ASSESSMENT OF DIVERSITY OF BLACKBERRY (*Rubus* L. subgenus *Rubus* Watson) ACCESSIONS IN KENYA USING MORPHOLOGICAL AND SSR MARKERS

OCHIENG' JARED ALFRED

A thesis submitted to the Graduate School in partial fulfillment for the requirements of the Master of Science Degree in Plant Breeding of Egerton University

EGERTON UNIVERSITY
OCTOBER 2018

DECLARATION AND RECOMMENDATION

Declaration

I declare that this thesis is my original work and has not been submitted for examination in any other institution.

Signature: …… Date: ……

Ochieng' Jared Alfred

Reg. No. KM21/14165/15

Recommendation

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

Signature: …… Date: ……

Dr. Maurice Oyoo, PhD

Department of Crops, Horticulture and Soils

Egerton University

Signature: …… Date: ……

Dr. Robert Gesimba, PhD

Department of Crops, Horticulture and Soils

Egerton University

Signature: …… Date: ……

Dr. Paul Korir, PhD

Department of Crops, Horticulture and Soils

Egerton University

ACKNOWLEDGEMENT

I would like to thank Egerton University for giving me an opportunity to pursue my studies. I am grateful to my university supervisors Dr. M.E. Oyoo, Dr. P.C. Korir and Dr. R.M. Gesimba for their guidance throughout this research and during writing of my thesis. I am also grateful to Dr. J.O. Owuoche and Dr. P.O. Okwiri for their academic advice, KALRO-Njoro staff for their insights and assistance during field and laboratory work, my colleagues at the university and staff from the department of Crops, Horticulture and Soils for their support during the research. This work was funded by RUFORUM Grant award (RUFORUM-RU/CGS/CRG/30/03/14).

COPYRIGHT
©2018

OCHIENG', J.A

All rights reserved. No part of this thesis may be used or reproduced or distributed in any form or by any means, or stored in a data base or retrieval system without prior written permission of the author.

ABSTRACT

Morphological and genetic diversity are important components for cultivar development and are a pre-requisite to cultivar improvement. The probability of producing unique genotypes increases in proportion to the number of genes for which parents differ (genetic distance). The objectives of this study were to determine the morphological and genetic diversity among native blackberry (*Rubus* L. subgenus *Rubus* Watson) accessions in selected counties in Kenya and their relationship with Plant Introductions (PIs) using morphological and SSR markers. Eleven out of thirteen available blackberry SSR primer sets were used to screen 90 blackberry accessions in this study. Molecular data were scored in binary fashion for SSR marker loci amplified and were analysed using DARwin 6.0, PowerMarker 3.25 and GenAlEx 6.5 software. Each individual blackberry accession was nested within its county of collection and morphological data were taken *in-situ* on all the accessions including erect, semi-erect and trailing types. Morphological data were analysed using GENSTAT 15th Edition programme, SAS ver. 9.1 (SAS Institute Inc., Cary, 2001) and *R* for statistical computing version 3.4.1 software. Both molecular and morphological data analysed detected considerable diversity within and among the blackberry accessions studied. Analysis of Molecular Variance (AMOVA) showed that much of the genetic diversity existed within the accessions (95%) with estimated genetic variation of 4.12. The expected heterozygosity (H_E) of the blackberry accessions ranged from 0.48 to 0.89. Principal component analysis (PCA) conducted on morphological data generated 10 axes, out of which, 7 had a cumulative variation of 96.30%, with the first two axes having a discriminatory variance of 52.71 % sufficient to identify variables able to differentiate blackberry accessions in Kenya. Further, out of the 10 important morphological traits subjected to PCA, 8 were able to differentiate the collected accessions and were considered as variables capable of discriminating them on the basis of morphology. Molecular data cluster analysis using the Jaccard's similarity coefficient grouped the accessions into three classes; I, II and III consisting of 31, 52 and 7 accessions, respectively, while a phylogenetic tree constructed for morphological data, using the Gower's coefficient, grouped the accessions into two classes; I and II consisting of 1 and 89 accessions, respectively. Both clusters were random and did not group the accessions according to their geographical origin, indicating that the accessions found in Kenya are closely related. This study revealed high levels of within genetic diversity in the blackberry genetic resources studied which can be used in blackberry breeding programs.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

USDA	:	United States Department of Agriculture
AMOVA	:	Analysis of Molecular Variance
PCR	:	Polymerase chain reaction
MAS	:	Marker Assisted Selection
NTSYS	:	Numerical Taxonomy and Multivariate Analysis System
UPGMA	:	Unweighted Pair Group Method with Arithmetic Mean
CTAB	:	Cetyl trimethylammonium bromide

DEFINITION OF TERMS

Accession: Accession refers to the basic working unit of conservation in the gene banks.

Germplasm: Germplasm is a valuable natural resource of plant diversity and contains the information for a species' genetic makeup. For plants, the germplasm is usually stored as a seed collection or (less common) another plant part - a leaf, a piece of stem, pollen or even just a few cells that can be turned into a whole plant.

Wild: native blackberry species in Kenya, not necessarily cultivated

CHAPTER ONE
INTRODUCTION

1.1 Background Information

Evaluation of crop species for genetic diversity in wild and introduced germplasm is essential in improvement of horticultural crops and is a precursor to knowledge of the inheritance of key traits, which is a basic requirement for cultivar development (Castro *et al*., 2013; Mason *et al*., 2015). Successful development of improved cultivars is dependent on variability among available genetic resources to act as a source of desirable genes and an increase in heterogeneity may improve the resistance against abiotic and biotic stresses due to new recombinants (van Esbroeck and Bowman, 1998). Eighty-four wild blackberry species which belong to 24 genera have been identified in Kenya (Chittaranjan, 2011). Worldwide, in addition to the wild blackberries, 15 species are in cultivation mainly in the USA. Traditionally, the plant has mainly been grown in the temperate regions although some cultivars also grow well in the tropical and sub-tropical regions of the world (Clark *et al*., 2007).

Some attempts have been made to characterize genetic variation within and among blackberry populations in native and introduced regions and among wild and introduced accessions (Amsellem *et al*., 2000; Miyashita *et al*., 2015). According to Miyashita *et al*. (2015), the percentage of polymorphic loci ranged from 41.5% in *Rubus buergeri* to 95.0% in *Rubus Idaeus* var. *aculeatissimus*. This study revealed high levels of polymorphisms which suggest a broad genetic base due to varied allele recombinations coupled with environmental effects and natural selection over time. The observed phenomenon could also be associated with interspecific hybridization. Amsellem *et al*. (2000) observed considerable genetic diversity between mainland native and island introduced populations of *Rubus alceifolius*. However, there are still only a few detailed studies comparing the genetic diversity between native and introduced germplasm (Amsellem *et al*., 2000; Dossett *et al*., 2011; Miyashita *et al*., 2015), especially in African populations, and with the current high deforestation rates being experienced, such useful germplasm is under threat of disappearance.

Fruit crops of the genus *Rubus* have had numerous uses throughout human history as documented in archaeological studies, as well as in art and herbals (Hummer and Janick, 2007; Hummer, 2010). For most of their history, they were fruits to be gathered from the wild. It was not until the mid to late 1800s that people started to select for better or, more typically in the early stages, novel characteristics in plants that were brought into cultivation (Clark *et al*., 2007). Wild relatives and landraces are the best sources for increasing diversity in the improved exotic introductions that are expected to be high yielders but less adapted to local conditions

(Hajjar and Hodgkin, 2007). Characters that show diversity within each species are commonly used in the characterization process. Attributes of the edible part of the plant such as leaf shape, length, persistence and total foliage cover are used in many crops (Chweya and Edmonds, 1997). For those crops whose fruit is the edible part, fruit size, texture, colour, length and weight are used. In addition to the nutritive aspects of each species, phenology and attributes related to storability of the harvested part for consumption are often considered (Human and Rheeder, 2004). Breeding in blackberry focusses on specific characters which include adaptation, pest and disease resistance, plant habit, primocane fruiting, thornlessness, fruit size and shape, fruit quality and yield (Clark *et al*., 2007).

Clear understanding of the germplasm diversity and relationships among blackberry germplasm is critical to its improvement, especially the high yielding accessions (Lewers *et al*., 2008). Wild relatives of blackberries are crucial reservoirs of natural diversity, often possessing abiotic stress tolerance, disease resistance, and other traits that are devoid in breeding material with narrow genetic base (Clark *et al*., 2007). Characterization of these collections is, therefore, crucial to identify blackberry germplasm diversity with well-adapted important agronomic traits that can be availed to farmers for cultivation (Ipek *et al*., 2009) and also further our understanding of the processes underlying the demographic establishment and evolutionary adaptation following invasion (Alice and Campbell, 1999). Artificial crossing and selection are usually done to improve on fruit characteristics to achieve specific uses (Human and Rheeder, 2004). Although traits related to plant architecture, phenology, fruit quality, pest resistance and environmental adaptation are essential in identification of wild blackberry and can be introgressed into cultivated germplasm (Finn and Clark, 2011), knowledge on interspecific hybridization is important in development of commercial cultivars (Mason *et al*., 2015).

Being an underutilized crop, challenges abound; inadequate breeding programmes and funding targeting blackberry in Kenya, little understanding of population structures within repositories and the available breeding programme, inaccurate identification of species and misclassification in gene banks, difficulty in identification of duplicate accessions in germplasm repositories and unavailability of improved local cultivars. In addition, the available cultivars experience pest and disease problems coupled with abiotic stresses that are not well understood and documented.

1.2 Statement of the Problem

Blackberry is a common wild fruit in many parts of Africa, including Kenya. However, the blackberry in Kenya has not been characterized both at molecular and morphological levels. This has led to taxonomical misclassification of blackberry accessions. The understanding of population structures within *in-situ* or *ex-situ* germplasm repositories in Kenya has not been documented. Consequently, processes underlying early demographic establishment and evolutionary adaptation which show levels of intra-population diversity and population differentiation of native and introduced counterparts have not been established. This phenomenon, coupled with disease and pest problems discourage smallholder farmers from growing blackberry. Lack of knowledge on blackberry reproduction, which varies from sexual, facultatively apomictic to obligately apomictic, has contributed to the minimal development of cultivars of this crop. Other blackberry genetic complexities are attributable to its cytological state (autoploidy and alloploidy) and inheritance (disomic and tetrasomic) for polyploid blackberry. Other challenges include logistical difficulties of collecting the germplasm over large spatial scales. As a consequence, its exploitation for breeding purposes is difficult. This implies that the efforts of plant breeders to develop better yielding cultivars through crossing of accessions with genetic distances are limited. This study will be a precursor to detection of accessions that breeders can use in improvement of the crop.

1.3 Objectives

1.3.1 Broad Objective

To contribute to blackberry breeding and conservation in Kenya by providing information on their genetic and morphological diversity.

1.3.2 Specific Objectives

i. To map blackberry germplasm occurrence in selected counties in Kenya.
ii. To determine the phenotypic variation of wild blackberry types in selected counties in Kenya and Plant Introductions (PIs) using morphological markers.
iii. To determine the genetic diversity of wild blackberry types in selected counties in Kenya and Plant Introductions (PIs) using SSR markers.

1.4 Hypotheses

i. There are no defined blackberry occurrences in selected counties in Kenya.
ii. There are no morphological differences among wild blackberry accessions in selected counties in Kenya and Plant Introductions (PIs).

iii. There are no genetic differences among wild blackberry accessions in selected counties in Kenya and Plant Introductions (PIs).

1.5 Justification of the study

The analysis of genetic relationships among and within crop species is an important component of crop improvement and a prerequisite to any successful plant breeding programme. Traditionally, germplasm characterization has been based on morphological descriptors (Fajardo *et al*., 2002) coupled with reactions to pest, diseases and other stresses existing within germplasm collections. Such quantitative phenotypic traits, however, tend to vary according to environment (Marinoni *et al*., 2003; Lewers *et al*., 2008) and are most useful for traits that are controlled by only a small number of genes (Brown-Guedira *et al*., 2000). Previous germplasm collectors searched only for characteristics based on phenotypic expression such as objective descriptions of tree and fruit characteristics discriminating against undesirable traits in the process (Marinoni *et al*., 2003). This preference for specific traits based on phenotypic descriptions led to the discarding of potentially important and advantageous germplasm (Castillo *et al.*, 2010). As such, classifying germplasm collections based solely on phenotyping protocols may not provide an accurate indication of genetic diversity (Menkir *et al*., 1997).

Characterization of germplasm aims to preserve useful genetic diversity for later introgression back into crop cultivars and for targeted breeding attempts in crop improvement. Characterization of germplasm can also reveal cases of species misclassification, providing useful genetic diversity information and confirming genome composition of the crop (Mason *et al*., 2015). With this in perspective, the combined use of both morphological and molecular markers in breeding is preferable because it provides useful complementary information. Morphological marker-assisted selection has been used by blackberry breeders for primocane-fruiting trait, implying that molecular marker-assisted selection has the potential for adoption (Lewers *et al*., 2008). Characterizing individuals and cultivars within blackberry germplasm collections is important to give insight into the evolutionary history of the crop in Kenya, as well as help breeders narrow the search for new alleles at loci of interest. This will assist in the identification of marker alleles from candidate genes that can then be introduced into new cultivars along with their associated desirable traits.

References

Alice, L. and Campbell, C. (1999). Phylogeny of *Rubus* (Rosaceae) based on nuclear ribosomal DNA internal transcribed spacer region sequences. *American Journal of Botany,* 86(1), 81–97.

Amsellem, L., Noyer, J. L., Bourgeois, T. L. E., and Hossaert-Mckey, M. (2000). Comparison of genetic diversity of the invasive weed *Rubus alceifolius* Poir. (Rosaceae) in its native range and in areas of introduction, using amplified fragment length polymorphism (AFLP) markers. *Molecular Ecology, 9*(4), 443–455.

Brown-Guedira, G.L., Thompson, J.A., Nelson, R.L. and Warburton M.L. (2000). Evaluation of genetic diversity of soybean introductions and North American ancestors using RAPD and SSR markers. *Crop Science,* 40, 815-823.

Castillo, N. R. F., Reed, B. M., Graham, J., Fernandez-Fernandez, F., and Bassil, N. V. (2010). Microsatellite Markers for Raspberry and Blackberry. *Journal of the American Society for Horticultural Science, 135*(3), 271–278

Castro, P., Stafne E.T., Clark, J.R. and Lewers, K.S. (2013). Genetic map of the primocane fruiting and thornless traits of tetraploid blackberry. *Theoretical Applied Genetics,* 126, 2521-2532.

Chittarajan, T.W. (2011). Blackberry and blueberry pests and diseases. *Indian Journal of Fruit Science,* 25, 54-68.

Chweya, J.A. and Edmonds, J. (1997). Nightshades and related species. http://www.underutilized-species.org/documents/publications/black_nightshades.pdf. Retrieved on 19th April 2016.

Clark, J.R., Stafne, E.T., Hall, H.K. and Finn, C.E. (2007). Blackberry breeding and genetics. *Plant Breeding Reviews,* 29, 82-101.

Dossett, M., Bassil, N. V., Lewers, K. S., and Finn, C. E. (2012). Genetic diversity in wild and cultivated black raspberry (*Rubus occidentalis* L.) evaluated by simple sequence repeat markers. *Genetic Resources and Crop Evolution*, *59*(8), 1849–1865.

Fajardo, D.S., La Bonte, D.R. and. Jarret, R.L. (2002). Identifying and selecting for genetic diversity in Papua New Guinea. *Genetic Resources and Crop Evolution,* 49, 463–470.

Finn C.E. and Clark J.R. (2011). Emergence of Blackberry as a World Crop. *Chronica Horticulturae,* 51(3), 13-18.

Hajjar, R. and Hodgkin, T. (2007). The use of wild relatives in crop improvement: a survey of developments in the last 20 years. *Euphytica,* 156, 1–13.

Human, C.F. and Rheeder, S. (2004). Mango breeding: Results and Successes. *Acta Horticulturae,* 645, 331 - 335.

Hummer, K. E. (2010). Rubus pharmacology: Antiquity to the present. *Horticultural Science,* 45, 1587–1591.

Hummer, K.E. and Janick, J. (2007). *Rubus* iconography: Antiquity to the Renaissance. *Acta Horticulturae,* 759, 89 – 106.

Ipek, A, Barut, E., Gulen, H., and Ipek, M. (2009). Genetic diversity among some blackberry cultivars and their relationship with Boysenberry assessed by AFLP Markers. *African Journal of Biotechnology*, *8*(19), 4830–4834.

Lewers, K.S., Saski, C.A., Cuthbertson, B.J., Henry D.C., Staton M.E., Main D.S., Dhanaraj A.L., Rowland L.J. and Tomkins J.P. (2008). A blackberry (*Rubus* L.) expressed sequence tag library for the development of simple sequence repeat markers. *BMC Plant Biology,* 8, 69-72.

Marinoni, D., Akkak, A., Bounous, G., Edwards, K.J. and Botta, R. (2003). Development and characterization of microsatellite markers in *Castanea sativa* (Mill.). *Molecular Breeding,* 11, 127-136.

Mason, A.S., Zhang, J., Tollenaere, R., Teuber, P.V., Dalton-Morgan, J., Hu, L., Yan, G., Edwards, D., Redden R. and Batley J. (2015). High-throughput genotyping for species identification and diversity assessment in germplasm collections. *Molecular Ecology Resources* 15(5): 1091–1101.

Menkir, A., Goldsbrough P. and Ejeta, G. (1997). RAPD based assessment of genetic diversity in cultivated races of sorghum. *Crop Science,* 37, 564-569.

Miyashita, T., Kunitake, H., Yotsukura, N. and Hoshimo, Y. (2015). Assessment of genetic relationships among cultivated and wild Rubus accessions using AFLP markers. *Scientia Horticulturae,* 193, 165-173.

van Esbroeck, G.A. and Bowman, D.T. (1998). Cotton germplasm diversity and its importance in cultivar development. *Journal of Cotton Science,* 2, 121-129.

CHAPTER TWO
LITERATURE REVIEW

2.1 Origin, Genetics and Adaptation of Blackberry

Blackberry (*Rubus* L. subgenus *Rubus* Watson) is a perennial plant that exhibits trailing to erect growth habits with canes attaining up to 5 meters tall (Clark *et al*., 2007). Blackberries were first mentioned in 370 B.C. by the Greek writer Theophrastus who reported that the plant was used in hedges to keep out invading forces 2000 years ago (Jennings, 1988). They were domesticated in Europe by the seventeenth century and in North America during the nineteenth century where they later considered as noxious weeds (Jennings, 1988). The earliest cultivars of blackberry were selected in 1800 during deforestation and, thereafter, numerous wild species propagated and hybridized (Darrow, 1937). The first released blackberry cultivar was 'Dorchester' in 1841 while the first cultivar to be widely planted was 'New Rochelle' (Hedrick, 1925).

The basic chromosome number of wild and cultivated *Rubus* accessions vary with ploidy levels and range from 2n=2x=14 to 2n=18x=126 including odd-ploids and aneuploids (Meng and Finn, 1999; 2002). Four diverse groups of blackberries have been domesticated; the European blackberries derived from a group of diploid and polyploid species (2n=2x=14, 2n=4x=28, 2n=6x=42 and 2n=8x=56); erect blackberries, trailing diploid and tetraploid dewberries and trailing blackberries from polyploidy species originally from Western America, predominantly *Rubus ursinus* (2n=8x=56, 2n=12x=84). Hybrids of *Rubus allegheniensis* Porter × *Rubus frondonsas* Bigelow played an important role in the domestication of the crop (Hedrick, 1925; Clark *et al*., 2007). The discovery and development of intersectional hybrid in blackberry between a pistillate *Rubus ursinus* cultivar '*Aughinbaugh*' and '*Red Antwerp'* was important in blackberry breeding. This breeding effort led to the first release of a blackberry cultivar from a breeding programme (Logan, 1955). Cultivar '*loganberry*' was later detected to be an allohexaploid resulted from a reduced gamete of an octoploid *Rubus ursinus* and an unreduced gamete of diploid *Rubus idaeus* (Jennings, 1981). Other interspecific polyploidy hybrids were selected in the late 1800s and early 1900s including '*Laxtonberry*' and '*Boysenberry*' (Clark *et al*., 2007). The first public blackberry breeding programme was initiated in Texas Agricultural Experiment Station with emphasis on developing blackberries with low chilling requirement whilst adapted to warmer climates (Darrow, 1937).

Blackberries are well adapted to a wide range of climates and soils pH (5.5 to 6.5) and its growth is improved under conditions of good soil drainage (Anderson and Crocker, 2014). Blackberry has been grown in the temperate regions, tropical and subtropical regions of the

world although, they have been considered as noxious weed in some areas of the world, especially Australia, Chile and North Western America (William and West, 2000). These shrubs in some regions invade natural ecosystems and are a serious threat to nature conservation, especially so in Australia and New Zealand, where biota have a high degree of endemism and have evolved to be particularly susceptible to environmental weeds (Evans and Weber, 2003). According to William and West, (2000) their classification as noxious weeds is because of their effect on most indigenous plant species which they have in fact, made extinct and the imminent threat they pose to ecosystem stability and functional complexity. In Kenya, commercial blackberry is grown around Naivasha (0.7172° S, 36.4310° E) and Limuru (1.1069° S, 36.6431° E) and in these regions, planting occurs during the rainy season of July and August, then a series of cultural manipulations are applied including defoliation with chemicals, pruning and application of growth regulators.

2.2 Taxonomy of Blackberry

Blackberry is a cross-pollinated plant of the genus *Rubus* that is characterized by considerable diversity which is in turn complicated by polyploidy, agamospermy, and frequent hybridization ranging from crops that are tiny and prostrate to semi-erect or erect (Strik, 1992; Alice, 2002) that form a very large bush of up to 5 metres tall (Clark *et al*., 2007). Blackberries are classified into three types based on their growth habit: erect, which produce self-supporting canes; semi-erect, where canes are partially erect but require a trellis for support; and trailing, where canes are not erect and require a trellis for support (Strik, 1992). Although all blackberries have thorny and thornless cultivars, trailing and semi-erect types have few root buds and usually produce primocanes from crown buds. Erect blackberries have many new shoots that are located at the base of the overwintering canes, and are usually out of sight called "crown buds" and readily produce primocanes from both roots and crowns (Poling, 1997). All blackberry plants have biennial canes in which the roots and crowns exhibit a seasonal pattern of growth. The flowers are showy, white to pink, perfect and self-fertile and are surrounded by a ring of stamens which in turn, are surrounded by a receptacle covered with pistils arranged in whorls. These flowers generally vary from 5 to 15 per lateral although some varieties have numerous flowers in each lateral (Lewers *et al*., 2010). The flowers are insect pollinated and when fertilization occurs, a drupelet is formed from each ovule resulting in an aggregate fruit that often ripens together according to the level of fruit with primary fruits ripening prior to the secondary, quaternary, or tertiary (Hummer and Janick, 2007). Each aggregate fruit is usually attached to a fleshy receptacle on fruit laterals and vary in size and shape among cultivars.

Blackberry cultivars have complex backgrounds, often composed of several ploidy levels which vary with progenitor species chromosome number and thus, lack an epitaph (Naess *et al*., 1998). This also implies that there is potential for further transfer of genetic information in every combination of maternal and paternal parents used from one species to another, hence, new sources of variability. The subgenus *Allegheniensis*, *Arguti*, *Rubus* and *Ursini* have been the primary contributors to the pedigree of cultivated blackberries (Finn and Clark, 2011). Other species in the *Caesii, Canadanses, Flagellares, Verotriviales, Idaeobatus* and *Lampobatus* have also been utilized in breeding (Jennings *et al*., 1992; Finn, 2008). Crosses among other ploidy levels within the *Rubus* subgenera are usually fertile. Crosses with members of the *Ideabatus* are also often successful and have produced numerous cultivars, however, crosses between diploids and tetraploids or with other sub-genera often lead to sterility (Finn *et al*., 2002; Clark *et al*., 2007). Blackberries have similar horticultural characteristics to raspberries and almost every region of the world where *Rubus* blackberries are, raspberries are also found (Potter *et al*, 2007). In contrast, they have lower production costs as compared to raspberries and this is attributed to their vigorous nature, greater disease tolerance, hence, longer-lived plantings (Finn and Clark, 2011).

This crop is indigenous to Europe and Asia (Alice, 2002) and most blackberries found in Kenya and Africa are indigenous to these areas. In fact, African production is only found in South Africa (Finn, 2008). Varieties of blackberry are classified as primocanes, floricanes and primocane fruiting (Clark, 2008). Primocanes are the first year canes that are vegetative only, floricanes are second-year canes and these flower, fruit and die while primocane-fruiting indicates that the fruits are borne on first year canes, without chilling requirements (Clark and Finn, 2011). Normally, the remaining buds that did not fruit on primocanes develop and fruit on floricanes the following year.

2.3 Importance of Blackberry

Blackberry is considered a minor crop in the world when compared to other fruits with respect to adoption, production and market niche (Finn and Clark, 2012). Their fruits, however, have long been collected and consumed worldwide regardless whether they are wild or from cultivated fields (Finn, 2008) or whether it is due to their health benefits since they provide natural phytochemicals (Rao and Snyder, 2010). Although the crop is moderately susceptible to extreme frost, it is drought resistant and can do well in areas with low soil moisture regimes (Clark *et al*., 2007). Chilling conditions in the temperate regions are necessary for vernalization

which shifts the primocanes to floricanes that subsequently, die after producing fruits. In spite of this, some species that exhibit primocane-fruiting, are able to flower without vernalization (Lopez-Medina and Moore, 1999). Blackberry fruits have great importance to humans: pharmacological history (Hummer, 2010) contains phenolic compounds that are secondary plant metabolites and integral part of human and animal diets (Siriwoharn *et al*., 2004; Lee *et al*., 2011). Phenolic compounds have long been considered as antinutrients. Recent studies and interest in food phenolics have proven otherwise, with these compounds being touted for their antioxidant properties (Hirsch *et al*., 2013), their remedy of cardiovascular diseases and other disorders (Hollman *et al*., 1996; Bravo, 1998). Blackberry fruits can also be consumed fresh or processed (Finn, 2008). An interspecific hybrid of blackberry × raspberry 'youngberry' that is mostly grown in South Africa is usually processed into wine for local and export markets (Finn and Clark, 2012).

The annual world production of blackberry is estimated to be approximately 154,644 tonnes, with Europe producing the bulk of it approximately 47,399 tonnes. Documentation of African production is only limited to South Africa with approximately 220 tonnes (Strik *et al*., 2007; Finn, 2008). Currently, blackberry production is expanding and despite this expansion, the demand for its products is ever increasing.

2.4 Released Cultivars of Blackberry in Different Countries

Various types of blackberries have been released from breeding programmes. These include erect-caned cultivars namely; 'Arapaho, Ouachita and Navaho; and the thorny Cherokee, Cheyenne, Choctaw, Kiowa and Prime Arkansas 45. Semi-erect cultivars include 'Chester Thornless', 'Thornfree', Triple Crown, Loch Ness, and Hull Thornless. Trailing cultivars include Marion, Silvan, and Thornless Evergreen and the blackberry-raspberry hybrids 'Boysen' and 'Logan' (Clark *et al*., 2007). There are also primocane-fruiting cultivars Prime-Jan and Prime-Jim, Prime Arkansas Traveler and Prime Arkansas 45 that are erect, thorny types (Clark, 2010). In addition, there is the 'Prime Arkansas freedom' (Clark and Finn, 2011). The erect blackberry cultivars produce primocanes from buds at the base of floricanes at the crown or from buds on roots, while trailing and semi-erect types only produce new primocanes from buds on the crown (Strik *et al*., 2007).

