2011).

Accession V1007108 ‘s breaker stage was the only one that expressed PR-1a1 gene which is known to confer resistance to powdery mildew and heat tolerance (Wang *et al*., 2013). This accession should be investigated further to find out to determine its resistance tolerance to powdery mildew.

5.5 Transcriptome analysis of the African eggplant accessions

Leaf samples were collected 3 weeks after germination from potted plants (Plate 5:2) , while the fruit samples were collected at the three fruiting stages; mature green, mature breaker and mature red (Plate 5:3).

5.5.1 Leaf samples used

Plate 5:2. Potted African eggplant (RV100332) Leaf sample at 3 weeks after planting

Table 5:4. African eggplants accessions used in this study

	CODE	ACCESSION NAME	NO	CODE	ACCESSION NAME
1	1aebf	RV100343	21	13aebf	RV100432
2	1bemg	RV100343	22	13bemg	RV100432
3	1aebrk	RV100343	23	13aebrk	RV100432
4	1ae mr	RV100343	24	13aemr	RV100432
5	3aebf	RV100201	25	14aebf	RV100246
6	3aemg	RV100201	26	14aemg	RV100246
7	3aebrk	RV100201	27	14aebrk	RV100246
8	3ae mr	RV100201	28	14aemr	RV100246
10	4aemg	RV100332	29	17aebf	RV100327
10	4ae mg	RV100332	30	17aemg	RV100327
11	4aebrk	RV100332	31	17aebrk	RV100327
12	4aemr	RV100332	32	17aemr	RV100327
13	6aebf	RV100445	33	23aebf	RV100330
14	6bemg	RV100445	34	23bemg	RV100330
15	6aebrk	RV100445	35	23aebrk	RV100330
16	6aemr	RV100445	36	23aemr	RV100330
17	10aebf	RV100265	37	28aebf	GBK50591
18	10bemg	RV100265	38	28bemg	GBK50591
19	10aebrk	RV100265	39	28aebrk	GBK50591
20	10aemr	RV100265	40	28aemr	GBK50591

Key: bf – Before fruiting, MG – mature green, mbrk- mature breaker,mr- mature red.

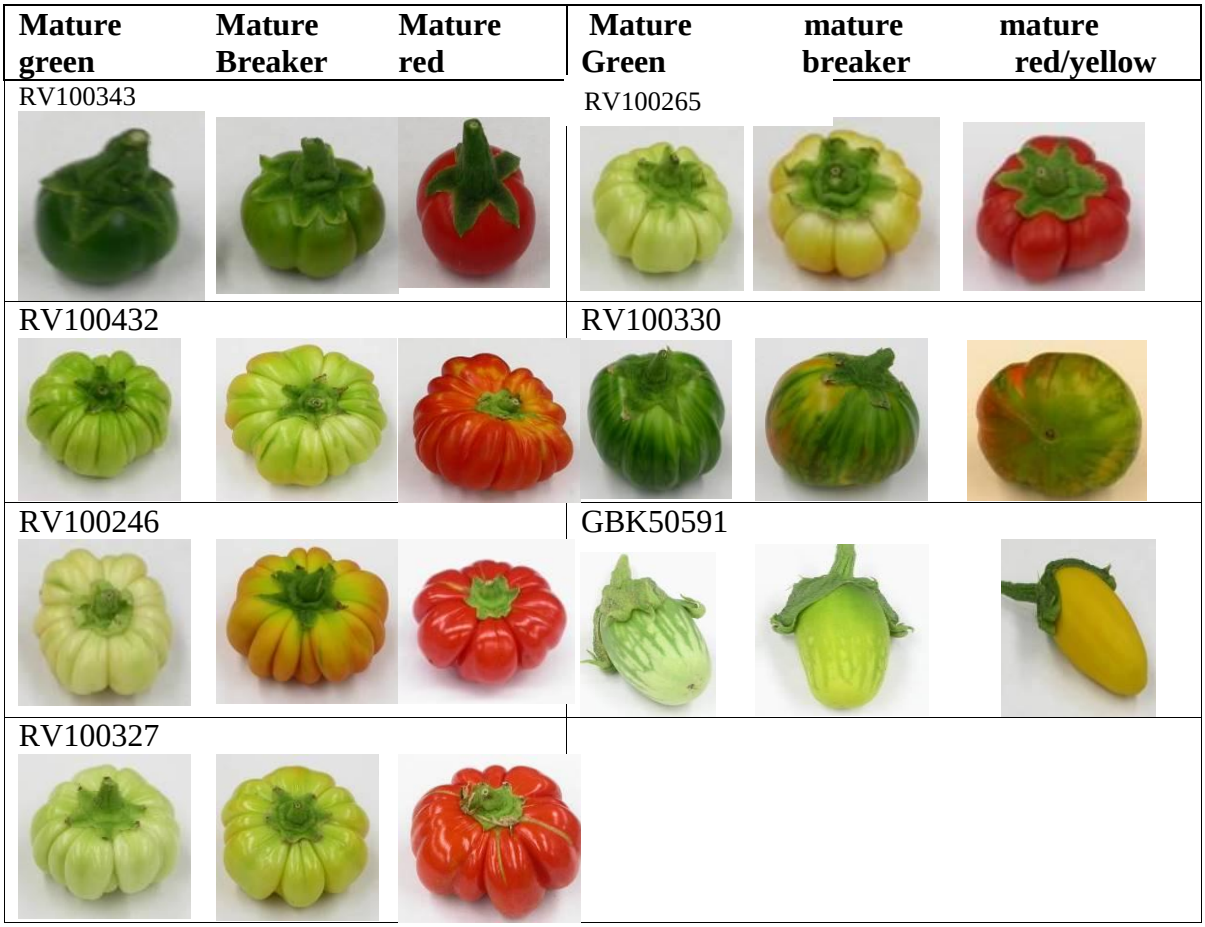


Plate 5:3. African eggplant fruits at the three fruiting stages

Key: (Mature green, mature breaker and mature red). The different stages are characterized by change in colour from green to red or yellow for different accessions.

5.5.2 Sequence quality control and validation

A total of 374 M raw reads were obtained from the illumina sequencer. FastQC software was used to check the quality of the sequence in terms of GC content, number of duplicate sequences, low quality reads, presence of adapters and barcode contaminations. NGSTool kit was used to filter the low quality sequences, remove adapters, duplicates and barcodes contaminants.

A read length for high quality was set at 70% while a cut off for quality score was set at 20 leaving a total of 329,018,858 filtered high quality reads with length of 101

base pairs. Percentage high quality reads ranged from 71.65% to 95.34% (Table 5:5 and Figure 5:7).

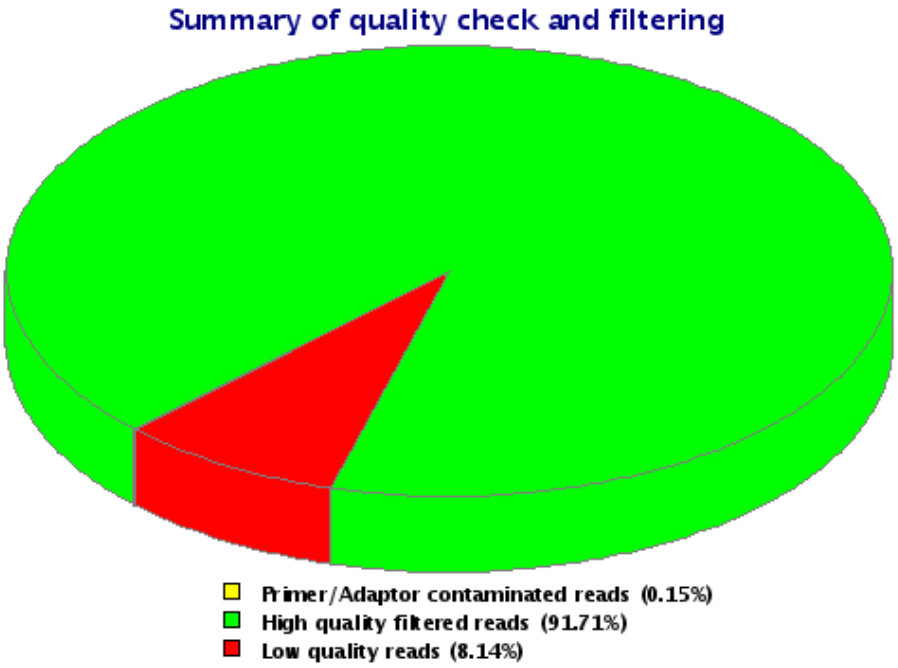


Figure 5:7. Distribution of reads based on quality. Low quality reads and primer/adaptor contaminated reads in RV100343 accession.

Table 5:5. Number of total raw reads, high quality reads after removal of low quality reads, high quality filtered reads after the primers and adapters have been removed from samples at four different stages in African Eggplant accessions

Accession	Total reads	HQ reads	HQ filtered reads	% HQ filtered reads	Accession	Total reads	HQ reads	HQ filtered reads	% HQ filtered reads
1 ae	2,839,753	2,608,641	2,604,283	91.71	6bemg	5,425,604	5,139,688	5,129,507	94.54
1aemg	6,850,010	6,476,591	6,466,870	94.41	6bebrk	4,159,086	3,835,714	3,822,264	91.90
1aebrk	2,687,189	2,311,845	2,310,930	86.00	6bemr	4,368,116	4,146,972	4,141,624	94.79
1aemr	4,211,797	3,881,917	3,841,930	91.22	10ae	4,456,758	4,075,011	4,020,153	90.20
1be	4,410,157	4,191,618	4,190,408	95.02	10aemg	4,994,270	4,757,162	4,735,280	94.81
1bemg	6,771,782	6,433,557	6,425,441	94.89	10bemg	6,922,743	6,586,354	6,555,934	94.70
1bebrk	3,935,823	3,364,426	3,362,126	85.42	10aebrk	4,073,855	3,506,087	3,503,911	86.01
1bemr	5,782,670	4,874,125	4,856,501	83.98	10aemr	5,016,561	3,872,728	3,838,892	76.52
3ae	6,067,518	5,536,485	5,533,037	91.19	10bemr	5,376,579	4,156,891	4,153,585	77.25
3aemg	4,953,523	4,672,544	4,649,918	93.87	10bebrk	4,106,410	3,511,082	3,505,016	85.35
3aebrk	2,267,545	1,874,085	1,873,073	82.60	13ae	4,372,434	3,990,916	3,987,744	91.20
3aemr	4,536,191	3,905,899	3,899,781	85.97	13aemg	5,988,081	5,112,406	5,107,006	85.29
3be	9,152,978	8,031,097	7,996,139	87.36	13aebrk	5,956,707	5,672,788	5,664,768	95.10
3bemg	7,609,093	7,198,397	7,165,187	94.17	13aemr	5,424,891	5,140,607	5,135,653	94.67
3bebrk	4,680,545	4,044,522	4,043,369	86.39	13be	4,747,599	4,241,897	4,232,807	89.16
3bemr	3,583,245	3,035,911	3,034,561	84.69	13bemg	3,810,841	3,257,816	3,252,827	85.36
4ae	6,491,692	5,801,146	5,794,093	89.25	13bebrk	3,693,645	2,992,536	2,986,112	80.84
4aemg	8,173,036	6,962,278	6,948,520	85.02	13emr	3,704,468	3,185,670	3,181,529	85.88
4aebrk	5,591,348	5,280,818	5,273,603	94.32	14ae	3,847,628	3,623,921	3,619,574	94.07
4aemr	4,848,105	4,471,344	4,464,834	92.09	14aemg	2,967,547	2,359,840	2,346,791	79.08
4be	5,765,895	4,934,042	4,887,970	84.77	14aebrk	3,611,595	2,692,885	2,587,966	71.65
4bemg	4,649,903	4,435,862	4,433,028	**95.34**	14aemr	4,797,533	3,958,108	3,924,579	81.80
4bebrk	4,442,362	4,209,053	4,205,003	94.66	14be	5,811,006	4,722,995	4,716,683	81.17

Accession	Total reads	HQ reads	HQ filtered reads	% HQ filtered reads	Accession	Total reads	HQ reads	HQ filtered reads	% HQ filtered reads
4bemr	4,572,652	4,193,016	4,173,328	91.27	14bemg	4,293,380	3,450,790	3,434,959	80.01
6ae	3,964,481	3,476,932	3,475,104	87.66	14bebrk	3,720,458	3,048,076	3,042,064	81.77
6aemg	5,502,895	5,226,295	5,221,933	94.89	14bemr	3,557,806	2,774,958	2,759,977	77.58
6aebrk	3,643,073	3,251,725	3,215,148	88.25	17ae	3,575,629	2,970,773	2,902,502	81.17
6aemr	3,683,152	3,146,298	3,144,764	85.38	17aemg	5,715,705	4,911,377	4,910,270	85.91
6be	5,786,325	5,501,400	5,495,729	94.98	17aebr	3,611,595	2,858,117	2,830,548	78.37
17aemr	3,640,714	2,917,131	2,906,557	79.83	23bebrk	5,041,633	4,339,438	4,335,697	86.00
17be	2,487,879	2,005,878	1,949,356	78.35	23bemr	4,682,234	4,441,579	4,419,706	94.39
17bemg	5,655,691	4,862,423	4,855,219	85.85	28ae	5,965,639	5,700,816	5,676,483	95.15
17bebrk	4,093,525	3,218,114	3,191,583	77.97	28aemg	4,036,107	3,428,979	3,407,975	84.44
17bemr	4,275,073	3,542,318	3,521,728	82.38	28aebr	5,456,267	4,636,931	4,635,515	84.96
23aemg	3,702,543	3,141,365	3,139,797	84.80	28aemr	4,238,713	3,584,245	3,575,240	84.35
23aebr	5,267,681	4,487,372	4,485,766	85.16	28be	6,790,071	6,200,499	6,189,728	91.16
23aemr	5,238,379	4,935,639	4,927,162	94.06	28bemg	3,868,185	3,281,956	3,280,085	84.80
23be	6,387,790	6,051,874	6,001,275	93.95	28bebrk	4,178,525	3,525,898	3,518,769	84.21
23bemg	4,482,385	3,809,989	3,807,546	84.94	28bemr	4,913,880	4,179,798	4,176,535	84.99
	373,966,182	**330,178,276**	**329,018,858**						

Key: Stages ae: before fruiting, mg- mature green, mbrk- mature breaker, mr- mature red

Raw reads means – sequences with all contaminants i.e. adaptors, chimeras, primers, ribosomal RNA, clean reads means – sequences without the contaminants but has rRNA, while high quality reads means, all contaminants have been removed and they have met the threshold set.

5.5.3. Dendogram representation and scatter matrix analysis for replicate validation

Similar to African tomato, the heatmap representation showed that each replicate of the different fruit development stage was true to type, since there were no introgressions between the replicates and stages. Similar to the case in tomato, the stage before fruiting however grouped far from the other three stages (Figure 5:8).

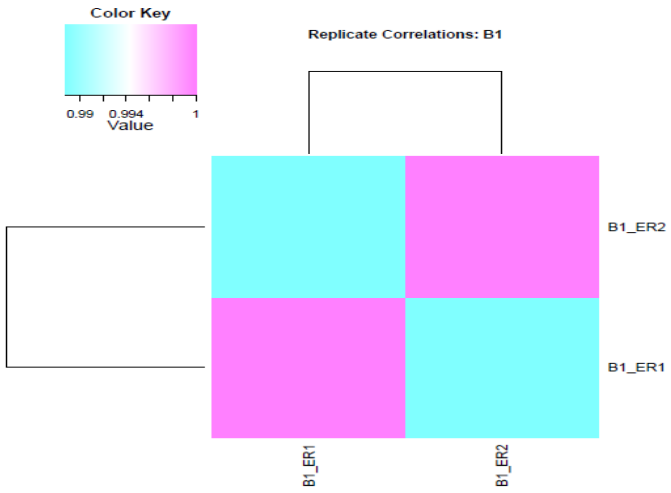


Figure 5:8. A heat map showing how the African eggplant's biological replicates used in this study related to each other.

5.5.4. Percentage GC Content

Similar to African tomato, the GC content percentage of the assembled African eggplant accessions unigenes and its distribution peaked between 40-45% (Figure 5:9).

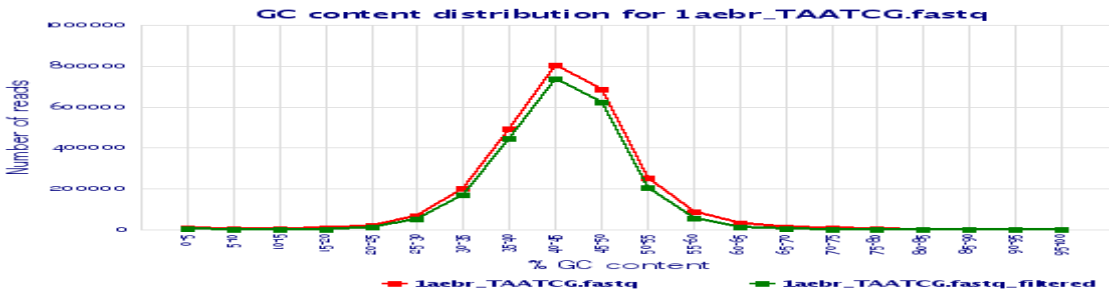


Figure 5:9. GC content concentration (%) in 1aebr in African eggplant

5.5.5. Transcript assembly

In African Eggplant, all the transcripts from the accessions were assembled using TRINITY, the longest consecious transcript was used as the reference gene for each African eggplant accession. A total of 173,194 genes were expressed from the African eggplant accessions.

5.5.6. Differential gene expression

In African eggplant, arbitrary numbers 4 and 6 were used, the gene clusters showed distinct patterns. For instance, there were genes expressed at the same level in all stages. However, some genes were expressed with increase and decrease at different stage (differential gene expression) (Figure 5:10 and 17, Appendix V and VI).

In sub - cluster 1, which had 316 transcripts, genes are expressed almost equally in all stages except for the stage before fruiting where there was down regulation of some genes (Figure 5:10).

In sub - cluster 2 (118 transcripts) and 3 (47 transcripts), some genes in mature red and mature breaker stage are upregulated than in other fruiting stages. In sub - cluster 4 (184 transcripts) there was an upregulation before fruiting (Figure 5:10).

In sub-cluster 5 (61 transcripts) there were some specific genes expressed in all the stages, but these genes were highly expressed before fruiting stage. In sub-cluster 6 (54 transcripts) some genes were expressed at the same level but were downregulated in mature red stage (Figure 5:10).

Similar to differential gene expression in African tomato, Scatter matrix analysis on the African eggplant showed that there was significant differential gene expression within the four fruiting stages examined. From the matrix (Figure 5:11), more genes were differentially expressed between the stage before fruiting and mature red stage, followed by mature breaker and mature green and vice versa.

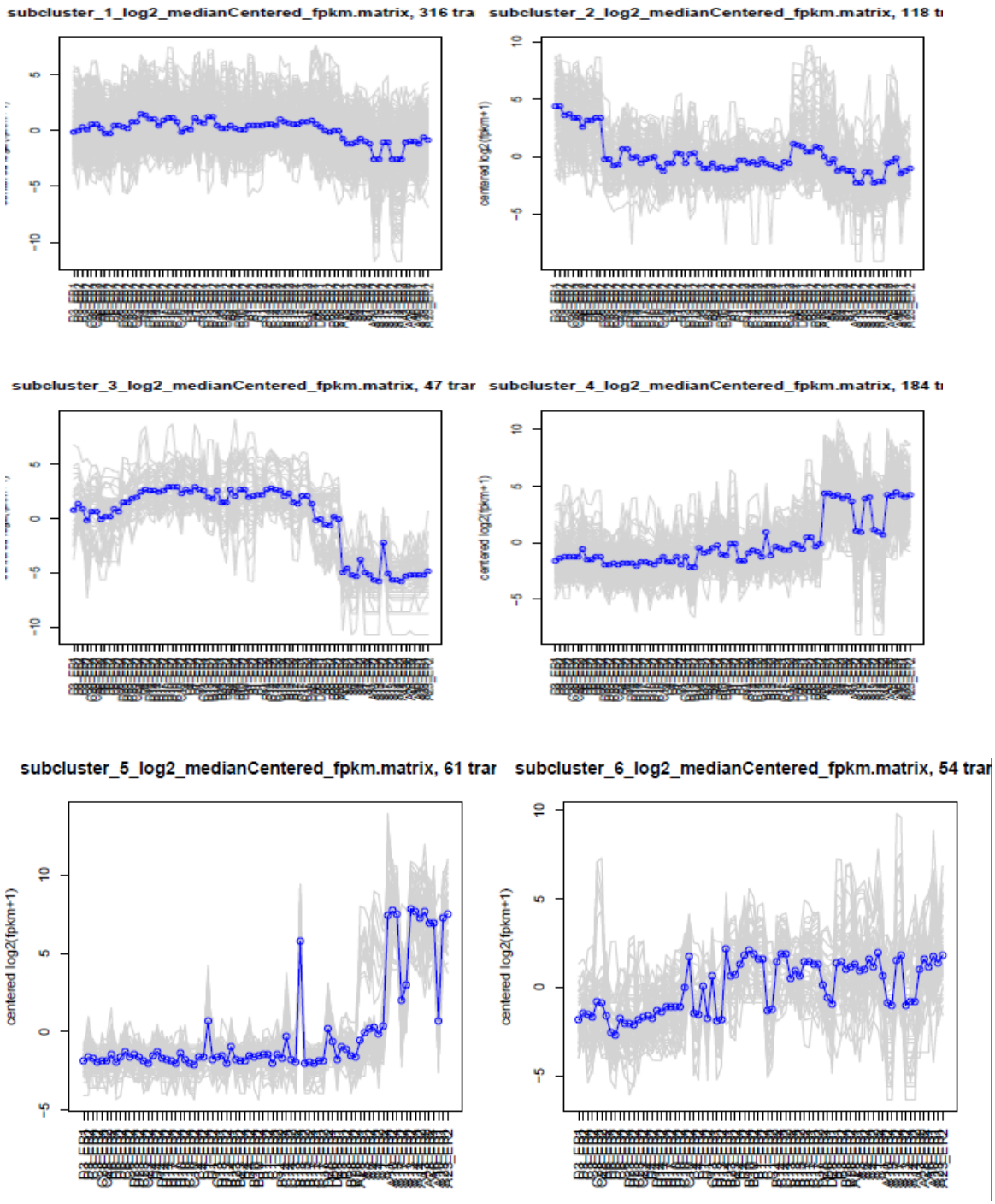


Figure 5:10. Showing arbitrary numbers 4 and 6 gene clusters in differential gene expression of the African eggplant accessions

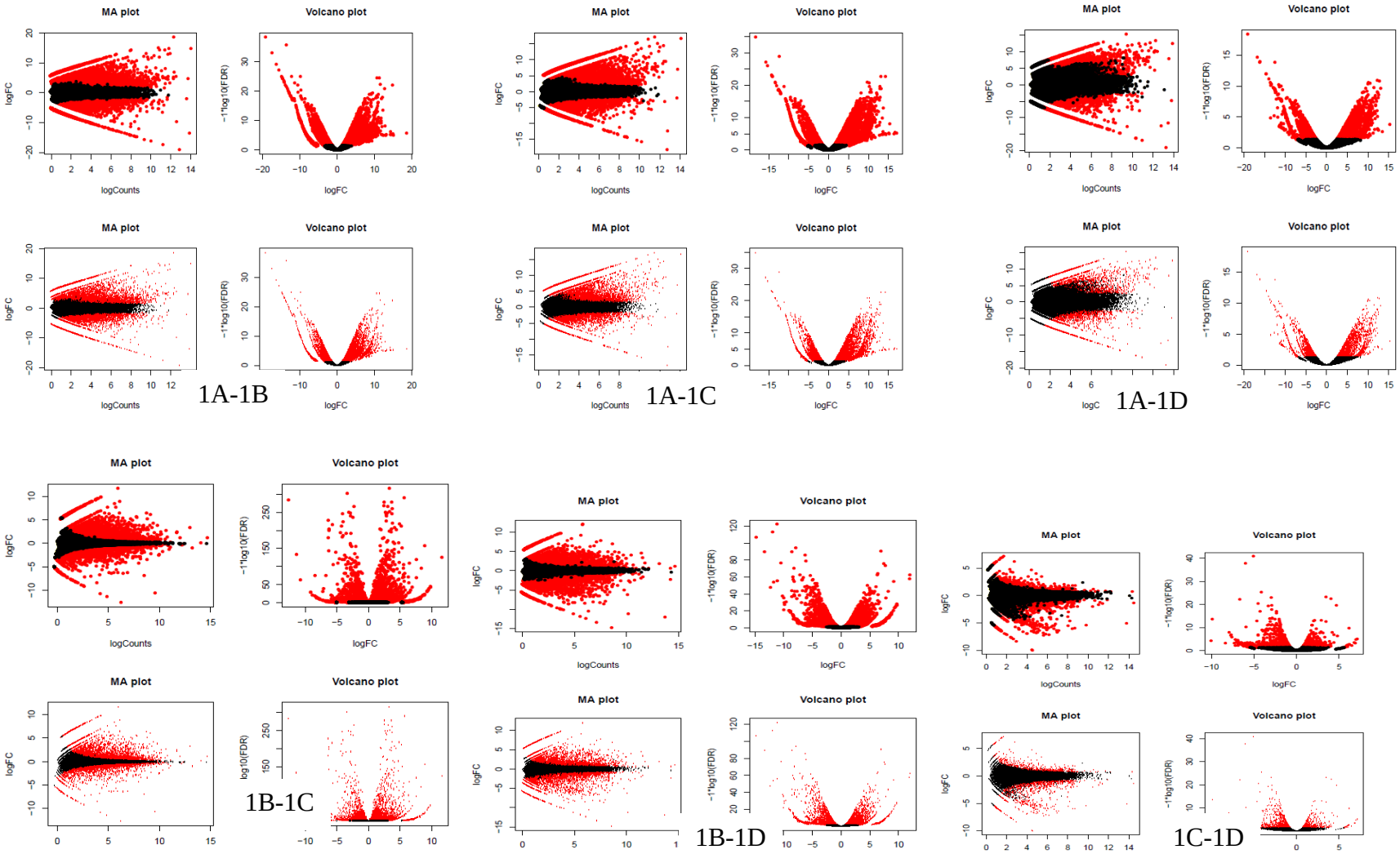


Figure 5:11. Scatter matrix analysis showing differential gene expression between different stages before fruiting, at mature green, mature breaker, and mature red stages in African eggplant accessions key A- Before fruiting, B- mature green, C- mature breaker, D- mature red.

