# PHENOTYPING AND QUANTITATIVE TRAIT LOCI MAPPING OF THE HARD-TO-COOK AND YIELD-RELATED AGRONOMIC TRAITS IN COMMON BEAN (*PHASEOLUS VULGARIS* L.)

**SAMUEL WANJOHI WAHOME**

**DOCTOR OF PHILOSOPHY**

**(Plant Breeding)**

**JOMO KENYATTA UNIVERSITY**

**OF**

**AGRICULTURE AND TECHNOLOGY**

**2023**

# Phenotyping And Quantitative Trait Loci Mapping of the Hard-To-Cook and Yield-Related Traits in Common Bean (*Phaseolus vulgaris* L.)

**Samuel Wanjohi Wahome**

**A Thesis Submitted in Partial Fulfilment of the Requirements for the Degree of Doctor of Philosophy in Plant Breeding of the Jomo Kenyatta University of Agriculture and Technology**

**2023**

## DECLARATION

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

Signature……..…… Date….....

**Samuel Wanjohi Wahome**

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

Signature…….…….. Date……

**Prof. Stephen Mwangi Githiri, PhD**

**JKUAT, Kenya**

Signature……. . Date……

**Prof. Geert Angenon, PhD**

**VUB, Belgium**

Signature…… Date……

**Dr. Peter K. Kinyanjui, PhD**

**JKUAT, Kenya**

## DEDICATION

I dedicate this thesis to my extended family for their encouragement, moral support, and love. And to my promoters and all scientist who made this work possible through their great contribution in various ways.

**ACKNOWLEDGEMENT**

Firstly, I wish to thank the almighty God for good health and strength during the period of my study. I would like to appreciate my supervisors Prof. Stephen Mwangi Githiri, Prof. Geert Angenon and Dr. Peter Kinyanjui Kahenya for their guidance and support from conceptualisation, experimentation, data analysis and write up of the research study. My utmost gratitude goes to 'Vlaamse Interuniversitaire Raad voor Universitaire Ontwikkelingssamenwerking' (VLIRUOS) for awarding me a study scholarship through through Legume Centre of Excellence for Food and Nutrition Security Project.

I am grateful to member of staff in the Department of Horticulture and Food Security for their guidance and support. I am grateful to Prof. Geert Angenon for his guidance and hosting me at Laboratory of Plant Genetics, VUB, Belgium. Special thanks to the technical staff of JKUAT Mr. Vodha Mumo and Mr. Patrick Amahema for their guidance in conducting various experiments during the study. Special thanks to my colleagues Dr. Esther Toili and Mr. John Kariuki for their moral support and ideas during this study.

I would like to acknowledge members of staff of SEQART AFRICA for genotyping services and guidance in Genome Wide Association Analysis. The assistance in linkage mapping and QTL analysis by Dr. Manje Gowda of CIMMYT is highly appreciated.

## TABLE OF CONTENT

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF APPENDICES

## LIST OF ABBREVIATIONS AND ACRONYMS

| | |
|---|---|
| **AEZ** | Agroecological Zone |
| **AFA** | Agriculture and Food Authority agency |
| **ANOVA** | Analysis of variance |
| **BecA** | Biosciences eastern and central Africa |
| **ACIAR** | Australian Centre for International Agricultural Research Basic Local |
| **BLASTn** | Alignment Search Tool for Nucleotides |
| **CMLM** | Compressed Mixed Linear Models |
| **CW** | Canadian Wonder |
| **CGIAR** | The consultative Group for International Agricultural Research |
| **CIAT** | International Center for Tropical Agriculture |
| **CRISPR** | Clustered Regularly Interspaced Short Palindromic Repeats |
| **CV** | Coefficient of Variation |
| **DArTseq** | Diversity Arrays Technology sequencing |
| **DNA** | Deoxyribonucleic acid |
| **ETC** | Easy-To-Cook |
| **FAO** | Food and Agriculture Organisation |
| **FAOSTAT** | Food and Agriculture Organisation Statistics |
| **GAPIT** | Genomic Association and Prediction Integrated Tool |
| **GBS** | Genotyping by Sequencing |
| **GCV** | Genotypic Coefficient of Variation |
| **GLM** | Generalised Linear Model |
| **GWAs** | Genome Wide Association studies |
| **HTC** | Hard-To-Cook |
| **IRLI** | International Research of Livestock Institute |
| **JKUAT** | Jomo Kenyatta University of Agriculture and Technology |
| **KALRO** | Kenya Agricultural and Livestock Research Organisation |
| **LCEFoNS** | Legume Centre of Excellence for Food and Nutrition Security |
| **LD** | Linkage Disequilibrium |

| | |
|---|---|
| **LSD** | Least Significant Difference |
| **LOD** | Logarithm of the odds |
| **MAF** | Minor Allele Frequency |
| **MAS** | Markers Assisted Selection |
| **MLM** | Mixed Linear Models |
| **$NO_3$** | Nitrate |
| **PME** | Pectin methylesterase |
| **PCR** | Polymerase Chain Reaction |
| **PC** | Principal component |
| **PCV** | Phenotypic Coefficient of Variation |
| **QTL** | Quantitative Trait Loci |
| **RAPD** | Random Amplified Polymorphic DNA |
| **NCBI** | National Center for Biotechnology Information |
| **NGS** | Next Generation Sequencing |
| **RCBD** | Randomized Complete Block Design |
| **RFLP** | Restriction Fragment Length Polymorphism |
| **RH** | Relative Humidity |
| **RILs** | Recombinant Inbred Lines |
| **SCA** | Specific Combining Ability |
| **SCAR** | Sequence Characterised Amplified Region |
| **SNP** | Single Nucleotide Polymorphism |
| **SSRs** | Simple Sequence Repeats |
| **RC** | Rosecoco |
| **SE** | Standard Error |
| **UM** | Upper Midland |

## ABSTRACT

Common bean (*Phaseolus vulgaris* L.) is a source of protein, carbohydrates, dietary fiber, and essential minerals to a large population worldwide. However, prolonged cooking time for dry beans is a factor that hinders the production and consumption of common bean seeds in many communities. The study aimed to assess common bean accessions for variation in cooking time of dry grains and identify regions in the genome that control the hard-to-cook (HTC) trait. Plant material used in this study comprised 257 common bean accessions sourced from the National Gene Bank of Kenya, Kenya Agricultural and Livestock Research Organisation (KALRO)-Embu, and from randomly selected local farmers' fields. $F_{2:6}$ recombinant inbred lines (RILs) were developed using common bean varieties GLP2, GLPx92, and accession GBK035420. Field experiments were carried out at Jomo Kenyatta University of Agriculture and Technology (JKUAT), Kenya. The experiments were laid down as a randomized complete block design (RCBD) with three replicates for over four seasons. Morphological data recorded included days to flowering, days to maturity, number of pods per plant, pod length, number of seeds for a pod, 100-seed weight, and grain yield. Cooking time was determined for each genotype using freshly harvested seeds and seeds stored at a temperature of 35$^{o}$C and 50% relative humidity for four months. Beans were soaked for 16 hours in distilled water and cooked at 96$^{o}$C in a thermostatically controlled water bath. DNA was isolated for each accession and RILs and genotyping were carried out using Single Nucleotide Polymorphism markers (SNPs) markers using Diversity Arrays Technology Sequencing (DArTseq). Data collected from field experiments was subjected to a normality test and analysis of variance (ANOVA). Pearsons phenotypic correlations analysis was conducted for the traits measured. Association analysis was conducted using the genome-wide analysis studies (GWAs) method to identify Quantitative Trait Loci (QTLs) associated with the hard-to-cook trait. Linkage mapping and QTL analysis were used to analyse data collected for RILs. The field experiment study revealed that days to flowering. 100-seed weight and grain yield had high broad-sense heritability and identified 19 common bean accessions that were both early maturing and high-yielding traits. The results revealed significant differences ($P \leq 0.05$) among accessions for all the traits evaluated, the seasonal and the interaction between accessions and seasons were also significant.

GWAS and QTL study identified QTLs region associated with all the agronomic traits. Agamous-like MADS-box transcripts like *Phvul.001G186400.1* locus co-localized with QTLs for days to flowering and maturity. The study revealed significant differences ($P \leq 0.05$) within and between fresh and aged bean accessions. Fresh seeds had a lower cooking time with a mean of 40.8 minutes and ranged from 28.1 to 72.2 minutes while aged seeds had a higher average cooking time of 54.1 minutes and ranged from 32.1 to 96.3 minutes. Genome wide association and QTL studies identified a region on chromosome 10 to be significantly associated ($P \leq 0.05$) with the cooking time of aged seeds. Consequently, two potential candidate genes *Phvul.010G038000* (galacturan 1,4-alpha galacturonidase) and *Phvul.010G038100* (polygalacturonase) were revealed. QTL analysis identified polygalacturonase/pectinase, pectin methylesterase, pectinesterase inhibitor, and galacturan 1, 4 alpha-galacturonidase enzymes to co-localize with the detected QTLs for cooking time. The characterized common bean accessions and the identified SNP markers could be utilized in breeding programs to improve common bean for cooking quality. The identified QTLs could be useful in introgressive hybridization of cooking time trait and implementation Markers assisted Selection (MAS) in the common bean.

# CHAPTER ONE
# INTRODUCTION

## 1.1 Background information

Common bean (*Phaseolus vulgaris* L.) holds a prominent position in human nutrition. It is a source of protein, carbohydrates, dietary fibre, vitamins, and essential minerals to a large population all over the world (Mecha *et al.,* 2018; Murube *et al.,* 2021). Common bean belongs to a large and diverse *Phaseolus* genus that contains approximately 70 to 80 species, of which only five are domesticated (Freytag and Debouck, 2002). Within this genus, common bean leads with wide geographical distribution, agronomic, nutritional, and economic value (Gepts, 2014). Studies have shown that the common bean has two distinct centres of genetic differentiation, namely the Middle American and Andean gene pools (Bitocchi *et al.,* 2012).

Common bean is cultivated in a wide range of environments which include tropics, sub-tropics and temperate regions, its adaptation is mainly limited to abiotic stresses such as extremes in temperatures, drought, Salinity in soils and photoperiod sensitivity (Sedlar *et al.,* 2019, Losa *et al.,* 2022). According to the database of the Food and Agriculture Organization (FAO) in 2023, the global production of dry common bean grains is 29 million, cultivated on 36 million hectares (FAOSTAT, 2023). Latin American is the largest producer of common bean, particularly Brazil and Mexico, with a production of approximately 5.5 million metric tons per year (Petry *et al.,* 2015). In Eastern African, Tanzania leads in production, followed by Uganda, Kenya and Ethiopia (Table 2.1) (FAOSTAT, 2023). Kenya produces about 722551 metric tons of dry common bean grains and has a deficit of 93100 metric tons (AFA, 2022).

Preference for common bean grains varies among consumers depending on taste, seed colour and size, ease of cooking, cooking quality and cultural factors (Wairimu, 2015; Swema and Mwinuka, 2021). Studies have shown that Kenyans still prefer the Grain Legume Project (GLP) varieties; GLP 585 (Red haricot), GLP2 (Rosecoco), GLPx92 (Pinto), GLP 24 (Canadian wonder). It was also revealed that consumers preference for each of these varieties ranked differently in urban and rural areas (Wairimu, 2015).

Common bean has a great number of varieties from the breeding programmes with differences in agronomic traits such as growth habit, duration to maturity, seed size and quality. These variations serve as genetic resources that have been extensively exploited in breeding programs to develop varieties (Elena *et al.,* 2010). In Eastern Africa, the Calima seed type (Red speckled or Rosecoco type) is the most popular followed by medium and small reds. Large reds including red kidney rank third in popularity, navy, whites, purples and black follow respectively (Okii *et al.,* 2014; Fisseha *et al.,* 2018).

The International Centre for Tropical Agriculture (CIAT), in collaboration with its partners in Africa, has been working to develop bean varieties that are tailored to consumer demands and preferences, considering attributes like pest and disease resistance, abiotic stress tolerance, and grain nutritional quality (Mukankusi *et al.,* 2019). Ongoing projects improving common bean in Kenya include the Australian Centre for International Agricultural Research (ACIAR) in collaboration with Pan-Africa Bean Research Alliance (PABRA), which is focusing on improving beans for rapid cooking and enhanced iron and zinc content in East Africa (Onyango, 2023). The Legume Centre of Excellence for Food and Nutrition Security (LCEFoNS) project of the Jomo Kenyatta University of Agriculture and Technology (JKUAT) is focusing on different stages of the legumes value chain, such as breeding, post-harvest storage, food processing and nutrition (Corporate Communications Office, 2023).

Common bean is commonly cooked by boiling in hot water above starch gelatinization temperature to produce a tender edible product and to develop aroma (Guzel and Sayar, 2012). The cooking duration varies among varieties and depends on previous storage environments, genetic differences and treatment before cooking. Cooking time has been reported to be a genetic trait which depends on seed characteristics such as the amount of insoluble pectins at cell wall and middle lamella, seed size, seed coat colour and thickness (Saha *et al.,* 2009).

Common bean seeds that have been stored under elevated temperatures (>25 ◦C) and relative humidity (>65%) develop the irreversible hard to cook trait characterised by prolonged cooking time (Perera *et al.,* 2023). Prolonged cooking consumes more fuel and

contributes to environmental degradation especially where firewood is the main source of energy (Nyamboki *et al.,* 2012). Development of easy to cook common bean varieties would not only improve consumer acceptance for dry beans but it would also save on cost of preparation and contribute to conservation of environment. Identification of the genomic regions controlling hard-to-cook (HTC) trait would assist in improving common bean varieties to produce easy to cook bean varieties.

### 1.2 Statement of the Problem

Common bean plays a critical role in nutrition security as a vital source of protein for a large population in third world countries and a source dietary fiber worldwide. The crop is rich in micronutrients such as vitamin B vitamins like thiamin, folic, niacin, and minerals such as iron and zinc (Mitchell *et al.,* 2009). The highest per capita consumption of common bean in Africa is found in Burundi, Kenya, and Rwanda ranging from 31 kg to 66 kg per year, equivalent to 180g per capita a day. It is a crucial dietary component in social institutions like schools and hospitals and for low-income households in rural and urban areas.

Culinary characteristics such as ease of preparation, the wholesomeness of grains after cooking, and digestibility are the main factors that influence consumers' choice of bean varieties (Wairimu, 2015). Dry beans require a long cooking time, especially at high altitudes, which is time and energy consuming. Common bean with HTC characteristic causes economic loss due to rejection by consumers for their poor texture and the need for increased energy required for cooking. HTC trait also affects the nutritional quality of common bean by lowering bioavailability of vitamins and protein (Perera *et al.*, 2023).

The main uses of energy in households are for cooking followed by lighting, firewood being the main source of this energy for cooking especially in rural areas. It has been reported that over-exploitation of this biomass fuel in rural households harms the environment, which indicates the importance of sustainable use of the scarce resource (Nyamboki *et al.,* 2012). There is a need to identify easy-to-cook common bean lines to provide breeding programs with raw materials and to understand the inheritance of the HTC trait assist breeders to choose a suitable breeding method.

### 1.3 Justification of the study

Common bean grains stored in adverse storage conditions develop hard-to-cook characteristics, which prolong cooking time and reduce digestibility (Perera *et al.*, 2023). The prolonged cooking time of stored common bean seeds leads to high fuel consumption and ultimately lowers the consumption of common beans. Studies have shown that cooking time is influenced by genetic differences among varieties, with some varieties being more susceptible to hardening during storage (Maryange *et al.*, 2010).

DNA analysis techniques make it possible to genotype large populations to identify alleles associated with the HTC trait. It would also allow the application of marker-assisted selection and other biotechnology techniques in the breeding process. This study utilised Next-generation sequencing (NGS) which uses Single Nucleotide Polymorphism (SNP) technology to identify Quantitative Trait Loci (QTLs) associated with common bean traits including cooking time. SNP technology is more practical, fast, inexpensive, and informative technique than the older generation markers that were gel-based (Gujaria-Verma *et al.*, 2016).

Reducing cooking time of common bean grains would contribute to environmental conservation through reduced overexploitation of firewood which is the main source of energy for cooking in the rural areas (Nyamboki *et al.*, 2012). Developing easier to cook varieties is the most economical solution to the HTC phenomenon. Easier to cook varieties will save on the cost of preparation, improve sensory qualities, and ultimately increase the consumption of common beans (Perera *et al.*, 2023).

Several studies have characterized a large population of common bean germplasm to identify genomic regions that control cooking time. Cichy *et al.*, (2015) used 206 common bean accessions of Andean origin and identified significant SNPs associated with cooking time on chromosomes 2, 3, and 6. Berry *et al.*, (2020) identified 10 QTLs on chromosomes 1, 2, 3, 5, 6, 10, and 11 using 146 recombinant inbred lines of common bean. The identified regions in these studies require further exploration to determine their robustness and stability across different genetic backgrounds and in a controlled storage condition. The identified easy-to-cook lines would serve as raw materials for breeding programs

developing easier-to-cook varieties. Understanding the inheritance of the HTC trait would assist breeders in selecting a suitable breeding method, and molecular markers identified associated with the HTC trait would allow the application of marker-assisted selection and identification of candidate genes that control the trait.

## 1.4 Objectives

### 1.4.1 Main objective

The overall objective of the study was to apply phenotyping and Quantitative Trait Loci mapping to elucidate the hard-to-cook trait in common bean

### 1.4.2 Specific objectives

1. To determine variation in yield-related agronomic traits within a panel of Kenyan common bean accessions
2. To investigate association of Single Nucleotide Polymorphisms (SNP) markers with cooking time in common bean accessions
3. To investigate SNP markers associated with variation in yield-related agronomic traits among common bean accessions
4. To assess Quantitative Trait Loci (QTL) associated with cooking time and selected yield-related agronomic traits in recombinant inbred lines of common bean

## 1.5 Null hypotheses

**$Ho_1$:** Locally conserved common bean accessions are not significantly different with respect to evaluated yield-related agronomic traits

**$Ho_2$:** Single Nucleotide Polymorphisms (SNP) markers are not significantly associated with cooking time in common bean

**$Ho_3$:** There are no Single Nucleotide Polymorphisms (SNP) markers significantly associated with variation in yield-related agronomic traits among common bean accessions

**$H_{O_4}$:** There are no quantitative trait loci (QTL) significantly linked to cooking time and selected yield-related agronomic traits of common bean

# CHAPTER TWO
# LITERATURE REVIEW

## 2.1 Botany, origin, and distribution of common bean

Common bean (*Phaseolus vulgaris* L.) belongs to the *Leguminosae* family that consists of approximately 600 genera. Within this genus, the common bean leads with wide geographical distribution, agronomic, nutritional, and economic value (Gepts, 2014). Common bean is a diploid species (2n=2x=22) with a haploid of 11 chromosomes and a genome of approximately 587 Mb (Schmutz *et al.,* 2014).