2.5 Pest and Diseases of Blackberry

Blackberry is generally free from pests and diseases as compared to red and black raspberries (Finn and Clark, 2012; Martin *et al*., 2013). Environmental conditions and type of cultivars grown are important determinants on type and severity of the pests and/or diseases that affect blackberry. However, new pest species or biotypes continue to be observed on *Rubus*,and this is attributed to modifications in pesticide usage, the introduction of new blackberry cultivars, or insect-host range. Some of the common insect pests in blackberry are; raspberry crown borer (*Pennisetia marginata*), redberry mite (*Acalitus essigi*), thrips (*Frankliniella tritici*) and flower thrips (*Frankliniella occidentalis*). Other arthropod pests of wild, cultivated and blackberry hybrids, for example, Tayberry and Loganberry include aphids (*Amphorophora idaei* Bomer and *Aphis idaei* Goot); raspberry beetle (*Byturus tomentosus* Degeer); clay-coloured weevil (*Otiorhynchus singularis* L.), raspberry cane midge (*Resseliella theobaldi* Barnes), raspberry moth (*Lampronia rubiella* Bjerkander) and two-spotted spider mite (*Tetranychus urticae* Koch.) (Gordon *et al*., 1997).

In regions where erect and/or semi-erect blackberries are dominant, anthracnose (*Elsinoe veneta*), botrytis fruit rot (*Botrytis cinerea*), blackberry cane canker (*Botryosphaeria dothidea*) and *Colletotrichum* spp. are prevalent (Converse, 1966). This is mainly in the continental regions. Mediterranean regions of Mexico, Chile, South Africa, USA and New Zealand are mostly affected by cane botrytis, cane spot (*Septoria rubi),* purple blotch (*Septocyta ruborum*), and spur blight (*Didymella applanata*). Since much of the ripening season is dry, fruit rots are not as much of a problem. Powdery mildew (*Sphaerotheca macularis*) is a common occurrence in areas where dry conditions are experienced throughout the growing seasons, although downy mildew (*Peronospora sparsa*) may result in case of wet conditions during blooming (Clark *et al*., 2007). Apart from the above mentioned diseases, bacterial diseases such as crown gall (*Agrobacterium tumefaciens*) can be problematic, however, they rarely result in economic crop losses (Ellis *et al*., 1991). Common disease problems regardless of environment and/or blackberry types also exist. These include; anthracnose (*Elsinoe veneta* Jenk.), botrytis fruit rot (*Botrytis cinerea*), cane blight (*Leptosphaeria coniothyrium*) and orange rust (*Gymnoconia peckiana*) (Ellis *et al*., 1991).

2.6 Yield of Blackberry

Blackberry yields vary greatly depending on growing region, blackberry type, cultivar, and cultural care (Fernandez and Ballington, 2012). The trailing blackberry is grown at a spacing of 0.9 to 1.5m by 3m and a trellis with the canes is usually wrapped around 2 to 3 wires (Strik and Finn, 2012). Its yield ranges between 8 and 12 t ha^1, depending on cultivar, growing region, and harvest method (Strik and Finn, 2012). Semi-erect blackberry is grown at a spacing of 1.0 to 1.8m by 3m trained to a multiple wire trellis (Strik and Finn, 2012). Its yield ranges between 20 to 30 t ha^1 and in regions where this type of blackberry is grown, shade or tunnels are used to protect the crop from the sun or rain damage. This is because they are usually high value, late-season crops (Strik and Finn, 2012). Erect blackberry is grown at a spacing of 0.8 to 2m by 3m and does not require any support (Strik and Finn, 2012). Its yield ranges from 8 to 10t/ha (Strik and Finn, 2012). Soft tipping is usually encouraged at about 1m to encourage better branching which has an effect on yields and control of pests and diseases.

2.7 Morphological Diversity of Blackberry

Morphological diversification of plants was the initial step in unparalleled innovation in the history and characterization of plant genetic resources (Kenrick and Crane, 1997). This is because, from a simple plant structure consisting of only a few cells, plants develop elaborate two-phase life cycles and an extraordinary array of complex organs and tissue systems. Terrestrial crops life cycles are characterized by alternating multicellular sexual (haploid gametophyte, n) and asexual phases (diploid sporophyte, $2n$) and in addition, phylogenetic studies show that the multicellular gametophyte is inherited from their algal ancestors whereas the sporophyte evolved during the transition to the land (Carrol, 2000). Together, these changes resulted in more highly differentiated plants with stomates, multicellular sexual and spore-bearing organs, water-conducting and other tissue systems. Consequently, there is morphological differentiation in the life cycles (gametophyte and sporophyte), although, speciation greatly favours sporophyte complexity among vascular plants over gametophyte phase. Previously, morphological diversification was based on apical growth and branching variations coupled with delayed initiation of spore-bearing organs were important innovations of vascular plants that led to more complex architectural frameworks in crops (Foote, 1997).

The characterization of germplasm has traditionally used morphological descriptors which consist of phenotypic traits like flower colour and growth habit (Fajardo *et al*., 2002). This method of classifying germplasm is the oldest and is considered as an initial step in classifying germplasm (Hedrick, 2005). Morphological markers are straightforward, easy,

cheap technique for plant identification and characterization, although, they are not as precise as DNA markers (Li *et al*., 2009). Variables of interest to the plant breeder are usually visually monitored and noted. This is because they are easily detectable plant characteristics like form and structure. However, there exist errors in scoring which may be attributed to environmental effects and, hence, observations made on some morphological descriptors tend to be subjective in nature. (Fajardo *et al*., 2002; Marinoni *et al*., 2003; Li *et al.*, 2009).

Additional limitations of morphological-marker assisted traits include phenological changes in plant phenotypes depending on the growth stage, insufficient variation and the length of time required for the appearance of informative traits particularly in tree crops (Castillo, 2010). Currently, there is no list of accepted morphological descriptors for blackberry DNA-informed breeding (Finn, personal communication). A standardized phenotyping protocol for blackberry is also presently being developed with the aim of standardizing blackberry phenotyping for the purpose of identifying horticulturally important quantitative trait loci (QTLs). Accurate identification, therefore, becomes difficult in the process, lowering the reliability of morphological markers for germplasm characterization (Finn *et al*., 2010).

2.8 Genetic Diversity of Blackberry

The analysis of genetic diversity in germplasm, wild or elite, is an important component in studies of plant genetics, breeding, conservation and evolution. The assessment of genetic diversity is an important aspect of plant breeding if there is to be an improvement by selection (Mason *et al*., 2017) as it provides a platform for stratified sampling structure available and breeding populations. Such analysis, however, depend on the genome or individual germplasm sampling with sufficient and informative genetic markers, as such, molecular markers are preferred. Molecular markers are generally superior to morphological, pedigree, heterosis, and biochemical data and are preferred for evaluation of genetic diversity of genotypes (Melchinger *et al.*, 1991). This genetic relatedness of cultivars is commonly measured by genetic distance (GD) or genetic similarity (GS = 1 - GD), both of which imply that there are either differences or similarities at the genetic level (Weir, 1990). Random Fragment Length Polymorphism-based GDs have been used in evaluating the genetic diversity of maize inbred lines and in determining their hybrid performances (Benchimol *et al*., 2000).

Advances in plant biotechnology offer novel techniques that greatly reduce breeding costs and importantly so, the time required to develop cultivars. DNA-based markers, simply detect differences in the genetic information carried by two or more individuals (Paterson *et al*., 1991). An array of molecular marker techniques has been developed for Rosaceae.

However, these molecular techniques have not been pursued as vigorously in blackberry. This may be because the crop is still considered a minor crop in the world (Strik, 2007). In spite of this, there are some molecular techniques available for blackberry. These include biochemical markers, amplified fragment length polymorphisms (AFLP), restriction fragment length polymorphisms (RFLP), random amplified polymorphic DNA (RAPDs) and simple sequence repeats (SSRs). These techniques involve DNA characterization and facilitate studies in the differentiation of genotypes (Stafne *et al*., 2005), mating systems (Kraft *et al*., 1996), conduct phylogeny analysis studies (Alice *et al*., 1997) and genetic diversity (Stafne and Clark, 2004). Phylogenetic insights in *Rubus* have also been studied using *In-situ* hybridization techniques (ISH) - Genomic *in-situ* hybridization (GISH) and fluorescence *in situ* hybridization (FISH) with an objective of determining clues to infer the role of *R. parvifolius* allegedly plays in speciation and polyploidization of the genus (Yan *et al*., 2015).

2.8.1 Use of Genetic Markers in Blackberry Breeding

DNA or molecular markers are variations in genotype amongst individuals that are closely linked to loci controlling or contributing to a trait. DNA markers act as “flags” or “signposts”, or “tags” and do not affect the phenotype of the individual. Further, the strength of a molecular marker is dependent on how closely linked it is to the gene of interest (St. Clair, 2010). DNA markers are extremely ubiquitous and are not influenced by phenological changes during plant growth, age, or environmental factors and can be represented by insertions, deletions, point mutations or errors in replication of tandemly repeated DNA (Collard *et al*., 2005).

Biochemical markers were introduced in the 1960s and involve protein and enzyme electrophoresis. These markers are useful in the analysis of genetic diversity as they reveal differences between seed storage proteins or enzymes encoded by different alleles at one (allozymes) or more gene loci (isozymes) (Rao, 2004). The enzymes are differently charged variants that are separable by electrophoresis. Visualization is achieved by supplying the bands with the substrates and co-factors and observing the formation of protein products encoded by different alleles/genes and provide co-dominant markers (Castillo *et al*., 2010). Allozymic polymorphism has been used to ascertain genetic diversity in almost all major crops and in identification of cultivars (Veasey *et al*., 2002). However, the level of isozyme variation is too low for cultivar identification and hence, some *Rubus* cultivars remain undistinguishable (Cousineau and Donnelly, 1992). This is a major constraint and a limitation in using isozyme

analysis for fingerprinting mostly because of lack of or low level of variation in many cultivars and species.

Restriction Fragment Length Polymorphism (RFLP) is the first DNA-based marker developed (Bostein *et al*., 1980) and resulted from differences in the sequences of nucleotides in different plants. This technique is based on the restriction enzymes that reveal the pattern difference between DNA fragment sizes in individual organisms (Semagn *et al*., 2006). DNA fragments are transferred by Southern blotting to a nitrocellulose or nylon membranes that are generally hybridized to a radioactively-labelled DNA probe. These markers require no sequence information, are co-dominant and analysis of band profiles is easy to score. This marker has been found to be effective in identifying *Rubus* cultivars (Waugh *et al*., 1990) and demonstrating genetic variability among the selected taxa (Nybom *et al*., 1992). This shows the ability of RFLP to reveal genetic differences among closely related *Rubus*' species or taxa. The disadvantages of RFLPs include the requirement of high quantity and quality of DNA and for radioactive labelling of specific probe libraries.

Random Amplified Polymorphic DNAs (RAPDs) are PCR-based markers (Williams *et al*., 1990). They were commonly used due to the simplicity and low cost of agarose gel electrophoresis. The RAPD protocol usually uses an oligonucleotide that is 10 bp long at constant annealing temperature, in a PCR reaction to amplify many copies of random genomic DNA sequences simultaneously. In *Rubus* RAPD markers have been used in identification of raspberry cultivars (Graham *et al*., 1997), establishing the genetic relationships (Weber, 2003). The approximation of the relatedness in pedigree analysis of RAPD data using cluster analysis can overestimate or underestimate percentage relationships. This results in an uncertainty of the relationship showed by pedigree analysis. RAPD markers also have limitations like the irreproducibility of banding patterns preventing comparisons to be made between studies (Nybom, 2004).

Amplified Fragment Length Polymorphism (AFLP) is based on the amplification of subsets of genomic restriction fragments using PCR (Vos *et al*., 1995). The first step of the AFLP protocol involves digestion of the DNA with two restriction enzymes, a rare cutter like *Eco*RI and a frequent cutter like *Mse*I. Polymorphisms are revealed after separating the amplified DNA fragments by electrophoresis on a sequencing gel and visualized by silver staining, radioactive or fluorescent detection. A large number of bands are generated that facilitates the detection of polymorphisms. AFLP reveals a high level of polymorphism has a high diversity index and can analyse a large number of bands (Russell *et al*., 1997). AFLP has been used in *Rubus* to demonstrate sexual recombination (Kollmann *et al*., 2000) and to

evaluate genetic diversity (Amsellem *et al*., 2001). AFLP markers are cost efficient, easy to use, require a small amount of DNA. The information generated is replicable, is of high quantity and resolution in comparison to other standard molecular markers (Mueller and Wolfenbarger, 1999). The technique also permits the detection of restriction fragments and can generate fingerprints of any DNA regardless of origin and complexity. It also has a broad taxonomic scope and can be developed in any organism with DNA without prior knowledge of the organism's genomic make-up.

Single nucleotide polymorphisms (SNPs) are informative, abundant genetic markers and are evenly distributed throughout the genomes of most plant species, however, challenges exist by using genotype by sequencing (GBS) especially in non-model species as many species of interest where SNP markers are lacking (Glaszmann *et al*., 2010; Smith and Maughan, 2015). SNPs have become markers system research involving genetic inference in many crops (Yan *et al*., 2010). This is because arrays of advanced high throughput platforms have been developed, and these platforms are capable of rapid and simultaneous genotyping of up to a million SNP markers. SNP markers have varied applications, including association studies, conservation genetics, genetic diversity analysis. Additionally, they are increasingly becoming the molecular marker of choice in marker-assisted plant breeding programmes SNP markers are developed through various approaches including chip hybridization and by targeting specific genomic regions although, such efforts are expensive, labour intensive and technically difficult in some plant species. This is due to the fact that some crops, for instance, blackberry, usually have complex genomes with abundant sequence repeats and genome duplications (Ward *et al*., 2013). Furthermore, most crop species do not have sequenced genomes and hence, make SNP discovery more challenging (Poland and Rife, 2012).

Genotyping-by-sequencing (GBS) is an advanced high throughput genomic approach for assessing genetic diversity of crops on a genome-wide scale as a result of the advances made technologies involving next-generation sequencing (NGS) technologies (Metzker, 2010). The GBS approach is a combined one-step process of marker discovery and genotyping and provides a rapid, high-throughput and cost-effective tool for a genome-wide analysis of genetic diversity for a range of non-model species and germplasm sets (Fu *et al*., 2014). Therefore, GBS is advantageous and encouraging in studies involving genetic diversity of crops with no informative markers available. Although, SNP methods for the identification and mapping of *Rubus* subgenus *Rubus* have yet to be observed, genotyping by Sequencing (GBS) has been used to produce highly saturated maps for a *Rubus idaeus* pseudo-testcross progeny. This resulted in low coverage and high variance in sequencing and in addition, a large number were

missing values for some individuals, which was corrected by imputation based on maximum likelihood marker ordering from initial marker segregation (Ward *et al*., 2013). In blackberry, 67,000 single-nucleotide polymorphisms (SNPs) have been detected (Garcia-Seco *et al*., 2015).

Simple sequence repeat markers (SSRs) are short oligonucleotide repeats, usually 6 to 10 base pairs long that vary in number (Rafalski *et al*., 1996). SSRs are highly polymorphic PCR-based markers and are found in coding and non-coding regions (Russell *et al*., 1997) and are occasionally transcribed, hence, may be identified in expressed sequence tags (ESTs). SSRs have many advantages which include requiring a small amount of starting DNA, are multi-allelic, co-dominant, high reproducibility, easily detected by PCR, relatively abundant and has extensive genome coverage (Powell *et al*., 1996). Reproducibility of SSR markers between laboratories as primer sequences is also easy and, therefore, provides a platform for collaborative research due to available common language. Since SSRs are highly reproducible and easily detected, they can distinguish between closely related crops that have a narrow genetic base like blackberry. In addition, the possibility of adding new data to an existing database, even when developed in a different laboratory has been a major advantage (Sehic *et al*, 2012). This, however, is dependent on the use of common SSR markers and suitable standardization procedures which have not been very successful in fruit tree characterisation studies as most research works develop their own SSR markers or choose various sets of SSRs from the literature. SSR markers have become particularly useful in the assessment of genetic diversity (Amsellem *et al*., 2001). SSR markers have also been used in fingerprinting and ecological-genetic studies (Li *et al*., 2009), marker-assisted selection and genetic linkage mapping studies (Stafne *et al*., 2005). Microsatellite markers for blackberry were recently developed from an expressed sequence tag library of 'Merton Thornless' (Lewers *et al*., 2008). Eight SSRs have been isolated from the invasive weed *R. alceifolius* Poir. (Amsellem *et al*., 2001) and in red raspberry (Graham *et al*., 2002; Graham *et al*., 2004). Primers for SSR loci in blackberry have been published (Castillo *et al*., 2010).

References

Alice, L. A. (2002). Evolutionary relationships in *Rubus* (Rosaceae) based on molecular data. In *Acta Horticulturae*, 585, 79–83.

Alice, L. A., Eriksson, T., Eriksen, B. and Campbell C. S. (1997). Inter subgeneric hybridization between a diploid raspberry, *Rubus idaeus*, and a tetraploid blackberry, *R. caesius* (Rosaceae). *American Journal of Botany*, 84, 171-176.

Andersen, P. C., and Crocker, T. E. (2014). The Blackberry. *University of Florida. IFAS Extension*, (HS807), 1–8.

Amsellem, L., Dutech, C. and Billotte, N. (2001). Isolation and characterization of polymorphic microsatellite loci in *Rubus alceifolius* Poir. (Rosaceae), an invasive weed in La Réunion Island. *Molecular Ecology Notes*, 1, 33-35.

Benchimol, L., De Souza, L., Garcia, A. F., Kono, M. S., Mangolin, A., Barbosa, M. M., Coelho, S. G. and De Souza, P. (2000) Genetic diversity in tropical maize inbred lines: heterotic group assignment and Hybrid performance determined by RFLP markers. *Plant Breeding*, 119, 491-496.

Bostein, D., White, R.L., Skolnick, M. and Davis, R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Genetics*, 32, 314-331.

Bravo, L. (1998). Polyphenols: Chemistry, Dietary Sources, Metabolism, and Nutritional Significance. *Nutrition Reviews*, 56, 317-333.

Carroll, S. B. (2000). Endless forms: the evolution of gene regulation and morphological diversity. *Cell*, *101*(6), 577–580.

Castillo, R.F., Reed, B. M., Graham, J. and Bassil, N.V. (2010). Microsatellite Markers for Raspberry and Blackberry. *Journal of the American Society for Horticultural Science*, 135(3), 271–278.

Castro, P., Stafne, E. T., Clark, J. R. and Lewers, K. S. (2013). Genetic map of the primocane-fruiting and thornless traits of tetraploid blackberry. *Theoretical and Applied Genetics*, *126*(10), 2521–2532.

Chakraborty, R. and L. Jin. (1993). Determination of relatedness between individuals using DNA fingerprinting. *American Journal of Human Biology*, 65, 875–895.

Clark, J.R., Stafne, E.T., Hall, H.K. and Finn, C.E. (2007). Blackberry breeding and genetics. *Plant Breeding Reviews*, 29, 82-101.

Clark, J.R. (2008). Primocane-fruiting blackberry breeding. *Horticultural Science,* 43, 1637-1639.

Clark, J.R. (2010). Prime-Ark 45®, a new blackberry variety. Small Fruit News. <http://www.smallfruits.org/Newsletter/Vol10-Issue1.pdf>. Retrieved on 19th April 2017.

Clark, J.R. and C.E. Finn. (2011). Blackberry Breeding and Genetics. In: H. Flachowsky and V-M. Hanke (eds.) Methods in Temperate Fruit Breeding. Fruit, Vegetable and Cereal Science and Biotechnology 5 (1), 27-43.

Collard, B. C. Y., Jahufer, M. Z. Z., Brouwer, J. B., and Pang, E. C. K. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica*, *142*(2), 169-196.

Converse, R. H. (1966). Diseases of raspberries and erect and trailing blackberries. Agricultural Research Service, U.S. Dept. of Agriculture. Retrieved from http://agris.fao.org/agris-search/search.do?recordID=US201300421343 on 21st December 2017.

Cousineau, J.C. and Donnelly, D.J. (1992). Genetic analysis of isoenzyme in raspberry. *Journal of the American Society for Horticulture Science,* 117(6), 996-999.

Darrow, G.M. (1937). Blackberry and Raspberry Improvement. *Yearbook,* 496-533.

Derek L. Jennings, Hugh A. Daubeny, J. N. M. (1992). Blackberries and Raspberries (*Rubus*). *Genetic Resources of Temperate Fruit and Nut Crops*, *290*, 331–372.

Ellis, M. A. Converse, R. H. Williams, R. N. and Williamson, B. (1991). Compendium of raspberry and blackberry diseases and insects. APS Press, St. Paul, MN.

Evans, K. J., and Weber, H. E. (2003). *Rubus anglocandicans* (Rosaceae) is the most widespread taxon of European blackberry in Australia. *Australian Systematic Botany*, *16*(4), 527–537.

Glaszmann, J. C., Kilian, B., Upadhyaya, H. D. and Varshney, R. K. (2010). Accessing genetic diversity for crop improvement. *Current Opinion in Plant Biology*, *13*(2), 167-173.

Gordon, S.C., Woodford, J.A.T. and Birch, A.N.E. (1997). Arthropod pests of Rubus in Europe: pest status, current and future control strategies. *Journal of Horticultural Science, 72*(6), 831-862.

Fajardo, D.S., La Bonte, D.R. and Jarret, R.L. (2002). Identifying and selecting for genetic diversity in Papua New Guinea. *Genetic Resources and Crop Evolution,* 49, 463–470.

Fernandez, G. and Ballington J.R. (2012). Growing blackberries in North Carolina. North Carolina Cooperative Extension Service, AG-401.

Finn, C.E., Swartz, H.J., Moore, P.P., Ballington, J.R. and Kempler, C. (2002). Use of 58 Rubus species in five North American breeding programs-breeders notes. *Acta Horticulturae,* 585, 113-119.

Finn, C. E. (2008). Blackberries. In: Temperate Fruit and Crop Breeding. (J.F. Hancock, ed.). Springer Science.

Finn, C. E. and Clark J.R. (2011). Emergence of Blackberry as a World Crop. *Chronica Horticulturae,* 51(3), 13-18.

Finn, C.E and Clark, J.R. (2012). Blackberry. In: Fruit Breeding, Handbook of Plant Breeding. (Badenes M.L. and Byrne, Eds). Springer Science.

Finn, C.E., Strik, B.C., Yorgey, B., Qian M., Martin R.R., Peterson M. (2010). 'Wild Treasure' Trailing Blackberry. *Horticultural Science,* 45(3), 434-436.

Foote, M. (1997). The Evolution of Morphological Diversity. *Annual Review of Ecology and Systematics*, *28*(1), 129–152.

Fu, Y. B., Cheng, B., and Peterson, G. W. (2014). Genetic diversity analysis of yellow mustard (*Sinapis alba* L.) germplasm based on genotyping by sequencing. *Genetic Resources and Crop Evolution*, *61*(3), 579–594.

Garcia-Seco, D., Zhang, Y., Gutierrez-Mañero, F. J., Martin, C., and Ramos-Solano, B. (2015). RNA-Seq analysis and transcriptome assembly for blackberry (*Rubus* sp. Var. Lochness) fruit. *BMC Genomics*, *16*(1), 5-11.

Gordon, S. C., Woodford, J. A. T., and Birch, A. N. E. (1997). Arthropod pests of Rubus in Europe: Pest status, current and future control strategies. *Journal of Horticultural Science*, *72*(6), 831–862.

Graham, J., Squire, G.R., Marshall, B. and Harrison, R.E. (1997). Spatially dependent genetic diversity within and between colonies of wild raspberry *Rubus idaeus* detected using RAPD markers. *Molecular Ecology,* 6, 1001-1008.

Graham, J. and Smith K. (2002). DNA markers for use in raspberry breeding. *Acta Horticulturae,* 585, 51-55.

Graham, J., Smith, K., MacKenzie, K., Jorgenson, L., Hackett, C. and Powell, W. (2004). The construction of a genetic linkage map of red raspberry (*Rubus idaeus* subsp. *idaeus*) based on AFLPs, genomic-SSR and EST-SSR markers. *Theoretical Applied Genetics,* 109, 740-749.

Hedrick, U.P. (1925). The Small Fruits of New York. J.B. Lyon. Albany, NY.

Hedrick, P.W. (2005). Genetics of populations, Jones and Bartlett, London, UK.

Hirsch, G.E., Vizzotto, M., Aboy, A.L., Henriques, A.T. and Emanuelli, T. (2013). Antioxidant activity of Blackberry (*Rubus sp*.) Genotypes from the southern region of Brazil. *B. CEPPA Curitiba*, 31(1), 83-98.

Hollman, P. C. H., Hertog, M. G. L. and Katan, M. B. (1996). Analysis and health effects of flavonoids. *Food Chemicals,* 57, 43-46.

Hummer, K. E. (2010). *Rubus* pharmacology: Antiquity to the present. *Horticultural Science,* 45, 1587–1591.

Hummer, K.E. and Janick, J. (2007). *Rubus* iconography: Antiquity to the Renaissance. *Acta Horticulturae,* 759, 89 – 106.

Jennings, D.L. (1981). A hundred years of Loganberries. *Fruit Varieties Journal,* 35, 34-37.

Jennings, D. L. (1988). *Raspberries and blackberries: their breeding, diseases and growth*. *Fruit Breeding*.

Jennings, D.L., Daubeny H.A. and Moore J.M. (1992). Blackberries and raspberries (Rubus). 331-389. In: Moore J.N., Ballington JR (eds) Genetic resources of fruit and nut crops. Vol 1. International Society for Horticultural Science, Wageningen. 329-330.

Kenrick, P., and Crane, P. R. (1997). The origin and early evolution of plants on land. *Nature*. https://doi.org/10.1038/37918.

Kollmann, J., Steinger, T. and Roy, B.A. (2000). Evidence of sexuality in European *Rubus* (Rosaceae) species based on AFLP and allozyme analysis. *American Journal of Botany,* 87, 1592-1598.

Kraft, T., Nybom H. and Werlemark G (1996). DNA fingerprint variation in some blackberry species (*Rubus* subgen. *Rubus,* Rosaceae). *Plant Systematics and Evolution,* 199, 93-108.

Lewers, K.S., Saski, C.A., Cuthbertson, B.J., Henry D.C., Staton M.E., Main D.S., Dhanaraj A.L., Rowland L.J. and Tomkins J.P. (2008). A blackberry (*Rubus* L.) expressed sequence tag library for the development of simple sequence repeat markers. *BMC Plant Biology,* 8, 69-72.

Lewers, K. S., Wang, S. Y., and Vinyard, B. T. (2010). Evaluation of blackberry cultivars and breeding selections for fruit quality traits and flowering and fruiting dates. *Crop Science, 50*(6), 2475–2491.

Lee, J., Dossett, M. and Finn, C.E. (2011). Rubus fruit phenolic research: The good, the bad, and the confusing. *Food Chemistry,* 130 (2012), 785-795.

Li P., Wang Y., Sun X. and Han J. (2009). Using microsatellite (SSR) and morphological markers to assess the genetic diversity of 12 falcata (*Medicago sativa* spp. *falcata*) populations from Eurasia. *African Journal of Biotechnology*, 8 (10), 2102-2108.

Logan, M.E. (1955). The *Loganberry.* Mary E. Logan (Mrs. J.H. Logan) Publisher, Oakland, California.

Lopez-Medina, J., and Moore J. N. (1999). Chilling enhances cane elongation and flowering in primocane-fruiting blackberries. *Horticultural Science*, 34, 638–640.

Marinoni, D., Akkak, A., Bounous, G., Edwards, K.J. and Botta, R. (2003). Development and characterization of microsatellite markers in *Castanea sativa* (Mill.). *Molecular Breeding*, 11, 127–136.

Martin, R., MacFarlane, S., Sabanadzovic, S., Quito, D. F., Poudel, B., and Tzanetakis, I. E. (2013). Viruses and virus diseases of Rubus. *Plant Disease, 97*(2), 168–182.

Mason, A.S., Zhang, J., Tollenaere, R., Vasquez Teuber, P., Dalton-Morgan, J., Hu, L., Yan, G., Edwards, D., Redden, R. and Batley, J. (2015). High-throughput genotyping for species identification and diversity assessment in germplasm collections. *Molecular ecology resources, 15*(5),1091-1101.

