MOLECULAR CHARACTERIZATION OF KENYAN COMMON BEAN GERMPLASM USING SSR AND PEROXIDASE GENE MARKERS

Pam Joshua Gyang (BSc, University of Jos)

Registration number: H56/75471/2014

A thesis submitted in partial fulfillment of the requirements for the award of the degree of Master of Science in Biochemistry of the University of Nairobi

December, 2018

DECLARATION

I, Pam Joshua Gyang declare that this thesis is my original work and has not been presented for a degree or any other award in any other university.

Name: Pam Joshua Gyang

Registration Number: H56/75471/2014

Signature..... **Date.....**

Supervisors

We confirm that the work reported in this thesis was carried out by the candidate under our supervision:

Dr. Evans Nyaboga

Department of Biochemistry, University of Nairobi

Signature..... **Date.....**

Dr. Edward Muge

Department of Biochemistry, University of Nairobi

Signature..... **Date.....**

DEDICATION

I wish to dedicate this thesis, which shows all the research work that I conducted here at the Department of Biochemistry, University of Nairobi, to the Almighty God, for His protection, abundant grace, endless mercies, surplus provision, and for being my source of inspiration throughout my period of study, and to my parents, for their unending support.

ACKNOWLEDGEMENTS

I acknowledge the source of my inspiration, the one who holds all things together; whose faithfulness and mercy kept me this far, God almighty.

My endless gratitude goes to my supervisor Dr. Evans Nyaboga for reading through every chapter meticulously and making useful inputs as well as being very understanding and caring at all times. I also appreciate him for providing all the materials and chemical reagents for this research work. I also wish to express my deep gratitude to my second supervisor, Dr. Edward Muge for his unanimous support, encouragement and guidance and finding time to read through this work and make corrections despite his tight schedule.

To my lovely and blessed dad, Prof. Gyang Pam, for the heroically role he played. Dad, thank you very much for your dogged determination to see I get the best education. I will also like to appreciate my mum, Chundung Gyang for her self-denial to see to my self-fulfillment. Mum, your prayers and moral support have been very fruitful to me. I also appreciate my siblings, Jerry and John, for always being there for me, encouraging me whenever I feel down.

I would like to convey my honest gratitude to all those who have been helpful to me in the course of my studies at the University of Nairobi, and during my work in the lab. I thank my colleagues Gilbert Osena, Irene Onyango, Keisha Njenga, Eric Osewo, Peter Oluoch, Francis Kariuki, Gladys Ada, and Damaris Ondieki for their support and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

AFLP	Amplified Fragment Length Polymorphism
AMOVA	Analysis of molecular variance
bp	base pairs
CIAT	Centro Internacionla de Agricultura Tropical
CTAB	Cetyltrimethyl-ammonium bromide
df	Degrees of freedom
DNA	Deoxyribonucleic acid
EDTA	Ethylenediamine tetra-acetic acid
g	Gram
He	Nei's gene diversity
Ho	Observed heterozygosity
I	Shannon's information index
ISDR	International Strategy for Disaster Reduction
ISSR	Inter Simple Sequence Repeat
MCMC	Markov Chain Monte Carlo
ml	Milliliters
MBP	Mega base pairs
MS	Mean of squares
N	Number of observed alleles
Ne	Effective number of alleles
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism

PCA	Principal Component analysis
PCR	Polymerase chain reaction
PIC	Polymorphic information content
POP	Population
ROS	Reactive oxygen species
POX	Peroxidase gene
Rpm	Revolutions per minute
SS	Sum of squares
SSAP	Sequence-specific amplification polymorphism
SSRs	Simple sequence repeats or microsatellite
TAE	Tris-acetate EDTA buffer
TE	Tris-EDTA buffer
UPGMA	Unweighted pair-group method with arithmetic mean
v/v	Volume per volume
w/v	Weight per volume
%	Percentage
µg	Microgram
µl	Microliter

ABSTRACT

Common bean (*Phaseolus vulgaris*) is a legume of high nutritional value and the most important legume worldwide for direct human consumption. It serves as a principal source of protein, starch, vitamins, folate and minerals (iron, potassium, phosphorus, magnesium, manganese and to a lesser extent zinc and calcium). There is limited information on the genetic diversity and population structure of common bean genotypes grown in Kenya. The objective of this study was to determine the genetic diversity and population structure of common bean genotypes from different growing regions (Eastern, Central, Rift Valley, Nyanza and Western) in Kenya using molecular markers. Seeds of 46 cultivars were collected from different regions and three seeds of each genotype were established in plastic pots containing sterile soil in the screenhouse. Young leaves from two-week-old plants were used for DNA extraction followed by PCR amplification using simple sequence repeat (SSR) and peroxidase gene (POX) molecular markers. A total of 366 alleles were amplified using five SSR primers across 40 genotypes, with an average of 4.5 alleles per locus. The polymorphism information content (PIC) of the SSR markers ranged from 0.48 to 0.74 with an average of 0.60. The pair wise genetic similarity between common bean genotypes ranged from 0.15 to 1.0 with an average of 0.54. A dendrogram based on the unweighted pair-group method with arithmetic mean (UPGMA) grouped the 40 genotypes into two major clusters. It was notable that the first major cluster was further divided into two-separate sub-clusters, representing genotypes from each of the regions. Principal component analysis (PCA) of the SSR markers showed that the first two principal components (PCs) explained a total of 28.79% of the genetic variation and failed to cluster the genotypes into distinct groups. Analysis of molecular

variance (AMOVA) revealed high levels of genetic variation (87%) within population, compared to the variation that exists among the populations. Using five POX markers, a total of 624 alleles were amplified ranging from 3 to 9 on every locus, with an average of 7.20. The PIC of the POX markers varied from 0.6204 (POX11) to 0.9110 (POX8), with an average of 0.7677. The range of the observed heterozygosity was from 0.6667 (POX 8) to 0.9150 (POX 12) with a mean of 0.7945, while the values of the mean genetic diversity obtained ranged from 0.3072 (POX11) to 0.4425 (POX8), having a mean of 0.3972. Un-weighted pair group method with arithmetic mean (UPGMA) was used to cluster the genotypes into two main genetic clusters, and the genotypes showed no grouping by geographical origins. The highest value of genetic variation was observed between the genotypes obtained from Western, Rift valley and Central regions of Kenya. Population structure analysis using the Bayesian model-based approach grouped the germplasms into 7 gene pools and showed a high genetic admixture within individual genotypes. Furthermore, population structure analysis showed that these 7 gene pools coexisted in genotypes belonging to different geographic regions. AMOVA revealed higher genetic diversity (99%) within population than among population (1%) and offers a reliable base for the design of genetic improvement schemes. This study demonstrates the existence of considerable genetic diversity in common bean genotypes cultivated in Kenya. The wider genetic diversity is important to future generations so that it copes with unpredictable climate changes and human needs. Therefore, these genotypes can be used as a foundation for future breeding programs to produce hybrids of desirable agronomical traits.

CHAPTER ONE

1.0 INTRODUCTION

1.1 Background information of the study

Common bean (*Phaseolus vulgaris* L.) is a grain legume with high nutritional value. It is the most important legume worldwide for direct human consumption, with India, Brazil, Myanmar, China and the USA being the largest world producers. In Africa, Uganda, Kenya, Burundi, and Tanzania the largest producers (FAO, 2014). The world production of common bean is 23 million tons (FAO, 2014). Common bean constitutes the most important food legume for more than 500 million people in Asia, Latin America and Africa. Many people in Africa rely on it for food security, nutrition and income (Beebe *et al*., 2013). The crop provides protein-rich food, restores and maintains the soil fertility by fixing atmospheric nitrogen, and also fits well in many cropping systems. Beans are also a rich source of essential vitamins and minerals, soluble fiber, starch and phytochemicals, and are also reported to have low fat content (Beebe *et al*., 2000; Svetleva *et al*., 2006; Nyombaire *et al*., 2007; Blair *et al*., 2013). It is the most important plant-based protein source for the people of Kenya, providing 25 % of the protein of the local diet (Broughton *et al*., 2003). More than half of the farmers in Kenya grow beans and it is widely adapted for growth in most agricultural regions (Katungi *et al*., 2011). A wide variety of beans are grown here; these include: Rose Coco, Mwitemania, Wairimu, Mwezi Moja, Canadian W, and KK 15. Despite being an important food crop in Kenya, there has been no focus on understanding the genetic diversity of the germplasm used in cultivation.

Characterization of the genetic diversity in the available germplasm provides essential information for its conservation, management of genebanks and utilization in genetic breeding programs (Arunga *et al*., 2015). To make the crossing program effective, parents should belong to different genetic cluster, hence, the need to know the genetic diversity of the existing genotypes before any hybridization studies.

The narrow genetic base of modern crop cultivars is problematic in breeding to sustain and improve crop productivity due to the vulnerability of genetically uniform cultivars to potentially new biotic and abiotic stresses (Abdurakhmonov and Abdukarimov, 2008). Moreover, assessment of genetic diversity is important in order to know the source of genes responsible for a particular trait (disease resistance, early maturity, and high yielding or drought tolerance) within the available germplasm (Nyakio *et al*., 2015). Therefore, it is necessary to investigate the genetic diversity and population structure in common bean germplasm grown in Kenya to understand and in the future broaden the genetic variation available for breeding.

Selection of diverse Kenyan common bean genotypes is fundamental for their efficient utilization in plant breeding programs and this can only be utilized if gene pool of local and introduced germplasm has sufficient amount of genetic variability. Morphological and agronomic traits are routinely used to assess genetic diversity but are greatly influenced by the environment, development stage and do not correctly reflect genetic relatedness between accessions. To overcome these problems, molecular markers represents a potential tool for effective characterization of genetic diversity and to aid in

the management of plant resources (Blair *et al*., 2006; Grisi *et al*., 2007; Laurentin, 2009; Blair *et al.*, 2011). These DNA markers, when closely linked to genes of interest can be used to select for the desirable allele/s in marker assisted breeding programs (Okii *et al*., 2014). Genetic diversity in common bean has been studied using different molecular markers such as allozymes (Singh *et al*., 1991; Santalla *et al*., 2002), amplified fragment length polymorphism (AFLP) (Lioi *et al*., 2005; Svetleva *et al*., 2006), random amplified polymorphic DNA (RAPD) (Ocampo *et al*., 2005, Martins *et al*., 2006; Marotti *et al*., 2007), restriction fragment length polymorphism (RFLP) (Nodari *et al*., 1992), inter-simple sequence repeat (ISSR) (Svetleva *et al*., 2006; Marotti *et al*., 2007) and simple sequence repeat (SSR) (Asfaw *et al*., 2009; Okii *et al*., 2014).

Of all DNA based markers, the most ideal for distinguishing closely related germplasm are microsatellites, which are informative markers that detect length polymorphisms at loci with simple sequence repeats (Powell *et al*., 1996). The advantages of SSR markers for genome fingerprinting include highly polymorphic, high allelic diversity, frequently co-dominant, highly reproducible, specific PCR-based assay, randomly and widely distributed in the genome (Pejic *et al*., 1998). Moreover, these markers are more closely connected with known function genes and possess higher transferability to related species (Ellis and Burke, 2007). Mutations in the motifs and flanking sequences as well as distribution of microsatellites in the genome of a species are exploited to reveal genetic variation and varietal identity. Microsatellites have been used to evaluate genetic diversity in dry bean genotypes from Italy (Marotti *et al*., 2007), Bulgaria (Svetleva *et al*., 2006), Nicaragua (Gomez *et al*., 2005), Slovenia (Maras *et al*., 2015), Uganda (Okii *et*

al., 2014) and Ethiopia (Fisseha *et al*., 2016). Peroxidase, a plant-specific oxidoreductase, plays an important role in many self-defense interactions in plants. It has become a model enzyme for studying the molecular mechanisms of self-defence in plants (Nemli, 2014). The peroxidase gene families are diverse in plants and can therefore be used as molecular markers to study genotypic diversity and provide new information for common bean breeding programs.

1.2 Problem statement

Despite being a very vital food for direct human consumption, serving as a good source of calories and proteins for many people around the world; common bean production rates have been declining in developing countries. Factors that can contribute to bean yield and quality losses are numerous and include pathogenic microorganisms (like various kinds of fungi, viral and bacterial diseases), plant nutritional deficiencies, drought, and insect pests (De Luque and Creamer, 2014). Therefore, there is a need to develop varieties with different attributes with resistance to major biotic and abiotic constraints, to suit the specific needs of different regions and cropping systems. The development of high yielding and stable varieties requires a continuous supply of new germplasm as a source of desirable genes and/or gene complexes. Use of genetically divergent parental materials can also enhance the level of heterozygosity and therefore hybrid vigor in common bean crossing progenies. However, genetic diversity and population structure of Kenyan common bean germplasm is limited and hence has not been fully and systematically exploited in breeding programs. It is thus important to identify genetically distinct groups (potential groups) that will then be exploited in

population crosses to increase heterozygosity and hybrid vigor in the progeny. In addition, assessing genetic relationships and diversity in Kenyan common bean populations using SSR and POX markers will be helpful in understanding the patterns of diversity of the populations in order to enhance populations with continuous genetic variation for association mapping studies.

1.3 Justification

Good knowledge of genetic variability and population structure is indispensable to effective management and use of genetic resources (Arunga *et al*., 2015). It provides farmers and plant breeders with options to develop through selection and breeding, new and more productive crops that are resistant to virulent pests and diseases as well as being adapted to changing environments (Nyakio *et al*., 2015). Molecular markers based on sequences of DNA can be great tools in accessing the genetic variability of common bean cultivars. SSR markers were used in this study because they are simple, specific, and reproducible and are able to detect polymorphisms even in closely related genotypes. Peroxidases play an important role in plant self-defence (Nemli, 2014). They are diverse in plants and therefore can be used as molecular markers to determine genetic diversity and offer information regarding plants defence mechanisms. The knowledge of genetic variation and relationships among genotypes will help the breeders in developing appropriate breeding strategies to solve problems of low yield in common bean (Khaidizar *et al*., 2012). Therefore, assessment of genetic diversity in the current common bean gene pool would facilitate in plant improvement in developing cultivars for

specific production constraints by providing an index of parental lines to be used in breeding programs through the use of genetic engineering.

1.4 Objectives

1.4.1 Main objective

The main objective of this study was to carryout genetic characterization of common bean germplasm grown in Kenya using molecular markers.

1.4.2 Specific objectives

The specific objectives of this study were:

1. To characterize Kenyan common bean cultivars using simple sequence repeat markers
2. To determine the genetic diversity and population structure of common bean cultivars using peroxidase gene-based markers

1.5 Null hypothesis

(i) SSR markers cannot be used to characterize Kenyan common bean germplasm.

(ii) Peroxidase-gene based markers cannot be used to determine the genetic diversity and population structure of common bean germplasm.