5.6. Discussion

In this study, all the transcripts obtained from the 10 African eggplant accessions sequenced were assembled using the TRINITY software to create an inbuilt sequence that was used as the gene of reference, this was because the eggplant database at NCBI had very many gaps. When alignment to this inbuilt (BioSample accession SAMN13046517) reference genome was done, the African eggplant aligned at 71.65-95.34%.

The percentage GC content for the African eggplant accessions unigenes peaked at 40-45%. This is similar to the GC content of African tomato (40-45%) and Arabidopsis transcripts (42.3%; TAIR version 10 cDNA).

In African eggplants, 374M raw reads were obtained from the 10 African eggplant accessions used in this study. After filtering, 353.3M high quality reads were left. All these transcripts were assembled together obtaining an inbuilt sequence which was used as the reference transcript. Other scholars like Haas *et al* ., (2013) and Waiho *et al.,* (2017) did similar work by generating a reference transcript by combining all clean reads of the illumina sequencing data sets and selecting only one (the longest) to represent the assembled component from each sample, to prevent redundancy. After aligning to this inbuilt reference transcript, a total of 173,194 genes was obtained.

Heat map presentation, scatter matrix and principal component analysis used to check the quality control and validate the biological replicates used in this study showed that the various fruit development stages clustered together, meaning that

there was differential gene expression among the accessions and between the different stages in leaf and fruit development. There was an indication that the biological replicates, sampled at the same time, showing limited suppression of genes thus no errors in the type of genes expressed in a specific fruit development stage gene, this was earlier reported by Wolf, (2013).

In African eggplant, gene expression clusters, using arbitrary numbers 4 and 6 the gene clusters showed distinct patterns, there were genes expressed equally in all stages. However there were changes in pattern in different stages showing how the genes were differentially expressed at specific fruiting stages (Table 5:10).

Differential gene expression at 6 arbitrary gene clusters observed in African eggplants indicates that there were genes expressed in the same way (housekeeping genes) across all the accessions while others showed down or upregulated in specific fruiting stages (Table 5:10). This was an expected occurrence in fruiting plants since there are specific genes that are either downregulated, upregulated, switched on or off at every stage in fruit development and maturity (Gapper *et al*., 2014).

The high coverage (95.34%) of the eggplant gene set by unigenes in this study indicates the broad representation of these unigenes. The relatively low coverage (71.65) of unigenes by the draft genome and the gene set could be due to incompleteness of the genome assembly in this study, and the novel genes not predicted in the genome and highly divergent unigene sequences from the African eggplants. This may suggest that African eggplant unigene set can serve as a valuable complementary resource for eggplant genomics and functional genomics (Chen *et al*., 2017). The African eggplant showed differential gene expression for

genes associated with biotic stresses (Appendix V), abiotic stress (Appendix VI), among other genes. This indicates that African eggplant have many important genes that can be utilized for development of improved varieities

CHAPTER SIX

6.0 EVALUATION OF TRANSCRIPTOME DIVERSITY TO DEVELOP MOLECULAR MARKERS IN AFRICAN TOMATO AND EGGPLANT

Abstract

RNA-seq reveals information on sequence variation at individuals' genomes and transcriptomes allowing inferring patterns of allele-specific expression that can be relevant to environmental response and adaptation. In this study, 17 African tomatoes and 10 eggplant were sequenced using illumina platform to establish novel SNPs that could be used to estimate transcriptome diversity among the African tomatoes and eggplant, using the ratio of the SNPs contributing to the phenotypic and transcriptome variation. The identified candidate SNPs were used to predict the loci responsible for agronomically important traits, e.g. fruit size and shape and plant architecture. SNPs were analyzed in four different categories of samples (before fruiting, mature green, mature breaker and mature red) of the African tomato and eggplant processed by RNA seq technology. A total of 115,965 SNPs and 689 multiallelic SNPs were established in African tomato. In African eggplants, a total of 965,908 SNPs and 2,944 multiallelic SNPs were established. These SNPs were as a result of transitions and transversions in the ratio of 1.40 to 1.57 in African tomato and 1.63 to 1.64 in African eggplant. The SNPs in African tomato and eggplant accessions revealed that variation among the accessions was more dependent on geographical locations than morphological descriptors. Variation in the African tomato and eggplant accession transcriptome was mainly dependent on the fruit developmental stage other than the accession. This study revealed that the

African tomato and African eggplant have enormous transcriptome diversity which is mainly contributed by environmental variations. This is only a preliminary analysis and further in depth data analysis and wet lab experiments would help for further validation.

6.1 Introduction

In the AVRDC Genebanks, over 1,500 African tomato and eggplants lines have been deposited from African countries. Both morphological and DNA-based genetic variation has not been characterized , this has led to underutilization of these solanaceous crops in as food and even concerning their utilization in plant breeding programs.

Using current sequencing and high throughput genotyping technologies, it is now possible to analyse genome-wide genetic diversity and transcriptome diversity in the large number among and within tomato and eggplant accessions currently available.

Associations between genetic and phenotypic variations can be identified in the genetic resources by using morphological traits recorded in objective 1 (section 1.6.2) and their corresponding SNPs. Single nucleotide polymorphisms (SNPs) are useful for characterizing allelic variation, for genome-wide mapping, and as a tool for marker-assisted selection in order to understand relationship between genetic and phenotypic variations in crops (Sathya *et al.*, 2014).

Evaluation of transcriptome diversity in the current study will enable the utilization of the identified SNPs in Solanaceae improvement. This study will also help to relate the Phenotypic variation which are caused by differences in gene sequences or

patterns of gene expression (Carroll, 2008). Evaluation of trancriptome diversity in African tomato and African eggplant help in determining adaptive polyphenism which shows how variation in gene expression is translated into different phenotypes (Dhaygude et al., 2017).

6.2 Materials and methods

6.2.1 Single nucleotide Polymorphisms (SNPs) Mining

6.2.2 Source of NGS data

After mRNA extraction and mRNA library construction the multiplexed samples for both African tomato and African eggplant were separately sent for sequencing at the Biotechnology Resource Centre- BRC, in Cornel University. For sequencing through pooled sequence approach using the Illumina Technology. The mRNA/transcriptome sequence of the African tomato and eggplant was obtained from the sequencing facility as Single – end sequences (workflow in Figure 6:1).

6.2.3 Pre-processing

Pre-processing was executed using Next generation sequencing (NGSQC) toolkit to assess the quality of the data, examine the distribution of nucleotide, and percolate the low quality reads based on sequence constitution (Patel and Jain, 2012). The NGSTool kit was also used to separate the adapter/ barcode trimmed sequences and high quality reads filtered and the ribosomal contamination was filtered from the high quality RNA – seq reads using Ribopicker v 0.4.3 (Lee *et al*., 2014; Nielsen *et al*., 2011) (Figure 5:1).

6.2. Mapping for African tomato to the reference transcript

The high quality clean reads from the African tomato were aligned to the *Solanum lycopersicum* SL2.50 genome from ensemblgenomes using STARv 2.3.0 using default settings (Lee *et al*., 2014; Nielsen *et al*., 2011; Pabinger *et al*., 2013; Li and Durbin, 2010).

6.2.5 Alignment and mapping for African eggplant

For the African eggplant, *De novo* assembly using the high quality clean reads of all the 80 samples (10 samples each with 4 fruiting stages and 2 biological replicates each) were assembled into 1 inbuilt transcript using TRINITY. This inbuilt sequence served as the reference genome for the African eggplant.

STAR v 2.3.0 software was used to align both African tomato and African eggplant to their respective reference transcripts yielding to a SAM alignment file, to the SAM file, read groups were added, duplicate reads removed, reads sorted by coordinates using the Filter SAM program in SAMtools and the file converted to BAM file and indexed using Picard – tools v2.1.1 according to instructions by Lee *et al*., 2014 ; Nielsen *et al.*, 2011 and Li et al., 2009 (Figure 6:1).

6.2.6 Variant calling, annotation and visualization

The Genome analysis tool kit unified genotyper v2.8-1 (GATK) was used to call SNPs in all the samples, resulting in mult-sample variant call format (VCF) files (Lee *et al*., 2014; Nielsen *et al*., 2011). Default parameters were used for SNP calling in GATK with HaplotypeCaller set at phred-scaled confidence threshold of 20. Annotation and prediction of effects and variants on genes in the VCF file was done using snpEff and SNP phylogenetic tree constructed with SNPhyo (Lee *et al*.,

2014; Dewey, 2011 and Wang *et al*., 2010). The generated trees were visualized using Figtree (Figure 6:1).

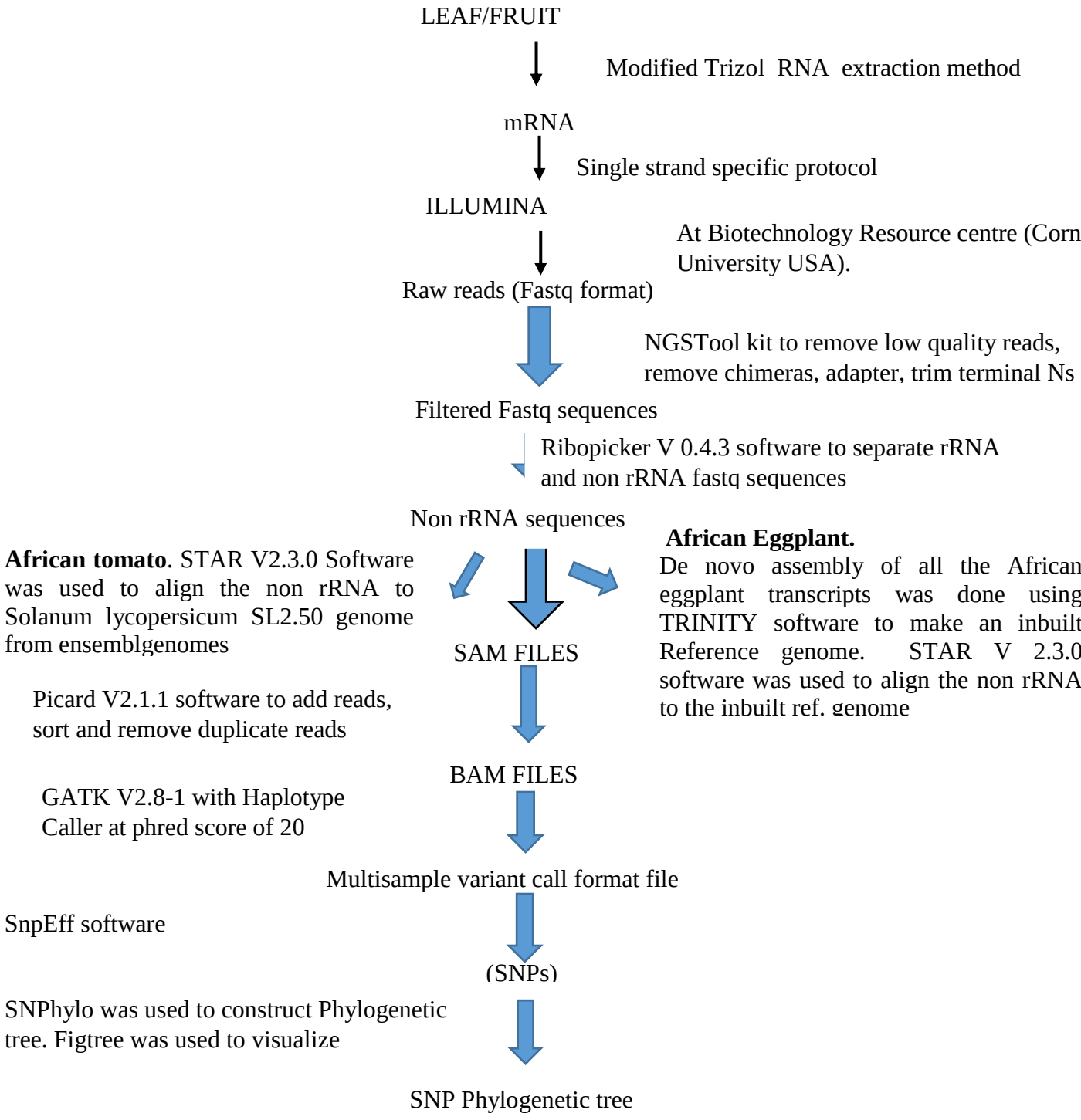


Figure 6:1. Steps involved in a workflow from mRNA extraction to SNP in African tomato and African eggplant accessions

6.3 Results

6.3.1 African tomato

A total of 115,965 SNPs and 689 Multiallelic SNPs were mined from all the 17 African tomato accessions used in this study (Table 3:5). The annotation was performed based on genomic location and the SNPs and were distributed in exonic and splicing region. The Significant variations was observed across all the chromosomes with SNPs being mined from all the twelve African tomato chromosomes as shown in Appendix VII.

Variations were observed as a result of transitions and transversions. The ratio between transitions: transversion ranged from 1.40 to 1.57 giving a difference of 0.1. (Appendix VII). Variations were also as a result of deletions and insertions. The number of insertions was triple the number of deletions. This was observed in all the accessions used in this study (Figure 6:1).

In the phylogenetic tree (Figure. 6:2) constructed using the various SNPs mined from African tomato accessions, the accessions grouped according to geographical locations, except for a few accessions. For instance, Accession RV102112 a pink and oval shaped fruit accession grouped together with V1030380 from Mauritius which has red and oval fruits. RV102112 also grouped closely to Tindi 050580, a Kenyan yellow and round fruit shaped accession (Figure 6:2).

V1007108, a red and oblong shaped fruit accession from South Africa grouped closely with Tindi 050580 (Yellow and round fruit) from Kenya and V1006838

a round and red fruit accession from Ethiopia (Figure 6:3). There seem to be population admixture between accessions from South Africa, Kenya and Ethiopia (Figure 6:3).

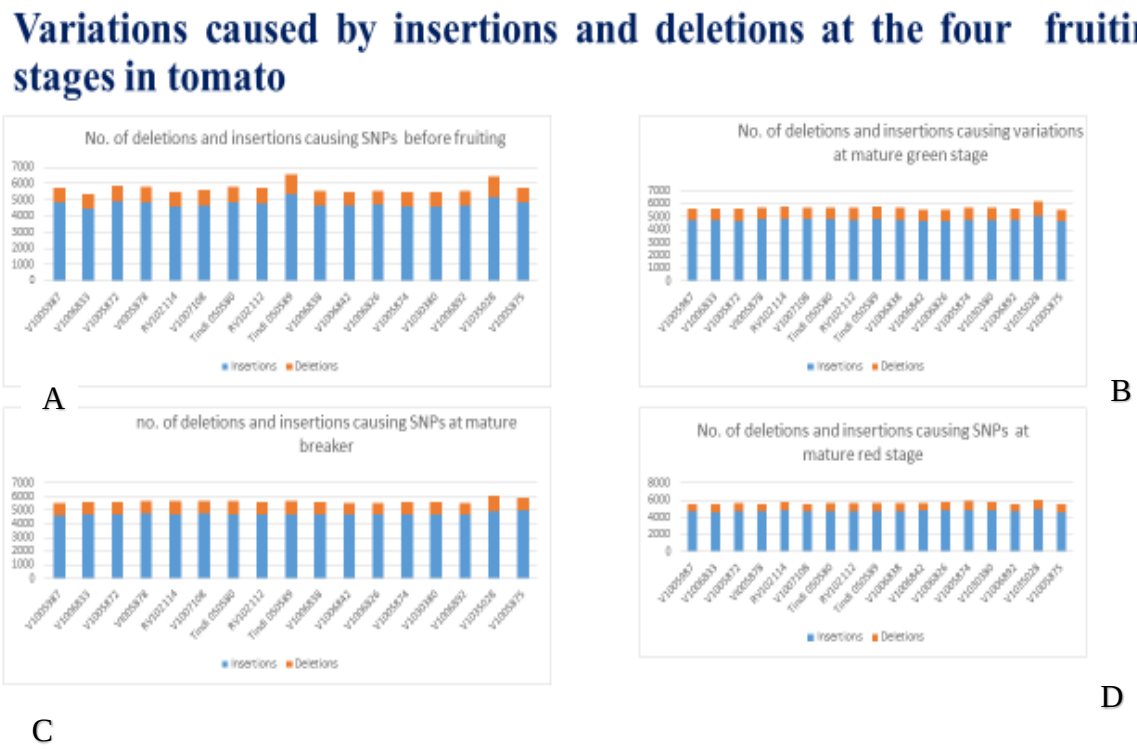


Figure 6:2. Deletions and insertions causing variations in the African tomato accessions A) Number of deletion and insertions observed before the fruiting stage, **B)** Number of deletions and insertions at the mature green stage, **C)** number of deletions and insertions at mature breaker stage and **D)** number of deletions and insertions at mature red stage.

Phylogenetic tree representation of the African tomato SNP diversity

Figure 6:2. Phylogenetic tree showing diversity in the 17 African tomato accessions as a result of variation in gene expressions at different fruiting stages.

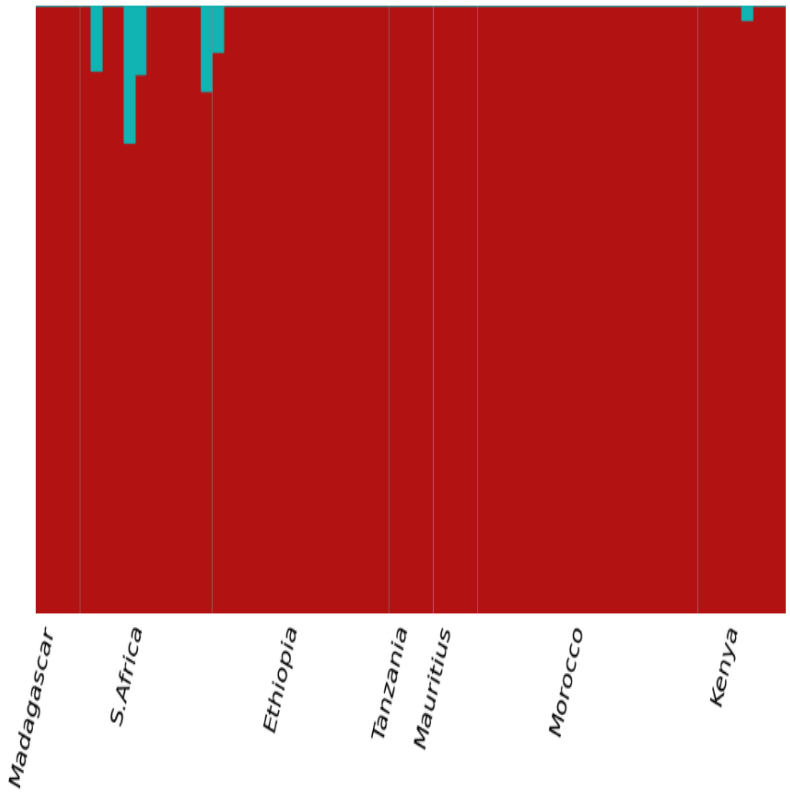


Figure 6:3. Population admixture among the 17 African tomato accessions. Blue colour represents the accessions admixture

6.3.2 SNPs obtained from the African eggplant

A Total of 965,908 SNPs and 2,944 multiallelic SNPs were mined within all the 10 African eggplants used in this study at all the fruiting stages and before fruiting stage at a depth of 1000. Variations were caused either by transitions or transversions (Ts/Tv) at a ratio of 1.63 and 1.64 giving a difference of 0.01 (Table 24).

Table. 6:1. Number of SNPs obtained from the African eggplant

Callset	SNPs			Multiallelic		Singletons(AC=1)	
	No.	Ts/tv	1st ALT	sites	SNPs	SNPs	Ts/tv
African eggplant	965,908	1.63	1.64	2,944	2,944	100%	1.63

Key: Ts- Transitions, Tv- Transversions, SNPs- Single nucleotide Polymorphism

6.4 Discussion

In this study, a considerable number of molecular markers in form of SNPs were obtained, for instance in African tomato, a total of 115,965 SNPs and 689 multiallelic SNPs were mined. This is a huge contribution to the study of the African *Solanaceae* crops especially in determining *Solanaceae* genetic diversity.

The SNPs observed in this study was as a result of transitions and transversion (Ti/Tv) in the ratio of 1.40 to 1.57 in African tomato yielding to a difference of 0.17 . This signifies that these SNPs were true nucleotide polymorphism. Previously, Ni *et al.,* (2012); Ding *et al.,* (2010) and Gopalakrishnan *et al.,* (2015), stated that Ti/Tv ratio for a random variation resulting from systematic errors in the sequencing technology, alignment artifacts and data processing failures should be close to 0.5.

The SNPs were mined across all the twelve chromosomes of the African tomato at varying numbers. Giovannoni, (2007) also observed similar results from his work on tomato fruit ripening with variation occurring in all the twelve chromosomes and that these variations were caused by either deletions or insertions (Appendix III).

SNPs diversity was associated with geographical locations unlike the morphological characterization which grouped the accessions according to fruit shape, colour and

size. SNPs diversity also revealed population admixture among specific accessions from Kenya, South Africa, Ethiopia, Morocco and Madagascar (Figure 6:3) This is in agreement with the findings of Wu *et al.,* 2015 and Hamilton *et al*., 2012 where environmental variables have an impact on the movement of gametes and individuals among natural populations hence affecting gene flow patterns. This may also lead to spatial and progressive dispersal of genetic variation and evolutionary advancement of regular populations.

There was significant variation within African tomato contributed by vegetative growth stages of the African tomato accession like plant height, leaf blade length, leaf blade width and fruit width. Substantial variation among the 17 African tomato accessions was observed in the reproductive stages i.e. fruit colour, fruit shape, fruit texture, leaf base and leaf lade colour. However, transcriptome SNP analysis revealed that the significant variation among these accessions was according to their geographical location indicating that morphological characterization of African tomato can only lay a foundation but it does not reveal genetic diversity. Transcriptome analysis goes beyond the phenotypic traits by showing which of the accessions from different geographical locations had been mixed.

This study revealed that environmental variables can have an impact on gene flow patterns, which may influence spatial and progressive dispersal of genetic variation and evolutionary advancement of regular populations. This study represents an important step forward in genomics, genetics, and for the breeding of cultivated tomato.

In African eggplants, a total of 965,908 SNPs and 2944 multiallelic SNPs were obtained. These SNPs observed in this study ranged between 1.63 to 1.64 and was as a result of transitions and transversion in the ratio in African eggplant, yielding to a difference of 0.01 in African eggplant. This signifies that these SNPs were true nucleotide polymorphism.

Significant variation among African eggplant was mainly contributed by vegetative and reproductive growth stages of the African eggplant accession like plant height, leaf blade length, leaf blade width and fruit width.

The construction of the African eggplant inbuilt reference genome in this study will be of great value in improving the current eggplant database (which has many errors) and future studies on the eggplant.

CHAPTER SEVEN

IDENTIFICATION OF GENES ASSOCIATED WITH DISEASE AND DROUGHT RESISTANCE, EFFECTIVE YIELD AND SHELF-LIFE QUALITY IN THE AFRICAN EGGPLANT AND AFRICAN TOMATO

Abstract

African tomato and eggplants are a resource for genes of interest. The high quality filtered sequences were blasted against different protein databases and annotated. Gene Ontology (GO) terms were assigned to African tomato and eggplant unigenes using the GOslim software into the different functional categories. Within the biological process category, cellular process, response to stress, biosynthetic process, nucleobase-containing compound metabolic process, and cellular component organization were among the most highly represented groups. Within the molecular function category, the top five most abundant groups were binding, nucleotide binding, hydrolase activity, catalytic activity and protein binding. Membrane, nucleus, plasma membrane, cytoplasm, and cytosol were the most represented groups within the cellular component category. GO annotations of the African tomato and eggplant revealed a large number of genes involved in important metabolic pathways such as degradation, utilization, assimilation and biological processes such as signal transduction, secondary metabolism and cell differentiation of secondary metabolites, folate and flavonoid biosynthesis. A large number of genes involved in responses to biotic and abiotic stresses were found to be differentially expressed in the African tomato and eggplants.

7.1 Introduction

7.1.1. Gene ontology

Gene ontology (GO) is a process involved in the assembly of the RNA polymerase complex at the promoter region of a DNA template resulting in synthesis of RNA from that promoter. Each GO has a name, a definition and an identification number (Gene ontology consortium, 2000).