Mason, A. S., Chauhan, P., Banga, S., Banga, S. S., Salisbury, P., Barbetti, M. J. and Batley, J. (2017). Agricultural selection and presence-absence variation in spring-type canola germplasm. *Crop and Pasture Science*, *69*(1), 55-64.

McCulloch, N. and Ota, M. (2002). Export Horticulture and Poverty in Kenya. *IDS Working Paper* 174. Brighton Institute of developmental studies.

Melchinger, A.E., Messmer, M.M., Lee, M., Woodman, W.L. and Lamkey, K.R. (1991) Diversity and relationships among U.S. maize inbreds revealed by restriction fragment length polymorphisms. *Crop Science,* 31, 669–678.

Meng, R. and Finn, C.E. (1999). Using flow cytometry to determine ploidy level in *Rubus. Acta Horticulture,* 505, 223–227.

Meng, R. and Finn, C.E. (2002). Determining Ploidy Level and Nuclear DNA Content in *Rubus* by Flow Cytometry. *Journal of the American Society for Horticulture Science*, 127(5), 767-775.

Metzker, M. L. (2010). Sequencing technologies the next, generation. *Nature Reviews Genetics, 11* (1), 31-46.

Mueller, U. and Wolfenbarger, L. (1999). AFLP genotyping and fingerprinting. *Tree*, 14, 389-394.

Naess, S., Swartz, H. and Bauchan, G. (1998). Ploidy reduction in blackberry. *Euphytica*, 57–73.

Nybom, H., Gardiner S. and Simon C. J. (1992). RFLPs Obtained from an rDNA probe and detected with enhanced chemiluminescence in apples. *Horticultural Science*, 27(4), 355-356.

Nybom, H. (2004). Comparison of Different Nuclear DNA Markers for Estimating Intraspecific Genetic Diversity in Plants. *Molecular Ecology*, 13, 1143–1155.

Paterson, A.H., Tanksley, S.D. and Sorrells, M.E. (1991). DNA markers in plant improvement. In *Advances in agronomy*, 46, 39-90.

Poland, J. A., and Rife, T. W. (2012). Genotyping-by-Sequencing for Plant Breeding and Genetics. *The Plant Genome Journal*, *5*(3), 92.

Powell, W., Machray, G. and Provan, J. (1996). Polymorphism revealed by simple sequence repeats. *Trends in Plant Science*, 1, 215-222.

Rao, N.K. (2004). Plant genetic resources: Advancing conservation and use through biotechnology. *African Journal of Biotechnology*, 3 (2), 136-145.

Rao, A. V. and Snyder, D. M. (2010). Raspberries and human health: A review. *Journal of Agricultural and Food Chemistry*, 58, 3871–3883.

Rafalski, J.A., Vogel, J.M., Morgante, M., Powell, W., Andre, C. and Tingey, S.V. (1996). Generating and using DNA markers in plants, p. 75–134. In: B. Birren and E. Lai (eds.). *Nonmammalian genomic analysis. A practical guide.* Academic, San Diego.

Russell, J., Fuller, J., Macaulay, M., Hatz, B., Jahoor, A., Powell, W. and Waugh, R. (1997). Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. *Theoretical Applied Genetics*, 95, 714-722.

Sehic, J., Garkava-Gustavsson L., Fernandez-Fernandez, F. and Nybom, H. (2012). Genetic diversity in a collection of European pear (Pyrus communis) determined with SSR markers chosen by ECPGR. *Scientia Horticulturae*, 145, 39-45.

Semagn, K., Bjørnstad, Å. and Ndjiondjop, M.N. (2006). An overview of molecular marker methods for plants, *African Journal of Biotechnology*, 5, 2540 -2568.

Siriwoharn, T., Wrolstad, R.E., Finn, C.E. and Pereira, C.B. (2004). Influence of cultivar, maturity and sampling on blackberry (*Rubus* L. hybrids) anthocyanins, polyphenolics, and antioxidant properties. *Journal of Agricultural and Food Chemistry*, 52, 8021–8030.

Smith, S. M., and Maughan, P. J. (2015). SNP genotyping using KASPar assays. *Methods in Molecular Biology*, *1245*, 243–256.

Stafne, E.T. and Clark, J.R. (2004). Genetic relatedness among eastern North American blackberry cultivars based on pedigree analysis. *Euphytica,* 139, 95–104.

Stafne, E T., Clark, J R., Weber, C A., Graham, J. and Lewers, K S. (2005). Simple Sequence Repeat (SSR) markers for genetic mapping of raspberry and blackberry. *Journal of the American Society for Horticulture Science,* 130, 722-728.

St.Clair, D. A. (2010). Quantitative Disease Resistance and Quantitative Resistance Loci in Breeding. *Annual Review of Phytopathology*, *48*(1), 247–268.

Strik, B.C. (1992). Blackberry Cultivars and Production Trends in the Pacific Northwest. *Fruit Varieties Journal,* 46(4), 202-206.

Strik, B. C., and Finn, C. E. (2012). Blackberry production systems-a worldwide perspective. In *Acta Horticulturae,* 946, 341–348.

Strik, B. C., Clark, J. R., Finn, C. E., and Ban, M. P. (2007). Worldwide blackberry production. *Horticulture Technology,* 17, 205–213.

Strik, B., Clark, J.R., Finn, C., Buller, G. (2008). Management of primocane-fruiting blackberry to maximize yield and extend the fruiting season. *Acta Horticulturae,* 777, 423–428.

Veasey, E., Vencovsky, R., Martin, P., and Bande, G. (2002). Germplasm characterization of *Sesbania* accessions based on isozyme analyses. *Genetic Resources and Crop Evolution,* 49, 449-462.

Vos, P., Hogers, R., Bleeker, M., Reijans, M, Van De Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J. Kuiper, M., and Zabeau M. (1995). AFLP: a new technique for DNA fingerprinting, *Nucleic Acids Research,* 11, 4407-4014.

Ward, J. A., Bhangoo, J., Fernández-Fernández, F., Moore, P., Swanson, J. D., Viola, R., and Sargent, D. J. (2013). Saturated linkage map construction in Rubus idaeus using genotyping by sequencing and genome-independent imputation. *BMC Genomics*, *14*(1).

Waugh, R., Ven, M.V.D., Millam, S., Brennan, R. and Powell, W. (1990). The potential use of restriction fragment length polymorphism in *Rubus* breeding. *Acta Horticulturae,* 280, 541-545.

Weber, C.A. (2003). Genetic diversity in black raspberry detected by RAPD markers. *Horticultural Science,* 38, 269-272.

Weir, B.S. (1990). Genetic data analysis. Sinauer Associates Publishers, Sunderland.

Williams, J. A., and West, C. J. (2000). Environmental weeds in Australia and New Zealand: issues and approaches to management. *Austral Ecology*, *25*(5), 425–444.

Williams, J., Kubelik, A., Livak, K., Rafalski, J. and Tingey, S. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research,* 18, 6531-6535.

Yan, J., Yang, X., Shah, T., Sánchez-Villeda, H., Li, J., Warburton, M., Yi, H., Crouch, J.H., and Xu, Y. (2010). High-throughput SNP genotyping with the Goldengate assay in Maize. *Molecular Breeding*, *25*(3), 441–451.

Yan, W., Wang, X., Chen, Q., Zhang, L., Tang, H., Luo, Y. and Liu, Z. (2015). Phylogenetic insight into subgenera *Idaeobatus* and *Malachobatus* (*Rubus*, Rosaceae) inferring from ISH analysis. *Molecular Cytogenetics,* 8:11-24.

CHAPTER THREE

Genetic diversity of blackberry (*Rubus* subgenus *Rubus* Watson) in selected counties in Kenya using SSR Markers

3.1 Abstract

Genetic diversity of blackberry (*Rubus* subgenus *Rubus* Watson) is essential for efficient breeding, and improvement of its pomological traits and yield. In this research, simple sequence markers (SSRs) were used to determine the genetic diversity of 90 blackberry accessions collected from six different counties in Kenya. From 11 SSR markers used to genotype the blackberry accessions, a total of 127 alleles were generated. The average number of alleles (*A*) per locus was 2.00 while the expected heterozygosity (H_E) of the SSR loci varied between 0.34 and 0.50, with a mean of 0.467. The blackberry *PIC* values ranged from 0.33 to 0.37 with a mean of 0.36. H_E of the blackberry accessions were higher than the observed heterozygosities (H_O), having 0.75 and 0.64, respectively. Analysis of Molecular Variance (AMOVA) revealed 95% variability within accessions and 5% ($P<0.01$) among accessions variation. Cluster analysis using the Jaccard's similarity coefficient grouped the accessions into three classes; I, II and III consisting of 31, 52 and 7 accessions, respectively. The clustering was random and did not group the accessions according to their geographical origin, indicating that accessions found in Kenya are closely related. This study detected considerably high levels of genetic diversity within analyzed accessions and could be used in blackberry breeding programs.

Keywords: Simple sequence repeats (SSR), *Rubus* subgenus *Rubus* Watson, genetic diversity, breeding

3.2 Introduction

Genetic diversity of plant species is important to their improvement and provides beginning to accruing benefits of genomics research, counteract genetic erosion and understand evolutionary relationships that leads to design of genetic conservation and breeding strategies (Mason *et al*., 2015; Jacob *et al*., 2017). As such, genetic diversity is vital for incorporation of informed breeding methods into crop breeding operations which is key to the improvement of plant genetic resources. Conventional breeding of blackberry (*Rubus* subgenus *Rubus* Watson) is expensive, time-consuming and labour intensive and hence, advances in molecular techniques would improve on the efficiency, accuracy and cost of breeding this fruit crop. Therefore, there is need for use of DNA information, simulate the available breeding utilities, identify efficient application schemes, have access to effective services in DNA-based diagnostics and integrate DNA information into breeding operations and decisions. (Brennan *et al*., 2014; Peace, 2017). Genetic diversity based on DNA-information has greatly improved breeding of crops by identifying relatedness and phylogeny and by unambiguously ascertaining germplasm identity, verifying and deducing its paternity/parentage, pedigree and distant ancestry (Alice *et al*., 1997; Ru *et al*., 2015).

As mentioned in the previous chapter, blackberries in Kenya are present as wild types. These wild types are the important sources of genetic diversity. They also act as potential sources of breeding materials for blackberry breeding programs, although sometimes they act as sources of natural pests and predators that affect the blackberry crop (Graham *et al*., 1997). Despite their importance in breeding, their status is under threat due to deforestation. The plant introductions (PIs) on the other hand influence the genetic diversity of natural populations by way of gene loss and transfer by pollen.

Blackberry is rich in antioxidants, flavonoids and phenolic compounds and is considered as anti-carcinogenic against oral, oesophageal and colon cancers (Ames *et al.*, 1993; Moyer *et al*., 2002; Bowen-Forbes *et al*., 2010; Overall *et al*., 2017). These beneficial health effects are associated with their antioxidant and anti-inflammatory properties and chemopreventative phytochemicals such as flavonols, phenolic acids, ellagic acid, vitamins C and E, folic acid and b-sitosterol (Tulio *et al*., 2008). There is growing interest in the fruit crop in diets due to its pharmacological properties and health benefits.

Stafne and Clark, (2005) conducted a study on the relatedness of North American blackberry species using the coefficient of relationships to determine the genetic similarity (GS) of these cultivars based on pedigree analysis and detected a coefficient of relationship of 0.00 to 0.74. The apparent high levels of maximum potential similarities and coefficient of

relationship in this research were attributed to higher levels of hybridization in the released cultivars. Most of the studies on genetic diversity of the *Rubus* species have been done in raspberry: *Rubus idaeus* (Parent and Fortin, 1993, Graham and McNichol, 1995, Graham *et al*., 1997), *Rubus occidentalis* (Parent and Page, 1998), *Rubus alceifolius* (Amsellem *et al*., 2000), hybrids of *Rubus idaeus* and *Rubus caesius* (Alice *et al*., 1997), *Rubus occidentalis* (Dossett *et al*., 2011) and *Rubus buergeri* (Miyashita *et al*., 2015). These studies used Random Amplified Polymorphic DNA (RAPD), Restriction Fragment Length Polymorphism (RFLP), and Sequence Characterized Amplified Region (SCAR), Internal Transcribed Spacer (ITS) and Single Sequence Repeats (SSR) markers. The use of markers has made it possible to confirm *Rubus* and hybrids phylogeny and understand their evolution (Alice, 2002). As a result, there has been increased interest in using molecular markers to facilitate blackberry breeding. Multiplexed DNA fingerprinting, characterization of germplasm, development of primers, genetic maps and blackberry expressed sequence tag (EST) libraries, marker-assisted seedling selection, and Quantitative Trait Loci (QTL) mapping have been used in different DNA based studies in blackberry (Lewers *et al*., 2008; Castillo *et al*., 2010; Castro *et al*., 2013; Bassil *et al*., 2016).

Challenges hindering breeding of blackberry include; lack of information on characterization of the genetic diversity and /or population structure within present breeding programs and repositories, difficulty in identifying duplicate accessions in germplasm repositories, searching for promising heterotic groups and selection of core collections. There are no improved cultivars as most blackberry types in Kenya are wild except for only two introductions. The objectives of this study were to (*i*) map blackberry germplasm occurrence in selected counties in Kenya and (*ii*) determine the genetic diversity of wild blackberry types in 6 counties in Kenya and 2 Plant Introductions (PIs) using SSR markers. It is postulated that the findings of this study will resolve the taxonomic uncertainty of duplicate accessions in *in-situ* and *ex-situ* blackberry gene banks.

3.3 Materials and Methods

3.3.1 Study site

Blackberry leaf samples for this research were collected *in-situ* as previously described by Oyoo *et al.* (2015). A total of 90 blackberry accessions (Table 1) were collected from six counties in Kenya. The counties represented were Nandi (0.1036° N, 35.1777° E) (14), Nakuru (0.3031° S, 36.0800° E) (26), Kericho (0.3689° S, 35.2863° E) (16), Uasin Gishu (0.5143° N, 35.2697° E) (7), Laikipia (0.3970° N, 37.1588° E) (6) and Baringo (0.4897° N, 35.7412° E)

(21). The sampling locations were randomly selected and had marked differences in altitude, climatic conditions or cropping systems, geographical features like rivers and mountains and ethnicity living in the area. This was to minimise sampling of duplicates. Global Positioning System (GPS) data was taken and each fruit tree sampled was catalogued.

3.3.2 Plant Material/Collection of germplasm

The blackberry samples taken were coded to reflect the county, district, division, subdivision, village and the collection number. Since most of the blackberry collected were either wild and named by farmers or by the communities at different times, it is difficult to explore their genuine distinct names and pedigree. Blackberry types from different nurseries were treated as independent cultivars in this research irrespective of the age of plants.

3.3.3 Isolation of Genomic DNA

Total deoxyribonucleic acid (DNA) was extracted from each dry young leaf using a modified CTAB protocol (Doyle and Doyle, 1990) for all the 90 accessions. The modification involved omission of the ammonium acetate step. This step is usually done for DNA recovery and removal of PCR inhibitors from the chloroform/aqueous interphase (crude extracts). However, its incorporation yields low quantities of DNA as compared to a longer precipitation time with ice-cold isopropanol. Overnight DNA precipitation time of 12 hours was preferred since blackberry leaf samples have a lot of phenolic compounds.

3.3.4 DNA Quantification

The concentration and purity of the extracted DNA samples were ascertained by using a NanoDrop spectrophotometer-ND 1000 (Thermo Fisher Scientific Inc., USA) and by resolving on 1% agarose gel (1 g of agarose powder in 100 ml of sodium borate buffer). This was to check for degradation and presence of contaminants. Samples with poor quality DNA were re-extracted. The DNA samples were then diluted to a working concentration of 50ng/µl.

3.3.5 PCR Amplification and Microsatellites Analysis

Eleven out of thirteen available blackberry SSR primer sets previously described by Castillo *et al*. (2010) were selected and used to screen 90 blackberry accessions in this study. Primer *RhM031* was uninformative while *RiG001* failed to amplify any blackberry and hybrid accessions and was used to identify raspberry genotypes. Subsequently, they were exempted in SSR data analysis. The SSR primer pairs and sequences are shown in Table 4. The extracted DNA was then subjected to polymerase chain reaction (PCR). PCR amplifications were

performed in a 10 µl volume consisting of 1.4 µl x10 PCR buffer (Thermo Fisher Scientific Inc., USA), 0.1 µl *Taq* polymerase (Thermo Fisher Scientific Inc., USA), 0.8 µl each of 10 pmol forward and reverse primers (Inqaba biotech, S.A), 0.60 µl of 25 mM $MgCl_2$, 4.3 µl of double distilled de-ionized water (ddH_2O) and 2 µl of genomic DNA. Amplification was performed in an Applied Biosystems 2720 thermocycler (Life Technologies Holdings Pte Ltd, Singapore). The amplification was performed under the following conditions: 94°C for 5 minutes, 35 cycles of denaturation at 94°C for 30 seconds, annealing temperature of 50-62°C (Ta, depends on the sequence of the primer) for 30 seconds, and initial extension at 72°C for 2 minutes, followed by a terminal extension at 72°C for 10 min.

3.3.6 Gel Electrophoresis of Polymerase Chain Reaction (PCR) Products

The PCR products were mixed with 6×Orange DNA loading dye (Thermo scientific Corp, Lithuania) and separated on 3% agarose gels (Duchefa Biochemie B.V., The Netherlands) stained with 3 µL ethidium bromide (Invitrogen Corp, U.S.A) in a 1× Sodium Borate (SB) buffer at 60 Volts and a current of 400mA for 2 hours. The separated amplicons were visualized under an Ebox-VX5 gel visualization system (Vilber Lourmat Inc, France). The alleles were scored as absent or present based on the size of the amplified product using a 100bp O'geneRuler ready to use DNA Ladder (Thermo Fisher Scientific Inc., USA).

3.4 Data Analyses

3.4.1 Cluster, Principal Component and Principal Coordinate Analyses

Distances between individual accessions were calculated as a proportion of shared alleles by using DARwin version 6.0 (Perrier *et al*., 2003; Perrier and Jacquemoud-Collet, 2006) using simple matching coefficient based on the following formula:

$$GS_{ij} = \frac{2N}{(N + N)} \qquad 3.0$$

Where; GS_{ij} – Observation of fragments shared by accession *i* and *j*, *Nij* - the number of fragments shared by accession *i* and *j*, *Ni* - amplified fragments in sample *i* and *Nj* - amplified fragments in sample *j* (Nei and Li, 1979). The dissimilarity coefficients were then used to generate an unweighted neighbour-joining tree (Saitou and Nei, 1987) with Jaccard's Similarity Coefficient with a bootstrapping value of 1,000 by using DARwin 6.0.

3.4.2 Analysis of Microsatellite Marker Data

Molecular data were recorded in binary fashion for SSR marker loci amplified. Individuals were scored for the presence (1) or absence (0) of each allele which was treated as

a separate locus. PowerMarker Version 3.25 (Liu and Muse, 2005) was used to calculate statistics on major allele frequencies (*M*) and polymorphism information content (*PIC*) (Botstein *et al.*, 1980) of the SSR primer sets; the genetic distance matrices were computed using PowerMarker with the proportion of shared alleles distance, Dsa (Chakraborty and Jin, 1993):

$$D = \frac{1}{m} \sum_{j=1}^{m} \sum_{i=1}^{a} \min(p, q) \qquad 3.1$$

Where, pij and qij are the frequencies of the i^{th} allele at the j^{th} locus, m is the number of loci examined, aj is the number of alleles at the j^{th} locus. GenAlEx 6.5 (Peakall and Smouse, 2012) and PowerMarker 3.25 (Liu and Muse, 2005) were used to calculate deviations from Hardy-Weinberg equilibrium (HWE), effective number of alleles (A_E) (Kimura and Crow, 1964), observed heterozygosity (*Ho*) and expected heterozygosity (H_E) (Nei, 1973), inbreeding coefficient (F_{IS}), pairwise genetic distance between populations (F_{ST}) (Nei, 1978), Shannon's diversity index (*I*) (Lewontin, 1972) and analysis of molecular variance (AMOVA). The genetic differentiation between the populations was also determined with PowerMarker 3.25 using *PIC*, a measure that allows intra-individual variation to be minimized.

Table 1: Sequences, annealing temperatures and size of bands of sets of 13 primers used to screen 90 blackberry accessions collected from different regions in Kenya.

Primers	Primer sequence (5'-3')y	SSR Motif	Ta (°C)	Expected size (bp)
RiM015F	CGACACCGATCAGAGCTAATTC	$(ATC)_5$	62	344–364
RiM015R	ATAGTTGCATTGGCAGGCTTAT			
RiM017F	GAAACAGGTGGAAAGAAACCTG	$(TG)_6$	59	181–201
RiM017R	CATTGTGCTTATGATGGTTTCG			
RiM019F	ATTCAAGAGCTTAACTGTGGGC	$(AG)_{12}$	59	146–196
RiM019R	CAATATGCCATCCACAGAGAAA			
RiM036F	AGCAACCACCACCTCAACTAAT	$(TG)_7$	51	227–335
RiM036R	CTAGCAGAATCACCTGAGGCTT			
RhM001F	GGTTCGGATAGTTAATCCTCCC	$(CA)_7$	51	229–282
RhM001R	CCAACTGTTGTAAATGCAGGAA			
RhM003F	CCATCTCCAATTCAGTTCTTCC	$(TG)_{10}$	50	173–264
RhM003R	AGCAGAATCGGTTCTTACAAGC			
RhM011F	AAAGACAAGGCGTCCACAAC	$(TC)_{18}$	56	252–346
RhM011R	GGTTATGCTTTGATTAGGCTGG			
RhM018F	CACCAATTGTACACCCAACAAC	$(CTT)_6$	54	363–381
RhM018R	GATTGTGAGCTGGTGTTACCAA			
RhM021F	CAGTCCCTTATAGGATCCAACG	$(TC)_6$	50	252–315
RhM021R	GAACTCCACCATCTCCTCGTAG			
RhM023F	CGACAACGACAATTCTCACATT	$(CAT)_5$	53	116–206
RhM023R	GTTATCAAGCGATCCTGCAGTT			
RhM031F	CAACCTAATGACCAATGCAAGA	$(CT)_9$	50	0, 391–433
RhM031R	GCAGAATCCATTCTCTTGTTGA			
RhM043F	GGACACGGTTCTAACTATGGCT	$(AC)_6$	56	332–386
RhM043R	ATTGTCGCTCCAACGAAGATT			
RiG001F	TGTCCGATCCTTTTCTTTGG	$(AT)_6$	55	–
RiG001R	CGCTTCTTGATCCTTGACTTGT			

3.5 Results

3.5.1 Field Observations

In Kenya, wild relatives of blackberry are found along roadsides, in sections of secondary growth, forest margins and in forest lands while cultivated types are grown in Naivasha (0.7172° S, 36.4310° E) and Limuru (1.1069° S, 36.6431° E) (Figure 1). In addition, the origins of the fruit trees species in Kenya are not well documented. This implies that the parental and progeny structure was mixed in the accessions sampled (Table 2). There is also very little information about the genetic background of the wild blackberry types except for the commercial names given to cultivated types.

3.5.2 DNA Quantification, Gel Electrophoresis and Analysis

All samples extracted exhibited good quality and quantity of DNA for PCR amplification. This was ascertained using a NanoDrop Spectrophotometer at a wavelength of 260/280nm and at an absorbance ratio of 1.8 – 2.0 (Table 3). Contamination by either proteins or phenolic compounds was minimal in this study.

Table 2. List of blackberry germplasm collected in selected counties in Kenya. Characters are based on Blackberry Standardized Phenotyping protocol (Yin, 2017).

SN.	Characterisation code	County	Phenotypic characteristics
1	NKR/NJR/MN/MN/KOR/01	Nakuru	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical stem
2	NKR/NJR/MN/MN/KOR/02	Nakuru	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical stem
3	NKR/NJR/MN/MN/KOR/03	Nakuru	Semi-erect architecture, small sparse thorns visible, asymmetrical stem symmetry
4	NKR/NJR/MN/MN/KOR/04	Nakuru	Semi-erect architecture, small, sparse thorns visible, somewhat symmetrical stem
5	NKR/NJR/MN/MN/SIG/01	Nakuru	Semi-erect architecture, small sparse thorns visible, asymmetrical stem
6	NKR/NJR/NES/NES/TRT/01	Nakuru	Semi-erect architecture, medium thorns spaced evenly, asymmetrical
7	NKR/NJR/NES/NES/TRT/02	Nakuru	Semi-erect architecture, medium to long thorns sporadic, asymmetrical
8	NKR/NJR/NES/NES/TRT/03	Nakuru	Semi-erect architecture, medium thorns spaced evenly, asymmetrical
9	NKR/NJR/NES/NES/KIM/01	Nakuru	Semi-erect architecture, medium to long thorns evenly spaced, asymmetrical
10	NKR/MOL/MS/MKJ/MAT/01	Nakuru	Semi-erect architecture, medium to long thorns evenly spaced, asymmetrical
11	NKR/MOL/MS/MKJ/MAT/02	Nakuru	Semi-erect architecture, medium to long thorns evenly spaced, asymmetrical
12	NKR/MOL/MS/MKJ/MAT/03	Nakuru	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical
13	NKR/MOL/MS/SAC/GSU/01	Nakuru	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical
14	NKR/MOL/MS/SAC/GSU/02	Nakuru	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical
15	NKR/ELB/ELB/SAL/ARI/01	Nakuru	Semi-erect architecture, small to medium thorns, sporadically spaced, somewhat asymmetrical
16	NKR/ELB/ELB/SAL/ARI/02	Nakuru	Semi-erect architecture, small sparse thorns visible, slightly symmetrical
17	NKR/ELB/ELB/SAL/ARI/03	Nakuru	Semi-erect architecture, medium to long thorns evenly spaced, asymmetrical
18	NKR/SE/BAH/DUN/GIT/01	Nakuru	Semi-erect architecture, medium to long thorns evenly spaced, asymmetrical
19	NKR/SE/BAH/DUN/GIT/02	Nakuru	Semi-erect architecture, medium to long thorns evenly spaced, somewhat asymmetrical
20	NKR/SE/BAH/DUN/GIT/03	Nakuru	Semi-erect architecture, medium to long thorns evenly spaced, slightly asymmetrical
21	KCO/CBA/CHY/CHY/UNL/01	Kericho	Semi-erect architecture, small, sparse thorns visible, somewhat symmetrical stem
22	KCO/CBA/CHY/CHY/UNL/02	Kericho	Semi-erect architecture, small sparse thorns visible, symmetrical