Common bean was domesticated by Middle American and South American cultures (Gepts, 1998) and dispersed to other regions of the world. Molecular, physiological, and morphological studies show that there exist two distinct centers of origin, namely Mesoamerican and Andean gene pools (Blair *et al.,* 2007; Burle *et al.,* 2010). The Andean gene pool is generally large-seeded and adapted to relatively higher altitudes and lower temperatures, while the Mesoamerican gene pool is small-seeded and adapted to lower altitudes and higher temperatures (Beebe *et al.,* 2011). Common bean has been adapted to a wide range of environments and is currently cultivated in many countries in the tropics, subtropics, and temperate regions (Burle *et al.,* 2010).

It is an annual herbaceous plant with oval-shaped leaves that are composite with three oval-shaped leaflets. The root system is made up of one main tap root and numerous secondary roots. The stem color can be green or may have anthocyanin coloration. Common bean has two major types of growth habits, bush and climbing types. The bush cultivars are day-neutral and early maturing. They grow to a height of 20 cm to 60 cm and have a lateral terminal inflorescence and determinate growth. Climbing types contain both day-neutral and short-day cultivars, they have indeterminate growth and grow up to 3 m in height, and they require staking for support. They have stems and fickle tendrils formed by the modification of terminal leaflets that allow their climbing habit (Kinyuru *et al.*, 2011).

Flowers are asymmetrical with petals of the color white or purple, they are hermaphrodite and predominately autogamous, with a 3% outcrossing. The standard of the flower is

reflexed, wings are of the same length or sometimes longer than the standard. The flower has 10 stamens, which are diadelphous with free vexillary stamens of equal length. Style is filiform, twisted, and bearded on the inner curve. Once the flower is pollinated it produces a pod that contains 3 to 12 seeds. The coloration of pods ranges from yellow to dark green (CIAT, 1986; Farrow and Andriatsitohaina, 2021). Seeds' colors vary from white, brown, yellow, red, black, purple, grey, and pinto; some are dotted, speckled, or have stripes. Common bean varieties vary in seed size, which depends on genetic variation and environmental conditions. Seeds can either be small, medium, or large-seeded depending on the weight of a random sample of a 100-seed (Angioi *et al.*, 2010).

### 2.2 Traits of common bean targeted for improvement through breeding

A wide diversity of traits in common bean exist in terms of growth habit, duration to maturity, resistance to biotic and abiotic stresses, seed size, seed color, cooking time, nutritional quality, and yield (Okii *et al*., 2014; Wairimu, 2015). Characterization and conservation of these traits in the common bean are crucial for improving the crop through breeding. As earlier mentioned, common bean shows a variation in growth habits ranging from determinate bush to indeterminate climbing. Schoonhoven and Pastor-Corrales (1987) categorized growth habits into five groups. The bush types are preferred because they do not require support and are hence convenient for market production (Okii *et al.*, 2014). Bush types are also popular for commercial production because they are early maturing and require less labor. However, climbing common beans is popular in highland areas because they are high-yielding and therefore ideal for small-scale farmers with a limited size of land (Okii *et al.*, 2014; Fisseha *et al.*, 2018).

The development of cultivars with improved resistance to biotic and abiotic stresses is an important goal in bean breeding throughout the world (Milkas *et al.,* 2006). The use of host resistance is a more economical method of controlling pests and disease, and therefore varieties that are resistant to disease pathogens and pests are preferred. It is expected that the distribution and severity of pathogens and pests are likely to be altered by climate change (Garret *et al.,* 2009). An increase in precipitation and humidity in some areas is likely to favor pathogenic fungi that cause foliar and root diseases. There have

been efforts to breed common bean varieties that are drought resistant using Mesoamerican and Durango races that are known to be good sources of drought stress resistance (Villordo-Pineda *et al.,* 2015). Plants that accelerate their cell cycle with early flowering and maturity and rapidly relocate their metabolites to grain production are known to escape drought (Beebe *et al.,* 2009).

Consumer preference for common bean grains depends on seed size and color. Seed size and color range from black, white, cream, yellow, brown-tan, red, and purple, they can be stripped, mottled or dotted, large-seeded or small-seeded. Seed color is determined by the presence and concentrations of flavanol glysides, anthocyanins, and condensed tannins on the seed while seed size depends on the genetic difference among varieties and environmental conditions (Reynoso *et al.,* 2006). Common bean varieties vary in seed size, those that weigh less than 25 g per 100-seed are classified as small-seeded, those that range from 25 to 40 g are medium-sized and those that weigh more than 40 g are large-seeded (Angioi *et al.,* 2010; Lei *et al.,* 2020). In eastern Africa, the Calima seed type (Red speckled) is highly popular and accounts for about 22% of common bean production. Medium and small reds follow in consumer preference accounting for approximately 20% of the production. Large reds including red kidney rank third in popularity accounting for about 10% of beans produced, navy, whites, purples, and black follow in popularity, respectively (Wortmann *et al.,* 1998).

Improving seed yield is a major objective for most common bean breeding programs (Vandemark *et al.,* 2014). The Andean bean types (large-seeded) are the most popular beans in Africa even though their yield is low compared to the middle American bean types (small-seeded) (Beebe, 2012). Seed yield is a polygenic trait that is conditioned by three yield components, number of pods per plant, number of seeds per pod, and seed weight (Negahi *et al.,* 2014; Kamfwa *et al.,* 2015). The knowledge of the association between these seed yield attributes may help in selecting a good donor and improving this trait.

Cooking is a fundamental part of bean preparation, it inactivates anti-nutritive factors, increases digestibility, and improves the sensorial quality of beans (Costa *et al.,* 2006).

Cooking beans is a high-energy demanding process due to their prolonged cooking time. Cooking of beans is known to be influenced by genetic differences, growth environment, post-harvest handling, storage conditions such as temperature and humidity, storage time, and treatments before cooking (Arruda *et al.,* 2012).

Common bean is a critical contributor to food security in the East African region and is a good source of nourishment second only to cereals. Common beans increase the protein content of the meal and improve the quality of the diet by a factor of 50% to 70% when served with cereals (Bressani *et al.,* 1988; Taptue, 2018). The common beans are also known to contain a small percentage of oligo and monosaccharides. The raffinose family of oligosaccharides, namely raffinose, stachyose, and verbascose are soluble carbohydrates known to contribute to flatulence in humans and animals. The amount of raffinose in common beans varies among common bean varieties (Reddy *et al.,* 1984). Common beans also contain Vitamin C, vitamin B and essential minerals (Mitova *et al.,* 2008; Mitchell *et al.,* 2009).

### 2.3 Nutritional quality and health benefits of common beans

Common bean is an important contributor to food security in the East African region, and it is a good source of nourishment second only to cereals but with a higher protein content ranging from 20% to 25% of their dry matter. The grain of the common bean plays a prominent role as a source of protein that supplements the low cereal proteins in developing countries and the diets of many vegetarians (Haddad and Tanzman, 2003). Common bean increases the protein content of the meal and improve the quality of the diet by a factor of 50 to 70% when served with cereals (Bressani *et al.,* 1988; Taptue, 2018). This is because beans are rich in lysine, complementing protein from cereals such as rice and corn. The ratio of essential and non-essential amino acids in immature seeds ranges from 0.82 (Mbithi-Mwikya *et al.,* 2000) to 0.93 (Slupski, 2010) in common dry beans. Common bean is an important source of protein where animal protein is limited among the low-income population, to whom beans are a daily food (Taptue, 2018). Nemeskeri (2012) reported that matured seeds have a higher protein content and mineral

composition than the green bean pods and dry bean seeds have a high level of serine, leucine, phenylalanine, and histidine contents (Gyori *et al.,* 1998).

Common bean contain about 60% carbohydrates of which two-thirds of it is in the form of starch, with a high ratio of amylose to amylopectin and a small percentage of oligo and monosaccharides (Hutchins *et al.,* 2012). The raffinose family of oligosaccharides, namely raffinose, stachyose, and verbascose are soluble carbohydrates found in legumes that have been identified as one of the important contributors to flatulence in humans and animals. The amount of raffinose depends on common bean varieties. Verbascose is predominantly in black seeds, red seeds, and mung beans. Stachyose is the major oligosaccharide in navy beans, pinto beans, red kidney beans, pink beans, and black eye beans (Mecha *et al.,* 2018).

Common beans contain a low saturated fat content and a high content of essential nutrients such as vitamin B components thiamin, folate and niacin, and essential minerals iron, zinc, magnesium, and potassium (Murube *et al.,* 2021). They are therefore used to substitute animal products in the diet as an alternative source of protein and the base for flours, starches, and fiber ingredients (Virginia, 2014). Common beans are rich in dietary fiber which is higher than that of other unrefined plant foods such as whole-grain meals (Galisteo, 2008). Dietary fiber includes indigestible polysaccharides such as cellulose, hemicellulose, oligosaccharides, and pectins, which affect gastrointestinal functions. A higher intake of dietary fiber reduces the risk of heart disease, diabetes, obesity, and some forms of cancers (Tosh and Yada, 2010).

Common bean is considered a good source of nourishment for people with diabetes. They have a low glycemic index, resulting to a slower and steadier rise in blood sugar levels. The protein and fiber content in beans help stabilize blood sugar levels, reducing the risk of spikes and crashes (Nchanji and Ageyo, 2021). Common has been reported to be more beneficial in weight management than in any other health issue, weight loss ranging from 2.24 kg to 2.93 kg in a period of four weeks has been reported (Wang *et al.,* 2020). However, promotion of consumption of beans for weight management faces challenge

due to the lack of standardization in terms of bean varieties, quantity consumed, and number of times administered (Nchanji and Ageyo, 2021).

Protease inhibitors such as trypsin inhibitors, decrease the digestibility of proteins in legumes (Van der Poel, 1990). Common bean genotypes vary greatly in trypsin inhibitor activity (Nemeskeri, 2012). Cruz *et al.,* (2003) reported that the digestibility varied among common bean varieties, the most coloured varieties showed low digestibility indicating that the pigments are related to the low protein quality of the varieties. The activity of trypsin was found to be high in coloured beans especially those of red kidney type (Nemeskeri, 2012).

Fertilizer application and soil types affect the ability of a plant to accumulate certain micronutrients. Organic fertilizer application was reported to increase vitamin C content in common beans (Mitova *et al.*, 2008). The development of HTC trait during storage has been reported to reduce the nutritional contents of common bean by lowering the availability of vitamins and protein (Perera *et al.*, 2023).

### 2.4 Common bean production and their ecological requirements

The global production of dry common bean grains is 29 million which is cultivated on 36 million hectares with an average yield of 1573 $kgha^{-1}$ (FAOSTAT, 2023). According to FAOSTAT data in 2020, Asia was the leading continent in common bean production, accounting for 50% of the global production. The top five dry bean-producing countries in the world from 2000 to 2019 were Myanmar, India, Brazil, China, and the United State (Nadeem *et al.,* 2021). In Eastern African, Tanzania leads in common bean production, followed by Uganda, Kenya and Ethiopia (Table 2.1) (FAOSTAT, 2023). According to Agricultural Foods Authority agency Kenya produces about 722551 metric tons of dry common bean grains and has a deficit of 93100 metric tons (AFA, 2022).

Recent statistics on common bean production in Eastern Africa are presented in Table 2.1. According to these statistics, Tanzania is the largest producer (1.3 million tons) of common bean with the land area devoted to common bean cultivation (1.01 million ha). Uganda follows with a production of 0.86 million tons with an area of 0.56 million ha under bean cultivation. Dry bean production in Kenya is practiced on approximately 1.2

million hectares with an actual average yield of 568.3 kg/ha under farmers' management (Table 2.1).

Common beans thrive in a wider range of soil types, ranging from light sand to heavy clays. The best soil for growth should be friable, well-drained, loam soils with high organic matter. Common bean requires the application of inorganic fertilizers especially potassium and phosphorous. Beans can fix atmospheric nitrogen using rhizobium bacteria which exist in root nodules. This bacterium is sensitive to cold and too much moisture (Anonimo, 1982; Alexandre and Oliveira, 2012).

**Table 2.1: Mean production figures of common bean in the eastern Africa region for the year 2021**

| Country | Area (ha) | Yield ($Kgha^{-1}$) | Production (Tons) |
|---|---|---|---|
| Tanzania | 1019495 | 1300.4 | 1325702.0 |
| Uganda | 461950 | 1852.6 | 855801.3 |
| Kenya | 1171869 | 568.3 | 666000.0 |
| Ethiopia | 355550 | 1741.2 | 619094.4 |
| Rwanda | 694481 | 708.4 | 491976.0 |
| Burundi | 827542 | 594.5 | 491967.8 |
| DRC | 483984 | 545.4 | 263961.0 |

FAOSTAT, 2023

Common beans are grown in areas with an average annual rainfall ranging from 900 to 2000 mm, which is well distributed during the growing season. Under moderate rainfall conditions, supplementary irrigation may be beneficial. Heavy rainfall adversely affects flower fertilization, resulting in a reduced pod set. The ideal altitude ranges between 1000 m to 2100 m above sea level (Greenlife, 2023). At higher altitudes the growth period is prolonged and there is an increased incidence of diseases because of colder conditions. Lower altitudes tend to have low rainfall, which is not ideal for common bean production unless irrigation water is available. The optimum temperature range is 16 to 24$^{\circ}$C. Below 10$^{\circ}$C bean plants are destroyed by chilling, while at temperatures above 30$^{\circ}$C blossom drop is very serious and may hamper pod formation and seed set (Burle *et al.*, 2010).

## 2.5 Seed yield

Seed yield improvement for Andean bean types has lagged compared to Mesoamerican beans, steady improvements in grain yield for Mesoamerican common beans have been reported resulting from genetic improvement and crop management (Singh *et al.,* 2007; Vandemark *et al.,* 2014). This is because greater genetic variability exists in the Mesoamerican gene pool than in the Andean gene pool (Bitochi *et al.,* 2013). As a result, more progress in improvement for seed yield and other traits has been reported in Mesoamerican bean types than in Andean bean types (Beebe *et al.,* 2012; Vandemark *et al.,* 2014). Transferring the favourable genes from Mesoamerican bean types into Andean bean types popular in Africa is a challenge due to incompatibility and linkage drag (Singh *et al.,* 2007, Beebe *et al.,* 2011).

Improving seed yield is a major objective for most common bean breeding programs, understanding the genetic architecture of this trait and its interaction with other yield components would lay a genetic foundation for improving seed yield (Vandemark *et al.,* 2014). Seed yield is a quantitative trait governed by multiple genes, and is conditioned primarily by three yield components namely, the number of pods per plant, number of seeds per pod, and seed weight (Negahi *et al.,* 2014). All the three yield components are quantitative and their interaction with seed yield is based on physiological and morphological features of the plant (Burbano-Erazo *et al.,* 2021). Understanding the interaction between these yield attributes and seed yield may help identify a suitable donor for this trait.

Yohannes (2020) reported a moderate broad-sense heritability (30-60%) for the number of seeds per pod. Several studies have observed a high heritability (>60%) for seed size (Henry *et al.,* 2019; Anunda *et al.,* 2019; Yohannes *et al.,* 2020). The number of pods per plant, seed weight, and biomass yield were reported as the first-order yield components with a positive direct effect on seed yield (Ghobary and Allah, 2010, Negahi *et al.,* 2014). High-yielding varieties were also reported to flower early, mature late, be taller, and had a higher number of primary branches per plant, pods per plant, and seeds per pod (Ashango and Alamerew, 2017). Several mapping studies on common beans have been carried out

to understand genomic regions contributing to seed yield and its components (Mukeshimana *et al.*, 2014; Kamfwa *et al.*, 2015; Briñez *et al.*, 2017; Sandhu *et al.*, 2018).

### 2.6 Storage, processing, and cooking quality of common beans

The condition and time of storage affect the cooking time of common bean grains. Storage in adverse conditions of high temperature and relative humidity makes beans susceptible to hardening phenomenon, loss of colour, and decrease water absorption capacity (Ousman *et al.,* 2013). Common bean grains stored under high temperatures above 25°C and high relative humidity above 65% developed the hard-to-cook trait with some varieties showing a higher increase in cooking time than others when stored in these adverse storage conditions (Perera *et al.*, 2023). Higher temperatures and relative humidity during storage reduce the hydration and swelling coefficient of common bean grains which in turn causes low water uptake (Wacu, 2016). Vindiola (1986) accelerated the hardening of beans by storing them at 100% relative humidity and a temperature of 45°C using a desiccator, he observed that the rate of hardening among varieties differed. Red beans tended to harden faster than brown beans while the white bean was the last to harden, but all beans eventually become uncookable after 7 to 9 days.

Due to the hard texture of common beans, they are generally consumed after soaking followed by cooking to produce acceptable sensory quality. Soaking is conducted at temperatures below starch gelatinization temperature to increase the water content. This accelerates the cooking process and helps leach out anti-nutritive compounds such as phytates and tannins (Fabbri and Crosbi, 2016). Soaking in de-ionized water and sodium carbonate is more effective in reducing the cooking time of common beans (Wacu, 2016). Guzel and Saya (2011) reported a significant increase in the percentage of splits for beans cooked under atmospheric pressure than those cooked under high pressure. They also found that higher pressure cooking increases the loss of solidness of legume seeds.

Hard to cook phenomenon is a textural defect that causes seeds to have poor soaking imbibition. Despite the prolonged cooking time, they do not achieve adequate texture due to the failure of cotyledon cells to separate upon cooking (Garcia *et al.*, 1998). Wacu (2016) reported that common bean variety red kidney showed a higher increase in cooking

time compared to rosecoco after having been subjected to temperatures of 27°C and relative humidity of 75%. Temperatures lower than 30°C and air humidity above 40% during grain filling lower the cooking time of beans (Zilio *et al.,* 2014). Long cooking time limits the utilization of common beans, it affects market price, processing cost, shelf life, and consumption patterns (Zaminder *et al.,* 2013).

The cooking time of dry beans is experimentally determined using an automated Matson pin dropper. In this method cooking time is defined as the time it takes to boil bean seeds in water for 80% of the seeds to be completely pierced by a 2 mm stainless steel pin (Wang and Daun, 2005). A cutting test using a Sun-Rheometer recorder is also used to determine the cooking time for beans. This method uses a cutting probe to measure the maximum force required to cut through a cooked bean at a speed of 100 mm per minute (Kinyanjui *et al.,* 2015). These methods are labour-intensive, slow, and expensive when evaluating many genotypes. Another method used is the use of bags; a plastic bag is used to hold other small bags containing bean seeds and hung in an upright position in boiling water (Maryange *et al.,* 2010). Finger pressing is used to determine the cooking time. In this method, the softness/hardness of the beans is determined by squeezing the cooked beans between the thumb and forefinger (Vindiola *et al.,* 1986). Beans are classified as cooked when the cotyledons are soft and free of graininess while hard beans are classified as not cooked. The percentage of cooked beans in the batch is determined as a function of time (Kinyanjui *et al.,* 2015). The bag method is preferred over the Matson cooker due to the low cost of material and the ability to cook many bean lines at a time (Maryange *et al.,* 2010). The use of a calibrated near-infrared spectroscopy and imaging is another method that can provide a high-throughput phenotyping method to assess cooking time where a large population of genotypes is involved (Mendoza *et al.,* 2018).