Gene annotation is a process of categorizing gene products using Gene ontology grouping them into three main ancestral classes. Gene ontology interprets the functional consequences of polymorphism. It helps in generation of genetic resources for species of biological interest due to their evolutionary significance or economic importance (Gan *et al*,. 2011).

GO has 3 main domains namely Biological processes, Molecular function and Cellular components (Gan *et al*,. 2011). In Biological process, GO seeks to answer which processes a gene product is involved in, the molecular function GO is involved in finding out which molecular functions a gene product have while Cellular component ontology seeks to know where does a gene product act (Gene ontology consortium, 2000).

7.1.2 RNA-sequencing and gene annotation

Next Generation Sequencing (NGS) has changed the scope and scale of transcriptome analysis and gene expression studies. RNA-sequencing (RNA-seq) technology, apply the principles of NGS to the complementary DNAs (cDNAs) derived from transcript populations. RNA-seq is designed to detect both extreme upper and lower

limits of gene expression allowing for more accurate quantification of differential transcript expression as well as identification of low-abundance transcripts (Filichkin *et al*.,2010).

Mining RNA-seq data in search of transcription start site (TSS) variation is also improving gene structure annotation and alternative TSSs have been detected in approximately 10,000 loci through analyses of full-length *Arabidopsis* and rice cDNAs (Tanaka *et al.*, 2009).

RNA- seq analysis has helped in the interpretation of full-length transcript sequences, as has been demonstrated in a study where approximately 10% of the untranslated region (UTR) boundaries of rice genes could be extended (Lu *et al.*, 2010).

An ideal genome annotation identifies both genes that show invariant transcript sequences and those that exhibit alternative splicing, linking these events to specific spatial, temporal, developmental and/or environmental cues. (Filichkin *et al.*, 2010). Transcriptome analysis and annotation of different species has helped in interpreting the functional consequences of polymorphism hence leading to gene prediction results in a more reliable annotated genome (Gan *et al*,. 2011).

RNA-seq has been used to examine transcriptional dynamics during various aspects of plant growth and development. For instance, an analysis of the transcriptome of grape (*Vitis vinifera*) berries during three stages of development identified >6,500 genes that were expressed in a stage-specific manner (Zenoni *et al.,* 2010), thereby detecting of 210 and 97 genes that undergo alternative spicing in one or two stages, respectively.

Wang *et al.*, (2012) analyzed the transcriptome of radish (*Raphanus sativum*) roots at two developmental stages and found >21,000 genes to be differentially expressed, including genes strongly linking root development with starch and sucrose metabolism and with phenylpropanoid biosynthesis.

In addition to studies focusing on transcriptional changes during development, RNA-seq is a very effective strategy to study plant responses and adaptations to abiotic and biotic stresses. For example, by analysing RNA-seq data derived from sorghum (*Sorghumbicolor*) plants treated with abscisic acid (ABA) or polyethylene glycol, in conjunction with published transcriptome analysis for *Arabidopsis*, maize, and rice, Dugas *et al* (2011) discovered >50 previously unknown drought- responsive genes.

Similarly, RNA-seq was used to reveal massive changes in metabolism and cellular physiology of the green alga *Chlamydomonas reinhardtii* when the cells become deprived of sulfur, suggesting molecular mechanisms that are used to tolerate sulfur deprivation (González-Ballester *et al.*, 2010). Equivalent high resolution gene expression information has also resulted from studies of plant responses to pathogens and the complexities of the metabolic pathways associated with plant defence mechanisms.

The broad dynamic range of transcript level detection allowed by RNA-seq profiling, and particularly the detection of low-abundance transcripts, facilitates meaningful discrimination between different strengths of association in correlation analyses (Iancu *et al.*, 2012). Gene ontology enrichment analysis of RNA-seq data often illustrates the complexity of interacting pathways.

For example, in a study of abiotic stress responses in maize, transcripts associated with numerous GO classifications were affected by drought treatment, including the categories “carbohydrate metabolic process,” “response to oxidative stress” and “cell division,” among others (Kakumanu *et al.,* 2012).

Gene ontology using the GOslim classification method was used in this study to establish the genes associated with disease and drought resistance, effective yield and improved shelflife.

7.2. Materials and Methods

Plant_Goslim classification method was used to identify Gene ontology terms that apply only in plants. Swisprot, a curated database was used to identify and group the differential gene annotations. Other databases used to group the GO terms include KEGG, Signalp, Pfam, Sprot Top BLASTP hit, protein coordinates, protein-id, RNAMMERS (Figure 7:1).

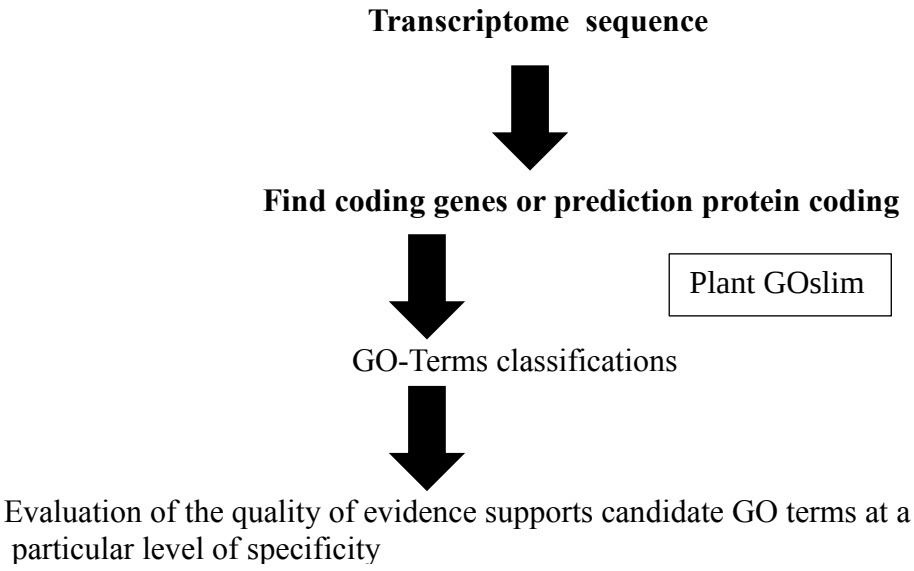


Figure 7:1. Workflow explaining the processes involved in gene ontology and classification in African tomato and African eggplant

7.3 Results

7.3.1 African eggplant Gene ontology

After the genes were grouped into the GO terms using the plant GOslim, to ensure that the grouping only used the plant databases (Figure 7:1). The expressed genes were grouped into the seventeen GO categories (Table 7:1) in the following percentages and multi-counts :- Biological process had 23.19 % with a total of 1,579,467 multi-counts, cellular components had 11.16% with a multicount of 760,197 while molecular function had 9.07% with a multicount of 617,935. Others include metabolic process with 10.6% with a multicount of 722,131. Cellular process had 8.93%, binding had 4.99%. The African eggplants showed response to stress at 0.62%, response to abiotic stimulus at 0.32%, response to biotic stimulus 0.25%, response to endogenous stimulus 0.23%, response to external stimulus 0.34%, response to extracellular stimulus 0.05%, ripening 0.01%, lipid metabolic process

0.33% cell wall 0.03% and reproduction 0.3% while other components had 29.9% (Table 7:1 and Figure 7:2).

Table 7:1. GO terms classification, their fractions in percentages and multi-counts for the African eggplant

GO terms classification	Fractions (%)	Multi counts
Biological process	23.19	1,579,467
Cellular components	11.16	760,197
Metabolic processes	10.6	722,131
Molecular function	9.07	617,935
Cellular process	8.93	608,220
Binding	4.99	728
Response to stress	0.62	42,179
Response to abiotic stimulus	0.32	21,477
Response to biotic stimulus	0.25	16,920
Response to endogenous stimulus	0.23	15,991
Response to external stimulus	0.34	22,928
Response to extracellular stimulus	0.05	3,352
Ripening	0.01	145
Lipid metabolic process	0.33	22,644
Cell wall	0.03	2,204
Reproduction	0.3	20,478
Others	29.9	2,354,989

Key: GO terms showed that Biological process had the highest percentage (23.19) followed by cellular component (1.116), metabolic process (10.6) and molecular function (9.07). African eggplant showed that African eggplant had genes which responds to stress, abiotic and abiotic stimulus, ripening lipid metabolic process and cell wall process at 0.62, 0.32, 0.25, 0.01, 0.33 and 0.03 respectively

From the heatmap (Figure 7:2) its clear expression of the gene associated with abiotic stress was different among the different accessions. However all the stages in accessions grouped together in case before fruiting, mature breaker and mature red stages were more closely related. Hydrogen peroxide was expressed in all accessions at varying expression levels except in accession RV100432 which did not have any expression. Accession RV100246 was the only one with differential expression of strictosidase gene. The gene was only expressed before fruiting, accession RV100445 at mature red stage had the highest salt tolerance gene expression while accession RV100246 had the highest expression of anthocyanin at mature green fruiting stage.

Accession RV100445 at mature red stage and accession RV100246 at mature green stage grouped differently from the other accessions. Grouping as exhibited by this heatmap was dependent on the fruiting stage and not on the accessions.

The heatmap showed how the genes of interest were differentially expressed among and within the African eggplant accessions. The fruiting stages determined which genes were expressed. For instance before fruiting stage for all accessions clustered together, also at mature green stage, a similar trend was observed. Mature breaker and red stages had similarities in the way the genes of interest were expressed (Figure 7:2).

From the heat map, accession RV100246 before fruiting had the highest strictosidase gene expressed accession RV100445 (mature red stage), had the highest levels of expression for genes against salt tolerance. mature red stage of accession RV100265 had the highest expression of genes resistant to drought. Accession RV100327 at breaker stage had the highest levels of expression of the histone related genes. Salicylic acid related genes were highest expressed before fruiting in accession RV100332.

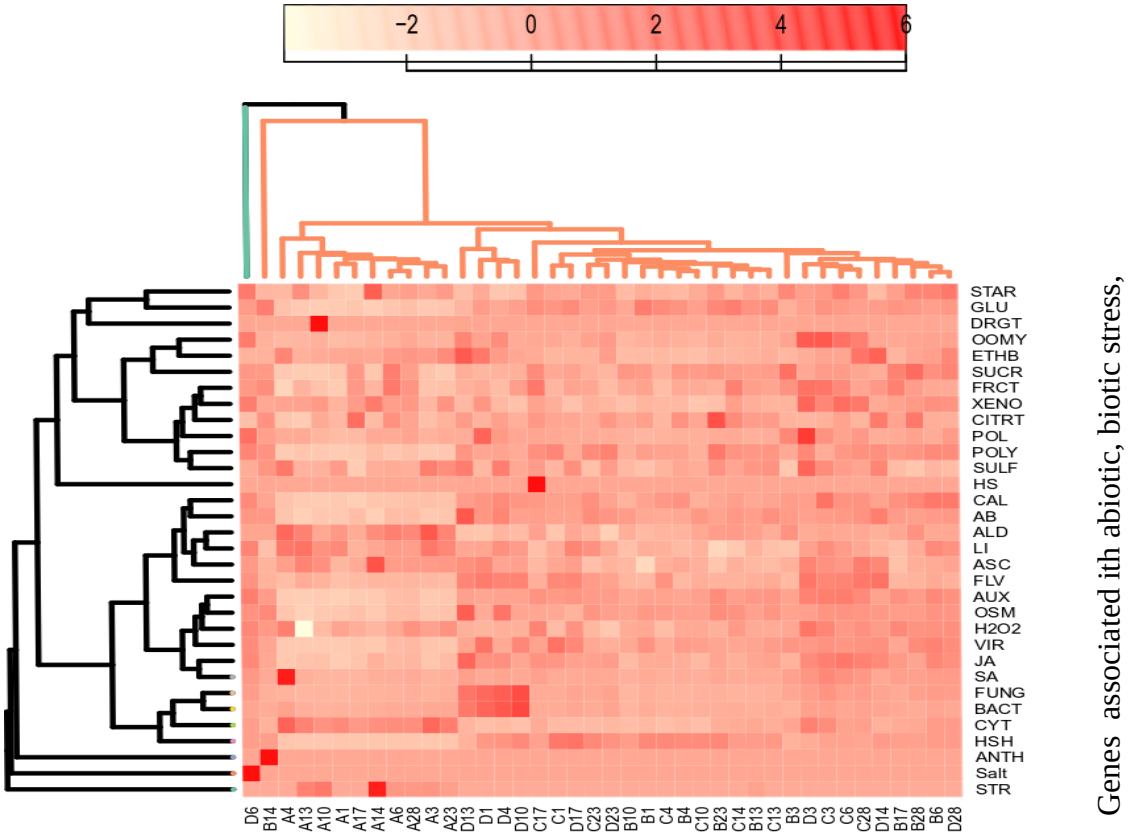


Figure 7:2. Heat map showing how the different genes associated with abiotic, biotic stress, improved yield and improved shelf life were expressed within and among the 10 African eggplant accessions

Key: 1-RV100343, 3-RV100201, 4-RV100332, 6-RV100445, 10-RV100265, 13-RV100432, 14-RV100246, 17-RV100327, 23-RV100330, 28-RV1001201, A- before fruiting , B- mature green, C- Mature breaker, D- Mature red. Genes associated with abiotic stress included from bottom: SALT, ANTH- Anthocyanin, CYT- Cytonin, BACT- Bacteria, FUNG- Fungi, SA- Salicyclic acid, JA- Jasmonic acid, VIR- Virus, H_2O_2- Hydrogen peroxide, OSM- Osmotic pressure, AUX- Auxin, FLV- Flavanoid, ASC- Ascorbic acid, Li, Lithium , ALD- , AB- Abscissic acid, CAAL- Calmodulin, HS-, SULF, Sulfur, POLY, POL, CITRT- Citric acid, FRCT- Fructose, SUCR- Sucrose, OOMY- Oomycetes, DRGT- Drought,

7.3.2 African tomato

From the transcripts , a total of 18,129 genes were detected from the 17 African tomato accessions. A total of 1,137 of these gene are known genes, 16,226 of these genes were predictive/ putative genes while 700 of the genes were unknown genes

The differentially expressed genes were subjected to the GO terms classification counter using the GO-slim classification method, GO terms were grouped according to the count occurrences in their respective ancestral classes (Table 7:2).

The differentially expressed genes had a total of 28,921 GO terms which were grouped into 127 of the "GO_slim" ancestral terms by single count. The differentially expressed genes were categorized into 25 gene ontologies with Biological process having highest percentage (21.64%), followed by metabolism (9.48%) (Table 7:2).

Table 7:2. African tomato gene ontology classification

Gene Ontology classification	Counts (single)	Fractions percentages)
Biological process	4,676	21.64%
Metabolism	2,049	9.48%
Molecular function	1,826	8.45%
Catalytic activity	1,138	5.27%
Cellular component	778	3.60%
Binding	366	1.69%
Response to stress	267	1.24%
Lipid metabolism	225	1.04%
Signal transduction	196	0.91%
Reproduction	183	0.85%
Response to external stimulus	139	0.64%
Carbohydrate metabolism	135	0.62%
Response to abiotic stimulus	123	0.57%
Response to biotic stimulus	83	0.38%
Response to endogenous stimulus	83	0.38%
Enzyme regulator activity	43	0.20%
Regulation of gene expression, epigenetic	33	0.15%
Lipid binding	23	0.11%
Viral life cycle	21	0.10%
Translation factor activity, nucleic acid binding	7	0.03%
Cell wall	6	0.03%
Actin binding	3	0.01%
Others	6,079	28.12%
Total	21,612	100.00%

Key: GO terms showed that Biological process had the highest percentage (21.64) followed by cellular component (4,676), metabolic process (9.48) and molecular function (8.45). African eggplant showed that African eggplant had genes which responds to stress, abiotic and biotic stimulus, lipid binding and cell wall process at 0.57, 0.38, 0.25,0.11 and 0.03 respectively

Using CateGOrizer which has an inbuilt REViGO software which summarizes and visualizes ling lists of Gene Ontology by removal of rendundant Go terms. The remaining GO terms are then visualized in Semantic similarity- based heatmaps (Figure 7:3).

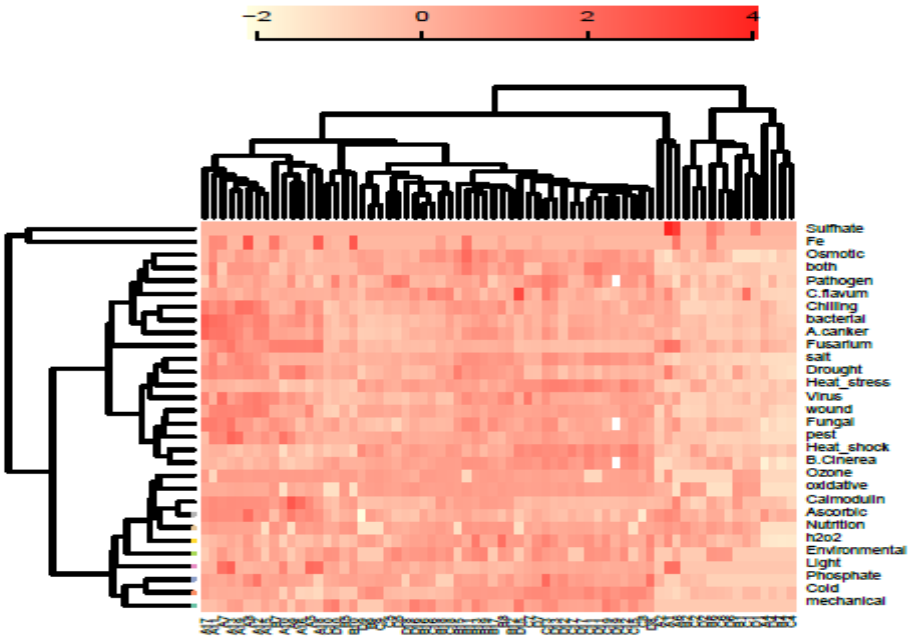


Figure 7:3. Heat map showing how the different genes of interest were expressed within and among the 10 African tomato accessions.

Key: 1-,V1005987 2-V1006833, 4-V1005872, 5-V1005878,6-RV102114, 7- V1007108, 8- Tindi 050580, 9- rv102112, 10-Tindi 050589, 11- V1006838,12- V1006842, 13- V1005874,16- V1030380. 17- V1006892, 18- V1035028, 19- V1005875 A- before fruiting , B- mature green, C- mature breaker, D- mature red. Genes of interest included the genes associated with:- mechanical, cold, phosphates, light, environmental, hydrogen peroxide, nutrition, ascorbic acid, calmodulin, oxidative , o-ozone, B. Cinerea, heat shock, pest , fungal ,wound, virus, heat stress, drought , salt, fusarium,A. canker, Bacterial , chilling, C. flavum, pathogens, osmotic pressure Iron, sulphate

The heatmap representation showed clearly that the differential expression depended on the fruiting stages, for instance accessions grouped together accornding to the stages for before the fruiting stage, mature green stage, mature breaker stage. The mature breaker and red clustered together meaning there was similar gene expression levels between this two stages (Figure 7:3).

Its clear from the heatmap that most of the accessions exhibited genes that confer to the biotic stresses, including fungal, viral and bacterial resistance as well as abiotic stresses at varying levels of expression (Figure 7:3).

7.4 Discussion

In this study, African eggplant, 2,276,995 GO terms were mapped to 103 of the "Plant_GOslim" ancestor terms by multi count. The genes were grouped into the 17 main GO categories (Table 7:1). Biological process had 23.19 % with a total of 1,579,467 multi-counts, cellular components had 11.16% with a multicount of 760,197 while molecular function had 9.07% with a multicount of 617,935. Others include metabolic process with 10.6% with a multicount of 722,131. Cellular process had 8.93%, binding had 4.99%.

The African eggplants showed that some accessions had genes associated with to both abiotic and biotic stresses. Among the abiotic stresses included heat shock stress, with accession RV100327 breaker stage having the highest expression levels. African eggplants had differencial gene expression for ascorbic acid, starch synthase, strictosidase and citrate synthase with Accession RV100246 had the highest gene counts (Appendix IV) for these abiotic stresses. Accession RV100265 had the highest gene count for drought resistance. Other abiotic stresses include response to sulfhur, calmodulin, osmotic stress, hydrogen peroxide, Jasmonic acid, salicylic acid, flavonoid and heat shock (Figure 7:2).

African eggplant accessions showed capacity to resist pest and diseases, for instance , accessions RV100343, RV100332, RV100265, RV100432 and RV100327 had the highest gene count for genes conferring resistance to fungal and bacterial attack. Other biotic stress included resistance to xenobacteria with all accessions having varying gene expressions (Figure 7:2) .

Other African eggplants responded to endogenous stimulus 0.23% (response to external stimulus 0.34% response to extracellular stimulus 0.05%) Some of the African eggplant had genes which are of importance to ripening 0.01%, lipid metabolic process 0.33% cell wall 0.03% and reproduction 0.3%. This is very significant especially in studies concerning shelf life improvement as well as in development of cultivars that are resistant to abiotic and biotic stresses (Table 7:1).

In African tomato, a total of 18,129 genes were expressed in the 17 African tomato accessions. Of these, 1,137 of these gene are known genes, 16,226 of these genes were putative genes while 700 of the genes were unknown genes (Table 7:1). The differentially expressed genes had a total of 28,921 GO terms which were grouped into 127 of the "GO_slim" ancestral terms by single count.

The three main ancestral classes included the biological processes with 40.30 % of the total value, molecular function had 13.72% while cellular component had 47.41% (Table 7:1).

Most of the accessions used in this study had differential gene expression for genes associated to respond to abiotic stress (0.57%) with a multicount of 123, biotic stress and endogenous stimuli at 0.38% and a count of 83 each, response to external stimulus at 0.64% with a count of 139. Some of the accessions showed viral life cycle at a count of 21 with a 0.1%. Others had significant cell wall and actin binding at 6 and 3 count at 0.03% and 0.01% respectively.

Accessions with differential gene expression in genes associated with response to

abiotic stress included accession V1005987, V1006833, V1005878, V1006892 and Tindi 050580. These accessions can be used further studies in resistant to harsh weather conditions or in improving cultivars which are higher yielding and are not resistant to abiotic stresses. Other abiotic stresses included response to iron, o-zone, calmodulin, chilling, ascorbic acid, light , salt, phosphate and heat stress (Figure 7:3).

The African tomato accessions showed significant resistance to biotic stress i.e. *Cladosporium flavum*, *A. canker*, Fusarium wilt, virus, *B.cinerea*, fungal infections included accessions V1005987, V1006833, V1005878, V1006892, Tindi 050580, V1007108, V1030380, V1006826 and V1005875 (Figure 7:3).

The heatmap representation showed that the differential gene expression depended on the fruiting stages, for instance accessions grouped together according to the stages for before the fruiting stage, mature green stage, mature breaker stage. The mature breaker and red clustered together meaning there was similar gene expression levels between this two stages (Figure 7:3). From the heatmap that most of the accessions exhibited genes that confer to the biotic stresses, including fungal, viral and bacterial resistance as well as abiotic stresses at varying levels of expression (Figure 7:3).

African tomato and African eggplant accessions differentially expressed genes associated to cell wall improvement, response to external stimuli, endogenous stimuli included at varying expression levels.

CHAPTER EIGHT

8.0 COMPARATIVE STUDY OF THE AFRICAN EGGPLANT AND TOMATO GENES

Abstract

The eggplant (*Solanum aethiopicum*.) and African tomato (*Solanum Lycopersicum*) are vegetable crop species belonging *Solanaceae* family. The two crops have a lot in common especially in morphological similarities; shape, colour and size of the fruit. However, the molecular similarities and or differences between the African tomato and eggplant accessions are not known or documented anywhere. The aim of this work was to compare the genes expressed in the two crops. Genes from the African tomato and African eggplants were compared with the protein Swissprot database to evaluate the extent of gene similarities between the African tomato and eggplant. A total of 349 and 353.3M high quality reads were generated from the African tomato and eggplants. From the African tomato, 18,129 unigenes were obtained after aligning the African tomato high quality reads to a reference genome (Solanum lycopersicum SL2.50) from ensemble genomes . The African eggplants transcripts were *de novo* assembled into 173,194 unigenes with a total length of 46.5 Mb. Resultant unigenes were compared against different Swissprot databases, nearly 60% of them were annotated and 50% could be assigned with Gene Ontology terms. A number of key metabolic pathways were predicted from the assembled African tomato and eggplants unigenes**.** Gene expression profiles similarities indicated that African tomato and eggplant are closely related in the way they respond to various biotic and abiotic stresses .

8.1 Introduction

To provide insights into relationship between African tomato and eggplant, the transcriptomic profiles of African tomato and eggplant, were analyzed using RNA-sequencing (RNA-seq) technology. To better understand the general molecular mechanism of African eggplant and tomato fruits with desirable agronomic trait in *Solanaceae* crops breeding, transcriptomic profiles were screened. The resultant unigenes and common Differentially expressed genes (DEGs) were used during the comparison process (Xu *et al*., 2015).

Comparative gene expression analysis have been done between eggplant lines with success (Chen *et al.,* 2017). RNA-Seq data was also used to compare gene expression profiles between cultivated and wild spinach (Xu *et al,*. 2015). RNA-Seq data was also used to compare two different *Salicina* cultivars by Lian *et al*., (2015) with success.