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 Origin, distribution and botany of common bean

Common beans (*Phaseolus vulgaris* L.) belong to the genus *Phaseolus*, family Leguminosae, subfamily Papilionoideae, tribe Phaseoleae and subtribe Phaseolinae (Gepts, 2001). Based on their centres of origin, common bean originated and was domesticated in the New World and has two major gene pools, the Andean and the Mesoamerican, respectively (Blair, 2006). Common beans originated mainly in Latin America, spread to other parts of the world and now it is widely cultivated in the tropics and subtropics as well as in temperate regions of the world (Ghamari *et al*., 2013). The common bean, a self-pollinated crop, is a true diploid (n = 11) with a small genome (Arumuganthan and Earle, 1991). All species of the genus are diploid and most have 22 chromosomes (2n = 2x = 22). A few species show an aneuploid reduction to 20 chromosomes. The genome of common bean is one of the smallest in the legume family at 625 Mbp per haploid genome (Gepts, 2001). Common bean is mostly grown in terrestrial habitats with flower petal colour of blue to purple, pink to red, and white. The leaf types are compound made up of two or more discrete leaflets with one leaf per node along the stem. The edge of the leaf blade is entire (has no teeth or lobes) (Gepts, 2001).

2.2 Production and economic nutritional importance of common bean

Common bean is grown in every continent except Antarctica, with Brazil and India being the largest producers, while China produces by far, the largest quantity of green bean. It is a legume with high nutritional value which originated from South and Central

America. The world production of common bean has been estimated at approximately 23 million tons and it is grown in nearly 150 countries on an estimated 27.7 million hectares (FAO, 2014). Latin America and Asia are the regions of greatest production, followed by Africa. Beans are becoming increasingly commercial with the trends of urbanisation and market globalization, with small farmers organizing themselves to tap into opportunities to export in other countries (Beebe *et al*., 2013). As expected, countries with technified agricultural systems present much higher yields than tropical and developing countries. In the USA, average yields in the past decade range from 1.64 to 1.96 t ha^{-1}, albeit with significant regional differences. Similarly, average yields in Argentina and Colombia are about 1.2 t ha^{-1} due to varietal selection, and in Brazil under intensive management and irrigation, yields average 1.8 t ha^{-1} (Beebe *et al*., 2013).

In Africa, most bean production is found in the eastern and southern highlands, extending from Ethiopia to South Africa, with Kenya, Uganda, Burundi and Tanzania being the largest producers. In West Africa, bean production is localized in specific environments, with Cameroon being the principal producer. Beans are a minor crop in Europe and North Africa, concentrated around the Mediterranean, in Spain, Italy, Morocco, Algeria, and the Balkan states. In Asia, common bean is spread in an extensive band from Turkey through Iran and the Himalayan foothills, and East through Myanmar and China. India is cited as a major producer of common bean, but these figures undoubtedly include other legumes (Beebe *et al*., 2013). Common bean is the most important grain legume for human consumption in the world. Since most protein consumed by the poor is from plant sources, being protein-rich, common bean plays a significant role in the human diet.

Although it is far less important than cereals as a source of calories, bean often supplies a significant proportion of carbohydrates (Beebe *et al*., 2013). It is consumed in many parts of the world and serves as a principal source of protein, starch, vitamins, folate and minerals (iron, potassium, phosphorus, magnesium, manganese and to a lesser extent zinc and calcium) (Anderson *et al*., 1999). The crop also plays an important role in sustaining soil fertility by adding atmospheric nitrogen to the soil through a symbiotic relationship with Rhizobia bacteria – it supplies the bacteria with carbohydrate and the bacteria in turn traps atmospheric nitrogen and adds it to the soil. Furthermore, it is used by most farmers as a cover crop because of its ability to suppress the growth of weeds on the farmland.

2.3 Constraints to common bean production

Despite their nutritional importance, production of common bean has been declining around the world, particularly in developing countries. Factors that can contribute to bean yield and quality losses are the numerous pathogenic microorganisms (like various kinds of fungi, viruses and bacteria) such as bean golden mosaic virus, bean yellow mosaic virus, bacterial brown spot, bacterial wilt, saran, and rost-foot, plant nutritional deficiencies, drought, and insect pests (Wu *et al*., 2016). Drought is by far the most important abiotic constraint to bean production, with a probability of occurrence estimated at 60% in Eastern Kenya. Common bean yield loss due to drought is substantial, with almost all varieties experiencing severe decline in yield when drought occurs, implying low levels of resistance among the cultivars grown by the farmers (Katungi *et al*., 2010).

2.4 Characterization of common bean germplasm

Markers for characterization of plant genetic resources are grouped into three main classes: (i) morphological and productive markers which are based on visually evaluated traits, (ii) biochemical markers which are based on gene product, and (iii) molecular markers which are founded on DNA analysis (Galal *et al*., 2013). Classical methods for characterization of plant germplasm involve the use of morphological and agronomic traits (Homar *et al*., 2011). However, the use of morpho-agronomic traits are influenced by the environment, development stage and do not correctly reflect genetic relatedness between different accessions. To overcome these problems, molecular markers represent a potential tool for effective characterization of genetic diversity and to aid in the management of plant resources (Blair *et al*., 2006, 2011; Grisi *et al*., 2007; Laurentin, 2009). These DNA molecular markers, when closely linked to genes of interest can also be used to select for desirable allele/s in marker-assisted breeding programs (Okii *et al*., 2014).

2.5 Use of molecular markers to characterize common bean germplasm

Molecular markers permit significant estimation of genetic diversity directly at the DNA level, reducing the interference of environmental variations. Genetic diversity in common bean have been studied using different molecular markers such as allozymes (Singh *et al*., 1991; Santalla *et al*., 2002), Amplified Fragment Length Polymorphism, AFLP (Lioi *et al*., 2005; Svetleva *et al*., 2006), Random Amplified Polymorphism, RAPD (Ocampo *et al*., 2005, Martins *et al*., 2006; Marotti *et al*., 2007), Restriction Fragment Length Polymorphism, RFLP (Nodari *et al*., 1992), Inter Simple Sequence Repeats, ISSR

(Svetleva *et al*., 2006; Marotti *et al*., 2007), Simple Sequence Repeats, SSR (Asfaw *et al*., 2009; Okii *et al*., 2014) and gene-based markers such as peroxidase gene (Nemli *et al*, 2014). The utility of SSR markers is highly desirable due to their abundant distribution and high polymorphism in the whole genome, their power to distinguish between closely related genotypes (Khaidizar *et al*., 2012), and because they are easily reproducible, multi-allelic and codominant genetic marker system (Saghai Maroof *et al*., 1994). Considering higher level of sequence diversity in peroxidase gene sequences among plant genotypes (Zhang *et al*., 2001), peroxidase markers can also be used to efficiently assess the genetic relationship among the common bean accessions.

2.5.1 Use of SSR markers

Simple sequence repeat markers or microsatellites are tracts of repetitive DNA in which certain DNA motifs (ranging in length from 2 – 5 base pairs) are repeated, typically 5-50 times (Turnpenny and Ellard, 2005). Microsatellites occur at thousands of locations in the genome and they are notable for their high mutation rate and high diversity in the population. The polymorphism is based on the number of repetitions of a short DNA motif at a given locus (Buso *et al*., 2006). These markers have been ideal for genetic mapping and characterizing genetic diversity in crop species due to the high mutation rates and resulting variability at SSR loci (Matus and Hayes, 2002). SSR markers have several advantages for genetic fingerprinting, including highly polymorphic and reproducible, enormous extent of allelic diversity, frequently co-dominant, strong discriminatory power specific PCR-based assay, randomly and widely distributed in the genome (Powell *et al*., 1996; Pejic *et al*., 1998). Simple sequence repeat (SSR) marker

analysis has been successfully used to evaluate genetic diversity in dry bean genotypes from several Countries including Italy, Bulgaria, Nicaragua, Slovenia, Uganda and Ethiopia (Marotti *et al*. 2007; Svetleva *et al*. 2006; Gomez *et al*. 2005; Maras *et al*. 2008; Okii *et al*., 2014; Fisseha *et al*., 2016). These markers have also been used in common bean to construct a PCR-based genetic map (Perseguini *et al*., 2011) and to evaluate intra-specific diversity within the genus and to fingerprint genetic diversity of common bean (Maras *et al*., 2008). There is limited information on molecular characterization of common bean germplasm grown in Kenya using SSR markers.

2.5.2 Use of peroxidase gene-based markers

Peroxidases make up highly class of enzymes in animals, plants and micro-organisms which are conserved (Zhang *et al*., 2001). They are proteins that contain heme and are capable of oxidizing compounds in the presence of oxygen (O_2) or hydrogen peroxide (H_2O_2) and consist of three highly conserved motifs namely distal, central and proximal domains (Hiraga *et al*., 2001). Based on their catalytic and structural properties, plant peroxidases are grouped into two major categories. The intracellular peroxidases are related to bacterial peroxidases (class I) (Gulsen *et al*., 2010), the second group of these enzymes (class III) aim the secretary pathway (Welinder, 1992). Class III peroxidases are plant-specific heme oxidoreductases made up of c. 300 amino acid residues. More advanced plant species have more peroxidase isoenzymes, encoded by multigene families (Yoshida *et al*., 2003); *Oryza sativa* has 138 (Passardi *et al*., 2004), 73 are found in *Arabidopsis thaliana* (Welinder *et al*., 2002) and *Populous trichocarpa* has 93 (Ren *et*

al., 2014). This implies that class III peroxidases have various functional properties, and there is a relationship between those function and their catalytic properties.

Plant peroxidases serve vital roles in various interactions related to stress tolerance. They also catalyze a wide range of physiological processes such as plant defense, insect tolerance, auxin catabolism, salt tolerance, lignin biosynthesis, cell wall proteins manufacture through deposition of callose, tissue suberization and plant senescence (Gulsen *et al*., 2010; Passardi *et al*., 2005). They act by aiding the deposition of macromolecules on the surface of the cell to strengthen plant tissues, thereby restricting expansion of the cell and invasion of pathogens (Almagro, 2008; Hiraga, 2001). Peroxidases also play a role in scavenging reactive oxygen species (ROS), which are partially reduced forms of atmospheric oxygen. They need to be reduced because they are can cause oxidative damage to the plant because of their high reactivity (Vicuna, 2005).

Plant peroxidases possess highly conserved domains across different plant species (Collard and Mackill, 2009), and the conserved DNA sequences within the genes can reveal how they function. DNA regions that are conserved and share the same priming site may be spread across the genomes of various germplasms in different ways; therefore it is possible to detect polymorphisms within species (Poczai *et al*., 2013). Degenerate oligonucleotide primers can be employed in amplifying DNA sequences that code for peroxidases from plants using these conserved domains (Collard and Mackill, 2009). Peroxidase-specific markers have previously been used in detecting polymorphisms of peroxidase genes among accessions of different plant species, including watermelon

(*Citrullus lanatus* (Thunb.), apple (*Malus domestica* Borkh.), wheat (*Triticum* specie.), citrus and beans (*Phaseolus vulgaris* L.); therefore, these markers can be utilized in studying evolutionary relationships and genetic diversity on an inter- and intra-specific level (Gulsen *et al*., 2010; Ceylan 2010; Ocal, 2014; Uzun *et al*., 2014; Nemli *et al*., 2014).

CHAPTER THREE

3.0 MATERIALS AND METHODS

3.1 Plant materials

Seeds of selected common bean germplasms were obtained from farmers in Central, Nyanza, Eastern, Western and Rift Valley regions of Kenya. A total of 46 cultivars representing a wide spectrum of phenotypic variability were selected for the present study (Figure 1; Table 1). After collection, they were placed in plastic bags and transported to the laboratory. Three bean seeds per genotype were planted in 2 liter plastic pots filled with a mixture of compost and loamy soil and placed on the bench in the glasshouse and watered once until germination, then watered once in two days till when ready for DNA extraction. After two weeks of planting, newly opened fresh young leaves of 40 genotypes were collected for DNA extraction. Seeds of 6 genotypes (codes B5, B6, B8, B9, B31 and B32) did not germinate and therefore molecular characterization was not carried out for these genotypes.

Figure 1**:** Phenotypic diversity of common bean genotypes collected in Kenya

Table 1: Local names, region of acquisition and characteristics of common bean genotypes used in the present study

No.	Code	Local name	Geographical region	100-seed weight $^{-1}$ (g)	Seed size	Seed colour
1	B1	Kikuyu1	Central (Muranga)	28.8	Medium	Brown
2	B2	Kikuyu2	Central (Maragua)	24.5	Small	Brown
3	B3	Kikuyu3	Central (Kiambu)	30.2	Medium	Cream with brown specks
4	B4	Amini	Nyanza (Keroka)	42.2	Large	Brown
5	B5	Yellow kidney	Nyanza (Kisii)	27.3	Medium	Yellow
6	B6	Makueni1	Eastern (Makueni)	31.1	Medium	Brown
7	B7	Makueni2	Eastern (Makueni)	47.8	Large	Light brown
8	B8	GLP 24	Eastern (Makueni)	49.0	Large	Dark brown
9	B9	Red haricot	Eastern (Makueni)	33.0	Medium	Brown
10	B10	Sugar1	Western (Kakamega)	39.6	Medium	Cream with red specks
11	B11	Makueni7	Eastern (Makueni)	32.2	Medium	Creamy
12	B12	Kiboko1	Eastern (Makueni)	29.8	Medium	Brown
13	B13	Masaku	Eastern (Machakos)	39.4	Medium	Cream with red specks
14	B14	Kibwezi1	Eastern (Makueni)	19.9	Small	Cream with red strips
15	B15	Rose coco	Eastern (Embu, Mbeere, Meru)	54.9	Large	Brown
16	B16	Royoo	Nyanza (Kisii)	23.1	Small	Dark brown
17	B17	Rose coco	Rift Valley (Cherangani)	33.6	Medium	Red
18	B18	Super Rose Coco	Eastern (Embu)	42.0	Large	Brown
19	B19	Mwitemania	Western (Kakamega)	46.7	Large	Cream with brown specks
20	B20	GLP 2	Central (Kiambu)	34.3	Medium	Red
21	B21	Unknown3	Central (Limuru)	33.9	Medium	Cream with brown specks
22	B22	GLP 24	Rift Valley (Kitale)	48.2	Large	Red
23	B23	New Rose Coco	Central (Kirinyaga)	33.5	Medium	Maroon
24	B24	Enyoro	Nyanza (Nyamira)	21.9	Small	Dark brown
25	B25	Nyaibu	Nyanza (Keumbu)	55.8	Large	Black
26	B26	Unknown1	Rift Valley (Njoro)	24.1	Small	Red
27	B27	Zaire	Nyanza (Mosocho)	22.9	Small	Light brown

28	B28	Bunda	Nyanza (Rongo)	65.5	Large	Black
29	B29	Unknown4	Western (Vihiga)	22.9	Small	Red
30	B30	Unknown7	Western (Vihiga)	19.0	Small	Maroon
31	B31	Wairimu	Central (Mwea)	33.3	Medium	Light brown
32	B32	Unknown11	Central (Mwea)	46.7	Large	Brown
33	B33	Kakamega1	Western (Kakamega)	31.8	Medium	Maroon
34	B34	Kakamega2	Western (Kakamega)	27.7	Medium	Maroon
35	B35	Kakamega3	Western (Kakamega)	35.8	Medium	Brown with red strips
36	B36	Morogi	Nyanza (Kisii)	20.0	Small	Black
37	B37	Canadian wonder	Eastern (Embu, Meru, Mbeere)	53.9	Large	Red
38	B38	Kisii3	Nyanza (Kisii)	31.4	Medium	Brown with red strips
39	B39	Morogi	Nyanza (Kisii)	20.6	Small	Black
40	B40	Migori1	Nyanza (Migori)	43.2	Large	Brown
41	B41	Raila	Nyanza (Migori)	32.4	Medium	Red
42	B42	GCP 004	Eastern (Machakos)	43.2	Large	Brown with white specks
43	B43	Yellow kidney	Eastern (Mbeere, Machakos)	40.8	Large	Yellow
44	B44	KAT 56	Eastern (Machakos)	48.4	Large	Red
45	B45	KAT B9	Eastern (Machakos)	48.2	Large	Red
46	B46	KAT 69	Eastern (Machakos)	40.1	Large	Red

Seed size = 100-seed weight^{-1}; Small = <25 g, Medium = 25 - 40 g, Large = >40 g.

