VARIABILITY STUDY AND GENOME-WIDE ASSOCIATION MAPPING FOR COOKING TIME IN COMMON BEAN (*Phaseolus vulgaris* L.) GENOTYPES

M.Sc. THESIS

ERMIAS ASSEFA EMAM

SEPTEMBER, 2016

JIMMA UNIVERSITY

VARIABILITY STUDY AND GENOME-WIDE ASSOCIATION MAPPING FOR COOKING TIME IN COMMON BEAN (*Phaseolus vulgaris L.*) GENOTYPES

M.Sc. THESIS

Submitted to the School of Graduate Studies, Jimma University College of Agriculture and Veterinary Medicine, Department of Horticulture and Plant Science

In Partial Fulfillment of the Requirements for the Degree of Master of Science in Plant Breeding

September, 2016

Jimma University, Ethiopia

DEDICATION

This thesis is dedicated to my mother, **Tenaye Ababu Yesuf**, who did a lot for my educational endeavor

STATEMENT OF THE AUTHOUR

I declare that this thesis is my original work and that all sources of materials used for this thesis have been duly acknowledged. This thesis has been submitted in partial fulfillment of the requirements for M.Sc. degree at Jimma University, College of Agriculture and Veterinary Medicine and put at the University Library to be made available to borrowers under the rules of Library. I declare that this thesis is not submitted to any other institution anywhere for the award of any academic Degree, Diploma, or certificate.

Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the Plant Science and Horticulture Department or the Director of the School of Graduate Studies when his/her judgment the proposed use of the material is in the interest of scholarship. In all other instances, however, permission must be obtained from the author.

Name: Ermias Assefa

Signature: __________

Date of Submission: ______________

Place: Jimma University

BIOGRAPHICAL SKETCH

The author was born on March 8, 1988 in Bale Goba (Oromia). He attended elementary secondary and preparatory school education at Butajira Elementary, secondary and preparatory School. After completion of his high school education, he joined Arbaminch University faculty of agriculture and graduated on July, 2010 with B.Sc Degree in horticulture.

Soon after his graduation, he was employed at Bonga Agricultural Research Center as a Junior Researcher, in pulse and oil crops improvement division under Southern Agricultural Research Institute. Finally, after serving for three years, he joined the School of Graduate studies in Jimma University College of Agriculture and Veterinary Medicine in October 2015 to pursue his M.Sc. in Plant Breeding.

ACKNOWLEDGEMENTS

I would like to praise the Almighty God and his blessed mother St. Mary for giving me endless opportunities and helps that enabled me to achieve all the successes I will achieve in my life.

I would like to express my deep gratitude to my major advisor Dr. Sentayehu Alamerew (Jimma University), for his constructive advice, guidance, encouragement and willingness to supervise my research and his invaluable comments during the course of this work. I would also like to thank Mr. Daniel Ambachew (Tennessee state university), for his guidance and encouragement during the course of my thesis work, especially providing research funding, and support the GAWS analysis.

My sincere gratitude is due to Dr. Karen Cichy (USDA, Sugar beet and Bean Research Unit leader, USA), for sharing us with the genotyping, data and allowing us to collect cooking time data in her lab free of charge. I am also very grateful to Dr Floure Deidre (ARC, South Africa) for sharing the diversity panel used for this study.

In addition, I am grateful to Dr. Tileye Feyissa, Dr. Gizachew Haile, Dr. Mandefro Nigussie, Mr. Temesgen Matiwos, Mr. Techale Birhan and Mr. Belachew Beyen for their help in the data analysis and constructive comments.

My deepest gratitude goes to Dr. Hailemariyam Gizaw, Zewditu Mulugeta, Fitsum Alemayehu, Yayis Rezene for they have helped me in one way or another from the preparation of the field material up to the timely completion of both field and laboratory work. My special thanks and appreciations also go to my wife Ms. Selamawit Teshome for their continuous support and encouragement.

Finally, I would like to acknowledge the finance support from Generation challenge program, Tropical legume I project (TLP) through Southern Agricultural Research Institute.

LIST OF ABBREVIATIONS

CSA	Central Statistical Agency
DArt	Diversity Array Technology
GA	Genetic Advance
GAM	Genetic advance as % of mean
GCV	Genotypic Coefficient of Variation
GWAS	Genome-Wide Association Study
HaARC	Hawassa Agricultural Research Center
LD	linkage disequilibrium
MAGIC	Multiple parent advanced generation inter cross population
MAS	Marker assisted selection
MLM	Mixed Linear Model
NAM	Nested Association Mapping
PCA	Principal Component Analysis
PCV	Phenotypic Coefficient of Variation
PH	Power of hydrogen
QTL	Quantitative trait loci
FDR	False Discovery Rate
SNP	Single nucleotide polymorphism
TASSEL	Trait Analysis by aSSociation, Evolution and Linkage

TABLE OF CONTENTS

Content **Page**

DEDICATION .... **i**
STATEMENT OF THE AUTHOUR .... **ii**
BIOGRAPHICAL SKETCH .... **iii**
ACKNOWLEDGEMENTS .... **iv**
LIST OF ABBREVIATIONS .... **v**
TABLE OF CONTENTS .... **vi**
LIST OF TABLES .... **ix**
LIST OF FIGURES .... **x**
LIST OF TABLE IN APPENDIX .... **xi**
LIST OF FIGURE IN APPENDICES .... **xii**
1. INTRODUCTION **2**
2. LITERATURE REVIEW **6**
2.1 Origin, Distribution and Taxonomy of Common Bean 6
2.2 Genetic Diversity of Common Bean 7
2.3 Cooking Time in Common Bean 8
2.4 Genotypic and Phenotypic Variability 9
2.5 Heritability 10
2.6 Genetic Advance 11
2.7 Correlation and Path Coefficient 12
2.7.1 Correlation among traits 12
2.8 Path Coefficient 13
2.9 Clustering 15
2.10 Principal Component 16
2.11 Molecular Markers and Mapping 17
2.11.1 Marker systems used in crop improvement strategies 17
2.11.2 Molecular markers 17
2.11.3 Single Nucleotide Polymorphism 18
2.11.4 Quantitative trait loci mapping (QTL) 18

TABLE OF CONTENTS (*Cont*..)

2.11.5 Association mapping 20

2.11.5.1 Candidate genes association mapping 21

2.11.5.2 Genome wide association mapping 22

3. MATERIALS AND METHODS 24

3.1 Experimental Site Description 24

3.2 Experimental Material 24

3.3 Field Experimental Design and Trial management 24

3.4 Data collected 25

3.4.3 Water uptake and cooking time 26

3.4.3.1 Water uptake 26

3.4.3.2 Cooking time 26

3.5 Statistical Analysis 27

3.5.1 Analysis of variance (ANOVA) 27

3.5.2.1 Estimation of variance components 28

3.5.2.2 Broad-sense heritability (h^2b) 28

3.5.2.3 Genetic advance under selection (GA) 29

3.5.2.4 Genetic advance as percent of mean 29

3.5.3 Correlation and path coefficient analysis 29

3.5.3.1 Correlation coefficient (r) 29

3.5.3.2 Path coefficient analysis 30

3.5.4 Cluster analysis 31

3.5.5 Genetic divergence analysis 31

3.5.6 Principal component analysis (PCA) 31

3.5.7 Genome-Wide Association analysis 32

4. RESULTS AND DISCUSSION 33

4.1 Variability Assessment 33

4.1.1 Analysis of variance (ANOVA) 33

4.1.2 Range and mean of different characteristics 33

4.1.3 Phenotypic and genotypic coefficient of variation 34

4.1.5 Estimates of expected genetic advance (GAM %) 36

TABLE OF CONTENTS (*Cont..*)

4.2 Association Studies 38

4.2.1 Correlation of grain yield with other traits 38

4.2.2 Correlation among other characters.. 39

4.3 Path Coefficient Analysis 40

4.4 Cluster Analysis 41

4.4.1 Mean cluster analysis.. 42

4.4.2 Genetic divergence 43

4.5 Principal Component Analysis.. 47

4.6 Phenotypic Diversity and Genome-wide Association Analysis 48

4.6.1 Phenotypic analysis for water uptake and cooking time.. 48

4.6.2 Heritability and Correlation Analysis.. 49

4.6.3 Fast cooking beans.. 50

4.7 Genome-wide Association Analysis.. 51

4.7.1. Genome-wide association analysis for water uptake.. 51

4.7.2 Genome-wide association analysis for cooking time 53

5 . SUMMARY AND CONCLUSION 56

6. REFERENCE..59

7. APPENDICES..72

LIST OF TABLES

Table **Page**

1: Skeleton of Analysis of variance (ANOVA) for augmented design 27
2. Estimate of mean, ranges, Phenotypic (PV) and genotypic (GV) coefficient of variation, bro ad sense heritability 37
3: Genotypic (above diagonal) and phenotypic correlation coefficients. 39
4: Path coefficients at genotypic level of direct (diagonal) and indirect effects of the characters 41
5: Mean values of 8 characters for the 21 cluster of 423 common bean genotype at Hawassa44
6: Mahalanobis distance between groups of common bean genotypes. 45
7: Eigenvectors and eigenvalues of the first four principal components (PCs) of common bean genotypes 48
8: Summary of the analysis of variance, heritability, phenotypic and genotypic variance for the cooking time and water absorption traits in common bean genotypes. 49
9: Identification of the eleven fastest cooking genotypes of the 398 common bean genotypes. 50
10. GWAS significant markers, genome position, *p* value, R^2, allele and allele frequency, and phenotypic effect associated with Water uptake 52
11. GWAS significant markers, genome position, p value, R^2, allele and allele frequency associated with cooking time 54

LIST OF FIGURES

Figure **Page**

1: Schematic diagram and contrast of genome-wide association mapping and candidate-gene association mapping. 22
2. QQ plot for the water uptake data and Manhattan plot of water uptake of 386 genotype with 4,416 SNP. 52
3. QQ plot for the cooking time data and Manhattan plot of cooking time of 386 genotypes with 4,416 SNP. 54

LIST OF TABLE IN APPENDIX

Appendix Table **Page**

1. Analysis of variance (Mean squares) for the 12 characters of 423 Common bean genotype. 73
2. Mean values of some of the studied 423 common bean genotypes 74
3. Distribution of 423 common bean landraces in different clusters group.......... 75

LIST OF FIGURE IN APPENDICES

Appendix Figure **Page**

1. Biplot scores of two principal components (PCA1 & PCA2).. 77

2. Linkage disequilibrium decay plot generated by 4,416 markers 77

Variability Study and Genome-Wide Association Mapping for Cooking Time in Some Common Bean (*Phaseolus vulgaris L.*) Genotypes

ABSTRACT

Common bean (Phaseolus vulgaris L.) is an important component of the production systems and a major source of protein for the poor in Eastern and Southern Africa. Information on genetic variability is prerequisite for further improvement of the crop. However, there is little information regarding variability study in the present genotypes of common bean. One of the major factors that limit greater utilization of beans is their long cooking times compared to other foods. So far, very little is known about the genomic regions involved in determining cooking time. The overall objective was to study the extent of genetic variation and association among grain yield and yield-related traits as well as to use genome-wide association analysis to identify genomic regions involved in determining cooking time. Four hundred twenty three genotype were tested in an augmented design at Hawassa Agricultural Research Centre in Southern Region of Ethiopia, in 2015. Analysis of variance revealed that the genotypes differ significantly for all the characters studied except for leaf chlorophyll content, pod harvest index and hundred seed weight. High phenotypic coefficient of variation (PCV) and genotypic coefficient of variation (GCV) were recorded for vertical root pulling force resistance and number of pod per plant. High GCV along with high heritability and genetic advance was obtained for number of pods per plant. Grain yield had positive and highly significant phenotypic and genotypic correlation with days to maturity, plant height, vertical root pooling force resistance and number of pod per plant. Path analysis revealed that plant height, days to maturity and pod per plant shows high and positive direct effect on grain yield. These three characters can be considered for selection. The D^2 analysis showed the 423 genotypes grouped into twenty-one clusters. This makes the genotypes to become moderately divergent. The x^2 test showed that all inter-cluster squared distance were highly significant at $P<0.01$. Principal component analysis showed that the first four principal components explained about 66.19 % of the total variation. In this study high variation for cooking time was observed and eleven common bean genotypes were identified which cook in less than 17 min. GWAS showed that, significant SNP associated with cooking time were found on chromosomes Pv04, Pv05, and Pv09.The associated markers are possible candidates for marker-assisted selection to improve cooking time trait. Further studies of common bean genotypes with larger sample size in different location should be conducted on common bean variability in order to give confirmative results. The significant markers need to be validated in different environments before their use in marker-assisted selection. This research serves as an important base for further studies to understand the genetic control of cooking time in common bean.

Keywords: Association mapping, Common bean, Correlation, Genetic Variability, Marker-assisted selection.

1. INTRODUCTION

Common bean (*Phaseolus vulgaris* L.) is one of the most important legumes for direct human consumption worldwide (Broughton *et al.,*2003).The genus *Phaseolus* contains approximately 70 species and within this genus, common bean (*Phaseolus vulgaris* L.) is a diploid ($2n = 2x = 22$) and predominantly self-pollinating species although 3% or more out crossing ratio has also been observed (Ibarra-Perez *et al*., 1997). It is found in two main centers of origin in the American continent: the Middle American and the Andean centers (Harlan, 1975). It was introduced to the old world by the Spaniards and Portuguese. It is now widespread and cultivated as a major food crop in many tropical, subtropical and temperate areas of the Americas, Europe, Africa and Asia (Wortmann, 2006).

Common bean is an important component of the production systems and contributes to the national economy as both a food and an export commodity, in both cases serving as a source of income and employment to a large supply chain (Tumsa *et al*., 2014). The country's export earnings from common bean (95.3 million USD) in 2012, exceeds that of other pulses (such as lentil, faba bean) (FAO, 2015). The crop provides vital nutrients as a food including vitamins, proteins (25%), Starch and minerals and is an excellent source of potassium, selenium, molybdenum, thiamine (Maiti and Singh, 2007). The stems of beans are also used as fodder for livestock, especially in the dry spell following the main cropping season (Wondatir and Mekasha, 2014). As a legume, common bean plants also contribute to soil fertility enhancement through atmospheric nitrogen fixation (Broughton *et al*., 2003).

Common bean grows from sea level up to 2200 m altitude in places where annual rainfall is between 300 and 4300 mm with optimum between 500 and 1500 mm, and where average temperatures range between 15°C and 23°C. The common bean grows well on a large variety of soils with pH ranging from 4 to 9. It does grow better on well-drained, sandy loam, silt loam or clay loam soils, rich in organic content (Ecoport, 2013).

According to FAO, the global bean production has risen from 15.4 million tonnes (Mt) in 1984-1986 (3-year-average) up to the record of 22.81 Mt obtained in 2013 (FAOSTAT, 2015). The current national production of common bean in Ethiopia is estimated at 0.32

million hectares; with a total production of 0.51 million tons and average productivity of 1.6 tons ha^{-1} (CSA, 2015). This is less than attainable yield (3 to 4 tons ha^{-1}) under good management practices (IFPRI, 2010).

The wide gap in yield is attributed to shortage of improved varieties, low soil fertility, drought, disease and insect pest damage especially from the bean stem maggot (BSM) and bean weevil (Bruchids). Diseases like common bacterial blight (*Xanthomonas campestris* pv*phaseoli)* and angular leaf spot (*Phaeoisariopsisgriseola*) pose a significant harvest loss in common bean at farmer's field in Ethiopia (Habtu *et al*., 1996; Fininsa and Tefera 2002; Fininsa and Yuen 2002).

Previous research work on genetic variability and associated traits within common bean genotype has been widely reported by different researchers. For instance, Kassaye (2006) reported high PCV (31.30%) and GCV (22.27%) for pods per plant and the lowest PCV (6.40 %) and GCV (5.93 %) for days to 90 % physiological maturity. High GCV for harvest index (123.98), number of nodes on the main stem (70.04), weight of pods per plant (54.13), grain yield per plant (50.33) also reported by Bagheri *et al*. (2015). Sentayehu (1997) reported that yield per plant was positively and significantly associated with days to maturity, number of primary branches, number of pods per plant, number of seeds per pod, pod length and hundred seed weight. Daniel *et al* (2015) reported that seed yield exhibited positive and significant correlations with leaf chlorophyll content, vertical root pulling resistance, pod harvest index, pods per plant and seeds per pod at both phenotypic and genotypic levels under stress and non-stress conditions.

In Ethiopia common bean research was started in 1970 (Imru, 1985) by introducing some germplasm. Since then much efforts has been made to improve production and productivity of the crop as a result many improved varieties have been released. However, for further improvement of the crop the knowledge of variability and character association is essential. Currently under Ethiopian lowland pulses improvement project, large numbers of common bean genotypes are introduced from the USA (USDA collections). Therefore, the present study is going to generate information on genetic variability and character association of these common beans based on morphological characterization and to fill the following research

gaps: The morphological variability and association among characters in these common bean genotypes is concerned little has been done.

Releasing of highly yielding varieties meeting consumer preferences and quality standards is an obligatory process. In terms of consumer preference, the most desirable traits are those related to the technical and nutritional qualities of the bean, such as the ease of cooking, a good taste, and a soft tegument texture, the ability to produce a thick sauce after cooking and a high protein and mineral content (Santos and Gavilanes, 2006). However, cooking time is certainly considered to be one of the factors that limit the consumption of beans at home. Reduction in cooking time could significantly increase the consumer's interest in the common bean as a food product and have direct and favorable consequences on both production and the commercial market. Furthermore, prolonged cooking times may lead to structural changes at the cellular level and nutrient loss as well as an increased capital cost for the process (Ribeiro *et al*., 2007; Mesquita *et al.,* 2007).

Cooking time is influenced by a number of factors, such as planting time, cultivation practices, high temperature, high or low humidity during bean growth, harvest time, post-harvest handling, storage conditions, and processing technology (Rodrigues *et al*., 2005; Bertoldo *et al.,* 2008; Coelho *et al.,* 2008). Studies have shown a high genetic variability in cooking time trait in common bean (Cichy *et al*., 2015; Elia *et al*., 1997).

Common beans are a food for which improvements in cooking time would especially be valuable (Akibode and Maredia, 2011). It takes 7–11 kg of fuel wood to cook one kg of beans, in contrast to one kg of maize flour, which requires less than one kg of fuel wood to cook (Adkins *et al.,* 2010). Decreasing the cooking times of dry beans would be especially important in areas where beans are consumed as a primary source of protein (Cichy *et al*., 2015).

Cooking time is controlled by a small number of genes (Jacinto-Hernandez *et al*., 2003) and is highly heritable with narrow sense heritability values between 0.74 and 0.90 (Elia *et al*., 1997). In the evaluation of bean accessions, five folds (Cichy *et al*., 2015) two folds (Elia *et al*., 1997) variation in cooking time has been noted. Very few studies have been conducted on

the genetic control of cooking time in beans. Accurate determination of cooking time is difficult and a restricted number of samples can be processed in a day. Faster evaluation methods and more information on the genetic control are needed by breeders (Cichy *et al.*, 2015).

QTL mapping is a key approach for understanding the genetic architecture of traits in plants (Holland, 2007). However, QTL mapping using bi-parental population mapping approach can evaluate only two alleles at a locus, low mapping resolution due to few recombinations, longer time required to develop mapping population (Erena, 2013). The arrival of association mapping approaches has overcome some of the limitations of bi-parental mapping populations. Genome-wide association (GWAS) is a promising approach for scanning the entire genome for detecting marker-trait associations with a large number of markers (Tabor *et al.,* 2002). Moreover, GWAS may assist the identification of genomic regions involved in determining cooking time (Cichy *et al*., 2015).

Knowledge of the genetic variability, naturally occurring diversity in a population helps to identify diverse groups of genotypes that can be useful for the breeding program**.** Little attention was given to cooking time improvement in common beans both at national and regional bean breeding programs. Keeping the above views in mind, the study was initiated to achieve the following objectives:

General Objectives

- To study the variability of common bean genotypes and identify genomic regions involved in determining cooking time using genome-wide association study.

Specific Objectives

1. To estimate variability of selected common bean genotypes.
2. To investigate the association among quantitative traits and
3. To identify simply inherited markers in close proximity to genes affecting cooking time.