Other comparative gene expression analysis by differential clustering was done successfully by Ihmels *et al.,* (2005) on the application to the *Candida albicans* transcription program. Provenzano *et al*., (2016) successfully compared the two mouse models for autism using transcriptome data. Comparative gene expression analysis was also done by Hwang *et al*., (2015); Du *et al.,* (2015) in comparing the expression profiles of the liver, kidney and blood vessels during renal injury.

In the current study, the DEGs of interest expressed in the two crops were compared to evaluate their similarities. This is the first comparative transcriptomic analysis in African tomato and eggplant.

For annotation, the African tomato and African eggplant predicted genes, their protein sequences were compared to GenBank nr, the Arabidopsis protein and UniProt (Swiss-Prot and TrEMBL; http://www.uniprot.org/) databases using BLAST (parameter '-evalue 1e-4'), as well as the InterPro database using InterProScan54 (v5.10–50.0). GO annotations were obtained using GOslim based on the BLAST results against the GenBank nr database and results from the InterProScan analysis. Functional descriptions were assigned to the African tomato and African eggplants genes.

8.2 Methodology

8.2.1 Comparative analysis of the African eggplant and tomato

Comparative analysis was performed between African eggplant and African tomato using the gene based strategy using the high-throughput Illumina sequencing technology. Strand-specific RNA-Seq libraries were constructed and sequenced for a total 16 African tomato and 10 African eggplants. The high- quality illumina reads from the African tomato (Table 5:2) were aligned to the reference genome at the NCBI. While the African eggplant high-quality Illumina transcripts reads (Table 5:5) were *de novo* assembled into unique transcripts. The unique transcripts were then extensively evaluated and annotated. Single nucleotide polymorphisms (SNPs) and differentially expressed genes among the 16 African tomato and 10 African eggplants were identified.

Comparative analysis was also done against *Solanum tuberosum* and *Arabidpsis thaliana*. Reciprocal best hit relationships were identified between the predicted genes in the inbuilt African eggplant reference genome and those of the African tomato using the method described by Fukoka *et al.*, (2012) who established a

routine procedure for SNP marker development, genotyping ans comparative studies. In his study he studied the effect of introducing mismatched nucleotides into allele-specific primers using alternative fluorescent dyes, and varying the DNA polymerase species on SNP discrimination in more more than 100 known SNP loci in a solanaceous crop, eggplant and tomato. Similarities searches were carried out with the Smith- waterman algorithm of the SSEARCH program. Our transcriptome data provide a valuable resource for future functional studies, further utilization and marker assisted breeding in *Solanaceae* crops.

8.3 Results

8.3.1 Comparative analysis of the African eggplant, tomato, potato and Arabidopsis genes

To find out the orthologous relationships between the African eggplant predicted genes and corresponding tomato, potato and Arabidopsis genes, Genes associated with Biotic stress , abiotic stress, improved cell wall and improved yield were seen differentially epressed in the two cops under study (Table 8:1, 8:2 and 8:3). Similarities between the nucleotide sequences were investigated. Of all the genes expressed in African eggplant, 39 genes had similarity with Arabidopsis at a 100% similarity, 429, 1989 and 4158 genes were between 90 to 99% , 80 to 89 % and 70 to 79% similarity respectively. Whereas 6,999 genes in African eggplants were between 80-89% similar to African tomato and 6,915 genes from African eggplant were 80-89% similar to *Solanum Tuberosum* (Table 8:4).

Table 8:1: Comparison of the Biotic stress related genes expressed by African tomato, African eggplants, potaotoes and Arabidopsis

Gene	QTL	A. tomato	potato	A. eggplant	A. thaliana
Phytoene Synthase	Biotic stress	2	1	0	1
Allene oxide synthase	Biotic stress	3	0	18	1
Abscissic acid	Biotic and abiotic stress	246	23	131	13
STG1	Biotic (Fusarium wilt)	1	1	0	2
Annexin	Biotic (fungal)	7	5	0	4
RRN4	Biotic	0	1	742	0
MAPK	Biotic, lycopene, abiotic	38	39	61	9
GME-GDP Mannose Epimerase	Biotic , cell wall	3	0	0	4
Subtilisin –like serine protease	Biotic	46	45	0	5
Aquaporin		18	24	0	0
SAMT	Biotic & cell wall	1	0	0	0
Terpene synthase	Biotic	6	0	70	4
phytopthora	Biotic	1	0	0	1
Thiamine biosynthetic pathway	Biotic	1	0	20	0
SAHH	Biotic and abiotic	1	0	0	1
NAC-NOR	Biotic (botrytis Cinerea)	1	0	0	0
Gamma glutamylhydrolase	Biotic (folate biosynthesis)	1	0	23	1
GRS	Biotic (disease resistance)	7	0	0	19
MPKA1	Biotic (microbes and insects)	1	0	0	0
MKK2	Biotic(resistance to Xanthomonas ampostris vesicatoria)	0	0	0	1
CNX61	Abiotic (heat shock)	5	3	0	7
BSL	Biotic (phytopthora infestans)	3	3	0	6
Beclin (atgc)	biotic and abiotic	5	0	0	9
ETR5	Biotic stress	1	0		0
Twi1	Biotic (responds to wound nd pathogen	4	8	22	0
CIP3	Biotic (mediated Rapomycin)	1	0	103	0
ASC	Biotic (alternaria stem Canker resistance)	1	0	0	0
RPL3	Biotic (trichothecare resistance)	2	0	0	3
Histone1	Biotic (response to bacterial infection)	1107	0	720	252
AGO4A	Biotic (defense against virus)	11	10	142	2
FK	Biotic (resistance to Begomoviruses)	15	19	37	13
RAN2	Biotic (suppresses schizosacchromyces)	2	2	28,955	830
Chitinase	Biotic (Chadosporin flavum)	21	22	56	9
WRKY3	Biotic and abiotic	1	0	0	3

Gene	QTL	A. tomato	potato	A. eggplant	A. thaliana
MEK1	Pto- mediated defense	1	1	0	2
Ve2	Biotic (Verticullum wilt)	1	0	58	1
Wfi	Biotic (whitefly induced)	1	0	0	0
mlo	Biotic	9	7	0	6
Ao2 (Aldehyde oxidase)	Biotic (ABA Biosynthesis)	2	0	65	2
CYP707	Biotic (overexpression reduces ABA)	1	2	0	3
Gene	QTL	A. tomato	potato	A. eggplant	A. thaliana
Srf	Biotic (induced by CMV)	0	0	0	3
ASC1	Biotic (Alternaria stem canker resistance)	1	0	0	0
SUT	Biotic (Brassinosteroid biosynthesis)	1	1	338	1
TCP20	Biotic (Pest resistance)	1	2	0	0
Subtilish	Biotic (Insect resistance)	5	0	0	0
		1586	219	31561	1218

Key: Table showing the genes expressed in the African eggplant and tomato with their biotic stress with their respective QTLs

Table 8:2: Comparison of differentially expressed genes of abiotic stress related genes

Gene	Environmental QTL	Solanum Lycopersicum	Solanum tuberosum	Solanum aethiopicum	Thaliana arabidopsis
Rab1a		1	0	0	0
Jasmonic acid	Responds to wounds	6	1	518	0
Salicylic acid		3	6	482	0
Starch synthase		8	2	10	1
Heatshock 70 (Hsp70)	Abiotic stress	4	2	26	2
GLox1		6	0	0	6
Citrate synthase		5	3	16	1
Strictosidase synthase		4	5	0	2
xyloglucan	Abiotic stress	21	24	110	14
LE16	Abiotic and biotic	1	0	2170	1
Auxin response factor	Abiotic	19	20	0	0
TSW12	Abiotic(drought stress)	1	0	0	0
Phytoene desaturase	Abiotic and flavor	1	0	0	0
Ascorbate peroxidase	Heat shock factor	8	8	0	0
Gene	Environmental QTL	Solanum	Solanum	Solanum	Thaliana

		Lycopersicum	tuberosum	aethiopicum	arabidopsis
Dehydroascorbate reductase	Abiotic and cellwall	5	4	9	0
MYC	Abiotic	5	5	207	6
Aquaporin		18	24	0	0
EREB	Cellwall, abiotic stress	1	0	13	0
DREB	Stress	2	0	0	3
Phosphosulfolactate synthase (PSR)	Abiotic	1	0	0	0
PR-5	Abiotic and Biotic	0	0	0	3
ARG1 (Arginase)	Response to Wound	1	0	2	1
SOS1	Abiotic stress (salt tolerance)	12	8	205	6
CZFP	Abiotic	1	0	0	0
SAHH	Biotic and abiotic	1	0	0	1
DHAR2	Abiotic	3	1	0	1
Gene	Environmental QTL	Solanum Lycopersicum	Solanum tuberosum	Solanum aethiopicum	Thaliana arabidopsis
HSC 70	Abiotic	1	0	27	4
GMP	Abiotic (salt stress)	2	2	60	3
Catalase Isoenzyme	Abiotic	5	2	20	0
SRG	Abiotic (salt tolerance)	10	12	0	0
Acid invertase (wiv)	Abiotic stress	1	0	0	0
UCP(Putative uncoupling protein	Abiotic	1	1	0	3
RS1	Abiotic (salicylic acid and heat)	0	0	35	9
COP	Abiotic(oxidative stresses)	4	2	666	88
UBC1	Abiotic stress	11	9	68	4
Alcohol dehydrogenase	Abiotic (alkaline stress)	3	0	63	0
HSP90-1	Abiotic stress	6	0	0	0
CNX61	Abiotic (heat shock)	5	3	0	7
Beclin (atgc)	biotic and abiotic	5	0	0	9
glyoxalase	Abiotic (enhances glycolysis in salt stressed plants)	8	0	4	1
BTF3	Abiotic (salt tolerance)	0	0	0	1
Pp2c	Abiotic(abscisic signalling)	1	0	0	0
Gene	**Environmental QTL**	**Solanum**	**Solanum**	**Solanum**	**Thaliana**

		Lycopersicum	tuberosum	aethiopicum	arabidopsis
RPL3	Biotic (trichothecare resistance)	2	0	0	3
GAL	Abiotic (nutritional and environmental stress)	324	61	665	0
PSBP	Herbivore resistance	8	4	0	1
GPX	Abiotic (salt stress)	3	0	0	3
CBF	Abiotic (response to cold)	1	3	0	1
TIL	Abioti (temperature induced)	52	52	191	15
Er SHSP	Abitic (small heat shock)	1	0	0	0
SQD1	Abiotic	2	2	0	1
GRX	Abiotic (oxidative drought and salt stress)	1	2	0	4
OPr2 OPr3	Jasmonic acid	1	0	0	12
		596	268	5567	217

Key: Table showing the genes expressed in the African eggplant and tomato with their abiotic stress respective QTLs

Table 8:3. Similarities in genes related to shelf life and aroma

Gene	Environmental QTL	A. tomato	potato	A. eggplant	A. thaliana
βgalactosidase	Shelflife	16	14	72	11
Mannose Epimerase	Biotic , cellwall	3	0	0	4
EREB	Cellwall, abiotic stress	1	0	13	0
SAMT	Biotic & cellwall	1	0	0	0
INT4	Shelf life	2	0	0	1
DHS	Shelf life (delayed fruit softening)	13	0	0	11
		36	14	85	27

Key: Table showing the genes expressed in the African eggplant and tomato with their shelf life and aroma respective QTLs

Table 8:4 . Similarity of African eggplant gene transcripts against the African tomato, potato and Arabidopsis.

Species	**Percentage similarity**			
	100%	**90-99%**	**80-89%**	**70-79%**
Arabidopsis	39	429	1,989	4,158
African tomato	0	0	6,999	0
Solanum Tuberosum	0	0	6,915	0

Key: Table showing percentage similarity between African tomato , African eggplant and Arabidopsis

Genes conferring the phenotypes of interest (biotic, abiotic stress, high yield, and better shelf life) to this study were compared between the African tomato, eggplant, Arabidopsis and potato. A total of 1586, 219, 31261,and 1218 genes responding to biotic stress phenotype were expressed in African tomato, potato, African eggplant and Arabidopsis respectively (Table 8:2).

A total of 598, 269, 5567 and 218 genes conferring to different abiotic stress were expressed in the African tomato, potato, African eggplant and Arabidopsis respectively (Table 8:2). In genes that can affect or are related to improved shelf life of the fruits, a total of 36, 14, 85 and 27 genes were expressed in African tomato, potato, African eggplant and Arabidopsis (Table 8:5).

Table 8.5. Number of important genes shared between African tomato, eggplant, potato and Arabidopsis

QTL	African tomato	African eggplant	Potato	Arabidopsis
Biotic factors	1,586	31,561	219	1,218
Abiotic factors	598	5,567	269	218
Shelf life	36	85	14	27
Aroma	254	43	251	56

Key: Tabulation of genes shared among the african tomato and eggplant accessions

8.4. Discussion

Using the known genes for the phenotypes of interest, the three Solanaceous crops and *Arabidopsis thaliana* were found to share important genes conferring to biotic disease resistance, drought resistance, fruit quality (Table 8:1, 8:2 and 8:3) among other important aspects and could be used as a feasible starting point for chromosomal walking experiments to isolate their orthologs in African eggplants and tomato**.**

African eggplant, African tomato, potato and Arabidopsis share key genes which are known for tolerance of abiotic stress (salt tolerance, drought resistance, increased temperature Table 8:3) and biotic stresses including resistance to fusarium wilt, Alternaria stem canker, verticulum wilt, *botrytis Cinerea, Xanthomonas ampostris vesicatoria, trichthecare schizosaccharomyces*, yellow curl virus- begomoviruses, whiteflies, cucumber mosaic virus (Table 8:1). Of the four crops compared, African eggplant and African tomato emerged more closely related in terms of the genes expressed**.**

In African eggplants, genes conferring to Beta carotene were overexpressed in accessions RV100445 and RV10050591. Lycopene was significantly upregulated in the mature red fruits whereas chlorophylls was overexpressed before fruiting and in mature green stage. Accessions RV100438, RV1050591, RV100332 and RV100259 had a high ascorbic acid gene expressions as compared to other accessions (Appendix VI).

There was significant differential expression in ascorbic acid binding genes in the mature red stages as compared to mature breaker and mature green in both African eggplant and tomato. In the African eggplant fruits, the levels of ascorbic acid RPKM increased as the fruit ripened while mature red fruits had higher levels as compared to the mature green stages.

However there was high expression of ascorbic acid before fruiting as compared to the fruiting stages. This is in agreement with several studies which have reported that, ascorbic acid content significantly increase during fruit maturation and ripening (Davey *et al.*, (2007); Gautier *et al.*, (2008). Accessions with high RPKM in terms of ascorbic acid expression included RV100343, RV100199 and RV1050591 (Appendix VI).

When compared with the African eggplant, the cherry type tomato accessions (with smaller fruits) had generally higher ascorbic acid contents. The breaker stage of some cherry type accessions such as V1005875, V1050580, V1050589, V1006838 and V1002114 had higher expressions of ascorbic acid as compared to the mature red stage. Gallie, (2013) in his findings reported that ascorbic acid levels in tomato varieties may vary with cultivars (Appendix VI).

In both African tomato and eggplant, ripening was dependent on accessions, for instance, RV100343, RV100199, RV100265 and GBK50591 African eggplant accessions and V1007108, V1002112, V1005874, V1030380, V1006892, Tindi 050589, Tindi 050580 and V1035028 African tomatoes ripened earlier than the other accessions. These accessions had the ascorbic acid genes overexpressed as compared to the other accessions. They could be selected as having better shelf life because ascorbic acid is known to prevent oxidative damage during ripening (Davey *et al*., 2007), (Appendix 4 and Appendix VI).

The yellow-fruited cherry type accessions Tindi 050580 and Tindi 050589 reported higher ascorbic acid gene expressions in the breaker stage as compared to the mature red fruits (Appendix IV). This results are in agreement with the findings of Mibei *et al* (2017), and Tembe *et al.* (2017), that cherry tomatoes had higher ascorbic acid concentrations as compared to the big accessions. Likewise, these results are in agreement with those of Adalid *et al.* (2010) and Vinkovic-Vrcek *et al.* (2011) who concluded that smaller fruits (cherry type tomatoes) have generally higher vitamin C content. Therefore, the increase in ascorbic acid concentration in the African eggplant and tomato leaves and fruits with growth and ripening stages demonstrates the nutritional importance of these crops. From the current study, the expression of ascorbic acid genes and drought resistance seemed to vary with accessions (Appendix IV and VI).

RV100343, RV100199, RV100265 and RV1050591 African eggplant accessions and V1007108, V102112, V1005874, V1030380, V1006892 and V1035028 African tomato accessions at mature red stage were found to be important in accumulating

ascorbic acid. These accessions overexpressed genes related to drought as compared to the other accessions (Appendix IV and VI).

There was significant differential β- carotene and lycopene gene expression between the fruiting stages. For instance, High β-carotene RPKMs were observed in breaker stages of Tindi 050580 and Tindi 050589 accessions. However, accessions V1006826, V1035028, V1005875, V1030380 and VI005878 had very low β-carotene gene expression. Findings in this study are consistent with the findings of previous studies which have found that cherry tomatoes are richest in β- carotene (Adalid *et al.*, 2010) and that β- carotene was high in the breaker stage as compared to the mature red stage (Appendix IV).

Although the lycopene differential gene expression was dependent on the ripening stage and accession, accessions with high lycopene differential gene expression included V1005875, V1007108, V1006892, V1006826, V1006838, V1030380, V1002112, V1002114, V1006828, V1035028 and V1005872 accessions. This also meant that accessions with the red and pink phenotype had higher lycopene expression as compared to the yellow and orange fruits in both African tomato and eggplants (Appendix IV and VI).

In African eggplant fruits, RV100343 had higher carotenoids expressions as compared to the other accessions. African eggplant accessions with lower carotenoid expression at breaker stage included RV100265, RV100432, RV100246 and RV100327. This may be attributed to the colour of the fruits since they were white in the breakers stage (Appendix VI).

In African tomato, V1006838, V1007108, V1006828 and V1002114 accessions had highest carotenoid expressions whereas Tindi 050580 and Tindi 050589 accessions had lower carotenoid RPKMs. Oval shaped accessions V1006838, V1006828 and V1002114 had significantly DEG in lycopene genes expression as compared to the cherry-type accessions, this in agreement with findings of Muratore *et al.* (2005) which reported that oval shaped tomatoes have the highest lycopene content when compared to cherry and round tomatoes (Appendix IV).

Accessions with higher expression of genes related or genes expressed due to drought included RV100343, V1006838, V1007108, V1006828 and V1002114 . These accessions had higher expressions of GABA which according to Sarvajeet and Narendra, (2010) is expressed as a response to drought stress. Accessions RV100273, RV100432, RV100327 and RV100445 had high Proline genes expressed . Other studies by Schafleitner *et al.*, 2007 showed that accumulation of proline in potato susceptible to abiotic stress.

Therefore RV100343, RV100199 and RV100265 African eggplant accessions and V1030380, V1006892 and V1035028 African tomato accessions may be perceived to be more tolerant to drought resistant as compared to the others. These accessions can be utilized for eggplant/tomato improved programs targeting nutritional quality and development of drought resistant plants (Appendix IV and VI).

Most of the African tomatoes had the 20ox-1 which is associated with have resistance to bacterial wilt (Appendix 3), which is known to reduce production and cause tremedous losses (Azzi *et al.*, (2015), Wang *et al.,(* 2012).

Most accessions expressed SGTI-1 and Itpg1 at differential genes expression levels (Appendix III). This gene is known to exhibit resistance to fusarium wilt disease Cantanzariti *et al* (2015), Nowicki *et al*,. (2013).

Most of the accessions expressed CIPK6 gene (Appendix III) which is known to be induced upon virus inducation, oomyceted and nematodes infestation. For instance, VI007108 Accession had very high levels of expression of this gene.

All accessions except 1 and 9 expressed CEL8 and SAR2 at different levels (Appendix III). These genes are known for resistance against Root Knot disease caused by nematodes. Zhang *et al.,* (2014). Azzi *et al.,* (2015).

African eggplants with the waterlogging tolerance (Appendix V) can be used as rootstocks for grafting tomato,this may improve waterlogging tolerance in tomato hybrids. Similar work was done successfully by Bahadur *et al*., (2015) using eggplants and tomato.

Some accessions had the NAC-NOR gene (Appendix V) at differential levels among all the accessions, this gene is known to be expressed in plants susceptible to the necrotrophic pathogen *Botrytis Cinerea* (Aoki *et al* 2010). This work confirms that African eggplants and tomatoes are potentially important crops in adapting to drought stress effect and disease resistance.

CHAPTER NINE

9.0 GENERAL DISCUSSION, CONCLUSION AND RECOMMENDATIONS

9.1 General discussion

The findings of this study show that African tomato and eggplants have significant phenotypic variations among the accessions at $P<0.05$. This variations is mainly contributed by plant growth habits and fruit morphology.

Comparisons done on morphologically between African tomato and eggplant clearly reveal close relationship between the two crops. This is in agreement with Romano *et al.*, (2014) and Zhou *et al*, (2009) who evaluated effects of grafting on tomato and eggplant.The African tomato and African eggplants are closely related being from the same family and hence can easily be grafted.

SNPs in African tomato and eggplant accessions revealed that variation among the accessions was more dependent on geographical locations and fruit development stage. This study revealed that environmental variables (geographical locations) can have an impact on gene flow patterns, which may influence spatial and progressive dispersal of genetic variation (Figure 6:3).

African tomato and African eggplant accessions shared genes which are known for abiotic stress (salt tolerance, drought resistance, increased temperature) biotic stresses including resistance to fusarium wilt, Alternaria stem canker, verticulum wilt, *botrytis Cinerea, Xanthomonas ampostris vesicatoria, trichthecare schizosaccharomyces*, yellow curl virus- begomoviruses, whiteflies, cucumber

mosaic virus indiating that these two crops shared an evolutionary pattern (Table 8:1 and 8:2).

African eggplant can been grafted onto African tomato rootstock to increase resistance to Verticulum wilt disease which is caused by *Verticulum dahlia*. This is because some of the African tomato and eggplants had genes which confer resistance to *V. dahlia*. According to *Zhou et al* ., (2009) tomatoes are better rootstock as compares to eggplants and production of wild eggplant rootstock seedlings can be very challenging as a result of low germination rate of seed, poor emergence, and slow early growth. In his work eggplant grafted onto tomato rootstock exhibited markedly higher disease resistance than non-grafted eggplant when challenged with *V. dahlia.*

Most of the African tomato and African eggplants accessions had genes that confer biotic stress resistance ie 20ox-1- resistance to bacterial wilt (Azzi *et al*., (2015), Wang *et al*., 2012), SGTI-1, Itpg1,-resistance to fusarium wilt disease (Cantanzariti *et al*., (2015) CIPK6 gene- induced upon virus inducation, oomyceted and nematodes infestation (Cantanzariti *et al.,* 2015), CEL8, SAR2, APS1- resistance against Root Knot disease caused by nematodes (Aoki *et al*., (2010); Zhang *et al*., (20140; Azzi *et al., (*2015), NAC-NOR- resistance to necrotrophic pathogen *Botrytis Cinerea* Aoki *et al.,* (2010), RPL3 gene- involved in trichothecare resistance (Aoki *et al* ., 2010), ASC-1 gene- confers resistance to Alternaria stem canker (Aoki *et al* 2010), HI gene- involved in altering host transcriptional responses to bacterial infection (Pham *et al* ., 2012), FK gene- Confers reisitance to Leaf curl virus and other begomoviruses, (Yang *et al.,* 2014a) among others.

African eggplants that had water-logging tolerance can be used as rootstocks for grafting tomato, this may improve water logging tolerance in tomato grafts. Similar work was done successfully by Bahadur *et al.,* (2015) who observed that grafting onto eggplant rootstock 'IC-111056' and 'IC354557' improved waterlogging tolerance in tomato scion 'Arka Rakshak' and 'Arka Samra eggplants rootstock.

Most of the accessions had genes which are for abiotic stress resistance or tolerance. For instance, Gamma aminobutyric acid (GABA) gene associated with drought stress; Sarvajeet and Narendra, (2010), Wiv-1 gene (gene regulated by jasmonate and salicyclic), RSI-1 gene (responds to Salicyclic acid and heat accumulation; Aoki *et al*., (2010), COP1 gene (involved in abiotic and oxidative stress; Mohan *et al* ., (2016), HSP17.4 gene (Small heat shock protein; Giorno *et al.,* (2010), GLX1 gene (enhances glycolysis in salt stressed plants; Aoki *et al*., (2010), BTF3 gene (involved in salt tolerance; Arafet *et al*., (2011), NAM 1 gene (involved in early salt stress response in tomato root; Aoki *et al*., (2010), among other genes.

9.2 General Conclusion

1. There was significant morphological variation within African tomato and within African eggplant contributed by vegetative growth stages like plant height, leaf blade length, leaf blade width and fruit width, length, fruit colour, fruit shape, texture and plant growth habit.

2. Both African tomato and African eggplant have genes associated with that can be used to improve the cultivated varieties.

3.The African tomato and eggplant are notably diverse genetically. Main source of their transcriptomic diversity is the stage of sampling rather than the accessions. The variations being caused by deletions and insertions. Other sources of variation include environment where they were grown and possible population admixture was observed among the accessions.