3.2 DNA extraction

Genomic DNA was extracted using cetyltrimethylammonium bromide (CTAB) protocol as described by Choudhary *et al*. (2008) with some modifications. Excluding the use of liquid nitrogen, the leaves were ground using a pre-chilled mortar and pestle to form homogenate in 700 µl of pre-warmed CTAB extraction buffer (2% CTAB, 5 M NaCl, 0.5 M EDTA, 1M Tris HCl pH 4.0, 1% polyvinylpyrocarbonate (PVP)). The homogenate was transferred to sterile Eppendorf tubes using a spatula and 200 µl of β-mecaptoethanol was added and incubated at 65 ^{o}C for 15 minutes with gentle mixing after every 5 minutes. Afterwards, the tubes were allowed to cool at room temperature for 5 minutes. The mixture was centrifuged using a bench top centrifuge at 12,000 rpm for 5 minutes, and the supernatant transferred to sterile Eppendorf tubes. A volume of 250 µl of chloroform:isoamyl alcohol (24:1) was added to each tube and mixed gently by inverting the tubes repeatedly for 5 minutes. The mixtures were centrifuged at 4 °C for 10 minutes at 8,000 rpm and 600 µl of upper aqueous layer was transferred to sterile Eppendorf tubes. The DNA was then precipitated by adding of 500 µl of ice-cold isopropanol and the tubes left at -20 ^{o}C for 30 minutes. To pellet the DNA, the tubes were centrifuged at 8000 rpm for 5 minutes at room temperature (25 °C). The supernatant was decanted and the DNA pellet was washed with 1 ml of 70% ethanol. Ethanol was carefully removed and the DNA pellets were dried by leaving the tubes open for 15 minutes in a laminar flow hood. The DNA was dissolved in 100 µl of TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA), 2 µl of RNaseA (10 mg/ml) was added and incubated at 37 ^{o}C for 1 hour.

3.3 DNA quantification and quality checking

The quality of the DNA was assessed using 0.8% agarose gel electrophoresis. The gel was prepared by melting 0.8% (w/v) agarose in 100ml of TAE buffer until it boiled. The agarose was then allowed to cool to about 50 ^{o}C (hand hot) and gel red added to a final concentration of 0.5 µg/ml. The gel was poured onto the gel tray fitted with combs and allowed to solidify. After the gel has solidified, it was placed in the electrophoretic tank, a buffer was loaded to slightly cover the gel, and the comb gently removed. After removing the comb, 10µl of the sample was carefully loaded with the first and last wells having a DNA ladder. The gel electrophoresis was run for 65 minutes at 60 volts. The bands were visualized under ultraviolet transilluminator, and images taken using a mobile device.

The DNA quantity of each sample was determined by spectrophotometry. The extracted DNA samples were thawed and 5 µl of each added to 995 µl of distilled water in a micro-centrifuge tube. It was properly mixed before reading the absorbance at 260 – 280 nm. This process works by measuring the absorption of light at 260-280 nm. The stock DNA was diluted to a final concentration of 5 ngμL^{-1} for PCR.

$$\text{DNA concentration } (\mu g/ml) = \frac{OD_{260} \text{ X } 50 \text{ (dilution factor) X } 50\ \mu g/ml}{1000}$$

(Sambrook *et al*., 1989)

The ratio A_{260}/A_{280} was used to determine the purity of the DNA samples.

3.4 Identification and selection of SSR and peroxidase gene (POX) primers

Five simple sequence repeat markers (Table 2) were used for characterizing common bean genotypes. Primer selection was based on previous studies which showed high amplification patterns and polymorphic characteristics (Blair *et al*., 2006; Isemura *et al*., 2012). Among the markers used, two (Bmd2 and Bmd17) were specific for common beans (Blair *et al*., 2006) and 3 (Vm71, Vm74 and Vm94) were specific for cowpea (Isemura *et al*., 2012). The choice of cowpea markers for this study is because cowpea and beans have similar genetic information and therefore share a considerable level of sequence conservation within the primer regions flanking the microsatellite loci. Their names, repeat types, predicted fragment length, base sequences and melting temperatures are presented in Table 2.

Peroxidase gene primers designed previously from peroxidase cDNA sequences of *Arabidopsis* and rice (Welinder *et al*., 2002; Gulsen *et al*., 2007; and Nemli *et al*., 2014) were synthesized by Inqaba Biotech, Ltd (South Africa). To target peroxidase gene sequences amplification from extracted DNA of common bean genotypes, five POX primers were used. The sequences of the marker pairs are presented in Table 2.

Table 2**:** Simple sequence repeat (SSR) and peroxidase gene (POX) markers used in the molecular diversity studies of Kenyan common bean genotypes

Marker	**Primer name**	**Repeat**	**Primer sequence (5'-3')**		**Size range (bp)**	**Tm ^{o}C**	**Reference**
			Forward	**Reverse**			
SSR	Vm71	(AG012(AAAG)3	TCGTGGCAGAGAATCAAAGACAC	TGGGTGGAGAAAACAAACC	100 - 250	58	Isemura *et al*., 2012
	Vm74	(AC)8(A)5	CTGCTACACCTTCCATCATTC	CCTTTGCGTTGTGGTGGTTT	100 – 400	55	Isemura *et al*., 2012
	Vm94	(CA)12(AAAG)3	TCGAACTTTGGCTTGAGG	TGTCGTTTTGTCCCCCATTA	100 - 350	61	Isemura *et al*., 2012
	Bmd2	(CGG)8	AGCGACAGCAAGAGAACCTC	CAACGTTTTGTCCCCCATTA	50 - 400	50	Blair *et al*., 2006
	Bmd17	(CGCCAC)6	GTTAGATCCCGCCCAATAGTC	CAACAAACGGAAGGGCGTGGTTT	100 - 900	46	Blair *et al*., 2006
POX	PM55		TTGTAGATTCTCGCTCGGAA	CTTGGCATAATTGTTATTTGGT	150 - 800	53	Nemli *et al*., 2014
	POX1		CTCGACCTACAAGGAC	ATGTAGGCGCTGGTGA	100 - 800	55	Nemli *et al*., 2014
	POX8b		CACCATCAAGAGCGTCATAAC	TTGCTAGAGCGAGCTGG	100 - 200	52	Nemli *et al*., 2014
	POX11		CCTTCTTCTTGCCATCTTGC	CATATCGCTCCACGACCTTT	150 - 750	50	Nemli *et al*., 2014
	POX12b		CTCTCTCCTGGGGGTTCTATGC	GCGAGCGTGGTGATGTC	100 - 750	53	Nemli *et al*., 2014

3.5 Polymerase chain reaction (PCR)

SSR- and POX-PCR reactions were done in a total volume of 20 µl, consisting of 5 µl 1× GoTaq Mix (Promega Corporation, Madison, USA), 1 µl of each of the forward and reverse SSR or POX markers (10 µM), 1 µl genomic DNA (20 ng), 12 µl nuclease-free water. All the PCR reactions were carried out in 200 µl thin-walled PCR tubes. Amplifications for SSR markers were performed in a MJ Mini^{TM} Thermal Cycler (Bio-Rad, Singapore) as follows: initial denaturation at 94 °C for 5 minutes, followed by 30 cycles of denaturation at 94 °C for 30 seconds, annealing at 46 - 61 °C depending on the primer pair) for 60 seconds, extension at 72 °C for 2 minutes with a final extension at 72 °C for 7 minutes. The samples were then maintained at 10 °C.

Amplifications for POX markers were done in an MJ Mini^{TM} Thermal Cycler machine (Bio-Rad, Singapore) as follows: initial denaturation at 94°C for 5 minutes, for 30 seconds, followed by denaturation (30 seconds) at 94°C, annealing at 46 - 56°C (45 sec), extension (1 minute) at 72°C, then a final extension at 4°C for 7 minutes.

The PCR reactions for each of the SSR and POX primers were performed at least twice using DNA from independent extractions and only clear and reproducible bands were used in data evaluation.

3.6 Separation of amplified PCR products by agarose gel electrophoresis

The amplified products were separated on 2% agarose gel in 1X Tris-Acetate EDTA (TAE) buffer. Gels were run for 65 minutes at 60 V. Amplified products were stained

with ethidium bromide and viewed under UV transilluminator and images were gotten. Estimation of SSR allele/band sizes were assigned based on the migration amplicons/fragments through the gel in comparison to that of 100 bp DNA ladder (Bioneer, South Africa).

3.7 Scoring of allele and statistical analysis of data

The alleles/bands were scored as present (1) and absent (0) and were recorded in a data matrix table as discrete variables. The summary statistics including the observed number of alleles per locus (na), number of effective alleles (ne), gene diversity (h) and Shannon's information index (I) were determined using GenAlEx6 software (Peakall and Smouse, 2012). The polymorphism information content (PIC) value was determined for each SSR and POX locus following the equation $PIC = 1 - \sum (pi)^2$ as described by Botstein *et al*. (1980) (where *pi* is the population carrying the *i*th allele). A similarity matrix was generated using the Nei's genetic distance (Nei and Li, 1979). Similarity data and generation of a dendrogram were processed through the unweighted pair group method using arithmetic mean (UPGMA) cluster analysis conducted using POPGENE32 program. To assess further the genetic relationships of common bean accessions as individual plants, a principal component analysis (PCA) was conducted based on the SSR variation patterns using XLSTAT software program. Analysis of molecular variance (AMOVA) among and within populations was performed using GenAlEx 6.5 program.

3.8 Population structure of common bean germplasm using POX markers

Population genetic structure analysis was done with a clustering approach of a Bayesian model-base, clustering approach in the STRUCTURE version 2.3.4 program (Pritchard, 2000). An analysis of all 40 cultivars was done using the number of clusters (*K*) ranging from 1-10, and a burn-in period of 5,000 iterations with 50,000 replications of Markov Chain Monte Carlo (MCMC). Results were not significantly affected with longer burn-in periods. The runs showing the maximum posterior probability for each *K* value was used. The *ad hoc* statistic ΔK was used to estimate the total sub-populations, and to determine *K* (Evanno *et al*., 2005). PCA was carried out depending on the variation patterns of the POX gene, and a two dimensional representation of relationships across the 40 common bean genotypes using XLSTAT program was generated. AMOVA within and among populations was done with the GenAlEx (v6.5) software (Peakall and Smouse, 2012).

CHAPTER FOUR

4.0 RESULTS

4.1 Molecular characterization of common bean germplasm using SSR markers

4.1.1 Polymorphism and diversity parameters revealed by SSR markers

All forty germplasms were successfully amplified with the five microsatellite primer pairs. A total of 366 reproducible and scorable alleles (a band represents an allele) were amplified with the 5 SSR primer pairs among the 40 genotypes. The number of alleles produced by each SSR primer ranged from 3 (Vm71, Vm74 and Vm94) to 5 (Bmd17) with an average of 4.5 alleles per locus (Table 3). A sample amplification pattern of the primer Bmd2 is shown in Figure 2. An average of 73.3 alleles were amplified with the highest number of alleles amplified being observed for marker Bmd17. Least number of alleles (29) was amplified by marker Vm94. The number of observed alleles ranged from 1.60 for (Vm71, Vm74 and Vm94) to 2.0 for Bmd17 with a mean of 0.8069 (Table 3). The number of effective allele values ranged from 1.1875 to 1.5502 with a mean value of 1.3530. It was observed that marker Vm71 had the lowest values while marker Bmd2 had the highest value.

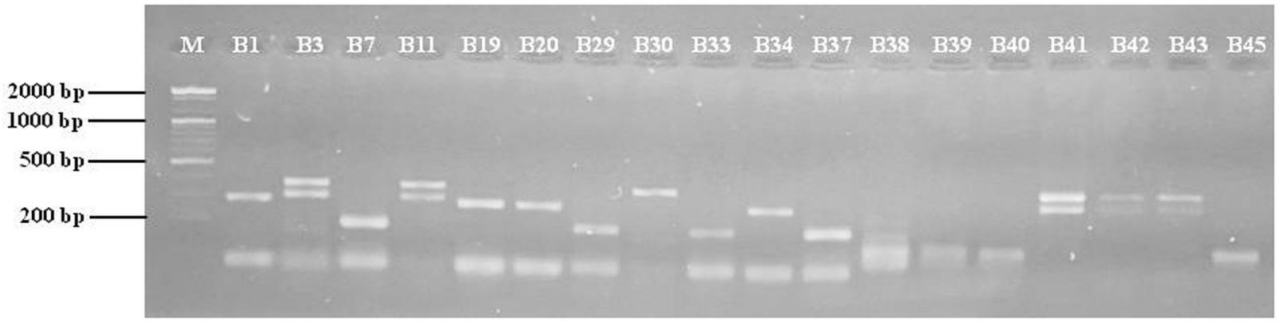

Figure 2**:** Electrophoretic SSR marker profile of 18 common bean genotypes generated by primer Bmd2. The 100 bp molecular weight marker is indicated in lane M (Bioneer, South Africa) and the codes indicate the common bean genotypes. More than one band indicates polymorphism

For all the genotypes the PIC values for the SSR loci ranged from 0.4818 for Vm94 to 0.7439 for Vm71, with an average PIC value of 0.5958 (Table 3). The mean Nei's (1973) gene diversity (h) of the loci producing polymorphic bands in the 40 common bean genotypes ranged from 0.1215 to 0.3212 with a mean value of 0.2129. Marker Vm71 had the lowest value while marker Bmd2 had the highest value of 0.3212. This observation was also confirmed by Shannon's information index at locus Bmd17 (p = 0.4811), which had the highest value as compared to the lowest value of p = 0.1988 at locus Vm71. The observed heterozygosity (Ho) calculated for each primer ranged from 0.5536 (Vm71) to 0.7750 (Bmd17) with a mean of 0.6537 (Table 3).