### 2.7 Mechanism of common bean cotyledon hardening

Several causes have been suggested to explain the hard-to-cook phenomenon which may occur also in combination. The widely accepted theory is the formation of insoluble pectates at the cell wall and middle lamella which renders the tissue more resistant to cell separation during cooking (Shomer *et al.,* 1990, Hentges *et al.,* 1990). Another theory

suggests degradation of cell membranes due to increased peroxidation within the cytoplasm leads to loss of membrane integrity (Richardson and Stanley, 1991). An increase in phenolic compounds that probably cause lignification of cells has also been proposed. Garcia *et al.,* (1998) confirmed this lignification by observing common bean seed cell walls that had been stored at 5°C and 40% relative humidity and 35°C at 75% relative humidity using scanning electron micrographs, which showed thickening of the middle lamella.

A dual enzyme mechanism has been proposed to explain the development of HTC conditions in storage at elevated temperatures and relative humidity (Jones and Boulter, 1983a). At high temperatures and relative humidity, pectin methylesterase (PME) hydrolyses pectin molecules forming pectic acid and methanol. Also, enzyme phytase hydrolyses phytic acid in the cells of the cotyledons to release inorganic phosphate and magnesium, while at the middle lamella, pectin methyl esterase hydrolyzes pectin to pectinic acid and methanol. The magnesium and calcium released in the cells migrate to the middle lamella and produce an insoluble magnesium pectinate and calcium pectinate that cements cells together hardening the cell wall (Jones and Boulter, 1983a; Kinyanjui *et al.,* 2015). According to this theory a Quantitative Trait Loci (QTLs) that control HTC could contain a locus related to the formation or breakdown of insoluble pectin.

This hypothesis is supported by the decrease in pectin solubility from 31.4% in fresh beans to 17.2% in hardened beans. Furthermore, the degree of pectin esterification decreased from 51% to 15% and phytic acid from 29 mg/g to 10 mg/g (Jones and Boulter, 1983a). In a similar study, Moscoso *et al.,* (1984) observed that cooking time increased by approximately 60% after incubating fresh beans in calcium ions with or without pectin methyl esterase. Varieties that have a shorter cooking time like rosecoco had more hot water-soluble pectin (8.44mg/g) than slow-cooking beans like pinto (5.51 mg/g) (Njoroge *et al.,* 2014). Steaming beans for a duration of two minutes at 120°C or ten minutes at 98°C retarded the rate of hardening during storage for a period of nine months at 90% relative humidity and temperature of 25°C (Molina *et al.,* 1976). Steaming could have reduced the enzyme activity that causes hardening. However, excessive heat treatment of

beans may cause hydrolysis of pectin and extra mobility of calcium and magnesium ions which reduce beans' cookability (Vindiola, 1986).

Irving (1980) found that fluoride ion, an inhibitor of phytase, prevented the hardening of pinto beans during soaking in acetate buffer at pH 4.7 and 41°C. A decrease of pH in the soaking solution of bean seeds from 4 to 1 causes the precipitated calcium and magnesium pectinate to dissolve as pectinic acid and which improves the cookability (Mattson, 1946). On the other hand, an increase in pH from 4 to 7, chelates phytate inside the cells to calcium and magnesium ions which keep pectinic acid in soluble form and allows beans to cook faster (Kon and Sanshuck, 2007). Kinyanjui *et al.*, (2015) demonstrated that beans soaked in calcium chloride and a solution of low pH (4) cooked slower than those soaked in high pH (8.5) solution. On the other hand, those soaked in low pH (4) solution cooked slower than those soaked in deionized water. He attributed this to the β-eliminative depolymerization of pectin favored by high temperatures and pH.

### 2.8 Genetics of cooking time

Gene action or heritability of a trait influences the breeding procedure applied to improve a trait, it aids a breeder to select a breeding procedure that will efficiently improve the performance of the genes (Dudley and Moll, 1969). Elia (2003) reported a 0.9 narrow sense heritability for the cooking time using $F_3$ and $F_4$ common bean populations developed using 16 varieties using North Carolina Design II mating scheme. Cooking time showed a degree of dominance less than one but larger than 0.0, which indicates that cooking time is governed by multiple genes with partial dominance for short cooking time over longer cooking time. Genes controlling cooking time for common bean were reported to be all nuclear genes with no influence from cytoplasmic genes (Elia, 2003).

Jacinto *et al.*, (2003) also estimated the heritability of cooking time at 0.74 using $F_6$ and $F_7$ recombinant inbred lines. Another study estimated narrow-sense heritability of cooking time at 0.47 using $F_2$ seeds (Mughi 2017). In latter study, some crosses showed a significant negative Specific Combining Ability (SCA) effect for cooking time. This confirmed the partial dominance of short cooking time over long cooking time reported earlier. The differences observed in these studies on the magnitude of heritability can be

attributed to the different populations used in these research studies. However, the results demonstrate that cooking time has a large genotypic effect which can be utilized to improve common beans through selection based on the trait itself (Elia, 2003).

In a similar study by Mashi (2006) using cowpea, it was observed that short cooking time was dominant over long cooking time and governed by two dominant alleles interacting at different loci. It has also been observed that high water absorption capacity is associated with short cooking time (Correa *et al.,* 2010). Elia (2003) estimated narrow-sense heritability for water absorption at 0.77 and a phenotypic correlation between water absorption and cooking time at -0.78, suggesting that the quantity of water absorbed during soaking can be used to predict the cooking time of bean accessions. Akinyele *et al.,* (1986) detected a positive correlation between cooking time and protein in cooked beans. In recent study that evaluated 242 recombinant inbred line (RIL) population, cooking time of pre-soaked seeds were found to exhibit high broad-sense heritability (0.68) (Bassett *et al.,* 2021).

### 2.9 Application of genetic markers

The application of DNA analysis techniques in common bean breeding programs has improved our understanding of genetic factors controlling various traits. Molecular markers have been utilized in breeding since the 1990s, the early generation markers were based on RAPDs and SCARs derived from these polymorphic fragments. Some of these markers are still in use in breeding programs like the SU91 marker for common bacteria blight (CBB) tolerance (Viteri *et al.,* 2014). The improvement of the DNA analysis techniques has led to the development of genetic maps with appropriate saturation degrees for mapping Quantitative Trait loci (QTLs).

Early generation markers are expensive and often cross-specific and therefore not ideal for high throughput marker screening. They also do not readily transfer across species and are hence limited in their use for comparative mapping. However, with the discovery of next-generation sequencing (NGS), Single Nucleotide Polymorphism (SNPs) has become more practical in genotyping and discovery of markers. Markers-assisted selection using SNP technology is much faster and inexpensive than the older generation markers that

were gel-based (Gujaria-Verma *et al.,* 2016). SNPs markers have been used in the construction of dense linkage maps that allow the identification of QTL associated with biotic and abiotic stress and other agronomic traits (Resende *et al.,* 2018; Sandhu *et al.,* 2018). Cortes *et al.,* (2011) reported that SNPs markers are useful in distinguishing Andean and Mesoamerican gene pools, he also validated 84 gene-based and 10 nongenic loci using KASPar technology in 70 genotypes. Sandhu *et al.,*, (2018) used SNP markers to map seed hardness using 85 $F_{2:7}$ RILs from a cross of hard and soft seeded black bean parents H68-4 and BK04-001, respectively, and revealed a major QTL on chromosome seven and two novel QTLs with significant effect on chromosome 1 and 2. A similar study identified QTL for cooking time on Pv8 and Pv10 using 242 recombinant inbred lines (RIL) population developed from a cross between Ervilha (Manteca) and PI527538 (Njano) using SNP markers (Bassett *et al.,* 2021)

**2.10 Quantitative Trait Loci (QTL) mapping**

Genetic control for most plant traits is reported to be predominantly due to genes with additive effects (Silva *et al.,* 2013). A more detailed understanding of traits at a molecular level is critical to improving beneficial traits of crops. Early genetic maps based on molecular and protein markers had estimated the size of the common bean genome at 1200 cM (Nodari *et al.,* 1993; Adam-Blondon *et al.,* 1994). A consensus of these maps was established in an integrated linkage map spanning 1226 cM and consisting of 563 markers that included random amplified polymorphic DNA (RAPD), restriction fragment length polymorphism (RFLP), sequence characterized amplified region (SCAR), isozyme and phenotypic markers (Freyre *et al.,* 1998). More linkage maps have subsequently been developed using different parents, segregating populations, traits, and molecular markers (Taran *et al.,* 2002; Beattie *et al.,* 2003; Blair *et al.,* 2006b; Sandhu *et al.,* 2018).

Many SNPs markers have been identified, which allow explorations of genetic diversity and population structure (Cichy *et al.,* 2015; Valdisser *et al.,* 2016). SNPs markers have been used to construct dense linkage maps that allow the identification of QTLs associated with various traits. SNPs markers have been utilized for the mapping of QTLs controlling for common bean traits like drought (Mukeshimana *et al.,* 2014), agronomic traits (Hoyos-

Villegas *et al.*, 2017), disease resistance (Nakedde *et al.*, 2016), grain yield (Resende *et al.*, 2018), and cooking time (Cichy *et al.*, 2015; Berry *et al.*, 2020).

A QTL for the number of pods per plant in common bean was reported by Koinange *et al.*, (1996) on Pv01 and Pv08 in a population of 65 $F_8$ recombinant inbred lines (RILs) from a cross of Mildas and G12873. Blair *et al.*, (2006b) reported a QTL of the same trait on Pv07, Pv09, and Pv11 in an inbred backcross population of 157 $BC_2$ $F_{3:5}$ from a cross between ICA Cerinza and G24404. Tar'an *et al.*, (2002) mapped the same QTL on Pv02 in 145 $F_{4:5}$ RILs from a cross of OAC Seaforth and OAC 95-4 navy bean. Kamfwa *et al.*, (2015) identified QTL for the number of pods per plant on Pv03 and Pv09 in 237 genotypes.

Several studies have reported QTLs associated with seed yield. Bettle *et al.*, (2003) found this QTLs for seed yield on Pv03 and Pv05 in a population of 110 $F_{5:7}$ RILs from a cross of W03391 and OAC Speedvale, Taran *et al.*, (2002) on Pv05, Pv09, and Pv11, Blair *et al.*, (2006b) found this QTL on Pv02, Pv04, and Pv09. Wright and Kelly (2011) identified QTLs for seed yield on Pv03, Pv05, Pv10, and Pv11 in a population of 96 $F_{4:5}$ RILs from a cross between Jaguar and 115 M. Mukeshimana *et al.*, (2014) mapped QTLs for seed yield on Pv03 and Pv09 in a population of 125 $F_{5:7}$ RILs from an inter gene cross of SEA 5 and CAL 96. Lastly, Kamfwa *et al.*, (2015) reported the QTLs for seed yield on Pv03 and Pv09 using 237 genotypes. Recently, through GWAs approach Resende *et al.*, (2018) detected two markers associated with grain yield on chromosome 3 using 188 common bean accessions.

Several studies have been conducted to map QTLs that control cooking time. Random amplified polymorphic DNA (RAPD) marker associated with cooking time was identified using 104 RILs, the marker explained 23% of the variation in cooking time (Jancinto-Hernandez *et al.*, 2003). Garcia *et al.*, (2012) mapped 6 QTLs that govern cooking time on chromosomes 1 and 9 using 105 polymorphic simple sequence repeats (SSRs) markers and 140 $F_{2:4}$ RILs. The most promising QTL was CT1.1 which explained 21% of the phenotypic variation.

In a recent study conducted by Berry *et al.,* (2020) using 146 RILs of common bean, 10 QTLs on chromosomes 1, 2, 3, 5, 6, 10, and 11 were identified, with the most robust QTLs being on chromosomes 3, 6, 10 and 11 that appeared in over two different environments. In a genome-wide association study, significant SNPs associated with cooking time were identified on chromosomes 2, 3, and 6 using 206 common bean accessions of Andean origin, the SNPs marker explained between 4 to 8.7% of the phenotypic variation (Cichy *et al.,* 2015).

The differences observed in the above examples could have resulted due to the limited number of markers, the type and size of the mapping populations used and the accuracy of the phenotyping techniques employed resulting in low resolution results in some studies, and the positioning of the candidate genes associated with the QTL become difficult.

### 2.11 Genome-Wide Association Studies

Genome-Wide Association Studies (GWAS) is a popular method used to identify QTLs associated with bean traits. GWAS invention was a result of several scientific discoveries early in the 21st century like the completion of the human genome project that provided much better context for the study of genetic variants (Hood and Rowen, 2013). GWAS has been used to identify QTL related to biotic stress (Zuiderveen *et al.,* 2016; Perseguini *et al.,* 2016), abiotic stress (Villordo-Pineda *et al.,* 2015), agronomic traits (Kamfwa *et al.,* 2015; Rasende *et al.,* 2018) and grain quality (Cichy *et al.,* 2015). It involves the application of molecular markers to plant breeding using statistical methods which enable breeders to estimate with accuracy the position and effects of genomic regions associated with variation in quantitative traits (Kafwa *et al.,* 2015). Genome-wide association method is dependent on genotype-environment interactions (GxE) (Beebe *et al.,* 2011), thereby giving an understanding of GxE at the molecular level, especially for a self-pollinating plant-like common bean that has been adapting to a constantly changing environment (Li *et al.,* 2003).

A few studies have used genome-wide association to find markers associated with cooking time. Cichy et al., (2015) used freshly harvested seeds, and identified SNPs that were

significantly associated with cooking time on chromosomes 2, 3, and 6. This study used 206 common bean accessions of Andean origin, the SNPs identified explained between 4 to 8.7% of the phenotypic variation (Cichy *et al.*, 2015). Another study combined GWAS and QTL analysis in a population of 922 lines of diverse origin to identify QTLs for cooking time (Diaz *et al.*, 2021).

GWAS is mainly concerned with determining alleles associated with various SNPs and making statistical comparisons to identify SNPs associated with a particular trait (Resende *et al.*, 2018). The presence of alleles in individuals with a particular trait is evidence that this allele may have an effect to some degree on this trait. Unlike in linkage mapping where only QTLs for which parents show differences can be identified, GWAS can be used to map various traits in a population at once. It has become a popular method that provides insight into explaining the total genetic variance, especially for traits with low heritability (Thorwarth *et al.*, 2017).

To determine the SNP-trait association, a mixed linear model (MLM) equation is used (Zhang *et al.*, 2008; Resende *et al.*, 2018)

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

Where Y= the vector of Phenotype, X= the vector of fixed effect of the SNP, P= the vector of fixed effect of population structure, K= Random effect of relative kinship, that is cryptic relatedness among genotypes from kinship matrix. ε= Error term which is assumed to be normally distributed. The Bonferonni correction for multiple tests with a global $\alpha= 0.05$ is used to determine the significance threshold for SNPs. Linkage Disequilibrium (LD) analysis is used to position candidate genes identified in the genomic regions surrounding significant SNPs (Zhang *et al.*, 2008; Resende *et al.*, 2018).

# CHAPTER THREE

# PHENOTYPING FOR YIELD-RELATED AGRONOMIC TRAITS IN A PANEL OF LOCALLY CONSERVED COMMON BEAN (*PHASEOLUS VULGARIS* L.) ACCESSIONS

## 3.1 Abstract

Characterization and conservation of common bean (*Phaseolus vulgaris* L.) germplasm is a critical step towards the genetic improvement of the crop. Seed yield improvement for the Andean bean types has lagged compared to Mesoamerican beans, thus improving seed yield in a major objective of common bean breeding programs in east Africa. This study assessed variation in 257 common bean genotypes which included 207 accessions obtained from the National Gene Bank of Kenya, 33 accessions from Kenya Agricultural and Livestock Research Organization (KALRO), 13 landraces collected from randomly selected local farmers' fields and four commercial varieties for yield-related agronomic traits which included number of pods per plant, number of seeds per pod and 100-seed weight. The experiments were laid out in a randomized complete block design with three replicates at Jomo Kenyatta University of Agriculture and Technology (Kenya) for four seasons between 2019 and 2020. Significant differences ($P \leq 0.05$) existed among the common bean accessions for all traits studied. Seed yield ranged from 220.6 kg/ha to 4641.9 kg/ha (KNB0106) among the accessions with a mean of 1267.0 kg/ha. Significant ($P \leq 0.05$) positive correlation was recorded for days to flowering and days to maturity (0.73), while 100-seed weight had a significantly negative correlation with the number of pods per plant (-0.66) and the number of seeds per pod (-0.65). High (>20%) broad-sense heritability was recorded for 100-seed weight (89.0%), days to flowering (76.8%), and grain yield (60.5%). Nineteen accessions that were both early maturity and high-yielding traits were identified. Higher seed yields were recorded for large-seeded and climbing genotypes compared to small-seeded and bush types. Seasonal differences were significant with higher yields during the long rain seasons. Common bean accessions characterized can be exploited in breeding programs.

## 3.2 Introduction

Common bean (*Phaseolus vulgaris* L.) holds significant importance for human nutrition due to its nutritional composition and various health benefits. The crop is a major source of protein, carbohydrates, dietary fiber, and essential minerals to a large population globally (Gepts *et al.*, 2008, Murube *et al.*, 2021). A wide diversity of traits exists in common bean regarding growth habit, duration to maturity, resistance to biotic and abiotic stresses, seed size, seed color, and yield (Okii *et al.*, 2014; Fisseha *et al.*, 2018). These variations serve as genetic resources that have been extensively exploited in breeding programs to develop varieties (Pérez-Vega *et al.*, 2010). Previous studies have demonstrated that the common bean has two distinct centers of genetic differentiation, namely the Middle American and Andean gene pools (Bitocchi *et al.*, 2012). The large-seeded types (Andean) are the most popular beans in Africa though their yield has been reported to be lower compared to the small-seeded types (Middle American) (Beebe 2012).

The growth habit of the common bean has been reported to range from determinate bush to indeterminate climbing (Farrow and Andriatsitohaina, 2021). The bush types are popular and preferred because they do not require support (Okii *et al.*, 2014), they are also early maturing, and can easily be mechanically harvested. On the other hand, climbing common beans have higher yields and are ideal for small-scale farmers in highland areas (Okii *et al.*, 2014; Fisseha *et al.*, 2018). Consumer preference for common bean grains depends on seed color and seed size. The most popular seed type in Eastern Africa is Calima (Red speckled or Rosecoco type) followed by medium and small red, while the large red including red kidney ranks third in popularity (Farrow and Andriatsitohaina, 2021).

Heritability estimates of a trait indicate how much variation can be attributed to genetic variation and the environmental influence in the expression of the trait. The heritability estimates therefore aids a breeder to select a breeding procedure that will efficiently improve the performance of the genes involved (Yohannes *et al.*, 2020).