4. Transcriptome analysis is able to reveal genes that are being actively expressed in specific tissue and species of interest. The inbuilt African eggplant reference genome will greatly improve the current eggplant database.

5. This study reveals that RNAseq is a powerful tool for SNP mining, transcriptome analysis, differential gene expression and comparison studies between African tomato and eggplant.

9.3 Recommendations

1. RV100343, RV100199 and RV100265 African eggplant accessions and V1030380, V1006892, V1007108 and V1035028 African tomato accessions should be investigated further for their preferred traits for improvement of the currently cultivated varieties (Table 7:2 and Appendix III and Appendix V).

2. Unknown genes expressed in African tomato should be considered for more investigation (Table 5:2).

3. The inbuilt African eggplant reference genome may be utilized more to improve the current eggplant database

REFERENCES

Abukutsa-Onyango, M., (2003). Unexploited potential of indigenous African vegetables in Western Kenya. *Maseno Journal of Education Arts and Science*, 4(1): 103–122.

Adeniji, O., and Agatha, A., (2012). Farmer's Knowledge of Horticultural Traits and Participatory Selection of African Eggplant Varieties (*Solanum aethiopicum*) in Tanzania. *Tropicultura*, 30(3): 185–191.

Ahmad, S., Tabasum, M., Swaraj B., Mohit, K., Aubid, H., Zahoor, A., Nazia, A., Anil, K., and Nasheeman, A. (2015). Comprehensive transcriptome analysis of Crocus sativus for discovery and expression of genes involved in apocarotenoid biosynthesis. *BMC Genomics*, 16:698 . DOI 10.1186/s12864-015-1894-5

Aoki, K., Kentaro, Y., Ayako, S., Shingo, K., Nozomu, S., Kunihiro, S., Atsushi, K., Tatsuya, S., Taneaki, T., Manabu, W., Kazuhide, O., Maiko, T., Takanori, N., Tadasu, S., Yuji, K., Naoki, Y., Hideki, T., Yuichiro, W., Mayumi, E., Motoichiro, K., Yuki, I., Mari, K., Sumire, F., Akiko, O., Tsutomu, A., Yuko, S., Katsumi, Y., Shinobu, S., Toshikazu O., Hiroshi E. and Daisuke S. (2010). Large-scale analysis of full-length cDNAs from the tomato (*Solanum lycopersicum*) cultivar Micro-Tom, a reference system for the Solanaceae genomics. *BMC Genomics* 201011:210 https://doi.org/10.1186/1471-2164-11-210

Areshchenkova, T., Ganal. (2002). Comparative analysis of polymorphism and chromosomal location of tomato microsatellite markers isolated from different sources. *Theor. Appl. Genet*. 104: 229-235 Crossref, Medline, ISI, Google Scholar.

Azzi, L., Deluche, C., Gevaudant, F., Frangne, N., Delmas, F., Hernould, M. and Chevalier, C. (2015). Fruit growth-related genes in tomato. J. Exp. Bot. 66, 1075–1086.

Bahadur, A., Wilkinson, E, Tanner T. (2015) Measuring resilience: an analytical review. ODI Working Paper. Overseas Development institute, London

Barakat, A., Denis, S., Yi, Z., Chris, S., Kathleen, B., William, A., Nicholas, W., Ron, S. and John, E. (2009). Comparison of the transcriptomes of American chestnut (*Castanea dentata*) and Chinese chestnut (*Castanea mollissima*) in response to the chestnut blight infection BMC Plant Biology20099:51 https://doi.org/10.1186/1471-2229-9-51

Bolger, A, Lohse, M., Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014 Aug 1;30 (15) :2114-20. Doi: 10.1093/bioinformatics/btu170. Epub 2014 Apr 1.

Bredemeijer, G., Cooke, R., Ganal, M., Peeters, R., Isaac, P., Noordijk, Y., Rendell, S., Jackson, J., Röder, M., Wendehake, K., Dijcks, (2002). Satellite database containing more than 500 tomato varieties. *Theor Appl Genet* 105: 1019-1026.

Bukenya, Z.., and Carasco, F. (1999). Ethnobotanical aspects of Solanum L. (Solanaceae) in Uganda. In: Nee, M., Symon, D. E., Lester, R. N., & Jessop, J. P. ed(s). Solanaceae IV. *Advances in botany and utilization*, pp. 345-360. Royal Botanic Gardens, Kew.

Catanzariti, A., Lim, G. (2015). The tomato I-3 gene: a novel gene for resistance to Fusarium wilt disease. *New Phytol*. 207(1):106-18. Doi: 10.1111/nph.13348. Epub 2015 Mar 4.

Chadha, M. and Oluocha, M. (2003). Home-based Vegetable Gardens and other Strategies to Overcome Micronutrient Malnutrition in Developing Countries. *Food, Nutrition and Agricultural Journal*. 32(1):17-23.

Chadha, M. and Mndiga, H. (2007). African Eggplant From Underutilized to a Commercially Profitable Venture. In: Proceedings of the 1st International Conference on Indigenous Vegetables and Legumes— Prospectus for Fighting Poverty, Hunger and Malnutrition. Eds. M.I. Chadha, G. Kuo, C. L.

L. Gowda, *Acta Horticulturalae. International Society for Horticultural Sciences (ISHS)*: Korbeek-Lo, Belgium, 2007; pp. 521–523.

Chatfield, C., Collins, J. (1980). Introduction to Multivariate Analysis. *Chapman & Hall, London and New York* 1980. 246 S

Chen, X., Zhang, M., Tan, J., Huang, S., Wang, C., Zhang, H. (2017). Comparative transcriptome analysis provides insights into molecular mechanisms for parthenocarpic fruit development in eggplant (*Solanum melongena L*.). *PLoS ONE* 12(6): 0179491. https://Doi.org/ 10.1371/journal.pone.0179491

Chen, B., Yu, R., Dong, J., Bihao, C. (2017). RNA-Seq-derived identification of differential transcription in the eggplant (*Solanum melongena*) following inoculation with bacterial wilt. *Elsievier* https://doi.org/10.1016/j.gene.2017.11.003

Chung, D., Barker, B. M., Carey, C. C., Merriman, B., Werner, E. R., Lechner, B. E., … Cramer, R. A. (2014). ChIP-seq and In Vivo Transcriptome Analyses of the Aspergillus fumigatus SREBP SrbA Reveals a New Regulator of the Fungal Hypoxia Response and Virulence. *PLoS Pathogens*, *10*(11).

Cingolani, P., Patel, V., Coon, M., Nguyen, T., Land, S., Ruden, D., Lu, X. (2012). Using Drosophila melanogaster as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. *Frontiers in genetics*, *3*, 35. Doi:10.3389/fgene.2012.00035

Cloonan, N., Forrest, A., Kolle, G., Gardiner, B., Faulkner, G., Brown, M., Taylor D., Steptoe, A., Wani, S., Bethel, G. (2008). Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nat Methods* 5*:* 613–619

Collonnier, C., Fock, I., Kashyap, V., Rotino, G., Daunay, M., Lian, Y. (2001). Applications of biotechnology in eggplant. *Plant Cell Tiss Org*. Cult. 65, 91–107.

Cook, T., Pichaud, F., Sonneville, R., Papatsenko, D., Desplan, C. (2003). Distinction between colour photoreceptor cell fates is controlled by Prospero in Drosophila. *Dev. Cel*l 4(6): 853--864.

Danhashet, E., Harris, E., Skipper, B., and Allen, S. (1993). Doing naturalistic inquiry: a guide to methods. *Newbury Park*, CA: Sage Publications, Inc.,

Daunay, M., Hazra, P. (2012). Eggplant, in Handbook of Vegetables, eds Peter K. V., Hazra P., editors. (Houston, TX: Studium Press ;), 257–322.

Daunay, M., Lester, R., Gebhardt, C., Hennart, J., Jahn, M., Frary, A., Doganlar, S. (2001). Genetic Resources of Eggplant (Solanum melongena L.) and Allied Species: a New Challenge for Molecular Geneticists and Eggplant Breeders. Nijmegan. The Netherlands: Nijmegen University Press; 2001.

Davey, R., Savva, G., Dicks, J., (2007). MPP: a microarray-to-phylogeny pipeline for analysis of gene and marker content datasets. *Bioinformatics* 23(8):1023-5

Denton, O., and Nwangburuka, C. (2011). Genetic variability in eighteen cultivars of *Solanum anguivi* Lam. using principal component analysis (PCA) and single linkage cluster analysis (SLCA). *Annals Biological Research*, 2(4): 62–67

Der, J., Barker, M., Wickett, N., dePamphilis, C., Wolf, P. (2011). De novo characterization of the gametophyte transcriptome in bracken fern, *Pteridium aquilinum* BMC Genomics , 2011 vol. 12 pg. 99

Dewey, and Li, B. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome BMC Bioinformatics 12:323. https://doi.org/10.1186/1471-2105-12-323

Dhaygude, K, Trontti, K., Paviala, J., Morandin, C., Wheat, C., Sundström, L., Helanterä, H. (2017). Transcriptome sequencing reveals high isoform diversity in the ant *Formica exsecta. PeerJ* 5: e3998 https://doi.org/10.7717/peerj.3998

Ding, D., Lawson, K., Kolbe-Alexander, T., Finkelstein, E., Katzmarzyk, P., van Mechelen, W., Pratt, M. (2016). The economic burden of physical inactivity: a global analysis of major non-communicable diseases. Lancet. 2016 Sep 24; 388 (10051):1311-24. Doi: 10.1016/S0140-6736(16)30383-X. Epub 2016 Jul 28.

Ding, L., Ellis, M., Li, S., Larson, D., Chen, K., Wallis, J., Harris, C., McLellan, M., Fulton, R., Fulton, L., Abbott, R., Hoog, J., Dooling, D., Koboldt, D., Schmidt, H., Kalicki, J., Zhang, Q., Chen, L., Lin, L., Wendl, M., McMichael, J., Magrini, V., Cook, L., McGrath, S., Vickery, T., Appelbaum, E., Deschryver, K., Davies, S., Guintoli, T., Lin, L., Crowder, R., Tao, Y., Snider, J., Smith, S., Dukes, A., Sanderson, G., Pohl, C., Delehaunty, K., Fronick, C., Pape, K., Reed, J., Robinson, J., Hodges, J., Schierding, W., Dees, N., Shen, D., Locke, D., Wiechert, M., Eldred, J., Peck, J., Oberkfell, B., Lolofie, J., Du, F., Hawkins, A., O'Laughlin, M., Bernard, K., Cunningham, M., Elliott, G., Mason, M., Thompson, D. , Ivanovich, J., Goodfellow, P., Perou, C., Weinstock, G., Aft, R., Watson, M., Ley, T., Wilson, R., Mardis, E. (2010). Genome remodelling in a basal-like breast cancer metastasis and xenograft. Nature. 2010 Apr 15;464 (7291):999-1005. Doi: 10.1038/nature08989.

Du, H., Wang, Y., Yang, J. and Yang, W. (2015). Comparative Transcriptome Analysis of Resistant and Susceptible Tomato Lines in Response to Infection by Xanthomonas perforans Race T3. Front. Plant Sci. 6:1173. doi: 10.3389/fpls.2015.01173

Dugas, V., Marcela, K., Andrew, O., Robert, R., Sunita, K., Doreen, W. and Patricia, E. (2011). Functional annotation of the transcriptome of *Sorghum bicolor* in response to osmotic stress and abscisic acid BMC Genomics 2011 **12**:514 https:// Doi.org/10.1186/1471-2164-12-514

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B., Bibillo, A., Bjornson, K., Chaudhuri, B.,

Christians, F., Cicero, R., Clark, S., Dalal, R., Dewinter, A., Dixon, J., Foquet, M., Gaertner,A., Hardenbol, P., Heiner, C., Hester, K., Holden, D., Kearns, G., Kong, X., Kuse, R., Lacroix, Y., Lin S., Lundquist, P., Ma, C., Marks, P., Maxham, M., Murphy, D., Park, I., Pham, T., Phillips, M., Roy, J., Sebra, R., Shen, G., Sorenson, J., Tomaney, A., Travers, K., Trulson, M., Vieceli, J., Wegener, J., Wu, D., Yang, A., Zaccarin, D., Zhao, P., Zhong, F., Korlach, J., Turner, S. (2009). Real-time DNA sequencing from single polymerase molecules. 323 (5910) :133-8. Doi: 10.1126/science.1162986.

Elekofehinti, O., Adanlawo, I., Fakoya, A., Saliu, J. and Sodehinde, S. (2012). Effect of saponin from *Solanum anguivi* Lam. fruit on heart and kidney superoxide dismutase, catalase and malondialdehyde in rat. *Current Research Journal of Biological Sciences*, 4(4): 530–533.

Ellegren, H., Smeds, L., Burri, R., Olason, P., Backström, N., Kawakami, T. (2012).. The genomic landscape of species divergence in Ficedula flycatchers. Nature. 2012; 491:756–760.

Elshire, R., Jeffrey, C., Glaubitz, S., Jesse, A., Poland, K., Edward, S., Buckler, E., Mitchell, (2011). A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species PLoS ONE 6(5): e19379. https://doi.org/10.1371/journal.pone.0019379

Filichkin, A., Priest, D., Givan, A., Shen, R., Bryant, W., Fox, E., Mockler, C. (2010). Genome-wide mapping of alternative splicing in *Arabidopsis thaliana*. *Genome Research*, *20*(1), 45–58. http://doi.org/10.1101/gr.093302.109

Franssen, S., Jenny, G., Ann-Kathrin, H., Isabell, W., Maximiliane, S., James, A., Jeanine, L., Thorsten, B., Erich, B. (2014). Genome-wide transcriptomic responses of the seagrasses *Zostera marina* and *Nanozostera noltii* under a simulated heatwave confirm functional types *Eur J Hum Genet.* ; 23(1):110-5. doi: 10.1038/ejhg.2014. Pages 65-73

Frary, A., Doganlar, S. and Daunay, M. (2007). Eggplant In: Genome mapping and molecular breeding in plants: *Vegetables Vol 5*, by C, Kole, 287-313 Verlag Berlin Heidelberg: Springer

Fukuoka, S., Mizobuchi, R., Saka, N., Suprun, I., Matsumoto, T., Okuno, K., and Yano, M. (2012). A multiple gene complex on rice chromosome 4 is involved in durable resistance to rice blast. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, *125*(3), 551-9.

Gadea, J., Mayda, V. Conejero, P. (1996). Characterization of defense-related genes ectopically expressed in viroid-infected tomato plants. *Molecular Plant-Microbe Interactions* 9(5):409-15

Gallie, D. (2013). Direct participation and the quality of work First Published April 4, 2013 Research Article https://doi.org/10.1177/0018726712473035.

Gan, X., Oliver S., Jonas, B., Joshua, G., Steffen, P., Drewe, K., Hildebrand, R., Lyngsoe, S., Schultheiss, E., Osborne, V., Sreedharan, A., Regina B,, Géraldine J., Paul, D., Paul, K., Eric, J., Belfield, N,. Harberd, E. Kemen, C., Paula, X., Kover, R., Clark, G., and Richard, M. (2011). Multiple reference genomes and transcriptomes for *Arabidopsis thaliana. Nature* **477**, 419–423 Doi: 10.1038/nature10414

Gapper, E., Giovannoni, J., Watkins, C. (2014). Understanding development and ripening of fruit crops in an 'omics' era. *Hortic. Res.* 1 14034 10.1038/hortres.2014.34

Gautier, A., Juillerat, A., Heinis, C., Corrêa, I., Kindermann, M., Beaufils, F., Johnsson, K. (2008). An engineered protein tag for multiprotein labeling in living cells. *Chem Biol*. 2008 Feb;15(2):128-36. doi: 10.1016/j.chembiol.2008.01.007

Giovannoni, J. (2007). Fruit ripening mutants yield insights into ripening control. *Plant Biol*. 10(3):283-9. DOI:10.1016/j.pbi.2007.

Giovannoni, J. and Harry, K. (2011). Genetics and Control of Tomato Fruit Ripening and Quality Attributes. *Annu. Rev. Genet*. 2011. 45:41–59 Doi: 10.1146/annurev-genet-110410-132507

González-Ballester, D., Casero, D., Cokus, S., Pellegrini, M., Merchant, S., Grossman, A. (2010). RNA-seq analysis of sulfur-deprived *Chlamydomonas* cells reveals aspects of acclimation critical for cell survival. *Plant Cell.* 2010;22:2058–2084. doi: 10.1105/tpc.109.071167

Gopalakrishnan, S., Sathya, A., Vijayabharathi, R. (2015). *Biotech* 5: 355. https://doi.org/10.1007/s13205-014-0241-x

Grandillo, S. and Chetelat, R. (2011). *Solanum lycopersicon* Wild Crop Relatives: Genomic and Breeding Resources. C. Kole, *Springer Berlin Heidelberg*: 129–215.

Guo, M., Rupe, M., Zinselmeier, C. (2004). Allelic variation of gene expression in maize hybrids. *The Plant Cell Online*, 16, 1707–1716.

Haas, B., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P., Bowden, J. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc.* 8:1494±512. Doi: 10.1038/nprot.2013.084 PMID: 23845962

Hamid, A., Tengku, A., Abdul, W., Omar, R. and Huyop, F. (2015). An S188V Mutation Alters Substrate Specificity of Non-Stereospecific α-Haloalkanoic Acid Dehalogenase E (DehE). *PLoS ONE*, 10(3), e0121687. http://doi.org/10.1371/journal.pone.0121687

Hamilton and Buell, (2012). Advances in plant genome sequencing. *Plant J.* 70(1):177-90. doi: 10.1111/j.1365-313X.2012.04894.x

Hanson, J. and Smeekens, S. (2009). Sugar perception and signaling - an update. *Current Opinion in Plant Biology*, 12: 562–567.

Haseneyer, G., Schmutzer, T., Seidel, M., Zhou, R., Mascher, M., Schön, C., Taudien, S., Scholz, U., Stein, N., Mayer,K.., Bauer, E. (2011). From RNA-seq to large-scale genotyping Genomics resources for rye (Secale cereale L.). *BMC Plant Biol*. 11: 131.

He, L., Wang, Q., Jin, X., Wang, Y., Chen, L., Liu, L. (2012). Transcriptome profiling of testis during sexual matu- ration stages in Eriocheir sinensis using Illumina sequencing. *PLoS One*. 2012; 7(3):e33735. Doi: 10.1371/journal.pone.0033735 PMID: 22442720

Hirakawa, H, Shirasawa, K., Miyatake, K, Nunome, T., Negoro, S, Ohyama, A. (2014). Draft genome sequence of eggplant (*Solanum melongena* L.): the representative solanum species indigenous to the old world. *DNA Res*. 2014; 21: 649±660. https://doi.org/10.1093/dnares/dsu027 PMID: 25233906

Hirakawa, H., Tsuchimoto, S., Sakai, H., Nakayama, S., Fujishiro, T., Kishida, Y., Kohara, M., Watanabe, A., Aizu, T., Toyoda, A., Fujiyama, A., Yamada, M., Tabata, S., Fukui, K., Sato, S. (2012). Upgraded genomic information of Jatropha curcas L. *Plant Biotechnol*.; 29:123–130

Hoeijmakers, W., Bártfai, R., Stunnenberg, H, (2012). Transcriptome Analysis Using RNA-Seq. In: Ménard R. (eds) Malaria. Methods in Molecular Biology (Methods and Protocols), vol 923. Humana Press, Totowa, NJHwang, P., Chou, S.W., Chen, Z., McDermott, B.M. 2015.the Stereociliary Paracrystal Is a Dynamic Cytoskeletal Scaffold In Vivo *Cell Reports* 13(7): 1287-94

International Network for the Improvement of Banana and Plantain, (INIBAP), Montpellier (France) (2003). ISBN 10:1029-2209 PG 40a

Jagatheeswari, D. (2014). Morphological studies on flowering plants (Solanaceae). *International Letters of Natural Sciences*, (15)10: 36–43.

Jochen, B., Wolf, (2013). Principles of transcriptome analysis and gene expression quantification: an RNA-seq tutorial. *Molecular Ecology Resources* 13, 559–572 Doi: 10.1111/1755-0998

Kakumanu, A, Ambavaram, M., Klumas, C., Krishnan, A., Batlang, U., Myers, E., Grene, R., Pereira, A. (2012). Effects of drought on gene expression in maize reproductive and leaf meristem tissue revealed by RNA-Seq. *Plant Physiol*. 2012 Oct;160(2):846-67. doi: 10.1104/pp.112.200444. Epub 2012 Jul 26.

Kaushik, P., Prohens, J., Vilanova, S., Gramazio, P. and Plazas, M. (2016). Phenotyping of Eggplant Wild Relatives and Interspecific Hybrids with Conventional and Phenomics Descriptors Provides Insight for Their Potential Utilization in Breeding. *Frontiers in plant science*, *7*, 677. doi:10.3389/fpls.2016.00677

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., & Salzberg, S. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology*, *14*(4), R36. http://doi.org/10.1186/gb-2013-14-4-r36

Kisua, J, Mwikamba, J, Muigai, A. (2014). Genetic diversity of sweet and grain sorghum populations using phenotypic markers *International Journal of Biosciences* 6: 34–46.

Knapp, S., Vorontsova, M., Prohens, J., (2013). Wild relatives of the eggplant (Solanum melongena L.: Solanaceae): new understanding of species names in a complex group. *PLoS One*, 8(2):e57039.

Knapp, S., Bohs, L., Nee, M. and Spooner, D. (2004). Solanaceae: A model for linking genomics with biodiversity. *Comparative Functional Genomics*, 5: 285–291

Kochieva, E., Ryzhova, I., Khrapalova, V., Pukhalskyi. (2002). Genetic Diversity and Phylogenetic Relationships in the Genus Lycopersicon (Tourn.) Mill. as Revealed by Inter-Simple Sequence Repeat (ISSR) Analysis. *Russian Journal of Genetics* 38(8):958-966 DOI: 10.1023/A: 1016896013986.

Kole, C. (2007). Genome Mapping and Molecular Breeding in Plants. Volume 5 Vegetables. *Springer-Verlag Berlin Heidelberg*. pp 1-95

Kwon, Y., Apostolidis, E. and Shetty, K. (2008). In vitro studies of eggplant (*Solanum melongena*) phenolics as inhibitors of key enzymes relevant for type 2 diabetes and hypertension. *Bioresource Technology*, 99: 2981–2988.

Lancu, O., Kawane, S., Bottomly, D., Searles, R., Hitzemann, R., McWeeney, S. **(**2012). Utilizing RNA-Seq data for de novo coexpression network inference *Bioinformatics*. Jun 15;28(12):1592-7. doi: 10.1093/bioinformatics/bts245. Epub 2012 May 3

Lawal, O., Lechner, W., Kulicke. (2007). Single and multi-step carboxymethylation of water yam (*Dioscorea alata*) starch: synthesis AND characterization, *Int. J. Biol*. Macromol., 42 (2008), pp. 429-435

Lee, J., Joung, J., McQuinn, R., Chung, M., Fei, Z., Tieman, D., Klee, H. and Giovannoni, J. (2012). Combined transcriptome, genetic diversity and metabolite profiling in tomato fruit reveals that the ethylene response factor SlERF6 plays an important role in ripening and carotenoid accumulation. *The Plant Journal*, 70: 191–204

Lee, S., Ramirez, J., Franco, M., Lectez, B., Gonzalez, M., Barrio, R., Mayor, U. (2014). Ube3a, the E3 ubiquitin ligase causing Angelman syndrome and linked to autism, regulates protein homeostasis through the proteasomal shuttle Rpn10. *Cell. Molec. Life Sci.* 71(14): 2747--2758

Lee, T., Guo, H., Wang, X., Kim, C. and Paterson, A. (2014). SNPhylo: a pipeline to construct a phylogenetic tree from huge SNP data. *BMC Genomics* 15,162.

Lenz, S., Karsten, P., Schulz, J.B., Voigt, A. (2013). Drosophila as a screening tool to study human neurodegenerative diseases. *J. Neurochem*. 127(4): 453--460.

Li H., Durbin R. (2010). Fast and accurate long-read alignment with burrows-wheeler transform. *Bioinformatics* 26, 589–595.10.1093/bioinformatics/btp698.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009; 25(15):2078±79.

Li, E., Li, C. (2014). Use of RNA-seq in aquaculture research. *Poult Fish Wildl Sci*. 2014;

Lian, X., Bao, X., Zilberter, M., Westman, M., Fisahn, A., Hsiao, C., Hazeltine, L. B., Dunn, K. K., Kamp, T. J., Palecek, S. P. (2015). Chemically defined, albumin-free human cardiomyocyte generation. *Nature methods*, *12*(7), 595-6.

Liu S, Wen-Yang L, Dang L, Chun-Gen H, Jin- Zhi Z. (2013). Development and Characterization of Genomic and Expressed SSRs in Citrus by Genome-Wide Analysis *Plos One* Doi.org/10.1371/journal.pone.0075149

Lu, W.J., Chapo, J., Roig, I., Abrams, J.M. (2010). Meiotic recombination provokes functional activation of the p53 regulatory network. *Science* 328(5983): 1278--1281.

Mardis, E. (2008). Next-Generation DNA Sequencing Methods Annual Review of Genomics and Human Genetics 2008 9:1, 387-402 *Science.* 323(5910):133-8. Doi: 10.1126/science.1162986. Epub 2008 Nov 20.