Table 3**:** Diversity parameters for 5 SSR loci used to analyze genetic diversity of common bean germplasm in Kenya

Locus	**No. of alleles**	**Total no. of alleles**	**na***	**ne***	**h***	**Ho**	**PIC**	**I***
Vm71	3	66	1.6000	1.1875	0.1215	0.5800	0.4918	0.1988
Vm74	3	62	1.6000	1.2332	0.1445	0.6250	0.5547	0.2282
Vm94	3	52	1.6000	1.2568	0.1715	0.5536	0.4818	0.2729
Bmd2	4	91	1.8000	1.5502	0.3212	0.7350	0.7069	0.4718
Bmd17	5	95	2.0000	1.5375	0.3150	0.7750	0.7439	0.4811
Mean	**3.6**	**73.2**	**1.7200**	**1.3530**	**0.2149**	**0.6537**	**0.5958**	**0.3306**
St. Dev	**0.9**	**18.8**	**0.4583**	**0.3497**	**0.1878**	**0.0969**	**0.1222**	**0.2653**

* na = Observed number of alleles * ne = Effective number of alleles [Kimura and Crow (1964)] * h = Nei's (1973) gene diversity* I = Shannon's Information index [Lewontin (1972)]; Ho = observed heterozygosity

4.1.2 Similarity coefficient among the 40 common bean genotypes

Genetic similarity matrix among all studied genotypes was obtained from fragments amplified with 5 SSR markers using Jaccard similarity coefficients. The similarity coefficient among the 40 bean genotypes ranged from 0.15 to 1.0 with an average of 0.54, which suggested that there was an abundant genetic diversity in the common bean accessions, grown in Kenya. The highest similarity coefficient was 1.00 between code B19 and B21 (Table 4). It indicated the two cultivars were genetically similar and no genetic distance (GD). However, the genetic similarities (0.15) between two genotypes B22 and B37 were the smallest. The low values of genetic similarity obtained indicated a high genetic diversity among the common bean genotypes.

Table 4**:** Pairwise genetic similarity index among 40 common bean genotypes based on SSR data

	B1	**B2**	**B3**	**B4**	**B7**	**B10**	**B11**	**B12**	**B13**	**B14**	**B15**	**B16**	**B17**	**B18**	**B19**	**B20**	**B21**	**B22**	**B23**	**B24**
B1	1	0.67	0.43	0.58	0.33	0.42	0.42	0.31	0.50	0.58	0.39	0.46	0.50	0.46	0.58	0.70	0.58	0.38	0.46	0.50
B2		1	0.40	0.43	0.31	0.50	0.39	0.29	0.46	0.43	0.36	0.31	0.46	0.42	0.54	0.50	0.54	0.20	0.42	0.36
B3			1	0.79	0.57	0.77	0.77	0.64	0.50	0.79	0.71	0.47	0.85	0.69	0.79	0.64	0.79	0.31	0.47	0.71
B4				1	0.50	0.69	0.69	0.57	0.43	0.85	0.64	0.62	0.77	0.62	0.85	0.83	0.85	0.33	0.50	0.77
B7					1	0.58	0.36	0.73	0.42	0.50	0.67	0.39	0.67	0.64	0.62	0.46	0.62	0.18	0.50	0.67
B10						1	0.67	0.54	0.50	0.69	0.62	0.46	0.75	0.73	0.69	0.54	0.69	0.27	0.36	0.62
B11							1	0.43	0.39	0.69	0.50	0.36	0.62	0.46	0.57	0.67	0.57	0.27	0.27	0.62
B12								1	0.50	0.47	0.62	0.46	0.75	0.58	0.69	0.54	0.69	0.17	0.46	0.62
B13									1	0.43	0.58	0.55	0.58	0.55	0.54	0.50	0.54	0.50	0.31	0.46
B14										1	0.53	0.62	0.64	0.62	0.71	0.69	0.71	0.33	0.40	0.64
B15											1	0.54	0.83	0.67	0.77	0.62	0.77	0.36	0.54	0.83
B16												1	0.54	0.50	0.62	0.58	0.62	0.44	0.39	0.54
B17													1	0.82	0.92	0.75	0.92	0.36	0.54	0.83
B18														1	0.75	0.58	0.75	0.44	0.39	0.67
B19															1	0.83	1.00	0.33	0.62	0.77
B20																1	0.83	0.40	0.46	0.75
B21																	1	0.33	0.62	0.77
B22																		1	0.18	0.36
B23																			1	0.54
B24																				1

Table 4: Continued

	B25	B26	B27	B28	B29	B30	B33	B34	B35	B36	B37	B38	B39	B40	B41	B42	B43	B44	B45	B46
B1	0.46	0.31	0.57	0.50	0.46	0.50	0.33	0.27	0.39	0.39	0.39	0.31	0.46	0.30	0.40	0.50	0.50	0.40	0.67	0.56
B2	0.42	0.29	0.33	0.46	0.42	0.36	0.18	0.25	0.36	0.36	0.36	0.29	0.42	0.27	0.25	0.46	0.46	0.50	0.46	0.50
B3	0.57	0.64	0.31	0.62	0.69	0.60	0.39	0.54	0.71	0.85	0.60	0.64	0.69	0.46	0.43	0.71	0.40	0.54	0.40	0.54
B4	0.50	0.57	0.33	0.54	0.62	0.77	0.42	0.46	0.64	0.77	0.53	0.57	0.62	0.50	0.46	0.64	0.43	0.46	0.54	0.46
B7	0.64	0.73	0.18	0.55	0.50	0.54	0.40	0.46	0.82	0.67	0.54	0.58	0.8	0.50	0.60	0.82	0.55	0.46	0.42	0.33
B10	0.58	0.54	0.27	0.64	0.58	0.50	0.36	0.42	0.75	0.75	0.62	0.67	0.73	0.60	0.42	0.75	0.39	0.70	0.39	0.42
B11	0.36	0.43	0.27	0.39	0.46	0.50	0.25	0.31	0.50	0.62	0.40	0.43	0.46	0.46	0.42	0.50	0.20	0.42	0.29	0.42
B12	0.46	0.67	0.17	0.50	0.58	0.50	0.36	0.55	0.62	0.75	0.62	0.67	0.73	0.46	0.42	0.62	0.39	0.42	0.29	0.42
B13	0.55	0.50	0.50	0.60	0.55	0.46	0.30	0.36	0.46	0.46	0.46	0.64	0.55	0.40	0.50	0.58	0.46	0.50	0.46	0.67
B14	0.50	0.47	0.33	0.54	0.50	0.64	0.42	0.36	0.64	0.64	0.53	0.57	0.62	0.50	0.46	0.64	0.43	0.46	0.54	0.36
B15	0.82	0.91	0.36	0.73	0.82	0.83	0.46	0.64	0.69	0.69	0.47	0.62	0.67	0.55	0.64	0.83	0.58	0.64	0.58	0.64
B16	0.50	0.46	0.30	0.55	0.50	0.67	0.56	0.46	0.43	0.43	0.43	0.58	0.50	0.50	0.46	0.54	0.42	0.46	0.55	0.33
B17	0.67	0.75	0.36	0.73	0.82	0.69	0.46	0.64	0.69	0.83	0.57	0.62	0.82	0.55	0.50	0.83	0.46	0.64	0.46	0.64
B18	0.64	0.58	0.44	0.70	0.64	0.54	0.40	0.46	0.67	0.67	0.54	0.58	0.80	0.50	0.46	0.82	0.42	0.60	0.42	0.46
B19	0.62	0.69	0.33	0.67	0.75	0.77	0.42	0.58	0.64	0.77	0.53	0.57	0.75	0.50	0.46	0.77	0.54	0.58	0.54	0.58
B20	0.46	0.54	0.40	0.50	0.58	0.75	0.36	0.42	0.50	0.62	0.40	0.43	0.58	0.46	0.55	0.62	0.39	0.42	0.50	0.55
B21	0.62	0.69	0.33	0.67	0.75	0.77	0.42	0.58	0.64	0.77	0.53	0.57	0.75	0.50	0.46	0.77	0.54	0.58	0.54	0.58
B22	0.30	0.27	0.60	0.33	0.30	0.36	0.29	0.22	0.25	0.25	0.15	0.27	0.30	0.25	0.38	0.36	0.20	0.22	0.33	0.38
B23	0.50	0.58	0.18	0.42	0.50	0.54	0.27	0.46	0.43	0.43	0.33	0.27	0.39	0.15	0.23	0.54	0.55	0.33	0.55	0.46
B24	0.67	0.75	0.36	0.58	0.67	0.83	0.46	0.50	0.69	0.69	0.47	0.50	0.67	0.55	0.64	0.83	0.46	0.50	0.58	0.50
B25	1	0.73	0.44	0.89	0.80	0.67	0.56	0.60	0.67	0.54	0.54	0.58	0.64	0.50	0.60	0.82	0.70	0.78	0.70	0.60
B26		1	0.27	0.64	0.73	0.75	0.36	0.55	0.62	0.62	0.40	0.54	0.58	0.46	0.55	0.75	0.5	0.55	0.50	0.55
B27			1	0.50	0.44	0.36	0.29	0.22	0.25	0.25	0.25	0.27	0.30	0.25	0.38	0.36	0.33	0.38	0.50	0.57
B28				1	0.89	0.58	0.63	0.67	0.58	0.58	0.58	0.64	0.70	0.56	0.50	0.73	0.60	0.88	0.60	0.67
B29					1	0.67	0.56	0.78	0.54	0.67	0.54	0.58	0.64	0.50	0.46	0.67	0.55	0.78	0.55	0.78
B30						1	0.46	0.50	0.57	0.57	0.38	0.50	0.54	0.55	0.64	0.69	0.58	0.50	0.73	0.50
B33							1	0.71	0.46	0.46	0.46	0.50	0.56	0.57	0.50	0.46	0.44	0.50	0.44	0.33
B34								1	0.50	0.64	0.50	0.55	0.60	0.44	0.40	0.50	0.50	0.56	0.36	0.56

B35									1	0.83	0.69	0.75	0.82	0.55	0.64	0.83	0.58	0.50	0.46	0.39
B36										1	0.69	0.75	0.82	0.55	0.50	0.69	0.46	0.50	0.36	0.50
B37											1	0.75	0.67	0.42	0.39	0.57	0.46	0.50	0.36	0.39
B38												1	0.73	0.60	0.55	0.62	0.50	0.55	0.39	0.42
B39													1	0.67	0.60	0.82	0.55	0.60	0.42	0.46
B40														1	0.63	0.55	0.40	0.63	0.40	0.30
B41															1	0.64	0.50	0.40	0.50	0.40
B42																1	0.58	0.64	0.58	0.50
B43																	1	0.50	0.78	0.50
B44																		1	0.50	0.56
B45																			1	0.50
B46																				1

4.1.3 Genetic relationships among common bean genotypes

The similarity coefficients matrix was used for UPGMA cluster analysis. The dendrogram constructed based on genetic similarities between genotypes showed that the 40 genotypes formed two major clusters (Figure 3, Table 5). The genotypes did not form specific groups according to geographic regions of acquisition/collection (Table 5). Cluster I was the largest and most diverse consisting of 38 genotypes from all the geographical regions. This cluster was further divided into 2 sub-clusters; sub-clusters A and B containing 7 and 31 genotypes, respectively. Both sub-clusters A and B were further divided into two groups each. Cluster II includes two genotypes from Rift Valley and Nyanza regions.

Table 5**:** Number of genotypes and their corresponding groupings based on 5 SSR loci data

Cluster	**Sub-cluster**	**Group**	**Genotypes**	**No. of genotypes in each cluster**	**Geographic region of collection**
1	A	I	B1, B2, B13, B46	4	Central and Eastern
		II	B23, B43, B45	3	Central and Eastern
	B	III	B3, B36, B19, B21, B17, B10, B7, B35, B42, B39, B18, B37, B38, B12, B11, B20, B4, B14, B24, B26, B15, B30, B25, B26, B29, B40, B41, B16, B44,	29	Central, Eastern, Rift Valley, Nyanza and Western
		IV	B33, B34	2	Western

2			B22, B27	2	Rift Valley and Nyanza

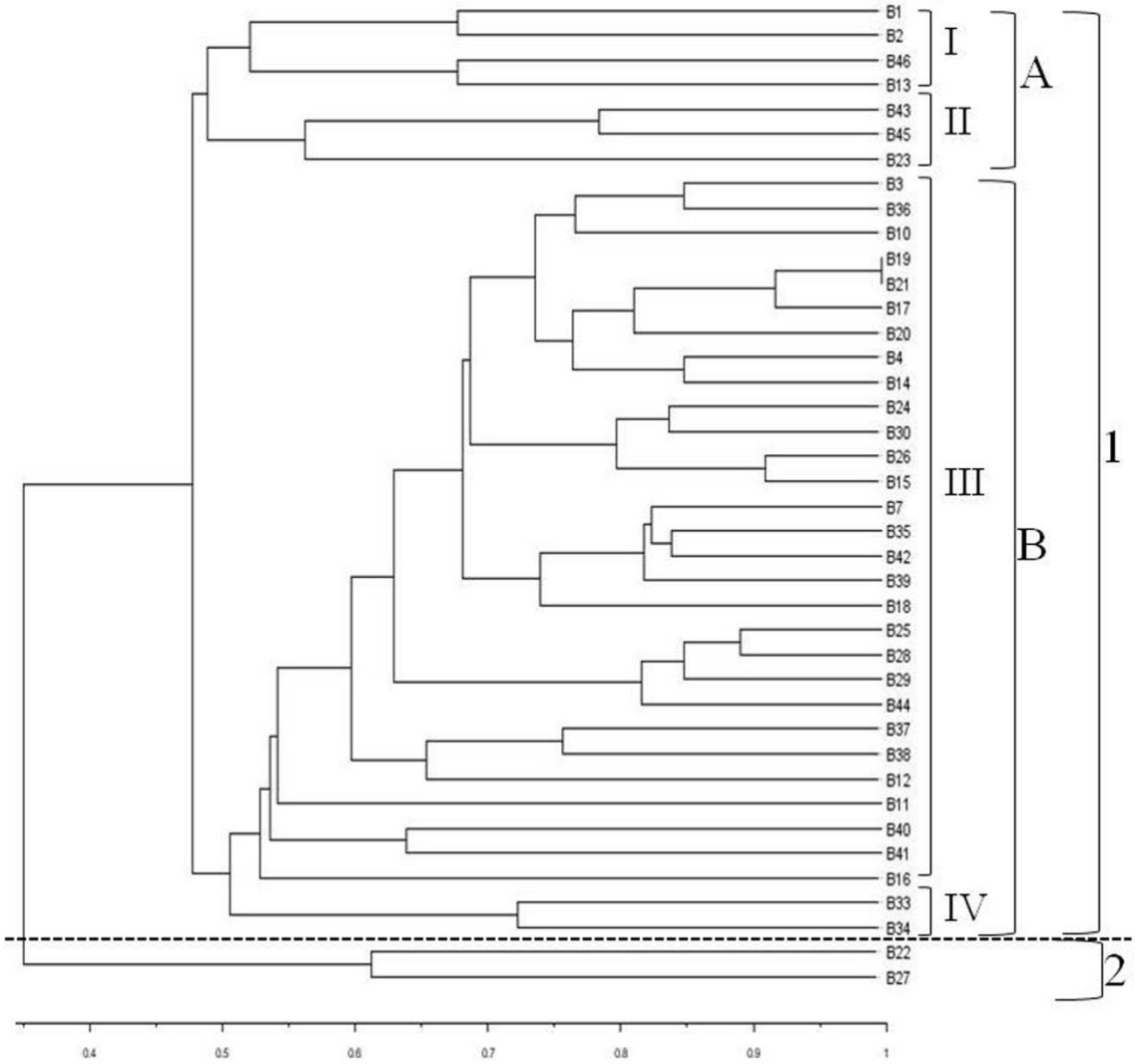

Figure 3**:** Dendrogram showing genetic relationship among 40 genotypes of common bean by SSR data using UPGMA

The genetic relationships among genotypes were also confirmed by scatter plot derived through principal component analysis (PCA). Principal component analysis based on allele frequencies generated using 5 SSR markers failed to detect significant grouping among the 40 common bean genotypes (Figure 4). The first and second principal components comprised 16.0% and 12.8% of the total variation (28.8%), respectively.