Improving seed yield is a primary objective for most common bean breeding programs (Vandemark *et al.,* 2014). Seed yield is a polygenic trait that is conditioned by three yield components, the number of pods per plant, the number of seeds per pod, and seed weight (Kamfwa *et al.,* 2015). The knowledge of the association between these seed yield attributes may help in the selection of a suitable donor to improve this trait. The objective of this study was to assess common bean accessions for variation in yield related agronomic traits which are essential for characterization, conservation, and variety improvement.

### 3.3 Materials and Methods

#### 3.3.1 Field experimental site

Field experiments were carried out for four seasons at Jomo Kenyatta University of Agriculture and Technology (JKUAT) in Kiambu County Kenya. The site is located at coordinates 3$^{o}$ 35'South and 36$^{o}$ 35' East at an elevation of 1520 m above sea level. The area falls within upper midland belt (UM) of agroecological zone (AEZ) IV (Jaetzold and Schmidt, 1983). The site experiences bimodal pattern of rainfall with an annual mean of 856 mm. Long rains occur between March and May while short rains occur between October and November with a monthly mean of 142 mm and 116 mm respectively. The mean annual maximum and minimum temperatures are 20$^{o}$C and 30$^{o}$C respectively. The monthly rainfall and temperature during 2018 and 2019 are shown in Appendix 4. The area has three types of soils namely, shallow clay soils, sandy clay soils, and deep clay soils (vertisols).

#### 3.3.2 Plant materials

Common bean genotypes in this study included 257 accessions sourced from the National Gene Bank of Kenya, 33 accessions from the Regional Agricultural Research Centre-Kenya Agricultural and Livestock Research Organization (KALRO)-Embu, 13 landraces collected from randomly selected local farmers' fields, and four commercial varieties (GLP-2; GLP-24, GLPx92 and GLP1192a). The growth habits of the genotypes were type

I & II, III and IV with 124, 84 and 49 accessions, respectively. The genotypes belonged to different market classes found in the region (Table 3.1 and Figure 3.1).


**Figure 3.1: A sample of seeds of various common bean accessions used in this study grouped into their respective seed classes based on seed colour and size**

**Table 3.1: Common bean accessions used in this study**

| Seed class | Description | Number of accessions |
|---|---|---|
| Pintos | Cream with brown specks-GLPx92 type | 22 |
| Sugars | Cream and can be speckled | 39 |
| Calima | Rosecoco type | 25 |
| Small reds | Red haricot type | 15 |
| Large reds | Canadian wonder type | 17 |
| Purples | Mwezimoja type | 11 |
| Medium whites | Medium and large whites | 13 |
| Brown and tan | Brown and orange | 19 |
| Cariocas | Red and Red specks | 28 |
| Yellow | Yellow coloured | 8 |
| Blacks | Black coloured | 23 |
| Navy | Small whites | 37 |
| **Total** | | **257** |

### 3.3.3 Experimental design and trial management

The experiment was conducted for four seasons during the long and short rains seasons of 2019 and 2020. The trial was laid out as a randomized complete block design with three replicates. The bean lines were grown in single rows of 5m in length with an inter-row spacing of 50 cm, the intra-row spacing of all the genotypes was 20 cm. Compound N.P.K (17.17.17) fertilizer was applied at a rate of 200 kg/ha, evenly spread, and thoroughly mixed with soil. The bean seeds were planted and lightly covered with soil.

The first manual weed control was conducted two weeks after emergence and the second one at 40-50 days thereafter (Figure 3.2). Insect pests and diseases were controlled by the application of chemical pesticides diazinon at a rate of 40 ml/20 litre and 500 g/litre pymetrozine at a rate of 400 to 600 g/ha. Before flowering, the climbing genotypes were supported with 1.5 m long sticks to prevent lodging.

### 3.3.4 Agronomic data collection

Data collection started one month after planting. Quantitative data recorded is described in Table 3. 2. Qualitative data collected included growth habits (climbing, semi climbing, and bush) and seed colour.


**Figure 3.2: Field experiments at Jomo Kenyatta University of Technology experimental farm**

**Table 3.2: Quantitative agronomic traits recorded in field trials**

| Trait | Units | Description |
|---|---|---|
| Days to flowering | d | Number of days from planting to the date when 50% of plants have one or more flowers |
| Days to maturity | d | Number of days after planting to the date when 50% of the plants have reached physiological maturity |
| Number of pods | no. | The average total number of pods from five randomly selected plants per plot at maturity. |
| Pod length | cm | Average pod length of five randomly selected pods from each plot measured using a ruler |
| Number of seed per pod | no. | The average number of seeds of five randomly selected pods from each plot |
| Seed weight | g | Weight of a random sample of 100 seeds from each plot |
| Grain yield | g | Total seed yield per plot which will be used to extrapolate Yield per hectare |

### 3.3.5 Statistical analysis

**Analysis of variance**

Qualitative data collected was used to group the accessions into their growth habits and market classes. Quantitative data collected from field experiments were combined over seasons and analyzed using R software (version 4.0.2). All traits' means were separated using Fisher's Least Significance Difference test (LSD) at 5% level. The significance of correlations was tested at 0.05 and 0.01 levels of probability.

**Phenotypic and genotypic coefficient of variation**

The estimates of phenotypic and genotypic coefficient of variation were calculated as described by Singh and Chaudhary (1985) as follows

$$PCV\ (\%) = \frac{\sqrt{\mathrm{Vp}}}{\mathrm{Mean}} X100, \qquad GCV\ (\%) = \frac{\sqrt{\mathrm{Vg}}}{\mathrm{Mean}} X100$$

Where PCV is the phenotypic coefficient of variance, Vp is the phenotypic variance, GCV is genotypic coefficient of variance, and Vg is the genotypic variance, GCV and PCV values were categorized as low (0-10%), moderate (10-20%) and high (20% and above) as indicated by Burton and de Vane (1953).

**Heritability**

Heritability was estimated as the ratio of genotypic variance to phenotypic variance as described by Singh and Chaudhary (1985).

$$H^2 = \frac{Vg}{Vp} X100$$

Where $H^2$ is broad-sense heritability, Vp is phenotypic variance and Vg is genotypic variance. Heritability percentage values were categorized as low (0-30%), moderate (30-60%), and high (60% and above) as described by Johnson *et al.,* (1955).

Cluster analyses were carried out based on Euclidean distance method. Complete clustering method was used to determine the genetic relationship among genotypes based on the agronomic data.

### 3.4 Results

#### 3.4.1 Descriptive statistics for agronomic and seed yield traits

The means, range, variance, and coefficient of variation for recorded traits are summarized in Table 3.3. The coefficient of variation ranged from 4.2% (days to flowering) to 36.1% (number of pods per plant). The highest coefficient of variation registered was for the number of pods per plant and grain yield at 36.1% and 32.3% respectively.

**Table 3.3: Descriptive statistics for days to flowering, days to maturity and yield-related for 257 common bean accessions grown at Juja in 2018 and 2019**

| Trait | Mean | Range | Min | Max | Variance | SE | CV% |
|---|---|---|---|---|---|---|---|
| Days to flowering | 37.8 | 12.6 | 32.5 | 45.1 | 16.4 | 0.09 | 4.2 |
| Days to maturity | 82.0 | 17.5 | 73.9 | 91.4 | 33.1 | 0.13 | 4.6 |
| Pods/plant (no.) | 12.6 | 24.0 | 5.9 | 29.9 | 56.6 | 0.19 | 36.1 |
| Pod length (cm) | 9.9 | 7.8 | 6.9 | 14.7 | 4.4 | 0.05 | 14.6 |
| Seed/pod (no.) | 4.6 | 3.7 | 3.0 | 6.7 | 1.1 | 0.02 | 17.3 |
| Seed weight (g) | 36.9 | 54.9 | 15.0 | 69.9 | 182.9 | 0.30 | 15.5 |
| Yield (kg/ha) | 1267.0 | 4421.3 | 220.6 | 4641.9 | 678470.6 | 18.23 | 32.3 |

n=257, SE=Standard error, CV=Coefficient of variation.

### 3.4.2 Estimation of genetic variables for traits measured

The extent of variance components and heritability estimates of seven common bean traits are presented in Table 3.4. The phenotypic coefficient of variation ranged from 6.3% (days to maturity) to 51.4% (grain yield). Days to flowering and days to maturity recorded a low phenotypic coefficient of variation (0-10%) of 8.7% and 6.3% respectively. On the other hand, pod length, seeds per pod, 100-seed weight, number of pods per plant, and grain yield showed a high phenotypic coefficient of variation (>20%) of 20.8%, 22.8%, 36.5%, 43.6%, and 51.4% respectively (Table 3.4).

Genotypic coefficient of variation ranged from low (0-10%), moderate (10-20%) to high (>20%). Days to maturity and days to flowering had a low genotypic coefficient of variation of 4.2% and 7.6%, respectively, while the number of seeds per pod and pod length had a moderate genotypic coefficient of variation of 12.6% and 14.6%, respectively. High genotypic coefficients of variation were recorded for the number of pods per plant, 100-seed weight, and grain yield of 24.4%, 34.3%, and 40.0%, respectively (Table 3.4).

Grain yield, days to flowering and 100-seed weight recorded high broad-sense heritability ($H^2$>0.6) of 60.5%, 76.8%, and 89.0%, respectively. In contrast, the number of pods per plant, number of seeds per pod, days to maturity, and pod length showed moderate broad sense heritability ($H^2$=0.3-0.6) of 32.2%, 38.7%, 51.1%, and 52.4%, respectively (Table 3.4).

**Table 3.4: Estimation of genetic variables for days to flowering, days to maturity and yield-related traits for 257 common bean accessions grown at Juja in 2018 and 2019**

| Components | DF (days) | DM (days) | Pods/plant (no.) | Pod length (cm) | Seed/ pod (no.) | 100-Seed weight (g) | Yield (kg/ha) |
|---|---|---|---|---|---|---|---|
| E variation | 2.5 | 14.8 | 20.8 | 2.1 | 0.8 | 21.8 | 303046.0 |
| P variation | 10.7 | 26.9 | 30.3 | 4.2 | 1.1 | 181.8 | 542948.4 |
| G variation | 8.2 | 12.1 | 9.5 | 2.1 | 0.3 | 160.0 | 239902.4 |
| $H^2$ (%) | 76.9 | 51.1 | 32.2 | 52.4 | 38.7 | 89.0 | 60.5 |
| PCV % | 8.7 | 6.3 | 43.6 | 20.8 | 22.8 | 36.5 | 51.4 |
| GCV % | 7.6 | 4.2 | 24.4 | 14.6 | 12.6 | 34.3 | 40.0 |

DF=Days to flowering, DM=Days to maturity, $H^2$=Broad sense heritability, E=Environment, P=Phenotypic, G=Genotypic, PVC=Phenotypic coefficient of variation, GCV=Genotypic coefficient of variation

### 3.4.3 Duration to flowering and maturity, and yield-related traits

There were highly significant ($P \leq 0.05$) differences for all the traits studied among the 257 common bean accessions (Appendix 1). The seasonal and the interaction effect between season and common bean accessions also significantly influenced all the evaluated traits. For example, the mean for days to flowering, days to maturity, 100-seed weight, and grain yield were higher in long rain seasons than in short rain seasons. However, the average pods per plant and pod length were higher during the short rain than in long rain seasons.

**Table 3.5: Mean values for days to flowering, days to maturity and yield-related traits for top 10, bottom 5 and checks of common bean accessions grown in Juja in 2018 and 2019 ranked based on grain yield**

| Name | Season | DF (days) | DM (days) | Pods/plant (no.) | Pod length (cm) | Seed/ pod (no.) | 100-Seed weight (g) | Yield ($Kgha^{-1}$) |
|---|---|---|---|---|---|---|---|---|
| **High yielding** | | | | | | | | |
| KNB0106 | S1 | 39.3 | 85.3 | 19.0 | 10.0 | 5.3 | 41.5 | 4926.3 |
| | S2 | 37.5 | 83.0 | 22.8 | 10.5 | 5.2 | 34.0 | 4357.5 |
| NUA700 | S1 | 40.8 | 88.0 | 12.5 | 11.5 | 3.5 | 42.5 | 2967.5 |
| | S2 | 38.5 | 85.3 | 16.3 | 11.9 | 3.8 | 37.3 | 2370.0 |
| GBK035092 | S1 | 38.8 | 82.0 | 14.0 | 10.3 | 5.3 | 37.5 | 2765.0 |
| | S2 | 39.0 | 79.3 | 13.5 | 11.1 | 4.9 | 31.5 | 2281.3 |
| GBK035025 | S1 | 38.5 | 82.5 | 12.5 | 11.3 | 5.0 | 44.3 | 3125.0 |
| | S2 | 39.8 | 78.8 | 12.8 | 11.8 | 3.6 | 43.8 | 1671.3 |
| GBK035051 | S1 | 37.5 | 83.5 | 13.5 | 11.3 | 5.5 | 25.5 | 2845.0 |
| | S2 | 37.0 | 84.0 | 16.3 | 10.2 | 5.1 | 27.3 | 1890.0 |
| KNB0107 | S1 | 35.0 | 83.8 | 12.5 | 9.5 | 4.8 | 24.8 | 1721.3 |
| | S2 | 34.8 | 79.8 | 17.8 | 11.3 | 5.2 | 26.3 | 3010.0 |
| NUA637 | S1 | 35.8 | 83.0 | 11.5 | 14.3 | 5.3 | 54.5 | 3078.8 |
| | S2 | 35.0 | 81.0 | 9.8 | 14.7 | 4.8 | 41.3 | 1645.0 |
| GBK035447 | S1 | 42.3 | 87.0 | 13.0 | 12.0 | 5.8 | 40.8 | 2913.8 |
| | S2 | 42.0 | 85.0 | 14.6 | 12.4 | 4.3 | 38.0 | 1801.3 |
| NUA662 | S1 | 39.5 | 84.5 | 9.0 | 10.8 | 4.0 | 62.8 | 3170.0 |
| | S2 | 38.0 | 84.0 | 13.1 | 10.2 | 3.2 | 36.3 | 1501.3 |
| NUA640 | S1 | 39.8 | 83.5 | 11.5 | 10.5 | 4.5 | 57.5 | 3241.3 |
| | S2 | 39.0 | 80.0 | 12.9 | 9.3 | 3.2 | 39.8 | 1223.8 |
| **Low yielding** | | | | | | | | |
| GBK034995 | S1 | 34.3 | 77.8 | 9.5 | 11.8 | 5.3 | 36.5 | 558.8 |
| | S2 | 33.7 | 81.3 | 12.7 | 9.9 | 3.7 | 29.3 | 176.7 |
| GBK035023 | S1 | 45.3 | 90.5 | 10.0 | 12.5 | 4.5 | 54.0 | 576.3 |
| | S2 | 41.3 | 86.8 | 3.8 | 10.9 | 3.3 | 48.3 | 75.0 |
| GBK035295 | S1 | 38.0 | 82.3 | 15.5 | 9.8 | 6.5 | 21.8 | 350.0 |
| | S2 | 37.8 | 82.0 | 13.3 | 8.7 | 4.3 | 34.8 | 227.5 |
| GBK035320 | S1 | 40.3 | 72.5 | 10.0 | 9.3 | 5.0 | 21.3 | 300.0 |
| | S2 | 38.5 | 83.5 | 8.3 | 8.6 | 3.9 | 32.8 | 200.0 |
| GBK035350 | S1 | 37.0 | 81.3 | 15.0 | 7.3 | 4.3 | 20.8 | 248.8 |
| | S2 | 37.0 | 87.8 | 14.5 | 6.6 | 3.8 | 35.5 | 192.5 |
| **Commercial varieties** | | | | | | | | |
| GLPx92 | S1 | 34.5 | 85.5 | 13.5 | 8.5 | 5.3 | 42.8 | 2306.3 |
| | S2 | 35.8 | 81.8 | 13.6 | 10.0 | 4.8 | 34.3 | 1683.8 |

| Name | Season | DF (days) | DM (days) | Pods/plant (no.) | Pod length (cm) | Seed/ pod (no.) | 100-Seed weight (g) | Yield ($Kgha^{-1}$) |
|---|---|---|---|---|---|---|---|---|
| GLP2 | S1 | 37.3 | 78.0 | 7.0 | 12.0 | 4.3 | 60.3 | 2093.8 |
| | S2 | 37.3 | 80.3 | 7.4 | 13.1 | 4.8 | 46.5 | 751.3 |
| GLP24 | S1 | 38.8 | 82.3 | 16.0 | 10.0 | 5.5 | 24.0 | 2197.5 |
| | S2 | 38.0 | 85.3 | 22.3 | 9.7 | 5.8 | 33.5 | 787.5 |
| GLP1127a | S1 | 35.3 | 80.8 | 15.5 | 11.5 | 4.3 | 45.8 | 1631.3 |
| | S2 | 37.3 | 77.8 | 8.1 | 12.8 | 4.8 | 32.0 | 905.0 |
| **Overall mean** | | **37.8** | **82.0** | **12.6** | **9.9** | **4.6** | **36.9** | **1267.0** |
| LSD Accessions (A)** | | 1.9 | 3.4 | 5.1 | 1.4 | 0.7 | 4.4 | 322.5 |
| LSD Seasons (S)** | | 0.2 | 0.4 | 0.6 | 0.2** | 0.1 | 0.5 | 65.2 |
| LSD AxS** | | 3.0 | 6.3 | 9.3 | 2.4 | 1.3 | 7.2 | 658.2 |
| CV% | | 5.2 | 4.2 | 35.5 | 14.0 | 16.5 | 12.1 | 32.3 |

**=Significant at P≤0.05 probability levels respectively, LSD=Least significance difference, AxS=Interaction between accession and seasons, CV=Coefficient of variation, DF=Days to flowering, DM=Days to maturity, S1=Season 1, S2=Season 2.

The period between flowering and maturity ranged from 35.5 (GBK035394) to 53.6 (GBK035007) days with a mean of 44.1 days, the grain filling period varied from 41.9 (GLP 2) to 48.5 days (GLPx92) among the commercial varieties. Among the commercial varieties evaluated in this study, GLPx92 was the earliest to flower (35.1 days) and had the highest grain yield (1995 kg/ha) (Table 3.5). On the other hand, GLP2 was the earliest to mature (79.1 days) had the longest pods (12.6cm) and the highest 100-seed weight (59.8g), while GLP24 had the highest number of pods per plant (20.2) and the highest number of seeds per pod (5.6) (Table 3.6). However, 57 accessions flowered earlier than GLPx92, 69 accessions matured earlier GLP2 and 20 accessions outyielded GLP2 in grain yield. Nineteen accessions had shorter duration to maturity and higher yields than the earliest maturing commercial variety GLP2 (Table 3.7).