Table 2: continued

23	KCO/CBA/CHY/CHY/UNL/03	Kericho	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical
24	KCO/CBA/CHY/CHY/UNL/04	Kericho	Semi-erect architecture, medium thorns, spaced evenly, asymmetrical
25	KCO/CBA/CHY/CHY/UNL/05	Kericho	Semi-erect architecture, medium to long thorns sporadic, asymmetrical
26	KCO/CBA/CHY/CHY/UNL/06	Kericho	Semi-erect architecture, small to medium thorns, sporadically spaced, symmetrical
27	KCO/CBA/CHY/CHY/UNL/07	Kericho	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical
28	KCO/CBA/KCO/AMI/CSD/1	Kericho	Semi-erect architecture, medium thorns, spaced evenly, very slightly asymmetrical
29	KCO/KLN/KCO/CSR/CSR/01	Kericho	Semi-erect architecture, medium to long thorns evenly spaced, slightly asymmetrical
30	KCO/KLN/KCO/CSR/CSR/02	Kericho	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical
31	KCO/KLN/KCO/CSR/CSR/03	Kericho	Semi-erect architecture, medium to long thorns evenly spaced, asymmetrical
32	KCO/KLN/KCO/CSR/CSR/04	Kericho	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical
33	KCO/LTN/BRT/CMB/KER/01	Kericho	Semi-erect architecture, small, sparse thorns visible, asymmetrical
34	KCO/LTN/BRT/CMB/KER/02	Kericho	Semi-erect architecture, medium to long thorns sporadic, slightly asymmetrical
35	KCO/LTN/BRT/CMB/KER/03	Kericho	Semi-erect architecture, medium to long thorns evenly spaced, very slightly asymmetrical
36	KCO/LTN/BRT/CMB/KER/04	Kericho	Semi-erect architecture, medium to long thorns sporadic, asymmetrical
37	BRG/TIN/TOR/LM/MAK/01	Baringo	Semi-erect architecture, medium to long thorns evenly spaced, asymmetrical
38	BRG/TIN/TOR/LM/MAK/02	Baringo	Semi-erect architecture, medium to long thorns evenly spaced, somewhat asymmetrical
39	BRG/TIN/TOR/LM/MAK/03	Baringo	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical
40	BRG/TIN/TOR/LM/MAK/04	Baringo	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical
41	BRG/TIN/TOR/LM/MAK/05	Baringo	Semi-erect architecture, medium thorns, spaced evenly, asymmetrical
42	BRG/TIN/TOR/LM/MAK/06	Baringo	Semi-erect architecture, small, sparse thorns visible, slightly asymmetrical
43	BRG/TIN/TOR/LM/MAK/07	Baringo	Semi-erect architecture, small, sparse thorns visible, asymmetrical
44	BRG/TIN/TIN/TOR/CHE/01	Baringo	Semi-erect architecture, small, sparse thorns visible, asymmetrical
45	BRG/TIN/TIN/TOR/CHE/02	Baringo	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical

Table 2: continued

46	BRG/TIN/TIN/TOR/CHE/03	Baringo	Semi-erect architecture, medium to long thorns evenly spaced, asymmetrical
47	BRG/TIN/TIN/TOR/CHE/04	Baringo	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical
48	BRG/ERN/LEM/LC/MSO/01	Baringo	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical
49	BRG/ERN/LEM/LC/MSO/02	Baringo	Semi-erect architecture, medium to long thorns evenly spaced, asymmetrical
50	BRG/ERN/IGE/MM/KIN/01	Baringo	Semi-erect architecture, medium to long thorns evenly spaced, asymmetrical
51	BRG/ERN/IGE/MM/KIN/02	Baringo	Semi-erect architecture, medium to long thorns sporadic, asymmetrical
52	BRG/ERN/IGE/MM/KIN/03	Baringo	Semi-erect architecture, small to medium thorns, sporadically spaced, slightly asymmetrical
53	BRG/ERN/IGE/MM/KIN/04	Baringo	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical
54	BRG/ERN/TIM/MBE/KMA/01	Baringo	Semi-erect architecture, small to medium thorns, sporadically spaced, slightly asymmetrical
55	BRG/ERN/TIM/MBE/KMA/02	Baringo	Semi-erect architecture, medium to long thorns evenly spaced, asymmetrical
56	BRG/ERN/TIM/MBE/KMA/03	Baringo	Semi-erect architecture, long thorns closely spaced, slightly asymmetrical
57	BRG/ERN/TIM/MBE/KMA/04	Baringo	Semi-erect architecture, small, sparse thorns visible, slightly asymmetrical
58	UG/KIP/ES/BF/BYT/01	Uasin Gishu	Semi-erect architecture, medium to long thorns closely spaced, slightly asymmetrical
59	NDI/NN/KUR/CKO/SUR/01	Nandi	Semi-erect architecture, small to medium thorns, sporadically spaced, slightly asymmetrical
60	NDI/NN/KUR/CKO/SUR/02	Nandi	Semi-erect architecture, medium to long thorns sporadic, symmetrical
61	NDI/NN/KUR/CKO/SUR/03	Nandi	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical
62	NDI/NN/KSB/KSB/KBA/01	Nandi	Semi-erect architecture, medium to long thorns sporadic, asymmetrical
63	NDI/NN/KSB/KSB/KBA/02	Nandi	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical
64	NDI/NN/KSB/KSB/KBA/03	Nandi	Semi-erect architecture, medium to long thorns evenly spaced, asymmetrical
65	NDI/KBY/SGO/SEP/SEP/01	Nandi	Semi-erect architecture, medium to long thorns evenly spaced, very slightly asymmetrical
66	NDI/KBY/SGO/SEP/SEP/02	Nandi	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical
67	NDI/KBY/SGO/SEP/SEP/03	Nandi	Semi-erect architecture, medium to long thorns sporadic, asymmetrical
68	NDI/NN/BAR/BAR/UEAB/01	Nandi	Semi-erect architecture, medium thorns, spaced evenly, asymmetrical

Table 2: continued

69	NDI/NN/BAR/BAR/UEAB/02	Nandi	Semi-erect architecture, small to medium thorns, sporadically spaced, slightly asymmetrical
70	NDI/ CSY/CMU/KKG/KKI/01	Nandi	Semi-erect architecture, small to medium thorns, sporadically spaced, slightly asymmetrical
71	NDI/ CSY/CMU/KKG/KKI/02	Nandi	Semi-erect architecture, medium thorns, closely spaced, asymmetrical
72	NDI/ CSY/CMU/KKG/KKI/03	Nandi	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical
73	LC/LKP/NYA/LKP/LU/01	Laikipia	Semi-erect architecture, medium to long thorns evenly spaced, asymmetrical
74	LC/LKP/NYA/LKP/LU/02	Laikipia	Semi-erect architecture, medium thorns, spaced evenly, asymmetrical
75	LC/LKW/LUM/RS/RS/01	Laikipia	Semi-erect architecture, medium to long thorns evenly spaced, asymmetrical
76	LC/LKW/LUM/RS/RS/02	Laikipia	Semi-erect architecture, medium to long thorns sporadic, asymmetrical
77	LC/LKW/GMA/KRI/RK/01	Laikipia	Semi-erect architecture, medium to long thorns evenly spaced, slightly asymmetrical
78	UG/KKB/ABK/KPG/CHES/01	Uasin Gishu	Semi-erect architecture, small to medium thorns, sporadically spaced, very slightly asymmetrical
79	UG/KKB/ABK/KPG/CHES/02	Uasin Gishu	Semi-erect architecture, medium thorns, closely spaced, symmetrical
80	UG/KKB/ABK/KPG/CHES/03	Uasin Gishu	Semi-erect architecture, small, sparse thorns visible, asymmetrical
81	UG/KKB/ABK/KPG/FLX/01	Uasin Gishu	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical
82	UG/KKB/ABK/CGA/ANG/01	Uasin Gishu	Semi-erect architecture, medium thorns, closely spaced, very slightly asymmetrical
83	UG/KSS/KPO/LEN/PLT/O1	Uasin Gishu	Semi-erect architecture, medium thorns, spaced evenly, very slightly asymmetrical
84	LC/LKW/GMA/KRI/RK/02	Laikipia	Semi-erect architecture, medium thorns, closely spaced, asymmetrical
85	CV/RBN/01	Introduction	Erect, small to medium thorns, sporadically spaced, asymmetrical
86	CV/BYN/01	Introduction	Trailing, small, sparse thorns visible, slightly asymmetrical
87	NKR/NJR/EGER/EGER/NGU/01	Nakuru	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical stem
88	NKR/NJR/EGER/EGER/F3/02	Nakuru	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical stem
89	NKR/NJR/EGER/EGER/F3/01	Nakuru	Semi-erect architecture, small to medium thorns, sporadically spaced, asymmetrical stem
90	NKR/NJR/EGER/EGER/F7/01	Nakuru	erect, long thorns, closely spaced, slightly asymmetrical

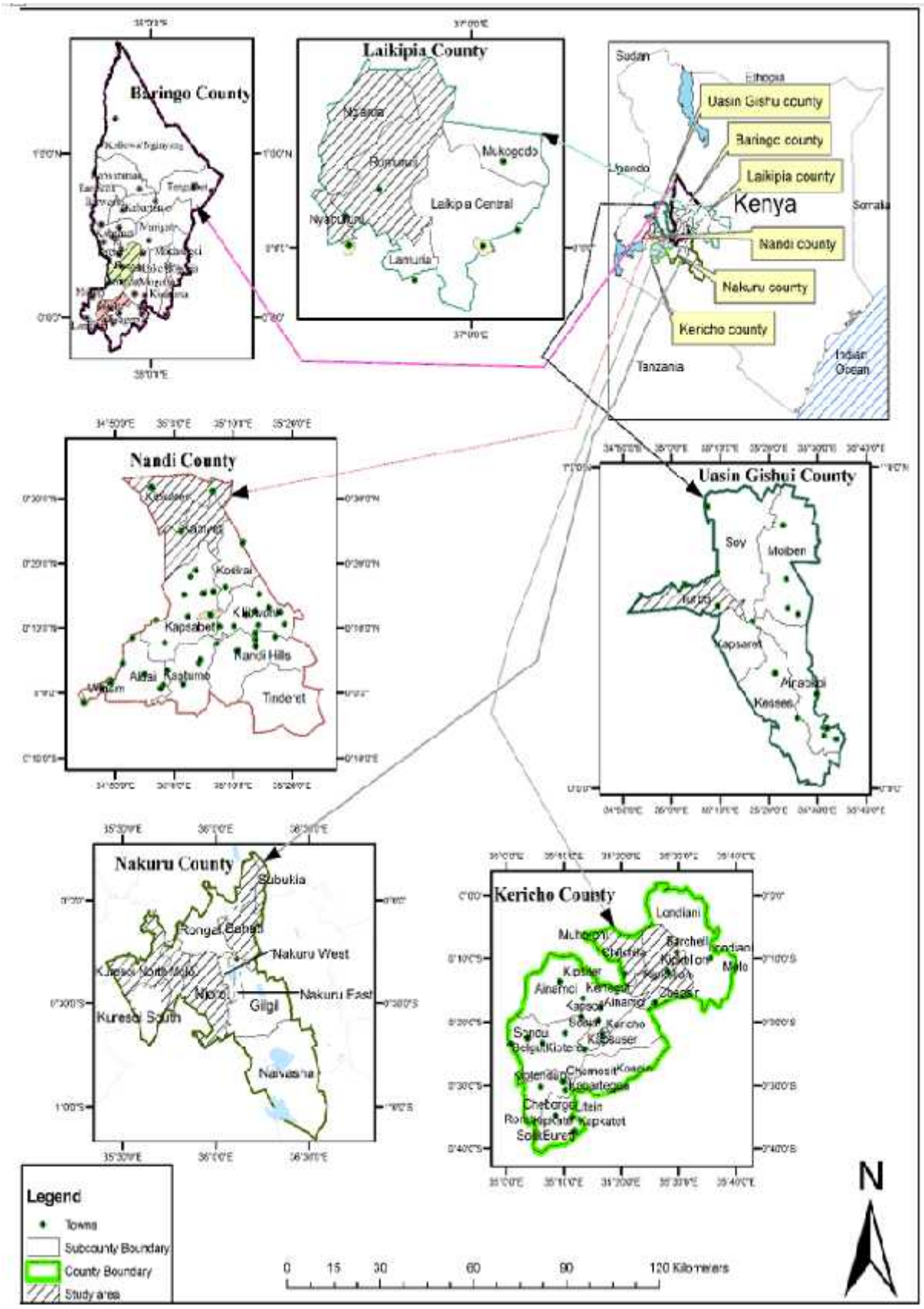

Figure 1: Map showing blackberry germplasm occurrence in 6 counties in Kenya as in Table 1 generated using *ArcGIS*.

Table 1: Concentration and quality of DNA isolated from blackberry as determined by a NanoDrop Spectrophotometer.

Accession	Nucleic Acid (ng/µl)	A260/280	Accession	Nucleic Acid (ng/µl)	A260/280	Accession	Nucleic Acid (ng/µl)	A260/280	Accession	Nucleic Acid (ng/µl)	A260/280
1	1219.8	1.87	28	947.6	1.52	55	628.3	1.28	82	32	2.19
2	1497.2	1.92	`29	1397.8	1.97	56	415.9	2.13	83	117.6	2.09
3	645.9	1.65	30	1132.2	1.91	57	1363.9	1.93	84	582.6	2.95
4	1290.6	2.11	31	665.2	1.87	58	1438.6	2.05	85	360.3	4.15
5	111.8	1.99	32	2659.2	2	59	507.4	1.98	86	652	2.09
6	1583.9	2.13	33	552.7	2.98	60	486.1	1.86	87	930.5	1.89
7	1540.1	2.06	34	336.1	2.06	61	1152.3	2.03	88	135.9	1.91
8	601.1	2.12	35	506.5	1.71	62	886	2.09	89	58.4	2.11
9	854.5	3.73	36	565.2	2	63	1135.6	1.99	90	1384.5	1.73
10	1384.5	1.73	37	400.4	2.18	64	374	1.8			
11	884.8	2.09	38	390.3	2.01	65	1762.9	1.67			
12	280.2	1.94	39	1855.2	1.69	66	1013.5	1.87			
13	383.1	1.85	40	1185.3	2.14	67	986.6	1.53			
14	799.6	1.84	41	1511.6	1.02	68	1258.7	1.95			
15	1182.2	1.82	42	141.1	2.01	69	1058.7	2.05			
16	567.9	1.87	43	324.3	2.19	70	580.8	2.76			
17	854.5	3.73	44	258.2	1.98	71	1456.6	2.03			
18	156.9	1.56	45	275.2	1.68	72	2091.1	1.96			
19	82.4	0.95	46	1184.4	2.1	73	1355	3.36			
20	1363.9	1.93	47	986.1	2.06	74	191	1.83			
21	11.4	1.96	48	415.3	1.67	75	142.3	2.09			
22	1152.4	1.49	49	1999.9	1.4	76	135.9	1.91			
23	11.4	1.96	50	603.8	2.13	77	142.3	2.09			
24	1271.3	1.93	51	31.9	1.84	78	87.1	2.15			
25	1055.3	2.04	52	50.9	1.65	79	191	1.83			
26	613.2	1.75	53	507.4	1.98	80	227.6	2.13			
27	2776.5	2.06	54	577.6	2.08	81	62.1	2.12			

3.5.3 Diversity Indices of SSR Loci in Blackberry Accessions

The effective number of alleles (A_E) per microsatellite locus varied from 1.51 (*RhM021*) to 1.99 (*RhM003* and *RiM019*) with an average of 1.88 (Table 2). The average value of Shannon's diversity index (*I*) across the primer sets was 0.66 and ranged from 0.52 (*RhM021*) to 0.69 (*RiM017* and *RhM043*). The average observed heterozygosity and expected heterozygosity values were 0.54 and 0.57, respectively. The least H_O was 0.19 (*RhM019*) while the highest was 0.88 (*RiM036* and *RhM003*). Among the blackberry accessions, H_E ranged from 0.34 (*RhM021*) to 0.50 (*RhM043*), respectively while inbreeding coefficient (F_{IS}) ranged from -0.86 (*RhM003*) to 0.71 (*RiM019*). The pairwise genetic distances (F_{ST}) ranged from 0.00 for *RhM003* to (0.17) *RiM017*. This study revealed moderate to significant differentiation ($0.05>F_{ST}\geq0.15$) within some blackberry accessions (Table 3). Additionally, high *PIC* values were observed for markers *RiM019, RiM017, RhM043, RiM015, RhM018* and *RhM001* (Table 4).

Table 4. Estimates of genetic diversity of SSR loci used to screen 90 blackberry accessions sampled from 6 counties in Kenya.

SSR loci	Range of bps	A_E	*I*	H_O	H_E	F_{IS}	F_{ST}	*M*	*PIC*
RhM011	252–346	1.87	0.66	0.31	0.47	0.31	0.14	0.63	0.36
RiM019	146–196	1.84	0.65	0.19	0.46	0.71	0.05	0.32	0.73
RiM017	181–201	1.99	0.69	0.28	0.50	0.36	0.17	0.39	0.61
RhM043	332–386	2.00	0.69	0.60	0.50	-0.05	0.03	0.32	0.69
RiM015	344–364	1.96	0.68	0.50	0.49	-0.16	0.04	0.26	0.75
RhM001	229–282	1.99	0.69	0.63	0.50	-0.36	0.04	0.48	0.63
RiM036	227–335	1.98	0.69	0.88	0.49	-0.80	0.01	0.58	0.49
RhM018	363–381	1.97	0.68	0.56	0.49	-0.64	0.01	0.64	0.51
RhM003	173–264	1.99	0.69	0.88	0.50	-0.86	0.00	0.77	0.37
RhM021	252–315	1.51	0.52	0.43	0.34	-0.38	0.10	0.57	0.37
RhM023	116–206	1.61	0.57	0.41	0.48	-0.37	0.01	0.51	0.38
Mean		**1.88**	**0.66**	**0.54**	**0.46**	**-0.23**	**0.06**	**0.51**	**0.52**
S.E		0.05	0.20	0.16	0.17				

Effective number of alleles (A_E), Shannon's diversity index (I), observed heterozygosity (H_O), expected heterozygosity (H_E), inbreeding coefficient (F_{IS}), pairwise genetic distance between populations (F_{ST}), major allele frequency (*M*) and Polymorphic Information Content (*PIC*).

3.5.4 Diversity Indices of Blackberry Accessions

The effective number of alleles per locus (A*E*) varied from 1.65 in accession NAK/NJR/NES/NES/TRT/01 to 7.56 in accession KCO/CBA/CHY/CHY/UNL/02 with a mean of 3.62 (Table 5). The observed number of alleles (*A*) varied from 6.00 (Accessions NKR/NJR/NES/NES/TRT/01, LC/LKW/LUM/RS/RS/01 and LC/LKW/GMA/KRI/RK/01) to 12.000 (Accessions KCO/CBA/KCO/AMI/CSD/01, BRG/TIN/TOR/LM/MAK/06, NDI/CSY/CMU/KKG/KKI/03, KCO/KLN/KCO/CSR/01, BRG/TIN/TIN/TOR/CHE/01) with a mean of 9.26. The observed heterozygosity (*HO*) ranged from 0.27 (Accession NAK/ELB/ELB/SAL/ARI/02) to 0.82 (Accessions KCO/KLN/KCO/CSR/CSR/01, BRG/ERN/TIM/MBE/KMA/03) with a mean of 0.54. The expected heterozygosity (H*E*) ranged from 0.41 (Accession NAK/NJR/NES/NES/TRT/01) to 0.91 (Accession KCO/CBA/CHY/CHY/UNL/02) with a mean of 0.7165. The *PIC* values varied from 0.28 (BRG/ERN/LEM/LC/MSO/02) to 0.38. Eighteen accessions had negative fixation indices (*FIS*), indicating high levels of intra-population genetic diversity (Table 3). The average number of alleles per locus (*A*) for all blackberry populations obtained from all regions was 9.26.

Table 2: Genetic diversity indices for a population of the 90 accessions of blackberry studied from 6 counties in Kenya.

Accession code	M	A	I	A_E	H_O	H_E	*PIC*	F_{IS}
NKR/NJR/MN/MN/KOR/01	0.70	7.00	1.07	1.85	0.55	0.48	0.33	-0.19
NKR /NJR/MN/MN/KOR/02	0.57	9.00	1.61	3.06	0.55	0.71	0.37	0.19
NKR /NJR/MN/MN/KOR/03	0.61	11.00	2.16	6.54	0.45	0.89	0.36	0.46
NKR /NJR/MN/MN/KOR/04	0.52	8.00	1.55	3.03	0.36	0.70	0.37	0.46
NKR /NJR/MN/MN/SIG/01	0.57	7.00	1.18	2.07	0.45	0.54	0.37	0.12
NKR /NJR/NES/NES/TRT/01	0.70	6.00	0.90	1.65	0.45	0.41	0.33	-0.16
NKR /NJR/NES/NES/TRT/02	0.52	9.00	1.51	2.69	0.64	0.66	0.37	-0.01
NKR /NJR/NES/NES/TRT/03	0.52	10.00	2.02	5.50	0.36	0.86	0.37	0.56
NKR /NJR/NES/NES/KIM/01	0.61	10.00	1.94	4.75	0.45	0.83	0.36	0.42
NKR /MOL/MS/MKJ/MAT/01	0.52	7.00	1.29	2.33	0.36	0.60	0.37	0.36
NKR /MOL/MS/MKJ/MAT/02	0.61	11.00	2.16	6.54	0.45	0.89	0.36	0.46
NKR /MOL/MS/MKJ/MAT/03	0.57	9.00	1.70	3.51	0.45	0.75	0.37	0.36
NKR /MOL/MS/SAC/GSU/01	0.61	9.00	1.70	3.51	0.45	0.75	0.36	0.36
NKR /MOL/MS/SAC/GSU/02	0.57	10.00	1.67	3.10	0.73	0.71	0.37	-0.07
NKR /ELB/ELB/SAL/ARI/01	0.61	7.00	1.29	2.33	0.36	0.60	0.36	0.36
NKR /ELB/ELB/SAL/ARI/02	0.57	8.00	1.64	3.46	0.27	0.74	0.37	0.62
NKR /ELB/ELB/SAL/ARI/03	0.65	8.00	1.45	2.66	0.45	0.65	0.35	0.27
NKR /SE/BAH/DUN/GIT/01	0.57	9.00	1.79	4.03	0.36	0.79	0.37	0.52
NKR /SE/BAH/DUN/GIT/02	0.52	7.00	1.07	1.85	0.55	0.48	0.37	-0.19
NKR /SE/BAH/DUN/GIT/03	0.61	10.00	1.86	4.10	0.55	0.79	0.36	0.28
KCO/CBA/CHY/CHY/UNL/01	0.65	7.00	1.29	2.33	0.36	0.60	0.35	0.36

Table 5: Continued

KCO/CBA/CHY/CHY/UNL/02	0.65	11.00	2.22	7.56	0.36	0.91	0.35	0.58
KCO/CBA/CHY/CHY/UNL/03	0.52	9.00	1.70	3.51	0.45	0.75	0.37	0.36
KCO/CBA/CHY/CHY/UNL/06	0.57	9.00	1.70	3.51	0.45	0.75	0.37	0.36
KCO/CBA/CHY/CHY/UNL/07	0.61	8.00	1.35	2.35	0.55	0.60	0.36	0.05
KCO/CBA/KCO/AMI/CSD/01	0.52	12.00	2.22	6.72	0.64	0.89	0.37	0.25
KCO/KLN/KCO/CSR/CSR/01	0.65	12.00	2.07	4.94	0.82	0.84	0.35	-0.03
KCO/KLN/KCO/CSR/CSR/02	0.65	9.00	1.51	2.69	0.64	0.66	0.35	-0.01
KCO/KLN/KCO/CSR/CSR/03	0.61	9.00	1.61	3.06	0.55	0.71	0.36	0.19
KCO/KLN/KCO/CSR/CSR/04	0.52	8.00	1.35	2.35	0.55	0.60	0.37	0.05
KCO/LTN/BRT/CMB/KER/01	0.52	11.00	2.16	6.54	0.45	0.89	0.37	0.46
KCO/LTN/BRT/CMB/KER/02	0.70	11.00	1.92	4.17	0.73	0.80	0.33	0.04
KCO/LTN/BRT/CMB/KER/03	0.52	10.00	1.77	3.56	0.64	0.75	0.37	0.11
KCO/LTN/BRT/CMB/KER/04	0.65	9.00	1.61	3.06	0.55	0.71	0.35	0.19
BRG/TIN/TOR/LM/MAK/01	0.65	11.00	2.16	6.54	0.45	0.89	0.35	0.46
BRG/TIN/TOR/LM/MAK/02	0.52	10.00	1.67	3.10	0.73	0.71	0.37	-0.07
BRG/TIN/TOR/LM/MAK/03	0.57	8.00	1.35	2.35	0.55	0.60	0.37	0.05
BRG/TIN/TOR/LM/MAK/04	0.52	9.00	1.51	2.69	0.64	0.66	0.37	-0.01
BRG/TIN/TOR/LM/MAK/05	0.70	11.00	1.92	4.17	0.73	0.80	0.33	0.04
BRG/TIN/TOR/LM/MAK/06	0.57	12.00	2.22	6.72	0.64	0.89	0.37	0.25
BRG/TIN/TOR/LM/MAK/07	0.65	8.00	1.45	2.66	0.45	0.65	0.35	0.27
BRG/TIN/TIN/TOR/CHE/01	0.65	12.00	2.22	6.72	0.64	0.89	0.35	0.25
BRG/TIN/TIN/TOR/CHE/02	0.52	8.00	1.35	2.35	0.55	0.60	0.37	0.05
BRG/TIN/TIN/TOR/CHE/03	0.52	10.00	1.94	4.75	0.45	0.83	0.37	0.42
BRG/TIN/TIN/TOR/CHE/04	0.52	9.00	1.61	3.06	0.55	0.71	0.37	0.19
BRG/ERN/LEM/LC/MSO/01	0.70	11.00	2.08	5.63	0.55	0.86	0.33	0.34
BRG/ERN/LEM/LC/MSO/02	0.78	10.00	1.77	3.56	0.64	0.75	0.28	0.11
BRG/ERN/IGE/MM/KIN/01	0.74	8.00	1.35	2.35	0.55	0.60	0.31	0.05
BRG/ERN/IGE/MM/KIN/02	0.57	9.00	1.79	4.03	0.36	0.79	0.37	0.52
BRG/ERN/IGE/MM/KIN/03	0.70	11.00	1.92	4.17	0.73	0.80	0.33	0.04
BRG/ERN/IGE/MM/KIN/04	0.57	10.00	1.86	4.10	0.55	0.79	0.37	0.28
BRG/ERN/TIM/MBE/KMA/01	0.61	8.00	1.35	2.35	0.55	0.60	0.36	0.05
BRG/ERN/TIM/MBE/KMA/02	0.65	9.00	1.70	3.51	0.45	0.75	0.35	0.36
BRG/ERN/TIM/MBE/KMA/03	0.61	11.00	1.83	3.61	0.82	0.76	0.36	-0.13
BRG/ERN/TIM/MBE/KMA/04	0.65	11.00	2.16	6.54	0.45	0.89	0.35	0.46
UG/KIP/ES/BF/BYT/01	0.52	11.00	2.16	6.54	0.45	0.89	0.37	0.46
NDI/NN/KUR/CKO/SUR/01	0.52	11.00	1.92	4.17	0.73	0.80	0.37	0.04
NDI/NN/KUR/CKO/SUR/02	0.61	9.00	1.51	2.69	0.64	0.66	0.36	-0.01
NDI/NN/KUR/CKO/SUR/03	0.57	10.00	1.77	3.56	0.64	0.75	0.37	0.11
NDI/NN/KSB/KSB/KBA/01	0.65	10.00	1.67	3.10	0.73	0.71	0.35	-0.07
NDI/NN/KSB/KSB/KBA/02	0.65	9.00	1.51	2.69	0.64	0.66	0.35	-0.01
NDI/NN/KSB/KSB/KBA/03	0.57	8.00	1.24	2.09	0.64	0.55	0.37	-0.22
NDI/KBY/SGO/SEP/SEP/01	0.65	10.00	1.77	3.56	0.64	0.75	0.35	0.11
NDI/KBY/SGO/SEP/SEP/02	0.61	10.00	1.86	4.10	0.55	0.79	0.36	0.28
NDI/KBY/SGO/SEP/SEP/03	0.70	9.00	1.79	4.03	0.36	0.79	0.33	0.52
NDI/NN/BAR/BAR/UEAB/01	0.61	10.00	1.77	3.56	0.64	0.75	0.36	0.11

Table 5: Continued

NDI/NN/BAR/BAR/UEAB/02	0.70	9.00	1.51	2.69	0.64	0.66	0.33	-0.01
NDI/ CSY/CMU/KKG/KKI/01	0.57	9.00	1.61	3.06	0.55	0.71	0.37	0.19
LC/LKP/NYA/LKP/LU/01	0.61	8.00	1.35	2.35	0.55	0.60	0.36	0.05
LC/LKP/NYA/LKP/LU/02	0.57	9.00	1.51	2.69	0.64	0.66	0.37	-0.01
LC/LKW/LUM/RS/RS/01	0.70	6.00	1.01	1.83	0.36	0.48	0.33	0.20
LC/LKW/LUM/RS/RS/02	0.65	11.00	2.08	5.63	0.55	0.86	0.35	0.34
LC/LKW/GMA/KRI/RK/01	0.70	6.00	1.01	1.83	0.36	0.48	0.33	0.20
UG/KKB/ABK/KPG/CHES/01	0.57	8.00	1.35	2.35	0.55	0.60	0.37	0.05
UG/KKB/ABK/KPG/CHES/02	0.70	8.00	1.35	2.35	0.55	0.60	0.33	0.05
UG/KKB/ABK/KPG/CHES/03	0.61	8.00	1.55	3.03	0.36	0.70	0.36	0.46
UG/KKB/ABK/KPG/FLX/01	0.70	11.00	1.92	4.17	0.73	0.80	0.33	0.04
UG/KKB/ABK/CGA/ANG/01	0.57	8.00	1.45	2.66	0.45	0.65	0.37	0.27
UG/KSS/KPO/LEN/PLT/O1	0.52	9.00	1.61	3.06	0.55	0.71	0.37	0.19
LC/LKW/GMA/KRI/RK/02	0.61	10.00	1.67	3.10	0.73	0.71	0.36	-0.07
CV/RBN/01	0.57	11.00	1.92	4.17	0.73	0.80	0.37	0.04
CV/BYN/01	0.70	10.00	1.86	4.10	0.55	0.79	0.33	0.28
NAK/NJR/EGER/EGER/NG/01	0.70	7.00	1.18	2.07	0.45	0.54	0.33	0.12
NAK/NJR/EGER/EGER/F3/02	0.52	9.00	1.51	2.69	0.64	0.66	0.37	-0.01
NAK/NJR/EGER/EGER/F3/01	0.65	7.00	1.18	2.07	0.45	0.54	0.35	0.12
NAK/NJR/EGER/EGER/F7/01	0.52	10.00	1.77	3.56	0.64	0.75	0.37	0.11
Mean	**0.61**	**9.26**	**1.66**	**3.62**	**0.54**	**0.72**	**0.36**	
S.E		0.98	0.17	0.38	0.06	0.08		

Effective number of alleles (A_E), Shannon's diversity index (I), observed heterozygosity (H_O), expected heterozygosity (H_E), inbreeding coefficient (F_{IS}), Allelic richness (A), major allele frequency (M) and Polymorphic information content (PIC).