2. LITERATURE REVIEW

2.1 Origin, Distribution and Taxonomy of Common Bean

Domestication of common bean took place in two regions distributed from northern Mexico to Colombia (Mesoamerican gene pool) and from southern Peru to northwestern Argentina (Andean gene pool) (Koinange and Gepts, 1992; Freyre *et al.*, 1996). Once domesticated, the common bean was introduced to other regions of the world, whereby both the Mesoamerican and the Andean cultivars were dispersed to lowland South America and Africa. The Mesoamerican cultivars became predominant in the southwestern United States, while the Andean cultivars in Africa, Europe and northeastern United States (Gepts and Debouck, 1991). Domestication in the two regions led to two distinct gene pools (Singh *et al*., 1991b; Becerra Velasquez and Gepts, 1994) because they arose from two already diverged gene pools and selection under domestication (Kwak and Gepts, 2009). The domestication of the common bean has altered the form, morphology, and phenology of the plant, especially the growth habit, seed size, seed retention, and maturity. During domestication, selection was inclined towards smaller, denser plants with short internodes, suppressed climbing ability, fewer and thicker stems and larger leaves (Debouck, 1991). The result of the selection was a compact growth habit of determinate and indeterminate common bean cultivars. However, the most distinct difference between the wild ancestors and the cultivated common bean, are the changes in pod size and the seed size, hence the diversity. The cultivated common beans are also quite diverse in seed size and edible parts such as the green immature pod and dry seed (Debouck, 1991).

The common bean (*Phaseolus vulgaris* L.) belongs to the family Fabaceae (Leguminosae) and the genus *Phaseolus.* The genus *Phaseolus* comprises 70 species. Four sections were classified as Chiapasana, Phaseolus, Minkelersia, and Xanthotricha (Debouck, 1991). The *Phaseolus* section includes four of the cultivated *Phaseolus* species: *P. vulgaris* L. (common bean); *P. coccineus* L. (runner bean); *P. lanatus* L. var. *lanatus* L. (lima bean); and *P. acutifolius* A. gray var. *acutifolius* (tepary bean). Of the four *Phaseolus* species, the common bean is the most widely grown occupying more than 85% of the production area sown to all *Phaseolus* species worldwide (Singh, 2001). Common beans are classified in the sub-phylum

dicotyledons (embryo with two cotyledons, parallel veined leaves and the stem with the vascular bundles arranged irregularly and cambium usually present), division Magnoliophyta, class Magnoliopsida, family Leguminosae, sub-family Papilionoideae or Fabaceae or Lotoideae (pulse family characterized by edible seeds and pods) and order Leguminales. Common beans are diploid (2n = 2x = 22) and are self-pollinated (Rutger and Beckham, 1970; Stoetzer, 1984).

2.2 Genetic Diversity of Common Bean

Genetic diversity is a level of biodiversity that refers to the total number of genetic characteristics in the genetic makeup of a species (Genetic diversity, n.d.). In addition, it refers to any variation in the nucleotides, genes, chromosomes, or whole genomes of organisms (Harrison *et al*, 2004).

The genetic diversity of common bean is mainly in the seed size, which is divided into three groups. The groups include large seeded Andean (>40 g 100-seed weight-1), small seeded Mesoamerican (<25 g 100-seed weight-1), and medium seeded/Middle American (25 to 40 g 100-seed weight-1) gene pools (Evans, 1980). The cultivated gene pools of Andean and Mesoamerican origin were further divided into six races: the Andean (all large seeded) have the races Chile, Nueva Granada, and Peru; Middle American has the races Durango, Jalisco (medium seeded); and Mesoamerican (all small seeded), each of which has its distinguishing characteristics and agronomic traits (Singh *et al.*, 1991a). Common bean is also divided into two groups based on their edible parts: snap beans (French beans or Haricot beans) are consumed as immature pods, and; dry beans are usually consumed as the mature dry seed after rehydration. The snap bean cultivars have a thick succulent mesocarp and have reduced or no fiber in the green pods and sutures (Myers, 2000). The green pods are used as fresh pods, or frozen or canned. There are different market classes of the snap bean cultivars determined by the pod shape (flat, oval or cylindrical), color (dark green, light green, yellow or purple), and the length of the pod. Among the snap bean cultivars, there is a large variation in their growth habits and their adaptation traits (Singh, 2001). Common bean cultivars have

also shown large variations in growth habit, phenological traits, seed color, seed size and shape, as well as canning and cooking qualities (Voysest and Dessert, 1991).

2.3 Cooking Time in Common Bean

Cooking time is one of the most important parameters in evaluating beans for processing quality. Fast and uniformly cooking beans are required both for processing and for traditional consumption of beans by local producers where firewood is the major source of fuel (Elia *et al*., 1996). Beans may fall into four categories according to their cooking times. These categories include soft (with cooking time of 15-19 minutes), normal (20- 30 minutes), semi hard (31-35 minutes), and hard (36 minutes or more) (CIAT, 1989).

For maximum utilization of beans by food-insecure consumers, it is essential to address the long cooking times required to make beans palatable. Food-insecure consumers often are also faced with limited cooking fuel and firewood. Often, when firewood supplies are limited, women change their food consumption patterns and this includes reducing food items that have long cooking times, including beans thereby compromising families' nutritional intakes (Browuer *et al*., 1989). Landraces or varieties that cook quickly are especially valued by consumers and fetch a premium price in the marketplace (Correa *et al.,* 2010). The direct effects are savings in cooking time and fuel costs, mainly for women, who are responsible for cooking (Biran *et al*., 2004). Adoption of fast cooking varieties reduces the quantities of firewood used and the time spent gathering it (Elia *et al.*, 1997; Brouwer *et al.,* 1989). For urban consumers, money will be saved on expensive fuels (kerosene, charcoal, electricity, and natural gas).

In many studies involving cooking of dry beans, seeds are often soaked in water to improve the hydration characteristics of the seed for uniform cooking. It has been reported that well-hydrated seeds generally cook more rapidly than the ones containing less water (Elia *et al*., 1996). Therefore, water absorption is another important attribute affecting the acceptability of beans for processing. In addition to improving the cooking characteristics, high water

absorption of bean seeds has economic implication to the processors. Specifically, the processors require beans with high hydration ratio for high 'can yield' (Ghaderi *et al*., 1984).

Elia *et al*. (1996) have reported significant genetic variations for cooking time, water absorption, protein content, and tannin content among bean genotypes. Biochemical analysis of different cultivars has also revealed highly significant differences for cooking time (Avila-Rodriguez *et al.,* 1996). Ghaderi et al. (1984) also reported highly significant differences among cultivars in 100 seed weight, color of dry and cooked beans, weight of soaked beans, texture, washed drained weight, and processed bean moisture.Elia *et al*. (1996) have studied the genetic control of cooking time and water absorption and found that these traits are under the additive genetic control.

2.4 Genotypic and Phenotypic Variability

Phenotypic variability is defined as the occurrence of differences among individuals due to differences in their genetic composition and of the environment in which they are raised (Falconer and Mackay, 1996; Welsh, 1990). Genetic variability is of prime interest to the plant breeder because proper management of this diversity can produce permanent gain in the performance of the plant and phenotypic variation is made of genotypic value and environmental deviation (Welsh, 1990).

Genotypic and phenotypic conditions are dominant factors for yield formation. Environmental conditions are partially controllable, but genotype of the plant can only be changed by breeding programs (Akçura *et al*., 2005).Phenotypic variability is the total variability, which is observable that includes both genotypic and environmental variation and hence changes under different environmental conditions. Such variation is measured in terms of phenotypic variance. In attempting to develop improved varieties, the plant breeder bases his/her observation often on the measurement of the phenotype. For plant, breeding to be effective there must be phenotypic variation of the desired trait and some of the variation must be heritable from generation to generation (Stoskopf *et al*., 1999).

Raffi & Nath (2004) evaluated thirty-one common bean genotypes to study genetic variability, heritability and genetic advance for yield and yield component traits. The highest genotypic and phenotypic variances were observed for days to maturity and the lowest variation was observed on pod length in common bean. In the case of number of seeds per plant the variation between genotypic and phenotypic variance and coefficient of variation were found high. According to Singh *et al.* (1994) yield per plant and days to flowering showed the highest and lowest phenotypic and genotypic coefficients of variations, respectively in common bean.

2.5 Heritability

Heritability is the extent of contribution of genotype to the phenotypic variation for a trait in a population and it is ordinarily expressed as the ratio of genetic variance to the total variance, i.e., phenotypic variance for a trait. In crop improvement, only genetic component of variation is important since only this component is transmitted to the next generation (Singh, 2001).

Heritability of a character is very high; selection for such character should be easy. This is because there would be a close correspondence between the genotypes and phenotypes due to the relatively small contribution of the environment effect to the phenotypes. However, for characters with low heritability, selection may be considerably difficult or virtually impractical due to the masking effect of the environment (Singh and Ceccerelli, 1996).

In a planned experiment, the broad-sense heritability is computed as: $h^2bs = V_G/V_P$. This reflects all the genetic contributions to a population phenotypic variance including additive, dominant, and epistatic (multi-genic interactions), where individuals are directly affected by their parents' genotype. Whereas narrow sense heritability due to additive (allelic) genetic effects could be computed as $h^2ns = V_A/V_P$ (Falconer, 1989).

Raffi and Nath (2004) reported high heritability values for days to 50% flowering, days to maturity, plant height, number of pods per plant, pod length, number of seeds per plant and 100 seed weight. Sentayehu (1997) reported very high heritability value for 100-seed weight and moderate heritability for number of pods per plant (51%) and number of seeds per plant (50%) was recorded.

According to Kassaye (2006) the highest heritability values recorded were for seeds per pod (65.50%) and hundred seed weight (93.10 %), moderately heritability for pods per plant (50.60%) and low value for biological yield (27.80%), seed yield per plant (24.50%) was recorded.

Bagheri *et al*. (2015) reported a high heritability for number of nodes on the main stem, biological yield per plant, weight of pods per plant, chlorophyll florescent, grain yield per plant, length seed, harvest index, number of days to 50% flowering, number of days to 50% pods and Straw weight.

2.6 Genetic Advance

Genetic advance tell us the estimate of the expected gain for a particular character through selection (Burton and DeVane, 1953).Genetic advance under selection is a genotypic value, which depends on three things (Allard, 1960); genetic variability, heritability and the selection intensity applied. Genetic progress would increase with increase in the variance. Therefore, the utility of estimates of heritability is increased when they are used in conjunction with the selection differential, the amount that the mean of the selected lines exceeds the mean of the entire group (Johnson *et al*., 1955). High amount of genetic advance are suggests of additive gene action while low amounts are suggests of non-additive gene action (Singh and Narayanan, 1993).

According to Kassaye (2006), high heritability estimate coupled with high genetic advance as percent of mean was observed for 100-seed weight, plant height and number of nodes on the main stem. Also Low heritability was accompanied with low genetic advance as percent of mean leaf dimensions, stem diameter, seed yield per plant and biological yield per plant .Higher genetic advance values for hundred seed weight (93.76%) and moderate genetic advance values for number of primary branches (50.86%) and yield per plant (41.9%) (Sentayehu, 1997).

According to Bagheri *et al*. (2015) high heritability estimates along with high genetic advance as percent of mean was observed for number of nodes on the main stem, weight of pods per plant, grain yield per plant, harvest index, straw weight and biological yield pert plant .

However, length seed, number of days to 50% flowering and number of days to 50% pods had high heritability coupled with low genetic advance.

2.7 Correlation and Path Coefficient

2.7.1 Correlation among traits

Correlation coefficient is a statistical measure, which is used to find out the degree (strength) and direction of relationship between two or more variable (Gomez and Gomez, 1984). Three types of correlations are conferring in quantitative genetics and these are phenotypic, genotypic and environmental correlations. The association between two characters that can be directly observed is the correlation of phenotypic values or phenotypic correlations (rp). Phenotypic correlations measure the extent to which the two observed characters are linearly related. It is determined from measurements of the two characters in a number of individuals of the populations. Genetic correlation (rg) is the associations of breeding values (i.e., additive genetic variance) of the two characters. Genetic correlation measures the extent to which degree the same genes or closely linked genes cause co-variation (simultaneous variations) in two different characters. The correlation of environmental deviations together with non-additive genetic deviations (i.e., dominance and epistatic genetic deviations) is referred to as environmental correlations (re) (Singh and Chaudhary, 1977; Falconer and Mackay, 1996).

Correlations between different characters of crop plants may arise either from genotypic or environmental factors. Environmental correlations arise from the effect of overall environmental factors that vary at different environments. Correlations due to genetic causes are mainly pleiotropic effects of genes and linkage (a phenomenon of genes inherited together) between genes affecting different characters. Pleiotropy is the property of a gene, which affects two or more characters; as a result, it causes simultaneous variations in the two characters when the genes are segregating (Singh, 1993; Falconer and Mackay, 1996).

Knowledge of correlations that exists between important characters may facilitate the interpretation of result obtained and provide the basis for planning more efficient program for future (Johnson *et al*., 1955). Inadequate knowledge of interrelationships among various traits and the practice of unilateral selection for agronomic traits frequently end up with less than

optimum result in plant breeding (Bhatt, 1973). The practical utility of selecting for a given character as a means of improving another depends on the extent to which improvement in major characters is facilitated by selection for the indicators. Such improvement depends on not only the genotypic correlation but also phenotypic correlation (Johnson *et al.*, 1955b). For selection based on yield component to be effective in increasing yield, Sidwell *et al.* (1967) stated that the components should fulfill the following: they should be highly heritable, the component should be genotypically independent or genotypic correlation among the component should be positive and the component should be physiologically related in a positive manner.

Daniel *et al* (2015) reported that, seed yield exhibited positive and significant correlations with leaf chlorophyll content, vertical root pulling resistance, pod harvest index, pods per plant and seeds per pod at both phenotypic and genotypic levels under stress and non-stress conditions. Kassaye (2006) reported the positive and significant correlation of seed yield with biomass yield (r_p = 0.94 and r_g = 0. 89) and number of pods per plant (r_p = 0.73 and r_g = 0.57).

Singh *et al.* (1995) reported that seed yield was positively associated with seed size. Sentayehu (1997) reported that, yield per plant was positively and significantly associated with days to maturity, number of primary branches, number of pods per plant, number of seeds per pod, pod length and hundred seed weight. They also reported highly significant and negative correlation between hundred seed weight and seed yield per plant.

Akhshi *et al.* (2015) showed that seed yield had a strong positive correlation with both seed number per plant and seed number per pod. A positive association between grain yield and biological yield pert plant were reported by Bagheri *et al*. (2015).

2.8 Path Coefficient

A path coefficient measures the direct influence of one variable upon another and permits the separation of correlation coefficient into components of direct and indirect effects. Path coefficient analysis specifies the cause and measures the relative importance of the characters, while correlation measures only mutual association without considering causation (Dewey

and Lu, 1959) and thus helps breeder in determining the yield components and understanding cause of association between two variables (Shukla *et al.,* 2006).

In any breeding program of complex characters such as yield for which direct selection is not effective, it becomes essential to measure the contribution of each of the component variables to the observed correlation and to partition the correlation into components of direct and indirect effect (Giriraji and Vijayakumar, 1974). Path analysis has proven useful in providing additional information that describes cause and effect relationships, such as between yield and yield components (Gravios and Helm, 1992). It is, therefore, essential to assess the importance as well as degree of association of various quantitative characters in order to initiate an effective selection program aimed at genetic improvement of crop yield.

To improve grain yield via selection of its components path coefficient analysis is a useful tool for understanding grain yield formation and provides valuable additional information about the traits (Garcia *et al*., 2003).

Pods per plant and seeds per pod had high positive direct effects on seed yield both under stress and non-stress condition; whereas, pods per plant had the highest indirect effect on seed yield through pod harvest index under stress condition as reported by Daniel *et al*. (2015).

Sentayehu (1997) reported that yield per plant was positively and significantly associated with days to maturity, number of primary branches; number of pods per plant, number of seeds per pod, pod length, and hundred seed weight. They also reported highly significant and negatively correlated between hundred seed weight and seed yield per plant.

Kassaye (2006) reported that number of pods per plant had the highest positive direct effect (0.96) on seed yield, followed by 100-seed weight and number of seeds per pod. Internode length had low and negative direct effect on seed yield. Pod length, plant height, number of nodes on the main stem, days to 50% flowering; days to 90% maturity and protein content recorded low positive direct effect, while harvest index exhibited a moderate direct effect.

Sadeghi *et al.* (2011) found that seed number per plant and harvest index had a high positive direct effect on seed yield in common bean. Seed number per plant had a positive and significant direct effect on seed yield in bean; whereas, seed weight per plant had a negative and significant direct effect on the seed yield (Cokkizgin *et al.,* 2013).

Bagheri *et al*. (2015) reported that weight of pods per plant had the highest (1.313) direct effect on grain yield. The indirect effects of biological yield per plant through weight of pods per plant were high (0.85).

2.9 Clustering

Clustering is defined as the process of organizing genotypes or individuals into groups whose members are similar in some way (Chahal *et al.*, 2002). There are broadly two types of clustering methods: 1) Distance-based methods, in which a pair wise distance matrix is used as input for clustering analysis. The result can be visualized as a tree or dendrogram in which clusters may be identified. And 2) Model-based methods, in which observation from each cluster are assumed to be random draws from some parametric model, and inference about parameters corresponding to each cluster and cluster membership of each individual are performed jointly using maximum-likelihood or Bayesian methods (Johnson and Wichern, 1992).

Another important aspect in cluster analysis is determining the optimal number of clusters or number of acceptable clusters. In essence, this involves deciding where to “cut” a dendrogram to find the true or natural groups. An “acceptable cluster” is defined as “a group of two or more genotypes with a within-cluster genetic distance less than the overall mean genetic distance and between cluster distances greater than their within cluster distance of the two clusters involved” (Mohammadi and Prasanna, 2003).

The resulting clusters of individuals should then exhibit high internal (within cluster) homogeneity and high external (between clusters) heterogeneity. Thus, if the classification is successful, individuals within a cluster shall be closer when plotted geometrically and different clusters shall be farther apart (Hair *et al.,* 1995).

Getachew (2010) grouped 36 common bean accessions into five clusters. Kassaye (2006) grouped 114 common bean genotypes into nine clusters, which makes them divergent. Lima *et al*. (2012) determined the genetic diversity of 100 genotypes of common bean and by using cluster analysis, *eight* clusters were determined. Molosiwa *et al*. (2014) clustered 9 bean accessions into three groups based on performance of genotypes. Awan (2014) obtained three groups from 13 common bean genotypes, each group having differing number of cultivars when analyzing using dendrogram.

2.10 Principal Component

Principal components analysis is a method for transforming the variables in a multivariate dataset into new variables which are uncorrelated with each other and which account for decreasing proportions of the total variance of the original variables. Each new variable is defined as a particular linear combination of the original variables. Full accounts of the method are given in Everitt and Dunn (2001). Given a data set with p numeric variables, you can compute p principal components. Each principal component is a linear combination of the original variables, with coefficients equal to the eigenvectors of the correlation or covariance matrix. The eigenvectors are customarily taken with unit length. The principal components are sorted by descending order of the eigenvalues, which are equal to the variances of the components (Michael, 2012). The first step in PCA is to calculate Eigenvalue, which define the amount of total variation that is displayed on the PC axes. The first PC summarizes most of the variability present in the original data relative to all remaining PCs. The second PC explains most of the variability not summarized by the first PC and uncorrelated with the first, and so on (Jolliffe, 2002).

Stoilova *et al.* (2013) reported variation in 15 Portuguese and 15 Bulgarian common bean landraces in which the first four PCs contributed 82.61 % of total variation.Assefa *et al.* (2014), reported that the combination of first three principal components explained more than 50% of the genotypic variations. Morojele and Mbewe (2015) by using principal component analysis identified the characters, which caused major variation among cultivars. Out of 10 principal components generated from 17 characters, only the first three components that constituted 54.57% of the total variation.

2.11 Molecular Markers and Mapping

2.11.1 Marker systems used in crop improvement strategies

Identification of different genotypes of crop species is essential in characterizing accessions of crop species, for registration of newly developed cultivars and to determine the purity of varieties (Malik *et al*., 2008) and this could be achieved through identification, characterization and quantification of traits or markers. Traits that serve as genetic markers are by definition polymorphic; the more polymorphic the trait, the greater its potential value to germplasm management (Seetharam *et al*., 2009). Markers are entities that are heritable as simple Mendelian traits and are easy to score (Schulman *et al*., 2004). Three marker systems could be recognized- morphological, biochemical and molecular marker systems- in crop improvement strategies.