**Table 3.6: Mean values for days to flowering, days to maturity and yield-related for top 10, bottom 5 and checks of common bean accessions grown in Juja in 2018 and 2019 ranked based on maturity**

| Name | Season | DF (days) | DM (days) | Pods/plant (no.) | Pod length (cm) | Seed/ pod (no.) | 100-Seed weight (g) | Yield ($Kgha^{-1}$) |
|---|---|---|---|---|---|---|---|---|
| **Early maturing** | | | | | | | | |
| GBK035322 | S1 | 34.8 | 74.8 | 12.0 | 8.3 | 4.8 | 33.5 | 616.3 |
| | S2 | 35.3 | 73.8 | 12.2 | 10.4 | 4.6 | 38.0 | 757.5 |
| GBK035394 | S1 | 39.3 | 78.5 | 17.0 | 8.0 | 5.0 | 17.5 | 1665.0 |
| | S2 | 38.3 | 70.0 | 10.3 | 8.7 | 4.5 | 18.8 | 487.5 |
| GBK035284 | S1 | 33.5 | 75.3 | 10.5 | 9.3 | 5.5 | 33.0 | 1871.3 |
| | S2 | 34.5 | 73.8 | 14.4 | 10.6 | 4.5 | 38.0 | 1142.5 |
| GBK035338 | S1 | 37.0 | 77.5 | 15.0 | 9.3 | 5.0 | 20.0 | 1320.0 |
| | S2 | 36.0 | 72.8 | 21.4 | 9.4 | 5.1 | 20.0 | 1086.3 |
| GBK035378 | S1 | 36.3 | 74.8 | 13.0 | 8.5 | 5.3 | 27.8 | 1970.0 |
| | S2 | 36.8 | 75.8 | 15.1 | 7.2 | 3.5 | 34.0 | 711.3 |
| GBK035318 | S1 | 33.8 | 75.8 | 8.5 | 10.5 | 4.5 | 47.0 | 740.0 |
| | S2 | 33.8 | 75.0 | 12.1 | 10.4 | 3.8 | 42.5 | 338.8 |
| GBK034983 | S1 | 35.4 | 76.8 | 9.5 | 9.1 | 5.1 | 27.5 | 1840.0 |
| | S2 | 35.8 | 74.1 | 13.8 | 9.5 | 4.8 | 32.5 | 1388.1 |
| GBK035078 | S1 | 35.5 | 75.0 | 6.5 | 7.8 | 4.0 | 48.8 | 1026.3 |
| | S2 | 32.3 | 76.5 | 11.0 | 8.5 | 3.3 | 33.8 | 951.3 |
| NUA730 | S1 | 33.0 | 76.0 | 10.0 | 9.5 | 4.5 | 47.5 | 1485.0 |
| | S2 | 34.0 | 75.8 | 7.7 | 8.5 | 3.8 | 40.0 | 1000.0 |
| GBK035001 | S1 | 34.3 | 77.5 | 14.5 | 9.8 | 4.8 | 25.8 | 2256.3 |
| | S2 | 35.0 | 74.5 | 16.4 | 7.4 | 3.8 | 36.8 | 686.3 |
| **Late maturing** | | | | | | | | |
| GBK035023 | S1 | 45.3 | 90.5 | 10.0 | 12.5 | 4.5 | 54.0 | 576.3 |
| | S2 | 41.3 | 86.8 | 3.8 | 10.9 | 3.3 | 48.3 | 75.0 |
| GBK034966 | S1 | 40.3 | 88.8 | 21.5 | 9.5 | 6.8 | 21.3 | 1218.8 |
| | S2 | 40.8 | 89.8 | 17.7 | 9.3 | 4.9 | 30.5 | 868.8 |
| GBK035377 | S1 | 43.5 | 91.0 | 10.5 | 11.3 | 6.3 | 25.0 | 601.3 |
| | S2 | 38.5 | 89.3 | 12.3 | 10.1 | 4.5 | 36.8 | 522.5 |
| GBK034981 | S1 | 43.8 | 93.8 | 14.5 | 8.3 | 6.0 | 16.8 | 1701.3 |
| | S2 | 36.8 | 88.8 | 15.0 | 9.4 | 4.8 | 21.0 | 716.3 |
| GBK035007 | S1 | 41.0 | 99.0 | 12.0 | 11.0 | 5.0 | 46.3 | 1343.8 |
| | S2 | 34.5 | 83.8 | 5.6 | 11.0 | 3.4 | 39.9 | 547.5 |
| **Commercial varieties** | | | | | | | | |
| GLPx92 | S1 | 34.5 | 85.5 | 13.5 | 8.5 | 5.3 | 42.8 | 2306.3 |
| | S2 | 35.8 | 81.8 | 13.6 | 10.0 | 4.8 | 34.3 | 1683.8 |

| Name | Season | DF (days) | DM (days) | Pods/plant (no.) | Pod length (cm) | Seed/ pod (no.) | 100-Seed weight (g) | Yield ($Kgha^{-1}$) |
|---|---|---|---|---|---|---|---|---|
| GLP2 | S1 | 37.3 | 78.0 | 7.0 | 12.0 | 4.3 | 60.3 | 2093.8 |
| | S2 | 37.3 | 80.3 | 7.4 | 13.1 | 4.8 | 46.5 | 751.3 |
| GLP24 | S1 | 38.8 | 82.3 | 16.0 | 10.0 | 5.5 | 24.0 | 2197.5 |
| | S2 | 38.0 | 85.3 | 22.3 | 9.7 | 5.8 | 33.5 | 787.5 |
| GLP1127a | S1 | 35.3 | 80.8 | 15.5 | 11.5 | 4.3 | 45.8 | 1631.3 |
| | S2 | 37.3 | 77.8 | 8.1 | 12.8 | 4.8 | 32.0 | 905.0 |
| **Mean** | | **37.8** | **82.0** | **12.6** | **9.9** | **4.6** | **36.9** | **1267.0** |
| LSD Accessions (A)** | | 1.9 | 3.4 | 5.1 | 1.4 | 0.7 | 4.4 | 322.5 |
| LSD seasons (S)** | | 0.2 | 0.4 | 0.6 | 0.2 | 0.1 | 0.5 | 65.2 |
| LSD AxS** | | 3.0 | 6.3 | 9.3 | 2.4 | 1.3 | 7.2 | 658.2 |
| CV% | | 5.2 | 4.2 | 35.5 | 14.0 | 16.5 | 12.1 | 32.3 |

**=Significant at P≤0.01 probability levels respectively, LSD=Least significance difference, AxS=Interaction between accession and seasons, CV=Coefficient of variation, DF=Days to flowering, DM=Days to maturity, S1=Season 1, S2=Season 2.

### 3.4.4 Cluster and correlation reaults

Cluster analysis based on the agronomic traits grouped the 257 genotypes into two major groups. The largest group constituted 82.1 % of the genotypes, which had the highest pod length, seed weight and yield with a mean of 10.4 cm, 45.5 g and 1322.1 $kgha^{-1}$ respectively. The second group constituted 17.9 % of the genotypes and had the highest days to flowering, days to maturity, number of pods per plant and number of seeds per pod of 39.7 days, 84.1 days, 15.9 and 5.1 respectively (Figure 3.1).

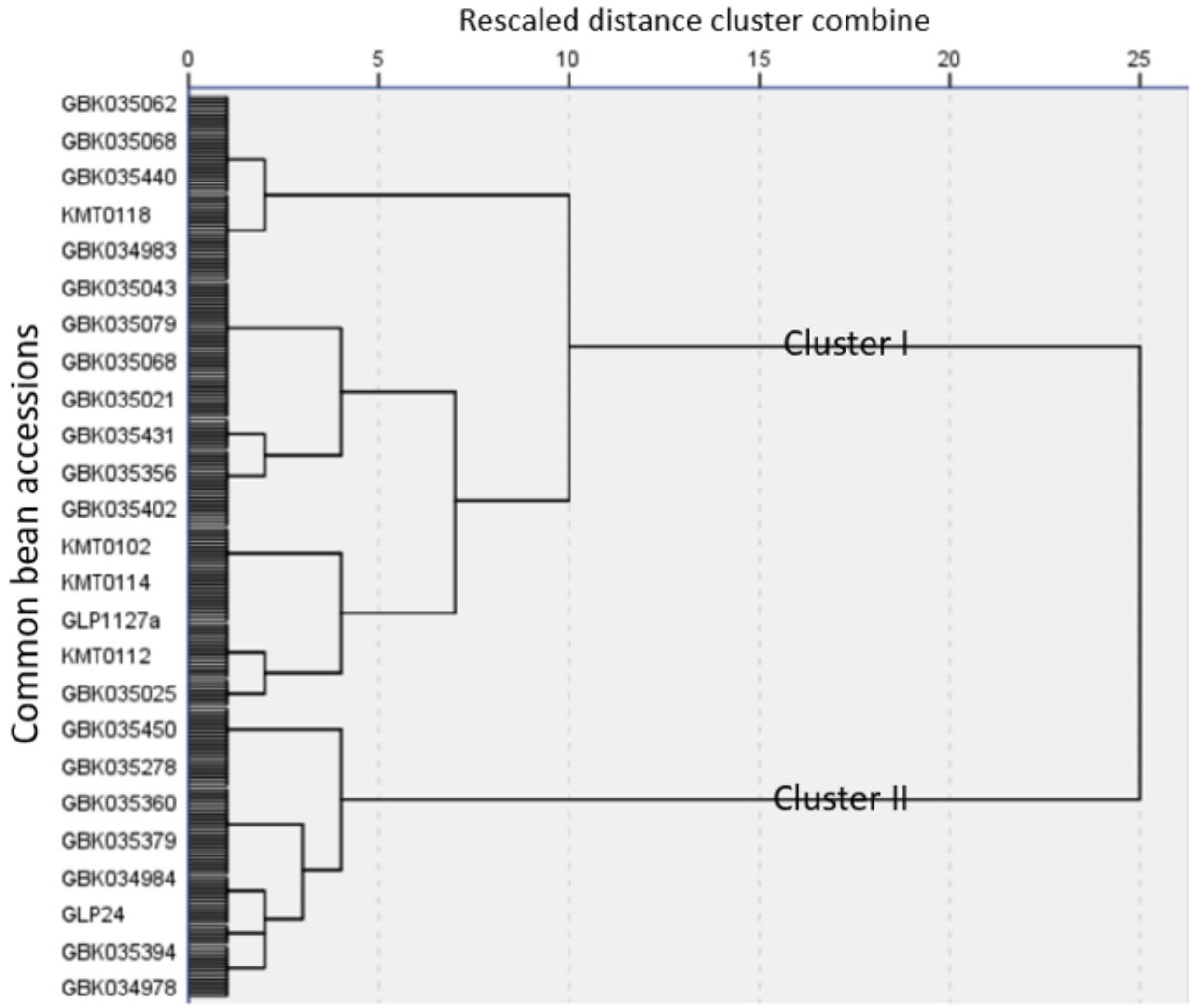


**Figure 3.3: Dendrogram showing relationship among 257 common bean genotypes evaluated in this study, only 25 accessions are labelled on the ruler**

The correlation coefficients among days to flowering, days to maturity, number of pods per plant, number of seeds per pod, 100-seed weight, and grain yield are presented in Table 3.7. A significant ($P \leq 0.05$) strong and positive correlation (0.73) was recorded between days to flowering and days to maturity. A moderate positive and significant association was revealed between pod length and 100-seed weight (0.51), between pods per plant and seed per pod (0.5) and between days to flowering and seeds per pod (0.43). A significant positive but weak relationship was revealed between days to flowering and

pods per plant (0.36), between days to maturity and number of seeds per pod (0.34), between pod length and grain yield (0.34), between 100 seed weight and grain yield (0.31) and between days to maturity and pods per plant (0.26). A significant extremely weak positive correlation of 0.17 was recorded between pods per plant and grain yield at P≤0.05 (Table 3.7).

A significant (P≤0.05) strong negative correlation was recorded between the number of pods per plant and 100-seed weight (-0.66) and between the number of seeds per pod and 100-seed weight (-0.65) among accessions. Furthermore, a significant moderate and negative relationship between days to maturity and 100-seed weight (-0.45) was recorded. Finally, a significant but weak negative association between pods per plant and pod length (-0.34) and between days to maturity and 100-seed weight (-0.28) were also recorded (Table 3.7).

**Table 3.7: Pearson correlation coefficient for days to flowering, days to maturity and yield-related traits for 257 common bean accessions grown at Juja in 2018 and 2019**

| | DF (days) | DM (days) | Pods/plant (no.) | Pod length (cm) | Seed/pod (no.) | 100-Seed weight (g) |
|---|---|---|---|---|---|---|
| DM | 0.73** | | | | | |
| Pods/plant | 0.36** | 0.26** | | | | |
| Pod length | -0.02ns | 0.05ns | -0.34** | | | |
| Seed/pod | 0.43** | 0.34** | 0.50** | 0.08ns | | |
| Seed weight | -0.45** | -0.28** | -0.66** | 0.51** | -0.65** | |
| Yield (kg/ha) | -0.04ns | 0.03ns | 0.17** | 0.34** | 0.04ns | 0.31** |

*, **=Significant at P≤0.05 and P≤0.01 probability levels respectively, DF=Days to flowering, DM=Days to maturity, ns=Not significant

### 3.4.5 Grain yields for common bean seed classes

The results revealed that the average grain yields varied with the seed size of the accessions. The large-seeded (>40g) accessions had the highest yield of 1406.5 kg/ha, followed by the medium-sized (25-40g) with an average of 1230.9 kg/ha, and lastly, small-seeded (<25g) had the lowest mean yields of 1039.3 kg/ha (Table 3.8).

### 3.4.6 Grain yields for large, medium, and small-seeded common bean lines

The results revealed that the average grain yields varied with the seed size of the accessions, the large-seeded (>40g) accessions had the highest yield of 1406.5 kg/ha, followed by the medium-sized (25-40g) with an average of 1230.9 kg/ha, lastly, Small-seeded (<25g) had the lowest mean yields of 1039.3 kg/ha (Table 3.8).

### 3.4.7 Grain yields for bush, semi climber, and climbing common bean lines

The result shows that the mean grain yield varied with the growth habit of the accessions. Accessions with climbing growth habit had the highest average grain yields of 1644.6 kg/ha, followed by semi climbers with an average yield of 1289.8 kg/ha and lastly, bush types had the lowest yields of 1102.3 kg/ha (Table 3.8).

**Table 3.8: Mean grain yield for common bean accessions grouped into their different growth habits, seed sizes and seed classes grown at Juja in 2018 and 2019**

| | | | Grain yield (kgha$^{-1}$) | | |
|---|---|---|---|---|---|
| **Trait** | **Description** | **Number of accessions** | **Long rain season** | **Short rain season** | **Mean** |
| **Growth** | | | | | |
| Type I&II | Bush bean and upright short | 124 | 1382.2 | 822.4 | 1102.3 |
| Type III | Vine type | 84 | 1565.5 | 1014.1 | 1289.8 |
| Type IV | Climbing type | 49 | 1835.6 | 1453.6 | 1644.6 |
| **Seed size** | | | | | |
| Large | >40g for 100 seeds | 127 | 1721.1 | 1091.8 | 1406.5 |
| Medium | 25-40g for 100 seeds | 62 | 1452.8 | 1009.0 | 1230.9 |
| Small | <25g for 100 seeds | 68 | 1237.9 | 840.6 | 1039.3 |
| **Market** | | | | | |
| Pintos | GLPx92 type | 22 | 1765.2 | 1306.3 | 1535.7 |
| Sugars | Cream, can be speckled | 39 | 1786.0 | 1192.6 | 1489.3 |
| Calima | Rosecoco type | 25 | 1986.5 | 991.9 | 1489.2 |
| Small reds | Red haricot type | 15 | 1585.9 | 1320.4 | 1453.1 |
| Large reds | Canadian wonder type | 17 | 1612.6 | 1133.5 | 1373.0 |
| Purples | Mwezimoja type | 11 | 1621.5 | 1122.1 | 1371.8 |
| Medium | Medium and large whites | 13 | 1536.3 | 864.0 | 1200.1 |
| Brown/tan | Brown and orange | 19 | 1235.1 | 981.5 | 1108.3 |
| Cariocas | Red and Red specks | 28 | 1379.8 | 781.4 | 1080.6 |
| Yellow | Yellow coloured | 8 | 1273.6 | 874.7 | 1074.2 |
| Blacks | Black coloured | 23 | 1303.8 | 727.0 | 1015.4 |
| Navy | Small whites | 37 | 1129.0 | 773.8 | 951.4 |

### 3.5 Discussion

In the current study, grain yield and number of pods per plant had the highest coefficient of variation, indicating a strong environmental influence among the genotypes evaluated for these traits. Grain yield is a polygenic trait conditioned by the number of pods, seeds per pod, and seed weight. On the other hand, pod yield is a product of successful pollination, fertilization and pod setting which are highly influenced by the environment. Comparable results were obtained for seed yield (50.7%) and the number of pods per plant (38.9) in a previous study conducted by Negahi *et al.,* (2014) using 284 genotypes.

The coefficient of variation is a scale that can be used to compare the extent of variation of different traits with different measurement units. According to Burton and de Vane (1953), phenotypic and genotypic coefficients of variation are categorized as 0-10% low, 10-20% high, and above 20% as high. The results show high genotypic coefficients of variations for 100-seed weight, the number of pods per plant, and grain yield, which indicates high genetic variation in these traits among the common bean accessions evaluated. Similar results have been reported for number pods per plant, 100-seed weight, and grain yield by Negahi *et al.,* (2014), who observed a high phenotypic coefficient of variation of 53.3%, 38.9%, and 50.7% for 100-seed weight, the number of pods per plant and seed yield, respectively. On the contrary, a low genotypic coefficient of variation for 100-seed weight, number of pods per plant, and seed yield of 4.6%, 4.67%, and 2.2%, respectively, have also been reported in a previous study that evaluated 52 common bean landraces. The low genotypic coefficient could be attributed to low genetic diversity among the accessions used in the study (Anunda *et al.,* 2019).

Heritability estimates indicate how much variation in a trait can be attributed to genetic variation and helps breeders to select based on the phenotypic performance of a trait. Based on the categorization of heritability by Johnson *et al.,* (1955), the traits days to flowering, 100-seed weight, and grain yield traits recorded a high broad-sense heritability (>60%). This indicates that the performance of these traits was majorly due to genetic differences and could be improved through selection based on the trait itself. Yohannes *et al.,* (2020) reported a high broad-sense heritability of days to maturity and 100-seed weight of 86.7% and 95.3%, respectively, and a moderate broad-sense heritability for days to flowering and number of seeds per pod of 40%, and 51.6% respectively.

The significant seasonal differences for various traits between long and short rains is due to differences in environmental conditions, especially temperatures and rainfall. Lower temperatures prevail during long rain season, causing a prolonged vegetative state that delay flowering and maturity. Heavy rainfall experienced during long rains adversely affects flower fertilization, resulting in reduced pod sets hence the lower number of pods in some cultivars during these seasons. On the other hand, short rain seasons tend to have

higher temperatures that lead to early termination of the vegetative state and initiation of the reproductive phase (Greenlife, 2023). Genotypes that flower and mature early tend to be more adapted to environment of growth than late maturing genotypes (Amanullah *et al.,* 2006). In this study nineteen accessions were found to combine early maturity and reasonable yield higher than that of earliest maturing commercial variety GLP2.