The genetic identity and distances (Nei, 1978) of the blackberry genetic resources for all the counties included in this study are shown in Table 6. There was variation in genetic identity in the accessions. A genetic identity value of 0.862 was noted between populations from Laikipia and Introductions from RSA. The highest genetic identity of 0.988 was observed between populations from Kericho and Baringo. This study revealed the greatest genetic distance (0.148) to be between populations from Laikipia and Kericho. This study showed significant genetic distances between the accessions from RSA and the wild types from different counties in Kenya.

Table 3: *Nei's* genetic identity (above diagonal) and genetic distance (below diagonal) among blackberry accessions.

Counties	Baringo	PIs	Kericho	Laikipia	Nakuru	Nandi	Uasin Gishu
Baringo	****	0.910	0.988	0.971	0.984	0.976	0.977
PIs (RSA)	0.094	****	0.923	0.862	0.931	0.869	0.875
Kericho	0.012	0.077	****	0.964	0.982	0.970	0.964
Laikipia	0.029	0.148	0.037	****	0.969	0.951	0.960
Nakuru	0.017	0.072	0.018	0.031	****	0.953	0.976
Nandi	0.024	0.140	0.031	0.051	0.048	****	0.977
Uasin Gishu	0.024	0.134	0.037	0.041	0.024	0.024	****

Plant Introductions (PIs) RSA: Republic of South Africa

The Analysis of Molecular Variance (AMOVA) for blackberry partitioned the genetic variance among and within the accessions and revealed that most of the variability was within the accessions (95%) (Table 7). The genetic variance was significant ($P<0.01$) among the accessions and accounted for 5% of the total variation. The hierarchical subdivision of the summary of Shannon's statistics indicated that most molecular variance was within populations accounting for 90.57% of the total genetic variation with only 9.43% of the molecular variation to the defined counties (Table 8). Only 9.43% of the molecular variance distinguished the six populations from Nakuru, Kericho, Nandi, Laikipia, Uasin Gishu and the RSA) ($P<0.01$).

Table 7: Analysis of Molecular Variance (AMOVA) of the diversity of 90 blackberry accessions collected from selected counties in Kenya.

Source of variation	DF	SS	MS	Estimated. Variation.	Total Variation. %	P(*f*)
Among accessions	6	39.06	6.51	0.20	5	0.01
Within accessions	83	341.57	4.11	4.12	95	0.01
Total	89	380.63		4.31	100	

3.5.5 Cluster Analysis and Population Structure

The Unweighted Pair Group Method with Arithmetic means (UPGMA) dendrogram generated from SSR marker information using Jaccard's similarity coefficient showed phylogenetic relationships among 90 blackberry accessions (Figure 5). The phylogenetic tree was divided into 3 distinct clusters. However, the cluster analysis failed to clearly cluster the accessions based on their regions of collection (Figure 4). The results showed that accessions collected from different counties clustered together, especially those from Kericho County (group II). Principal Coordinate Analysis (PCoA) confirmed results from the cluster analysis and showed that most accessions overlapped (Figure 2 and 3). The first three axes accounted

for 55.48% of the total variations with each axes explaining 30.04%, 13.53% and 11.91% of the variation, respectively at 95% confidence interval (Figure 3).

Table 8: Shannon statistics summary of the 90 blackberry accessions sampled from selected counties in Kenya.

Source of variation	Degrees freedom DF	Log-like. Chi-Sq G-Test	Shannon inform. sH	Percent of total inform.	Diversity estimate exp(sH)	Estimated probability P(r)
Among accessions	6	9.37	0.05	9.43	1.05	0.01
Within accessions	83	89.97	0.50	90.57	1.65	0.99
Total	89	99.34	0.55	100.00	1.74	

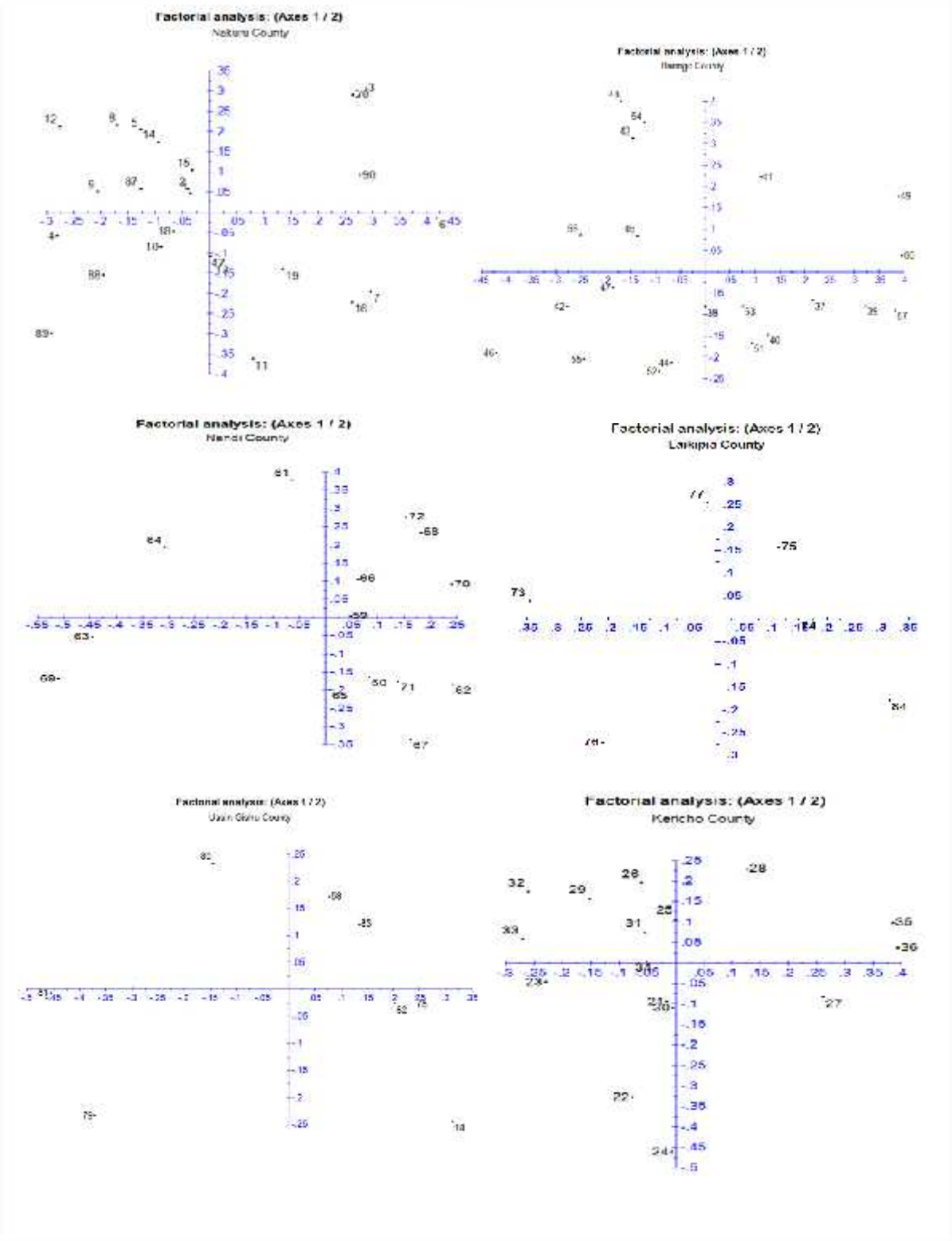

Figure 2: PCoA of axes 1 and 2 based on dissimilarity of 11 SSR markers across 90 blackberry accessions from different regions in Kenya.

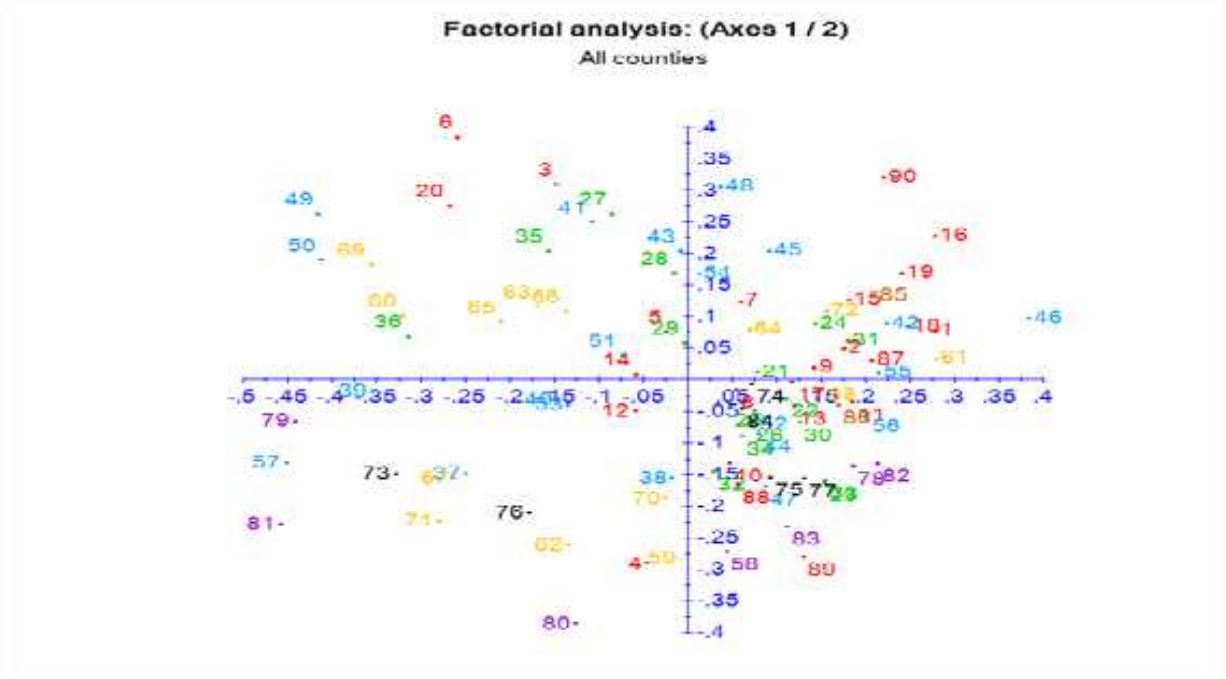

Figure 3: PCoA of axes 1 and 2 based on the dissimilarity of 11 SSR loci across 90 blackberry individuals from different counties. The accessions are depicted using the following colour codes: Red = Nakuru; Green= Kericho; Blue= Baringo; Yellow = Nandi; Black = Laikipia; Purple = Uasin Gishu and Orange = introductions.

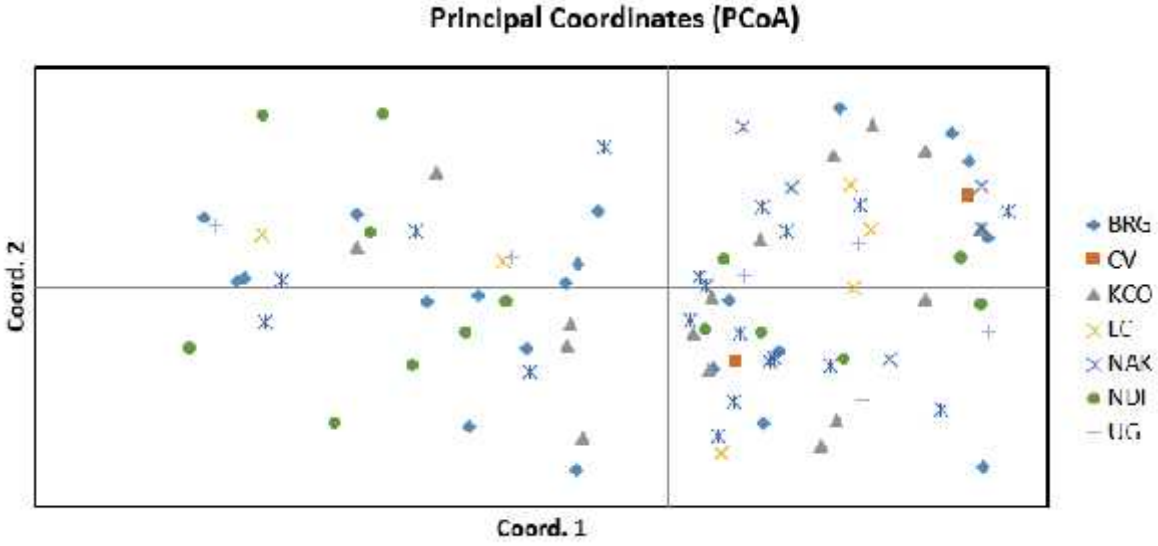

Figure 4: Distribution blackberry accessions by region of collection under principal component axes 1 and 2. BRG: Baringo; CV: Plant Introductions; KCO: Kericho; LC: Laikipia; NAK: Nakuru; NDI: Nandi; UG: Uasin Gishu

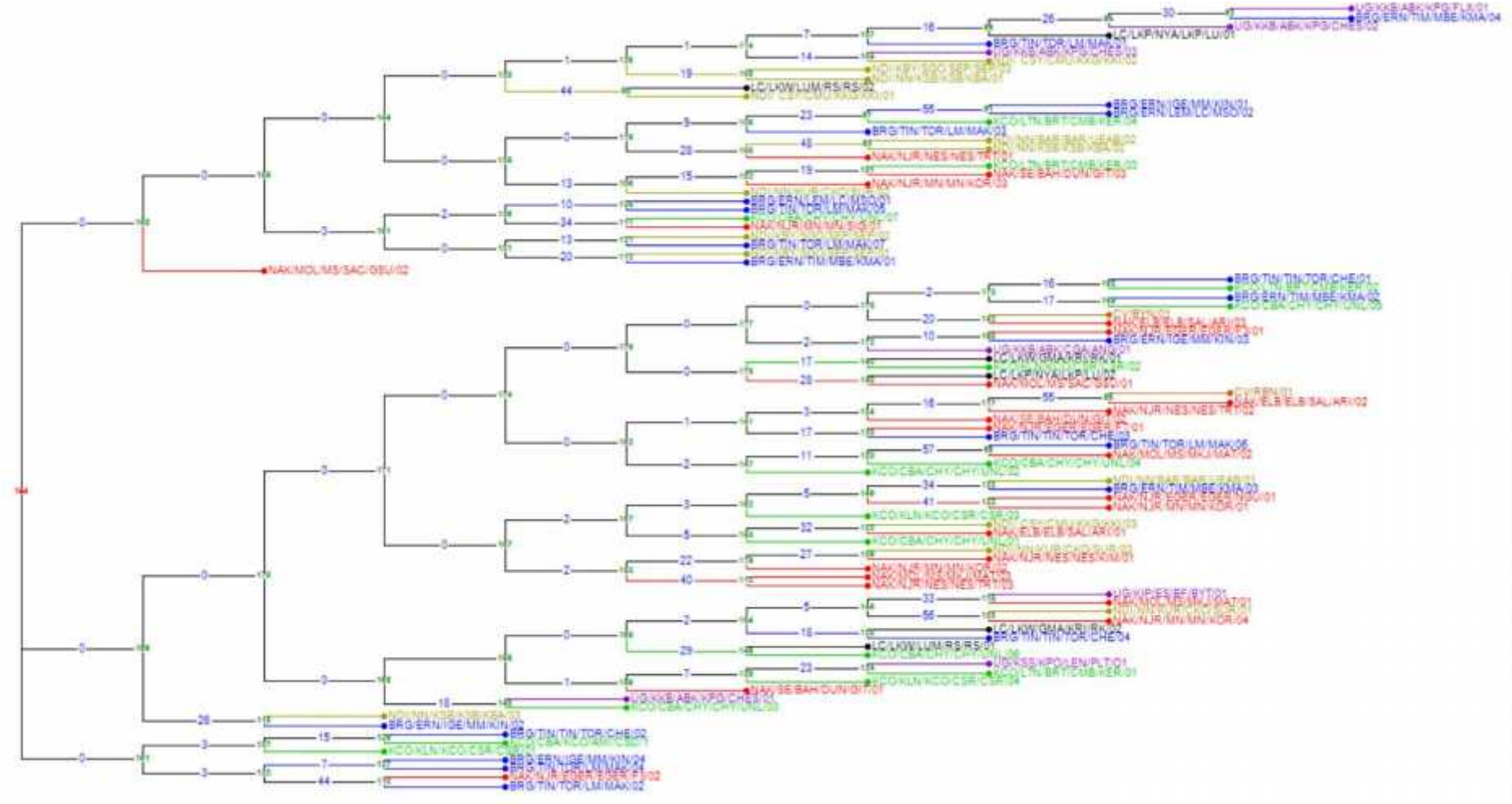

Figure 5: Dendrogram generated by Jaccard's similarity coefficients among 90 blackberry accessions. The accessions are depicted using the following colour codes: Red = Nakuru; Green= Kericho; Blue= Baringo; Yellow = Nandi; Black = Laikipia; Purple = Uasin Gishu and Orange = Plant Introductions (PIs).

3.6 Discussion

The observed heterozygosity (*HO*) and expected heterozygosity (*HE*) were estimated to show the level of polymorphism and usefulness of the SSR markers used in this study (Table 4). The *HE* ranged from 0.34 (*RhM021*) to 0.50 (*RhM043*) which had highest heterozygosity index whilst observed heterozygosity (*HO*) ranged from 0.19 (*RhM019*) to 0.88 (*RiM036* and *RhM003*). Studies independently conducted by Marulanda *et al*. (2007) and Castillo *et al*. (2010) had *HE* values vary from 0.00 to 0.33 and 0.21 to 0.98, respectively. The *HE* range of 0.41 to 0.91 in the accessions revealed genetic diversity in almost all the populations of the blackberry studied although low *HO* were also observed (*HO*=0.27) in some accessions. Often, high *HE* values are observed when wild populations are grown in close proximity to cultivated populations, and this may explain the high *HE* values obtained in the cultivated types. The *HE* values obtained in this study ranged between 0.41 to 0.91 and according to Nybom (2004), these values are within the range of long lived perennials (*HE*=0.68) although some may be endemic to their areas of collection, hence, limited distribution (*HE*=0.42) and others, dispersed by gravity (*HE*=0.47). The low *HO* values obtained in some blackberry accessions could be as a result of imbalanced population sampling among the different regions of germplasm collection both in wild and cultivated blackberries.

Although there is no standard multi-locus estimate of genetic distance (F_{ST}) universally accepted to distinguish distinct plant species, an F_{ST} value of 0.05–0.15 is considered to infer moderate to significant genetic differentiation, while any value more than 0.30 indicate highly differentiated populations (Hartl and Clark, 1997; Frankham *et al*., 2002, 2010). F_{ST} values range from 0 to 1. Observations close to zero indicate a larger number of heterozygotes while observations close to 1 indicate a larger number of homozygotes. The average F_{ST} value of 0.057 obtained in this study shows the presence of heterozygotes in the blackberry accessions. The pairwise genetic distance between populations (Nei 1978) ranged from 0.003 (*RhM003*) to 0.171 (*RiM017*) based on the SSR markers (Table 4). The pairwise genetic variation (F_{ST}) generated from this study indicate moderate to significant differentiation ($0.05>F_{ST}\geq0.15$) within the blackberry accessions or, in this case, between and within wild and cultivated blackberry types. The multi-locus estimate of genetic distance (F_{ST}) based on SSR loci also revealed that there were genetically distinct accessions with *RiM017* (F_{ST} = 0.17) and *RhM011*

(F_{ST} = 0.14) being the best markers for identification of admixtures. The hybridity in these accessions can be maintained if the accessions are propagated using clones.

The inbreeding coefficient, determined by the Wright's (1978) fixation index (F_{IS}), which is a measure of heterozygote deficiency or excess ranged from -0.863 (*RhM003*) to 0.711 (*RiM019*). Inbreeding levels (F_{IS}) are considered high in plants at $F_{IS} \geq 0.5$ and moderate at $F_{IS} \geq 0.25$, and this is where inbreeding depression can substantially impact any population (Ritland, 1996). Only one marker (*RiM019*) showed some evidence of excessive inbreeding ($F_{IS} \geq 0.5$). Most of the accessions showed moderate to high inbreeding levels (Table 5). This may be explained by the reproductive and invasive nature of the blackberry genotypes. Most invasive plants are clonally propagated and are usually self-compatible which could also lead to increased inbreeding levels and decreased variations (Amsellem *et al*., 2000; Liu *et al*., 2006). Inbreeding levels in invasive species can sometimes be synonymous with clonal propagation, where, a species grows vigorously enabling faster spread. In such cases, the molecular variations obtained in the clonal invasives can be due to characteristics other than genetic diversity. This could also infer availability of polyploids among the accessions and subsequent dispersal mechanism across the counties of germplasm collection. Some of the regions of collection were geographically adjacent and could be considered as one large single population. Apart from the reproductive nature (clonal), blackberry genetic variability is also determined by the effect of cross-pollination between polyploid species which in turn, influences the seed and fruit quality, whilst increasing the ploidy levels and taxonomic proximity (Kollmann, 2000). Some outcrossing accessions were observed (those with negative F_{IS}) (Table 6). These accessions also had the highest H_E and H_O indices (genetic diversity) and could be selected as parents in a breeding program as they have the greatest genetic diversities.

The analysis of molecular variance (AMOVA) revealed significant differences ($P \leq 0.05$) in partitioning genetic variances within and among the accessions. SSR markers showed greater divergence within than among the accessions (Table 8). The genetic variance within the blackberry accessions was 95% with an estimated variation of 4.12. Summary of Shannon diversity statistics also showed greater variability within than among population genetic diversity, accounting for 90.57% and 9.43% respectively (Table 7). This illustrates that much of the genetic diversity in blackberry accessions found in Kenya resided within the blackberries. A research conducted by Oyoo *et al*. (2015) for coastland coconut (*Cocos nucifera* L.) populations had a total variance 98% within the genotypes. Additionally, in a study to evaluate the genetic diversity of wild and cultivated *Rubus* species in Colombia using AFLP and SSR markers, Marulanda *et al*. (2007) detected a considerable within population SSR

variation of 80.4%. The analysis of molecular variance (AMOVA) showed less estimated variation among accessions in different regions (0.19) accounting for only 5% of the total variation. The lower genetic diversity among the accessions may be attributed to a limited number and frequency of plant introductions, method of reproduction, in this case, clonal, frequent self-fertilization and method of dispersal that can result in redundancies especially in geographical locations of close proximity. Blackberry is often an invasive plant in nature and with multiple introductions, invasive plants tend to exhibit high levels of genetic diversity (Roman and Darling, 2007) and thus, among accessions estimated variance may be due to fewer introductions into their native habitat. This was demonstrated in common ragweed (*Ambrosia artemisiifolia*) in which, as a result of multiple sources of introductions, had a high genetic diversity and where a pattern of isolation between native and introduced ragweeds existed, lower among-population differentiation was observed (Genton *et al*., 2005).

A UPGMA dendrogram generated by Jaccard's similarity coefficient grouped the accessions into three clusters; I, II and III consisting of 31, 53 and 7 accessions, respectively (Figure 5). All the three clusters had sub-clusters, indicating high levels of intra-accessions heterogeneity. Group II consisted mainly of the accessions from Nakuru. The cultivated blackberry cultivars were also clustered in this group. There was no grouping in all accessions on the basis of area of collection. This can be explained by the diverse folk nomenclature in the collection areas which in turn influences redundancies in germplasm distribution, method of dispersal of the germplasm, and outcrossing nature of some blackberry species. Geographically adjacent areas may have had the same types of blackberry accessions, with discriminant differences used during germplasm for molecular characterization sampling being due to environmental effects. This is shown by the *Nei's* genetic identity and genetic distance among blackberry accessions (Table 6) where, the significant genetic distances were observed between the accessions from RSA and the wild types from different counties in Kenya. Additionally, the invasive nature of the blackberries could have been a major driver in the results aforementioned.

Pattern recognition using Principal Coordinate Analysis (PCoA) failed to group accessions according to their areas of origin suggesting high levels of uniformity across the geographical locations of germplasm collections (Figures 2, 3 and 4). This is however not synonymous with higher homozygosity or narrow genetic bases for the blackberry species found in Kenya. This is because PCoA conducted solely on accessions from each region where the accessions were collected revealed considerably genetic diversity within the accessions

(Figure 2). Blackberries have a varied genetic base that includes numerous species and there could be a selective advantage of heterozygotes as shown by the results obtained in this study.

3.7 Conclusion

There exists considerable genetic diversity in each county on the blackberry accessions studied. However, between one county and the other, low indices of diversity were observed. Findings from this research revealed that even with hybridizations and inbreeding depression, there is still a wide array of genes to be explored in breeding blackberry in Kenya. The best markers (based on PIC) for genotyping blackberry from this study were *RiM017*, *RiM019, RhM043*, *RiM015, RhM018* and *RhM001*. Genetic diversity also exists in the blackberry accessions with eighteen outcrossing types identified (Those with negative F_{IS}). This study is a prerequisite to blackberry breeding and offers insights into the genetic diversity if blackberry in Kenya.

References

Alice, L. A. (2002). Evolutionary relationships in *Rubus* (Rosaceae) based on molecular data. In *Acta Horticulturae* (Vol. 585, pp. 79–83).

Alice, L. A., Eriksson, T., Eriksen, B., and Campbell, C. S. (1997). Intersubgeneric hybridization between a diploid raspberry, *Rubus idaeus*, and a tetraploid blackberry, *R. caesius* (Rosaceae). *American Journal of Botany*, *84*, 171.

Ames, B. N., Shigenaga, M. K., and Hagen, T. M. (1993). Oxidants, antioxidants, and the degenerative diseases of ageing. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(17), 7915–7922.