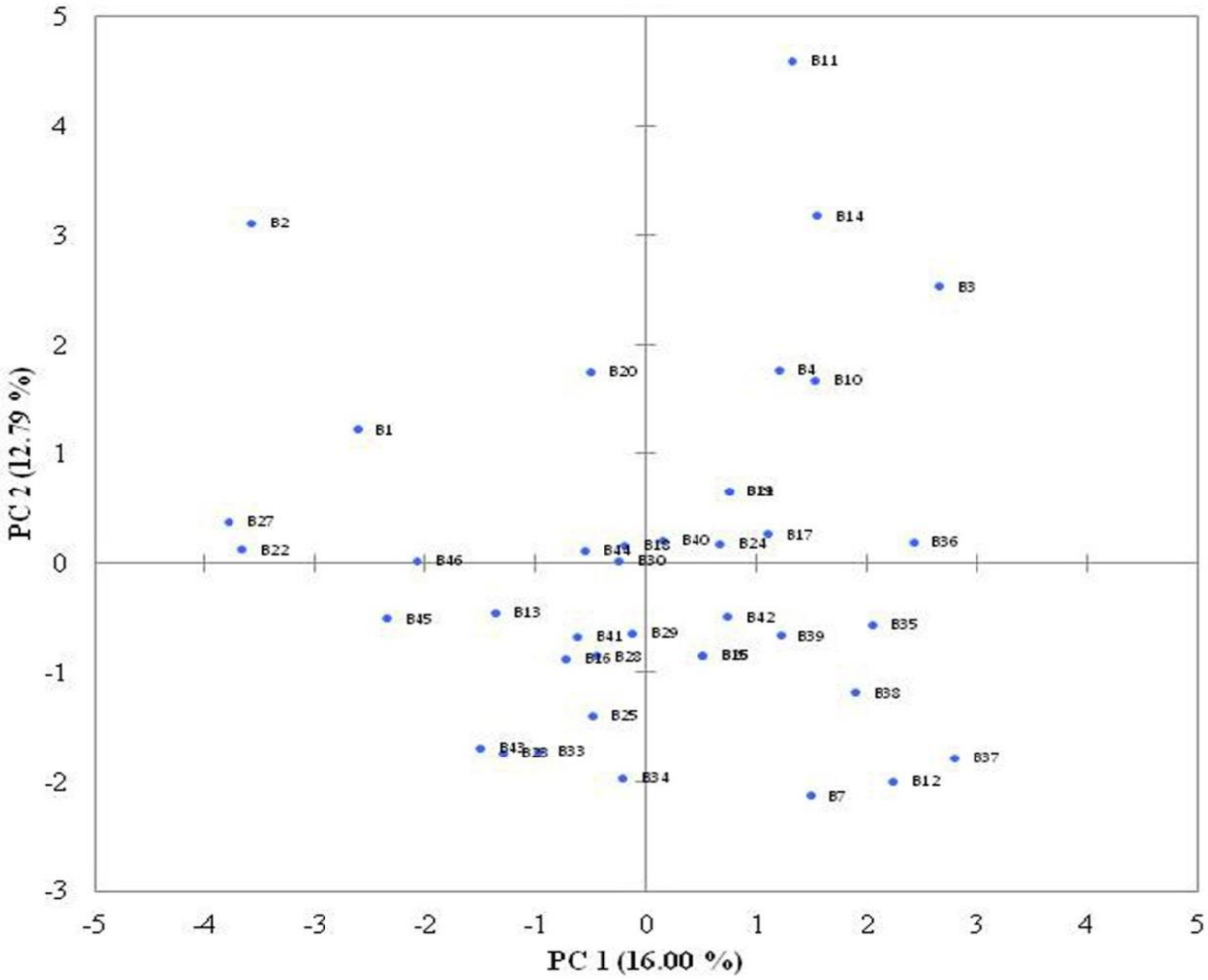

Figure 4**:** Principal component analysis (PCA) of 40 common bean genotypes based on 5 SSR data. PC 1 and PC 2 refer to the first and second principal components, respectively. The numbers in parentheses refer to the proportion of variance explained by the corresponding axes

4.1.4 Analysis of molecular variance (AMOVA)

Analysis of Molecular Variance (AMOVA) was used to estimate the partitioning of genetic variance among and within populations (Table 6). AMOVA results based on SSR data revealed that the vast majority of the total genetic variance was due to within population variation (87%) and only 13% of the genetic variation was among populations. Most of the genetic diversity of *P. vulgaris* resides within the populations. Both the diversity between and within populations was statistically significant at $p<0.01$.

Table 6**:** Analysis of molecular variance (AMOVA) of 40 common bean genotypes based on 5 SSR markers

Source of variation	**Df**	**SSD**	**MSD**	**VC**	**TVP (%)**	**P-Value**
Among populations	5	24.475	4.895	0.378	13%	< 0.01
Within populations	34	83.000	2.441	2.441	87%	< 0.01
Total	39	107.475		2.819	100%	

Degrees of freedom (df), sum of squares (SSD), mean squares (MSD), variance component (VC) and total variance percentage (TVP), $p< 0.01$

4.2 Genetic diversity and population structure of common bean germplasm using POX markers

4.2.1 Characteristics of the peroxidase gene (POX) primers

The values of the polymorphic information content (PIC) which reflects how effective the five POX markers are ranged from 0.6204 (for marker POX11) to 0.9110 (for marker POX8), having an average of 0.7677 (Table 7). The POX primers used in this study had PIC values above 0.5, thus were very polymorphic. The high average PIC values indicate that the POX markers had high discriminatory ability and were very informative; therefore, the loci were suitable for use in analysis of relationship and genetic diversity.

4.2.2 Relationship and genetic diversity among common bean cultivars

A statistical summary for various parameters of genetic diversity are shown in Table 3. All primers produced distinct/scorable fragments with high percentage of polymorphism. The peroxidase gene polymorphism yielded 624 reproducible and distinct/scorable fragments (90% showed polymorphism). The total fragments amplified varied from three (POX8) to nine (POX1) per locus, with an average of 7.2 (Table 7). An example of amplification patterns for POX1 is shown in Figure 5. An average of 124.8 fragments was amplified with the highest number of fragments being observed for POX1 (174) and the least number of fragments (68) was amplified using POX8. Number of observed alleles (N) varied from 1.8750 (for primer POX11) to 2.0000 (for primers PM55, POX12, POX1 and POX8) and an average of 1.9750. Values for number of alleles (Ne) were from 1.4907 (POX11) to 1.8179 (POX8), having a mean of 1.7066.

The values for Shannon's diversity (I) ranged from 0.4657 to 0.6327 with an average value of 0.5784. Markers POX8 and POX11 had the highest and lowest values, respectively (Table 7), suggesting that the cultivars used in the present study were highly diverse. The Nei's gene diversity (He) gotten varied from 0.3072 (POX11) to 0.4425 (POX8) with an average of 0.3972. The average observed heterozygosity (Ho) was 0.7945, with POX8 having the lowest value of 0.6667and POX12 having the highest value of 0.9150.

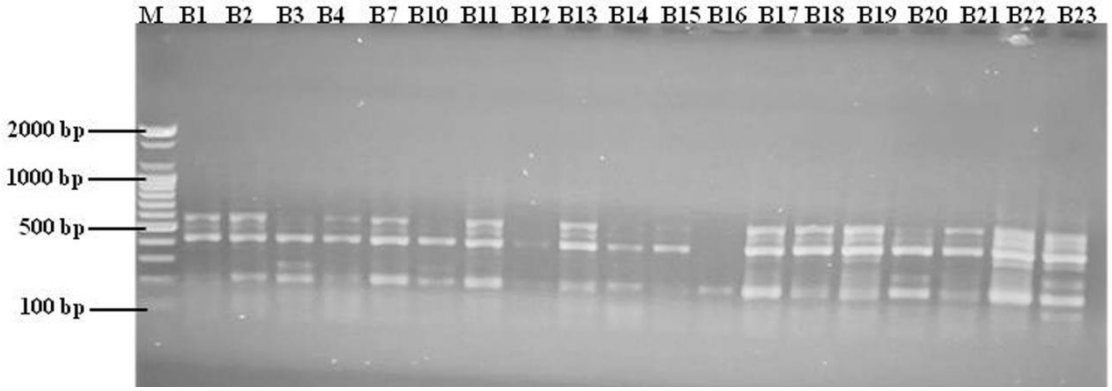

Figure 5**:** Peroxidase gene amplification pattern of 19 bean DNA samples using POX1 primer pair. The 100 bp molecular weight marker is indicated in lane M (Bioneer, South Africa) and the codes indicate the common bean genotypes as shown in Table 1. More than one band indicates polymorphism

Table 7**:** Diversity indices for each of the 5 peroxidase gene (POX) markers studied

Marker	**Number of alleles**	**Total number of alleles**	**Molecular parameters**					
			N	**Ne**	**I**	**He**	**Ho**	**PIC**
PM55	8	88	2.0000	1.6253	0.5439	0.3647	0.7733	0.7372
POX 1	9	174	2.0000	1.7853	0.6229	0.4331	0.7644	0.8419
POX 8	3	68	2.0000	1.8179	0.6327	0.4425	0.6667	0.9110
POX 11	8	138	1.8750	1.4907	0.4657	0.3072	0.8532	0.6204
POX 12	8	156	2.0000	1.8139	0.6268	0.4385	0.9150	0.7280
Mean	**7.20**	**124.80**	**1.9750**	**1.7066**	**0.5784**	**0.3972**	**0.7945**	**0.7677**
St. Dev	**2.39**	**45.14**	**0.1667**	**0.2689**	**0.1451**	**0.0135**	**0.0944**	**0.1121**

N = observed number of alleles, Ne = effective number of alleles [Kimura and Crow (1964)], I = Shannon's Information index [Lewontin (1972)], He = Nei's (1973) gene diversity, Ho = observed heterozygosity, PIC = polymorphic information content.

4.2.3 Genetic distance and identity

The average Nei's unbiased genetic distance (Nei, 1973) showed among and within the regions of sample collections is presented in Table 8. The analysis indicated the most distance for genotypes sampled between the regions of Western and Rift valley (0.6858) followed by Rift valley and Central (0.6723) (Table 8). The lowest genetic distance was between Eastern and Central (0.5840). Also, genetic identity among cultivars and regions ranged from 0.3143 to 0.4170 (Table 8). The highest identity (0.4170) was between Eastern and Central and lowest (0.3143) was between Western and Rift valley.

Table 8**:** Nei's genetic distance (shown below diagonal) and genetic similarity (shown above diagonal)

Regions	**Central**	**Nyanza**	**Eastern**	**Western**	**Rift Valley**
		Genetic identity			
Central	****	0.3763	0.4170	0.3693	0.3279
Nyanza	0.6240	****	0.3734	0.3646	0.3628
Eastern	0.5840	0.6290	****	0.3609	0.3344
Western	0.6310	0.6385	0.6393	****	0.3143
Rift Valley	0.6723	0.6376	0.6660	0.6858	****
		Genetic distance			

4.2.4 Clustering of populations

Cluster analysis was performed by generating a dendrogram using the UPGMA, resulting in the genotypes separating into two major groups (Figure 6). The genotypes did not form specific groups according to geographic regions of collection. Group 1 was the largest and the most diverse consisting of 37 genotypes from all the geographical regions. This group was divided into 2 clusters; clusters A and B containing 24 and 13 genotypes, respectively. Each of the clusters A and B were further divided into two sub-clusters. Group 2 was the smallest and contained only 3 genotypes namely B15, B22 and B33 from Eastern, Western and Rift valley regions of Kenya respectively.