2.11.2 Molecular markers

Molecular markers are based on naturally occurring polymorphisms in DNA sequences (i.e. base pair deletion, substitutions, additions or patterns) (Gupta *et al*., 1999). They are superior to both morphological and biochemical markers because they are relatively, abundant through the genome even in a highly- inbred cultivars, completely independent of environmental conditions and can be detected at virtually any stage of plant development (O' Neill *et al*., 2003). DNA based markers provide very effective and reliable tool for measuring genetic diversity in crop germplasm and studying evolutionary relationships than other available techniques for assessing the genetic variability and relatedness among crop germplasm (Malik *et al*., 2008). They can be applied in the identification of cultivars and clones, genetic mapping, marker assisted selection (MAS), population genetics, molecular systematics etc (Weising *et al*., 2005). According to Vithanage *et al.* (1995) molecular markers are 'land marks' which can be identified on the genome and, therefore, offer the best possible means of identifying individuals from biological samples. In recent years, different marker systems such as RFLP, RAPD, ISSR, STS, AFLP, SSR, SNPs, and others have been developed and applied to a range of crops (FAO, 2003).

2.11.3 Single Nucleotide Polymorphism

SNPs have become the marker of choice for crop genetics and breeding applications because of their high abundance in genomes, and the availability of a wide array of genotyping platforms with various multiplex capabilities for SNP analysis (Rafalski, 2002).

The particular germplasm and the number of populations used for the SNP discovery panel will affect the levels of SNP polymorphism and allele frequency distribution detected in independent populations; this is known as ascertainment bias (Clark *et al.,* 2005). In general, the principles underlying the SNP assays involve generating allele-specific products in biochemical reactions and identifying the products using detection procedures (Chen and Sullivan, 2003; Syvanen, 2001). Currently, a wide variety of SNP genotyping systems that use different chemistries and detection systems to assay SNPs are commercially available. Moreover, the number of SNP genotypes scored per reaction for each sample ranges from one to over one million on different platforms (Perkel, 2008; Ragoussis, 2006). The choice of a suitable genotyping platform, thus, will depend on the user's criteria, including the flexibility of the systems, desired data throughput, the applications, and per sample cost (Chen and Sullivan, 2003).

The technological improvements in sequencing and SNP genotyping have resulted in the generation of a wealth of information. As the cost continues to drop for whole-genome sequencing (Podolak, 2010), genetic mapping and genotyping by re sequencing are now possible for plants with complete genomic sequences (Huang *et al.,* 2009). For the majority of the crops, however, it is more feasible to develop genome-wide SNP markers using the NGS techniques. Such marker resources will be highly valuable for dissecting the genetic architecture of complex agronomic traits and facilitating GS. The development of highly parallel SNP assays, such as Illumina's Golden Gate assay with 1536-plex platform (Fan *et al.,* 2003), enabled the genome-wide studies previously not feasible for economically important crops.

2.11.4 Quantitative trait loci mapping (QTL)

In crop plants, the standard mapping populations are derived from crosses between two

parents which have contrasting characters of a trait under investigation; for example, drought tolerant versus drought susceptible parents. These bi-parental cross populations have been used for determining the number, effect size and chromosomal locations of QTL underlying agriculturally important quantitative traits. Some of the advantages of bi-parental populations include the requirement of relatively fewer markers for genome coverage, no population structure and ability to locate QTL regions along chromosomes (Sorrells and Yu, 2009). The disadvantages of bi-parental population mapping approach are: only two alleles can be evaluated at a locus, low mapping resolution due to few recombinations, longer time required to develop mapping population (Erena, 2013).

Several researchers identified the QTL of different traits in common bean. Daniel *et al*. (2015) identified a total of 11 Quantitative trait loci for pod harvest index, root pulling resistance, grain yield and bean stem maggot damage scores in recombinant inbred lines from a cross of BAT881 and G21212 using SSR and SNP markers. Asrat *et al*. (2012) identified a total of 15 putative QTL for seven rooting pattern traits and four shoot traits from a recombinant inbred line population of DOR364 and BAT477. Augusta *et al.* (2012) worked out QTL mapping for the cooking time of common beans. Six significant QTLs with an LOD C 3.0 were found for the cooking time of the F2:4 and F2:5 generations. Mukeshimana *et al.* (2014) recorded a total of 14 QTL associated with yield and yield components (number of pods per plant, SW), pod harvest index, and phonology (number of days to flower and maturity) were consistently identified mainly on six chromosomes (Pv01, Pv03, Pv04, Pv07, Pv08, and Pv09) of the eleven common bean chromosomes. Beattie *et al.* (2003) mapped QTL for number of pods per plant on Pv03. Blair *et al*. (2006) reported the identification of QTL in an advanced backcross population from a cross of an Andean bean ICA Cerinza with a wild bean accession G24404, Quantitative trait loci for pods per plant was mapped on Pv07, Pv09, and Pv11. Kamfwa *et al.* (2013a) conducted the Identification of QTL for Fusarium Root Rot Resistance in a resistant line MLB-49-89A of common bean. The major QTL appears to map close to the same region of Pv03. A major QTL with a logarithm of odds (LOD) score of 6.1 and R^2 of 34% was detected between PVBR87 and PVBR109 markers in the K132 population.

2.11.5 Association mapping

The classical method of QTL identification is conducted by a bi-parental QTL mapping approach. Association analysis which does not require development of a bi-parental mapping population is becoming a common method of QTL mapping mainly due to its high resolution, broader allele coverage and cost effectiveness. In this method, diverse lines or cultivars can be used for obtaining information on marker-trait associations. It has the potential to identify QTL associated with a desired trait and even to detect the causal polymorphisms within a gene that are responsible for the difference in two alternative phenotypes (Gupta *et al*., 2005).

The resolution of QTL is high as only closely linked alleles are in LD due to a long history of recombination (Ingvarsson and Street, 2011). Association mapping is also useful for establishing associations between haplotype blocks and traits of interest. However, genomic locations of QTL detected by the association mapping approach need to be inferred from a consensus genetic map and/or physical map for the crop under study. Special mapping populations known as Nested Association Mapping (NAM) populations allow simultaneous QTL detection and chromosomal position determination (Ersoz and Buckler, 2009).

Association mapping broadly falls into two major classes: (1) genome-wide association mapping, which surveys genetic variation in the whole genome using a large number of markers to detect regions associated with the phenotype (Zhu *et al*., 2008); and (2) candidate-gene association mapping, which relates within candidate gene polymorphisms with phenotypic variations of the traits. The choice between whole genome scanning and candidate gene approaches depends on the extent of LD in the population and the availability of markers. Although genome-wide association is a promising approach for scanning the entire genome for detecting marker-trait associations with a large number of markers, the candidate gene approach is also important to map targeted genes with known function (Tabor *et al.,* 2002).The association mapping approach has been used for several crops to identify QTL and also to characterize candidate genes (Erena, 2013).

Shi *et al.* (2011) carried out Association mapping of common bacterial blight resistance QTL in Ontario bean breeding populations. Eighteen and 22 markers were significantly associated with CBB rating at 14 and 21 days after inoculation, respectively. Fourteen markers were significant for both dates and the markers UBC420, SU91, g321, g471, and g796 were highly significant. Nemli *etal*. (2014) conducted Association mapping for five agronomic traits in the common bean. Sixty-two marker–trait associations were identified for the five traits.

2.11.5.1 Candidate genes association mapping

Candidate-gene association mapping is a hypothesis driven approach to complex trait dissection, with biologically relevant candidates selected and ranked based on the evaluation of available results from genetic, biochemical, or physiology studies in model and non-model plant species (Risch and Merikangas, 1996). Because SNPs offer the highest resolution for mapping QTL and are potentially in LD with the causative polymorphism, they are the preferential candidate-gene variant to genotype in association studies (Rafalski, 2002).

Candidate-gene association mapping requires the identification of SNPs between lines and within specific genes. Therefore, the most straightforward method of identifying candidate gene SNPs relies on the resequencing of amplicons from several genetically distinct individuals of a larger association population. Fewer diverse individuals in the SNP discovery panel are needed to identify common SNPs, whereas many more are needed to identify rarer SNPs. Promoter, intron, exon, and 5′/3′-untranslated regions are all reasonable targets for identifying candidate gene SNPs, with non-coding regions expected to have higher levels of nucleotide diversity than coding regions. The rate of LD decay for a specific candidate gene locus dictates the number of SNPs per unit length (e.g., kb) needed to identify significant associations (Whitt and Buckler, 2003). Therefore, the number and base-pair length of amplicons required to sufficiently sample a candidate gene locus is almost entirely dependent on LD and SNP distribution, with a higher density of SNP markers needed in regions of relatively low LD and high nucleotide diversity (Zhu *et al.,* 2008).

Burt *et al*. (2015) studied the Identification of Candidate Gene with SNP Marker-Based Fine Mapping of Anthracnose Resistance Gene Co-4. *Co-42* is the most effective anthracnose

resistance gene characterized to date (Balardin and Kelly, 1998). Alleles at the Co–4 locus confer resistance to a number of races of *C. lindemuthianum*. A population of 94 F4:5 recombinant inbred lines of a cross between resistant black bean genotype B09197 and susceptible navy bean cultivar Nautica was used to identify markers associated with resistance in bean chromosome 8 (Pv08) where *Co–4* is localized. Thirty-two unique annotated candidate genes were identified that spanned a physical region of 936.46 kb.

2.11.5.2 Genome wide association mapping

Genome-wide association mapping is becoming a widespread method to identify quantitative trait loci (QTL). Its benefit over traditional bi-parental mapping approaches depends on the extent of linkage disequilibrium (LD) in the mapping population and dense marker coverage across the genome (Zhu *et al*., 2008).

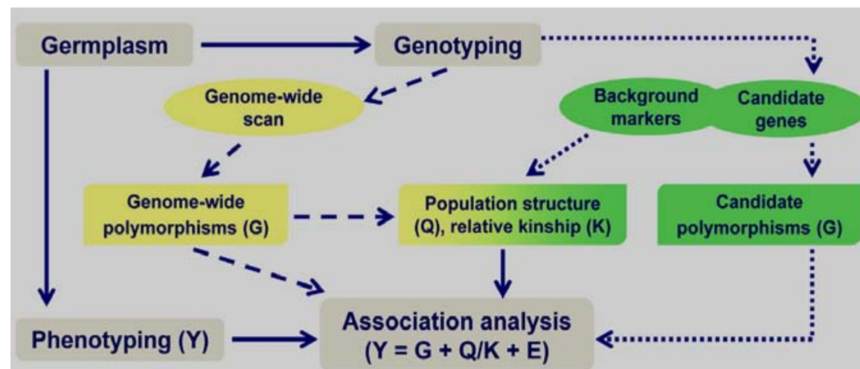


Figure 1: Schematic diagram and contrast of genome-wide association mapping and candidate-gene association mapping.

The inclusion of population structure (Q), relative kinship (K), or both in final association analysis depend on the genetic relationship of the association-mapping panel and the divergence of the trait examined. E stands for residual variance (Zhu *et al.,* 2008).

Very few studies have been conducted on genome wide study in common bean. Cichy *et al* .(2015), conducted genome- wide association analysis (GWAS) to identify genomic regions influencing cooking time trait, and to test the ability to predict cooking time by raw seed characteristics on a panel of 206 *P. vulgaris* accessions. Based on the result, GWAS revealed regions on chromosomes Pv02, Pv03, and Pv06 associated with cooking

time. Significant SNPs associated with cooking time were found on chromosomes Pv02, Pv03, and Pv06. Kamfwa *et al*. (2015c) conducted a genome-wide association study (GWAS) to explore the genetic basis of variation for symbiotic nitrogen fixation (SNF) and related traits in the Andean Diversity Panel (ADP). He reported a Significant SNPs and candidate genes for symbiotic nitrogen fixation (SNF) and related traits were identified on Pv03, Pv07 and Pv09 chromosomes of common bean.

Perseguini *et al.* (2016), conducted Genome-Wide Association Studies of Anthracnose and Angular Leaf Spot Resistance in Common Bean and reported Using SNPs, 21 and 17 significant statistically associations were obtained for Anthracnose and Angular Leaf Spot, respectively, providing more associations with this marker. The SSR-IAC167 and PvM95 markers, both located on chromosome Pv03, and the SNP scaffold00021_89379, were associated with both diseases. The other markers were distributed across the entire common bean genome, with chromosomes Pv03 and Pv08 showing the greatest number of loci associated with ANT resistance. The chromosome Pv04 was the most saturated one, with six markers associated with ALS resistance. The telomeric region of this chromosome showed four markers located between approximately 2.5 Mb and 4.4 Mb. The results demonstrate the great potential of genome-wide association studies to identify QRLs related to ANT and ALS in common bean.

Shi *et al.* (2011) reported Fourteen markers were significant for both days after inoculation (DAI) and five markers were highly significant, including UBC420 and SU91 after conducting genome-wide association mapping in a bean breeding population to detect the markers associated to CBB resistance. Kamfwa *et al*. (2015b), after conducting Genome-Wide association study of agronomic traits in common bean, significant SNP markers associated with several agronomic traits were identified. Significant SNPs for seed yield were identified on Pv03 and Pv09.

3. MATERIALS AND METHODS

3.1 Experimental Site Description

The field experiment was conducted at Hawassa Agricultural Research Center, which is located in South Nations, Nationalities and People's Regional State (SNNPR) of Ethiopia. The site is located 275 km away from Addis Ababa and found at $7^{o}03$'N and $38^{o}30$' E at the elevation of 1650m.a.s.l. The bimodal rainfall has an annual mean of 1021 mm with 36.4 and 51.3% received as "Belg and Meher" seasons rainfall, respectively. The daily average maximum and minimum temperatures of the site during the growing season were 26° and 12.7^{o}, respectively. The soil of the experimental site is clay loam with pH 7.0.

3.2 Experimental Material

The experimental material comprised a total of 418 genotypes of Common been including five standard checks (REMEDA, TATU, IBADO, DAB302 and DAB489). The experimental materials were obtained from Hawassa Agricultural Research Center. The genotypes were USDA collections; 94 entries from Africa, 31 from CIAT core collection, 71 from US core collection, 69 from Caribbean, 50 from Ecuador, 60 from CIAT Africa, and 48 from Northern America.

3.3 Field Experimental Design and Trial management

The experiment was carried out from August 2015 to December 2015 of cropping season in an Augmented Block Design (ABD) comprising of 21 blocks, where each block contains 21 test entries and 5 checks (randomly allocated) with the total of 26 germplasm in each block. Each germplasm was sown in two rows with a spacing of 0.4 meter and 0.1m inter row and between plant spacing, respectively**.** Each incomplete block was spaced 1 meter. As per the national recommendation, a seed rate of 90-kg/ha i.e. 6.48 g per plot was used. Across both the treatments, a total of, 100 kg/ha DAP (Di ammonium Phosphate) fertilizer was applied at planting. Hand weeding was used for weed control, and all other agronomic practices were undertaken uniformly to the entire plot.

3.4 Data collected

3.4.1. Data collected on plant basis

3.4.1.1 Number of pods per plant: The number of pods per plant were counted in five randomly taken plants and expressed as an average for each plot.

3.4.1.2 Number of seeds per pod: Counted from five random pods from each of five randomly taken plants per plot and expressed as an average of five plants per plot.

3.4.1.3 Pod harvest index (%): For recording pod harvest index pods were collected from five plants per plot at harvest and separated into seeds and pod wall. The pod wall was oven dried at 80 ºC for 48 hours while the moisture content of the seeds was determined with seed moisture tester and adjusted to 0% moisture content and the weights recorded. PHI was then calculated as suggested by Beebe *et al.* (2013).

$$\frac{\text{seed weight (0\% moisture)}}{\text{dry weight of pod (seed + pod wall)}} \times 100$$

3.4.1.4 Plant height (cm): Height of five plant of the main stem from the ground level to the top of the main stem was measured.

3.4.1.5 Vertical root pulling force resistance (Ib): Was measured on 3 plants per plot using a DS2 digital force gauge (IMADA Inc).

3.4.1.6 Leaf chlorophyll content: Was measured on ten fully expanded mature plants using a SPAD-502 chlorophyll meter

3.4.2 Data collected on plot basis

3.4.2.1. Days to 50% flowering: This was taken as the number of days from planting to when 50% of the plants in the plot were in bloom.

3.4.2.2. Days to 90% maturity: Number of days from sowing to the stage when 90% of the plants in a plot have changed the color of their pods from green to lemon yellow.

3.4.2.3. Hundred seed weight: Determined from the average 100-seeds mass at (12-14%) moisture content of the seed and expressed in grams.

3.4.2.4. Seed yield per hectare: A dried plant was threshed separately and seeds obtained from them were weighted. Yield was corrected based on seed moisture content determined with a seed moisture meter. The plot yield was converted to yield per hectare after adjusting to 12% moisture content.

3.4.3 Water uptake and cooking time

3.4.3.1 Water uptake

A sub sample of approximately 200 seeds was randomly selected and placed in a 6.5 × 14 cm enveloped and stored at room temperature. To equilibrate the moisture content before cooking, samples were held in hermetic storage over a saturated salt solution (63 % relative humidity) within a controlled atmosphere storage cabinet prior to cooking. When the moisture content reached 10–14 %, as it is determined with/by a moisture meter (Moisture Check Plus, Deere and Company, Moline, IL), 30 seeds were selected per entry, which were weighed, and soaked in (1:4) w/w distilled water for 12 h. Seeds were drained, blotted dry, and weighed again to determine percentage of water uptake during soaking with the equation of Cichy *et al.* (2015).

$$\frac{\text{Seed Weigt After Soak} - \text{Seed Weight Before Soak}}{\text{Seed Weigt After Soak}} \text{X100}$$

3.4.3.2 Cooking time

Cooking time was measured on 25 soaked seeds with a Mattson cooker in boiling distilled water. The Mattson Cooker consists of a plate with 25 wells for individual seeds and on top of each well rests a metal pin (Wang and Daun, 2005). Cooking time was recorded as the time it

took for 80 % of the seeds to be completely pierced with an 85 g stainless steel rod with a 2-mm pin (Cichy *et al.*2015).

3.5 Statistical Analysis

3.5.1 Analysis of variance (ANOVA)

The quantitative data collected from the location for all the parameters were subjected to analysis of variance (ANOVA) using the SAS (v 9.4). TUKEY was employed to identify genotypes that are significantly different from each other. The analysis was carried out according to the following model (Federer, 1956).

$$Y_{ij} = \mu + gi + cj + \beta j + \varepsilon ij$$

Where: Y_{ij} is the observation of treatment i in j^{th} block μ is the general mean, g_i is the effect of test treatment, c_j is the effect of control treatments in j^{th} block, βj is the block effects, (ε) is the error term.

Table 1: Skeleton of Analysis of variance (ANOVA) for augmented design

Source of variation	Df	SS	MS	F-value
Block(adj)	(b-1)	SSB	MSB	MSB/MSE
Trt (adj)	(c+g)-1	SSt	MSTrt	MSt/MSE
Among-controls	(c-1)	SSC	MSc	MSC/MSE
Among-test	(g-1)	SSG	MSG	MSG/MSE
Test-v-Control	1		SSE/(c-1)(b-1)	
Error	(b-1)(c-1)			

Where: b = number of block, C = check varieties, g = genotype, df = degree of freedom, SS = sum square, MS = mean square, SSB and MSB are sum square and mean square of blocks, respectively; SSG and MS_G are sums of squares of genotypes and mean square of genotype, respectively; SSC and MSC are sum square and mean square of check variety, respectively; SSt and MSt are sum square and mean square of treatment, respectively.

3.5.2 Analysis of genetic parameters

3.5.2.1 Estimation of variance components

The phenotypic, genotypic and environmental variances and coefficients of variation were estimated based on the method suggested by (Singh and Chaudhury 1985).

Environmental variance (σ^2_e) = Error mean square

Genotypic variance $(\sigma^2 g) = \frac{MSG - \sigma^2 e}{r}$

Where, MS_G is mean square due to genotypes

(σ^2_e) is mean square of error

Phenotypic variance $\delta^2 P = \delta^2 g + Mse$

Where, $\sigma^2 g$ is genotypic variance

$\sigma^2 e$ is environmental variance.

Phenotypic coefficient of variation $PCV = \frac{\sqrt{\delta^2 P}}{\overline{X}} * 100$

Genotypic coefficient of variation $GCV = \frac{\sqrt{\delta^2 g}}{\overline{X}} * 100$

Where: $\sigma^2 p$ =Phenotypic variation

$\sigma^2 g$= Genotypic variation and

$\bar{x}$= Grand mean of the characters under study

3.5.2.2 Broad-sense heritability ($h^2 b$)

Broad sense heritability for all characters was estimated as the ratio of genotypic variance to the phenotypic variance and expressed in percentage according to the methods suggested by Allard (1999).

$$h^2 b = \frac{\sigma 2g}{\sigma 2p} X100$$

Where, h^2b = Heritability in broad sense, σ^2p =Phenotypic variation; σ^2g= Genotypic variation

3.5.2.3 Genetic advance under selection (GA)

The expected genetic advance for each character at 5% selection intensity was estimated using the methodology described by Allard (1999) as:

$$(GA) = h^2b\ \sigma pK$$

Where: GA = expected genetic advance.

h^2b = heritability in broad sense and

K = the selection differential at 5% selection intensity (K = 2.063),

$\sigma_{p=}$ phenotypic variance

3.5.2.4 Genetic advance as percent of mean

Genetic advance as percent of mean was calculated to compare the extent of predicted advance of different traits under selection, using the following formula:

$$(GAM) = \frac{GA}{X} X100$$

Where: GAM=Genetic advance as percent of mean.