Amsellem, L., Noyer, J. L., Bourgeois, T. L. E., and Hossaert-Mckey, M. (2000). Comparison of genetic diversity of the invasive weed *Rubus alceifolius* Poir. (Rosaceae) in its native range and in areas of introduction, using amplified fragment length polymorphism (AFLP) markers. *Molecular Ecology*, *9*(4), 443–455.

Bassil, N. V. V, Nyberg, A. M. M., Finn, C. E. E., Clark, J. R. R., Peace, C. P. P., Iezzoni, A., and Pearl, B. (2016). Development of a multiplexed fingerprinting set in blackberry. *Acta Horticulturae*, (1133), 89–96.

Bostein. D., White, R.L., Skolnick, M. and Davis, R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Genetics*, 32, 314-331.

Bowen-Forbes, C. S., Zhang, Y., and Nair, M. G. (2010). Anthocyanin content, antioxidant, anti-inflammatory and anticancer properties of blackberry and raspberry fruits. *Journal of Food Composition and Analysis*, *23*(6), 554–560.

Brennan, R. M., Caligari, P. D. S., Clark, J. R., De Oliveira, P. N. B., Finn, C. E., Hancock, J. F., Jarret, D., Lobos, G. A., Raffle, S., and Simpson, D. (2014). Berry crops. In *Horticulture: Plants for People and Places, Volume 1: Production Horticulture* (pp. 301–325).

Castillo, N. R. F., Reed, B. M., Graham, J., Fernandez-Fernandez, F., and Bassil, N. V. (2010). Microsatellite Markers for Raspberry and Blackberry. *Journal of the American Society for Horticultural Science*, *135*(3), 271–278

Castro, P., Stafne, E. T., Clark, J. R., and Lewers, K. S. (2013). Genetic map of the primocane-fruiting and thornless traits of tetraploid blackberry. *Theoretical and Applied Genetics*, *126*(10), 2521–2532.

Clark, J.R., Stafne, E.T., Hall, H.K. and Finn, C.E. (2007). Blackberry breeding and genetics. *Plant Breeding Reviews*, 29, 82-101.

Chakraborty, R., and Jin, L. (1993). Determination of relatedness between individuals using DNA fingerprinting. *Human Biology*, *65*(6), 875–95.

Dossett, M., Bassil, N. V., Lewers, K. S., and Finn, C. E. (2011). Genetic diversity in wild and cultivated black raspberry (*Rubus occidentalis* L.) evaluated by simple sequence repeat markers. *Genetic Resources and Crop Evolution*, *59*(8), 1849–1865.

Doyle, J.J., and Doyle. J.L. (1990). Isolation of plant DNA from fresh tissue. *Focus*, 12:13-15.

Frankham, R., Ballou, J.D. and Briscoe, D.A. (2002). Introduction to Conservation Genetics. Cambridge University Press, Cambridge, UK.

Frankham, R., Ballou, J.D. and Briscoe, D.A., (2010). Introduction to Conservation Genetics, 2nd edition. Cambridge University Press, Cambridge, U.K.

Graham, J., and McNicol, R. J. (1995). An examination of the ability of RAPD markers to determine the relationships within and between Rubus species. *Theoretical and Applied Genetics*, *90*(7–8), 1128–1132.

Graham, J., Squire, G. R., Marshall, B., and Harrison, R. E. (1997). Spatially dependent genetic diversity within and between colonies of wild raspberry Rubus idaeus detected using RAPD markers. *Molecular Ecology*, *6*, 1001–1008.

Genton, B. J., Shykoff, J. A., and Giraud, T. (2005). High genetic diversity in French invasive populations of common ragweed, Ambrosia artemisiifolia, as a result of multiple sources of introduction. *Molecular Ecology*, *14*(14), 4275–4285.

Hartl, D.L. and Clark, G.C. (1997) Principles of Population Genetics. Sinauer Associates, Sunderland.

Jacob, P., Avni, A. and Bendahmane, A. (2018). Translational Research: Exploring and Creating Genetic Diversity. *Trends in Plant Science*, *23*(1), 42-52.

Kimura, M., and Crow, J. F. (1964). The number of alleles that can be maintained in a finite population. *Genetics*, *49*, 725–738.

Kollmann, J., Steinger, T., and Roy, B. A. (2000). Evidence of sexuality in European Rubus (Rosaceae) species based on AFLP and allozyme analysis. *American Journal of Botany*, *87*(11), 1592–1598.

Lewers, K. S., Saski, C. A., Cuthbertson, B. J., Henry, D. C., Staton, M. E., Main, D. S., and Tomkins, J. P. (2008). A blackberry (*Rubus* L.) expressed sequence tag library for the development of simple sequence repeat markers. *BMC Plant Biology*, *8*.

Lewontin, R. C. (1972). The Apportionment of Human Diversity. In *Evolutionary Biology* (pp. 381–398).

Liu, K., and Muse, S. V. (2005). PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics*, *21*(9), 2128–2129.

Liu, J., Dong, M., Miao, S. L., Li, Z. Y., Song, M. H., and Wang, R. Q. (2006). Invasive alien plants in China: Role of clonality and geographical origin. *Biological Invasions*, *8*(7), 1461–1470.

Marulanda, M. L., López, A. M., and Aguilar, S. B. (2007). Genetic diversity of wild and cultivated Rubus s pecies in Colombia using AFLP and SSR markers. *Crop Breeding and Applied Biotechnology*, *7*, 242-252.

Miyashita, T., Kunitake, H., Yotsukura, N. and Hoshimo, Y. (2015). Assesment of genetic relationships among cultivated and wild Rubus accessions using AFLP markers. *Scientia Horticulturae,* 193, 165-173.

Moyer, R. A., Hummer, K. E., Finn, C. E., Frei, B., and Wrolstad, R. E. (2002). Anthocyanins, phenolics, and antioxidant capacity in diverse small fruits: Vaccinium, Rubus, and Ribes. *Journal of Agricultural and Food Chemistry*, *50*(3), 519–525.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America*, *70*(12), 3321–3323.

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89(3):583-90.

Nei, M. and Li, W.H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases (molecular evolution/mitochondrial DNA/nucleotide diversity), *Proceedings of the National Academy of Sciences*, USA, 76(10), 5269-5273.

Nybom, H. (2004). Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology*, *13*(5), 1143–1155.

Oyoo, M.E., Muhammed, N., Githiri, S.M., Ojwang, P.O., Muniu, F.K., Masha, E. and Owuoche, J. (2015). *In-situ* morphological characterization of coconut in the coastal lowlands of Kenya. *African Journal of crop science*, 9(2), 65-74.

Oyoo, M.E., Najya, M., Kimani, N. C., and Simon, M. G. (2016). Assessment of the genetic diversity of Kenyan coconut germplasm using simple sequence repeat (SSR) markers. *African Journal of Biotechnology*, *15*(40), 2215–2223.

Overall, J., Bonney, S. A., Wilson, M., Beermann, A., Grace, M. H., Esposito, D., Lila, M.A. and Komarnytsky, S. (2017). Metabolic effects of berries with structurally diverse anthocyanins. *International Journal of Molecular Sciences*, *18*(2).

Parent, J. G., Fortin, M. G., and Page, D. (1993). Identification of Raspberry Cultivars by Random Amplified Polymorphic DNA (RAPD) Analysis. *Canadian Journal of Plant Science*, *73*(4), 1115–1122.

Parent, J. G., and Pagé, D. (1998). Identification of raspberry cultivars by sequence characterized amplified region DNA analysis. *HortScience*, *33*(1), 140–142.

Peakall, R., and Smouse, P. E. (2012). GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, *28*(19), 2537–2539.

Perrier, X., Flori, A. and Bonnot, F. (2003). "Methods of data analysis." In *Genetic diversity of cultivated tropical plants*, 47-80.

Perrier X, Jacquemoud-Collet JP (2006). DARwin software http://darwin.cirad.fr/

Roman, J., and Darling, J. A. (2007). Paradox lost: genetic diversity and the success of aquatic invasions. *Trends in Ecology and Evolution*. https://doi.org/10.1016/j.tree.2007.07.002

Ritland, K. (1996). Estimators for pairwise relatedness and individual inbreeding coefficients. *Genetics Research*, *67*(2), 175-185.

Ru, S., Main, D., Evans, K., and Peace, C. (2015). Current applications, challenges, and perspectives of marker-assisted seedling selection in Rosaceae tree fruit breeding. *Tree Genetics and Genomes*, *11* (1), 7-12.

Saitou, N. and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology Evolution*, 4(4),406-425.

Stafne, E. T., and Clark, J. R. (2004). Genetic relatedness among eastern North American blackberry cultivars based on pedigree analysis. *Euphytica*, *139*(2), 95–104.

Tulio J.R., Reese, R.N., Wyzgoski, F.J., Rinaldi, P.L., Fu, R., Scheerens, J.C., and Miller, A.R. (2008). Cyanidin 3-rutinoside and cyanidin 3-xylosyl-rutinoside as primary phenolic antioxidants in black raspberry. *Journal of Agriculture and Food Chemistry, 56,* 1880–1888.

Wright, S. (1978). Evolution and the Genetics of Populations. 4: *The University of Chicago Press, Chicago, Vol.IV,* 157.

Yin, M. (2017). Studies in Blackberry : Development and Implementation of a Phenotyping Protocol for Blackberry Seedling Populations and Impact of Time of Day of Harvest on Red Drupelet Reversion for University of Arkansas blackberry genotypes. *Theses and Dissertations*. 2417. http://scholarworks.uark.edu/etd/2417 retrieved on 10th January 2018.

CHAPTER FOUR

In-situ Morphological Characterization of blackberry (*Rubus* subgenus *Rubus* Watson) in Rift Valley region of Kenya.

4.1 Abstract

The variation of *in-situ* morpho-physiological traits of blackberry (*Rubus* subgenus *Rubus* Watson) is important for consideration in improvement in a breeding programme. Morphological characterization is the initial step in determining the genetic diversity of any crop. The objective of this study was to characterize wild blackberry accessions in Kenya and Plant Introductions (PIs) using morphological descriptors. Each blackberry accession was nested within its county of collection and a phylogenetic tree that was constructed, using the Gower's coefficient, grouped the accessions into class I and II consisting of 1 and 89 accessions, respectively. Clustering of accessions did not show an association between the origin of collection and the accessions. Principal Component Analysis (PCA) revealed 10 axes of which 7 had a cumulative variation of 96.30% with the first two axes having a discriminatory variance of 52.71 %. This suggested that variables identified in this study could be used to differentiate blackberry accessions morphologically. This study demonstrated that number of internodes per average growing shoots, thorniness of the plant and length of internode were associated with the first axis with Eigenvalue of 27.79%. Plant thorniness was also associated with the second axis with Eigenvalue of 24.92%. These results suggest that there exists qualitative and quantitative variation among blackberry accessions in Kenya that can be utilized in breeding blackberry.

Keywords: Morphological diversity, *Rubus* subgenus *Rubus* Watson, plant genetic resources, cluster analysis.

4.2 Introduction

The assessment and monitoring of diversity of plant genetic resources *in-situ* and *ex-situ* is essential for germplasm management and for establishing core breeding stocks (Orobiyi, 2017). Knowledge on morphological variability of germplasm collections improves understanding of the relationship between the structural morphology of plants and their corresponding functional botany (Lauri and Normand, 2017).

Blackberry (*Rubus* subgenus *Rubus* Watson) is a cross-pollinated, fruiting plant species formerly of subgenus *Eubatus*. The fruits aggregate around a receptacle and consist of fleshy drupelets, each with a single seed (Finn, 2008). Blackberry is a perennial plant with biennial canes and is of three types in reference to cane architecture; erect, semi-erect and trailing (Clark *et al*., 2007). There are 84 wild species of blackberry in Kenya (Chittaranjan, 2011) and only two plant introductions; one hybrid berry (definitive genetic origin unknown but is believed to be an interspecific cross between European Raspberry, *Rubus idaeus* and another European blackberry, *Rubus fruticosus*) (Wood *et al*., 1999) and the other, a European berry, *Rubus fruticosus* cultivated mainly for export market. Blackberries have a complex reproductive (sexual, facultatively apomictic to obligately apomictic), ploidy (autoploidy and alloploidy) and inheritance strategies (disomic and tetrasomic) (Clark *et al*., 2007). Thus, there is difficulty in identifying superior berry and also designation to definite groups and are sometimes misclassified. Blackberry fruits have varied health benefits and are rich in natural phytochemicals (Rao and Snyder, 2010), vitamin C and E (Hirsch, 2013), contain phenolic compounds that are secondary plant metabolites integral in human and animal diets (Siriwoharn *et al*., 2004; Lee *et al*., 2011) due to their antioxidant properties (Hirsch, 2013). They are also used to prevent lifestyle diseases like diabetes, cancer, cardiovascular diseases and other pathogens (Bravo, 1998; Hollman *et al*., 1996). Blackberry fruits are consumed fresh or processed as individually quick frozen (IQF), canned, pureed, juiced or freeze-dried (Finn, 2008). The crop is thus gaining prominence in Kenya and Africa at large due to its possible health benefits and the influx of a more informed, aggressive middle-class population.

Modern breeding objectives emphasize on the evaluation of the characteristics of importance to production and productivity within genetic resources and concentration the same in one cultivar (Bozovic *et al*., 2016). Analysis of genetic diversity can be achieved through molecular markers and morphological markers. Some of the molecular markers that have been used in the assessment of genetic diversity of blackberry are random amplified polymorphic DNA (RAPD), amplified fragment length polymorphisms (AFLP), restriction fragment length polymorphisms (RFLP), ISSR-EST, simple sequence repeats (SSRs) (Clark et al., 2007). *In-*

situ hybridization techniques (ISH) such as Genomic *in-situ* hybridization (GISH) and fluorescence *in-situ* hybridization (FISH) have also been used to infer blackberry phylogeny genus (Yan *et al*., 2015). Morphological markers are still useful in phenotypic descriptions of plant populations. Some morphological traits have been associated with influencing some other trait that is of great economic importance but difficult to measure such as disease susceptibility (Karimi *et al*., 2009). Phenotypic descriptors are widely used to classify cultivars, genotypes and landraces based on discriminant variables for the plant genetic resources (PGR) studied (Orobiyi *et al*., 2017). Consequently, a comparative analysis is done of the composition of PGR with those of the classes obtained from principal component analysis and correlation analysis. This can better reveal the constitution of each group with respect to the landraces, cultivars or genotypes studied.

In Kenya, blackberry is still a minority crop and grown mainly for the export market or the suburban population. Smallholder farmers are still few as the crop gains prominence in the region. Statistics for yield in Africa are only available for South Africa (220 tonnes) and still, this is very low in comparison to other regions of the world (Strik *et al*., 2007; Finn, 2008). Being a minority crop, challenges abound; inadequate breeding programs and funding targeting blackberry in Kenya, little understanding of population structures within repositories and the available breeding program, inaccurate identification of species and misclassification in gene banks, difficulty in identification of duplicate accessions in germplasm repositories and unavailability of improved local cultivars. In addition, the available varieties experience pest and disease problems coupled with abiotic stresses that are not well documented. There is need for genotyping of blackberry resources in Kenya. Phenotypic expression such as objective descriptions of tree and fruit characteristics discriminating against undesirable traits in the process is unreliable and may not provide an accurate indication of genetic diversity (Menkir *et al*., 1997). This preference for specific traits based on phenotypic descriptions has also previously led to the discarding of potentially important and advantageous germplasm. In addition, expression of morphological data are greatly influenced by environment, phenological stages of development and can be subjective in nature resulting in errors (Marinoni *et al*., 2003). This assessment of the morphometric diversity of the fruit tree species in core germplasm collections, wild and introduced, whether *in-situ* or *ex-situ* is, however, necessary. It offers prerequisite remedies to the challenges above mentioned and is vital for the thorough understanding of these genetic resources, breeding options and subsequent conservation efforts. The objective of this study was to characterize wild blackberry types in selected counties in Kenya and Plant Introductions (PIs) using morphological markers.

4.3 Materials and Methods

4.3.1 Experimental Site

The study was conducted *in-situ* in selected counties in Kenya. These counties included Kericho (0.3689° S, 35.2863° E), Nakuru (0.3031° S, 36.0800° E), Uasin Gishu (0.5143° N, 35.2697° E), Nandi (0.1036° N, 35.1777° E), Laikipia (0.3970° N, 37.1588° E), and Baringo (0.4897° N, 35.7412° E). In each county, five random districts or sub-counties were selected out of which five random locations and villages were chosen for germplasm sampling.

4.3.2 Germplasm Sampling

Sampling of blackberry for morphological trait analysis was carried out in the areas mentioned in section 4.3.1. Fruit trees were coded to reflect the county, district, division, subdivision, village and the collection number (Oyoo *et al*., 2015). If the collection was from Baringo County, Tinet district, Torongo division, Lembus Mosop location, Makutano village and it was the first blackberry sampled, the code given was; BRG/TIN/TOR/LM/MAK/01. Sampling was done to reflect different agro-ecological zones in the counties where blackberry is reportedly growing. The selected agro-ecological zones were different and are designated pyrethrum (*Chrysanthemum cinerelifolium*)-wheat (*Triticum aesitivum*) zone (UH2), tea-dairy zone (LH1), wheat-maize-pyrethrum zone (LH2) wheat-barley zone (LH3), cattle-sheep-barley (*Hordeum vulgare*) zone (LH4), coffee (*Coffea arabica*) zone (UM2) and sunflower (*Helianthus annus*)-maize zone (UM4). The altitude varied from 1650 m to 2743 m above sea level. Samples were taken in areas where the fruit trees are morphologically different and there are marked changes in altitude, cropping systems, a formidable barrier such as a mountain, river, valley or local people are ethnically different (dialect) from previous collection sites. Here, quantitative and qualitative attributes of the plants were taken along edaphic, topographic and climatic gradients. Data stations in a location were within 200 m intervals. This was done to minimise redundancies. For each fruit tree sampled, Global Positioning System (GPS) data were taken and the plant photographed. This was vital for mapping of these areas. Plants with similar features growing in ecologically distinct locations were assumed to be of different eco-strain and hence, were sampled and characterized. Blackberry accessions were evaluated for population structure, architecture and fruit tree characteristics

4.4.3 Evaluation of Traits

Seven qualitative and three quantitative important traits to blackberry breeding were characterised in this study (Yin, 2017). These included tree, stem, leaf, reproductive

characteristics and stress severity assessment. The descriptors of blackberry used are provided below.

Pre-harvest

Vegetative observations

The cane architecture, showing the degree of creeping for each plant (denoted for each individual as erect “E,” semi-erect “S,” or trailing “T”); stem type, whether malformed, symmetrical or asymmetrical; thorniness, indicating whether the plant is thorny or not (denoted for each individual as thorny “T” or thornless “N”); overall plant health, showing the degree of infestation, apparent nutrient deficiencies and general abiotic stress susceptibility (subjectively assessed from 1 to 10, where 10 = excellent health); overall plant vigor, examining leafiness, length of current season’s growth, and relative number of actively growing shoots (subjectively assessed from 1 to 10, where 10 = extremely vigorous).

Reproductive observations

Number of internodes/actively growing shoot; internode length (mm), the average length of the fourth internode of four plants; pubescent colour (white, varied or purple); flower colour (white, varied or purple).

Table 4: Measured variables and observation criteria used to characterize blackberry genotypes

Code	Qualitative traits	Phenotypic scale
PT	Plant Thorniness	1 = thornless, 2 = thorns very small and sparse, only detectable upon touch, 3 = small, sparse thorns visible, 4 = small to medium thorns, sporadically spaced, 5 = medium thorns, spaced evenly, 6 = medium thorns closely spaced, 7 = medium to long thorns sporadic, 8 = medium to long thorns evenly spaced, 9 = long thorns, closely spaced
OPH	Overall Plant Health	1 = dead, 2 = extreme biotic/abiotic stress (B/AS), with < 50% leaves green, 3 = obvious symptoms of B/AS with 50% of leaves green, 4 = obvious symptoms of B/AS with > 50% of leaves green, 5 = apparent majority of leaves green but with obvious symptoms of B/AS, 6 = leaves mostly green with minor symptoms of B/AS, 7 = sporadically spaced symptoms of B/AS, 8 = B/AS only observable upon close inspection, 9 = no symptoms of B/AS
OPV	Overall Plant Vigour	1 = no PC or new leaf growth, 2 = leaf growth with no primocane (PC) development, 3 = 1-2 PC or PC < 15 cm in length, 4 = PC > 15 cm, shorter than floricane (FC), 5 = PC growth approximately same length as FC, 6 = 3-5 PC, length similar to FC, 7 = 3-5 PC slightly longer than FC, 8 = PC significantly longer than FC, vigorous growth, 9 = 5 or more PC significantly longer than FC with overly vigorous growth
SS	Stem Symmetry	1 = extremely malformed and asymmetrical, 2 = very asymmetrical, 3 = significantly asymmetrical, 4 = asymmetrical, 5 = somewhat asymmetrical, 6 = slightly asymmetrical, 7 = very slightly asymmetrical, 8 = symmetrical, 9 = completely symmetrical
PFC	Plant Flower colour	1 = white, 2 = purple, 3 = pink
PPC	Plant Pubescent colour	1 = white Light Green, 2 = brown, 3 = green

4.5 Data Analyses

4.5.1 Multivariate Analysis

Multivariate analysis was carried out by using GENSTAT 15th Edition programme on morphological data to identify discriminant variables amongst the 90 accessions. Means of quantitative traits were first obtained using PROC GLM in Statistical Analysis System (SAS)

version 9.1 (SAS Institute Inc., Cary, 2001) to determine the significant differences among the accessions. The following statistical model was used;

$$Y_{ijk} = \mu + \alpha_i + \beta_{(i)j} + \varepsilon_{ijk} \quad 4.0$$

Where: μ: overall mean, α_i: Effect of the i^{th} county, $\beta_{(i)j}$: a random effect due to the j^{th} blackberry accession nested within the i^{th} county, ε_{ijk}: random error component associated with each observation.

Factorial Component Analysis (FCA) based on discriminant variables obtained, was performed using DARwin 6.0 (Perrier *et al*., 2003; Perrier and Jacquemoud-Collet, 2006) to illustrate the different grouping of the germplasm available. R program for statistical computing version 3.4.1 (R Development Core Team, 2015) was used for the construction of a hierarchical dendrogram to show the overall similarity between the morphological data by plotting the results in homogenous groups. Since the data under consideration are of mixed types (qualitative and quantitative characters), the UPGMA dendrogram was constructed using Gower's dissimilarity coefficient (Gower, 1971) as shown below;

$$S_{Gower} = \frac{\sum_{i=1}^{m} Si}{\sum_{i=1}^{m} W} \quad 4.1$$

Where; $S_i = 1$ if $x_i = y_i$ (binary or qualitative data); $S_i = 0$ if $x_i \neq y_i$ (binary or qualitative data); $S_i = 1 - |x_i - y_i| / R_i$ (quantitative data): where; $W_i = 1$ if x_i can be compared to y_i and $W_i = 0$ if x_i cannot be compared to y_i. Therefore;

$$D_{Gower}(x,y) = 1 - S_{Gower}(x,y) \quad 4.2$$

The silhouette width was determined using R for statistical computing ver. 3.4.1 as follows;

$$S_i = \frac{b - a}{\max\{a, b\}} \quad 4.3$$

Where; S_i = the silhouette width, a_i = the average dissimilarity ($d_{i,k}$) between the i^{th} accession and all others in its cluster; b_i = the average dissimilarity ($d_{i,k}$) between the i^{th} accession and its neighbour cluster. Identification of the discriminant variates by the PCA (Principal Component Analysis) and correlation was done using GENSTAT 15th Edition programme.

4.6 Results

4.6.1 Qualitative Phenotypic Variability

The accessions exhibited a wide range of differences in qualitative morphological features in plant architecture, which refers to the growth habit of canes. Blackberry genotypes are often classified into erect, semi-erect, and trailing according to their cane architecture. The

semi-erect type was the most dominant with 97% of the fruit trees sampled, with trailing and erect types being a distant 1% and 2%, respectively. For the fruit trees sampled, 36% had small to medium thorns, sporadically spaced; 22% had medium to long thorns, evenly spaced; 13% had small, sparse visible thorns; 12% had medium to long thorns, sporadically spaced; 9% had medium thorns, evenly spaced; 6% had medium thorns, closely spaced while only 2% had long thorns that are closely spaced. White was the most frequent flower colour representing 62% of all sampled fruit trees, followed by purple (33%). Pink flower colour accounted only for 4%. The plant (leaf) pubescence colours were categorized intro 3; white was the most dominant (54%), with brown and green types being 37% and 9% respectively.

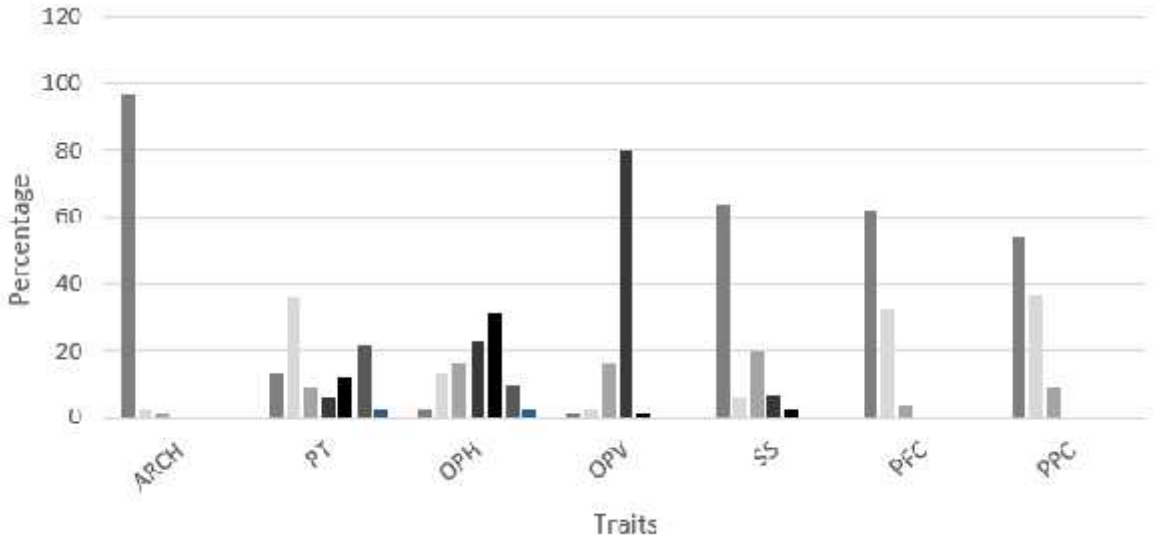

Figure 6: Distribution of some qualitative morphological features among the blackberry accessions collected from different regions in Kenya. ARCH: architecture, PT: plant thorniness, OPV: Overall plant vigour SS: stem symmetry, PFC: plant flower colour PPC: plant pubescence colour.

Due to the continuing interests in blackberry worldwide, breeding for adaptation to different environments in the country is important. In this respect, overall plant health (showing the degrees of infestation, apparent nutrient deficiencies and general abiotic stress susceptibility) and overall plant vigour (examining leafiness, length of current season's growth, and a relative number of actively growing shoots) were considered in this study. For all the germplasm sampled, 31% had their leaves mostly green with minor symptoms of biotic and abiotic stress (B/AS), 23% had apparent majority of leaves green but with obvious symptoms of B/AS, 16% had obvious symptoms of B/AS with more than 50% of leaves green, 13% had obvious symptoms of B/AS with 50% of leaves green, 10% had B/AS only observable upon close inspection; 3% had extreme B/AS with less than 50% leaves green and only 2% had no

symptoms of B/AS (Figure 6). In addition, 80% of the plant genetic resources studied had vigorous growth with long primocanes. In terms of stem type, 64% were asymmetrical; 20% slightly asymmetrical; 7% very slightly asymmetrical; 6% somewhat asymmetrical while 3% were symmetrical. No malformed stems were observed.

4.6.2 Correlation Among Traits

The correlations among quantitative traits are presented below (Table 10). There was a significant positive correlation (*r*=0.81**) between Internode length and number of internodes per average growing shoots.