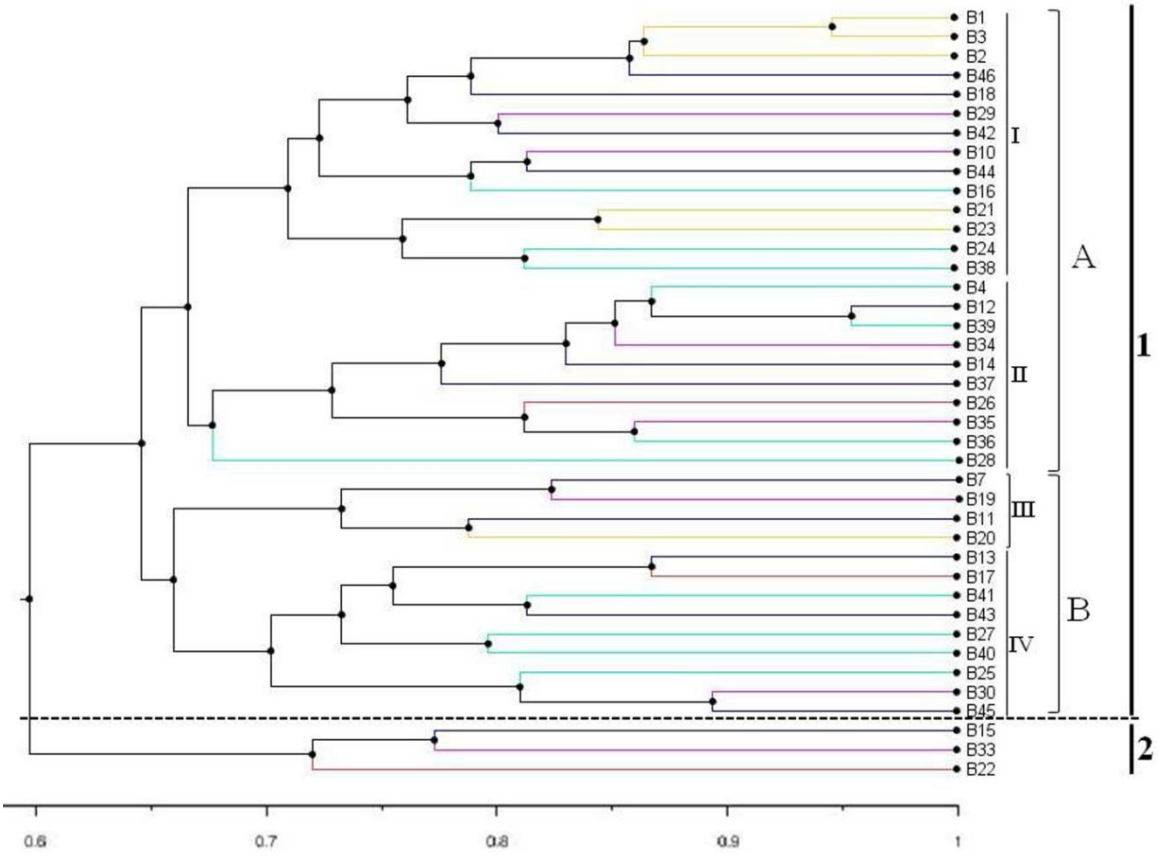

Figure 6**:** UPGMA dendrogram showing genetic relationship among 40 Kenyan common bean genotypes using 5 POX markers

4.2.5 Population genetic structure estimation

The 40 Kenyan bean genotypes were analyzed for population structure using Bayesian base method without any prior classification to know the highest populations (K). In the result, the most suitable *ΔK* value was analyzed using the STRUCTURE harvester from a range of 1 to 10. K = 7 showed the maximum value (Figure 7). Analysis using STRUCTURE indicated the presence of seven sub-populations (Pop1 – Pop7) by the model-based method in the Kenyan common bean germplasm (Figure 8) from the five geographical regions. The genotypes in all the sub-populations clearly shared more admixture memberships. The 40 genotypes, however, had no obvious geographical distribution characteristics.

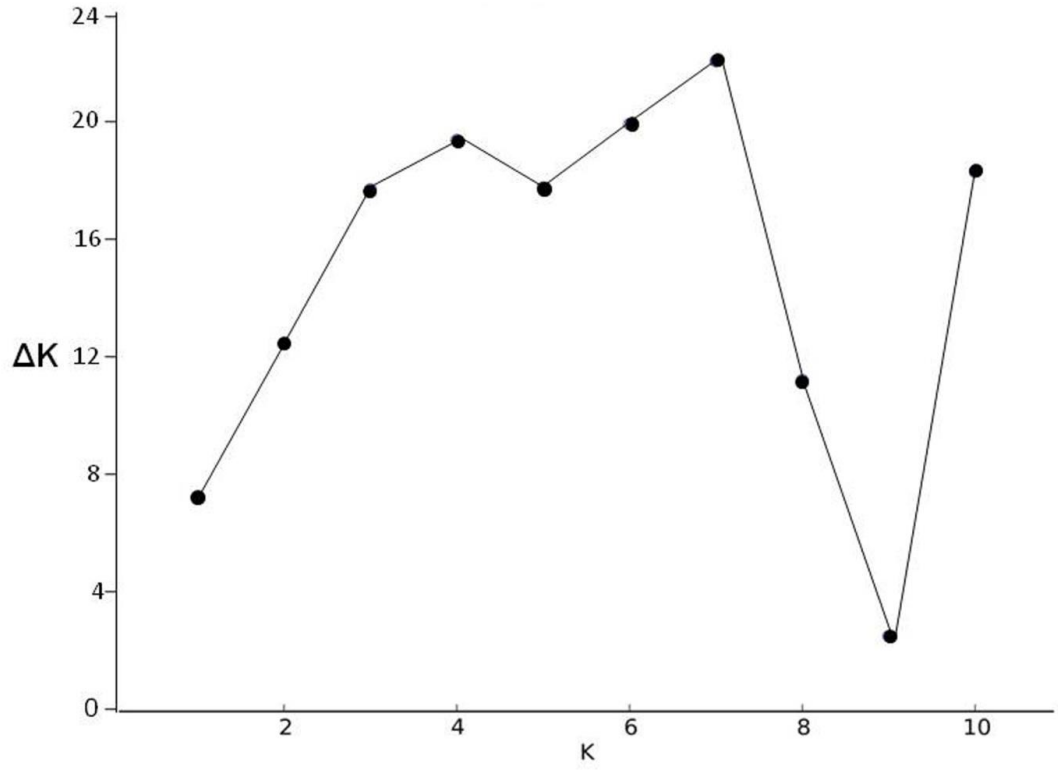

Figure 7**:** STRUCTURE analysis of the total genetic clusters for values of *K* (*K*=1 to 10), using delta *K* (*ΔK*) values

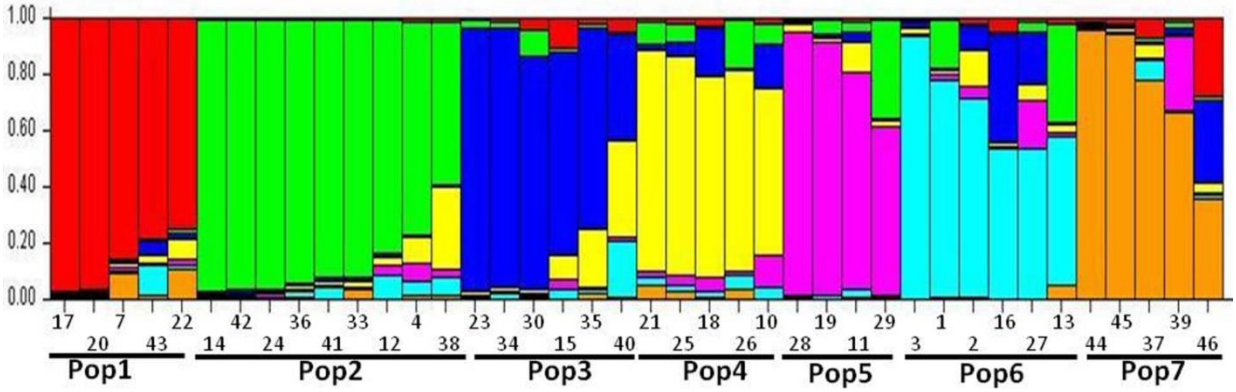

Figure 8**:** Population structure of 40 common bean genotypes based on peroxidase-gene (POX) based primers for *K* = 7. The colors represent single sub-population and the colored segment length indicates the analyzed membership proportion of every sample to designed population. The maximum K value was determined by structure harvest to be 7 meaning that the 40 common bean genotypes consisted of 7 sub-populations

4.2.6 Principal component analysis (PCA)

Genetic relationship of individual genotypes was analyzed using PCA. The analysis showed that grouping of the 40 common bean genotypes was not on the basis of a particular geographical region (Figure 9). The components are made up of 21.6% and 14.6% of total molecular variance.

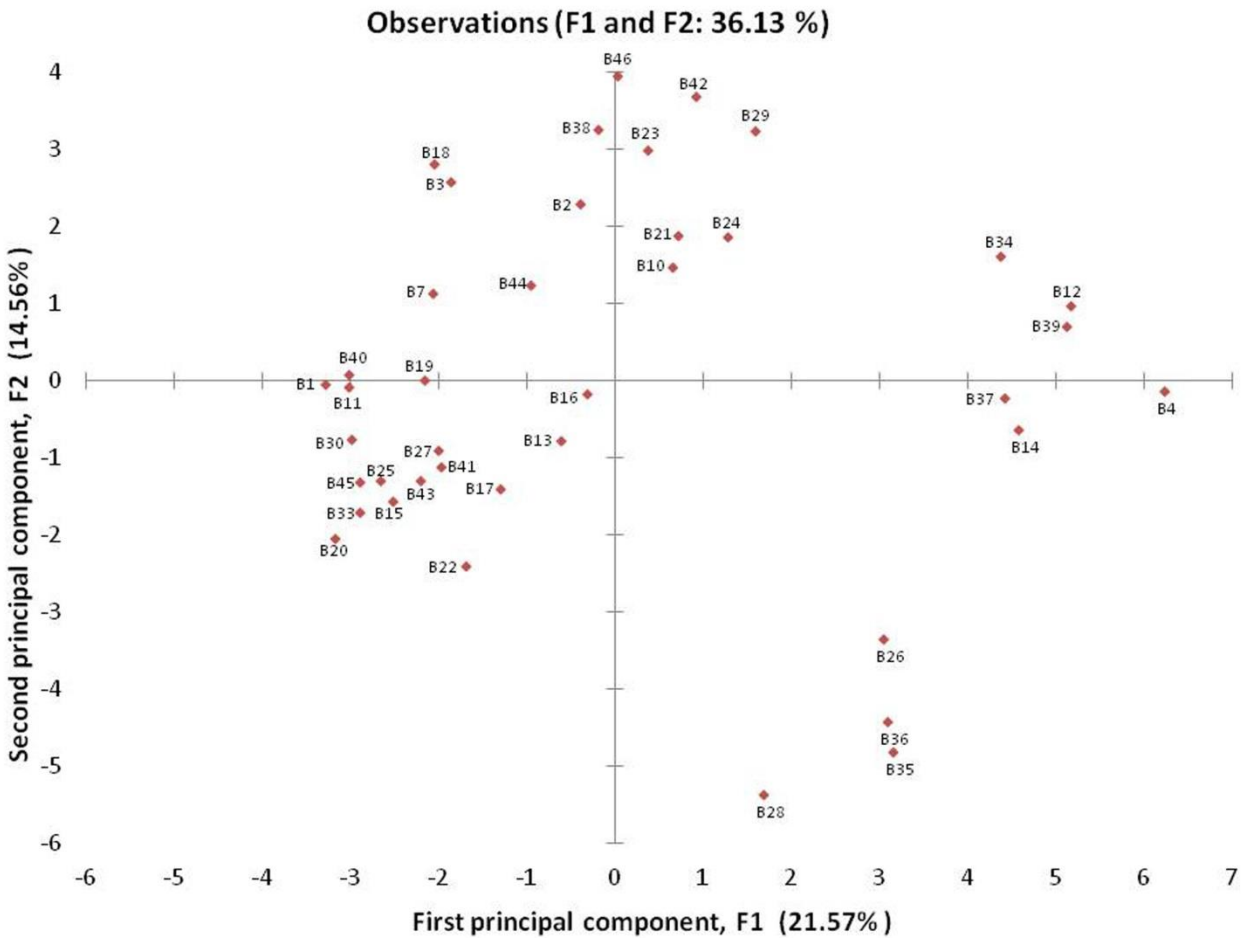

Figure 9**:** Two-dimensional PCA of 40 Kenyan common bean genotypes using peroxidase-gene primers

4.2.7 Analysis of molecular variance (AMOVA)

Analysis of molecular variance was calculated in order to estimate the partitioning of variation genetically into two components; among and within populations. The genetic partitioning within population was 99% and among population was 1% . There was a difference ($P \leq 0.001$) in molecular variance within geographical regions of collection (Table 9).

Table 9**:** Analysis of molecular variance of the POX markers among and within 40 common bean genotypes

Source	**df**	**SS**	**MS**	**Estimated variation**	**Percentage variation**	**P-Value**
Variation among populations	5	38.825	7.765	0.090	1%	< 0.001
Variation within populations	34	244.200	7.182	7.182	99%	< 0.001
Total	39	283.025		7.272	100%	

df = degrees of freedom, SS = sum of squares, MS = mean of squares

CHAPTER FIVE

5.0 DISCUSSION

Knowledge of genetic diversity is a crucial determinant of germplasm utilization in crop improvement strategies to meet the demand for future food security. Germplasm with high level of genetic diversity is a valuable resource for broadening the genetic base in any breeding program. Limited genetic diversity poses a threat to the survival of a species as this limits ability to respond to changes in climate, pathogen populations and agricultural practices (Manifesto *et al*., 2001). Hence, evaluating different sets of genetic materials with appropriate tools would be useful for identifying diverse genotypes to be incorporated in different breeding programs. A number of tools including morphological and molecular markers have been used for revealing the genetic diversity in crop plants. Using morphological markers is not an easy task because these traits can be affected by environmental factors and cultivation conditions, which reduce the accuracy of the results (Lee, 1995). Therefore, use of molecular markers has become the most reliable technique for analysis of the genetic diversity of germplasm resources.

5.1 Molecular characterization of common bean germplasm using SSR markers

In the present study, SSR markers were successfully used to determine genetic diversity among 40 common bean genotypes grown in Kenya. The three SSR primers specific to cowpea produced amplification products in common bean showing that a considerable level of sequence conservation exists within the primer regions flanking the microsatellite loci. This was the first time that SSRs developed for *V. unguiculata* were used in the species *P. vulgaris*. The five SSR markers were able to discriminate between the different

genotypes. Studies have shown that SSR loci give good discrimination between closely related individuals in some cases even, when only a few loci are employed (Powell *et al*., 1996). A total of 366 alleles were amplified with 4.5 alleles per SSR loci. The average number of alleles per locus (4.5) was higher compared to previous reports using AFLP (1.45) and SSAP markers (1.68) (Liu and Hou 2010; Ouji *et al*., 2012). This suggests that SSR markers are suitable tools for assessing genetic diversity of common bean. Asfaw *et al*. (2009) found 389 alleles with an average of 10 alleles per locus using 38 SSR markers from a collection of 192 common bean collections from East Africa. Blair *et al*. (2010) reported 301 alleles with an average of 10 alleles across 30 SSR markers in 365 common bean genotypes from Central Africa. Okii *et al*. (2014) also found 423 alleles with an average of 19 alleles per locus using 22 SSR markers in 100 common bean genotypes from Uganda. The marked differences of alleles recorded in this study and other previous studies in common bean can be attributed to the differences in the number and type of polymorphic markers used, sample sizes and collection sites.