Marioni, J., Mason C., Mane S., Stephens M., Gilad Y. (2008). RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res.* 18(9):1509-17. Doi: 10.1101/gr.079558.108.

Mead, J, Bruning, A., Gill, M., Steiner, A., Acton, T., Vershon, A. (2002). Interactions of the Mcm1 MADS box protein with cofactors that regulate mating in yeast. *Mol Cell Biol* 22(13):4607-21

Mibei, E. K., Ambuko, J., Giovannoni, J., Onyango, A., and Owino, W. (2017). Carotenoid profiling of the leaves of selected African eggplant accessions

subjected to drought stress. *Food science & nutrition*, *5*(1), 113-122. doi:10.1002/fsn3.370

Mishra, S., Aastik, J., Kushwah, M. and Mishra, V. (2013). Genetic divergence in eggplant (*Solanum melongena* L.) Vegetable Science 40 (2) : 219-220

Mizrahi, R., Kenk, M., Suridjan, I., Boileau, I., George, T. P., McKenzie, K., Wilson, A. A., Houle, S., Rusjan, P. (2014). Stress-induced dopamine response in subjects at clinical high risk for schizophrenia with and without concurrent cannabis use. *Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology*, *39*(6), 1479-89.

Mortazavi, A., Williams, B., McCue, K., Schaeffer, L., Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods*. 2008 Jul; 5(7):621-8. Doi: 10.1038/nmeth.1226. Epub 2008 May 30.

Muratore, M., Srsen, V., Waterfall, M., Downes, A. and Pethig, R. (2012). 'Biomarker-free dielectrophoretic sorting of differentiating myoblast multipotent progenitor cells and their membrane analysis by Raman spectroscopy.' *Biomicrofluidics*, vol. 6, no. 3, 034113. DOI: 10.1063/1.4746252

Nagalakshmi, U., Wang, Z., Waern, K,, Shou, C., Raha, D., Gerstein, M, Snyder, M. (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing, *Science* 6;320(5881):1344-9. Doi: 10.1126/science.1158441.

Naurin, S., Hansson, S., Bensch, and D. Hassequist, D. (2010). Why does dosage compensation differ between XY and ZW taxa? *Trends Genet.* 26:15–20.

Ni, Y., Amelia, W., Anna, B., and Vishwanath, R. (2012). SNP identification and assessment of allele-specific bias from ChIP-seq data. *BMC Genetics* 13:46 https://doi.org/10.1186/1471-2156-13-46.

Nielsen, R., Findley, M., Davis, Z., Candland, T. and Nielson, D. (2011), Foreign Aid Shocks as a Cause of Violent Armed Conflict. *American Journal of Political Science*, 55: 219–232. doi:10.1111/j.1540-5907.2010.00492.x

Nookaew, I., Papini, M., Pornputtapong, N., Scalcinati, G., Fagerberg, L, Uhlén, M., Nielsen, J., (2012). Comprehensive comparison of RNA-Seq-based transcriptome analysis from reads to differential gene expression and cross-comparison with microarrays: a case study in Saccharomyces cerevisiae. *Nucleic Acids Res*. 2012 Nov 1; 40 (20):10084-97. doi: 10.1093/nar/gks804. Epub 2012 Sep 10.

Osei, C., Dontwi, K., Otchere, G. and Singye, A. (2014). Curriculum relevance and graduate career: a tracer study of cemba, cempa and industrial mathematics graduates, *Journal of Science and Technology*, Vol. 35, No. 2 (2015), pp89-99. http://dx.doi.org/10.4314/just.v35i2.9

Pabinger, S., Dander, A., Fischer, M., Snajder, R., Sperk, M., Efremova, M., (2014). A survey of tools for variant analysis of next-generation genome sequencing data. *Brief. Bioinformatics* 15, 256–278.10.1093/bib/bbs086

Patel, R., Jain, M. (2012). NGS QC Toolkit: A Toolkit for Quality Control of Next Generation Sequencing Data. *PLoS ONE* 7(2): e30619. https://doi.org/10.1371/journal.pone.0030619

Peter, R., Leavitt, S., Fritz, N., Anderson, P., Baker, T., Blenckner, L., Bunting, J., Catalan, D., Conley, W.,. Hobbs, E., Jeppesen, A., Korhola, S., McGowan, K., Ru¨ hland, J., Rusak, G., Simpson, N., Solovieva, and Werne, J. (2008). Paleolimnological evidence of the effects on lakes of energy and mass transfer from climate and humans Limnol. *Oceanogr*. 54(6, part 2), 2009, 2330–2348 2009

Provenzano, G., Corradi, Z., Monsorno, K., Fedrizzi, T., Ricceri, L., Scattoni, M. and Bozzi, Y. (2016). Comparative Gene Expression Analysis of Two Mouse

Models of Autism: Transcriptome Profiling of the BTBR and En2−/− Hippo campus. *Front. Neurosci*.10:396.

Rao, A. and Rao, L. (2007). Carotenoids and human health. *Pharmacological research*, 55(3): 207–216

Rao, R., Dash, A., Rao, G. , Verma, R. , Katara, J., Mukherjee, A., Singh, O., Bagchi, T. (2016). Phenotypic and Marker-Assisted Genetic Enhancement of Parental Lines of Rajalaxmi, an Elite Rice Hybrid. *Frontiers in plant science*, *7*, 1005. doi:10.3389/fpls.2016.01005

Ray-Yu, Y. and Ojiewo, C. (2011). African Nightshades and African Eggplants: Taxonomy, Crop Management, Utilization, and Phytonutrients. African Natural Plant Products Volume II: Discoveries and Challenges in Chemistry, Health, and Nutrition. *ACS Symposium Series; American Chemical Society: Washington, DC,* pp 137–165.

Raz, N., and Linden, B. (2010). News of cognitive cure for age-related brain shrinkage is premature: a comment on Burgmans et al. (2009). *Neuropsychology*, *24*(2), 255-7.

Reddy, A.R., Chaitanya, K.V. and Vivekanandan, M. (2004). Drought induced responses of photosynthesis and antioxidant metabolism in higher plants. *Journal of Plant Physiology,* 161: 1189–1202

Romano, D. and Paratore, A. (2014). Effects of grafting on tomato and eggplant. *Acta Hortic*. 559, 149-154 DOI: 10.17660/ActaHortic.2001.559.21

Sabatini, D., Erdjument-Bromage H., Lui M., Tempts P., Snyder, S. (2006). RAFT1: a mammalian protein that binds to FKBP12 in a rapamycin-dependent fashion and is homologous to yeast TORs. *Cell*, 78 (1994), pp. 35-43

Saito, T., Ariizumi, T., Okabe, Y., Asamizu, E., Hiwasa-Tanase, K., Fukuda, N., Mizoguchi, T., Yamazaki, Y., Aoki, K, Ezura, H. (2011). TOMATOMA: a

novel tomato mutant database distributing Micro-Tom mutant collections. *Plant & cell physiology*, *52*(2), 283-96.

Saliba-Colombani, V., Causse, M., Gervais, L., Philouze, J. (2000). Efficiency of RFLP, RAPD, and AFLP markers for the construction of an intraspecific map of the tomato genome. *Genome*. 43(1):29-40.

Sarah, M. and Maina, M. (2008). Vegetables in East Africa. Elewa Publications, Farming Resources Series.

Sarvajeet and Narendra (2010). Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants *Narendra Tuteja* https://doi.org/10.1016/j.plaphy. 2010.08.016 *Plant Physiology and Biochemistry* pp 909-930

Sathya, S., Reuben, J., Prathebha, K., Usha, G., and Jovita, J. (2014). (E)-3-(4-Hydroxy-3-meth-oxy-phen-yl)-1-(4-hy-droxy-phen-yl) prop-2-en-1-one. *Acta crystallographica. Section E, Structure reports online*, *70*(Pt 5), o593-4. Doi: 10.1107/S1600536814008757

Schafleitner R, Raymundo Oscar Gutierrez Rosales Ricardo Espino, Merideth Bonierbale (2007). Field Screening for Variation of Drought Tolerance in Solanum tuberosum L. *Agronomical, Physiological and Genetic Analysis.* DOI: 10.1007/s11540-007-9030-9 Potato Research 50(1):71-85

Schatz, M., Jan,W., Richard, W., (2012). Current challenges in *de novo* plant genome sequencing and assembly. *Genome Biology* **13**:243 https://doi.org/10.1186/gb-2012-13-4-243.

Scherr, S. and McNeely, J. (2007). Biodiversity conservation and agricultural sustainability: towards a new paradigm of 'Eco agriculture' landscapes. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *363*(1491), 477-94.

Schippers, R.R. (2000). African indigenous vegetables: an overview of cultivated species. Natural resource institute, *Chatham, London*, UK. ISBN 0 85954 515 6. pp. 214.

Sękara, A., Cebula, S. and Kunicki, E. (2007). Cultivated eggplants–origin, breeding objectives and genetic resources, a review. *Folia Horticulturae*, 19(1): 97–114.

Shafiei, S. (2015). Apolarity for determinants and permanents of generic matrices *J. Commut.* pp 89-123.

Shalom, N. , Abayomi, C., Olasumbo, Okwuchukwu, K., Eboji, O, Olajum oke, K. and Damilola, I. (2011). Proximate and Phytochemical Analyses of Solanum aethiopicum L. and Solanum macrocarpon L. *Fruits. Res. J. Chem. Sci.1* (3): 63-71

Shirasawa, K., Fukuoka, H., Matsunaga, H., Kobayashi, Y., Kobayashi, I. (2013). Genome-wide association studies using single nucleotide polymorphism markers developed by re-sequencing of the genomes of cultivated tomato. *DNA Res*. 20: 593–603.

Silin, Z., Je-Gun, J., Yi, Z., Yun-ru, C., Bao,L., Ying, S., Jenny, Z., Xiang, Zhangjun, F., and Giovannoni, J. (2014). High-Throughput Illumina Strand-Specific RNA Sequencing Library Preparation. *Cold Spring Harb Protoc;* Doi:10.1101/pdb.prot5652

Slewinski, T., Anderson, A., Zhang, C., Turgeon, R. (2012). Scarecrow plays a role in establishing Kranz anatomy in maize leaves. *Plant Cell Physiol.* Doi: 10.1093/pcp/pcs147. Epub 2012 Nov 4.

Sol Genomics Network: International Tomato Genome Sequencing Project. (*http://solgenomics.net/organism/Solanum_lycopersicum/genome).*

Stàgel, A., Portis, E., Toppino, L., Rotino, G.L. and Lanteri, S. (2008). Gene-based microsatellite development for mapping and phylogeny studies in eggplant. *BMC genomics*, 9(1): 357–370

Sun, F., Liu, S., Gao, X., Jiang, Y., Perera, D., Wang, X. (2013). Male-biased genes in catfish as revealed by RNA-Seq analysis of the testis transcriptome. *PLoS One*. 2013; 8:e68452. Doi: 10.1371/journal.pone.0068452 PMID: 23874634

Sun, J., Kawakami, H., Zech, J., Speck, C., Stillman, B., Li, H. (2012). Cdc6-induced conformational changes in ORC bound to origin DNA revealed by cryo-electron microscopy. *Structure* 20(3):534-44

Tam, S., Mhiri, C., Vogelaar, A., Kerkveld, M., Pearce, S., Grandbastien, M. (2005). Comparative analyses of genetic diversities within tomato and pepper collections detected by retrotransposon based SSAP, AFLP and SSR. *Theor. Appl. Genet*. 110: 819-831.

Tanaka, T., Shen J, Abecasis, G., Kisialiou, A., Ordovas, J., Guralnik, J. (2009). Genome-Wide Association Study of Plasma Polyunsaturated Fatty Acids in the InCHIANTI Study. *PLoS Genet* 5(1) doi.org/10.1371/journal. pgen. 1000338

Tanigaki, Y., Higashi, T., Takayama, K., Nagano, A., Honjo, M,, Fukuda, H. (2015). Transcriptome Analysis of Plant Hormone-Related Tomato (*Solanum lycopersicum*) Genes in a Sunlight-Type Plant Factory. *PLoS ONE* 10(12): e0143412. https://doi.org/10.1371/journal.pone. 0143412 Tanzania. Tropicultura 30, 185–191

Tembe, K., Chemining'wa, G., Ambuko, J., and Owino, W. (2017). Effect of water stress on yield and physiological traits among selected African tomato (Solanum lycopersicum) land races. *International Journal of Agronomy and Agricultural Research*, 10(2): 78–85

The Gene Ontology Consortium, Ashburner, M., Ball, C., Blake, J., Botstein, D., Butler, H., Sherlock, G. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, *25*(1), 25–29. http://doi.org/10.1038/75556.

Tirosh, I., Reikhav, S., Levy, A.A, Barkai, N. (2009). A yeast hybrid provides insight into the evolution of gene expression regulation. *Science*, 324, 659–662.

Varshney, R., Chen, W., Li, Y., Bharti, A., Saxena, R., Schlueter, J., Donoghue, M., Azam, S,. Fan, G., Whaley, A., Farmer, A., Sheridan, J., Iwata, A., Tuteja, R., Penmetsa, R., Wu, W., Upadhyaya, H., Yang, S., Shah, T., Saxena, K., Michael, T., McCombie, W., Yang, B., Zhang, G., Yang, H., Wang, J., Spillane, C., Cook, D., May, G., Xu, X.(2012). Draft genome sequence of pigeonpea (Cajanus cajan), an orphan legume crop of resource-poor farmers. *Nat Biotechnol.*, 30: 83-89

Vera, J., Wheat, C., Feesceonfer, H., Frilinder, M., Crawford, D. (2008). Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. *Mol.Ecol*.17:1636-1647.

Vrček, I., Mirza Bojić Irena Ž., Marica, Medić-Š. (2011). Phenol content, antioxidant activity and metal composition of Croatian wines deriving from organically and conventionally grown grapes. *Food Chemistry* 124(1) :354-361DOI: 10.1016/j.foodchem.2010.05.118

Waiho, K., Fazhan, H., Shahreza, M., Moh, J., Noorbaiduri, S., Wong, L. (2017). Transcriptome Analysis and Differential Gene Expression on the Testis of Orange Mud Crab, *Scylla olivacea*, during Sexual Maturation. *PLoS ONE* 12(1): e0171095. Doi: 10.1371/journal. pone.0171095

Wang, C., Li, S., Januschke, J., Rossi, F., Izumi, Y., Garcia-Alvarez, G., Gwee, S., Soon, S.B., Sidhu, H.K., Yu, F., Matsuzaki, F., Gonzalez, C., Wang, H. (2011). An ana2/ctp/mud complex regulates spindle orientation in Drosophila neuroblasts. *Dev. Cell* 21(3): 520--533.

Wang, K., Li, M, Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data, *Nucleic Acids Research*, Volume 38, PPe164, https://doi.org/10.1093/nar/gkq60

Wang, Y. (2012). Coiled-coil networking shapes cell molecular machinery. *Mol Biol Cell* 23(19):3911-22

Wang, Z., Wang, J. and Chan, P. (2013). Treating type 2 diabetes mellitus with traditional Chinese and Indian medicinal herbs. *Evidence-Based Complementary Alternative Medicine.* doi:10.1155/2013/343594.

Wencai, Y., Yu, F., Lin, S., Chia, W., Yang, X. (2003). Apical complex genes control mitotic spindle geometry and relative size of daughter cells in Drosophila neuroblast and pI asymmetric divisions. *Cell* 112(1): 51--62.

Wolf, J., Schliebs, W., Erdmann, R. (2010). Peroxisomes as dynamic organelles: peroxisomal matrix protein import. *FEBS J* 277(16):3268-78.

Wu, G., Sun, J., Yu, S., Dong, Q., Zhuang, G., Liu, W., Lin, J. (2015). Improved activity of the Cel5A endoglucanase in Saccharomyces cerevisiae deletion mutants defective in oxidative stress defense mechanisms. *Biotechnol Lett* 37(5):1081-9

Xu, C., Jiao, C., Zheng, Y., Liu, W., Cai, X., Wang, X. (2015). De novo and comparative transcriptome analysis of cultivated and wild spinach. *Sci Rep.* 5:17706. Doi: 10.1038/srep17706 PMID: 26635144.

Yang, S., He, J., Sun, C., Chan, S. (2014a) Characterization of the shrimp neuroparsin (MeNPLP): RNAi silencing resulted in inhibition of vitellogenesis. *FEBS Open Bio.* 2014; 4:976–86. doi: 10.1016/j.fob.

Yang, X., Yu-Fu, C., Cao, D., Yan, M., Zhi-Wen, W., Xue-Hao, C. and Lin-Bao, X. (2014). Comparative transcriptome analysis of eggplant (Solanum melongena

L.) and turkey berry (Solanum torvum Sw.): Phylogenomics and disease resistance analysis. *BMC Genomics* 2014 15:412.

Zhang, L., Li, D., Gao, P. (2012). Expulsion of selenium/protein nanoparticles through vesicle-like structures by Saccharomyces cerevisiae under microaerophilic environment. *World J Microbial Biotechnology* 28(12):3381-6

Zhang, Y., Moqtaderi, Z., Rattner, B., Euskirchen, G., Snyder, M., Kadonaga, J., Liu, X., Struhl, K. (2009). Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions in vivo. *Nat Struct Mol Biol* 16(8):847-52.

Zhou, X. and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nature genetics*, *44*(7), 821-4. doi:10.1038/ng.2310

Zouine, M. Latché, A., Rousseau, C., Regad, F., Pech, J., Philippot, M., Bouzayen, M., Delalande, C., Frasse, P., Schiex, T. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*, 485:635–641.

APPENDICES

Appendix I. RNA extraction

RNA was extracted from the leaf samples, and fruit samples at different fruiting stages.

Extraction protocol

1. Samples were collected from the greenhouse, placed immediately in well labelled test tubes containing liquid nitrogen.
2. The leaf samples were ground undere liquid nitrogen using a mortar and pestle while the fruit samples were ground using an electrical mortar.
3. The powder was return into the falcon tubes and placed in dry ice and later kept in -20 C.
4. 200-300 mg of the leaf powder was weighed and placed in a 2 ml Eppendorf tube on liquid nitrogen or dry ice
5. 0.8ml (800µl) of the modified RLT buffer (in the hood)
6. 0.8 ml (800µl) of chloroform was added (with IAA)
7. Mixing was done by vortexing till all the tissue was dissolved, and stored in ice for 5 sec.
8. Spinning was done for 5 min at 4° C maximum speed (or at room temperature at 12000 rpm for 5min).
9. The upper phase was transferred to a 2ml new tube (at room temperature)
10. The chloroform step was repeated.

11. 500µl of 100% ethanol was added to the upper phase and mixed by pipetting.
12. 550 µl of the sample including the precipitate have formed a RNaesy spin column (pink) was transfered into a 2ml collection tube supplied).
13. The sample was spinned for 1 min at 8000 rpm at room temperature Twice
14. After spinning, the supernatant was washed with 500µl RWI one and centrifuge for 1 Min at rm temp.
15. 500µl 80% ethanol (RPE) was added once and centrifuged at 8000 rpm at rm. temp. for 1 min
16. extra ethanol (RPE) was pipetted out with a tip, spinned 1 more time for 2min at 13000 rpm. The column was transfered to a 1.5ml Eppendorf tube.
17. The column was eluted with 55µl of DEPC treated water (55µl same twice).

Appendix II. Specific RNA-Seq Library Construction Protocol (Silin et al., 2014)

Poly A RNA isolation and fragmentation (1hour).

1. Binding buffer (1x and 2X) was prepared
2. Appropriate amount of oligodT25 Dynabeads (15µl for 10µg, 10µl for < 2µg sample) was prepared.

	1 x10µg RNA	4 samples	8 samples	12 samples
Beads	15µl	60µl	120µl	180µl

PCR tubes with Dynabeads were placed on the magnetic rack for 1-2 min, The storage solution was removed, take the tube away from the magnet, 150µ l of 1 x binding buffer (1ml 2x binding buffer (with mercaptoethanol) + 1ml DEPC treated H_2O), was added and mixed well by pipetting.

3. The wash one more time was repeated.
4. The beads on the magnet were recovered. Re-suspended with appropriate volume of 2 x binding buffer A (50µl per sample), and transfered to a 1.5 ml tube.

	1 x10 µg RNA	4 samples	8 Samples	12 samples
2 x binding buffer A	50µl	100µl	400µl	600µl

5. 50µl of beads was aliquoted to each PCR tubes and then 50µl was addeded to total RNA (5-10µg).

(Mix 8-10 times)

6. The tubes were closed, heated to 65°C for 1 min on a thermocycler with heated lid and incubated at room temperature for 10min with occasional shaking.
7. PCR strip was briefly spinned to collect the beads on the cap, placed on magnetic stand and remove / discard the solution.
8. The strip was taken away from the magnet and 150µl of wash buffer B (well mixed to break the bead clumps), was added and the wash repeated.
9. mRNA was eluted by adding 50µl TE (add 1% β- ME to TE) and incubated at 70°C for 1 min. The PCR strip immediately placed back on ice.
10. 150µl 2 X Binding Buffer A was added to each well (so the final concentration is 1 x Binding buffer). This was mixed well by pipetting.
11. The tubes were closed, heated to 65°C for 1 min on a thermocycler with a heated lid and incubated at RT for 10 min with occasional shaking. Followed by a snap spin. **(**meanwhile, step 14 was prepared).
12. 150µl Buffer B (without β mercaptoethanol) was used to wash 2 times as step 9.
13. The strip was placed on the magnet for 1-2 min, all the solution was removed carefully! The beads were re –suspend in 10µl of superscript buffer.

	1 x	4x	8 x	12x
5 x first- strand buffer	4µl	16	32µl	48µl
Hexamer (1µg/µl	0.5µl	2	4µl	6µl
Oligo dTVN (100ng/µl	0.5µl	2	12µl	6µl
Water	5µl	20	40µl	60µl

14. This was incubated at 94°C (with heated lid) for exactly 5 min to fragment the mRNA,and immediately it was placed on ice. (meanwhile, 1st strand was prepared).
15. The strip was snap spinned, placed on magnetic stand, and the solution containing fragmented mRNA transfered to a new strip. Note: stopping point, eluted RNA can be stored at – 80°C.

First – strand cDNA synthesis

16. The ReverseTranscription (RT) reaction was assembled and mixed on ice and add to each tube

	1X	4x	8X	12X
Water	6 µl	24	48 µl	72 µl
dNTP(10mM)	1 µl	4	8 µl	12 µl
DTT (100Mm)	2 µl	8	16 µl	24 µl
Rnase inhibitor	0.5 µl	2	4 µl	6 µl
NEB protoScript II	0.5 µl	2	4 µl	6 µl
	Ix by pipetting			

17. RT reaction was Performed as follows :

    25°C 10min

    42°C 50min

18. 36µl RNA Clean XP was immediately added to each tube and incubated the mixture on ice for 15 min. *Note: the solution is very viscous; pipetted up and down at least 10 times to mix*. **(** meanwhile 2nd strand was prepared*)*
19. The SPRI-beads were collected on magnetic stand for 3-5 minutes
20. The beads were washed twice with 150 µl 75% ethanol (EtOH) without disturbing the beads.
21. The beads were air – dried for 1 min and RNA/cDNA hybrid was eluted with 10µl H_2O

This was mixed 8- 10 times with the micropipette and for 2 min

Second – strand synthesis with dUTP

22. The 2nd strand reaction master mix was prepared on ice as follows: very important to set this reaction up on ice. Place on ice immediately after reaction also.

	1 x	**4x**	**8x**	**12x**
Water	2.4µl	9.6	19.2µl	28.8µl
dUTPmix (10mM)	0.5µl	2.0	4µl	6µl
10 x Blue Buffer	1.5µl	6.0	12µl	18µl
RNase H(5U/µl)	0.1µl	0.4	0.8µl	1.2µl
DNA Pol1(10 U/µl)	0.5µl	2.0	4µl	6µl

23. 5µl of the master mix was added to each 10µl RNA /cDNA, snap spinned. Incubate at 16°C for 2.5 hours. Note: the completed 2nd strand reaction can be held in the PCR machine at 4°C overnight**.**

24. dsDNA was Purified using 1.8 (27µl) volumes of Sera- mag beads (AMPure XP) and incubated for 15 min at RT , washed twice with 150µl 75% ethanol, air dried for 1min, and eluted with 10µL of H_2O incubated for 1 min at RT.