Table 5: Pearson correlation coefficients showing a pair-wise association of quantitative traits among 90 blackberry accessions.

	Internode Length (cm)	Number of Internodes\Average Growing Shoots	No. of Leaflets
Internode Length (cm)	1	0.80948**	-0.10579
No. Internodes\Average Growing Shoots			-0.01634
No. of Leaflets			1

** = Significant correlation at P< 0.01 probability level

4.6.3 Principal Component Analysis (PCA)

Principal component analysis allowed the association of axes to the variables and out of the 10 axes generated, 7 had a cumulative variation of 96.30% with the first two axes having a discriminatory variance of 52.71% (Table 11). Three variables are associated with the first axis with Eigenvalue of 27.79%. These are number of internodes per average growing shoots, plant thorniness and internode length. Plant thorniness was also associated with the second axis with Eigenvalue of 24.92%). Plant health was associated with the third (Eigenvalue = 14.29%), fourth (Eigenvalue = 11.77%), fourth (Eigenvalue = 8.38%) and sixth (Eigenvalue = 6.78%) axes. These are overall plant health, number of leaflets, symmetry of stem and the colour of plant during pubescence. The number of internodes per average growing shoots and length of internode were associated with the fourth axis with Eigenvalue of 11.77%. The principal components and variables factor biplots(s) for the first two principal components of the morphological data were also generated (Figures 7 and 8).

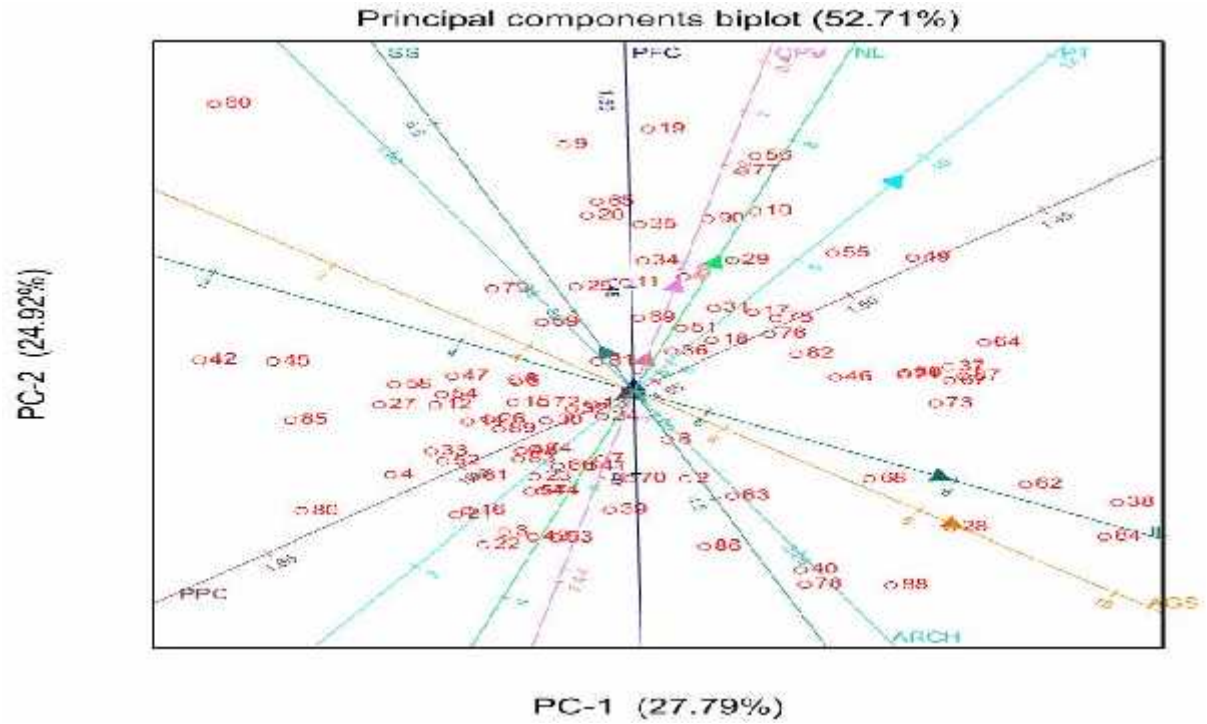

Figure 7: Principal components biplot for the first two principal components of a morphological principal components analysis of blackberry accessions (SS: Stem Symmetry, PFC: Plant Flower Colour, PPC: Plant Pubescence Colour, ARCH: Architecture, OPV: Overall Plant Vigour, OPH: Overall Plant Health, AGS: Number of internodes per average growing shoot).

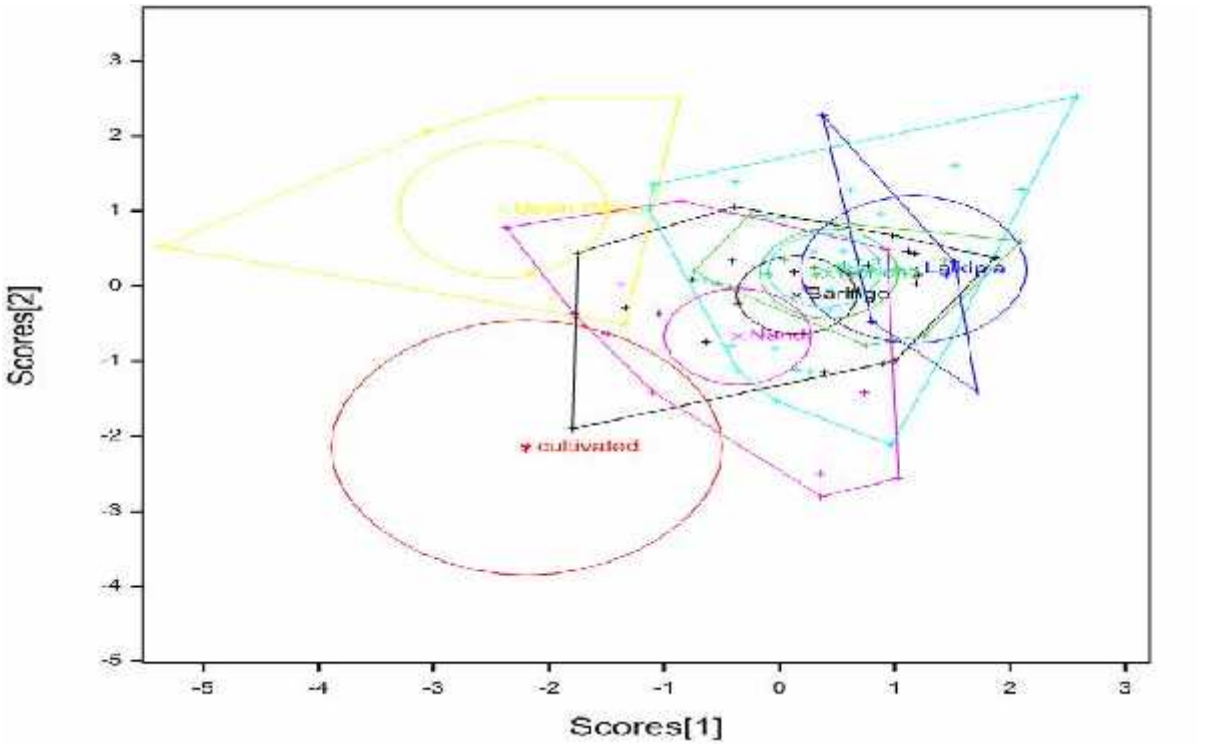

Figure 8: Variables factor biplot for the first two principal scores of a discriminant analysis for the Blackberry Introductions (which is cultivated) and other accessions from 6 different regions in Kenya.

Table 6: Principal Component loadings of 10 traits for 90 blackberry accessions collected from different regions in Kenya.

Variables	Principal component loadings						
	Axis 1	Axis 2	Axis 3	Axis 4	Axis 5	Axis 6	Axis 7
Architecture	0.009	-0.015	0.012	0.039	0.000	0.032	0.006
No. of internodes per shoot	**0.609**	-0.425	0.015	**0.558**	0.339	0.129	0.012
No. of leaflets	0.156	0.407	-0.319	0.419	**0.595**	0.412	-0.079
Overall Plant Health	0.079	0.331	**0.860**	0.272	-0.129	-0.173	-0.059
Overall Plant Vigour	0.023	0.097	0.083	0.004	-0.024	-0.071	-0.242
Plant Flower Colour	0.000	0.014	0.057	-0.020	0.125	0.113	-0.108
Plant Pubescent colour	-0.015	-0.011	0.055	0.078	-0.147	-0.051	**0.947**
Plant Thorniness	**0.500**	**0.667**	-0.216	-0.219	0.389	-0.210	0.107
Stem Symmetry	-0.053	0.118	0.270	-0.258	0.302	**0.846**	0.104
Internode length	**0.586**	-0.274	0.160	**-0.563**	-0.484	0.037	-0.032
Eigen value	4.843	4.344	2.491	2.052	1.460	1.181	0.414
Variation (%)	**27.790**	**24.920**	**14.290**	**11.770**	**8.380**	**6.780**	**2.370**
Cumulative variation (%)	27.790	52.710	67.000	78.770	87.150	93.930	96.300

Factorial component analysis carried out using dissimilarity coefficients obtained from the usual Euclidean distance was conservative and split the accessions into 4 classes (Figure 9 and 10). Most of the accessions overlapped, demonstrating redundancies in the morphology of the characterized germplasm. From the PCoA plot generated (Figure 9 and 10), principal axes 1 and 2 showed that NKR/NJR/EGER/EGER/F7/01 (90), NKR/NJR/NES/NES/KIM/01 (9) from Nakuru County; BRG/TIN/TOR/LM/MAK/06 (42), BRG/TIN/TOR/LM/MAK/02 (39), BRG/ERN/IGE/MM/KIN/04 (53), BRG/ERN/TIM/MBE/KMA/O2 (56) from Baringo County; NDI/NN/KUR/CKO/SUR/03 (60) from Nandi County, LC/LKN/GMA/GMA/KBI/RK/01 (77) from Laikipia County, UG/KKB/ABK/KPG/CHES/03 (80) and UG/KKB/ABK/KBG/CHES/01 (78) from Uasin Gishu and an introduced germplasm CV/RBN/01 (85) (from South Africa) were the most distinct from the other accessions.

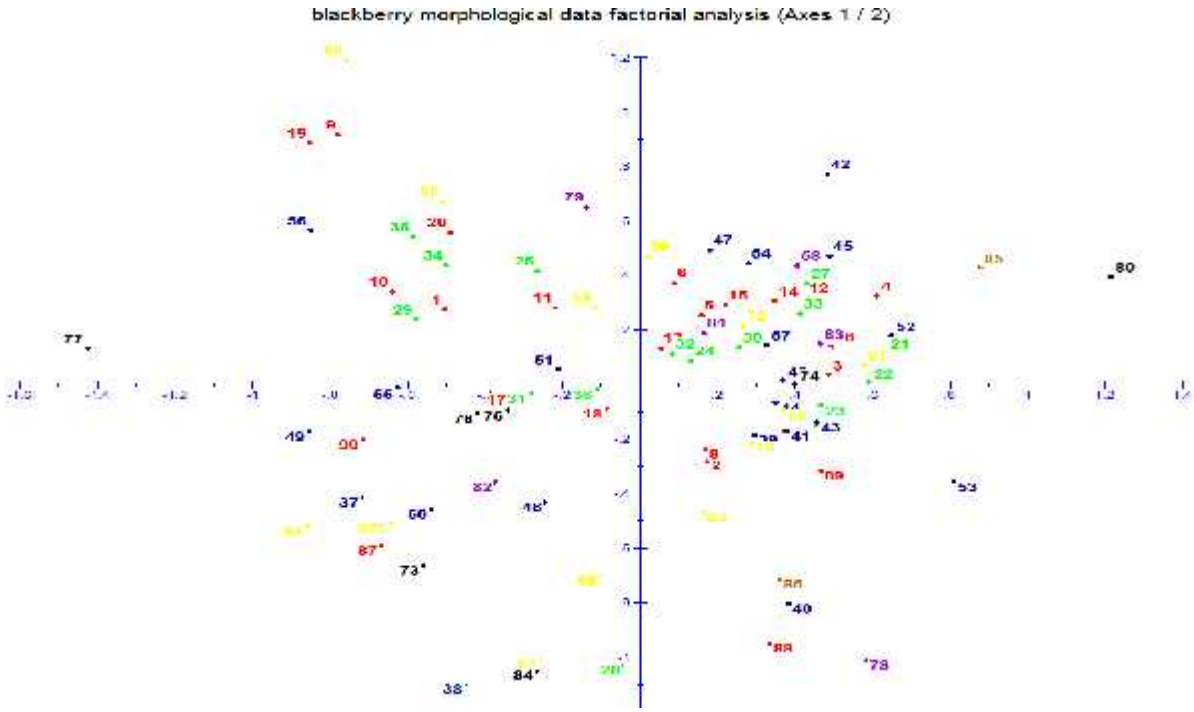

Figure 9: PCoA of axes 1 and 2 based on the dissimilarity of 90 blackberry accessions. The accessions are depicted using the following colour codes: Red = Nakuru; Green= Kericho; Blue= Baringo; Yellow = Nandi; Black = Laikipia; Purple = Uasin Gishu and Orange = introductions.

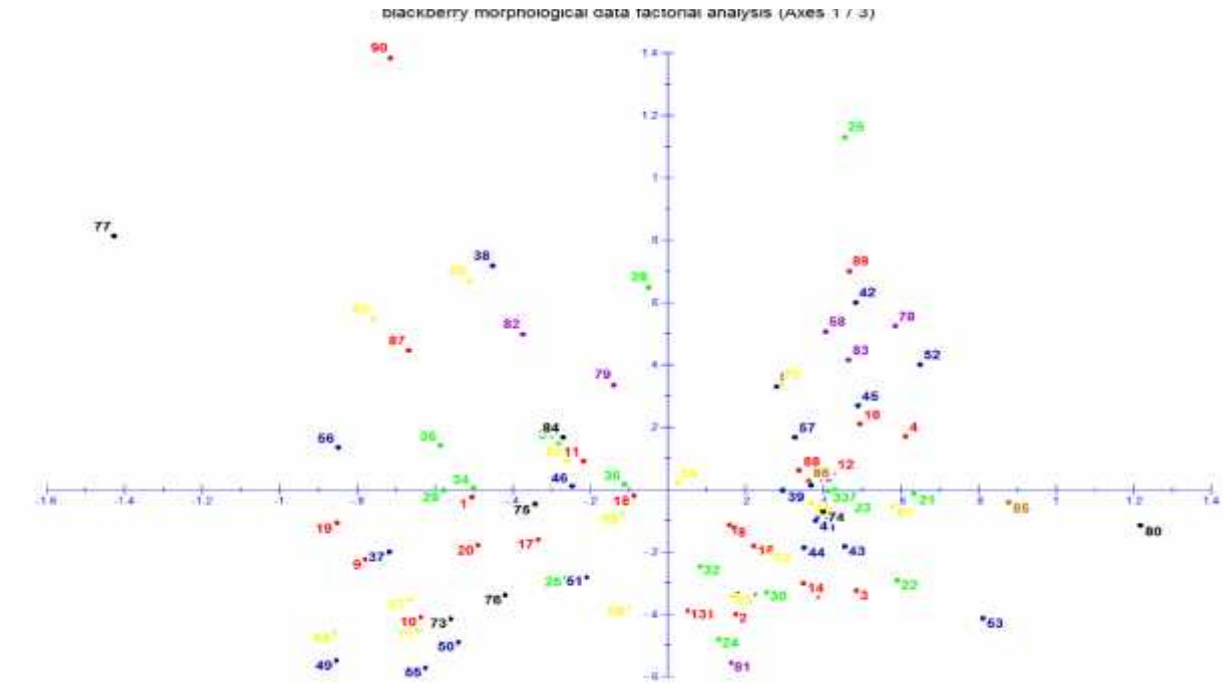

Figure 10: PCoA of axes 1 and 3 based on the dissimilarity of 90 blackberry accessions. The accessions are depicted using the following colour codes: Red = Nakuru; Green= Kericho; Blue= Baringo; Yellow = Nandi; Black = Laikipia; Purple = Uasin Gishu and Orange = introductions.

4.6.4 Cluster Analysis

Cluster analysis refers to the task of grouping individuals such that more similar ones are placed in one group in comparison to those in other groups (or "cluster"). Cluster analysis split the accessions into two clusters, I and II (Figure 11). The Plant Introductions (CV/RBN/01) had its own cluster, I, while the other (CV/BYN/01) clustered with the rest of the wild accessions collected from different regions of Kenya. The accessions were clustered together according to the traits (quantitative and qualitative) taken. The accessions did not cluster according to counties of origin. However, most of the accessions from all the six geographical regions clustered in group II. Cluster II had the highest number of genotypes of 89 based on the morphological descriptors used. Cluster II also had sub-clusters. Grouping of these accessions into these sub-clusters indicated a substantial level of intra-polymorphism within the wild blackberry population in the country. Cluster one only had one genotype indicating inter-polymorphism with the rest of the accessions collected. Structure analysis was illustrated using a silhouette plot.

A silhouette is used for pattern recognition and compares the minimum average dissimilarity of each accession to other clusters with the average dissimilarity to accessions in its own cluster. There are two main groups within this set of germplasm, therefore, objective determination in the number of stable clusters. Observations close to 1 (large s_i) indicate that the individual (s) is very well clustered. Clusters with observations close to 0 (small s_i) indicate that the germplasm lie between two clusters. Based on the silhouette plot generated, the natural number of clusters in this particular germplasm, given by the traits analysed, is $k = 2$. The average silhouette width (ASW) from this structure analysis is 0.32 (Figure 12). This shows that the structure of the population of the accessions under study is weak and could, therefore, be artificial.

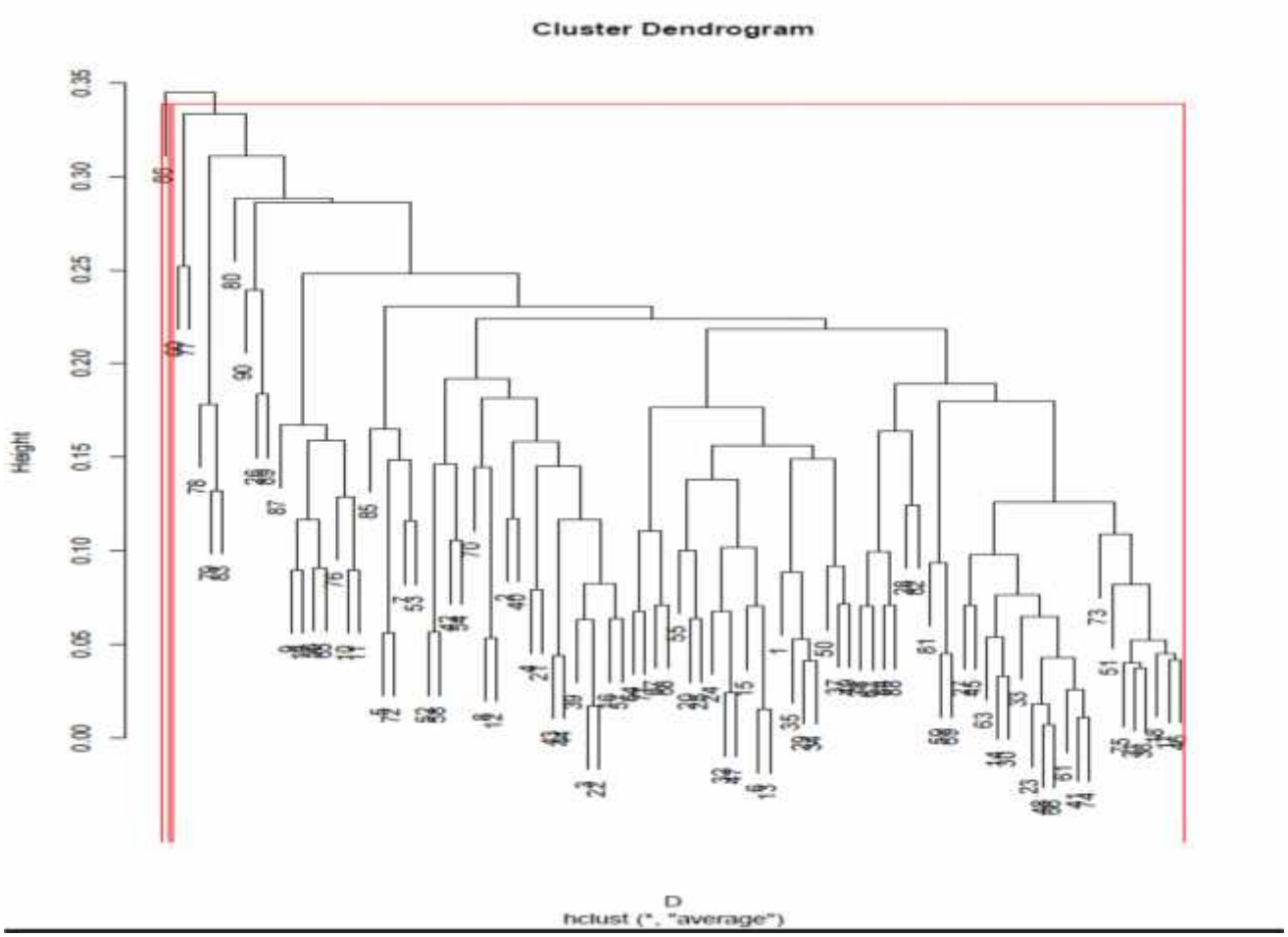

Figure 11: Hierarchical dendrogram based on Gower's dissimilarity matrix calculated from the dataset of 90 blackberry accessions clustered in two groups as revealed by Gowers based clustering model. The two groups are demarcated using the red border line.

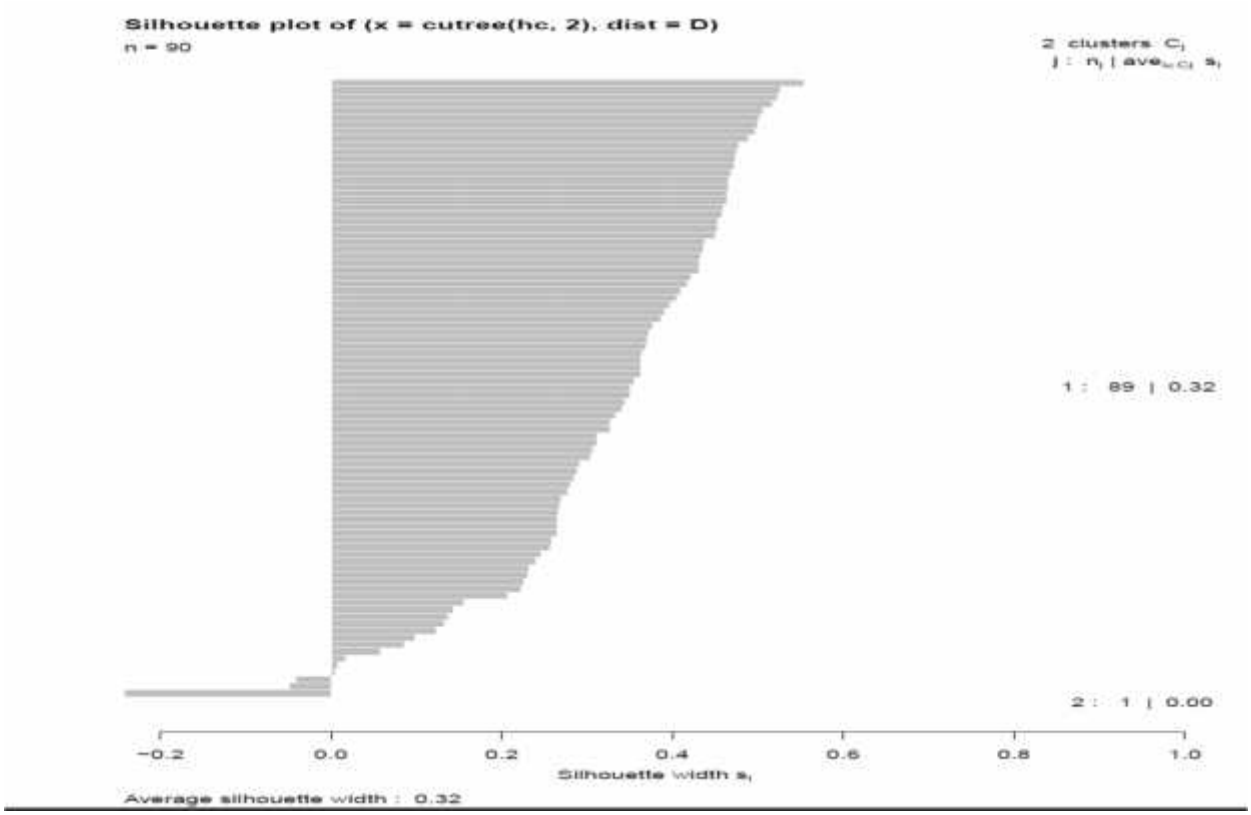

Figure 12: A silhouette showing a graphical display of the structure of the blackberry germplasm.

4.7 Discussion

The assessment of variation in morphological traitsin germplasm is the first step in the determination of genetic diversity. It is a prerequisite for conservation and utilization of plant genetic resources (Mason *et al*. 2017). Morphological characterization can also be useful in selection of parents for breeding (Orobiyi *et al*., 2017; Kagimbo *et al*., 2017). Therefore, there is a need to assess the diversity of any crop prior to selection and crossing to better utilize the resources in any breeding programme. In this work, blackberries were studied at different agro-ecological areas and therefore, had differences in morphological expressions (Figure, 8). According to tree characteristics, most of the wild blackberry accessions were semi-erect (87%) and (80%) of genotypes studied had vigorous growth. This study anticipated higher morphological diversity due to the inclusion of introduced germplasm from South Africa. However, a narrow diversity was observed (ASW=0.32) as per the silhouette plot (Figure 12). This shows that the structure of the population of the accessions under study is weak and could be artificial. The introduced germplasm from South Africa had their origins from Europe and North America. Neither Kenya nor South Africa is a centre of origin for blackberry.

The observed low genetic diversity might be due a number of reasons including nature of propagation of the crop, method of dispersal of the crop, effect of environment and farmer to farmer exchanges of germplasm in the case of cultivated types. Blackberry reproductive nature is complex. This varies from sexual to facultatively apomictic to obligately apomictic. Blackberry (*Rubus* subgenus *Rubus* Watson) are often hermaphrodites (Nybom 1986) and outcrossing has also been observed (Antonius and Nybom, 1995). Additionally, self-fertilization is frequent (Nybom 1988). Infertility or partial fertility may also occur in some plants, and this is attributed to genetic factors such as poor pollen production, unattractive nature to pollinators, lack of pollinators and environmental effects. Open pollinators are likely to have higher diversities compared to inbreds or vegetatively propagated berries (Stafne and Clark, 2004). Most wild blackberries are clonally propagated by way of root sprouts, underground stems (rhizomes) and branches that root at the tips (stolons). Therefore, the number of breeding parents may be few, thus, low diversity. Also since the crop is an invasive species, and propagate vegetatively very vigorously, enabling the clonal spread of single individuals in a patch of habitat. This, however, does not mean that there is a narrow genetic base or higher homozygosity of blackberries in Kenya. In retrospect, the genetic base of fruit tree crop is varied and can be attributed to the different species available. Cross-pollination could also occur and this may explain the morphological redundancies which could be attributed to intra and interspecific hybridization coupled with environmental influence.

A number of morphological and physiological traits were measured for the accessions tested. Thornlessness was one of the morphological traits studied. The progress of breeding blackberry genotypes is directly affected by the plant thorniness. Thornlessness is the most bred qualitative trait in blackberry. Four genes have been detected to be responsible for thornlessness, and they can vary from recessive to dominant for the qualitative trait. Breeding progress is thus, hampered by the source of the thornless genotypes and the ploidy of the blackberry type in question. Plant flower colour varied from white, pink to purple in which white was the most dominant color (62%). Differences in plant flower colour and plant pubescence colours are commonly noticed in natural and introduced blackberry germplasms. This was evident in the data taken and is important as it influences pollination and diversity of the accessions. Some accessions, for instance CV/RBN/01, was more divergent than others and this may be attributed to their outcrossing nature (Figure 8 and 11).