Variety GLP 2 was the earliest to mature among the commercial varieties. However, there were 19 bean accession that had higher seed yield and matured earlier than this variety. These bean accessions are ideal for cultivation in areas with a short rainy season. On the other hand, the variety GLPx92 was the earliest to flower, the latest to mature, and had the highest seed yield among the commercial varieties. Therefore, GLPx92 had a prolonged grain filling period that led to higher seed yield. Beebe *et al.,* (2013) found that that drought tolerant lines with improved yields also presented shorter period to maturity. Cluster analysis grouped the common bean genotypes into two major groups. The cluster with majority of genotypes (82.1%) contained the large-seeded (> 45 g 100-seed weight) genotypes of Andean gene pool which are reported to be adapted to higher altitude and cooler environments. The other groups consisted of small-seeded accessions with a mean 100-seed weight of 23.9 g of Mesoamerican gene pool which is adapted to lower altitudes and higher temperatures (Beebe *et al.,* 2011).

Grain yield is a polygenic trait that is conditioned by three yield components, the number of pods per plant, the number of seeds per pod, and seed weight (Kamfwa *et al.,* 2015). Consequently, the knowledge of the association between these seed yield attributes may help in selecting an excellent donor to improve this trait through indirect selection. A target trait can be improved through indirect selection via other traits. A strong positive relationship (0.73) between days to flowering and days to maturity indicates that days to flowering can be used to predict days to maturity for common bean accessions. Strongly associated traits are usually under the influence of the same gene or genes located close together on the chromosome and can both be selected simultaneously (Lobo 2008). Days to flowering have been reported to be under the control of dominance and additive gene effect with the dominance effect being lower, and when present it reduces the number of

days to flowering (Mendes *et al.,* 2008). A similar correlation result (0.7) between days to flowering and days to maturity was reported in a previous study conducted by Kamfwa *et al.,* (2015).

The weak and moderate positive correlation between days to flowering and number of pods per plant (0.36) and between days to flowering and number of seeds per pod (0.43) indicates that the number of pods per plant and number of seeds per pod is, to an extent, influenced by duration to flowering or time of flowering. For a crop that is largely self-pollinated like the common bean, pollination vectors may not be a limiting factor but the survival of pods and seeds after pollination may be affected by the competition of photosynthetic assimilates, soil nutrients, and water. These results agree with a previous study conducted by Marzooghian *et al.,* (2014), who reported a positive correlation between days to flowering with both the number of pods per plant (0.36) and the number of seeds per pod (0.29).

The significant positive but weak relationship between grain yield and number of pods per plant (0.17), pod length (0.34) and 100-seed weight (0.31) indicate that these traits influence grain yields and should be put into consideration during selection to improve grain yield. Strong positive correlations between seed yield per plant and the number of pods per plant (0.67) (Anunda *et al.,* 2019) have also been reported. A significant correlation of 0.51 between 100-seed weight and pod length suggests that large-seeded genotypes tend to have longer pods. Similar correlation result of 0.48 between 100-seed weight and pod length was reported by Okii *et al.,* (2014).

A significant strong negative correlation between 100-seed weight and the number of pods per plant and between 100-seed weight and number of seeds per pod indicate compensation among yield components (Negahi *et al.,* 2014). This negative association between yield components means that selecting for a greater number of pods per plant would lead to small-seeded plants, and selection for large-seeded genotypes would lead to plants with low seed locules per pod. It is critical to understand the nature of this negative relationship if it is independent of competition or due to competition for a limited resource. Similar negative correlation results between 100-seed weight with both number

of pods per plant and number of seeds per pod were reported by Negahi *et al.,* (2014) and Kamfwa *et al.,* (2015). However, Kamfwa *et al.,* (2015) reported weak negative correlation between seed weight and the number of pods per plant (-0.17) and between the number of seeds per plant (-0.38) and days to maturity (-0.27).

The results show that pinto, sugars, calima, small reds, large reds, and purples seed classes were more adapted to the environment in which the experiment was conducted compared to medium white, brown and tan, cariocas, yellows, blacks, and navy seed classes. It has been reported that consumer preference for common beans depends on seed size and color among other characteristics. In eastern Africa, the calima seed type (red speckled) is highly popular and accounts for about 22% of common bean production. Medium and small reds follow in consumer preference accounting for approximately 20% of the production. Large reds including red kidney rank third in popularity accounting for about 10% of common beans produced, navy, whites, purples, and black follow in popularity, respectively (Wortmann *et al.,* 1998; Farrow and Andriatsitohaina, 2021). The results show that the popular seed classes in the region were the highest yielding in this study, which could have resulted from continuous selection by local common bean farmers that improved their adaptability.

Common bean varieties vary in seed size, those that weigh less than 25 g per 100 seeds are classified as small-seeded while those that range from 25 to 40 g are classified as medium-sized, and those that weigh more than 40g are classified as large-seeded (Angioi *et al.,* 2010; Lei *et al.,* 2020). The result in this study indicates that large-seeded accessions are more adapted in this region, unlike the medium and small-seeded genotypes. Based on seed size, the common bean has been categorized into two distinct centers of origin, namely Mesoamerican and Andean gene pools (Blair *et al.,* 2007; Burle *et al.,* 2010). Andean gene pool is generally large-seeded and adapted to relatively higher altitudes and lower temperatures. In contrast, the Mesoamerican gene pool is small-seeded and adapted to lower altitudes and higher temperatures (Beebe *et al.,* 2011).

The high yielding potential of climbing common bean was revealed in this study. However, climbing genotypes are labor-intensive as they require staking and may not be

ideal for mixed cropping. The results agree with Okii *et al.,* (2014), and Farrow and Andriatsitohaina, (2021), who reported that common beans with climbing growth habits are higher-yielding and hence ideal for small-scale farmers with a limited size of land in highland areas. Oppositely, the bush types are preferred because they do not require support and are early maturing, hence convenient for commercial production (Okii *et al.*, 2014).

### 3.6 Conclusion

The study evaluated a panel of locally conserved common bean accessions for variation in yield-related traits over four seasons. The results showed significant differences among the common bean accession for all the evaluated traits. Significant higher yields were recorded during the long rain seasons. High phenotypic and genotypic variation for 100-seed weight, the number of pods per plant, and grain yield was observed. The traits days to flowering, 100-seed weight, and grain yield showed high broad-sense heritability. Grain yield had weak positive correlations with the number of pods per plant, pod length, and 100-seed weight. Large-seeded, climbing, and popular (pinto, calima, small reds, and purples) bean accessions had higher yields. The study identified nineteen common bean accessions that were significantly ($P \leq 0.05$) early maturing and had higher yields than the commercial varieties. Majority of common bean accessions evaluated in this study were of Andean origin and showed heritable variation that could be exploited in breeding programs.

# CHAPTER FOUR

# GENOME WIDE ASSOCIATION STUDY OF VARIATION IN COOKING TIME AMONG COMMON BEAN (*PHASEOLUS VULGARIS* L.) ACCESSIONS USING DIVERSITY ARRAYS TECHNOLOGY MARKERS

## 4.1 Abstract

Stored grains of common bean (*Phaseolus vulgaris* L.) develop the hard to cook trait (HTC) which is manifested in a prolonged cooking time thereby imposing time and energy constraints. The objective of this study was to determine variation in cooking time among common bean genotypes and to identify Single Nucleotide Polymorphism (SNP) markers associated with cooking time. Seeds of 222 common bean accessions sourced from Kenyan institutions were multiplied in JKUAT farm in 2019. Freshly harvested seeds and those stored at 35°C and 50% RH for four months for accelerated aging were soaked in distilled water for 16 hours and evaluated for cooking time using finger pressing method. The accessions were also genotyped to determine variation in single nucleotide polymorphism (SNP) markers using Diversity Arrays Technology Sequencing (DArTseq). Genome-wide association study (GWAS) analysis was conducted to identify SNPs significantly associated with cooking time. The study revealed that significantly differences ($P \leq 0.05$) within and between fresh and aged bean accessions. Storage significantly ($P \leq 0.05$) influenced the cooking time of common bean accessions. Fresh seeds had a lower cooking time with a mean of 40.8 minutes and ranged from 28.1 to 72.2 minutes while aged seeds had a higher average cooking time of 54.1 minutes and ranged from 32.1 to 96.3 minutes. Among the aged seeds genotype GBK034996 that took the shortest time to cook (32.1 minutes) compared to 48.6 minutes of the easy to cook commercial variety GLP2. The aging process affected the cooking time of bean accessions differently with the least affected being NUA Ciankui (0.3% increase). GWAS identified a region on chromosome 10 to be significantly ($P \leq 0.05$) associated with the cooking time of aged seeds. Consequently, two potential candidate genes *Phvul.010G038000* (galacturan 1,4-alpha galacturonidase) and *Phvul.010G038100* (polygalacturonase) were

revealed. The characterized common bean accessions and the identified SNP markers can be utilized in breeding programs to improve the cooking quality of the common bean.

### 4.2 Introduction

Common bean plays a critical role in the nutrition security of a large segment of the world population, especially in third world countries. Its dry grains are used as a major source of dietary protein. Cooking is a fundamental part of bean preparation, and it inactivates anti-nutritive factors, increases digestibility, and improves the sensorial quality of beans (Costa *et al.,* 2006). Some samples of dry beans have been found to require a long cooking time which is time and energy consuming (Kinyanjui *et al.,* 2015). The cooking time of beans has been reported to be influenced by a diversity of factors including, genetic differences, growth environment, post-harvest handling, storage conditions such as temperature and humidity, storage time, and treatments before cooking (Arruda *et al.,* 2012).

Farmers and traders commonly store the grains for long periods before availing them to end-users. When storage occurs in adverse conditions of high temperature and high relative humidity, the grains develop the hard-to-cook (HTC) phenomenon, which increases the cooking time of beans. In addition, the improperly stored grains become discolored and decrease water absorption capacity (Ousman *et al.,* 2013). Further, Vindiola (1986) reported that whereas all bean samples were affected by the HTC phenomenon, the rate of hardening differed among varieties.

The formation of insoluble pectates at the cell wall and middle lamella that renders the tissue more refractive to cell separation during cooking is believed to be the cause of HTC (Shomer *et al.,* 1990; Hentges *et al.,* 1990). Pectin is made up of complex acid polysaccharides with a backbone of galacturonic acid residue with an alpha 1,4 glycosidic linkage. Homogalacturonan-rich pectin is commonly found in the middle lamella region of plant cell walls where two cells border (Atkinson *et al.,* 2002). The increase of phenolic compounds that cause lignification of cells has also been proposed to cause HTC (Elisabeth *et al.,* 1998; Garcia *et al.,* 1998). A dual enzyme mechanism has been proposed to explain the development of HTC conditions in storage at elevated temperatures and

relative humidity (Jones and Boulter, 1983a). This theory suggests that at high temperatures and relative humidity, the pectin methylesterase (PME) enzyme hydrolyses pectin molecules in the middle lamella, forming pectic acid and methanol. Concurrently, the phytase enzyme hydrolyses phytic acid in cotyledons cells to release inorganic phosphate and magnesium ions. The magnesium and calcium ions released in the cells migrate to the middle lamella and produce an insoluble magnesium pectinate and calcium pectinate that cements cells together hardening the cell wall (Jones and Boulter, 1983a). This hypothesis was supported by the presence of more water-soluble pectin (8.44 mg/g) in varieties with shorter cooking time than slow cooking beans (5.51 mg/g) (Njoroge *et al.,* 2014).

The use of deoxyribonucleic acid (DNA) analysis techniques in common bean breeding programs has improved the understanding of genetic factors controlling various traits. Genome-wide association studies (GWAS) is a popular method used to identify quantitative trait loci (QTLs) associated with bean traits. GWAS studies are mainly concerned with determining alleles associated with various single nucleotide polymorphisms (SNPs) and making statistical comparisons to identify SNPs associated with a particular trait (Resende *et al.,* 2018). SNPs markers are more practical in genotyping and hence preferred in the construction dense linkage map (Gujaria-Vema *et al.,* 2016). In a previous GWAS study, significant SNPs associated with cooking time were identified on chromosomes 2, 3, and 6 using 206 common bean accessions of Andean origin; the significant SNPs identified explained between 4 to 8.7% of the phenotypic variation (Cichy *et al.,* 2015). In a recent study conducted by Berry *et al.,* (2020) using 146 recombinant inbred lines of common bean, 10 QTLs on chromosomes 1, 2, 3, 5, 6, 10, and 11 were identified, with the most robust QTLs being on chromosome 3, 6, 10 and 11 that appeared in over two different environments. The identified regions require further exploration to determine their robustness and stability across different genetic backgrounds, growth environments, and storage conditions. This study aimed at identifying common bean accessions that are easier to cook and investigate Single Nucleotide Polymorphism (SNP) markers associated with cooking time.

## 4.3 Materials and Methods

### 4.3.1 Plant materials and field multiplication

A total of 222 of the 257 common bean accessions were successfully phenotyped and genotyped in this study. These included 169 accessions from the National Gene Bank of Kenya based at Kenya Agricultural and Livestock Research Organisation (KALRO) - Muguga, 38 accesions from KALRO-Embu, 11 landraces collected from selected farmers' fields, and four commercial varieties (GLP-2, GLP-24, GLPx92 and GLP1192a). The accessions belonged to different market/seed classes including small whites, blacks, yellows, cariocas, brown and tan, medium whites, purples, larges reds, small reds, calima, sugars, and pintos. A single seed was randomly picked from each accession and grown in the screenhouse. The harvested seeds were then multiplied during the short rain season of the year 2019 (September 2019 – January 2020) at the Jomo Kenyatta University of Agriculture and Technology (JKUAT) trial farm in Kenya. The site used for the multiplication of the plant materials is describe in section 3.3.1 in details.

### 4.3.2 Incubation of seed and determination of cooking time

Cooking time was determined on each accession using freshly harvested seeds and aged seeds. The aging process involved storing the seeds in a thermostatically controlled incubator (HP300 G, China) at a temperature of 35°C and 50% relative humidity for a period of four months (Figure 4.1). Freshly harvested and seeds removed from the incubator after ageing treatment were stored at -20°C to prevent further aging during the cooking experiment which took a period of two months. A sample of 100 whole grain seeds of each accession was rinsed and soaked in distilled water for 16 hours. Hard-shelled seeds that failed to absorb water were sorted out and excluded from the cooking experiment to avoid the effect of impermeable seed coat on the cooking processes. Soaked seeds were then subjected to standard cooking using distilled water at 96$^{o}$C in a thermostatically controlled water bath (WBU-45; Memmert, Schwabach, Germany). Ten (10) seeds were sampled from the cooking water bath after 20 minutes and at 5 minutes intervals thereafter without interrupting the boiling. The 10 cooked bean seeds were

cooled in cold water for a minute and their softness/hardness determined using a subjective finger pressing method (Vindiola *et al.*, 1986; Kinyanjui *et al.*, 2015). The bean seeds were considered cooked when the cotyledon disintegrated on pressing and felt soft (lack of graininess). Cooking of the beans was done in duplicate by two people and the percentage of cooked beans in a batch was expressed as a function of time.


**Figure 4.1: Incubators used for storage of seed in a controlled temperature and relative humidity chamber to artificially age the common bean accessions**

### 4.3.3 DNA extraction and genotyping

DNA isolation and genotyping were conducted at SEQART Africa laboratories housed in the Biosciences eastern and central Africa (BecA) hub, in International Livestock Research Institute (ILRI) Campus, Nairobi. Briefly, the procedure involved germinating the seeds and growing the young seedlings in vermiculite for a period of 10 days. The Nucleomag® Plant DNA extraction kit (Macherey-Nagel AG, Switzerland) was used for DNA isolation from young leaves tissue of each common bean accession. The quality and quantity of DNA was visualized using gel electrophoresis on a 0.8% agarose. Genotyping by Sequencing (GBS) using DArTseqTM technology was used to identify variability in

single nucleotide polymorphism (SNP) markers. DNA libraries were constructed according to Diversity Arrays Technology Sequencing (DArTseq) complexity reduction method through the digestion of genomic DNA using a combination of two restriction enzymes (PstI and MseI) and ligation of barcoded adapters followed by PCR amplification of adapter-ligated fragments. Libraries were sequenced using single-read sequencing runs for 77 cycles (Kilian *et al.*, 2012). Illumina Hiseq2500 platform was used for high-throughput sequencing and scoring of markers was achieved using DArTsoft14 Software Version 1.5.2. beta (Diversity Arrays Technology 2017) as SilicoDArT markers and SNP markers. Markers were scored as binary for presence /absence (1 and 0, respectively) of the restriction fragment with the marker sequence in the genomic representation of the sample. Both SilicoDArT markers and SNP markers were aligned to the reference genomes of common bean (*Phaseolus vulgaris* 442 version 2.1) to identify chromosome positions (Goodstein *et al.*, 2012).

### 4.3.4 Data analysis

Logistic regression modeling was used to describe the relations between cooking time and the percentage of cooked bean seeds as described by Wafula *et al.*, (2020). Cooking time was defined as the time corresponding to the probability that 95% of the bean seeds would be cooked. The intercept and regression coefficients were used to generate graphs of cooking time for different common bean accessions. Analysis of variance was performed on the obtained cooking time using R software (version 4.0.2) and the mean values of different accessions were compared using the least square difference (LSD) at 0.05% significance level. Complete clustering analysis was conducted using NbClust package of the R software considering Euclidean distances to determine the genetic relationship among genotypes.

### 4.3.5 Linkage disequilibrium

A total of 19188 SNPs markers with minor allele frequency (MAF) > 0.05 and integrity > 0.8 were selected and used for subsequent analysis. Principal component (PC) (Q matrix) and the relative kinship (K matrix) analyses were conducted using VanRaden method

within R based GAPIT package version 0.3.4 to account for the population structure. The first three principal components were used to construct the PC matrix. Linkage disequilibrium was estimated between SNPs on each chromosome by calculating the square value of correlation ($r^2$) between the pair of markers using KDCompute software 0.6.1.

### 4.3.6 Marker-trait association

Based on the 19188 SNPs markers and cooking time of 194 fresh and 222 aged common bean accession, association analysis was performed using genome-wide association mapping (GWAS) to identify SNPs associated with cooking time. Marker-trait association analysis was conducted using the Generalized Linear Model (GLM), and Compressed Mixed Linear Models (CMLM) of the default settings of Genomic Association and Prediction Integrated Tool (GAPIT) software version 0.3.4. via the KDCompute interface (https://kdcompute.seqart.net/kdcompute/login). The GWAS threshold for the significant marker-trait association was the F-test for testing the null hypothesis that there is no association between the SNP and trait.

### 4.3.7 Identification of potential candidate gene

Potential candidate genes that were flagged by the SNPs significantly associated with cooking time, were identified using the jBrowse search tool against the reference genome of common bean (*Phaseolus vulgaris* 442 version 2.1) which is available on the Phytozome database (www.phytozome.net) (Goodstein *et al.,* 2012). The maximum threshold for identification of the potential candidate genes was set at 100kb around the position of the trait-associated SNPs.