GA=Genetic advance under selection and

$\bar{x}$=Grand Mean of the trait

3.5.3 Correlation and path coefficient analysis

3.5.3.1 Correlation coefficient (r)

The phenotypic and genotypic correlation coefficient were computed by the method described by Singh and Chaundry (1985).

Phenotypic correlation is given by

$$r_p = \frac{pcov\ x.y}{\sqrt{\delta^2 px.\delta^2 py}}$$

Genotypic correlation is given by

$$r_g = \frac{g\text{cov } x.y}{\sqrt{\delta^2 gx.\delta^2 gy}}$$

Where: r_p and r_g are phenotypic and genotypic correlation coefficients, respectively.

pcov x.y and g cov x.y are phenotypic and genotypic and covariance between variables x and y, respectively.

δ^2px and δ^2gx are phenotypic and genotypic, variances for variable x

δ^2py and δ^2gy are phenotypic and genotypic variances for the variable

The coefficients of correlation were tested using 'r' tabulated value at n-2 degrees of freedom, at 5% and 1% probability level, where n is the number of treatments (genotypes).

3.5.3.2 Path coefficient analysis

Path coefficient analysis was computed with the formula provided by Dewey and Lu (1959).

$$rij = Pij + \sum rik * Pkj$$

Where: r_{ij} = mutual association between the independent character i (yield-related trait) and dependent character, j (grain yield) as measured by the genotypic correlation coefficients P_{ij} = components of direct effects of the independent character (i) on the dependent character (j) as measured by the path coefficients and

$\sum r_{ik} p_{kj}$ = summation of components of indirect effects of a given independent character (i) on a given dependent character (j) *via* all other independent characters (k).

Whereas the contribution of the remaining unknown characters is measured as the residual which is calculated as:

$$P_R = \sqrt{(1 - \sum P_{ij} \ r_{ij})}$$

3.5.4 Cluster analysis

Clustering was performed using the proc cluster procedure of SAS version 9.4 software (SAS, 2014) by employing the method of average linkage clustering strategy of the observation. The numbers of clusters were determined by following the approach suggested by Copper and Miligan (1988). It was done by looking into three statistics namely Pseudo F, Pseudo t^2 and cubic clustering criteria.

3.5.5 Genetic divergence analysis

Genetic divergence between clusters was determined using the generalized Mahalanobis's D^2 statistics (Mahalanobis, 1936). The analysis was based on all yield contributing characters. In matrix notation, the distance between any two groups was estimated from the following relationship.

$$D^2_{ij} = (Xi - Xj)\, S^{-1} (Xi - Xj)$$

Where D^2_{ij} = the square distance between any two genotypes i and j.

X_i and X_j = the vectors for the values for accession i^{th} and j^{th} genotypes and

S^{-1} = the inverse of pooled variance covariance matrix within groups.

Testing the significance of the squared distance values obtained for a pair of clusters was taken as the calculated value of χ^2 (chi-square) and tested against the tabulated χ^2 values at n-2 degree of freedom at 1% and 5% probability level, where n= number of characters used for clustering genotypes.

3.5.6 Principal component analysis (PCA)

The genetic divergence calculation of principal component analysis was carried out using the proc princomp procedure of SAS version 9.4 software (SAS, 2014). Statistical inference was computed by taking into account all the factors at a time. In this study investigation of suitable multivariate technique for analyzing data for all the characters were proposed. The general formula to compute scores on the first component extracted (created) in a principal component analysis.

$$PC1 = b11(X1) + b12 + \cdots b1p = (Xp)$$

Where: PC1 = the subject's score on principal component 1 (the first component extracted).

b1p = the regression coefficient (or weight) for observed variable p, as used in creating principal component 1and

Xp = the subject's score on observed variable p.

3.5.7 Genome-Wide Association analysis

Marker genotype data from SNP markers was used in the analysis. Publicly available GBS data were used. The 418 genotype were genotyped with the Illumina (Illumina Inc., San Diego, CA, USA) BARC Bean6K_3 SNP array of 4,416 SNP markers distributed across the 11 pairs of common bean chromosomes as described by Cichy *et.al*, (2015). After filtering a minor allele frequency of 2 % or less from 5398 SNP markers, 4,416 SNPs remained for GWAS analysis with a mixed linear model (MLM).

Population structure was accounted for in the model with Principal Component Analysis (PCA) using a correlation matrix in the program TASSEL 5.0 (Bradbury *et al*. 2007) and two PCs were included, which explained 39.8% of the variance (PC1 26.5 %, PC2 13.3%). A kinship matrix (K) was also included in the association analysis to account for relatedness. The QQ plot was generated from the observed and expected LOD scores for each trait. Manhattan plot and QQ plot graphics were developed in qqman in R (Turner, 2014). The phenotypic data for cooking time and water uptake were collected 2016 field seasons prior to Genome-Wide Association. The following MLM equation was used:

$$Y = X\alpha + P\beta + K\mu + e$$

Where Y is phenotype, X is SNP, p is the PCA matrix and both X and p represents fixed effects, K is the relative kinship matrix value, and e is for residual effects. The cut off used for significant SNP markers for each trait were the False Discovery Rate (FDR) (Benjamini and Hochberg, 1995) determined in using Bioconductor in R (Gentleman *et al*. 2004). Candidate genes were identified based on proximity to significant SNPs using the *P. vulgaris* reference genome (Schmutz *et al*., 2014).

4. RESULTS AND DISCUSSION

4.1 Variability Assessment

4.1.1 Analysis of variance (ANOVA)

The analysis of variance (ANOVA) for different characters is presented in Appendix Table 1. There were very highly significant differences (P< 0.001) among genotypes for days to flowering, days to maturity and grain yield hectare^{-1} . Characters, like plant height, number of seeds pod^{-1} and root pulling force resistance was significant (p< 0.05) and leaf chlorophyll content, pod harvest index and hundred seeds weight were non-significant.

Among tests, significant difference was observed for days to flowering, days to maturity, plant height, number of seeds pod^{-1} and grain yield hectare^{-1}. Whereas among control group, significant differences were observed for all characters except number of pod plant^{-1}, leaf chlorophyll content and pod harvest index. Among test versus control, significant difference was also observed for days to flowering, days to 90% maturity, leaf chlorophyll, and root pulling force resistance and hundred seed weight grain yield hectare^{-1}.

Rezene *et al.* (2013) evaluated a total of 49 small red seeded common bean genotypes under stresses and none stresses environment. Similar results were reported for days to flowering, days to maturity and pods per plant, but opposite results were reported for pod harvest index (%), leaf chlorophyll content and hundred seed weight. Semiha and Huseyin (2007) reported significant differences among 10 genotypes for 7 morphological traits such as seed yield, 100 seed weight, pod number plant^{-1}, plant height , seed number pod^{-1}, seed number plant^{-1} and grain yield plot^{-1}.

4.1.2 Range and mean of different characteristics

Range and mean for the eight characters are presented in Tables 2. The mean performance of some common bean genotype for eight characters is presented in Appendix Table 2. The mean values for days to flowering and days to maturity ranged from 42 (ADP0579) to 55 (ADP0647), 68 (ADP0579) to 78 (ADP0588), respectively. Plant height was varies from

18.41 (ADP0785) to 86.7 (ADP0780), number of pod per plant was ranged from5.66 (ADP0454) to 29.6 (ADP0793). Number of seeds pod^{-1} was ranged from 2 (ADP0596) to 6.6 (ADP0508), vertical root pulling force resistance was ranged from 4.27 (ADP0546) to 73.77 (DAB489), whereas seed yield hectare^{-1} (kg) was ranged from 481.07 (ADP0508) to 3873.2 (ADP0383). About forty six per cent (46.33%) of the genotypes gave above the grand mean (1904.497 kg/ha). The broad spectrum in the genotypes mean of days to maturity implies great possibility for the development of early maturing variety to be used by the bean growers in short rainy season areas such as Rift Valley areas of Ethiopia. In agreement with this finding Rezene *et al*. (2013) reported wide range of variation for days to maturity and pods per plant. The finding of Ahmed and kamaluddin (2013) indicated that plant height exhibited high variation; whereas days to 50% flowering, 100 seed weight and seed yield (kg/ha) showed moderate variation and number of pods/plant, pod length and number of seeds/pod showed low variation.

4.1.3 Phenotypic and genotypic coefficient of variation

Understanding the amount of genotypic and phenotypic variability that exists in a species is of utmost importance in breeding programs. Estimate of phenotypic ($\sigma^2{}_p$), genotypic variance ($\sigma^2{}_g$), phenotypic (PCV) and genotypic coefficient of variation (GCV) are presented in Table 2. In this study, genotypic and phenotypic coefficients of variation were used to measure the variability that exists in common beans. Phenotypic coefficient of variations (PCV) ranged from 3.34 for days to 90% maturity to 51.3 for root pulling force resistance. According to Deshmukh *et al.* (1986) PCV and GCV values > 20% regarded as high, PCV and GCV values between 10 and 20% medium, PCV and GCV values < 10% low. Accordingly, high values of PCV were observed for root pulling force resistance (51.3), number of pod plant^{-1} (23.7) and grain yield (32.85). Moderate PCV (%) values were recorded for plant height (18.24) and number of seeds pod^{-1} (15.2). whereas low PCV values were recorded for days to 90% maturity (3.34) and days to flowering (5.62). High PCV were reported by Ahmed and kamaluddin (2013) for pant height (32.85), number of pod per plant (24.45) and grain yield (41.52) and similarly low PCV value recorded for days to flowering(5.45) in 57 germplasm lines of rajmash beans.

The genotypic coefficient of variations (GCV) ranged from 2.73 for maturity date to 27.66 for root pulling force resistance. High values of GCV were observed for root pulling force resistance (27.66) and number of pod plant^{-1} (20.22). Also moderate GCV value was observed for grain yield (17.7). The high values of GCV for root pulling force resistance and number of pod plant^{-1} as well as moderate GCV for grain yield are evident for the presence of high genetic variability among the entries that in turn offers good scope for genetic gain by selection. Bagheri *et al*. (2015) reported high GCV for weight of pods per plant (54.13) and grain yield per plant (50.33). However, the result disagrees with the results of Raffi and Nath (2004) who report low GCV for weight of pods per plant and grain yield per plant. This may be due to difference in environments and genetic materials studied. Low GCV values were recorded for days to flowering, days to harvest maturity and seed per pod. The low GCV values indicate the presence of limited improvement through selection for these characters.

Phenotypic coefficients of variation were generally higher than genotypic coefficients of variation for all traits studied indicating that the influence of growing environments. In most of cases, the two values differ slightly indicating less influence of environmental factors. In the present study, among all characters, high PCV and GCV values (> 20%) were observed for root pulling force resistance and number of pod per plant.

4.1.4 Estimates of heritability (h^2) in broad sense

In this study, the heritability estimate ranged from 25% number of seeds pod to 72.9% for number of pod plant^{-1} (Table 2). Robinson *et al.* (1949) classified heritability values as high (>60%), moderate (31-60%) and values less than 30% low. Based on this delineation high heritability estimates were obtained for days to 90% maturity (66.67%) and number of pod per plan (72.9%). According to Singh (2001) if heritability of a character is very high, selection for such characters could be fairly easy. This is because there would be a close correspondence between the genotype and the phenotype due to the relative small contribution of the environment to the phenotype. Nechifor *et al*. (2011) reported that numbers of pod per plant and numbers of seed pod-1 have medium heritability values.

Medium heritability estimate was observed for days to flowering (54.93%). Whereas, low heritability values were observed for plant height (29.21%), number of seeds pod (25%), root pulling force resistance (29.06%), and grain yield (29.04%). For characters with low heritability, selection may be considerably difficult or virtually impractical due to the masking effect of the environment. Contradictory to the current findings Ahmed and kamaluddin (2013) reported high heritability estimates for plant height, pods plant^{-1} and seed yield under temperate conditions in common bean.

4.1.5 Estimates of expected genetic advance (GAM %)

Genetic advance expressed as a percentage of the mean ranged from 4.59% for maturity date to 35.63% for number of pod per plant (Table 2). Falconer and Mackay (1996) classified genetic advance as percent of mean as low (0-10%), medium (10 - 20%) and high (20% and above). Accordingly, genetic advance as percentage of mean was high for number of pod per plant (35.63) and root pulling resistance (32.06). In addition, the medium value was obtained for plant height and grain yield. The low genetic advance as percent of mean was observed for days to harvest maturity, days to flowering and number of seed per pod.

Selection based on those traits with a high genetic advance as percent of means will result in the improvement of the performance of the genotypes for the traits. A case in point, it is number of pod per plant. This trait also had high heritability. Mudesir *et al*. (2012) reported similar findings for pod per plant in common bean.

Generally, high GCV along with high heritability and genetic advance was obtained from number of pod per plant. High GCV coupled with high heritability and high genetic advance for these traits indicated that selection is effective. Similarly, Ahmed and Kamaluddin (2013), reported high GCV (23.37%) along with high heritability (91.40%) and genetic advance (46.03%) for pod per plant in common bean. High GCV along with high heritability and genetic advance provide better information than other parameters alone (Manju *et al.,* 2002).

Table 2. Estimate of mean, ranges, Phenotypic (PV) and genotypic (GV) coefficient of variation, broad sense heritability and genetic advance as percent of mean for 7 characters of 423 Common bean genotype tested at Hawassa

Character	Mean	Range	δ^2g	δ^2p	H^2b (%)	GCV (%)	PCV (%)	GA (5%)	GAM
Number of pod plant^{-1}	12.83	5.80-29.60	6.74	9.24	72.9	20.22	23.7	4.57	35.63
Days to harvest maturity (number)	73.37	68.0-78.00	4	6	66.67	2.73	3.34	3.37	4.59
Days to flowering (number)	47.42	42.0-55.00	3.9	7.1	54.93	4.16	5.62	3.02	6.37
Plant height (cm)	42.87	8.40-86.70	18.3	61.5	29.76	9.98	18.29	4.81	11.23
Root pulling force resistance (lb)	17.56	4.27-73.77	25.6	88.1	29.06	27.66	51.3	5.63	32.06
Seeds per pod (number)	4.16	2.00-6.600	0.1	0.4	25	7.6	15.2	0.33	7.84
Grain yield (Kg ha^{1-})	1939.8	529.08-3873.2	117900	406028	29.04	17.7	32.85	381.75	19.68

σ^2p, Phenotypic variation; σ^2g, Genotypic variation; PCV, Phenotypic coefficient of variation; GCV, Genotypic coefficient of variation; h^2b, Broad sense heritability; GA, genetic advance; GAM, Genetic advance as percent of mean.

4.2 Association Studies

4.2.1 Correlation of grain yield with other traits

The analysis of the relationship among characters and their association with yield is essential to establish selection criteria (Singh *et al.,* 1990). Therefore, understanding of interrelationships of grain yield and of the magnitudes of genotypic and phenotypic correlations of grain yield and its components among yield related traits are highly crucial to utilize the existing variability through selection.

Phenotypic and genotypic correlation estimates between the various characters are indicated in Table 3. Grain yield had positive and significant correlation with days to maturity, plant height, root pulling force resistance and number of pod per plant at both phenotypic and genotypic levels. Grain yield manifested significant and positive association with seeds per pod at genotypic level only Therefore, any improvement of these characters would result in a substantial increment on grain yield. These results are in accordance with the findings of Salehi *et al,* (2010).The finding of Daniel *et al* (2015) showed that common bean grain yield was positively and significantly correlated with root pulling force resistance, pod per plant and seeds per pod. Akashi *et al*. (2015) also reported a strong positive correlation of seed yield with seed number per plant. Plant height showed significant and positive relation with seed yield as reported by Kassaye (2006). However, Daniel *et al.* (2015) reported a negative and significant correlation of days to harvest maturity and plant height with seed yield across locations and over stress regimes.

Grain yield showed significant and negative correlation with days to flowering at both phenotypic and genotypic levels. Negative correlation was indicated inverse relationship between earliness characters and grain yield that is desirable if stresses such as terminal heat and drought are expected. This negative correlation between grain yield and days to flowering is in harmony with the finding of Daniel *et al.* (2015).

Table 3: Genotypic (above diagonal) and phenotypic correlation coefficients.

	DF	DHM	PLHT	SDPD	PDPL	RPS	GYLD
DF	**1**	0.913**	-0.589**	-0.426**	0.084	0.08	-0.272**
DHM	0.684**	**1**	0.04	-0.313	0.330**	0.588**	0.422**
PLHT	-0.236**	0.018	**1**	0.933**	0.511**	0.709**	0.804**
SDPD	-0.158**	-0.128**	0.325**	**1**	0.160**	0.313**	0.115*
PDPL	0.212**	0.058	0.236**	0.068	**1**	0.541**	0.777**
RPS	0.032	0.259**	0.207**	0.084	0.249**	**1**	0.776**
GYLD	-0.109*	0.186*	0.234**	0.031	0.592**	0.225**	**1**

t =0.098 ($P<0.05$) and t =0.128 ($P<0.01$) for df=n-2, where n is the number of genotypes, DF, days to flowering (number); DHM, days to maturity (number); PLHT, plant height (cm); SDPD, seeds per pod (number); PDPL, pods per plant (number); RPS, root pulling force resistance (lb); GYLD, grain yield (Kg ha^{1-}).

4.2.2 Correlation among other characters

Days to flowering showed positive and significant association with days to maturity at both phenotypic and genotypic levels. However, it displayed negative and significant association with plant height and number of seed per pod at both phenotypic and genotypic levels. It had non-significant correlation with the rest of the characters.

Days to maturity also revealed positive and significant association with vertical root puling force resistance and number of pod per plant at both levels. Plant height showed positive and significant association with seed per pod, pod per plant, vertical root pooling resistance and grain yield hectare $^{-1}$ at both phenotypic and genotypic levels. Similarly, Ahmed and Kamaluddin (2013) reported that plant height showed positive and significant association with number of pods $plant^{-1}$ and number of seeds pod^{-1}.

Seed per pod showed positive and significant association with vertical root pooling resistance at genotypic level. The higher resistance to the upward pulling force should be correlated with

better anchoring of the root system to the soil, possibly indicating higher root density and deeper rooting system. For those traits, which were positively associated, the improvement for one trait will simultaneously improve the other trait.

4.3 Path Coefficient Analysis

As correlation does not permit partitioning of genotypic and phenotypic correlation coefficients in to direct and indirect effects they are further analyzed by path coefficient analysis using grain yield as a dependent variable. The genotypic direct and indirect effects of different characters on grain yield are presented in Tables 4.

Path coefficient analysis revealed that day to maturity had the highest positive direct effect on yield (1.958) followed by plant height (1.863). The genotypic correlations were also positive and significant (r = 0.422and 0.804), respectively. This justifies that there is a true relationship between grain yield, plant height and days to harvest maturity and direct selection through these traits will be effective. Our finding is in agreement with the findings of Karasu and Oz, (2010), where path coefficient analysis revealed that plant height exhibited positive direct effect (0.301) on seed yield. However, a negative direct effect was resulted from days to flowering (-1.794), seed per pod (-1.388) and vertical root puling resistance (-1.092). The positive genotypic correlation of grain yield with seed per pod (0.115) and vertical root puling resistance (0.776) was due to the indirect effect of plant height with seed per pod (1.738) and root pulling resistance (1.321). This showed there was no true relationship between number of seed per pod, root pulling resistance and grain yield. The present finding disagrees with the findings of Daniel *et al*, (2015), where path coefficient analysis revealed that seeds per pod showed a positive direct effect of on seed yield under stress conditions. The disagreement of the finding might be associated with the stressed environmental condition.