Cluster analysis on the characteristics split the accessions into two groups, I and II. These groups were in a random manner irrespective of their geographical origin. The two distinct groupings were due to the availability of introduced germplasm (CV/RBN/01) that

singly constituted group I. The rest of the wild types (landraces) grouped together, albeit with subgroups. Although there was no clear association between the subgroups and counties of origin, most germplasm from Nakuru County tended to group together. This may be explained in terms of gene pool concept where the wild types formed the primary gene pool, which consists of the crop species itself and other species that can be easily crossed with it. The cultivated type may have grouped alone (CV/RBN/01) as variability in cultivated plant species depends on how evolutionary forces impact on natural populations. The sole grouping of CV/RBN/01 over the rest of the accessions in this study may also be due to the selective advantage it has over the rest of the accessions. This can be by way of mutation, genetic drift that is as a result of random changes in allele frequencies for generations due to the finite size of populations, gene exchanges or gene flow among populations and selections (both natural and artificial selection). The high similarity of the wild types in morphometry and agro-morphological traits across the different agro-ecological zones may be attributed to the invasive nature of the accessions characterized.

As stated above, cluster analysis for qualitative and quantitative grouped the accessions in a random manner irrespective of their geographical origins. This is in agreement with findings of the Principal Component Analysis (PCA). PCA also did not associate the accessions with their regions of origin. Out of the 10 traits subjected to PCA, 8 were able to differentiate the collected accessions and are considered as variables that are capable of discriminating accessions on the basis of morphology. It was evident that principal component analysis also categorized assessed phenotypic traits in the population into several related groups (Figure 7). This can also be explained by the reproductive and often invasive nature of the fruit tree species over wide ethno-geographical regions as well as the folk nomenclature that exists in these areas. Apomixis (referring to seed formation without fertilization) also occurs in some blackberry species (wild and introduced). Therefore, clones dispersed by man or birds spatially across habitats can be a cause. This often results in misclassification of genotypes and existence of duplicates (Agre *et al*., 2017; Mason *et al*., 2015).

4.8 Conclusion

This work is important and can be used with DNA genotyping information to understand the morphological variations that are present in blackberries. There exist morphological variations in the germplasm studied. Selections can be made in the various clusters obtained. Although a more detailed study is required, this work is of great significance in management of plant genetic resources and in blackberry breeding programmes.

References

Agre, A. P., Bhattacharjee, R., Dansi, A., Becerra Lopez-Lavalle, L. A., Dansi, M., and Sanni, A. (2017). Assessment of cassava (*Manihot esculenta* Crantz) diversity, loss of landraces and farmers preference criteria in southern Benin using farmers' participatory approach. *Genetic Resources and Crop Evolution*, 64(2), 307–320.

Antonius, K., and Nybom, H. (1995). Discrimination between Sexual Recombination and Apomixes Automixis in a Rubus Plant Breeding Programme. *Hereditas, 123*(3), 205-213.

Bozovic, D., Lazovic, B., Ercisli, S., Adakalic, M., Jacimovic, V., and Sezer, I. (2016). Morphological characterization of autochthonous apple genetic resources in Montenegro. *Erwerbs-Obstbau*, 58(2), 93–102.

Bravo, L. (1998). Polyphenols: Chemistry, Dietary Sources, Metabolism, and Nutritional Significance. *Nutrition Reviews,* 56, 317-333.

Chittaranjan, T.W. (2011). Blackberry and blueberry pests and diseases. *Indian Journal of Fruit Science,* 25, 54-68.

Clark, J.R., Stafne, E.T., Hall, H.K. and Finn, C.E. (2007). Blackberry breeding and genetics. *Plant Breeding Reviews,* 29, 82-101.

Finn, C. E. (2008). Blackberries. In: Temperate Fruit and Crop Breeding. (J.F Hancock, ed.). Springer Science.

Genstat Committee. (2015). Genstat v17. 0. *VSNI, Hemel Hempstead, UK*.

Gower, J. C. (1971). A General Coefficient of Similarity and Some of Its Properties. *Biometrics* 27(4), 857.

Harlan, J., and de Wet, J. (1971). Toward a Rational Classification of Cultivated Plants. *Taxon*, 20(4), 509–517.

Hirsch, G.E., Vizzotto, M., Aboy, A.L., Henriques, A.T. and Emanuelli, T. (2013). Antioxidant Activity of Blackberry (*Rubus sp*.) Genotypes from the Southern Region of Brazil. *B. CEPPA Curitiba,* 31(1), 83-98.

Hollman, P. C. H., Hertog, M. G. L. and Katan, M. B. (1996). Analysis and health effects of flavonoids. *Food Chemicals,* 57, 43-46.

Kagimbo, F. M., Shimelis, H., and Sibiya, J. (2017). Diversity assessment of sweet potato germplasm collections for yield and yield-related traits in western Tanzania. *Acta Agriculturae Scandinavica, Section B - Soil and Plant Science,* 1–9.

Karimi, H. R., Zamani, Z., Ebadi, A., and Fatahi, M. R. (2009). Morphological diversity of Pistacia species in Iran. *Genetic Resources and Crop Evolution, 56*(4), 561–571.

Lauri, P. and Normand, F. (2017). Are leaves only involved in flowering? Bridging the gap between structural botany and functional morphology, *Tree Physiology*, *37* (9), 1137-1139.

Lee, J., Dossett, M. and Finn, C.E. (2011). *Rubus* fruit phenolic research: The good, the bad, and the confusing. *Food Chemistry*, 130 (4), 785-796.

Marinoni, D., Akkak, A., Bounous, G., Edwards, K.J. and Botta, R. (2003). Development and characterization of microsatellite markers in *Castanea sativa* (Mill.). *Molecular Breeding*, 11, 127–136.

Mason, A.S., Zhang, J., Tollenaere, R., Teuber, P.V., Dalton-Morgan, J., Hu, L., Yan, G., Edwards, D., Redden R. and Batley J. (2015). High-throughput genotyping for species identification and diversity assessment in germplasm collections. *Molecular Ecology Resources*, 10, 1755-1798.

Mason, A.S., Chauhan, P., Surinder, B., S, Phil., Barbetti, M.J., and Batley, J. (2017). Agricultural selection and presence-absence variation in spring-type canola germplasm. *Crop and Pasture Science*, doi.org/10.1071/CP17161.

Menkir, A., Goldsbrough P. and Ejeta, G. (1997). RAPD based assessment of genetic diversity in cultivated races of sorghum. *Crop Science*, 37, 564-569.

Nybom, H. (1986). Chromosome Numbers and Reproduction in *Rubus* subgen. *Malachobatus*. *Plant Systematics and Evolution*, 152, 211-218.

Nybom, H. (1988). Apomixis versus sexuality in blackberries. *Plant Systematics and Evolution*, *160*(Christen 1950), 207–218.

Orobiyi, A., Loko, L. Y., Sanoussi, F., Agré, A. P., Korie, N., Gbaguidi, A., Adjatin, A., Agbangla, C., Dansi, A. (2017). Agro-morphological characterization of chilli pepper landraces (Capsicum annuum L.) cultivated in Northern Benin. *Genetic Resources and Crop Evolution*, 9, 1–15.

Oyoo, M.E., Muhammed, N., Githiri, S.M., Ojwang, P.O., Muniu, F.K., Masha, E. and Owuoche, J. (2015). *In-situ* morphological characterization of coconut in the coastal lowlands of Kenya. *African Journal of crop science*, 9(2), 65-74.

Perrier, X., Flori, A. and Bonnot, F. (2003). "Methods of data analysis." In *Genetic diversity of cultivated tropical plants*, 47-80.

Perrier X, Jacquemoud-Collet JP (2006). DARwin software http://darwin.cirad.fr/

R Development Core Team. 2015. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing Vienna Austria* 0: {ISBN} 3-900051-07-0.

Rao, A. V. and Snyder, D. M. (2010). Raspberries and human health: A review. *Journal of Agricultural and Food Chemistry,* 58, 3871–3883.

SAS Institute. (2001). SAS procedure for personal computers. Version 8 SAS Institute, Cary, NC, USA.

Siriwoharn, T., Wrolstad, R.E., Finn, C.E. and Pereira, C.B. (2004). Influence of cultivar, maturity and sampling on blackberry (*Rubus* L. hybrids) anthocyanins, polyphenolics, and antioxidant properties. *Journal of Agricultural and Food Chemistry*, 52, 8021–8030.

Stafne, E. T., and Clark, J. R. (2004). Genetic relatedness among eastern North American blackberry cultivars based on pedigree analysis. *Euphytica*, *139*(2), 95–104.

Strik, B.C., Clark, J.R., Finn, C.E. and Banados, M.P. (2007). Worldwide blackberry production. *Horticulture Technology,* 17, 205–213.

Wood, G. A., Andersen, M. T., Forster, R. L., Braithwaite, M. and Hall. H. K. (1999). History of Boysenberry and Youngberry in New Zealand in relation to their problems with Boysenberry decline, the association of a fungal pathogen, and possibly a phytoplasma, with this disease. *New Zealand Journal of Crop and Horticultural Science*, 27, 281–295.

Yan, W., Wang, X., Chen, Q., Zhang, L., Tang, H, Luo, Y. and Liu, Z. (2015). Phylogenetic insight into subgenera *Idaeobatus* and *Malachobatus* (*Rubus*, Rosaceae) inferring from ISH analysis. *Molecular Cytogenetics,* 8:11-24.

Yin, M. (2017). Studies in Blackberry : Development and Implementation of a Phenotyping Protocol for Blackberry Seedling Populations and Impact of Time of Day of Harvest on Red Drupelet Reversion for University of Arkansas Blackberry Genotypes. *Theses and Dissertations*. 2417. http://scholarworks.uark.edu/etd/2417. Retrieved on 10th January 2018.

CHAPTER FIVE
GENERAL DISCUSSION, CONCLUSION AND RECOMMENDATIONS

5.1 General discussion

Blackberry is a minority crop in the world. This is in terms of adoption, research activities, awareness and use. However, the crop is gaining prominence in the world. This can be attributed to its health and pharmacological properties. These include anti-inflammatory, antioxidant, chemopreventive and anti-carcinogenic properties especially against colon, oesophageal and oral cancers. Although, important, little research has been done in the crop as molecular genetics in blackberry has remained largely unexplored. This is the first work of its kind in Kenya. So far, research on blackberry has largely focused on pedigree analysis and identification of polymorphic regions. The number of polymorphic markers determines the ability to identify recombination break points in organisms. This information can be used to identify each individual distinctly from the other and develop a phenotype and genotype data set inferring probable locations of genes controlling or contributing to a trait. This can also be used in identifying marker polymorphisms that co-segregate or are associated with significant phenotypic variation, especially, in segregating populations. Various data analysis techniques conducted on both molecular and morphological data were random and failed to group the accessions according to their regions of collection. However, discriminant analysis of the accessions based on regions of collection, grouped cultivated types differently from the wild types. This may be attributed to variations due to crossing based on genetic distances in the cultivated types resulting in changes in phenotypic expressions. Only accessions from Nakuru and Kericho tended to group together (group II), although some germplasm from these counties were found in other clusters. From this research, selection of blackberry accessions for breeding purposes, based on area of origin could potentially be misleading as some wild types from different regions can be morphologically similar. The use of molecular data in this case is imperative as it is non-ambiguous and not influenced by environment.

5.2 Morphological characterization

The morphological traits taken were subjected to Principal component analysis (PCA) and much of the variation observed were accounted for by the number of internodes per average growing shoots, plant thorniness and internode length, for PC1 and PC2. Higher loadings were observed for plant thorniness, an important qualitative trait that is under improvement in most blackberry breeding programmes. Plant thorniness is the most bred trait in blackberry breeding

programmes. However, this trait varies from recessive to dominance and is usually considered a modifying trait. Additionally, this trait has a complex inheritance strategy that varies from disomic to tetrasomic in diploid and poplyploid blackberry species, respectively. This in essence means that the source of the thornless genotype and the ploidy of the blackberry species to be improved have to be determined before crossing. Other important morphological traits identified from this study are number of internodes per actively growing shoot and the internode length. Biplot analysis for the first two principal components of the morphological data grouped the accessions into two distinct groups. However, it was random and did not group the accessions based on either folk nomenclature or regions of geographical origin, and this has been attributed to be due to the effect of environmental conditions in the natural environments, plant dispersal mechanisms, method of reproduction and farmer to farmer seed exchanges. This is also evident in the variables factor biplots analysis based on the first two principal scores of a discriminant analysis of the blackberry accessions. The grouping of the accessions based on the cluster dendrogram can be used for idetifying different accessions can be identified per cluster and used as potential parents for varietal development.

5.3 Comparison between genetic diversity based on SSR markers with morphological traits

This study employed 11 out of the available 13 SSR markers previously described by Castillo *et al*., 2013 and 7 important breeding morphological traits adapted from the Standardized blackberry phenotyping protocol (Yin, 2017). Genetic distance estimates were generated using dendrograms which showed that the morphological markers had lower genetic diversity in comparison to molecular markers which showed relatively higher within accessions genetic diversity, although among accessions variation exhibited lower genetic diversity estimates. This however, should not be interpreted to infer lower diversities due to nature of the crop and the large number of species in blackberry. In retrospect, the first three axes using SSRs accounted for 55.48% of the total variations with each axes explaining 30.04%, 13.53% and 11.91% of the variation, respectively at 95% confidence while on the other hand, morphological traits had the first two axes having a discriminatory variance of 52.71%. This implies that there are varied blackberry species in the accessions collected and furthermore, there exists considerable diversity in morphological traits. This inference based on morphological markers could however, be potentially misleading as some accessions from different regions could have been morphologically similar and in such situations, molecular markers are preferable. The amount genetic variation in plant introductions (cultivated crops)

is influenced by the interaction of evolutionary forces as natural populations. These forces include selection, mutation, genetic drift characterized by random changes in allele frequencies among generations due to the finite size of populations, gene exchanges and gene flow among populations. This, in addition to population structure, has been used to analyse the evolutionary mechanisms at work in natural populations has be used to study wild and cultivated populations and to identify the main forces at work before and after establishing breeding programs in different crop species. Therefore, the heterogeneity in genetic variation within and among accessions used in this study is in fact determined by mutation, genetic drift, and natural selection across time and regions of collection. This is evident in the results obtained above whereby the selection of the heterozygote in blackberry is preferable. Additionally, the SSR markers used in this study are considered to be selectively neutral (Kimura, 1983) and therefore, the genetic level and population genetic structure revealed in this research reflect the effect of demographic factors such as migration and genetic drift.

5.4 Impact of the findings

This research explores the genetic diversity in *in-situ* and *ex-situ* blackberry genetic resources. It is a precursor to crossing blackberry accessions in the country based on the genetic distances/similarity estimates for various regions of germplasm collection. It also detected hybridization in the wild and these hybrids can be collected and maintained by wat of clonal propagation. This research also demystifies the ambiguities that existed in the blackberry germplasm repositories and redundancies in collections in gene banks and hence, informed selection of core blackberry germplasm collections. The complementarity in the use of morphological and molecular markers is vital for differentiation of accessions as some crops tend to have similarity in morphology but different in genotypic composition. For instance, SSR markers employed in this research did not cluster the accessions based on areas of their origin, and hence, plant breeders can track selected accessions that are genetically diverse to develop crosses. In this study, there was synergy between the morphological and SSR markers suggesting better inference of diversity estimates found in the research.

5.5 Conclusion

This study detected diversity in the blackberry accessions and identified important outcrossing blackberry types from both morphological and DNA analysis. These accessions also exhibited higher heterozygosities and can be used as parents for crossing based on genetic distances.

5.6 Recommendations

1. *Ex-situ* characterization and chemical studies is also recommended especially for blackberry fruits.
2. The outcrossing accessions identified in this study should be selected as parents and crossed to develop new hybrids that can be evaluated on important blackberry quality traits.
3. Selections can also be done both in molecular and morphological study based on the dendrograms generated. Few accessions can be sampled in each clusters and more detailed studies done on the chemopreventive, anti-oxidant and phytochemical properties of the blackberry.

5.7 References

Castillo, N. R. F., Reed, B. M., Graham, J., Fernandez-Fernandez, F., and Bassil, N. V. (2010). Microsatellite Markers for Raspberry and Blackberry. *Journal of the American Society for Horticultural Science*, *135*(3), 271–278.

Yin, M. (2017). Studies in Blackberry : Development and Implementation of a Phenotyping Protocol for Blackberry Seedling Populations and Impact of Time of Day of Harvest on Red Drupelet Reversion for University of Arkansas Blackberry Genotypes. *Theses and Dissertations*. 2417. http://scholarworks.uark.edu/etd/2417. Retrieved on 10th January 2018.

APPENDICES

Appendix 1: Preparation of standard solutions

CTAB: for 1L of CTAB buffer
100 ml of 1 M Tris, pH 8.0
280 ml of 5 M NaCl
40 ml of 0.5 M EDTA
20 g of CTAB (Cetyl Trimethyl Ammonium Bromide)
to 1L with H_20

1 M Tris, pH 8.0: for 1 L
121.1 g Tris
700 ml ddH_2O
Dissolve tris and bring to 900 ml.
pH to 8.0 with concentrated HCl (will need ~50ml)
Bring to 1 L.

0.5 M EDTA pH 8.0: for 1 L
186.12 g of EDTA
750 ml ddH_2O
Add about 20 g of NaOH pellets
Slowly add more NaOH until pH is 8.0, EDTA will not dissolve until the pH is near 8.0.

5 M NaCl: for 1 L
292.2 g of NaCl
700 ml ddH_2O
Dissolve and bring to 1 L.

7.5 M Ammonium acetate: for 250 ml
144.5 g ammonium acetate. Add to volume with ddH_20

0.1 M sodium borate pH 8.0: for 1 L
4.76 g Boric Acid (Cat # B-1934)
2.54 g Borax (Cat # B-0127)
Add 1000 mL of ddH2O
Slowly add more NaOH until pH is 8.0

Appendix 2: Names and sources of blackberry germplasm used in this study

SN No.	County	District	Division	Location	Village	Dialect
1	Nakuru	Njoro	Mau Narok	Mau Narok	Korofio	Dorobo
2	Nakuru	Njoro	Mau Narok	Mau Narok	Korofio	Dorobo
3	Nakuru	Njoro	Mau Narok	Mau Narok	Korofio	Dorobo
4	Nakuru	Njoro	Mau Narok	Mau Narok	Korofio	Kikuyu
5	Nakuru	Njoro	Mau Narok	Mau Narok	Sigaon	Tugen
6	Nakuru	Njoro	Nessuit	Nessuit	Tritagon	Tugen
7	Nakuru	Njoro	Nessuit	Nessuit	Tritagon	Tugen
8	Nakuru	Njoro	Nessuit	Nessuit	Tritagon	Tugen
9	Nakuru	Njoro	Nessuit	Nessuit	Kimundu	Kikuyu
10	Nakuru	Molo	Mau Summit	Molo-Kericho junction	Matumaini	Kisii
11	Nakuru	Molo	Mau Summit	Molo-Kericho junction	Matumaini	Kisii
12	Nakuru	Molo	Mau Summit	Molo-Kericho junction	Matumaini	Kisii
13	Nakuru	Molo	Sachangwan	Sachangwan	GSU camp	Kisii
14	Nakuru	Molo	Sachangwan	Sachangwan	GSU camp	Kipsigis
15	Nakuru	Elburgon	Elburgon	Salama	Arimi	Kipsigis
16	Nakuru	Elburgon	Elburgon	Salama	Arimi	Tugen
17	Nakuru	Elburgon	Elburgon	Salama	Arimi	Tugen
18	Nakuru	Subukia East	Bahati	Dundori	Githiori	Kikuyu
19	Nakuru	Subukia East	Bahati	Dundori	Githiori	Kikuyu
20	Nakuru	Subukia East	Bahati	Dundori	Githiori	Kikuyu
21	Kericho	Cheboswa	Cheymen	Cheymen	Unilever	Nandi
22	Kericho	Cheboswa	Cheymen	Cheymen	Unilever	Nandi
23	Kericho	Cheboswa	Cheymen	Cheymen	Unilever	Nandi
24	Kericho	Cheboswa	Cheymen	Cheymen	Unilever	Nandi
25	Kericho	Cheboswa	Cheymen	Cheymen	Unilever	Nandi
26	Kericho	Cheboswa	Cheymen	Cheymen	Unilever	Nandi
27	Kericho	Cheboswa	Cheymen	Cheymen	Unilever	Nandi
28	Kericho	Cheboswa	Kericho	Ainamoi	Chesinende	Nandi
29	Kericho	kIpkelion	Kericho	Chepsir	Chepsir	Nandi
30	Kericho	kIpkelion	Kericho	Chepsir	Chepsir	Nandi
31	Kericho	kIpkelion	Kericho	Chepsir	Chepsir	Nandi
32	Kericho	kIpkelion	Kericho	Chepsir	Chepsir	Nandi
33	Kericho	Letien	Buret	Chemoiben	Kerega	Nandi
34	Kericho	Letien	Buret	Chemoiben	Kerega	Nandi
35	Kericho	Letien	Buret	Chemoiben	Kerega	Nandi
36	Kericho	Letien	Buret	Chemoiben	Kerega	Nandi
37	Baringo	Tinet	Torongo	Lembus Mosop	Makutano	Kipsigis

Appendix 3: continued

38	Baringo	Tinet	Torongo	Lembus Mosop	Makutano	Kipsigis
39	Baringo	Tinet	Torongo	Lembus Mosop	Makutano	Kipsigis
40	Baringo	Tinet	Torongo	Lembus Mosop	Makutano	Kipsigis
41	Baringo	Tinet	Torongo	Lembus Mosop	Makutano	Kipsigis
42	Baringo	Tinet	Torongo	Lembus Mosop	Makutano	Kipsigis
43	Baringo	Tinet	Torongo	Lembus Mosop	Makutano	Kipsigis
44	Baringo	Tinet	Tinet	Chemosus	Chemosus Forest	Kipsigis
45	Baringo	Tinet	Tinet	Chemosus	Chemosus Forest	Kipsigis
46	Baringo	Tinet	Tinet	Chemosus	Chemosus Forest	Kipsigis
47	Baringo	Tinet	Tinet	Chemosus	Chemosus Forest	Kipsigis
48	Baringo	Eldama Ravine	Lembus	Lembus Central	Metipso	Kipsigis
49	Baringo	Eldama Ravine	Lembus	Lembus Central	Metipso	Kipsigis
50	Baringo	Eldama Ravine	Igure	Maji Mazuri	Kinare	Kipsigis
51	Baringo	Eldama Ravine	Igure	Maji Mazuri	Kinare	Kipsigis
52	Baringo	Eldama Ravine	Igure	Maji Mazuri	Kinare	Kipsigis
53	Baringo	Eldama Ravine	Igure	Maji Mazuri	Kinare	Kipsigis
54	Baringo	Eldama Ravine	Timboroa	Mumberes	Kirima	Kikuyu
55	Baringo	Eldama Ravine	Timboroa	Mumberes	Kirima	Kikuyu
56	Baringo	Eldama Ravine	Timboroa	Mumberes	Kirima	Kikuyu
57	Baringo	Eldama Ravine	Timboroa	Mumberes	Kirima	Kikuyu
58	Gishu	Kipkabus	Eldoret Soth	Burnt Forest	Bayete	Kipsigis
59	Nandi	Nandi North	Kurgung	Chepkoiyo	Surungai	Nandi
60	Nandi	Nandi North	Kurgung	Chepkoiyo	Surungai	Nandi
61	Nandi	Nandi North	Kurgung	Chepkoiyo	Surungai	Nandi
62	Nandi	Nandi North	Kapsabet	Kapsabet	Kambolowa	Nandi
63	Nandi	Nandi North	Kapsabet	Kapsabet	Kambolowa	Nandi

Appendix 3: continued

64	Nandi	Nandi North	Kapsabet	Kapsabet	Kambolowa	Nandi
65	Nandi	Kabiyet	Sangalo	Septonok	Septonok	Nandi
66	Nandi	Kabiyet	Sangalo	Septonok	Septonok	Nandi
67	Nandi	Kabiyet	Sangalo	Septonok	Septonok	Nandi
68	Nandi	Nandi	Baraton	Baraton	UEAB	luhya
69	Nandi	Nandi	Baraton	Baraton	UEAB	Luhya/Nandi
70	Nandi	Chesumei	Chemundu	Kapng'etuny	Kipkigei	Luhya
71	Nandi	Chesumei	Chemundu	Kapng'etuny	Kipkigei	Nandi
72	Nandi	Chesumei	Chemundu	Kapng'etuny	Kipkigei	Kikuyu
73	Laikipia	Laikipia	Nyahururu	Laikipia	Laikipia University	
74	Laikipia	Laikipia	Nyahururu	Laikipia	Laikipia University	
75	Laikipia	Laikipia West	Lumuria	Riverside	Riverside	Masai
76	Laikipia	Laikipia West	Lumuria	Riverside	Riverside	Masai
77	Laikipia	Laikipia West	Gituamba	Karandi	River Katito/Forest	Masai
78	Gishu Uasin	Kipkabus	Ainabkoi	Kaptagat	Chesogor	Nandi
79	Gishu Uasin	Kipkabus	Ainabkoi	Kaptagat	Chesogor	Nandi
80	Gishu Uasin	Kipkabus	Ainabkoi	Kaptagat	Chesogor	Tiriki
81	Gishu Uasin	Kipkabus	Ainabkoi	Kaptagat	Flax	Tiriki
82	Gishu Uasin	Kipkabus	Ainabkoi	Chagaiya	Anguina	Nandi
83	Gishu	Kesses	Kipchumo	Lengut	Plateau	Nandi
84	Laikipia	Laikipia West	Gituamba	Karandi	River Katito/Forest	
85	cultivated	Var. Roben				
86	cultivated	Var. Boysen				
87	Nakuru	Njoro	Egerton	Njokerio	R.Ndaragu	Kikuyu
88	Nakuru	Njoro	Egerton	Egerton	Field 3	.
89	Nakuru	Njoro	Egerton	Egerton	Field 3	.
90	Nakuru	Njoro	Egerton	Egerton	botanic garden	Ndorobo

Appendix 4: UPGMA clustering using R program for statistical computing

```
#Own data from common population for  beanfly resistance project
#Using - UPGMA clustering and visualization with gower metric
#set a working directory

setwd("C:/Users/kenya13/Desktop/BECA_ANAL/R_cbp_Gower/Pacal_cbp/Pascal_cbp")

#Data importation and structure of data
berry
str(berry)
Accession <- read.csv("berry.csv")

# the command below shows the first top 4 rows, you can choose any number

head(Accession,4)

#check the dimentions of the table

dim(Accession)
str(Accession)

#obtain summaries of missing data point presented as NA
#for example trait (morphological data) FLw_CLR, Growth habit etc
summary(Accession$Stem.Symmetry)

#With function daisy() in package 'cluster':
berry <- Accession[,-1]
row.names(berry) <- Accession[,1]
str(berry)
#Need the library cluster which should be activated

library(cluster)
D <- daisy(berry, metric = "gower")
hc <- hclust(D, method = "average")

#Visualizations

plot(hc)
heatmap(as.matrix(D),Rowv=as.dendrogram(hc),Colv=as.dendrogram(hc))

#Cluster number determination
#Silhouette plots
plot(silhouette(cutree(hc,2),D))
```

```
plot(silhouette(cutree(hc,3),D))
plot(silhouette(cutree(hc,4),D))
plot(silhouette(cutree(hc,5),D))
plot(silhouette(cutree(hc,20),D))


plot(hc)
rect.hclust(hc, k=2, border="red")

#Overlaying additional information

table(cutree(hc,2),ecozone$AGRO_ECOL)
```