Additionally, the significant SNP sequences were utilized in a basic local alignment search tool for nucleotides (BLASTn) within the National Center for Biotechnology Information (NCBI) database to find potential matches (www.ncbi.nlm.nih.gov).

## 4.4 Results

### 4.4.1 Phenotypic evaluation

The results show that cooking time for fresh and aged seeds significantly ($P \leq 0.05$) varied among common bean accessions (Appendix 2). The cooking time of fresh seeds ranged from 28.1 to 72.2 minutes with a mean of 40.8±0.4 minutes, while that of aged seeds ranged from 32.1 to 96.3 minutes with a mean of 54.9±0.7 minutes (Table 4.1). Storage of the seeds increased cooking time of the accessions, but the increase varied among common bean accessions. Accessions NUA Ciankui had the least increase in cooking time due to storage of 0.3% in comparison to commercial variety GLP2 (Rosecoco) with the least increase in cooking time of 0.8 % among the commercial varieties (Table 4.1).

Commercial varieties evaluated in this study were categorized into two groups based on the least significant differences in cooking time, GLP2 and GLP24 had a shorter cooking time while GLPx92 and GLP1127a had a longer cooking duration. GLP2 had the shortest cooking duration of 47.2 minutes while GLP1127a had a longer cooking time of 61.1 minutes for aged seeds. Among the bean accessions evaluated, GBK034996 had the shortest cooking time (32.1 minutes) while GBK035370 had the highest cooking time of 96.3 minutes for aged seeds (Table 4.1). A total of 133 and 65 bean accessions had a shorter cooking time for fresh and aged seeds respectively than commercial variety GLP24 (Table 4.1).

**Table 4.1: Mean values of cooking time of common bean accessions for fresh and aged seeds ordered based on cooking time of aged seeds**

| | Cooking time | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Name | Fresh (min) | Aged (min) | Mean (min) | Increase (%) | Growth habit | Market class | DM | 100-seed weight (g) | Yield $Kgha^{-1}$ |
| GBK034996 | 29.8 | 32.1 | 31.0 | 7.8 | C | BT | 84.0 | 28.0 | 1436.9 |
| GBK035012 | 29.8 | 35.5 | 32.6 | 19.0 | SC | Pinto | 83.3 | 42.4 | 1097.5 |
| GBK035279 | 34.7 | 35.8 | 35.2 | 3.0 | SC | Black | 87.3 | 21.6 | 1497.5 |
| GBK035024 | 33.7 | 36.9 | 35.3 | 9.4 | SC | Sugar | 87.1 | 47.9 | 903.1 |
| GBK035353 | 31.8 | 37.0 | 34.4 | 16.6 | B | Navy | 86.5 | 21.5 | 591.9 |
| GBK035341 | 28.1 | 37.1 | 32.6 | 32.0 | C | BT | 83.9 | 28.6 | 850.6 |
| NUA611 | 31.4 | 37.5 | 34.4 | 19.4 | SC | Calima | 87.4 | 49.8 | 1096.3 |
| KMT0103 | 36.0 | 37.7 | 36.9 | 4.7 | B | Calima | 78.0 | 50.3 | 1355.6 |
| GBK035052 | 28.3 | 38.2 | 33.3 | 35.2 | SC | Carioca | 81.3 | 38.5 | 1540.6 |
| KNB0104 | 34.8 | 38.6 | 36.7 | 11.0 | SC | Small red | 85.6 | 28.6 | 1006.9 |
| GBK034975 | 34.1 | 38.9 | 36.5 | 13.8 | C | Black | 84.6 | 21.4 | 581.3 |
| GBK035419 | 33.0 | 39.3 | 36.1 | 18.9 | B | Black | 86.5 | 18.9 | 690.6 |
| GBK035334 | 30.2 | 39.5 | 34.8 | 30.6 | SC | Navy | 87.8 | 20.6 | 801.3 |
| KNB0118 | 39.5 | 40.1 | 39.8 | 1.6 | SC | BT | 81.8 | 20.5 | 1413.5 |
| NUA640 | 36.0 | 40.2 | 38.1 | 12.0 | B | Calima | 81.8 | 56.8 | 2232.5 |
| NUA631 | 36.6 | 40.3 | 38.4 | 10.0 | B | Calima | 79.6 | 56.0 | 2140.0 |
| NUA Ciankui | 40.4 | 40.5 | 40.5 | 0.3 | B | Calima | 83.6 | 49.4 | 939.4 |
| GBK035060 | 31.5 | 40.7 | 36.1 | 29.4 | SC | Carioca | 86.0 | 40.0 | 578.1 |
| GBK035397 | 33.5 | 41.4 | 37.5 | 23.7 | B | BT | 78.3 | 41.5 | 841.9 |
| GBK035276 | 31.6 | 41.7 | 36.7 | 31.6 | C | Purple | 85.4 | 35.0 | 1766.9 |
| KNB0119 | 36.4 | 41.9 | 39.1 | 15.1 | B | Carioca | 81.2 | 43.3 | 1069.2 |
| GBK035005 | 36.6 | 42.1 | 39.4 | 14.9 | SC | Large red | 86.6 | 22.0 | 1395.6 |
| GBK035079 | 38.1 | 42.2 | 40.2 | 10.7 | B | Carioca | 81.3 | 46.9 | 868.1 |
| GBK035432 | 32.2 | 42.4 | 37.3 | 31.6 | B | Purple | 79.0 | 45.3 | 1570.6 |
| GBK035036 | 31.6 | 42.9 | 37.2 | 35.9 | SC | Small red | 83.9 | 22.6 | 1559.4 |
| KNB0101 | 37.0 | 42.9 | 39.9 | 16.1 | C | Small red | 82.5 | 24.5 | 2082.5 |
| GBK035055 | 39.4 | 43.1 | 41.3 | 9.4 | B | Carioca | 76.9 | 47.8 | 1109.4 |
| GBK035315 | 37.4 | 43.2 | 40.3 | 15.4 | B | Black | 83.4 | 16.6 | 617.5 |
| GBK035019 | 37.4 | 43.2 | 40.3 | 15.5 | B | Calima | 84.0 | 37.8 | 1231.3 |
| GBK035319 | 36.9 | 43.7 | 40.3 | 18.4 | B | Pinto | 78.3 | 40.1 | 830.0 |
| GBK035444b | 33.3 | 43.9 | 38.6 | 31.8 | C | Sugar | 84.3 | 46.5 | 788.5 |
| GBK035062 | 39.3 | 44.0 | 41.6 | 11.9 | B | Carioca | 77.9 | 47.8 | 1027.5 |
| GBK035026 | 28.3 | 44.0 | 36.1 | 55.5 | SC | Pinto | 76.6 | 35.3 | 1280.0 |
| GBK035374b | 35.5 | 44.1 | 39.8 | 24.1 | SC | BT | 81.0 | 33.8 | 948.0 |
| GBK035400 | 39.7 | 44.3 | 42.0 | 11.5 | C | Navy | 86.3 | 16.1 | 1463.8 |
| GBK035448 | 41.4 | 44.3 | 42.9 | 7.0 | C | Small red | 85.5 | 24.4 | 1715.0 |

| Name | Cooking time | | | | Growth habit | Market class | DM | 100-seed weight (g) | Yield Kgha$^{-1}$ |
|---|---|---|---|---|---|---|---|---|---|
| | Fresh (min) | Aged (min) | Mean (min) | Increase (%) | | | | | |
| GBK035074 | 41.2 | 44.4 | 42.8 | 7.7 | B | Carioca | 80.1 | 46.5 | 1375.0 |
| GBK035035 | 35.0 | 44.5 | 39.7 | 27.0 | B | Small red | 81.5 | 48.4 | 933.1 |
| GBK034966 | 31.5 | 44.5 | 38.0 | 41.3 | SC | Black | 89.3 | 21.4 | 1043.8 |
| GBK035068 | 36.6 | 44.5 | 40.6 | 21.8 | B | Carioca | 78.5 | 47.8 | 803.8 |
| GBK035409b | 33.2 | 44.6 | 38.9 | 34.4 | C | Yellow | 81.3 | 34.0 | 861.9 |
| KNB0111 | 41.9 | 44.7 | 43.3 | 6.5 | B | Large red | 81.2 | 56.7 | 1479.0 |
| GBK035046a | 34.4 | 44.7 | 39.6 | 29.9 | B | Calima | 76.0 | 54.3 | 123.0 |
| GBK047121 | 35.3 | 44.8 | 40.0 | 27.1 | B | Large red | 80.9 | 44.5 | 1108.1 |
| NUA700 | 41.6 | 44.9 | 43.3 | 7.9 | C | Calima | 86.6 | 43.6 | 2668.8 |
| GBK035065 | 41.6 | 45.1 | 43.4 | 8.5 | B | Carioca | 87.5 | 40.8 | 834.4 |
| GBK035330 | 43.0 | 45.4 | 44.2 | 5.6 | B | Black | 84.5 | 28.3 | 1224.4 |
| GBK035072 | 30.0 | 45.5 | 37.8 | 51.4 | C | Carioca | 78.5 | 44.5 | 1350.6 |
| GBK035047 | 36.2 | 46.0 | 41.1 | 27.2 | B | Carioca | 77.5 | 48.7 | 1163.1 |
| GBK035449 | 39.1 | 46.1 | 42.6 | 17.8 | SC | Black | 86.8 | 20.9 | 1346.9 |
| GBK035381 | 34.4 | 46.4 | 40.4 | 35.0 | C | Pinto | 89.9 | 26.9 | 1393.1 |
| GBK035022 | 38.0 | 46.5 | 42.2 | 22.4 | B | Purple | 76.4 | 46.0 | 1113.1 |
| NUA596b | 39.3 | 46.5 | 42.9 | 18.5 | B | Large red | 81.5 | 43.2 | 708.5 |
| GBK034978 | 33.9 | 46.6 | 40.3 | 37.4 | C | Large red | 80.9 | 26.0 | 1320.6 |
| GBK035085 | 37.7 | 46.6 | 42.1 | 23.8 | B | Sugar | 80.8 | 69.1 | 1875.0 |
| NUA637 | 37.2 | 46.7 | 41.9 | 25.7 | B | Large red | 82.0 | 52.4 | 2361.9 |
| GBK035009 | 31.5 | 46.8 | 39.2 | 48.9 | B | Pinto | 79.0 | 20.6 | 1376.9 |
| GBK035406 | 42.1 | 46.9 | 44.5 | 11.4 | SC | Small red | 85.9 | 25.4 | 1316.3 |
| GBK035033 | 42.6 | 47.1 | 44.9 | 10.4 | SC | Purple | 77.5 | 43.5 | 1328.8 |
| GBK035456 | 37.9 | 47.6 | 42.7 | 25.6 | SC | Large red | 84.0 | 47.0 | 1846.3 |
| GBK035280b | 41.8 | 47.8 | 44.8 | 14.1 | C | BT | 85.7 | 24.2 | 1016.0 |
| GBK035020 | 31.5 | 47.9 | 39.7 | 51.9 | C | Small red | 82.8 | 22.6 | 994.4 |
| KMT0105 | 39.2 | 48.0 | 43.6 | 22.3 | B | Calima | 78.1 | 46.4 | 993.8 |
| GBK035025 | 39.6 | 48.2 | 43.9 | 21.5 | SC | Sugar | 80.6 | 43.6 | 2398.1 |
| GBK035367b | 38.2 | 48.4 | 43.3 | 26.7 | B | Black | 84.5 | 25.5 | 389.0 |
| GBK035042 | 39.8 | 48.5 | 44.1 | 21.6 | SC | Pinto | 79.3 | 36.6 | 976.3 |
| KMT0110 | 38.6 | 48.5 | 43.6 | 25.4 | SC | Sugar | 78.1 | 55.3 | 1631.3 |
| GBK035450 | 32.7 | 48.7 | 40.7 | 49.1 | SC | Small red | 85.0 | 23.8 | 1539.4 |
| GBK034977b | 41.2 | 48.7 | 45.0 | 18.3 | B | Carioca | 76.2 | 56.5 | 954.0 |
| GBK034968 | 43.7 | 48.9 | 46.3 | 12.0 | B | BT | 81.0 | 34.9 | 615.0 |
| GBK035425 | 45.6 | 49.0 | 47.3 | 7.3 | B | Purple | 77.1 | 52.1 | 1413.8 |
| GBK035437 | 39.3 | 49.0 | 44.1 | 24.9 | SC | Sugar | 79.6 | 45.9 | 1141.3 |
| GBK035377 | 40.8 | 49.0 | 44.9 | 20.1 | SC | BT | 90.1 | 25.5 | 561.9 |
| GBK035285 | 33.6 | 49.4 | 41.5 | 46.9 | C | Small red | 81.8 | 32.3 | 1795.6 |
| GBK035030 | 44.9 | 49.5 | 47.2 | 10.2 | B | BT | 79.5 | 32.1 | 862.5 |

| Name | Cooking time | | | | Growth habit | Market class | DM | 100-seed weight (g) | Yield Kgha$^{-1}$ |
|---|---|---|---|---|---|---|---|---|---|
| | Fresh (min) | Aged (min) | Mean (min) | Increase (%) | | | | | |
| GBK035442 | 36.8 | 49.7 | 43.2 | 35.0 | C | Sugar | 84.5 | 50.5 | 2048.1 |
| GBK035025b | 38.3 | 49.7 | 44.0 | 29.7 | B | Sugar | 82.5 | 51.2 | 620.0 |
| GBK035076 | 37.4 | 49.8 | 43.6 | 33.1 | B | Carioca | 77.3 | 51.1 | 1073.8 |
| GBK035438 | 42.6 | 49.9 | 46.3 | 17.1 | B | Large red | 86.4 | 34.8 | 665.6 |
| GBK035284b | 39.4 | 50.0 | 44.7 | 26.9 | C | BT | 79.3 | 34.7 | 1143.5 |
| NUA739 | 41.3 | 50.0 | 45.7 | 21.1 | SC | Calima | 81.0 | 46.9 | 1488.1 |
| GBK035073 | 40.4 | 50.2 | 45.3 | 24.5 | B | Carioca | 79.1 | 47.1 | 1437.5 |
| GBK035446 | 41.9 | 50.4 | 46.2 | 20.2 | SC | Sugar | 86.8 | 43.3 | 1399.4 |
| KNB0112 | 46.1 | 50.5 | 48.3 | 9.7 | B | Large red | 85.7 | 37.5 | 712.5 |
| GBK035374 | 36.6 | 50.6 | 43.6 | 38.4 | C | Small red | 86.4 | 23.4 | 1558.1 |
| GBK035431 | 39.2 | 50.6 | 44.9 | 29.3 | B | Navy | 84.6 | 21.6 | 370.6 |
| GBK035420b | 42.7 | 50.7 | 46.7 | 18.8 | B | Carioca | 76.5 | 52.0 | 867.5 |
| GBK035409b | 49.7 | 50.9 | 50.3 | 2.4 | B | Yellow | 81.3 | 34.0 | 862.0 |
| NUA666 | 41.2 | 50.9 | 46.1 | 23.3 | B | Sugar | 83.5 | 51.5 | 1917.5 |
| GBK035439 | 48.2 | 51.0 | 49.6 | 5.7 | B | Sugar | 83.0 | 63.9 | 1630.6 |
| GBK035324 | 31.8 | 51.0 | 41.4 | 60.6 | SC | Purple | 83.6 | 24.4 | 766.9 |
| GBK035392 | 38.3 | 51.1 | 44.7 | 33.4 | C | Small red | 85.8 | 23.4 | 1585.0 |
| GBK035011 | 43.2 | 51.5 | 47.4 | 19.1 | B | Calima | 81.1 | 48.6 | 906.9 |
| GBK035021 | 36.5 | 51.6 | 44.1 | 41.4 | SC | Sugar | 80.3 | 52.3 | 1398.1 |
| GBK035444 | 45.6 | 51.6 | 48.6 | 13.1 | SC | Sugar | 86.4 | 42.1 | 1260.0 |
| GBK034099 | 38.0 | 51.7 | 44.8 | 35.8 | SC | BT | 81.4 | 28.4 | 1520.6 |
| GBK035376 | 38.3 | 51.9 | 45.1 | 35.5 | B | Yellow | 77.3 | 47.6 | 1702.5 |
| KMT0104 | 40.0 | 51.9 | 45.9 | 29.9 | SC | Calima | 80.4 | 50.5 | 1761.9 |
| GBK035037 | 30.7 | 52.1 | 41.4 | 69.6 | C | Small red | 87.9 | 23.5 | 1334.4 |
| GBK035310 | 49.2 | 52.2 | 50.7 | 6.1 | SC | Pinto | 84.4 | 44.9 | 1601.9 |
| NUA596 | 40.9 | 52.5 | 46.7 | 28.3 | B | Large red | 82.6 | 43.4 | 1635.6 |
| GBK035381 | 43.4 | 52.6 | 48.0 | 21.1 | B | Purple | 77.9 | 53.4 | 1471.3 |
| KMT0114 | 47.9 | 52.6 | 50.3 | 9.8 | SC | Sugar | 78.9 | 63.0 | 1428.8 |
| GBK035078 | 45.7 | 52.8 | 49.3 | 15.4 | B | Carioca | 75.8 | 46.3 | 988.8 |
| KMT0113 | 49.8 | 52.8 | 51.3 | 6.1 | SC | Sugar | 80.8 | 46.8 | 1153.1 |
| GBK035057 | 35.3 | 53.0 | 44.2 | 49.9 | B | Carioca | 76.9 | 51.5 | 1279.4 |
| KNB0114 | 44.1 | 53.0 | 48.5 | 20.4 | SC | Small red | 81.3 | 36.5 | 903.5 |
| GBK035006 | 39.0 | 53.1 | 46.1 | 36.2 | B | BT | 76.8 | 47.9 | 1426.3 |
| GBK034992 | 43.1 | 53.4 | 48.2 | 23.9 | SC | Sugar | 81.9 | 44.3 | 971.9 |
| GBK035331 | 44.6 | 53.4 | 49.0 | 19.6 | B | Yellow | 79.5 | 51.1 | 848.1 |
| GBK034997 | 34.2 | 53.6 | 43.9 | 56.5 | C | Pinto | 80.8 | 34.8 | 1624.4 |
| NUA680 | 33.7 | 53.8 | 43.8 | 59.5 | SC | Sugar | 85.9 | 57.9 | 1542.5 |
| NUA686 | 45.8 | 54.2 | 50.0 | 18.1 | B | Calima | 81.9 | 40.5 | 1000.0 |
| KMT0106 | 37.4 | 54.3 | 45.8 | 45.3 | B | Purple | 80.3 | 63.4 | 2052.5 |