Days to flowering (1.788), number of pod per plant (0.646) and root puling resistance (1.151) exhibited positive genotypic indirect effects via days to harvest maturity. Which suggest the merit of days to harvest maturity for improving seed yield. These indirect effects had considerable contribution to their total correlations. The genotypic path coefficient analysis also revealed that seed per pod (1.738), pod per plant (0.952) and root puling resistance (1.321) had

positive indirect effects on seed yield through plant height. Similarly, days to harvest maturity (0.101), plant height (0.157) and root puling resistance (0.161) exhibited positive genotypic indirect effects via number of pod per plant. Plant height (0.977) and seed per pod (0.706) exhibited positive genotypic indirect effects via days to flowering. Thus, both the direct and indirect effects revealed the importance of days to harvest maturity, plant height and seed per pod for the improvement of seed yield per plot. Similarly Daniel *et al,* (2015) reported that root pulling force resistance showed high indirect effects on seed yield through pods per plant and seeds per pod under drought stressed and bean stem maggot infested growing conditions. However, plant height (-1.295) and root puling resistance (-0.814) had negative indirect effects via number of seed per pod. Similarly, days to flowering (-1.145) had negative indirect effects via days to maturity. The path analysis revealed the residual value of 0.45 that means the characters in the path analysis expressed the variability in grain yield by 55%.

Table 4: Path coefficients at genotypic level of direct (diagonal) and indirect effects of the characters

	DF	DHM	PLHT	SDPD	PDPL	RPS	TIE	r_g
DF	**-1.658**	1.788	-1.097	0.591	0.026	-0.087	1.386	-0.272**
DHM	-1.514	**1.958**	0.075	0.434	0.101	-0.642	-1.536	0.422**
PLHT	0.977	0.078	**1.863**	-1.293	0.157	-0.774	-1.059	0.804**
SDPD	0.706	-0.613	1.738	**-1.386**	0.049	-0.342	1.501	0.115*
PDPL	-0.139	0.646	0.952	-0.222	**0.307**	-0.591	0.47	0.777**
RPS	-0.133	1.151	1.321	-0.434	0.166	**-1.092**	1.868	0.776**

Residual = 0.45 DF, days to flowering (number); DHM, days to harvest maturity (number); PLHT, plant height (cm); SDPD, seeds per pod (number); PDPL, pods per plant (number); RPS, root pulling force resistance (lb); TIE, total indirect effect; r, is the correlation coefficient with seed yield.

4.4 Cluster Analysis

Clustering was made to classify quantitative traits into components for the sake of understanding the share components contribute to major variation in the study (Appendix Table 3). The dendrogram obtained from the cluster analysis grouped the 423 genotypes into 19 group and 2

solitaries. This indicates that the tested common bean genotypes were moderately divergence. . Darkwa *et al.* (2016) grouped 64 common bean genotype into four clusters; Kassaye (2006) grouped 114 common bean genotypes into nine clusters, which makes them divergent.

Cluster II had the largest member of all clusters, included 49 (11.58%) genotypes, followed by cluster IV that included 47 (11.11%) genotypes. Similarly, cluster 6 and 7 had 38 (8.98%) and 39 (9.21%) genotypes, respectively. The genotypes found in cluster 1, 2, 3 and 8 that comprised the standard checks, might be regarded as having the overall characteristics of the checks variety.

In contrast, cluster 11, 18, 20 and 21 had the smallest component, constituted of 4 (1%) for cluster 18and 11) and 1 genotype (0.24% for cluster 20 and 21). The genotypes used under this study were collected from the potential common been producing areas. The pattern of clustering revealed that the genotypes collected from same country did not form a single cluster. This indicates that geographic diversity is not always related to genetic diversity.

4.4.1 Mean cluster analysis

Mean value of the 8 quantitative characters for each cluster group is presented in Table 5. Cluster 1 was characterized by a medium value of the traits. Cluster 2, 3 and 4 were characterized by all traits with a medium value that lied between the lowest and the highest values. Cluster 5 was characterized by the shortest plant height with combination of earliest maturity, highest number of pod with lowest values seed yield and highest rot puling force resistance value. Cluster 10 was characterized by the lowest number of pods per plant with a medium value of other traits that lied between the lowest and highest value. Cluster 11 possessed desirable combinations of characters; namely, latest maturing and flowering, the highest root puling force resistance, plant height, many number of emerged seedlings and seed per pod, pod per plant, highest seed yield. Cluster 14 is characterized by the earliest flowering with a medium value that lied between the lowest and highest value. Cluster 15 is characterized by the tallest plant height, the latest flowering and maturing with fewest numbers of pods and highest grain yield. Cluster 18 and 19 characterized by the lowest value of root puling resistance and the highest number of pods, respectively. Generally, this study revealed that the genotypes included in this study are

moderately divergent. Similarly, Getachew (2010) and Kassaye (2006) obtained differences for cluster means for different accessions in common bean.

4.4.2 Genetic divergence

Based on D^2 values the inter cluster distance ranged from 104.80 to 3691.39 (Table 7). The x^2 test for the 21 clusters indicated that there was statistically accepted difference between clusters (Table 6). The highest inter-cluster distance was between cluster 5 and 11(D^2=3691.39) followed by cluster 5 and 14 (D^2 = 3404.59), cluster 10 and 11(D^2=3249.13), cluster 3 and 14 (D^2 = 3208.10). The minimum being between by cluster 16 and 17 (D^2 = 104.80), cluster 17 and 20(D^2=129), cluster 6 and 13 (D^2 = 132.8). Parents for hybridization could be selected on the basis of large inter-cluster distance (in the present case from cluster 5 and 11) for isolating useful recombinants in the segregating generations. Increasing parental distance implies a greater number of contrasting alleles at the desired loci, and then to the extent that these loci recombine in the F_2 and F_3 generations following a cross of distantly related parents, the greater will be the opportunities for successful selection for any character of yield interest (Ghaderi *et al*., 1984).

Table 5: Mean values of 8 characters for the 21 cluster of 423 common bean genotype at Hawassa

Cluster	DF	DHM	PLHT	SDPD	PDPL	RPS	GYLD
1	46.88	74.5	43.38	4.13	12.22	23.87	3219.02
2	47.73	73.81	39.56	4.37	13.08	19.31	2007.79
3	48	73.5	37.17	3.8	13.5	19.76	378.29
4	46.73	72.27	36.44	4.02	12.01	15.93	1807.61
5	47	70	29	4.2	14.82	31.73	181.83
6	47.43	72.84	37.92	4.15	13.21	16.94	1495.99
7	48.21	72.93	38.61	4.04	11.79	16.43	825.51
8	47.22	74.02	40.17	4.34	11.41	17.88	2182.3
9	47.69	73.38	36.17	4.22	13.42	14.76	1184.37
10	47.43	72.71	35.71	3.74	13.19	10.68	624.14
11	51	77	40	4.2	12.54	32.1	3873.19
12	46.61	73.65	39.91	4.17	14.47	20.98	2598.14
13	47.79	73.18	39.03	4.1	12.81	18.42	1628.81
14	45.86	74	40	4.11	11.82	22.31	3586.38
15	48	76	44.33	4.03	11.82	23.51	3364.18
16	46.3	72.85	39.85	4.21	12.92	15.24	2801.84
17	46.6	74.2	41.7	4.08	13.13	29.02	2905.67
18	47.55	73.17	36.1	4.01	12.59	13.78	1365.75
19	47.41	72.36	36.95	4.47	12.7	16.76	1010.8
20	48.25	74.25	38.92	4.15	13.66	20.18	3034.45
21	46.7	73.08	40.47	4.1	12.34	18.43	2404.55

DF, days to flowering (number); DHM, days to harvest maturity (number); PLHT, plant height (cm); SDPD, seeds per pod (number); PDPL, pods per plant (number);RPS, root pulling force resistance (lb); GYLD, grain yield (Kg ha^{1-}).

Table 6: Mahalanobis distance between groups of common bean genotypes.

Cluster	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1211.25	2840.74	1411.47	3037.25	1723.06	2393.53	1036.76	2034.69	2594.93	654.25	620.92	1590.23	367.40
2		1629.51	200.25	1826.05	511.81	1182.28	174.53	823.45	1383.68	1865.45	590.36	378.98	1578.59
3			1429.33	197.15	1117.71	447.25	1804.02	806.10	246.03	3494.93	2219.86	1250.53	3208.10
4				1625.88	311.62	982.10	374.72	623.25	1183.48	2065.66	790.57	178.84	1778.79
5					1314.29	643.97	2000.56	1002.73	442.93	3691.39	2416.37	1447.09	3404.59
6						670.48	686.32	311.64	871.88	2377.25	1102.16	132.83	2090.40
7							1356.79	358.87	201.50	3047.72	1772.64	803.30	2760.87
8								997.95	1558.19	1690.96	415.87	553.49	1404.09
9									560.24	2688.89	1413.80	444.47	2402.03
10										3249.13	1974.04	1004.71	2962.26
11											1275.11	2244.42	287.04
12												969.34	988.24
13													1957.57

x^2 = 12.592, 16.812and 22.46 at 5%, 1% and 0.1% probability level respectively.

Table 6 .Mahalanobis distance between groups of common bean genotypes (continued).

Cluster	15	16	17	18	19	20	21
1	145.19	417.30	313.45	1853.32	2208.25	184.71	814.51
2	1356.41	794.06	897.93	642.07	997.00	1026.66	396.76
3	2985.90	2423.56	2527.40	987.49	632.52	2656.17	2026.27
4	1556.62	994.24	1098.15	441.86	796.81	1226.86	596.96
5	3182.41	2620.09	2723.88	1184.10	829.16	2852.67	2222.80
6	1868.21	1305.85	1409.73	130.30	485.20	1538.47	908.56
7	2538.68	1976.33	2080.20	540.25	185.29	2208.94	1579.04
8	1181.91	619.55	723.46	816.57	1171.51	852.16	222.26
9	2179.85	1617.48	1721.37	181.39	173.59	1850.10	1220.20
10	2740.08	2177.70	2281.61	741.62	386.71	2410.33	1780.43
11	509.11	1071.50	967.55	2507.51	2862.44	838.83	1468.72
12	766.07	203.79	307.64	1232.42	1587.36	436.32	193.63
13	1735.39	1173.03	1276.90	263.12	618.02	1405.64	775.74
14	222.27	784.58	680.75	2220.65	2575.59	551.94	1181.84
15		562.43	458.57	1998.47	2353.41	329.80	959.66
16			104.80	1436.09	1791.05	232.68	397.31
17				1540.00	1894.92	129.15	501.24
18					354.97	1668.72	1038.82
19						2023.66	1393.76
20							629.91

x^2 = 12.592, 16.812and 22.46 at 5%, 1% and 0.1% probability level respectively.

4.5 Principal Component Analysis

Principal component analysis (PCA) is one of the multivariate statistical techniques, which is a powerful tool for investigating and summarizing underlying trends in complex data structures (Legendre and Legendre, 1998). Principal component analysis reflects the importance of the largest contributor to the total variation at each axis for differentiation (Sharma, 1998). The first step in PCA is to calculate eigenvalues, which defines the amount of total variation that is displayed on the PC axes. The PCs with Eigenvalue > 1.0 were used as criteria to determine the number of PCs (Kaiser, 1960).

The principal component analysis revealed that four principal components PC1, PC2, PC3 and PC4 with eigenvalues 1.724, 1.547, 1.037 and 1.00, respectively (Table 7). They have accounted for 66.19 % of the total variation among genotypes for the eight quantitative traits. The relative magnitude of eigenvectors from the first principal component was 21.6% showing that days to maturity and days to flowering were the most contributing traits for the total variation. In the second principal component (PC2), which contributed 19.3% of the total variation, the most predominant characters were vertical root pulling force resistance, plant height, number of seed per pod and grain yield. In the same way, 13.0% of the total variability among the tested genotypes accounted for the third principal component (PC3) originated from variation in number of pod per plant. Number of pod per plant mainly in the fourth principal component (PC4) was the major contributors for the total variation.

The principal component analysis indicated the involvement of a number of traits in contributing towards the overall observed diversity. Similarly, Stoilova *et al.* (2013) reported that variation in 15 Portuguese and 15 Bulgarian common bean landraces in which the first four PCs contributed 82.61 % of total variation .Similarly Assefa *et al* (2014), reported that the combination of first three principal components explained more than 50% of the genotypic variations

Table 7: Eigenvectors and eigenvalues of the first four principal components (PCs) of common bean genotypes

Variable	PC1	PC2	PC3	PC4
Days to 90% maturity	0.656	-0.046	-0.010	0.038
Days to flowering	0.632	0.273	-0.079	0.067
Plant Height	-0.237	0.541	0.027	0.061
Number of seeds pod^{-1}	-0.284	0.313	0.110	0.189
Number of pod plant^{-1}	-0.043	0.004	-0.576	0.798
Vertical root pulling force	0.142	0.535	-0.099	-0.154
Grain yield hectare^{-1}	0.028	0.477	0.278	0.095
Eigenvalue	1.724	1.547	1.037	1.02
Proportion	21.6	19.3	13.0	12.3
Cumulative	21.59	40.89	53.86	66.19

4.6 Phenotypic Diversity and Genome-wide Association Analysis

4.6.1 Phenotypic analysis for water uptake and cooking time

A total of 398 genotypes of common bean were evaluated for water uptake and cooking time traits. The analysis of variance (ANOVA) reveled highly significant differences (P< 0.001) among the genotypes for water uptake and cooking time (Table 8). The water uptake ranged from 51.92% (ADP0722) to147.2 % (ADP0457) with the average value of 112.36%. Whereas, cooking time ranged from 14.76 (ADP0367) to 44.86 min (ADP0722). It was observed that the overall mean for cooking time was 25.51 min. The result indicated reasonably sufficient variation. Similarly, Cichy *et al*. (2015) reported wide range of variation for cooking time in common bean. In addition, Elia *et al*. (1996) reported significant variations for cooking time and water absorption. The current finding disagrees with the finding of Bulti, (2007), who reported non-significant differences and inadequate variation among genotypes for cooking time. The low variation of cooking time might be due to small number of genotypes (Eight) tested by the researcher.

4.6.2 Heritability and Correlation Analysis

The heritability value of the cooking time and water uptake were 39.27 and 64.1. Based on Robinson *et al.* (1949) delineation the estimated heritability values was high for cooking time, which suggests cooking time may be under the control of gene. Similarly, Augusta *et al*. (2012) demonstrated in their study a high heritability value for cooking time.

The correlation analysis revealed that there was a negative correlation between water uptake and cooking time trait. Such genotypes that received higher water during soaking got shortest time to cook. The causes of slow water absorption might be due to the hard skin of the bean seed. Hard-shell takes into account the phenolic compounds found in the seed coat that can be oxidized into complex polymers that may interact with proteins to yield a hydrophobic material that repels water imbibitions (Stanley and Aguilera, 1985). Selection based on the water absorption of breeding lines an indirect estimation of its cooking time; it is rapid save resources (Elia *et al*., 1996). Several researchers report inverse relationship between cooking time and water uptake (Castellanos *et al.,*1995; Bertoldo *et al*., 2008 Cichy *et al.,* 2015.). Similarly, Elia *et al*. (1996) reported a negative phenotypic correlation between water uptake and cooking time (-0.78).

Table 8: Summary of the analysis of variance, heritability, phenotypic and genotypic variance for the cooking time and water absorption traits in common bean genotypes.

Source of variation	DF	Water uptake mean Square	DF	Cooking time mean Square
Genotype	397	262.09**	397	48.82**
Error	398	114.19	398	8.12
C V (%)		9.51		11.30
$\sigma^2 g$		73.9		19.65
σ2p		188.19		30.27
$H^2 b$ (%)		39.27		64.91
Mean range		51.92 -147.2 %		14.76 - 44.86
Grand mean		112.36		25.21

DF; degrees of freedom, C.V; Coefficient of variation; ** Statistical significance of the F test (P <0.01); σ2p =Phenotypic variation; σ^2g= Genotypic variation; $H^2_{b=}$ Broad sense heritability

4.6.3 Fast cooking beans

The fast cooking time of the 11 genotypes is presented in Table 9. Out of 398 Common bean genotype, 11 genotypes showed fast cooking time which were cooked in less than 17 min. The fast cooking genotype originated from CIAT (5), Africa (5) and Ecuador (1). Those genotypes which, showed fast cooking time had eight different seed colors, namely: purp cran, yellow, cran, white, purple speckled and red mottled. The growth habits of the fast cooking genotypes are vine, bush and climber. Cichy *et al*. (2015) reported that five bean accessions, which cook fast. The fast cooked genotypes can be used as parents to develop fast cooking breeding populations (recombinant inbred line) for selection.

Table 9: Identification of the eleven fastest cooking genotypes of the 398 common bean genotypes.

ADP ID	Source	Genotype	Type	Habit	Region	Cooking time(min)	Water uptake (%)
ADP0367	05IS-0886	G23086	purp cran	Vine	CIAT	14.76	107.79
ADP0037	OT1122-162	W6 16488	brn	Bush	Africa	14.89	126.99
ADP0271	05IS-0506	G 13167	white	Bush	CIAT	14.99	124.25
ADP0093	OT1122-100	MORO	yellow	Vine	Africa	15.77	120.86
ADP0180	05IS-0007	G 433	cranberry	Bush	CIAT	15.85	110.89
ADP0066	OT1122-54	NJANO	yellow	Vine	Africa	16.26	109.76
ADP0092	OT1122-99	MORO	yellow	Vine	Africa	16.43	122.87
ADP0207	05IS-0170	G 4564	j. cattle	Bush	CIAT	16.53	117.98
ADP0063	OT1122-46	Soya	purp. Spec	Climber	Africa	16.63	104.69
ADP0206	05IS-0167	G 4499	white	Vine	CIAT	16.74	115.37
ADP0458	-	INIAP 483	red mottled	Bush	Ecuador	16.79	111.22

4.7 Genome-wide Association Analysis

4.7.1. Genome-wide association analysis for water uptake

The genotypic data used in this study was dense enough to capture the variability and to study the genome wide association mapping (Appendix Fig. 2) as described by Pearson and Manolio. (2008). In this study, a total of 10 significant SNP markers were detected above the FDR cutoff and the phenotypic effect of SNP marker alleles on the associated characters and the number of genotypes carrying each significantly associated marker allele were identified (Table 10 and Fig.2). The significant marker trait associations were found on chromosomes Pv04, Pv05, and Pv09. The significant markers-trait associations were identified using 10 different SNP markers for water uptake with R^2 ranging from 5.7 % (ss715646204) to 10.41 % (ss715649806).

The most significant set of SNPs associated with water uptake were found on chromosome Pv04.This group of five SNPs had an R^2 of from 5.7 to 10.2 % of the phenotypic variation for this trait (Table 10). The significant SNP have minor allele frequencies of less than 0.005.

The significant markers for water uptake detected on Pv05 explained 14.1% of the variation and have minor allele frequencies of less than 0.002. The associations on Pv09 were the most significant and also explained the phenotypic variation with marker R^2 values of 10.4. The significant SNPs detected on Pv09 had minor allele frequency of 0.002. Cichy *et al.* (2015) reported SNP markers that control water uptake in common bean on chromosomes Pv01, Pv03, Pv06, and Pv07 by using Genome wide association mapping. Perez-Vega *et al.* (2010) identified a QTL for water uptake on Pv03, based on the location of the nearest marker on that map with known sequence (Bng21) the QTL was located to the same chromosome arm.

The first candidate gene identified was PF14547 on Pv09 that codes for Hydrophobic seed protein (Hydrophob_seed). The location of the significant SNP on Pv09 (ss715649806) is approximately 0.06 Mb away from the estimated location of the gene, which was consistently significant for water absorption (Stanley and Aguilera, 1985).

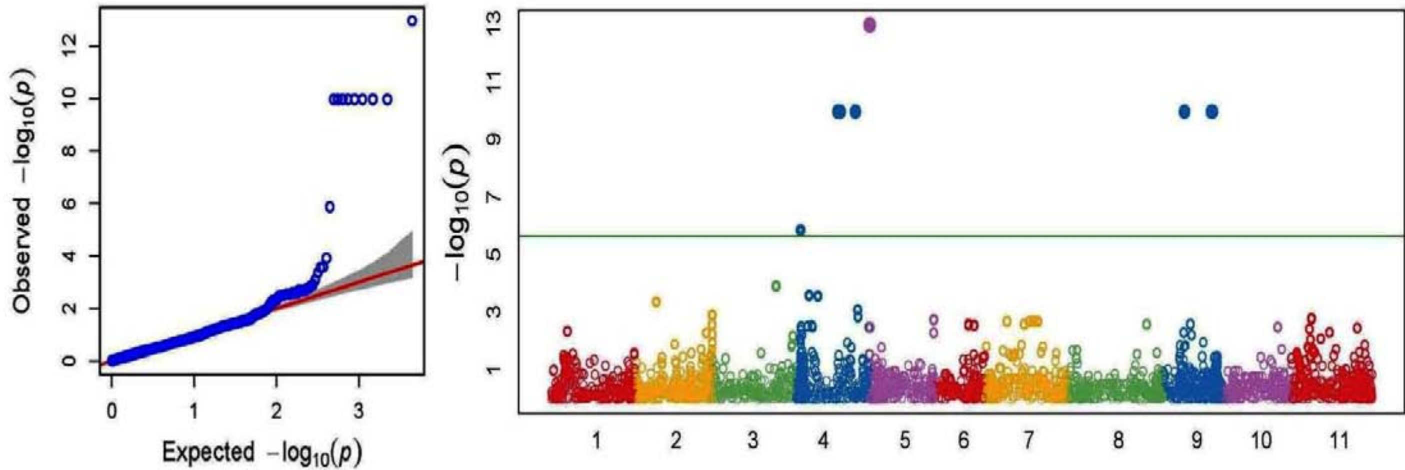


Figure 2. QQ plot for the water uptake data and Manhattan plot of water uptake of 386 genotype with 4,416 SNP.