The polymorphic information content (PIC) demonstrates the informative capability of the SSR loci and their potential to detect differences among the varieties based on their genetic relationships (Al-Badeiry *et al.*, 2014). In this study, the PIC values for the SSR loci ranged from 0.4818 for Vm94 to 0.7439 for Bmd17, with an average PIC value of 0.5958, which confirms that SSR markers used in this study were highly informative, because PIC values higher than 0.5 indicate high polymorphism. The high level of polymorphism is due to diverse genotypes and more variation of SSR loci used in the present study. Markers with PIC values of 0.5 or higher are highly informative for genetic

studies and are extremely useful in distinguishing the polymorphism rate of a marker at a specific locus (DeWoody *et al.*, 1995). Benchimol *et al.* (2007) assessed the genetic diversity of 20 common bean genotypes with SSRs and found PIC values ranging from 0.05 to 0.83. Perseguini *et al.* (2011) obtained PIC values varying from 0.03 to 0.70 for a set of 60 common bean genotypes, suggesting that PIC is strongly influenced by the number and diversity of the genotypes under evaluation. Lower PIC value may be the result of closely related genotypes and the high values of PIC indicate that the markers used showed that the varieties were of high diversity. In addition, the number of alleles amplified by a primer and its PIC values depends on the repeat number and the repeat sequence of the microsatellite sequences (Blair *et al*., 2011, 2013).

Gene diversity or expected heterozygosity can be used as a general indicator of the amount of genetic variability in a population (Nassiry *et al*., 2009). The mean Nei's (1973) gene diversity (h) of the loci producing polymorphic bands in this study ranged from 0.1215 to 0.3212 with a mean value of 0.2129. Marker Vm71 had the lowest value while marker Bmd2 had the highest value, suggesting that Bmd2 loci could be useful in revealing genetic diversity of common bean genotypes in Kenya. This observation was also confirmed by Shannon's information index at locus Bmd17 (p = 0.4811), which had the highest value as compared to the lowest value of p = 0.1988 at locus Vm71. The observed heterozygosity (Ho) calculated for each primer ranged from 0.5536 (Vm71) to 0.7750 (Bmd17) with a mean of 0.6537. The genetic similarity coefficients ranged from 0.15 to 1.0 with an average of 0.54, which indicate substantial diversity (0 to 85%)

among the genotypes used in the present study. These results reveal an abundance of genetic diversity in the common bean genotypes cultivated in Kenya.

The genetic diversity of a population in a species is affected by a number of factors, including the seed dispersal, gene flow, natural selection, geographic range, and the diversity center (Hamrick and Godt, 1989). In the present study, the dendrogram constructed using UPGMA method suggested occurrence of two major clusters. The UPGMA cluster analysis of the genotypes based on the SSR data illustrated no clear grouping of genotypes by geographical region. The observed low divergence of common bean genotypes from different growing regions could be explained by the high gene flow rate or the extensive germplasm exchange within Kenya and in most cases farmers grow common beans, either from seeds collected in their neighborhood or from seeds purchased at the market.

Principal component analysis (PCA) is a widely used tool in analyzing genetic variation among plant accessions and provides information about associations between genotypes, which are useful to formulate better strategies for breeding (Price *et al.*, 2006). The common bean genotypes did not cluster into distinct groups on the scatter plots. In addition, there was no obvious relationship between geographical origin and distribution of the genotypes on the scatter plot. In the PCA scatter plot, the distances among the genotypes reflected the genetic distances among them, hence varieties that were clustered close together were interpreted to be closely related and sharing similar genetic traits whereas those clustered far apart were distantly related. Clustering of the bean genotypes

by UPGMA and PCA methods revealed that there was no association in the observed pattern of variations with their geographical origin. Such non-congruence between the clustering pattern and geographical origin could be due to exchange of germplasm among the different geographical regions. Therefore, the artificial transfer of genotypes from one region to another resulted in a false determination of the geographic origin.

5.2 Genetic diversity and population structure of common bean germplasm using POX markers

In the current study, 40 genotypes of Kenyan common bean were characterized using five POX loci in order to detect genetic diversity and population structure among the five common bean-growing regions of Kenya. The POX markers were previously made using conserved motifs of rice and *Arabidopsis* peroxidase by Gulsen *et al*. (2007). All the POX markers used in this study were polymorphic. They had an average PIC of 0.77, implying that the POX primers were highly informative since any value of PIC greater than 0.5 shows high polymorphism (Bolstein *et al*., 1980). This affirms the potential application of POX markers in evaluating the genetic resources of common bean through genetic diversity studies. The mean PIC value reported in this study compared favorably with that reported by Wittayawannakul *et al*. (2010), who studied the genetic diversity among *Garcinia* species using POX polymorphisms and obtained an average PIC value of 0.79. However, the mean value of PIC observed in this study was greater than that reported by Nemli *et al*. (2014) who estimated variation among Turkey's common bean genotypes and got a mean PIC of 0.40. PIC is among the most vital characteristics of genetic markers and can be utilized in indicating differentiation abilities of the primers

(Ni *et al.*, 2002). Findings in this research, therefore, suggest that Kenyan common bean germplasms harbor high genetic divergence.

The total amplified alleles on each locus varied from 3 - 9, with an average score of 7.2. Nemli *et al*. (2014) reported 1 - 8 alleles on each locus and had an average value of 4.0 when characterizing common bean genotypes from Turkey using POX markers. The high number of alleles indicates high level of variation among Kenyan bean cultivars, which can be utilized in breeding programs. Values for Shannon's diversity value were between 0.4657 and 0.6327, with 0.5784 being the mean value. Gene diversity (He), which measures the heterozygosity at every loci was moderate at 0.40. This is a common phenomenon, particularly for species that are self-pollinated such as beans. The findings here correspond with previous reports carried out in Spain by Santalla *et al*. (2002), and also in Portugal by Martins *et al*. (2006).

Cluster analysis using the Jaccard's similarity coefficient grouped the 40 cultivars into two main groups. The germplasms were not grouped based on their origin of collection, suggesting genetic variation of collections from the same origin. These findings indicate that no relationship exist between geographical distribution and molecular divergence of common bean genotypes grown in Kenya. Similar to this research, lacking geographic associations with the collection source among genotypes was also reported by Okii *et al*. (2014), Fisseha *et al*. (2016) and Nyakio *et al*. (2015) for bean germplasms from Uganda, Ethiopia and Kenya, respectively. The lack of differentiation among geographical groups may be due to human activities involving transportation to different geographical regions

and seed trade leading to exchange of germplasm among and between farmers of different regions (Buah *et al*., 2017). Knowledge of genetic diversity of common bean will benefit breeding programs and will facilitate greater use of cultivars to design new improved cultivars having wider genetic base. This is advantageous in common bean breeding programs because such materials can be used to maximize the level of variation and to assess these genotypes in varying environments that make it possible to broaden their diversification.

Principal component analysis is an important tool in detecting variation among genotypes of plants and showing the most vital variables leading to such difference (Price *et al*., 2006). The common bean genotypes did not cluster into distinct groups on the scatter plots and intermixing of populations was observed with the PCA plot. The population structure of the 40 bean germplasms was evaluated using STRUCTURE software and it showed a high value of K (K = 7). Thus, the 40 common bean genotypes were divided by STRUCTURE analysis into seven subgroups and did not indicate any geographical distribution pattern. Genetic similarities among the sub-populations were not similar with the clusters identified using UPGMA and PCA analysis. The seven sub-populations showed different levels of admixture by STRUCTURE analysis probably due to the mixing of accessions across regions or previous breeding programs through the years. In addition, populations overlapping in this study also show that exchange of common bean genotypes between regions might be more, as most genotypes from these regions were distant from each other geographically. There was no clear clustering of genotypes based

on geographical regions, which is mostly due to high gene flow caused by material traffic and seed exchange among farmers between different regions.

Analysis of molecular variance showed a higher genetic variance within populations than among populations, meaning there is a low genetic differentiation existing among populations. The results from this study are similar to previous reports in common bean, which showed a greater genetic differentiation within than among the populations (Fisseha *et al*., 2016; Nyakio *et al*., 2015; Gyang *et al*., 2017). Results here are also in agreement with findings from other grain legumes like faba beans, reported by Oliveira *et al*., (2016); and Rebaa *et al*., (2017) and also cowpea (Badiane *et al*., 2012).

Plant peroxidases serve vital functions in various abiotic and biotic stress resistance (Ocal *et al*., 2014; Seda *et al*., 2014), and they might be utilized in defining relationships among plant genotypes in relation to their adaptive conditions in different geographical locations. Common bean genotypes from various geographical locations experience several abiotic and biotic stresses that could increase developmental variations in their peroxidases, which may result to differences in estimating genetic variability and relationships. In the present study peroxidase genes were found to be vital in determining diversity and the population structure among bean and findings here could be exploited for managing germplasms and using it in breeding programs.

CHAPTER SIX

6.0 CONCLUSIONS AND RECOMMENDATIONS

Genetic variability is important for the development of new and improved cultivars. In this study, two molecular markers (SSR and POX) were successfully used to reveal characterization of the genetic identities and relationships among 40 common bean cultivars grown in different regions in Kenya. Therefore, they provided effective genetic information of common bean germplasm for future breeding programs.

The use of SSR markers demonstrated the existence of a considerable amount of genetic diversity among common bean germplasms grown in different regions of Kenya. This indicates the potential application of such genotypes in common bean breeding programs by exploiting SSR markers for selection of specific traits. The cluster analysis results of SSR amplification can still be used by common bean breeders to guide crossings and to evaluate the need to incorporate greater genetic variability in their breeding programs. The results of the current study show that SSR markers can be reliably used for common bean genetic diversity studies, which is key in conducting breeding programs in order to obtain new biotic and abiotic-tolerant common bean varieties.

This study also shows that the five POX markers used were highly polymorphic and they distinguished the 40 common bean genotypes investigated sufficiently. The cultivars showed the presence of high genetic variability in Kenyan common bean germplasm which could be useful for future genetic improvement such as selection of parental lines for breeding schemes so as to obtain improved abiotic and biotic-tolerant common bean

genotypes. Population structure analysis determined seven groups and indicated that the Kenyan common bean germplasm has a high level of admixture. Results of this study indicate that POX gene-based primers would be important tools for genetic diversity research and breeding programs of common bean.

REFERENCES

Abdurakhmonov IY and Abdukarimov A. Application of association mapping to understanding the genetic diversity of plant germplasm resources. International Journal of Plant Genomics. 2008; 92(6): 478-487.

Al-Badeiry NA, Al-Saadi AH, Merza TK. Analysis of genetic diversity in maize (*Zea mays* L.) varieties using simple sequence repeat (SSR) markers. Pure and Applied Science. 2014; 6: 1768- 1774.

Almagro L, Gomez RLV, Belchi-Navarro S, Bru R, Ros BA and Pedreno MA. Class III peroxidases in plant defense reactions. Journal of Experimental Botany. 2008; 60: 377-390.

Anderson JW, Smith BM and Washnock CS. Cardiovascular and renal benefits of dry bean and soybean intake. American Journal of Clinical Nutrition. 1999; 70: 464S-474S.

Arunga EE, Kinyua M, Ochuodho J, Owuoche J, Chepkoech E. Genetic diversity of determinate French beans grown in Kenya based on morpho-agronomic and simple sequence repeat variation. Journal of Plant Breeding and Crop Science. 2015; 7(8): 240-250.

Arumuganathan KEE and Earle ED. Nuclear DNA content of some important plant species. Plant Molecular Biology Reporter. 1991; 3: 208-218.

Asfaw A, Blair MW, Almekinders C. Genetic diversity and population structure of common bean (*Phaseolus vulgaris* L.) landraces from the East African highlands. Theoretical and Applied Genetics. 2009; 120: 1-12.

Badiane FA, Gowda BS, Cissé N, Diouf D, Sadio O, Timko MP. Genetic relationship of cowpea (*Vigna unguiculata*) varieties from Senegal based on SSR markers. Genetics and Molecular Research (GMR). 2012; 11(1): 292-304.

Beebe SE, Rao IM, Blair MW and Acosta-Gallegos JA. Phenotyping common beans for adaptation to drought. Frontiers in Physiology. 2013; 4: 35.

Beebe SE, Skroch PW, Tohme J, Duque MC, Pedraza F, Nienhuis J. Structure of genetic diversity among common bean landraces of Mesoamerican origin based on correspondence analysis of RAPD. Crop Science. 2000; 40: 264-273.

Benchimol LL, Campos T, Carbonell SAM, Colombo CA, Chioratto AF, Formighieri EF, Souza AP. Structure of genetic diversity among common bean (*Phaseolus vulgaris* L.) varieties of Mesoamerican and Andean origins using new developed microsatellite markers. Genetic Resources and Crop Evolution. 2007; 54: 1747-1762.

Blair MW, Giraldo MC, Buendía HF, Tovar E, Duque MC, Beebe SE. Microsatellite marker diversity in common bean (*Phaseolus vulgaris* L.). Theoretical and Applied Genetics. 2006; 113(1): 100-109.

Blair MW, Gonzalez LF, Kimani PM, Butare L. Genetic diversity, inter-gene pool introgression and nutritional quality of common beans (*Phaseolus vulgaris* L.) from Central Africa. Theoretical and Applied Genetics. 2010; 121: 237-248.

Blair MW, Hurtado N, Chavarro CM, Munoz-Torress CM, Giraldo MC, Pedraza F, Tomkins J, Wing R. Gene-based SSR markers for common bean (*Phaseolus vulgaris* L.) derived from root and leaf tissue ESTs: an integration of the BMc series. Plant Biology. 2011; 10: 1186/1471-2229-11-50.

Blair MW, Paulo I, Astudillo C, Grusak MA. A legume biofortification quandary: variability and genetic control of seed coat micronutrient accumulation in common beans. Frontiers in Plant Science. 2013; 10: 3389.

Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. American Journal of Human Genetics. 1980; 32(3): 314-331.

Broughton WJ, Hernandez G, Blair M, Beebe S, Gepts P, Vanderleyden J. Beans (*Phaseolus* spp.) – model food legumes. Plant Soil. 2003; 252, 55-128.

Buah S, Buruchara R, Okori P. Molecular characterization of common bean (*Phaseolus vulgaris* L.) accessions from Southwestern Uganda reveal high levels of genetic diversity. Genetic Resources and Crop Evolution. 2017; 1 – 14.

Buso GS, Amaral ZP, Brondani RP and Ferreira ME. Microsatellite markers for the common bean *Phaseolus vulgaris*. 2006; 6: 252-254.

Ceylan H. Detection of peroxidase gene polymorphism in wheat from Turkey. Erciyes University. 2010.