25. End- Repair

26. An appropriate amount of the end-repair mastermix was prepared on ice as follows: *very important reaction, was set this up on ice.*

	1 X	**4x**	**8 X**	**12 X**
Water	2.75µl	7.0µl	22µl	33µl
dNTP mix (10mM)	0.5µl	2.0µl	4µl	6µl
10 x End –Repair buffer	1.5µl	6.0µl	12µl	18µl
End repair enzyme LC	0.25µl	1.0µl	2µl	3µl

27. 5µl of the mastermix was added to 10µl of the dsDNA snap spinned, incubated at 22°C for 30min, snap spinned. Meanwhile Da-tailing and ligation buffers but no enzymes were taken out the from the fridge.
28. The dsDNA was Purified using 1.8 volumes (27µl) Sera- Mag beads (AMPure XP) incubated for 15 minutes at room temperature, washed twice

with 150µl 75% ethanol, air dried fro 30s-1min and eluted with 10µl H_2O incubated for 1 min at room temperature.

dA – tailing

29. An appropriate amount of the mastermix on ice was prepared as follows:

	1 x	**4x**	**8 x**	**12 x**
Water	2.5µl	10µl	20µl	30µl
dATP mix (10mM)	0.5µl	2.0µl	4µl	6µl
10x Blue Buffer	1.5µl	6.0µl	12µl	18µl
Klenow exo-	0.5µl	2.0µl	4µl	6µl

30. 5µl of the mastermix was added to each sample and incubated at 37°C for 30min.
31. dsDNAwas purified using 1.8 volumes (27µl) of (sera mag) AMPure XP beads, incubated 15 minute at room temperature, washed twice with 150µl 75% ethanol air dried for 1min, eluted with 9.0µ l of H_2O

Tru-Seq. adapter ligation

32. The mastermix was prepared on ice as follows

	1 x	**4x**	**8 x**	**12 x**
2 x Rapid ligation buffer	10µl	40µl	80µl	120µl
TruSeq adapter (2.5µM)	0.5µl	2.0µl	4µl	6µl
T4 Ligase (NEB)	0.5µl	2.0µl	4µl	6µl

Ligation program

33. 11.5µl of the mastermix was added to each well. the mastermix was pipetted up and down multiple times to make a viscous solution and incubated at RT (or 22°C on PCR) for 15min (kept on ice for a while after 15 min). *Note: stopping point, the ligated library can now be stored at -20°C*

Site selection and UDG digestion

34. UDG digestion buffer was prepared as follows:

	1 x	4x	**8x**	**12x**
TE (Ph 7.5)	14.5µl	**58µl**	116µl	174µl
UNG	0.5µl	**2.0µl**	4µl	6µl

35. 1.4 volumes (28µl) of (AMPure XP) Sera- mag beads was added and incubated 10min at RT to purify the ligation product, this was eluted with 15µl TE supplemented with 0.5µl uracil DNA Glycosylase, and incubated at 37°C for 15min snap spin

36. 1.4 volume (21.7µl) of (AMPure XP) Sera-mag beads was added, incubated 10min at RT to purify the ligation product, the DNA was eluted with 10µl water. *the size fractionated DNA was stored at -20°C.*

PCR Enrichment

37. PCR reaction was prepared as follows: **4x**

- UDG digested DNA 5.0µl
- Index Primer mix (10M each) 1.0µl
- 5 x buffer 5.0µl 20.0µl

• 10mM dNTP	0.5µl	2.0µl
• H_2O	13.0µl	52µl
• Herculase / Phusion	0.5µl	2.0µl

38. Initial denaturation was done at 94°C for 2 min, followed by 12-16 cycles of amplification (98°C for 30s, 65°C for 30s, 72°C for 20s). Final extensions of 72°C for 2min, was followed by a 4°C soak.

39. few libraries were Randomly select a and run 2µl or 2.5µl of their pre-Purified PCR on agarose gel. 1% 3 kb fast running buffer (Lithium acetate at Ph 5.8).

40. 1.4 volumes of (AMPure XP) Sera Mag beads was used to purify the dsDNA,and eluted with 20µl of TE

Mix barcoded libraries for multiplexed sequencing

41. DNA concentration of each library was measured using Quan-IT DNA HS assay kit (single –tube)
42. Equal amount (e.g. 20ng) of each barcode library was combined.
43. The library was concentrated using 1.4 volume of Sera mag (AMPure XP) and eluted with 10µ l TE.

General Procedure for Using Ampure/ Spri Beads

1. An appropriate amount of beads was added to the sample as indicated in the protocol.
2. This was well mixed well by pipetting up and down 10 times.
3. The plate was incubated / tube at RT for 15 min.
4. Placed on the magnet stand for 2-5 minutes to collect the beads .
5. The solution was gently removed without disturbing the bead pellets.
6. 150µl of 75% ethanol was Added keeping the plate on the magnet
7. this was kept for 30 s and the ethanol removed without disturbing the beads
8. 75% ethanol wash was Repeated once more. inspecting the plate carefully to remove any remaining ethanol droplets.
9. The beads were Air-dried for 1-2 minutes and add appropriate amount of water (15µl) or TE to elute the DNA.
10. Placed on the magnet, transfer elute to a new plate/tube.

Appendix III. Biotic stress resistance differentially expressed gene in African tomato

	FUSARIUM	BACTERIAL	C.FLAVUM	VIRUS	FUNGAL	B.CINEREA	PEST	A.CANKER	PATHOGEN	WOUND
DATASET	2.615444	1.180787	1.208642	1.531706	1.601202	0.534656	1.530629	0.603391	0.705568	2.029551
B1	0.810407	0.703223	0.663322	1.143373	1.370747	0.791011	0.377849	0.229718	0.176118	0.996955
C1	0.251431	0.469254	1.573217	1.616106	1.100747	1.299443	0.510833	0.096057	0.477383	1.012993
D1	0.14405	0.413002	0.547248	0.490334	0.983984	1.316178	0.449969	0	0.53744	1.014958
A2	1.586533	1.143826	0.775332	2.569898	0.948774	0.527001	0.605945	0.474734	0.755755	0.782901
B2	0.762131	0.508791	0.524952	0.475586	1.094987	0.961293	0.830542	0.314134	1.054091	0.718911
C2	0.324206	0.532459	0.190914	0.528237	1.089519	1.270171	0.791291	0.266044	0.298006	0.797271
D2	0.168445	0.685586	0.5225	0.495162	0.990982	1.229976	0.842011	0.429867	0.799084	0.802948
A4	0.973518	0.973004	0.295912	1.347498	0.444963	0	0	0.446541	0.056576	1.317063
B4	0.400975	0.275293	0.22589	0.744762	0.525447	0.191885	0	0.112622	0.195843	0.598283
C4	0	0.218796	0.245094	0.64759	0.574169	0.398862	0	0.218796	0.354563	0.690727
D4	0.529111	0.484614	0.268261	1.194863	0.719184	0.591143	0	0.227423	0.87473	0.868651
A5	2.101232	1.085742	0.712355	1.61372	1.603165	1.072897	1.225786	0.563966	1.033024	2.739696
B5	0.999676	0.774992	0.669667	1.02076	1.260187	1.486615	1.141864	0.183774	0.841043	1.692745
C5	0.493758	0.51214	0.835429	1.166219	1.221168	1.759482	1.569973	0.175922	1.495161	1.837102
D5	0.170285	0.438558	0.882662	1.219062	0.784749	1.583986	2.131199	0.041583	2.073793	1.949972
A6	2.463806	1.013435	0.621771	2.561967	1.994931	0.940737	0.766289	0	1.11E-16	0.97852
B6	0.805899	0.263401	0.507051	1.115192	1.585599	1.824807	0.477294	0	0.374869	1.194475
C6	0.43317	0.371903	0.549765	0.670539	1.317706	1.613637	0.312235	0	0.870384	0.990079
D6	0.143595	0.325179	0.384955	0.656012	1.358951	1.904116	0.385132	0	0.662908	0.944054
A7	2.044757	2.618101	0.937186	4.699999	2.360498	1.155086	2.673786	1.850754	0.587367	2.879566
B7	2.143294	1.224798	0.711176	3.5408	2.112914	1.71617	1.843872	0.853086	0.754639	1.846481
C7	0.265792	1.177309	0.821082	4.512181	1.906321	2.173872	1.781493	0.755196	1.527309	1.979409
D7	0.275781	0.831599	0.618278	2.255503	1.812477	2.033639	1.706363	0.51527	1.712647	2.050699

A8	1.430189	1.792223	0.811008	4.075422	1.972521	1.177057	1.933781	1.023952	0.97628	2.900523
	Fusarium	**bacterial**	**C.flavum**	**Virus**	**Fungal**	**B.Cinerea**	**pest**	**A.canker**	**Pathogen**	**wound**
B8	B16	0.308836	0.779212	0.859124	1.587392	1.329996	1.654618	1.392668	0.493612	0.847984
C8	0.263506	0.714663	0.622341	2.404546	1.750183	2.024171	1.491957	0.714663	0.960001	2.221804
D8	0.386753	0.56668	1.018277	1.966107	1.601313	2.156958	1.493277	0.806228	1.813229	2.340126
A9	2.156266	1.661942	0.747997	3.174214	1.738036	0.846019	2.603269	0.80172	1.539711	1.879486
B9	0.365344	0.301893	0.892212	1.693976	1.152461	1.715113	1.517523	0.018089	0.720249	1.978862
C9	0.279823	0.422641	0.927272	1.969405	1.042281	1.788868	1.594795	0.095653	0.794238	1.964126
D9	0.134874	0.416722	0.507748	1.198441	1.261949	1.497185	1.357387	0	0.466341	1.934492
A10	2.178108	1.497488	0.809987	2.486533	1.926546	1.003902	3.306976	0.880812	0.751937	1.717068
B10	1.038301	1.225677	0.714462	1.273014	1.585878	1.493663	1.343005	0.569877	1.341253	1.008243
C10	0.285916	0.714834	0.692629	0.87836	1.636083	1.77184	1.160673	0.35065	0.842724	0.933457
D10	0.254355	0.400021	0.878374	1.293967	1.850327	1.995426	1.16145	0.158667	1.833421	0.863469
A11	1.655264	2.731439	0.845263	1.844406	2.532445	1.376599	2.760462	1.818207	1.679986	2.803746
B11	0.742779	1.106576	0.939023	3.565025	2.255377	2.194526	1.708695	1.074021	1.173501	1.952937
C11	0.279999	1.133874	0.756592	1.94901	1.661385	2.450796	1.835762	0.87929	1.256298	2.062926
D11	0.125441	0.902048	0.531511	1.777096	1.688776	2.100963	1.740907	0.642711	1.803711	1.920645
A12	1.470188	2.177821	0.935503	4.37466	1.558938	1.453058	1.875249	1.537081	1.30157	3.086561
B12	0.859955	1.135638	0.756874	3.897399	1.694707	2.074775	1.721305	0.902506	1.048311	2.167237
C12	0.279143	0.981117	0.749469	2.731349	1.584004	2.227021	1.717638	0.747343	1.44125	2.263918
D12	0.238606	0.516963	0.645464	2.219543	1.370327	2.36214	1.824832	0.266803	1.50135	2.271263
A13	2.485112	2.196411	0.714458	2.698025	2.667227	1.378079	3.790599	1.667039	1.078045	2.83424
B13	0.63951	1.351605	0.86029	2.885933	2.204867	2.177967	2.155852	1.2393	0.949725	2.163736
C13	0.399947	1.357857	1.419943	3.26868	2.133484	2.413759	2.165539	1.219162	2.012689	2.146743
D13	0.264398	1.003539	1.076856	2.300846	2.132016	2.16205	1.986863	0.896387	1.871541	2.034496
A15	1.579445	2.248266	0.665776	3.891432	1.733856	1.37926	2.37302	1.507862	1.042219	2.906324
B15	0.760752	0.898941	0.577701	2.625433	1.869302	1.911776	1.800948	0.487961	1.165135	1.864146

C15	0.29229	1.002597	0.783746	2.654373	1.895255	2.345032	2.022801	0.594149	1.465886	1.9954
D15	0.174661	0.647003	2.049991	1.976081	1.563545	2.184034	1.903155	0.346939	0.985941	1.986138
	Fusarium		**C.flavum**	**Virus**	**Fungal**	**B.Cinerea**	**pest**	**A.canker**	**Pathogen**	**wound**
A16	2.144347	1.872777	0.729384	2.666343	1.170984	1.085561	1.62117	0.803925	0.539886	1.880871
B16	B16	0.308836	0.779212	0.859124	1.587392	1.329996	1.654618	1.392668	0.493612	0.847984
C16	0.200595	0.69989	1.35852	2.158224	1.361118	1.790462	1.451366	0.53064	0.874	2.004045
D16	0.284522	0.589975	1.102459	1.614392	1.5266	1.80824	1.542605	0.592743	0.758285	1.856431
A17	0.766907	3.134869	0.787486	3.90907	2.514335	1.342645	0.98664	2.202785	0.20694	1.93272
B17	0.679542	1.304971	0.800153	2.716514	2.155368	1.892329	1.230641	1.215092	0.32663	2.185629
C17	0.281788	1.244397	0.662415	2.180005	1.602962	1.992692	1.429018	1.013157	0.694942	2.062871
D17	0.186944	1.208002	0.848429	2.228806	1.732947	1.88757	1.495669	0.992951	0.745759	2.08543
A18	2.322059	1.925618	0.706893	2.071687	1.368643	1.435132	1.535229	1.167583	0.266736	1.458908
B18	0.6644	0.796222	1.194065	1.10797	1.303195	1.924411	1.527804	0.617975	0.61136	1.664517
C18	0.510699	1.344996	1.299928	1.017082	1.332799	1.925506	1.62882	0.817793	1.111895	1.577929
D18	0.305448	0.765124	0.809639	0.932808	1.26104	1.902716	1.392561	0.350586	1.835645	1.468141
A19	1.7301	2.109168	0.628921	3.350628	1.949081	1.535849	3.46282	1.446598	1.135319	2.754056
B19	0.619909	1.600495	0.502522	3.014224	1.591791	2.150817	1.943727	1.190317	1.510421	1.842753
C19	0.293403	1.116842	0.39623	2.015063	1.319138	2.136322	1.809606	0.744711	1.115505	1.802896
D19	0.206985	0.993274	0.530023	2.615342	NA	NA	1.782311	0.593525	NA	1.836598

Appendix IV. Abiotic stress genes differentially expressed in African tomato at different stages

Dataset	Ascorbic	Calmodulin	Chilling	Cold	Drought	Environmental	Fe	h2o2	Heat_shock	Heat_stress	Light	mechanical	Nutrition	Osmotic	oxidative	Ozone	Phosphate	Sulfhate	salt
A1	0.3461	0.2374	1.2547	0.1699	4.5340	0	0	0.2531	0.9528	0.9528	1	0	0.307	0	0.184	0.29	0	0.65	1.35
	74033	94212	01363	94554	44817			04189	27048	27048			837		875	7287		9708	6823
A2	0.3193	0.2158	1	0.1859	2.7383	0	0	0.1672	1.2479	1.4775	0	0	0.258	0.23	0.195	0	0.338	0	1.83
	1646	20881		17142	64688			16951	59099	96772			873	2712	64		345		703
A4	0.3131	0.2038	1.0308	0.1936	2.2238	0	0	0.0386	0.9372	0.9372	0	0	0	0.19	0	0	0	0	0.65
	11315	59442	45124	36758	47084			78971	29578	29578				0738					7425
A5	0.3432	0.3191	1.3561	0.1686	3.4261	0.241348	0	0.1318	0.7255	1.0732	1.0697	0.1363	0.269	0.49	0.233	0.26	0.429	0	2.41
	65241	76775	80149	52272	02983	619		77073	04236	24303	30191	03071	144	7736	471	8992	869		3237
A6	0.3271	0.3281	0.9555	0.1235	2.9113	0.172719	0.5057	0.1069	0.7953	1.0038	1	0	0.274	0.69	0	0	0	0.54	0.88
	97917	69742	55474	18967	56527	201	12151	47553	92603	40437			473	9883				6978	3718
A7	0.3316	0.3059	2.1409	0.4013	3.9338	0.279441	0.3394	0.1673	1.5466	2.1210	0.9910	0.1752	0.257	0.95	0.219	0.28	0.454	0	2.63
	10451	04155	96872	03002	28037	425	76471	60894	06879	14568	53861	12402	089	5094	113	2423	5		2422
A8	0.4082	0.3081	2.5566	0.4319	4.2657	0.252475	0.7680	0.2290	1.3885	1.9371	0.1879	0.1280	0.251	1.18	0.231	0.26	0.767	0	4.07
	75649	02753	44681	88084	41597	858	58858	96386	36335	06009	46135	0591	877	9362	134	7461	523		1627
A9	0.3734	0.4289	0.9867	0.4038	4.2238	0	0	0.1658	1.2704	1.4960	0.1829	0	0	0.94	0.228	0.27	0.496	0	2.46
	20691	11124	60106	6516	40691			41069	46771	85828	88956			8007	102	8679	592		8202
A10	0.3112	0.3170	0.7427	0.1759	4.2728	0.156530	0.1163	0.0913	0.4198	0.9756	0.7418	0.3561	0.268	0.59	0	0.27	0.603	0	2.54
	63683	86371	91742	65515	7033	227	16826	39558	37541	93071	19304	90731	131	3539		9096	063		7351
A11	0.3384	0.3171	1.7291	0.3746	4.3745	0.123458	0.3563	0.1450	1.5259	2.4976	0.2049	0.0980	0.281	1.60	0.222	0	0.473	0	4.54
	06752	84805	83381	65665	77428	645	48383	12815	26692	07306	668	01687	526	6425	388		027		7725
A12	0.3159	0.3085	2.1631	0.4038	3.9091	0.212837	0	0.1493	1.2102	2.1070	0.2434	0.1216	0.241	0.90	0.225	0.27	1.072	0	3.63
	56816	76205	41627	02312	1014	912		62731	19662	09615	57576	94564	344	0691	301	9965	789		9334
A13	0.3304	0.3115	1.7394	0.3942	3.8655	0.165531	0	0.1909	1.2179	2.0497	1.1611	0.3717	0.257	0.88	0.233	0.28	0.582	0	3.79
	95991	73269	76538	93134	15943	345		8044	62464	43568	48275	13896	14	6842	583	5513	727		9424
A15	0.3281	0.3013	1.7010	0.4630	3.9762	0	0	0.1299	1.5660	2.3264	0.2939	0.3184	0	1.04	0.228	0.28	0.480	0	3.00
	57794	23394	60376	46383	93064			72472	3054	30659	67679	41231		0439	685	1009	852		0263
A16	0.3252	0.3812	2.0197	0.1509	4.6856	0.113203	0	0.1285	1.0786	1.6393	0.1747	0.3272	0.295	0.92	0.195	0.28	0.498	0	2.50
	43088	31231	08046	03396	54108	719		67134	10955	26935	63808	84206	708	414	703	4353	283		537
A17	0.3216	0.2994	2.3058	0.4512	4.0731	0	0	0.1304	1.3633	2.1769	0.2069	0.1048	0.253	0.77	0.225	0.27	0.541	0	2.97
	69941	20202	63254	38523	46545			9541	72512	82005	40417	73525	417	2312	352	4953	939		7153
A18	0.3180	0.2701	1.5348	0.1241	3.2841	0.192469	0.6982	0.1670	1.2403	1.7180	1.0185	0.1341	0.296	0.72	0.212	0.27	0	0	2.40
	97585	67175	98973	42085	90452	387	57389	7405	61898	25988	04884	0406	531	1789	051	9748			9175
A19	0.3248	0.3100	2.6129	0.3849	3.9562	0.114206	0	0.1658	1.0252	1.7821	0.2343	0.3952	0.250	1.35	0.240	0.28	0.444	0	4.40
	38012	62899	56666	32277	94648	276		17709	39028	54948	9567	91486	832	3996	826	1497	875		9324
B1	0.2402	0.2851	0.4035	0.2053	3.3243	0	0	0.1511	0.8977	0.8977	0	0	0.288	0	0.303	0.25	0	0	0.98
	82697	36553	73984	03102	15511			54633	07461	07461			518		72	7168			5023
B2	0.2449	0.2073	0	0.2022	2.5370	0	0	0.2074	1.1232	1.3313	0	0	0.317	0.96	0.314	0	0.275	0	1.56
	10179	3994		44067	44561			11215	25326	55223			873	3695	024		208		0667
B4	0.2544	0.1952	0.1166	0.2137	1.2917	0	0	0.0277	0.8395	1.0140	0	0	0	0.48	0	0	0	0	0.54
	28931	55899	17614	23224	24972			08488	25008	51066				0185					1749

B5	0.2457	0.2590	0.9407	0.2276	2.9522	0.154731	0	0.1198	1.0849	1.3780	0.1046	0.1991	0.301	0.79	0.290	0.27	0.550	0	2.13
	42856	87377	17244	03017	66042	151		25051	83587	48883	50065	13123	001	6435	93	3026	013		1189
B6	0.2589	0.2569	0.2337	0.3123	2.6074	0.297022	0.3787	0.1539	0.8488	1.0880	0	0	0.293	1.13	0	0	0	0.24	1.35
	11768	33054	44382	81064	92302	482	88644	12204	94565	85459			327	8892				0885	5141
B7	0.2164	0.2709	0.8786	0.4337	2.5798	0.215635	0.4874	0.1944	1.5988	2.2721	0.2373	0.1776	0.314	1.25	0.293	0.26	0.533	0	2.40
	27139	77164	35124	31785	05406	842	21902	95403	46012	65542	94376	09214	932	2205	021	9056	664		3247
B8	0.2062	0.2703	1.8694	0.4851	3.4424	0.280865	0.1326	0.3313	1.5719	2.3001	0.0834	0.2524	0.303	1.05	0.285	0.25	0.520	0	3.57
	13444	33611	45117	87118	14778	256	84799	53277	39469	32664	06615	75944	34	7666	927	522	16		4016
B9	0.2939	0.2021	0.6232	0.5903	2.4756	0	0	0.2188	2.0364	2.5347	0.2754	0	0	1.16	0.273	0.27	0.506	0	2.47
	75449	31748	4633	04941	3855			44448	717	41371	58223			3607	452	3979	709		6227
B10	0.2618	0.2815	0.5576	0.2216	3.2532	0.242355	0.7436	0.1365	0.6633	1.1285	0.4889	0.4340	0.316	1.08	0	0.26	0.550	0	2.43
	21554	1938	52114	18165	56292	562	15305	18986	95779	15501	17917	47464	184	2867		6595	521		3666
B11	0.2791	0.2670	1.4943	0.5126	3.6624	0.280032	0.4952	0.1764	1.5428	2.7091	0.0963	0.2460	0.311	2.26	0.294	0	0.519	0	3.93
	62057	55379	11819	88879	32155	553	91254	73062	46134	79276	40336	6883	371	2297	959		932		0962
B12	0.2560	0.2700	0.7256	0.4774	3.2138	0.235342	0	0.2166	1.7549	2.8247	0.1101	0.2326	0.302	1.55	0.272	0.26	0.660	0	3.75
	43741	93806	8497	45707	47928	651		20759	67457	91826	75903	49132	844	7201	054	9998	171		7506
B13	0.2462	0.2609	1.6907	0.5153	3.9384	0.361042	0	0.3052	1.9059	2.8674	0.1197	0.6063	0.289	1.49	0.268	0.25	0.536	0	3.96
	29544	44785	00803	61018	47344	259		5191	26757	2626	72988	68655	351	3449	928	6419	786		1158
B15	0.2641	0.2700	1.6827	0.4603	4.5881	0	0	0.2269	2.1060	2.9027	0.1086	0.4895	0	1.66	0.266	0.25	0.596	0	3.41
	80532	82285	78335	42444	81998			99949	80276	49431	90931	38725		4014	068	501	26		6224
B16	0.2726	0.1953	0.8836	0.3029	2.3261	0.286410	0	0.2019	1.9787	2.6852	0.3271	0.5011	0.253	1.38	0.270	0.28	0.519	0	2.95
	33191	6171	91714	06497	1363	432		53912	96305	1908	79339	98401	23	7223	136	1239	8		8155
B17	0.2437	0.2779	1.4074	0.5145	3.6260	0	0	0.1948	1.6958	2.6712	0.1433	0.2648	0.294	1.41	0.262	0.27	0.255	0	3.42
	59923	71027	97224	78118	20958			62623	88338	72331	84615	8316	241	1113	486	8989	987		9404
B18	0.2570	0.2866	1.0449	0.3031	3.3798	0.271024	0.1941	0.2052	1.7960	2.5608	0.2263	0.2773	0.286	1.55	0.266	0.26	0	0	2.72
	64728	52553	23204	64327	17766	789	71564	9887	50172	56981	28735	42494	136	1048	582	925			7567
B19	0.2534	0.2794	1.3311	0.5033	3.6848	0.390968	0	0.1876	1.3217	2.0817	0.2220	0.4398	0.309	1.76	0.275	0.27	0.595	0	3.71
	73364	19787	49196	95574	98267	626		32852	92718	91863	95012	66877	556	2524	861	1634	734		5565
C1	0.2367	0.2078	0.4326	0.3265	2.2677	0	0	0.1966	1.1793	1.1793	0	0	0.237	0	0.269	0.24	0	0	0.87
	261	88122	38677	32321	10872			0353	16227	16227			402		041	2911			6528
C2	0.2541	0.1565	0	0.3380	2.1588	0	0	0.1881	1.9814	2.6766	0	0	0.243	0.39	0.258	0	0.191	0	1.47
	63546	73675		78951	75798			77664	63	08522			433	1141	67		112		1452
C4	0.2388	0.1548	0	0.3308	1.1041	0	0	0.0304	1.1999	1.5536	0	0	0	0.28	0	0	0	0	0.60
	77358	45435		93141	08986			70081	97694	08449				2305					5128
C5	0.2443	0.2149	0.9308	0.3255	2.8608	0.264520	0	0.1099	1.7848	2.3812	0.4468	0.3132	0.252	0.83	0.258	0.25	0.546	0	2.39
	00607	45301	25317	6315	29672	794		06371	13542	44953	56056	48701	223	9301	428	9488	848		7184
C6	0.2599	0.2211	0.1129	0.2987	1.8919	0.283479	1.11E-	0.1124	1.2736	1.5639	0	0	0.257	0.95	0	0	0	0.15	1.04
	94208	93952	92013	54171	21408	5	16	1651	17077	16342			247	9198				6017	3091
C7	0.2245	0.2036	0.7485	0.6028	2.2873	0.190139	0	0.1645	2.3922	3.3719	0.3061	0.3179	0.243	1.20	0.259	0.23	0.490	0	2.40
	68102	96872	15733	53658	48772	235		21507	47909	0587	38909	65133	924	6507	933	5325	452		7743
C8	0.1086	0.2177	1.3383	0.6208	2.5280	0.161185	0	0.1760	1.9667	2.8160	0.3182	0.3214	0.246	0.89	0.268	0.20	0.532	0	3.43
	97044	93836	21772	67142	21711	506		17552	11145	94864	19603	1843	331	4092	956	5634	285		814
C9	0.2678	0.2014	0.5155	0.5731	2.5288	0	0	0.1840	2.2923	2.9280	0.2777	0	0	1.17	0.281	0.27	0.568	0	2.44
	56433	01306	90823	68339	86324			03153	46737	44909	34421			4933	01	7667	164		2138