Table 10. GWAS significant markers, genome position, *p* value, R^2, allele and allele frequency, and phenotypic effect associated with Water uptake

SNP	Chr	Position	P.value	MAF	R^2 (%)	Allele	Effect	Obs
ss715650990	4	26314760	1.08E-10	0.0025907	10.4	T	-9.6	385
						C	0.0	1
ss715650115	4	27464168	1.08E-10	0.0025907	10.4	T	-9.6	385
						C	0.0	1
ss715649688	4	27781563	1.08E-10	0.0025907	10.4	A	-9.6	385
						G	0.0	1
ss715647337	4	37280582	1.08E-10	0.0025907	10.4	A	-9.6	385
						C	0.0	1
ss715646204	4	2757991	1.39E-06	0.0051813	5.7	A	-4.9	384
						G	0.0	2
ss715647828	5	183765	1.09E-13	0.0025907	14.1	A	10.8	385
						G	0.0	1
ss715646633	9	13714826	1.08E-10	0.0025907	10.4	A	-9.6	385
						G	0.0	1
ss715646638	9	13773940	1.08E-10	0.0025907	10.4	T	-9.6	385
						G	0.0	1
ss715646279	9	30645670	1.08E-10	0.0025907	10.4	A	-9.6	385
						C	0.0	1
ss715649806	9	31246392	1.08E-10	0.0025907	10.4	A	-9.6	385
						G	0.0	1

$\mathbf{R^2}$ is the percent of phenotypic variation explained by the SNP marker; **MAF,** minor allele frequency.

4.7.2 Genome-wide association analysis for cooking time

Significant SNPs associated with cooking time were found on chromosomes Pv04, Pv05, and Pv09 (Fig. 3). The most significant set of SNPs associated with cooking time were found on chromosome Pv04 and Pv09. on Pv04, the significant SNPs have the same effects on cooking time. The four SNPs markers (ss715650990, ss715650115, ss715649688 and ss715647337) had 3 % of the phenotypic variation. The major allele reduced cooking time by about 9 min and was found in 99 % of the genotype (Table 11).

Few candidate genes of interest potentially are involved in determining cooking time on PV 04. One of these genes is CHX10 (CATION/$H^{(+)}$ ANTIPORTER 10-RELATED). The function of this gene in Arabidopsis has been related to the directed movement of potassium ions (K^+) and sodium ions (Na^+) into, out of or within a cell, or between cells. It is evidence that monovalent cations (Na and K) to interact with pectin, and leading to a structure that is heat-labile (Stanley and Aguilera, 1985)

The significant SNPs on chromosome Pv05 had R^2 values of 3.4 % (Table 11) and the major allele decreased cooking time by 3 min. The minor allele was found in 28 genotypes. From this SNP marker (ss715640177) the minor allele identifies in some genotypes with extended cooking times ranging from 28 to 44 min.

The other majority of significant SNPs associated with cooking time were found on chromosome Pv09. This group of four SNPs had an R^2 of 9.21 % of the phenotypic variation for this trait (Table11) and the major allele, which was found in 99 % of the genotypes, decreased cooking time by 9.21 min.

A candidate gene, which potentially involved in determining cooking time is found on PV 09. This gene is pectin esterase/pectin esterase inhibitor 13 (PME 13). It has been related to three functions in Arabidopsis. The first function related to the series of events leading to chemical and structural alterations of an existing cell wall that can result in loosening, increased extensibility or disassembly. The second function is the chemical reactions and pathways

resulting in the breakdown of pectin and the third one is, catalysis of the hydrolysis of an ester bond by a mechanism involving a catalytically active aspartic acid residue (NCBI).

The other candidate gene on Pv09 near SNPs associated with cooking time is Naringenin-chalcone synthase / Flavonone synthase genes. Chlcone synthase (CHS, EC 2.3.1.74) is a key enzyme of the flavonoid/isoflavonoid biosynthesis pathway. CHS expression causes acumilation of flavonoid and isoflavonoid phytoalexins (Dao *et al*., 2015). The role of flavonoids is responsible for the development of the major pigments (red, blue, and purple pigments in plants) in plants. Chalcone synthase gene is involved in flavonoid biosynthesis (Bowles *et al.* 2005) and this gene is of interest because seed coat color (via market class) was correlated with cooking time (Cichy *et al*., 2015)

For all the above significant SNP markers, the minor allele was found in ADP 0722 accession with cooking times 45.

A little study has been conducted previously for cooking time. In GWAS for cooking time, significant SNPs associated with cooking time were identified on chromosomes Pv02, Pv03, and Pv06 (Cichy *etal*.,2015). Garcia *et al*. (2012) identified six QTL on chromosome Pv01and Pv09 from the evaluation of cooking time on 140 F2:4 segregating families.

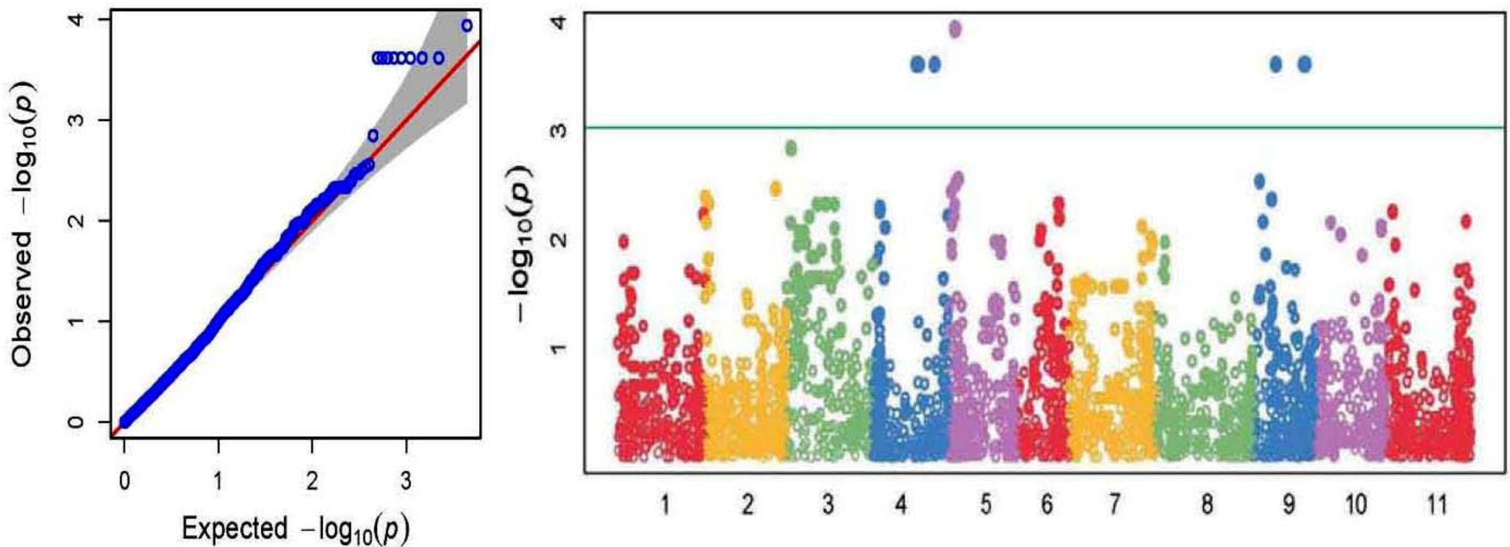


Figure 3. QQ plot for the cooking time data and Manhattan plot of cooking time of 386 genotypes with 4,416 SNP.

Table 11. GWAS significant markers, genome position, p value, R^2, allele and allele frequency associated with cooking time

SNP	Chr	Position	*P*.value	MAF	R^2(%)	Allele	Effect	Obs
ss715650990	4	26314760	0.000241	0.002591	3	T	9.21	385
						C	0.00	1
ss715650115	4	27464168	0.000241	0.002591	3	T	9.21	385
						C	0.00	1
ss715649688	4	27781563	0.000241	0.002591	3	A	9.21	385
						G	0.00	1
ss715647337	4	37280582	0.000241	0.002591	3	A	9.21	385
						C	0.00	1
ss715646633	9	13714826	0.000241	0.002591	3	A	9.21	385
						G	0.00	1
ss715646638	9	13773940	0.000241	0.002591	3	T	9.21	385
						G	0.00	1
ss715646279	9	30645670	0.000241	0.002591	3	A	9.21	385
						C	0.00	1
ss715649806	9	31246392	0.000241	0.002591	3	A	9.21	385
						G	0.00	1
ss715640177	5	3636455	0.000115	0.067358	3.4	G	3.20	358
						T	-8.85E-03	28

R^2 is the percent of phenotypic variation explained by the SNP marker; **MAF,** minor allele frequency.

5 . SUMMARY AND CONCLUSION

The information regarded with the extent and pattern of genetic variability in a population, its genetic diversity associated with its quantitative characters and developing high yielding common bean varieties with acceptable consumer preference and meeting quality standards is one of the top priorities for any breeding strategy and improvement program. In order to generate such information, 418 common bean genotype including five standard checks were tested at Hawassa Agricultural Research Center with the objective of assessing the genetic variability and character associations for 8 morphological characters and identify genomic regions involved in determining cooking time using genome-wide association study.

The analysis of variance revealed the genotypes were significantly different at($p<0.05$) for all characters except leaf chlorophyll content, pod harvest index and hundred seed weight. The ranges of mean values for most of the characters were large showing the existence of variations among the tested genotypes. The value of PCV ranges from 3.34 - 51.3 % while GCV ranges between 2.73 - 27.66 %. High PCV and GCV values were observed for vertical root pulling force resistance and number of pod per plant. On the other hand, low GCV and PCV values were observed for days to flowering and maturity, which suggests the limitation of selection for this trait.

Very high heritability estimates resulted in days to 90% maturity and number of pod per plant. However, the broad sense heritability was low for plant height (29.21%), number of seeds pod (25%), root pulling force resistance (29.06%), and grain yield (29.04%). High heritability coupled with relatively high genetic advance as percent of the mean was observed for number of pod per plant. On the contrary, Low heritability was coupled with low genetic advance values were observed for seed per pod. Generally, high GCV along with high heritability and genetic advance were obtained from a number of pod^{-1}.

It was observed that yield had positive and significant phenotypic and genotypic association with days to maturity, plant height and root pulling force resistance and number of pod per plant. Generally, for all characters studied, the genotypic correlation coefficients were greater than the phenotypic correlation coefficients. By selecting for these traits, showing positive and significant

correlation coefficient with grain yield there is a possibility to increase grain yield of common bean.

Genotypic correlation coefficients of various characters were partitioned in to direct and indirect effects. Path coefficient analysis revealed that plant height and days to maturity showed high and positive direct effect on grain yield. The correlation coefficient was also positive and significant. Since plant height and days to maturity had positive correlation with yield in the process of selection much attention should be given to them, as these characters are helpful for indirect selection.

The cluster analysis based on D^2 analysis on pooled mean of genotypes classified the 423 genotypes into twenty-one clusters, which makes them to be moderately divergent. There was statistically approved difference between most of the clusters. Each cluster had its own characteristic feature for its cluster formation. Cluster II had the largest member of all clusters, included 49 (11.58%) genotypes. In contrast, cluster 11, 18, 20 and 2 had the smallest component, constituted of 4 (0.95%for cluster 18and 11) and 1(0.24% for cluster 20 and 21) genotypes respectively.

The principal component analysis revealed that four principal components PC1, PC2, PC3 and PC4 with eigenvalues 1.724, 1.547, 1.037 and 1.00, respectively, have accounted for 66.19 % of the total variation. This result further confirmed the presence of genetic diversity for use in improvement plan.

High variation of cooking time was observed and eleven common bean genotypes (ADP0367, ADP0037, ADP0271, ADP0093, ADP0180, ADP0066, ADP0092, ADP0207, ADP0063, ADP0206 and ADP0458) were identified for cooking less than17 min. The identified fast cooking accession is useful in breeding program. GWAS reveled significant SNPs that associated with cooking time was found on chromosomes Pv04, Pv05, and Pv09 for water uptake on Pv04, Pv05and Pv09.

This study, generally, indicated that there was sufficient genetic variability among the genotypes. Thus, there is enormous opportunity in the improvement program of the common bean genotypes

Therefore; the information generated from this study needs to be used by breeders who are interested in improving different traits of the crop.

For future improvement program, superior genotype for yield and different yield components can be selected. Moreover, multiple parent advanced generation inter cross population (MAGIC) can bring further improvement in common bean (Andean gene pool) adaptation to drought and low soil fertility, with market value for Ethiopia in the future.

However, all the above conclusions were derived from results of studies conducted within one season. Thus, further studies of common bean genotypes with larger sample size in broad environments and seasons should be conducted on the common bean variety in order to give confirmative results.

The associated markers are possible candidates for future marker-assisted selection to improve cooking time trait. However, the significant markers need to be validated in different environments before their use in marker-assisted selection. This research serves as an important base for further studies to understand the genetic control of cooking time in pulse crop.

.

6. REFERENCE

Adkins, E., E.Tyler, J.Wang, D. Siriri and V. Modi, 2010. Field-testing and survey evaluation of household biomass cook stoves in rural sub-Saharan Africa. *Energy for sustainable development*, **14** (3): 172-185.

Ahmed, S. and Kamaluddin., 2013. Correlation and path analysis for agro-morphological traits in rajmash beans under Baramulla-Kashmir region. *African Journal of Agricultural Research*, **8** (18): 2027-2032.

Akçura, M., Y. Kaya and S.Taner, 2005. Genotype- Environment Interaction and Phenotypic Stability Analysis for Grain Yield of Durum Wheat in the Central Anatolian Region Turk *J. Agric*. **29**:369-375

Akhshi, N., F. Firouzabadi & K. Cheghamirza, 2015. Coefficient Analysis and Association between Morpho-Agronomical Characters in Common Bean (Phaseolus vulgaris L.). *Cercetari Agronomice in Moldova*, **48**:29-37

Akibode, S. and M. Maredia, 2011. Global and regional trends in production, trade and consumption of food legume crops. Department of Agricultural, Food and Resource Economics, Michigan State University.

Allard, R.W., 1960. Principles of Plant Breeding. John Wiley and Sons Inc. New York.

Allard, R.W., 1999. Principles of Plant Breeding 2nd (ed). John Wiley and sons, Inc. New York. 95-115p.

Asrat, A., W. B. Matthew & C. S. Paul, 2012. Multienvironment quantitative trait loci analysis for photosynthate acquisition, accumulation, and remobilization traits in common bean under drought stress. *G3: Genes| Genomes| Genetics*, **2**(5):579-595.

Assefa, M., B. Shimelis., S. Punnuri , R. Sripathi, W. Whitehead, and B. Singh, 2014. Common Bean Germplasm Diversity Study for Cold Tolerance in Ehtiopia. *American Journal of Plant Sciences*, *5*(13):1842.

Augusta, G.R. V., P.N. Rangel, P.Z. Bassinello, C. Brondani, L.C. Melo, S.T. Sibov, and R.P. Vianello-Brondani, 2012. QTL mapping for the cooking time of common beans. *Euphytica*, **186**(3):779-792.

Avilla-Rodriguez, E. H., M. P. Barra- Perez., Intriago- Ortega and M. I. Valderrama Chairez, 1996. Genetic variation for grain quality and nutritional characteristics in common bean (*Phaseolus vulgaris* L.) cultivars of Mesoamerican and Andean origin. Annual Report of the Bean Improvement Cooperative. **39**:258-259.

Awan F.K.,Y.M. Khurshid & O. Afzal, 2014. Agro-morphological evaluation of some exotic common bean genotypes under rainfed conditions of Islamabad, Pakistan. *Pak.* J. Bot. **46**(1):259

Bagheri, M., D. Kahrizi and A. Zebarjadi, 2015. Study on genetic variation and morpho-phenologic traits in Common bean. *Biharean Biologist*.

Balardin, R.S. & J.D. Kelly, 1998. Interaction among races of *Colletotrichum lindemuthianum* and diversityin *Phaseolus vulgaris*. *J Am Soc Hortic Sci* **123**:1038–1047.

Baldridge, D. E. and G. L. Tibke, 1990. 1990. Common bean. Montana State University, Animal and Range Sciences Extension Service, Forage extension program, Bozeman, USA.

Beattie, A.D., J. Larsen., T.E. Michaels & K.P. Pauls , 2003. Mapping quantitative trait loci for a common bean (*Phaseolus vulgaris* L.) ideotype. Genome **46**:411–422.

BecerraVela´squez, V.L. and P. Gepts., 1994. RFLP diversity in common bean (*Phaseolus vulgaris* L.). Genome,**37**:256-263.

Benjamini, Y. and Y. Hochberg, 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B (Methodological), **57**:289–300

Berhanu, M., 2004. Genetic variability and character associations in bread wheat (Triticum aestivum L.) genotypes developed for semiarid areas (Doctoral dissertation, MSc Thesis, Haramaya University, Ethiopia).

Bertoldo ,J.G, J. Coimbra. , C.B Silveira and D. Toaldo, 2008. Consequences of common bean domestication for plant breeding . Biotemas **21**:39–44.

Bhatt, G.M. 1973. Significance of path coefficient analysis in determining the nature of character association. *Euphytica*, **22**: 338-343.Biotechnology and Conventional Approaches. Narosa Publishing House, New Delhi.604P.

Biran, A., J. Abbot and Mace., R., 2004. Families and firewood: A comparative analysis of the costs and benefits of children in firewood collection and use in two rural communities in sub-Saharan Africa. Human Ecology **32**:1-25.

Blair, M.W., M. C. Giraldo., H. F. Buendía., E. Tovar., M. C. Duque and S. E. Beebe, 2006. Microsatellite marker diversity in common bean (Phaseolus vulgaris L.). *Theoretical and Applied Genetics*, **113**(1):100-109.

Bowles D, Isayenkova J, Lim E-K, Poppenberger B (2005) Glycosyltransferases: managers of small molecules. Curr Opin Plant Biol8:254–263

Bradbury, P.J., Z. Zhang., D.E. Kroon., T.M. Casstevens and E.S. Buckler, 2007. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics **23**:2633–2635.

Bressani, R., 1993. Grain quality of common beans. *Food Reviews International*, **9**(2):237-29.

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

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

Brouwer, I.D., L.M. Nederveen., A.P. Denhartog and A.H.C. Vlasveld, 1989. Nutritional impacts of an increasing fuelwood shortage in rural households in developing countries. Progress in Food and Nutrition Science **13**:349-361

Bulti Tesso, 2007. Yield and canning quality attributes of navy beans (*Phaseolus vulgaris L.*) as influenced by genotype and environment. An M.Sc. Thesis Presented to Haramaya University, Ethiopia.

Burt, A.J., H.M. William., G. Perry., R. Khanal., K.P. Pauls and J.D. Kelly, 2015. Candidate Gene Identification with SNP Marker-Based Fine Mapping of Anthracnose Resistance Gene Co-4 in Common Bean. *PLoS ONE* **10**(10).

Burton, G.A. and E.H. Devane, 1953. Estimating heritability in tall Fescue (Festuca arundinacea) from replicated clonal material. Journal of Agronomy **45**:481-487.

Castellanos, J.Z,.H.Guzman-Maldonado,J.A. Acosta-Gallegos, and J.D.Kelly, 1995.Effects of hardshell character on cooking time of common beans grown in the semiarid highlands of mexico. *J.Sci.Food Agr.* **69**:437-443.

Central Statistics Agency., 2015. Report on Area and Production of Crops (Private peasant Holdings. Meher Season). Volume I, Statistical Bulletin, Addis Ababa, Ethiopia.

Centro Internacional de Agricultura Tropical (CIAT).,1989. Bean Program Annual Report. Working Document No. 68, 1989, Cali, Colombia. Pp. 165-171.

Chahal, G.S. and S.S. Gosal, 2002. Principles and Procedures of Plant Breeding: Biotechnology and Conventional Approaches. Narosa Publishing House, New Delhi, India, pp: 604.