Choudhary K, Mathur N, Choudhary OP, Pillai U. Protocol for isolation of genomic DNA from dry and fresh leaves of *Vigna* Species suitable for RAPD and restriction digestion. Advances in Biological Research. 2008; 2(5-6): 83-86.

Collard BCY, Mackill DJ. Conserved DNA-derived polymorphism (CDDP): a simple and novel method for generating DNA markers in plants. Plant Molecular Biology Reporter. 2009; 27: 558–562.

De Luque JJR and Creamer B. Major constraints and trends for common bean production and commercialization; establishing priorities for future research. Agronomía Colombiana. 2014; 32(3): 423-431.

DeWoody JA, Honeycutt RL, Skow LC. Microsatellite markers in white-tailed deer. Journal of Heredity. 1995; 86: 317–319.

Ellis JR, Burke JM. EST-SSRs as a resource for population genetic analyses. Heredity. 2007; 99: 125-132.

Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology. 2005; 14: 2611–2620.

FAO (Food and Agriculture Organization of the United Nations) Statistics Division. 2015. Available: http://faostat.fao.org/

Fisseha Z, Tesfaye K, Dagne K, Blair MW, Harvey J, Kyallo M, Gepts P. Genetic diversity and population structure of common bean (*Phaseolus vulgaris* L) germplasm of Ethiopia as revealed by microsatellite markers. African Journal of Biotechnology. 2016; 15(52): 2824-2847.

Galal OA, Rehan M and Abd El-Karim RE. Analysis of genetic diversity within and among four rabbit genotypes using biochemical and molecular genetic markers. African Journal of Biotechnology. 2013; 12(20): 2830-2839.

Gepts P. *Phaseolus vulgaris* (Bean). Academic Press. 2001; 95: 616-851.

Ghamari H, Ahmadvand G. Growth analysis of dry bean (*Phaseolus vulgaris L.*) in different weed interference situations. Notulae Scientia Biologicae. 2013; 5(3): 394-399.

Gomez OJ, Blair MW, Frankow-Lindberg BE, Gullberg U. Comparative study of common bean (*Phaseolus vulgaris* L.) landraces conserved *ex situ* in genebanks and *in situ* by farmers. Genetic Resources and Crop Evolution Journal; 2005; 52: 371-380.

Grisi MC, Brondani RP, Blair M, Gepts P, Brondani C, Pereira PA. Genetic mapping of a new set of microsatellite markers in a reference common bean (*Phaseolus vulgaris*) population. Genetics and molecular research. 2007; 891-900.

Gulsen O, Kaymak S, Ozongun S and Uzun A. Genetic analysis of Turkish apple germplasm using peroxidase gene-based markers. Scientia Horticulturae. 2010; 125: 368–373.

Gulsen O, Shearman RC, Heng-Moss TM, Mutlu N, Lee DJ and Sarath G. Peroxidase gene polymorphism in buffalograss and other grasses. Crop science. 2007; 47: 767–772.

Gyang JP, Evans NN, Edward KM. Molecular characterization of common bean (*Phaseolus vulgaris* L.) genotypes using microsatellite markers. Journal of Advances in Biology & Biotechnology. 2017; 13(2): 1-15.

Hamrick JL, Godt MJW. Allozyme diversity in plant species. In: Brown AHD, Clegg MT, Kahler AL and Weir BS (Eds). Plant Population Genetics, Breeding and Genetic Resources. Sinauer Press, Sunderland, Mass. 1989: pp. 43–63.

Homar RG, Jose SM, Patricia MLV, Rigoberto R, Netzahualcoyotl M. Genetic diversity analysis of common beans based on molecular markers. Genetic and Molecular Biology. 2011; 34(4): 595-605.

Hiraga S, Sasaki K, Ito H, Ohashi Y and Matsui H. A large family of class-III plant peroxidases. Plant Cell Physiology. 2001; 42: 462–468.

Isemura T, Kaga A, Tabata S, Somta P, Srinives P, Shimizu T, Jo U, Vaughan DA, Tomooka N. Construction of a genetic linkage map and genetic analysis of domestication related traits in mungbean (*Vigna radiata*). PLoS ONE. 2012; 7(8): e41304.

Katungi E, Farrow A, Mutuoki T, Gebeyehu S, Karanja D, Alamayehu F, Sperling L, Beebe S, Rubyogo JC, Buruchara R. Improving common bean productivity. An Analysis of Socio-economic factors in Ethiopia and Eastern Kenya. CIAT Working. 2010; document #217.

Katungi E, Horna D, Gebeyehu S, Sperling L. Market access, intensification and productivity of common bean in Ethiopia: a microeconomic analysis. African Journal of Agricultural Research. 2011; 6(2): 476-487.

Khaidizar MI, Haliloglu K, Elkoca E, Aydin M, Kantar F. Genetic diversity of common bean (*Phaseolus vulgaris* L.) landraces grown in Northeast Anatolia of Turkey assessed with simple sequence repeat markers. Turkish journal of field crops; 2012, 17(2): 145-150.

Kimura M and Crow J. The number of alleles that can be maintained in a finite population. Genetics. 1964; 49: 725-738.

Laurentin H. Data analysis for molecular characterization of plant genetic resources. Genetic Resources and Crop Evolution. 2009; 56: 277-292.

Lee M. DNA markers and plant breeding programs. Advance in Agronomy. 1995; 55: 265-344.

Lewontin RC. The apportionment of human diversity. Evolutionary Biology. 1972; 6: 381-398.

Lioi L, Piergiovanni AR, Pignone D, Puglisi S, Santantonio M, Sonnante G. Genetic diversity of some surviving on-farm Italian common bean (*Phaseolus vulgaris* L.) landraces. Plant Breeding. 2005; 124: 576-581.

Liu Y and Hou W. Genetic diversity of faba bean germplasms in Qinghai and core germplasm identified based on AFLP analysis. Legume Genomics & Genetics. 2010; 1: 1.

Manifesto MM, Schlatter AR, Hopp HE. Quantitative evaluation of genetic diversity in wheat germplasm using molecular markers. Crop Science. 2001; 41(3): 682–690.

Maras M, Pipan B and Jelka S-V. Examination of genetic diversity of common bean from the Western Balkans. Journal of American Society for Horticultural Science. 2015; 140(4): 308-316.

Maras M, Sustar-Vozlic J, Javornik B, Meglic V. The efficiency of AFLP and SSR markers in genetic diversity estimation and gene pool classification of common bean (*Phaseolus vulgaris* L) Acta Agriculturae Slovenica. 2008; 91: 87-96.

Marrotti I, Bonetti A, Minelli M, Catizone P, Dinelli G. Characterization of some Italian common bean (*Phaseolus vulgaris* L.) landraces by RAPD, semirandom and ISSR molecular markers. Genetic Resources and Crop Evolution. 2007; 54: 175-188.

Martins SR, Vences FJ, Miera LE, Barroso MR, Carnide V. RAPD analysis f genetic diversity among and within Portuguese landraces of common white bean (*Phaseolus vulgaris* L.). Scientia Horticulturae. 2006; 108: 133-142.

Matus IA, Hayes PM. Genetic diversity in three groups of barley germplasm assessed by simple sequence repeats. Genome. 2002; 45(6): 1095–1106.

Nassiry MR, Javanmard A, Tohidi R. Application of statistical procedures for analysis of genetic diversity in domestic animal populations. American Journal of Animal and Veterinary Sciences. 2009; 4(4): 136-141.

Nei M. Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences of the USA. 1973; 70(12): 3321-3323.

Nei M and Li WH. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proceedings of the National Academy of Sciences of the USA. 1979; 76(10): 5269-5273.

Nemli S, Kaya HB, Tanyolac B. Genetic assessment of common bean (*Phaseolus vulgaris* L.) accessions by peroxidase gene based markers. Journal of the Science of Food and Agriculture. 2014; 94(8): 1672-1680.

Ni J, Colowit PM, Mackill DJ. Evaluation of genetic diversity in rice subspecies using microsatellite markers. Crop Science. 2002; 42(2): 601–607.

Nodari RO, Koinange EM, Kelly JD, Gepts P. Towards an integrated linkage map of common bean: Development of genomic DNA probes and levels of restriction fragment length polymorphism. Theoretical and Applied Genetics. 1992; 84: 186-192.

Nyakio KM, Katherine AS, Valerie AP. Genetic diversity of dry bean (*Phaseolus Vulgaris* L.) accessions of Kenya using SSR markers. American Journal of Experimental Agriculture. 2015; 5(4): 306-319.

Nyombaire G, Siddiq M, Dolan K. Effect of soaking and cooking on the oligosaccharides and lectins of red kidney beans (*Phaseolus vulgaris* L.). Bean Improvement Cooperative. 2007; 50: 31-32.

Ocal N, Mikail A, Osman G, Halit Y, Ilknur S, Nebahat S. Genetic diversity, population structure and linkage disequilibrium among watermelons based on peroxidase gene markers. Science Horticulturae. 2014; 176: 151-161.

Ocampo CH, Martin JP, Sanchez-Yelamo MD, Ortiz JM, Toro O. Tracing the origin of Spanish common bean cultivars using biochemical and molecular markers. Genetic Resources and Crop Evolution. 2005; 52: 33-40.

Okii D, Tukamuhabwa P, Odong T, Namayanja A, Mukabaranga J, Paparu P, Gepts P. Morpholgical diversity of tropical common bean germplasm. African Crop Science Journal. 2014; pp. 59-67.

Ouji A, Bok SE, Syed NH, Abdellaoui R, Rouaissi M, Flavell AJ, Gazzah ME. Genetic diversity of faba bean (*Vicia faba* L.) populations revealed by sequence specific amplified polymorphism (SSAP) markers. African Journal of Biotechnology. 2012; 11(9): 2162-2168.

Oliveira HR, Toma´s D, Silva M, Lopes S, Viegas W, Veloso MM. Genetic diversity and population structure in *Vicia faba* L. landraces and wild related species assessed by nuclear SSRs. PLoS ONE. 2016; 11(5): e0154801.

Passardi F, Longet D, Penel C and Dunand C. The class III peroxidase multigenic family in rice and its evolution in land plants. Phytochemistry. 2004; 65: 1879-1893.

Passardi F, Cosio C, Penel C, Dunand C. Peroxidases have more functions than a Swiss army knife. Plant Cell Reports. 2005; 24: 255-265.

Peakall R and Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research – an update. Bioinformatics. 2012; 28(19): 2537-2539.

Pejic I, Ajmone-Marsane P, Morgane M, Kozumplick V, Castaglioni P, Taramino G, Motto M. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, SSRs and AFLPs. Theoretical Applied Genetics. 1998; 97: 1248-1255.

Perseguini JM, Chioratto AF, Zucchi MI, Colombo CA, Carbonell SA, Mondego JM, Gazaffi R, Garcia AA, de Campos T, de Souza AP, and Rubiano LB. Genetic diversity in cultivated carioca common bean based on molecular marker analysis. Genetics and Molecular Biology Journal. 2011; vol.34 (1): 88-102.

Poczai P, Varga I, Laos M, Cseh A, Bell N, Valkonen JPT, Hyvönen J. Advances in plant gene-targeted and functional markers: a review. Plant Methods. 2013; 9: 6.

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Molecular Breeding. 1996; 2: 225-238.

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nature Genetics. 2006; 38: 904-909.

Pritchard JK. Inference of population structure using multilocus genotype data. Genetics. 2000; 155: 945-959.

Rebaa F, Abid G, Aouida M, Abdelkarim S, Aroua I, Muhovski Y, Baudoin J, Mahmoud M, Sassi K, Jebara M. Genetic variability in Tunisian populations of faba bean (*Vicia faba* L. var. *major*) assessed by morphological and SSR markers. Physiology and Molecular Biology of Plants. 2017; 23(2): 397-409.

Ren LL, Liu YJ, Liu HJ, Qian TT, Qi LW, Wang XR, Zeng QY. Subcellular relocalization and positive selection play key roles in the retention of duplicate genes of Populus class III peroxidase family. Plant Cell. 2014; 26: 2404–2419.

Saghai-Maroof MA, Biyashev RM, Yang GP, Zhang Q, Allard RW. Extraordinarily polymorphic microsatellites DNA in barley: Species diversity, chromosomal locations and population dynamics. Proceedings Natural Academy Science (USA). 1994; 91: 5466 - 5470.

Sambrook J, Fritsch EF, Maniatis F. Molecular cloning. A laboratory manual, 2nd edition. Cold Spring Harbor, New York; Cold Spring Laboratory. 1989.

Santalla M, Rodino AP, De Ron AM. Allozyme evidence supporting southwestern Europe as a secondary centre of genetic diversity for common bean. Theoretical and Applied Genetics. 2002; 104: 934-944.

Singh SP, Nodari R, Gepts P. Genetic diversity in cultivated common bean: I. Allozymes. Crop Science. 1991; 331: 19-23.

Svetleva D, Pereira G, Carlier J, Cabrita L, Leitao J, Genchev D. Molecular characterization of *Phaseolus vlugaris* L genotypes included in Bulgarian collection by ISSR and AFLP analyses. Scientia Horticulturae. 2006; 109: 198-206.

Turnpenny P. and Ellard S. Emery's elements of medical genetics. 12th Edition, Elsevier, London. 2005.

Uzun A, Gulsen O, Seday U, Yesiloglu T, Kacar YA. Peroxidase gene-based estimation of genetic relationships and population structure among Citrus spp. and their relatives. Genetic Resources and Crop Evolution. 2014; 61: 1307-1318.

Vicuna D. The role of peroxidases in the development of plants and their responses to abiotic stresses. Doctoral thesis. Dublin Institute of Technology. 2005; 10: 21427.

Welinder KG. Super family of plant, fungal and bacterial peroxidases. Current Opinion in Structural Biology. 1992; 2: 388-393.

Welinder KG, Justesen AF, Kjaersgard IVH, Jensen RB, Rasmussen SK, Jespersen HM. Structural diversity and transcription of class III peroxidases from Arabidopsis thaliana. European Journal of Biochemistry. 2002; 269: 6063–6081.

Wittayawannakul W, Garcia RN, Yllano OB, Borromeo TH, Namuco LO and Tecson-Mendoza EM. Assessment of genetic diversity in *Garcinia*species usingperoxidase, RAPD and gene sequence specific amplification polymorphism (GSSAP). Philippine Agric Science. 2010; 93: 31-41.

Wu J, Wang L and Wang S. Comprehensive analysis and discovery of drought-related NAC transcription factors in common bean. BMC Plant biology. 2016; 16: 193.

Yoshida K, Kaothien P, Matsui T, Kawaoka A, and Shinmyo A. Molecular biology and application of plant peroxidase genes. Applied Microbiology and Biotechnology. 2003; 60: 670-675.

Zhang L, Pond SK and Gaut BS. A survey of the molecular evolutionary dynamics of twenty-five multigene families from four taxa. Journal of Molecular Evolution. 2001; 52: 144–156.