C10	0.2531	0.2114	0.2132	0.3233	2.5604	0.315135	0.0608	0.1059	0.8571	1.3613	0.3385	0.5712	0.224	0.92	0	0.25	0.414	0	2.10
	75682	90064	98896	08234	55346	472	26628	66284	60301	87319	0056	39035	881	0424		5079	166		5641
C11	0.2274	0.2072	0.9370	0.6298	2.7315	0.318559	0.0700	0.2456	2.6970	4.1311	0.3661	0.3583	0.243	1.25	0.263	0	0.563	0	3.96
	42334	56078	97449	97105	94239	501	88339	40563	06756	70349	52049	11909	609	2554	693		228		6317
C12	0.2519	0.2079	0.4769	0.6078	2.8142	0.332634	0	0.2141	2.7423	4.0461	0.3028	0.3173	0.260	1.11	0.261	0.26	0.419	0	3.68
	56864	35908	62481	01634	16281	51		49045	87345	75457	83194	28605	426	9571	506	0144	604		9504
C13	0.2377	0.2249	1.3551	0.6097	3.3071	0.271150	0	0.2828	2.4804	4.0160	0.3541	0.5299	0.267	1.08	0.272	0.26	1.204	0	4.01
	49646	96624	66891	09821	62518	228		77	47307	48408	91752	07065	91	1305	28	0056	758		7852
C15	0.2501	0.2218	1.2165	0.6235	2.6867	0	0	0.1808	2.8368	3.5909	0.3738	0.5333	0	1.26	0.267	0.25	0.546	0	3.12
	94361	069	98659	33787	3919			11118	27985	88047	48709	37217		9253	79	9322	515		4244
C16	0.2343	0.2109	0.8234	0.2697	2.5757	0.332739	0	0.2263	1.8411	2.6511	0.3564	0.5647	0.254	1.36	0.264	0.28	0.461	0	2.97
	19199	57557	47117	95599	1548	561		14312	84632	33221	42939	563	871	5594	221	1097	944		3946
C17	0.2458	0.2063	0.8862	0.5495	2.8421	0	0	0.2184	2.7331	3.8098	0.3146	0.3258	0.240	1.34	0.258	0.24	0.564	0	3.08
	47409	23981	58789	36222	44022			79824	93598	51273	63887	45889	536	6357	116	1944	822		0619
C18	0.2496	0.2430	1.5592	0.3078	3.2237	0.317648	0	0.2100	2.2047	3.0424	0.3282	0.3055	0.252	1.25	0.277	0.26	0	0	2.50
	34045	72602	00928	97315	68186	809		18599	67296	69249	64953	96583	969	2338	358	4047			0034
C19	0.2582	0.2050	1.2088	0.6254	2.9291	0.304077	0	0.2469	2.5534	3.3171	0.2726	0.6734	0.236	1.49	0.262	0.23	0.495	0	3.92
	33106	7614	50558	52749	54099	555		26879	5735	41159	00214	7471	621	322	124	6994	106		7016
D1	0.2075	0.2105	0.3803	0.3230	1.8008	0	0	0.1469	1.0911	1.0911	0	0	0.198	0	0.274	0.23	0	0.22	1.14
	5516	84679	40424	5695	74463			75961	75962	75962			532		144	8167		982	7413
D2	0.2150	0.1666	0	0.3006	1.7145	0	0	0.2106	1.7134	2.4976	0	0	0.213	0.82	0.264	0	0.224	0	1.51
	85497	30251		89642	28246			9436	75744	73902			26	7038	142		948		2781
D4	0.2279	0.2084	0.6759	0.2911	1.3930	0	0	0.0324	1.1370	1.5619	0	0	0	0	0	0	0	0	0.60
	17771	3021	81992	13769	15095			39094	98634	29761									0429
D5	0.1971	0.2097	0.8223	0.3052	2.9701	0.363064	0	0.1531	1.4803	2.2500	0.3013	0.3675	0.211	0.87	0.252	0.23	0.532	0	2.68
	2663	79459	30286	05706	47903	541		44928	91192	59709	40098	9086	917	5523	944	4727	784		5665
D6	0.1845	0.1976	0.1261	0.2856	1.6162	0.277098	1.11E-	0.1395	1.1497	1.4453	0	0	0.209	0.64	0	0	0	0	1.16
	01262	84404	1166	62612	15125	802	16	1292	4191	81562			329	813					7046
D7	0.2595	0.2229	0.8177	0.6188	2.2539	0.343364	0.1085	0.2420	2.5851	3.3859	0.3773	0.3506	0.217	1.16	0.263	0.24	0.557	0	2.45
	40738	47678	05614	62052	5104	259	67323	86415	74515	44255	20147	5656	94	3952	035	9191	686		2602
D8	0.2846	0.2090	1.2426	0.5260	2.8091	0.332464	0	0.2704	2.1936	3.0935	0.4079	0.3182	0.233	1.32	0.249	0.30	1.111	0	3.89
	44723	37716	4199	62688	54568	845		31416	96579	19389	82765	60105	621	1955	751	5378	276		8685
D9	0.0632	0.1917	0.3907	0.4962	1.8573	0	0	0.2037	2.5454	3.1853	0.2955	0	0	1.27	0.253	0.20	0.472	0	2.22
	37815	73704	59783	24454	88938			63339	10525	03898	79133			9708	192	391	294		096
D10	0.2072	0.1970	0.4350	0.3068	3.0396	0.311978	0	0.1743	1.1018	1.6441	0.3818	0.6847	0.224	0.91	0	0.23	0.466	0	2.47
	777	82856	34488	61579	71043	671		68443	22946	15063	86166	22758	291	6972		5381	927		2163
D11	0.1838	0.2062	0.8332	0.5402	2.2120	0.298079	1.11E-	0.1986	2.3345	3.8332	0.3382	0.3081	0.195	1.29	0.254	0	0.475	0	3.85
	88109	81829	60648	54053	40779	903	16	46507	6389	58554	04929	21802	701	3599	137		908		833
D12	0.2094	0.2144	0.6202	0.5736	2.1131	0.251415	0	0.1898	2.3434	4.1171	0.3575	0.3450	0.230	0.93	0.276	0.22	0.793	0	3.32
	62142	97746	10623	12534	00844	52		63242	46169	78264	11382	90462	403	1295	908	5641	391		0337
D13	0.2182	0.2091	1.3702	0.5428	3.0701	0.227223	0	0.2432	2.4967	4.1635	0.3036	0.5478	0.220	1.07	0.261	0.23	0.661	0	3.95
	17625	17035	30423	04945	73223	583		20564	74565	05195	66955	02841	882	9569	415	3994	008		6986
D15	0.1883	0.2146	1.5132	0.5194	3.6402	0	0	0.2244	2.6055	3.3968	0.2358	0.7011	0	1.59	0.273	0.24	0.405	0	3.06
	44705	64368	89226	44764	2554			6584	2921	8006	79583	44795		9294	458	0945	746		8603

D16	0.1994	0.2170	1.2340	0.3026	2.4788	0.284753	0	0.2046	2.1427	3.0922	0.1595	0.6540	0.231	0.81	0.302	0.18	0.572	0	2.71
	6352	67616	34172	75397	35711	787		44698	92641	58675	16449	13318	722	0904	913	5517	837		6501
D17	0.2223	0.2096	0.9214	0.5513	2.4596	0	0	0.2190	2.3086	3.5311	0.3542	0.3187	0.247	1.00	0.289	0.24	0.610	0	2.77
	28781	62076	69915	55367	42945			98501	51116	91793	80473	3317	58	1299	572	0113	995		2045
D18	0.2084	0.2182	0.8612	0.2853	2.3131	0.250728	0	0.1902	1.9195	2.9002	0.3662	0.3057	0.197	0.99	0.278	0.22	0	0	2.31
	07633	81791	93517	01725	12903	379		63412	10541	72164	4818	54255	41	3746	904	2562			922
D19	0.1952	0.2064	0.9163	0.5462	2.5808	0.200527	0	0.1695	2.1525	2.9781	0.3051	0.5438	0.237	1.49	0.257	0.24	0.510	0	3.61
	46719	80704	70757	48735	81184	28		74578	75779	99138	73689	92697	755	3276	588	5861	387		0538

Appendix V. Biotic stress genes Differential gene expression in African eggplant

Dataset	ASC	BACT	OOMY	VIR	XENO	FUNG	Dataset	ASC	BACT	OOMY	VIR	XENO	FUNG
A1	2222.77	16061.3	479.09	1812.045	368.98	8208.775	**C1**	3061	18051.86	1087.245	12759.16	398.575	13165.9
A3	3721.08	19332.19	797.365	2575.935	357.035	10191.47	**C3**	4245.19	22882.12	3593.535	8782.035	588.415	22935.91
A4	3487.29	21782.9	623.165	2040.345	447.645	10431.86	**C4**	2576.535	15903.56	958.67	5671.21	343.305	9900.805
A6	3554.675	18568.45	822.19	3644.47	475.54	10819.88	**C6**	3865.705	20114.77	2821.455	8382.845	705.485	18539.87
A10	4346.775	15965.08	822.19	1833.485	497.16	9310.4	**C10**	2501.955	16394.8	725.195	7068.185	292.375	10727.55
A13	5001.925	17076.72	684.36	1764.3	456.08	9070.735	**C13**	1917.19	17536.18	904.63	7493.805	370.985	11124.8
A14	7811.405	17259.78	779.325	1931.585	635.67	10204.82	**C14**	2538.17	17608.2	771.14	6243.11	468.3	10413.43
A17	2770.73	16215.66	609.29	2332.7	486.145	10001.93	**C17**	2006.595	18533.2	1121.04	6709.26	453.71	13022.7
A23	3893.375	19306.2	818.545	2284.845	524.635	11930.71	**C23**	1672.31	18125.31	1260.895	5953.04	385.69	14696.53
A28	3686.055	18401.04	1084.385	4077.685	505.06	11762.19	**C28**	5397.775	21485.3	2401.095	8664.815	658.08	19042.66
B1	224.88	17714.38	828.515	8762.09	362.85	11611.18	**D1**	4838.435	45214.59	1315.065	12586.11	513.765	41828.56
B3	2452.74	15410.7	1238.855	8336.595	369.595	15193.95	**D3**	5003.68	22342.92	3527.42	8005.78	746.885	18769.14
B4	2924.48	16116.12	844.915	7398.61	349.375	10265.03	**D4**	3150.72	50349.35	1976.355	5822.785	348.365	46016.1
B6	2798.34	20636.6	1314.17	10612.7	546.355	19689.17	**D6**	3773.61	20292.1	2486.265	9658.58	632.955	19359.87
B10	2077.24	14636.93	662.875	6079.54	317.795	8469.625	**D10**	3354.72	56071.58	1012.17	10161.29	307.205	53527.82
B13	2086.375	16483.37	836.385	6537.2	400.47	10378.22	**D13**	4325.795	37681.25	2187.045	6339.295	345.42	36746.59
B14	2307.625	18531.41	825.775	6480.585	428.575	11492.69	**D14**	4796.37	20960.89	1165.075	8985.105	419.095	16213.89
B17	1585.87	19153.39	1096.36	6158.54	455	13444.01	**D17**	3507.14	20036.26	885.115	10025.2	308.585	15652.14
B23	1334.77	17357.15	993.12	6296.265	316.195	10988.48	**D23**	2348.295	21265.05	1525.225	5203.55	357.48	17004.98
B28	2621.47	18020.82	1228.375	9582.42	495.66	17811.26	**D28**	3217.39	19593.63	1286.1	9322.725	523.865	18866.6

Appendix VI. Abiotic stress resistance genes in African eggplant at different stages

Dataset	ASC	OSM	DRGT	Salt	HSH	HS	LI	H2O2	AB	ETHB	CYT	AUX	CAL	SA	JA	ALD	SULF	STR
A1	2222.77	3874.25	72.76	3874.25	2076.8	82.52	4682.205	8487.955	2115.11	1116.56	15627.17	3950.845	1331.355	2660.72	3235.71	3999.37	314.595	0.5
A3	3721.08	3982.615	93.45	3982.615	1486.275	100.605	5298.285	8707.48	2073.555	1370.465	19817.87	4687.115	1691.58	3135.325	3540.14	5955.91	402.735	9.635
A4	3487.29	4012.615	86.295	4012.615	2276.035	104.495	5319.345	11606.66	2359.42	1937.415	20581.05	4393.525	1620.86	30400.08	3357.41	5501.355	420.715	0
A6	3554.675	4886.715	121.655	4886.715	1810.64	196.04	3941.51	8186.08	2745.82	1409.31	15129.83	5778.315	1550.855	3842.32	4326.54	4779.655	268.645	9.315
A10	4346.775	3647.615	7848	3647.615	1524.62	52.54	4416.84	6525.985	2206.27	1029.465	14498.86	4272.505	1250.96	2835.225	3835.79	3514.175	253.535	23.635
A13	5001.925	2916.95	98.875	2916.95	1491.95	58.67	5513.85	666.225	1789.935	879.215	16461.39	4225.69	1299.08	3477.18	3776.095	4707.89	248.605	18.795
A14	7811.405	3641.615	123.535	3641.615	1775.99	68.685	3579.89	6937.3	1741.66	1072.69	14896.1	4746.995	1498.37	3892.985	4769.845	4346.49	262.995	64.005
A17	2770.73	4474.295	70.355	4474.295	1960.94	64.055	3131.55	7605.455	2761.665	1153.27	13756.43	4456.81	1122.655	2620.71	3644.27	3370.615	175.155	0.505
A23	3893.375	4350.39	95.96	4350.39	2186.615	93.17	4798.085	10045.35	2302.59	1764.35	17100.49	4895.7	1644.98	3603.575	4174.485	4573.855	355.85	5.33
A28	3686.055	5005.755	116.62	5005.755	2198.875	97.11	3910.395	10029.18	2718.38	1541.64	15087.72	5941.76	1674.32	4313.355	4491.505	4469.38	260.12	8.485
B1	224.88	8292.225	119	8292.225	21678.66	265.16	3317.83	8034.47	3803.035	665.72	9583.185	8628.69	2920.88	4388.005	5112.19	2917.425	208.425	0
B3	2452.74	7334.7	109.295	7334.7	8122.03	280.83	2520.28	8204.4	4606.735	758.755	12087.55	8503.885	3155.625	6243.115	4757.23	4156.54	188.345	0
B4	2924.48	7539.565	111.845	7539.565	21156.02	231.8	2880.6	7946.075	3858.47	629.48	9566.72	8192.555	3610.105	3938.17	4956.72	3125.675	283.365	0
B6	2798.34	8698.885	149.42	8698.885	15510.17	257.285	4688.44	10145.75	4586.2	1302.465	13463.43	11168.3	4718.4	7278.75	7357.055	3509.225	243.655	0
B10	2077.24	7148.875	138.955	7148.875	15290.32	259.14	2680.625	7802.22	4564.065	613.09	9030.35	8139.565	2304.83	3572.695	4239.285	3295.145	234.425	0
B13	2086.375	7384.545	147.83	7384.545	15769.51	217.52	3176.24	7911.75	4417.59	838.5	11699.87	9629.23	2711.025	4353.085	5381.995	3713.04	331.885	4.16
B14	2307.625	9534.475	127.075	9534.475	14967.39	236.17	2677.28	9322.145	4660.44	856.515	10484.52	9353.34	2789.955	4429.345	5267.51	3576.185	334.63	0
B17	1585.87	7799.295	127.02	7799.295	14509.18	199.745	2721.2	10786.62	4483.32	807.89	11315.04	10353.39	3680.76	6478.955	6195.205	3517.915	225.825	0
B23	1334.77	9240.83	118.42	9240.83	21707.12	159.185	1844.215	7444.365	4156.615	764.18	10930.24	10783.24	2744.645	4834.365	5535.725	3845.19	341.755	0
B28	2621.47	8035.59	117.17	8035.59	14049.37	231.57	3415.65	9970.06	4270.35	1145.31	11778.93	9297.01	4042.615	6363.015	5595.99	4176.455	195.335	0
C1	3061	6678.54	132.605	6678.54	27367.7	312.6	3424.175	6486.12	4125.03	776.385	9370.28	9509.07	3041.5	4369.785	6398.085	2608.15	242.49	0
C3	4245.19	8606.235	171.015	8606.235	13245.74	250.9	4579.69	11172.91	4568.19	1200.41	15137.13	12251.62	4933.32	10720.6	7800.21	3566.53	358.91	0
C4	2576.535	7830.56	152.815	7830.56	21820.71	183.71	3246.4	6350.555	3708.145	723.145	8932.125	8531.96	3326.97	3702.075	5238.04	3132.1	244.695	0
C6	3865.705	8724.75	211.53	8724.75	13349.61	215.905	4057.865	9411.075	4704.56	1141.945	12498.92	12466.97	3636.465	9010.48	8246.155	3457.915	312.715	0
C10	2501.955	7542.145	130.91	7542.145	20802.44	262.67	3147.42	8057.84	4782.35	588.29	8900.99	8827.77	2370.27	3654.595	4347.48	3509.955	267.055	0

C13	1917.19	8407.29	165.03	8407.29	16976.87	239.69	2645.945	7893.135	5221.385	966.215	10988.99	9824.225	2882.695	4402.295	5669.025	3335.455	345.275	0
C14	2538.17	8684.155	124.745	8684.155	15338.47	207.79	2355.105	9389.575	4704.685	763.175	11214.6	9017.725	3062.545	4205.35	4867.36	3893.145	290.45	0
C17	2006.595	6919.805	141.91	6919.805	13840.06	31197	2721.945	11215.17	4380.855	1026.315	11319.07	10156.38	3398.715	6114.855	6659.63	3704.555	318.365	0
C23	1672.31	8497.37	153.675	8497.37	21639.17	224.54	4403.695	6894.435	5350.535	814.98	9419.885	9412.6	3751.445	4700.655	5759.545	3152.175	212.695	0
C28	5397.775	8585.83	169.385	8585.83	12344.18	163.555	3582.04	10160.98	5066.3	2323.405	12904.46	11059.63	3598.875	7896.015	7673.38	3502.96	292.495	0
D1	4838.435	6441.545	144.405	6441.545	16452	270.01	3651.305	7145.565	4848.56	2247.005	8927.39	8212.39	3475.735	4450.155	6661.575	2549.955	235.68	0
D3	5003.68	7592.935	232.59	7592.935	10198.68	236.99	3806.59	12467.2	4240.02	989.265	16686.23	11927.3	3073.72	8315.42	7116.895	3359.78	478.5	0
D4	3150.72	11582.53	171.22	11582.53	18269.77	173.395	4440.5	7486.82	5718.17	1420.16	9308.725	8561.29	3896.98	5350.97	6856.97	3059.78	404.415	0
D6	3773.61	8791.725	225.19	8791725	14935.4	173.395	4893.575	10976.83	4766.35	843.185	13022.08	12356.86	3979.345	7840.86	7296.365	3570.015	267.4	0
D10	3354.72	7052.67	166.885	7052.67	21339.51	192.95	4015.75	7419.445	4241.375	1133.485	8542.21	8556.14	3461.985	4817.485	5460.3	2911.815	281.62	0
D13	4325.795	13694.12	225.205	13694.12	8103.155	206.215	3591.075	7431.075	8101	2944.38	10108.89	8829.98	3139.515	6348.03	9268.705	2647.74	407.21	0.105
D14	4796.37	8303.27	210.125	8303.27	15697.27	262.285	3105.705	9171.175	6105.325	2755.165	10541.25	8507.9	2915.785	4820.76	7150.535	3241.75	381.795	0
D17	3507.14	5739.42	169.37	5739.42	28552.36	253.17	4836.87	9928.065	4157.41	823.28	9267.82	10008.43	3031.555	4431.51	5761.91	3461.9	351.72	0
D23	2348.295	7562.6	164.305	7562.6	18761.68	213.14	3493.98	5411.735	4819.615	1329.615	9520.26	9227.845	2889.44	4684.53	5980.19	2278.675	311.89	0
D28	3217.39	8584.54	127.83	8584.54	15420.59	189.205	4375.745	10310.78	4722.72	1872.07	13310.41	10405.56	4638.655	7328.635	6810.33	3498.43	218.25	0

Appendix VII: SNPs mined from the African tomato

Sample	SNPs	SNP (TI/TV)	SNP (het/hom)	Insertions	deletions	Same as reference
A1	23,156	1.34	1.18	4,841	902	59,134
A2	23,540	1.36	1.36	4,480	881	58,306
A4	26,072	1.37	1.37	4,872	984	56,105
A5	27,053	1.39	1.53	4,857	925	55,198
A6	24,524	1.34	1.37	4,557	908	58,044
A7	24,781	1.38	1.36	4,666	947	57,639
A8	24,577	1.36	1.36	4,849	969	57,638
A9	24,822	1.37	1.35	4,790	951	57,470
A10	26,864	1.41	1.02	5,383	1,186	54,600
A11	23,711	1.36	1.24	4,661	900	58,761
A12	23,422	1.34	1.29	4,625	877	59,109
A13	23,316	1.35	1.25	4,684	882	59,151
A15	25,312	1.37	1.48	4,587	916	57,218
A16	24,256	1.36	1.36	4,558	913	58,306
A17	24,067	1.35	1.30	4,641	903	58,422
A18	34,166	1.44	1.19	5,175	1,276	47,416
A19	23,165	1.34	1.25	4,820	886	59,162
B1	23,425	1.35	1.28	4,738	913	58,957
B2	24,474	1.35	1.34	4,739	900	57,920
B4	25,635	1.37	1.51	4,666	950	56,782
B5	23,550	1.35	1.26	4,801	898	58,784
B6	25,665	1.37	1.35	4,799	979	56,590
B7	24,287	1.36	1.28	4,780	908	58,058
B8	25,460	1.37	1.40	4,788	958	56,827
B9	25,115	1.36	1.40	4,754	956	57,208
B10	27,151	1.38	1.50	4,809	1,017	55,056
B11	23,820	1.36	1.25	4,768	912	58,533
B12	23,667	1.34	1.34	4,615	899	58,852
B13	23,649	1.34	1.29	4,645	898	58,841
B15	24,125	1.35	1.34	4,760	929	58,219
B16	26,881	1.38	1.51	4,712	960	55,480
B17	24,360	1.35	1.32	4,719	926	58,028
B18	31,859	1.42	1.77	5,049	1,136	49,989
B19	23,577	1.34	1.31	4,689	888	58,879
C1	23,448	1.34	1.29	4,610	888	59,087
C2	24,021	1.34	1.32	4,719	912	58,381
C4	24,856	1.36	1.45	4,665	944	57,568
C5	23,537	1.34	1.26	4,743	907	58,846
C6	25,570	1.35	1.40	4,708	938	56,817
C7	23,801	1.36	1.25	4,772	922	58,538
C8	25,942	1.37	1.49	4,695	946	56,450
C9	24,909	1.36	1.40	4,681	931	57,512
C10	24,954	1.36	1.33	4,698	965	57,416
C11	23,738	1.35	1.27	4,677	897	58,721
C12	23,855	1.34	1.35	4,652	888	58,638
C13	23,587	1.34	1.30	4,662	882	58,902
C15	24,095	1.35	1.38	4,662	904	58,372
C16	25,105	1.36	1.36	4,696	937	57,295
C17	24,267	1.35	1.34	4,656	901	58,209
C18	30,277	1.40	1.50	4,916	1,108	51,732
C19	25,372	1.36	1.38	5,023	905	56,733
D1	23,528	1.34	1.29	4,678	882	58,945
D2	24,104	1.34	1.35	4,627	917	58,385
D4	26,757	1.37	1.59	4,676	965	55,635
D5	23,703	1.34	1.26	4,679	894	58,757
D6	25,556	1.36	1.39	4,783	955	56,739
D7	24,127	1.36	1.31	4,677	897	58,332
D8	25,000	1.37	1.36	4,686	984	57,399
D9	25,304	1.36	1.43	4,696	938	57,095
D10	24,866	1.36	1.33	4,697	948	57,522
D11	23,774	1.35	1.28	4,699	909	58,651
D12	23,698	1.35	1.33	4,799	914	58,622
D13	23,292	1.35	1.25	4,841	895	59,005
D15	29,276	1.39	1.65	4,842	1,067	52,908
D16	25,595	1.37	1.40	4,811	959	56,668
D17	24,391	1.35	1.36	4,664	922	58,056
D18	30,373	1.41	1.77	4,959	1,094	51,607
D19	27,659	1.40	1.69	4,592	913	54,869