Chen, X. and P.F. Sullivan, 2003. Single nucleotide polymorphism genotyping: biochemistry, protocol, cost and throughput. *The pharmacogenomics journal*, **3**(2):77-96.

Cichy ,K.A., T. Porch., J.S. Beaver., P. Cregan., D. Fourie., R. Glahn., M. Grusak., K. Kamfwa., D. Katuuramu., P. McClean., E. Mndolwa., S. Nchimbi- Msolla., MA. Pastor-Corrales and P.N. Miklas, 2015. A *Phaseolus vulgaris* diversity panel for Andean bean improvement. *Crop Science*, ***55***(5):2149-2160.

Cichy, K.A., J.A. Wiesinger and F.A. Mendoza, 2015. Genetic diversity and genome-wide association analysis of cooking time in dry bean (Phaseolus vulgaris L.). *Theoretical and Applied Genetics*, **128**(8):1555-1567.

Clark, A.G., M.J. Hubisz., C.D.Bustamante., S.H. Williamson and R. Nielsen, 2005. Ascertainment bias in studies of human genome-wide polymorphism. *Genome research*, **15**(11):1496-1502.

Coelho, C., C. de Mattos Bellato., J. Santos., E.Ortega and S.M. Tsai, 2007. Effect of phytate and storage conditions on the development of the 'hard-to-cook' phenomenon in common beans. *J Sci Food Agric* **87**:1237–1243

Cokkizgin, A., M. Colkesen., L. Idikut., B. Ozsisli and U. Girgel, 2013. Determination of relationships between yield components in bean by using path coefficient analysis. *GAgri Scie*, **3** (2):085-089.

Copper, M.C and G.W. Milligan , 1988. The effect of error on determining the number of clusters. Proceeding of International Workshop on Data Analysis, Decision Support and Expert Knowledge Representation in Marketing and Related Areas of Research pp: 319-328.

Correa M.M, L.M. Jaeger de Carvalho., M.R. Nutti., J.L. Viana de Carvalho., A.R.H. Neto and E.M.G. Ribeiro, 2010. Water absorption, hard shell and cooking time of common beans (*Phaseolus vulgaris* L.). *African Journal of Food Science and Technology* **1**: 13-20.

Daniel, A., Firew, M., Asrat, A., Stephen, E. B., Matthew, W.B., 2015. Trait associations in common bean genotypes grown under drought stress and field infestation by BSM bean fly. *The Crop Journal*, **3**(4):305-316.

Daniel, A., Asrat, A., Firew, M., Raatz, B., Beebe, S. & Blair, M., 2014. QTL Mapping of Resistance to Drought and Bean Fly in Common Bean (Phaseolus vulgaris L .). Generation Challenge Program, poster presentation.

Dao, T.T.H., Linthorst, H.J.M. and Verpoorte, R., 2011. Chalcone synthase and its functions in plant resistance. Phytochemistry Reviews, 10(3), pp.397-412.

Darkwa, K., D. Ambachew, H. Mohammed , A. Asfaw, & M. W. Blair, 2016. Evaluation of common bean (Phaseolus vulgaris L.) genotypes for drought stress adaptation in Ethiopia. *The Crop Journal*. http://dx.doi.org/10.1016/j.cj.2016.06.007

David, S., R. Kirkbyand and S. Kasozi, 2000. Assessing the impact of bush bean varieties on poverty reduction in sub-Saharan Africa: Evidence from Uganda (Vol. **31**). CIAT. Network on bean research in Africa.

Debouck, D., 1991. Systematics and morphology, *In* A. Van Schoonhoven and O. Voysest, eds. Common beans: Research for crop improvement. CIAT, Cali, Colombia. p. 55-118,

Deshmukh, S.N., M.S. Basu and P.S. Reddy, 1986. Genetic variability, character association and path coefficient analysis of quantitative traits in Viginia bunch varieties of ground nut. *Indian Journal of Agricultural Science,* **56**:515-518.

Dewey, D.R. and K.H. Lu, 1959. A correlation and path coefficient analysis of components of crested wheat grass seed production. *Agronomy Journal.* **51**: 515-518.

Ecoport, 2013. Ecoport database. Ecoport A database devoted to inter-disciplinary integration of information to manage biodiversity.

Elia, F. M., G. L. Hosfield and M. A. Uebersax, 1996. Inheritance of cooking time, water absorption, protein and tannin content in dry bean and their expected gain from selection. Annual Report of the Bean Improvement Cooperative. **39**:266-267.

Elia, F.M., G.L. Hosfield., J.D. Kelly and M.A. Uebersax, 1997. Genetic analysis and interrelationships between traits for cooking time, water absorption, and protein and tannin content of Andean dry beans. *J Am Soc Hortic Sci* **122**:512–518

Erena, E.A., 2007. Association mapping for yield, yield components and drought tolerance-related traits in spring wheat grown under rainfed and irrigated conditions (Doctoral dissertation, Colorado State University. Libraries).

Ersoz, E., J. Yu and E. Buckler , 2009. Applications of Linkage Disequilibrium and Association Mapping in Maize. A.L. Kriz, B.A. Larkins (eds.), Molecular Genetic Approaches to Maize Improvement 173. *Biotechnology in Agriculture and Forestry*, Vol. **63**: 173-195.

Evans, A.M. 1980. Structure, variation, evolution, and classification in *Phaseolus.*, p. 337–347, *In* R. J. Summerfield and A. H. Bunting, eds. Advances in legume science. Royal Botanic Gardens, Kew, UK.

Everitt, B. S. and G. Dunn, 2001. Applied multivariate data analysis (No. 519.5076 E9). Arnold, London.

Falconer, D. S. and F. C. T Mackay, 1996. Introduction to Quantitative Genetics. pp. 122-125. 4th (eds). Longman Group Ltd, England.

Falconer, D.S., 1989.Introduction to Quantitative Genetics. (3rdEd). Logman Scientific and Technical, Logman House, Burnt Mill, Harlow, Essex, England.

Fan, J.B., A .Oliphant., R .Shen., B.G. Kermani., F. Garcia., K.L. Gunderson., M.Hansen., F. Steemers., S.L. Butler., P.Deloukas and L. Galver, 2003, January. Highly parallel SNP genotyping. In *Cold Spring Harbor symposia on quantitative biology* (Vol. 68, pp. 69-78). Cold Spring Harbor Laboratory Press.

FAO, 2015. Analysis of price incentives for haricot beans in Ethiopia. Technical notes series, MAFAP, by Workao T. K., MasAparisi A., Lanos B., Rome.

FAO.,2003. *http://www.fao.org/BIOTECH/docs/Korzun.pdf. Accessed on June 2008.*

FAOSTAT, 2015. Food and Agriculture Organisation of the United Nations (FAO) Statistics online website, http://faostat.fao.org (accessed June 2015).

Federer, W. T.,1956. Augmented designs.*Hawaiian Planters Record*,**55**, 191-208.

Fininsa, C. and J. Yuen, 2002. Temporal progression of bean common bacterial blight (*Xanthomonas campestris pv. Phaseoli*) in sole and intercropping systems. *Eur. J. Plant Pathol*., **108**:485-495.

Fininsa, C. and T. Tefera, 2002. Inoculum sources of bean anthracnose and their effect on bean epidemics and yield. Trop. Sci., **42**: 30-34.

Freyre, R., R.Rı´os., L. Guzma´n., D. Debouck and P . Gepts, 1996. Ecogeographic distribution of *Phaseolus* spp. (Fabaceae) in Bolivia. *Economic Botany* **50**:195-215.

Garcia del Moral, L.F., Y. Rharrabti,. D. Villegas and C. Royo, 2003. Evaluation of grain yield and its components in durum wheat under Mediterranean conditions: An ontogenic approach. *Agron. J*., **95**: 266-274.

Geil, P. B. & J. W .Anderson. 1994. Nutrition and health implications of dry beans: A review. *Journal of the American College of Nutrition*, **13**, 549–558.

Genetic Diversity. In: Wikipedia, the Free Encyclopedia [discussion list on the Internet]. 2009 Mar. 13; 1 59 am. Available from: http://en.wikipedia.org/wiki/Genetic_diversity

Gentleman, R.C., V.J. Carey., D.M. Bates., Bolstad., B. Dettling., M. Dudoit., S. Ellis., B. Gautier., L. Ge., and J. Gentry, 2004. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5:R80

Gepts, P. and Debouck, D.,1991. Origin, domestication and evolution of common bean (*Phaseolus vulgaris* L.), p. 7-54, *In* A. Van Shoonhoven and O. Voysest, eds. Common beans: Research for crop improvement. CIAT, Cali, Colombia.

Getachew, E., 2010. Phenotypic diversity in haricot bean(*phaseolus vulgaris l.)* Germplasm a thesis submitted to the School of Graduate Studies, Jimma University College of Agriculture and Veterinary Medicine.

Ghaderi, A., M. W .Hosfield., A .MUebersax, 1984. Variability in culinary quality, component interrelationships, and breeding implications in navy and pinto beans. *Journal of American Society of Horticultural Sciences*. **109** (1): 85-90.

Ghaderi, A., M.W . Adams and A.M. Nassib, 1984. Relationship between genetic distance and heterosis for yield and morphological traits in dry edible bean and faba bean. *Crop Science,* **24**:37-24.

Giriraj, K. and S.Vijayakumar, 1974. Path-coefficient analysis of yield attributes in mung bean. *Indian Journal of Genetics and Plant Breeding (The)*, **34**(1):27-30.

Giriraji, K. and S .Vijayakumar,1974. Path coefficient analysis of yield attributes in mung bean. Ind. J. Genet. **34**: 27-30.

Gomez, K.A. and A.A. Gome, 1984. Statistical Procedures for Agricultural Research. pp294-297, 467-469. 2nd (eds). John Wiley and sons, Inc. New York.

Gravois, K.A. and R.S. Helms, 1992. Path analysis of rice yield and yield components as affected by seeding rate. *Agronomy Journal* **84**: 1–4

Gupta P.K., S. Rustgi., P.L. Kulwal, 2005. Linkage disequilibrium and association studies in higher plants: Present status and future prospects. *Plant Molecular Biology* **57**:461-485.

Gupta, H. & P. Gupta, 2004. Neural tube defects and folic acid. *Indian Pediatr*, Vol. **41**, (6): 577-586.

Gupta, P., R. Varshney., P. Sharma and B. Ramesh, 1999. Molecular markers and their applications in wheat breeding. *Plant Breeding,* **118**: 369-390.

Habtu, A., I. Sache and J. C. Zadoks, 1996. A survey of cropping practices and foliar diseases of common bean in Ethiopia. *Crop Protection*, **15:** 179−186.

Hair, J. F. R. E. Tatham and W. Black. 1995. Multivariate data analysis. 5th Edition, Prentice-Hall. New Jersey. USA. 768 p.

Harlan, J.R.,1997. Geographic patterns of variation in some cultivated plants. *Journal of heredity*, **66**:184-191.

Harrison, I., M. Laverty and E. Sterling, 2004. Genetic Diversity. *Retrieved from the Connexions Web site: http://cnx. org/content/m12158/1.2*.

Huang, X., Q. Feng., Q. Qian., Q. Zhao., L .Wang., A. Wang., J. Guan., D. Fan., Weng,., Q. Huang and G. Dong, 2009. High-throughput genotyping by whole-genome resequencing. *Genome research*, ***19***(6), 1068-1076.

Holland, J.B., 2007. Genetic architecture of complex traits in plants. *Current opinion in plant biology,* **10**(2):156-161.

Ibarra-Perez, F and Ehdaie., G .Waines,1997. Estimation of out crossing rate in common bean. *Crop Sci* ,**37**:60-65

IFPRI (International Food Policy Research Institute), 2010. Pulses value chain potential in Ethiopia: constraints and opportunities for enhancing exports. Pulses diagnostics,

Ingvarsson, P.K., N.R.Street,., 2011. Association genetics of complex traits in plants. New Phytologist 189:909-22. DOI: 10.1111/j.1469-8137.2010.03593.x.

Imru Assefa, 1985. Bean Production in Ethiopia. Centro Internacional de Agricultura .'opical. 1985. Potentialfor Field Beans (Phaseolus vulgaris L) in West Asia and North Africa.

proceedings ofa regional workshop in Aleppo, Syria, 21-23 May, 1983. Cen'ro Internacional de Agricultura Tropical, Call, Colombia. 144p.

Jacinto-Hernandez, C., S. Azpiroz-Rivero., J.A. Acosta-Gallegos., H.Hernandez- Sanchez and I .Bernal-Lugo,., 2003. Genetic analysis and random amplified polymorphic DNA markers associated with cooking time in common bean. *Crop Sci* **43**:329–332

Johnson, H.W., H.F. Robinson and R.F. Comstock, 1955. Estimates of genetic and environmental variability in Soya bean. *J.Agronomy,* **47**: 314-318

Johnson, H.W., H.F. Robinson and R.F. Comstock, 1955b. Genotypic and phenotypic correlation in Soya bean and their implication in selection. *Agronomy J* **47**: 477-483.

Johnson, R. and D . Wichern, 1992. Applied multivariate statistical methods.*3rd Edi., Prentice Hall, Englewood Cliffs, NJ*.

Jolliffe, I., 2002. *Principal component analysis*. John Wiley & Sons, Ltd.

Kaiser, H. F., 1960. The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, **20**:141-151.

Kamfwa K., M .Mwala., P. Okori ., P .Gibson & C . Mukankusi, 2013a. Identification of QTL for *Fusarium* Root Rot Resistance in Common Bean Crop Improv. **27**, (4):406-418.

Kamfwa, K., Cichy, K.A. and Kelly, J.D., 2015c. Genome-wide association analysis of symbiotic nitrogen fixation in common bean. *Theoretical and Applied Genetics*, **128**(10):1999-2017.

Kamfwa, K., K.A. Cichy and J.D. Kelly, 2015b. Genome-wide association study of agronomic traits in common bean. *The Plant Genome*, **8**(2).

Karasu, A. and M. OZ, 2010. A study on coefficient analysis and association between agronomical characters in dry bean (*Phaseolus vulgaris* L.). *Bulg. J. Agric. Sci.,* **16**: 203-211.

Kassaye, N., 2006. Studies on Genetic Divergence in common bean (*Phaseolus Vulgaris* L.) introductions of Ethiopia. M.Sc. Thesis submitted to the school of graduate studies of Addis Ababa University.

Koinange, E.and P. Gepts, 1992. Hybrid weakness in wild *Phaseolus vulgaris* L. *Journal of Heredity,* **83**:135-139.

Kwak, M. and P. Gepts., 2009. Structure of genetic diversity in the two major gene pools of common bean (*Phaseolus vulgaris* L., Fabaceae). Theoretical and Applied Genetics **118**:979-992.

Pearson, T.A. and T.A. Manolio, 2008. How to interpret a genome-wide association study. *Jama*, *299*(11), pp.1335-1344.

Legendre, P. and L. Legendre , 1998. Numerical Ecology. 2nd ed. Amster dam: Elsevier. 853

Lima, M. S. De, Eustáquio, J., Carneiro, D. S., Crescêncio, P., Carneiro, S., & Santana, C., 2012. Characterization of genetic variability among common bean genotypes by morphological descriptors, 76–84.

Mahalanobis, P. C., 1936. On the generalized distance in statistics. *Proc. Natl. Inst. Sc. India.*, **2**: 49-55.

Malik, A., H. Zahida and S. Muhammad, 2008. Genetic diversity analysis of traditional and improved cultivars of Pakistani rice (*Oryza sativa* L.) using RAPD markers. *Electronic Journal of Biotechnology*, **11** (3).

Manju, P. R. and I. Sreelathakumary, 2002. genetic variability, heritability and genetic advance in hot chilli *(Capsicum Chinense* JACQ*.) Journal of Tropical Agriculture.* **40**: 4-6.

Mattson, S., E. Åkerberg., E. Eriksson., E. Koutler-Andersson., K. Vahtras, 1950. Factors determining the composition and cookability of peas. *Acta Agric Scand* **1**:40–61

Maiti, R.K. and V.P. Singh, 2007.Advances in Common Bean and Related Species, 1st edit. Publisher: Agro bios (International) *Plant Physiology*. **84**: 835-840.

Mesquita, F.R., A.D. Corre., C . Abreu., R . Lima., A. Abreu, 2007. Bean (Phaseolus vulgaris L.) lines: chemical composition and protein digestibility. *Cienc Agrotec* **31**:1114–1121.

Michael, E., 2012. *Introduction to the Practice of Statistics*, Third Edition University of Toronto, Inc.

Mohammadi, S. A. And B. M. Prasanna,2003. Analysis of genetic diversity in crop plants-salient statistical tools and considerations. *Crop Science.* **43**:1235–1248.

Molosiwa, O.O., S.B. Kgokong., B. Makwala,., C. Gwafila and M.G. Ramokapane, 2014. Genetic diversity in tepary bean (Phaseolus acutifolius) landraces grown in Botswana. *Journal of Plant Breeding and Crop Science*, **6**(12), 94-199.

Morojele M. E., D. N. M. Mbewe, 2015. Genetic Diversity Of Common Beans As Determined Using Morphological Markers. *Asian J Agri Biol*, **3**(1): 2-7.

Mudasir, S., P. A. Sofi., M. N. Khan., N. R. Sofi & Z. A. Dar, 2012. Genetic diversity, variability and character association in local common Bean (*Phaseolus vulgaris L.*) germplasm of Kashmir, **3**(3), 883–891.

Mukeshimana, G., L. Butare., P. B. Cregan., M. W. Blairm & J. D. Kelly, 2014. Quantitative trait loci associated with drought tolerance in common bean. *Crop Science*, **54** (3):923-938.

Myers, J.R., 2000. Tomorrow's snap bean cultivars, p. 39-51, *In* S. P. Singh, ed. Bean research production and utilization. University of Idaho, USA

Nechifor, B., F. Raluca., S. Lizia, 2011. Genetic variability, heritability and expected genetic advance as indices for yield components selection in common bean (Phaseolus vulgaris L.)‘, *Scientific Papers, UASVM Bucharest Series* A., **54** : 332-337

Nemli, S., K. Asciogul., B. Kaya., A. Kahraman., E . Dursu & B. Tanylac ,2014. Association mapping for five agronomic traits in the common bean (Phaseolus vulgaris L.). *Journal of the Science of Food and Agriculture*, **94**(15):3141-3151.

O’Neill, R., R. Snowdon and W. Kohler,2003. Population genetics aspects of biodiversity. Progress in Botany, **64**: 115-137. Springer Berlin Heidelberg.

Perez-Vega E, A. Paneda, C. Rodriguez-Suarez, A. Campa, R. Giraldez & J.J. Ferreira ,2010. Mapping of QTLs for morpho-agronomic and seed quality traits in a RIL population of common bean (*Phaseolus vulgaris* L.). *Theor Appl Genet* **120**:1367–1380.

Perkel, J., 2008. SNP genotyping: six technologies that keyed a revolution.*Nature Methods*, **5**(5), pp.447-454.

Perseguini, J.M.K.C., P.R. Oblessuc, J.R.B.F. Rosa, K.A. Gomes, A.F. Chiorato, S.A.M. Carbonell, A.A.F. Garcia, R.P. Vianello, and L.L. Benchimol-Reis, 2016. Genome-Wide Association Studies of Anthracnose and Angular Leaf Spot Resistance in Common Bean (Phaseolus vulgaris L.). *PloS one*, **11**(3), p.e0150506.

Podolak, E., 2010. Applied Biosystems' SOLiD sequencing system was commercialized in 2008, vastly improving upon their Sanger method-based automated sequencing system (Click to enlarge). *BioTechniques*, **48**(2):105-111.

Rafalski, A., 2002. Applications of single nucleotide polymorphisms in crop genetics. Current Opinion in Plant Biology **5**:94-100. DOI: Doi 10.1016/S1369-5266(02)00240-6.

Raffi, S. A. and U. K.Nath , 2004. Variability, heritability, genetic advance and relationship of yield and yield contributing characters in dry bean (*Phaseolus vulgaris L.*), J.l of biological science **4**(2) 157-159.

Ragoussis, J. ,2006. Genotyping technologies for all. *Drug Discov Today Technol* **3**:115–12

Rezene, Y., S. Gebeyehu & H. Zelleke, 2013. Morpho-physiological Response to Post-flowering Drought Stress in Small Red Seeded Common Bean (Phaseolus vulgaris L .) Genotypes, **2**(1). 42.

Ribeiro, N.D., P. Londero, A. Cargnelutti Filho, E. Jost,., Poersch, N.L. and Mallmann, C.A., 2007. Amino acid composition in common bean cultivars and applications for genetic breeding. *Pesquisa Agropecuária Brasileira*, **42**:1393–1399.

Risch, N. and K. Merikanga, 1996. The future of genetic studies of complex human diseases. Science **273**:1516–1517.

Rodrigues, J.A., N.D. Ribeiro., A. Filho., M. Trentin and P. Londero, 200. Cooking quality of common bean grain obtained in differents sowing periods. *Bragantia*, **64**: 369–376.

Rutger, J.N. and L.S.Beckham, 1970. Natural hybridisation of *Phaseolus vulgaris* x *Phaseolus coccineus L. Journal of the American Society for HortScience* **95**:659-661.

Semiha Guler and O. Huseyin, 2007. Relationships Between Leaf Chlorophyll and Yield Related Characters of Dry Bean (*Phaseolus vulgaris* L.). *Asian Journal of Plant Sciences,* **6***:* 700-703.

Sadeghi, A., K.Cheghamirza., H.R. Dorri ,2011. The study of morpho-agronomic traits relationship in common bean (*Phaseolus vulgaris* L.). *Biharen Biologist* **5**(2): 102-108.

Salehi, M., A. Faramarzi & N. Mohebalipour, 2010. Evaluation of Different Effective Traits on Seed Yield of Common Bean (Phaseolus vulgaris L .) With Path Analysis, *American Eurasian J. Agric. And Environ. Sci*, *9*, pp.52-54.***9***(1), 52–54.

Santos, J.B., M. L. Gavilanes, 2006. Botânica. *Feijão. Viçosa: UFV*.
SAS Institute Inc., 2014. Statistical analysis Software version 9.4, Cary, NC: SAS Institute Inc. USA.

Schmutz, J., P.E. McClean., S. Mamidi., G.A.Wu., S.B. Cannon., J. Grimwood., J. Jenkins., S . Shu., Q .Song., C .Chavarro., M . Torres-Torres., V. Geffroy., Moghaddam, S.M, Gao, D., Abernathy, B., Barry, K., P.N . Miklas., J.M .Osorno., J . Rodrigues., V. Thareau., C.A . Urrea., M . Wang., Y. Yu., M . Zhan., R.A. Wing., P.B . Cregan., D.S . Rokhsar., S.A . Jackson, 2014. A reference genome for common bean and genome-wide analysis of dual domestications. Nat Genet **46**:707–713.

Schulman, A., A. Flavell and T . Ellis,2004. The application of LTR retrotransposons as molecular markers in plants. *Methods Mol Biol*. **260**: 145-173

Seetharam, K., S. Thirumen and K . Paramasivam, 2009. Estimation of genetic diversity in rice (*Oryza sativa* L.) genotypes using SSR markers and morphological characters. *African Journal of Biotechnology*, **8** (10):.2050-2059.

Sentayehu, Alamirew, 1997. Variability and association of some agronomic characters with seed yield in haricot bean (*Phaseolus vulgaris L.*), a thesis submitted to the school of Graduate studies, Haramaya University of Agriculture.

Shi, C., A . Navabi and K. Yu, 2011. Association mapping of common bacterial blight resistance QTL in Ontario bean breeding populations. *BMC Plant Biology*, **11**(1):1

Shukla, S., A. Bhargava, A. Chatterjee and S.P. Singh, 2006. Estimates of genetic parameters to determine variability for foliage yield and its different quantitative and qualitative traits in vegetable amaranth (*A. tricolor*). *J. Genet. Breed*, **58**: 169-176.

Sidwell, R.J., E.L. Smith and R.W. McNew, 1967. Inheritance and interrelationships of grain yield and selected yield related traits in a hard red winter wheat cross. *Crop. Sci.* **16**: 650-654.

Singh, B.D., 2001. Plant breeding. Hindu university, Varanasi, Kalayani pubilisher, Ludhiana, New Delhi.

Singh, G. P., K. R. Maurya., B. Prasad and A. K. Singh, 1994. Genetic variability in *Capsicum annuum L. J. appl. Biol.* **4**: 19-22.

Singh, P. and R.B. Ceccerelli,1996. Estimation of heritability of crop traits from variety trial data. Technical manual No. 20, ICARDA, Aleppo, Syria.

Singh, P. and S.S. Narayanan, 1993. Biometrical Techniques in Plant Breeding. Kalyani, Publishers New Delhi.

Singh, R.K. and B.D. Chaudhar, 1985. *Biometrical Methods in Quantitative Genetic Analysis.* Kalyani Publishers, New Delhi, p.318

Singh, R.K., D.N . Singh., S.K. Singh and H.N. Singh, 1994. Genetic variability and correlation studies in foreign commercial hybrids of sugarcane. *Agric. Sci. Dig.*, **14**:103-107.

Singh, S.P., A . Molina., P . Gepts, 1995**.** Potential of wild common bean for seed yield improvement of cultivars in the tropics. *Can. J. Plant Sci*., **75**: 807-813

Singh, S.P., P. Gepts and D.G. Debouck, 1991a. Races of common bean (*Phaseolus vulgaris*, Fabaceae). *Economic Botany* **45**:379-396.

Singh, S.P., R. Nodari and P. Gepts, 1991b. Genetic diversity in cultivated common bean. I. Allozymes. Crop Science **31**:19-23.

Singh, S.P.,2001. Broadening the genetic base of common bean cultivars: A review. Crop Science **41**:1659 - 1675.

Singh,K.B., Geletu Bejiga and R.S. Malhorta, 1990.Associatition of some characters with seed yield in chick pea collection*. Euphytica* **49**: 83-88.

Sorrells, M., J . Yu, 2009. Linkage disequilibrium and association mapping in the Triticeae. In *Genetics and genomics of the Triticeae* (pp. 655-683). Springer US.

Stoetzer, H.A., 1984. Natural cross-pollination in bean in Ethiopia. Annual Report of the Bean Improvement Cooperative **27**:99-100.

Stoilova, T., G. Pereira & M. T.D. Sousa, 2013. Morphological characterization of a small common bean (Phaseolus vulgaris L.) collection under different environments. *Journal of Central European Agriculture*, **14**(3).

Stoskopf, N.C., D.T. Tom es and B.R . Christie, 1999. Plant Breeding: Theory and Practice. Scientific publishes, India, 435p.

Syvänen, A.C., 2001. Accessing genetic variation: genotyping single nucleotide polymorphisms. *Nature Reviews Genetics*, **2**(12), pp.930-942.

Tabor,H.K., N.J. Risch and R.M. Myers, 2002. Candidate-gene approaches for studying complex genetic traits: Practical considerations. Nat. Rev. Genet. **3**:391–397.

Teshale Assefa; Habtu Assefa, and P. Kimani, 2006. Development of improved Haricot bean germplasm for the mid-and low-altitude sub-humid agro-ecologies of Ethiopia. In Proceedings of the Workshop on food and forage legumes. 87-94.

Tumsa K., R. Buruchara, S.E. Beebe, 2014. Common Bean Strategies and Seed Roadmaps for Ethiopia, in: E.S. Monyo, G.C.L. Laxmipathi (Eds.), Grain Legumes Strategies and Seed Roadmaps for Selected Countries in Sub Saharan Africa and South Asia, TL-II Project Report, ICRISAT, India. 3–11.

Turner, S.D., 2014. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. bioRxiv. p.005165.

Vithanage, V., K. Anderson and M. Thomas,1995. Use of molecular markers in crop improvement of Macadamia. In *The Sixth Conference of the Australasian Council on Tree and Nut Crops* (pp. 11-15).

Voysest, O. and M. Dessert,1991. Bean cultivars: Classes and commercial seed types, p. 119-162, *In* H.F. Shwartz and M.A. Pastor-Corrales, eds. Bean production problems in the tropics. CIAT, Cali, Colombia.

Wang ,N. and J.K. Daun , 2005. Determination of cooking times of pulses using an automated Mattson cooker apparatus. *J Sci Food Agric*, **85**:1631–1635.

Weising, K., H . Nybom., K. Wolff and G. Kahl,2005. *DNA Fingerprinting in Plants: Principles, Methods and Applications*. Taylor and Francis Group, USA. 444pp.

Welsh, J.R., 1990. Fundamentals of Plant Genetic and Breeding. John Wiley and Sons. New York.

Whitt, S.R. and E.S. Buckler, 2003. Using natural allelic diversity to evaluate gene function. Methods. *Mol. Biol.* **236**:123–140.

Wondatir Z., Y. Mekasha, 2014. Feed resources availability and livestock production in the central rift valley of Ethiopia, *Int. J. Livest. Prod.* **5**:30–35.

Wortmann, C. S., 2006. *Phaseolus vulgaris* L. (common bean). Record from PROTA4U. Brink, M. & Belay, G. (Editors). PROTA (Plant Resources of Tropical Africa / Ressources végétales de l'Afrique tropicale), Wageningen, Netherlands.

Wortmann, C.S., R. A. Kirby., C.A. Eledu and D.J. Allen, 1998. Atlas of common bean production in Africa, CIAT publication No. 297.

Zhu, C., M .Gore., E.S. Buckler and J . Yu, 2008. Status and prospects of association mapping in plants. *The plant genome*, 1(1), pp.5-20.

7. APPENDICES

Appendix Table 1. Analysis of variance (Mean squares) for the 12 characters of 423 Common bean genotype.

Characters	Block (adj) (df=20)	Error (df=80)	Trt (adj) (df=422)	Among-controls (df=4)	Among-test (df=417)	Test v Control (df=1)	CV (%)
Days to flowering(days)	4.91	3.20	7.10**	92.90**	6.20**	37.10**	10.81
Days to 90% maturity(days)	5.91	2.00	6.00**	81.70**	22.90**	18.10**	1.93
Leaf chlorophyll	30.00	17.10	19.20^{NS}	0.40^{NS}	19.30^{NS}	83.50*	16.75
Plant Height (cm)	212.10	43.20	61.50* .	709.40**	53.10^{NS}	899.30**	15.32
Number of seeds pod^{-1}	0.30	0.30	0.40*	2.30**	0.40*	0^{NS}	13.07
Number of pod plant^{-1}	6.93	2.50	9.24*	36.80*	8.90*	33.50*	12.32
Vertical root pulling force resistance	55.20	62.50	88.10*	409.90**	83.70^{NS}	608.30**	43.20
Pod harvest index	51.80	44.20	58.10^{NS}	60.60^{NS}	58.20^{NS}	0.10^{NS}	8.91
Hundred seed weight(g)	31.40	410	48.20^{NS}	310.90*	44.60^{NS}	563.20*	18.24
Grain yield hectare^{-1} (kg)	1921950.40	288128	406027.90**	1206972.80**	389723.70**	3864133.40**	27.67

Appendix Table 2. Mean values of some of the studied common bean genotypes

ACCESSION CODE	DF	DHM	PLHT	PDPL	SDPD	RPS	GYLD
ADP0383	51	77	40.91	12.54	4.2	32.1	3873.19
ADP0555	47	73	55.71	12.1	4.2	25.83	3658.6
ADP0726	45	74	48.11	15.82	4	21.43	3644.67
ADP0166	43	73	57.91	14.9	6.4	42.7	3618.23
ADP0613	46	74	45.71	9.06	3.6	8.83	3609.73
ADP0074	44	75	34.11	10.62	3	25.53	3537.3
ADP0520	48	75	50.31	12.78	3.8	15.03	3518.69
ADP0608	48	74	40.51	7.46	3.8	16.8	3517.44
ADP0122	42	77	39.91	11.9	4	19.7	3461.5
ADP0620	51	76	46.71	10.26	3.8	24.2	3373.01
………	……	……	……	……	……	……	……
………	……	……	……	……	……	……	……
ADP0570	48	72	33.11	13.34	3.8	12.6	760.242
ADP0303	43	69	43.51	7.06	5.4	14.37	759.183
ADP0515	51	77	26.31	7.58	3.8	5.27	748.708
ADP0588	42	72	32.51	13.26	3.6	12.03	720.733
ADP0743	53	76	49.71	18.42	3.6	5.67	713.133
ADP0510	48	71	44.31	12.18	4	8.73	677.667
ADP0760	52	76	50.71	13.98	3.6	11.73	610.378
ADP0010	42	71	34.31	13.58	3.8	10.3	584.511
ADP0626	47	71	41.71	11.46	3.8	9	533.492
ADP0206	48	72	36.91	9.5	3.8	17.3	529.083

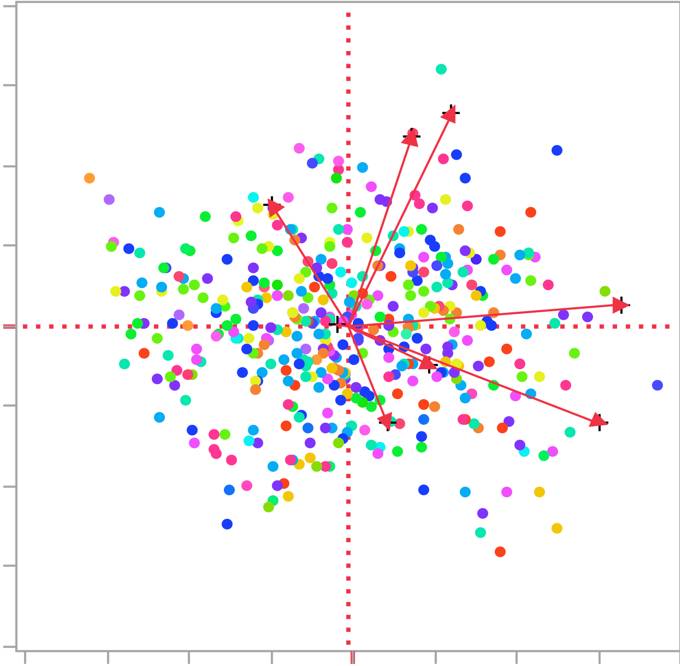
DF, days to flowering (number); DHM, days to harvest maturity (number); PLHT, plant height (cm); SDPD, seeds per pod (number); PDPL, pods per plant (number);RPS, root pulling force resistance (lb); GYLD, grain yield

Appendix Table 3. Distribution of 423 common bean genotypes in different clusters group

Cluster	Number of Genotypes	Accession Code										
		ADP0651	ADP0662	ADP0459	ADP0648	ADP0110	ADP0458	ADP0018	ADP0607	ADP0441	ADP0730	
		ADP0716	ADP0269	ADP0750	ADP0470	ADP0595	ADP0481	ADP0770	ADP0102	ADP0453	ADP0076	
1	34	ADP0001	ADP0680	ADP0639	ADP0591	ADP0757	ADP0040	ADP0537	ADP0042	ADP0466	ADP0579	
		ADP0442	ADP0019	IBADO	ADP0614							
		ADP0417	ADP0676	ADP0026	ADP0677	ADP0092	ADP0650	ADP0049	ADP0232	ADP0435	ADP0751	
		ADP0479	ADP0560	ADP0572	ADP0054	ADP0737	ADP0089	ADP0060	ADP0084	ADP0208	ADP0279	
2	49	ADP0460	ADP0433	ADP0656	ADP0660	ADP0023	ADP0008	ADP0610	ADP0519	ADP0571	ADP0592	
		ADP0583	ADP0774	ADP0598	ADP0740	ADP0587	ADP0576	ADP0053	ADP0427	DAB302	ADP0469	
		ADP0527	ADP0584	ADP0785	ADP0522	ADP0793	ADP0752	ADP0100	ADP0602	ADP0780		
3	20	ADP0271	ADP0611	ADP0090	ADP0604	ADP0557	ADP0725	ADP0015	ADP0029	ADP0450	ADP0395	
		ADP0718	ADP0392	ADP0741	ADP0212	ADP0512	ADP0633	TATU	ADP0242	ADP0428	ADP0472	
		ADP0623	ADP0654	ADP0566	ADP0635	ADP0580	ADP0629	ADP0443	ADP0455	ADP0526	ADP0666	
4	47	ADP0272	ADP0071	ADP0255	ADP0061	ADP0077	ADP0081	ADP0004	ADP0738	ADP0267	ADP0052	
		ADP0622	ADP0106	ADP0030	ADP0586	ADP0675	ADP0661	ADP0590	ADP0720	ADP0105	ADP0657	
		ADP0543	ADP0536	ADP0655	ADP0605	ADP0354	ADP0643	ADP0739	ADP0006	ADP0779	ADP0530	
		ADP0005	ADP0782	ADP0085	ADP0597	ADP0551	ADP0111	ADP0523				
5	11	ADP0617	ADP0670	ADP0471	ADP0620	ADP0436	ADP0464	ADP0225	ADP0775	ADP0125	ADP0121	ADP0556
		ADP0017	ADP0631	ADP0561	ADP0747	ADP0045	ADP0445	ADP0047	ADP0528	ADP0742	ADP0080	
		ADP0461	ADP0093	ADP0207	ADP0064	ADP0065	ADP0517	ADP0784	ADP0324	ADP0525	ADP0095	
6	38	ADP0050	ADP0609	ADP0101	ADP0063	ADP0683	ADP0638	ADP0792	ADP0529	ADP0119	ADP0753	
		ADP0011	ADP0777	ADP0168	ADP0562	ADP0532	ADP0118	ADP0769	ADP0679			
		ADP0599	ADP0658	ADP0663	ADP0107	ADP0719	ADP0524	ADP0577	ADP0062	ADP0681	ADP0075	
7	39	ADP0437	ADP0190	ADP0103	ADP0003	ADP0665	ADP0721	ADP0545	ADP0664	ADP0678	ADP0736	
		ADP0346	ADP0538	ADP0036	ADP0746	ADP0214	ADP0558	ADP0521	ADP0434	ADP0456	ADP0022	
		ADP0514	ADP0513	ADP0578	ADP0097	ADP0618	ADP0180	ADP0058	ADP0642	ADP0048		

8	25	ADP0027	ADP0044	ADP0099	ADP0649	ADP0367	ADP0439	REMEDA	ADP0012	ADP0550	ADP0098	
		ADP0109	ADP0438	ADP0776	ADP0034	ADP0429	ADP0127	ADP0581	ADP0059	ADP0108	ADP0687	
		ADP0778	ADP0013	ADP0731	ADP0037	DAB489						
9	19	ADP0368	ADP0668	ADP0247	ADP0452	ADP0070	ADP0041	ADP0056	ADP0205	ADP0057	ADP0729	
		ADP0682	ADP0035	ADP0596	ADP0473	ADP0024	ADP0641	ADP0021	ADP0732	ADP0554		
10	32	ADP0014	ADP0213	ADP0544	ADP0659	ADP0516	ADP0625	ADP0647	ADP0467	ADP0606	ADP0646	ADP0480
		ADP0632	ADP0083	ADP0735	ADP0645	ADP0032	ADP0667	ADP0025	ADP0621	ADP0744	ADP0117	ADP0567
		ADP0431	ADP0094	ADP0120	ADP0559	ADP0644	ADP0534	ADP0569	ADP0535	ADP0585	ADP0574	
11	4	ADP0520	ADP0608	ADP0074	ADP0122							
12	17	ADP0123	ADP0379	ADP0601	ADP0672	ADP0619	ADP0733	ADP0391	ADP0126	ADP0390	ADP0016	
		ADP0518	ADP0722	ADP0086	ADP0612	ADP0673	ADP0224	ADP0067				
13	33	ADP0087	ADP0748	ADP0468	ADP0531	ADP0624	ADP0773	ADP0483	ADP0745	ADP0199	ADP0781	ADP0475
		ADP0758	ADP0009	ADP0462	ADP0096	ADP0564	ADP0630	ADP0007	ADP0114	ADP0066	ADP0454	ADP0430
		ADP0482	ADP0031	ADP0628	ADP0474	ADP0043	ADP0640	ADP0432	ADP0310	ADP0051	ADP0540	ADP0674
14	15	ADP0345	ADP0717	ADP0186	ADP0734	ADP0055	ADP0549	ADP0038	ADP0072	ADP0088	ADP0457	
		ADP0636	ADP0355	ADP0376	ADP0546	ADP0615						
15	18	ADP0082	ADP0478	ADP0449	ADP0476	ADP0303	ADP0570	ADP0575	ADP0653	ADP0637	ADP0727	
		ADP0515	ADP0589	ADP0113	ADP0791	ADP0588	ADP0743	ADP0028	ADP0510			
16	5	ADP0206	ADP0626	ADP0010	ADP0760	ADP0508						
17	6	ADP0754	ADP0771	ADP0033	ADP0652	ADP0553	ADP0724					
18	4	ADP0555	ADP0726	ADP0166	ADP0613							
19	5	ADP0465	ADP0759	ADP0603	ADP0728	ADP0211						
20	1	ADP0783										
21	1	ADP0383										

Appendix Figure 1. Biplot scores of two principal components



Appendix Figure 2 Linkage disequilibrium decay plot generated by 4,416 markers