| Name | Cooking time | | | | Growth habit | Market class | DM | 100-seed weight (g) | Yield Kgha$^{-1}$ |
|---|---|---|---|---|---|---|---|---|---|
| | Fresh (min) | Aged (min) | Mean (min) | Increase (%) | | | | | |
| GBK035047b | 35.5 | 54.4 | 44.9 | 53.4 | B | Calima | 79.3 | 44.8 | 836.0 |
| GBK035043 | 34.3 | 54.4 | 44.3 | 58.7 | SC | Sugar | 82.0 | 44.1 | 1619.4 |
| GBK035402b | 38.1 | 54.6 | 46.4 | 43.4 | B | Large red | 83.1 | 33.8 | 649.5 |
| GBK035379 | 49.3 | 54.8 | 52.1 | 11.1 | C | Navy | 86.9 | 17.1 | 816.9 |
| GBK035346 | 45.3 | 54.8 | 50.1 | 20.9 | SC | black | 82.4 | 26.1 | 590.0 |
| GBK035434 | 43.0 | 54.9 | 48.9 | 27.7 | SC | Sugar | 80.6 | 50.9 | 1866.3 |
| GBK034965 | 48.9 | 54.9 | 51.9 | 12.2 | B | Sugar | 82.5 | 38.9 | 938.8 |
| GBK035332 | 42.3 | 55.1 | 48.7 | 30.3 | C | Black | 87.3 | 20.3 | 1318.1 |
| GBK035360 | 32.4 | 55.2 | 43.8 | 70.3 | B | Navy | 83.0 | 20.6 | 513.1 |
| NUA 619 | 39.3 | 55.2 | 47.3 | 40.6 | B | Large red | 84.8 | 58.1 | 1200.6 |
| GBK035362 | 42.7 | 55.4 | 49.0 | 29.8 | B | Navy | 88.3 | 20.1 | 3618.8 |
| KNB0116 | 30.7 | 55.7 | 43.2 | 81.6 | B | large red | 80.2 | 34.5 | 620.0 |
| GBK035063 | 41.5 | 55.8 | 48.6 | 34.3 | SC | Large red | 86.6 | 48.0 | 1904.4 |
| GBK035082 | 38.3 | 55.8 | 47.1 | 45.9 | B | Pinto | 78.6 | 45.0 | 1276.3 |
| NUA718 | 50.5 | 56.1 | 53.3 | 10.9 | SC | Calima | 78.9 | 45.1 | 1543.8 |
| KMT0102 | 40.4 | 56.4 | 48.4 | 39.7 | SC | Purple | 80.3 | 47.5 | 1985.0 |
| NUA604 | 38.3 | 56.6 | 47.5 | 48.0 | B | Small red | 83.5 | 44.8 | 1080.0 |
| GBK035071 | 43.9 | 57.0 | 50.4 | 30.0 | B | Carioca | 80.0 | 51.5 | 1229.4 |
| GBK035339 | 34.8 | 57.2 | 46.0 | 64.1 | B | BT | 83.3 | 30.8 | 987.5 |
| GBK035351 | 43.0 | 57.2 | 50.1 | 32.9 | SC | Navy | 86.6 | 16.8 | 603.1 |
| GBK034983b | 41.3 | 57.2 | 49.3 | 38.5 | C | Sugar | 83.5 | 33.7 | 1056.5 |
| GBK035402 | 39.5 | 57.4 | 48.5 | 45.2 | SC | Purple | 83.1 | 33.8 | 644.4 |
| NUA612 | 44.7 | 57.5 | 51.1 | 28.4 | B | Small red | 78.5 | 48.6 | 1538.1 |
| GBK035324a | 35.9 | 57.8 | 46.8 | 61.0 | C | Pinto | 83.8 | 25.2 | 1284.0 |
| GBK035416 | 50.5 | 58.1 | 54.3 | 14.9 | C | Navy | 84.0 | 30.6 | 1678.1 |
| GBK035090 | 48.8 | 58.5 | 53.7 | 19.8 | SC | Sugar | 81.3 | 39.9 | 2001.9 |
| GBK034345 | 36.8 | 59.1 | 47.9 | 60.6 | C | Purple | 85.3 | 42.4 | 1440.0 |
| NUA695 | 28.5 | 59.1 | 43.8 | 107.3 | B | Calima | 80.9 | 49.6 | 1605.6 |
| KNB0117 | 40.2 | 59.1 | 49.7 | 46.8 | SC | Pinto | 83.8 | 22.3 | 399.0 |
| KMT0117 | 57.9 | 59.3 | 58.6 | 2.3 | SC | Sugar | 82.6 | 51.9 | 1407.5 |
| NUA692 | 44.6 | 59.5 | 52.1 | 33.4 | B | Calima | 84.1 | 56.3 | 2021.9 |
| GBK035367 | 43.6 | 59.5 | 51.6 | 36.4 | B | Large red | 81.6 | 45.3 | 1660.0 |
| GBK035395 | 44.4 | 59.8 | 52.1 | 34.6 | C | MW | 85.1 | 32.4 | 2211.3 |
| KMT0112 | 38.0 | 60.4 | 49.2 | 59.0 | SC | Sugar | 87.3 | 56.5 | 1709.4 |
| GBK035021b | 47.9 | 60.7 | 54.3 | 26.7 | SC | Pinto | 79.5 | 35.7 | 896.0 |
| GBK035396 | 35.3 | 61.0 | 48.1 | 72.9 | B | MW | 74.0 | 31.6 | 1673.1 |
| KNB0115 | 44.7 | 62.6 | 53.6 | 40.2 | SC | Navy | 84.8 | 18.7 | 660.0 |
| GBK035039 | 31.5 | 62.8 | 47.1 | 99.4 | B | BT | 82.6 | 44.8 | 966.9 |
| GBK035439b | 45.6 | 63.0 | 54.3 | 38.1 | B | Sugar | 80.6 | 51.2 | 1215.0 |

| Name | Cooking time | | | | Growth habit | Market class | DM | 100-seed weight (g) | Yield Kgha$^{-1}$ |
|---|---|---|---|---|---|---|---|---|---|
| | Fresh (min) | Aged (min) | Mean (min) | Increase (%) | | | | | |
| KMT0109 | 44.9 | 63.0 | 54.0 | 40.4 | B | Sugar | 82.1 | 64.4 | 1923.1 |
| GBK035395a | 50.2 | 64.2 | 57.2 | 27.9 | SC | Navy | 83.4 | 20.1 | 713.1 |
| KMT0108 | 45.3 | 65.2 | 55.3 | 43.7 | B | Yellow | 77.1 | 47.6 | 860.0 |
| GBK035092 | 40.9 | 65.6 | 53.2 | 60.5 | SC | Sugar | 80.6 | 41.6 | 2523.1 |
| GBK035391 | 52.7 | 65.7 | 59.2 | 24.8 | B | Navy | 83.4 | 22.6 | 628.8 |
| GBK035378 | 40.8 | 66.3 | 53.5 | 62.7 | SC | MW | 75.3 | 27.3 | 1340.6 |
| NUA709 | 44.5 | 67.8 | 56.2 | 52.5 | B | Yellow | 76.1 | 48.5 | 1226.3 |
| GBK035342 | 45.7 | 67.9 | 56.8 | 48.7 | SC | Navy | 84.8 | 23.4 | 1353.8 |
| GBK035034 | 39.0 | 68.6 | 53.8 | 76.0 | B | Navy | 80.3 | 18.8 | 466.3 |
| GBK035053 | 44.6 | 68.9 | 56.8 | 54.5 | C | Navy | 87.5 | 26.8 | 1322.5 |
| KNB0113 | 47.9 | 69.1 | 58.5 | 44.3 | SC | BT | 83.5 | 25.2 | 833.5 |
| GBK034987 | 63.8 | 69.1 | 66.5 | 8.4 | B | Black | 85.1 | 17.8 | 1150.0 |
| GBK035282 | 36.9 | 69.2 | 53.1 | 87.6 | SC | Navy | 79.6 | 24.1 | 1132.5 |
| GBK034307 | 61.9 | 69.8 | 65.9 | 12.8 | SC | Navy | 85.6 | 16.9 | 1560.6 |
| GBK035295 | 41.6 | 70.4 | 56.0 | 69.1 | B | Navy | 82.1 | 20.4 | 288.8 |
| GBK035354 | 39.8 | 70.9 | 55.3 | 78.3 | SC | Navy | 83.5 | 22.3 | 1026.3 |
| GBK035337 | 46.4 | 71.3 | 58.9 | 53.6 | C | Navy | 83.8 | 19.6 | 858.8 |
| GBK035413 | 42.4 | 72.8 | 57.6 | 71.6 | B | Navy | 78.9 | 19.1 | 775.0 |
| GBK035284 | 62.6 | 73.7 | 68.1 | 17.7 | C | MW | 74.5 | 31.5 | 1506.9 |
| GBK035341 | 44.7 | 74.6 | 59.7 | 67.1 | B | Navy | 81.1 | 23.6 | 963.1 |
| GBK035384 | 53.5 | 75.9 | 64.7 | 41.6 | B | Sugar | 80.5 | 48.1 | 1548.1 |
| GBK034999 | 34.8 | 76.2 | 55.5 | 118.6 | SC | Navy | 79.0 | 24.1 | 1888.1 |
| GBK035395a | 39.5 | 77.1 | 58.3 | 94.9 | SC | MW | 85.1 | 32.4 | 2211.5 |
| GBK035277 | 66.3 | 77.2 | 71.8 | 16.4 | SC | Navy | 80.4 | 19.1 | 1015.0 |
| GBK035394 | 44.1 | 77.9 | 61.0 | 76.6 | B | Navy | 74.3 | 18.1 | 1076.3 |
| GBK035356 | 40.2 | 78.4 | 59.3 | 95.2 | SC | Navy | 84.3 | 19.9 | 733.8 |
| GBK035337 | 72.2 | 80.6 | 76.4 | 11.6 | SC | MW | 75.0 | 32.1 | 1470.6 |
| GBK035340 | 32.2 | 81.7 | 57.0 | 153.9 | SC | Navy | 81.0 | 22.0 | 514.4 |
| GBK035368 | 37.4 | 82.1 | 59.7 | 119.7 | C | Navy | 84.3 | 16.1 | 1971.9 |
| GBK035286 | 48.2 | 85.8 | 67.0 | 77.8 | SC | MW | 76.1 | 28.8 | 1157.5 |
| GBK035408 | 55.7 | 85.9 | 70.8 | 54.0 | B | MW | 78.8 | 31.6 | 659.4 |
| GBK035305 | 50.8 | 86.7 | 68.8 | 70.6 | C | MW | 78.0 | 35.6 | 1118.1 |
| GBK035355 | 61.2 | 87.3 | 74.2 | 42.5 | B | Navy | 82.8 | 21.6 | 702.5 |
| GBK035359 | 48.7 | 93.3 | 71.0 | 91.6 | SC | Navy | 87.1 | 25.5 | 1148.8 |
| GBK035338 | 51.4 | 94.8 | 73.1 | 84.2 | B | Navy | 75.1 | 18.8 | 1203.1 |
| GBK035322 | 54.5 | 96.3 | 75.4 | 76.7 | B | MW | 74.3 | 31.3 | 686.9 |
| GBK035370 | 54.9 | 96.3 | 75.6 | 75.6 | C | Navy | 86.9 | 21.3 | 1724.4 |
| **Commercial varieties** | | | | | | | | | |

| Name | Cooking time | | | | Growth habit | Market class | DM | 100-seed weight (g) | Yield Kgha$^{-1}$ |
|---|---|---|---|---|---|---|---|---|---|
| | Fresh (min) | Aged (min) | Mean (min) | Increase (%) | | | | | |
| GLP2 | 46.8 | 47.2 | 47.0 | 0.8 | B | Calima | 79.1 | 59.8 | 1422.5 |
| GLP24 | 43.0 | 48.4 | 45.7 | 12.5 | C | Small red | 83.8 | 23.8 | 1492.5 |
| GLPx92 | 49.7 | 60.8 | 55.3 | 22.3 | C | Pinto | 83.6 | 41.0 | 1995.0 |
| GLP1127a | 52.2 | 61.1 | 56.7 | 17.1 | B | Purple | 79.3 | 43.9 | 1268.1 |
| **Mean** | **40.8** | **54.5** | **47.7** | **34.5** | | | | | |
| LSD Accessions (A)** | | | 4.8 | | | | | | |
| LSD Storage (T)** | | | 0.5 | | | | | | |
| LSD Interaction (AxT)** | | | 7.3 | | | | | | |
| CV% | | | 7.2 | | | | | | |

LSD= Least significant difference, min= Minutes, CV= Coefficient of variation, **=Significant at 0.01 level, DM= Days to maturity, BT= Brown and Tan, B=Bush, SC=Semi climber, C=Climber, MW=Medium white

The cooking time profile of fresh and aged seeds was sigmoid; the difference between the cooking time of fresh and aged seeds was due to the prolonged lag phase followed by the slightly less steep exponential phase observed in the cooking time of aged seeds (Figure 4.2).

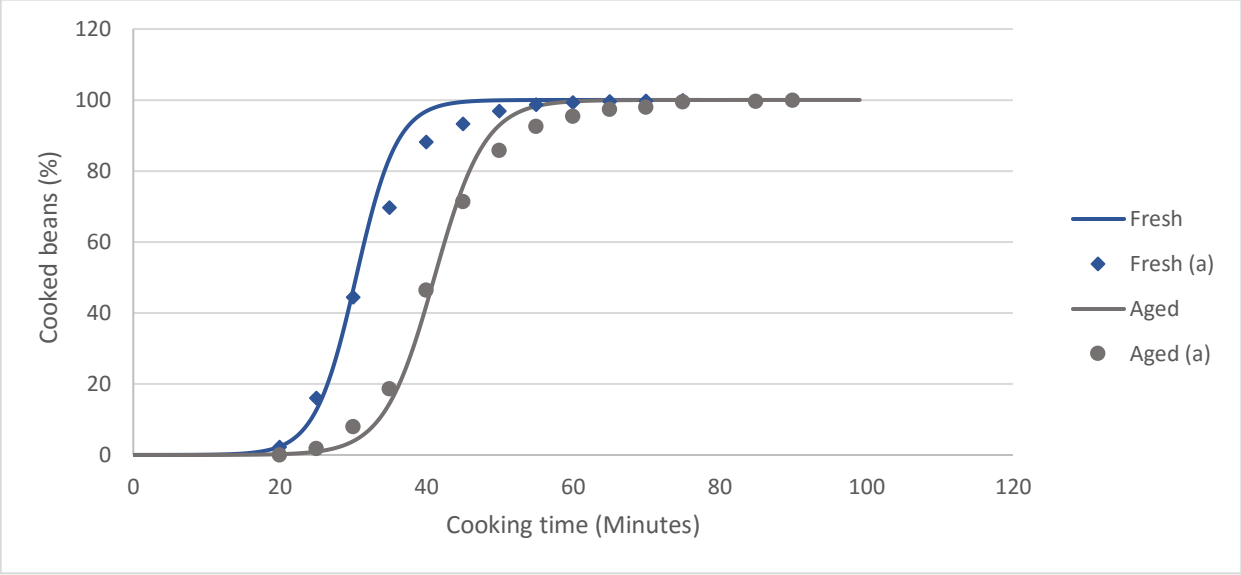


**Figure 4.2: Cooking time profile of fresh and aged common bean seeds using mean averages of all accessions evaluated**

There was a significant moderate positive correlation (0.59) between the cooking time of fresh and aged seeds. Cooking time for both fresh and aged seeds had a significant extremely weak negative association with days to maturity (-0.2), whereas cooking time of aged seeds had a significant but weak negative correlation with the number of pods per plant (-0.18), pod length (-0.18), and seed weight (-0.25) (Table 4.2).

**Table 4.2: Pearson correlation coefficient between cooking time and seven agronomic traits of common bean**

| | Aged CT | Fresh CT | Mean CT | DF | DM | PN | PL | SP | SW |
|---|---|---|---|---|---|---|---|---|---|
| | (Min) | (Min) | (Min) | (days) | (days) | (no.) | (cm) | (no.) | (g) |
| ACT | 1 | 0.59** | 0.95** | -0.11 | -0.2* | 0.18* | -0.18* | 0.07 | -0.25** |
| FCT | | 1 | 0.82** | -0.12 | -0.16* | 0.05 | -0.03 | -0.03 | -0.03 |
| Mean | | | 1 | -0.13 | -0.2** | 0.14 | -0.14 | 0.04 | -0.19* |

CT= Cooking time, DF=Days to flowering, DM=Days to maturity, PN=Number of pods, PL=Number of pods, SP= Number of seeds per pod, SW= 100 seed weight, ACT=Aged cooking time, FCT=Fresh cooking time, *, **=Significant 0.05 and 0.01 respectively.

The common bean accessions used in this study belonged to twelve seed classes commonly found in the eastern Africa region. The results show significant ($P \leq 0.05$) differences in cooking time for fresh and aged seeds among seed classes. Carioca seed class had the shortest cooking time while medium whites had the longest cooking time for both fresh and aged seeds (Table 4.3). The seed classes were categorized into seven groups based on the least significant difference in cooking duration (Table 4.3).

**Table 4.3: Cooking time of fresh and aged soaked seed of common bean market classes**

| Seed classes | No. of accessions | Cooking time | | | Increase (%) | Range of cooking time | |
|---|---|---|---|---|---|---|---|
| | | Fresh | Aged | Mean | | Fresh | Aged |
| Medium Whites | 7 | 51.9 | 78.6 | 65.2[a] | 51.3 | 21.8 | 30.0 |
| Navy | 32 | 44.6 | 70.1 | 57.4[b] | 57.4 | 42.0 | 56.9 |
| Yellow | 4 | 41.5 | 56.0 | 48.8[c] | 34.9 | 11.4 | 17.0 |
| Sugar | 29 | 41.7 | 53.9 | 47.8[c] | 29.4 | 26.4 | 31.9 |
| Pinto | 12 | 40.3 | 52.9 | 46.6[cd] | 31.2 | 17.5 | 18.5 |
| Purples | 9 | 40.9 | 52.0 | 46.4[cd] | 27.1 | 20.6 | 19.5 |
| Calima | 22 | 40.3 | 50.8 | 45.6[cd] | 26.0 | 27.2 | 48.4 |
| Large Reds | 16 | 39.4 | 50.5 | 45.0[cde] | 28.3 | 15.4 | 17.4 |
| Small Reds | 12 | 38.1 | 49.4 | 43.8[def] | 29.7 | 11.5 | 18.1 |
| Brown and tan | 18 | 37.4 | 50.1 | 43.7[def] | 33.9 | 19.8 | 36.9 |
| Black | 15 | 38.1 | 47.3 | 42.7[ef] | 24.0 | 35.5 | 33.4 |
| Carioca | 18 | 37.4 | 45.8 | 41.6[f] | 22.4 | 17.5 | 21.5 |
| **Overall mean** | | **40.8** | **54.9** | **47.8** | **34.5** | | |
| LSD Seed class** | | | 1.1 | | | | |
| LSD Storage** | | | 0.5 | | | | |
| LSD Seed class x Storage** | | | 7.3 | | | | |
| CV% | | | 7.2 | | | | |

Means followed by the same letter are not significantly different, **=Significant at 0.001 level

The profile of cooking time for fresh and aged seeds for small and medium whites was more distinct from the rest due to their prolonged cooking time mainly at the lag phase of the curve (Figure 4.4 and 4.5).

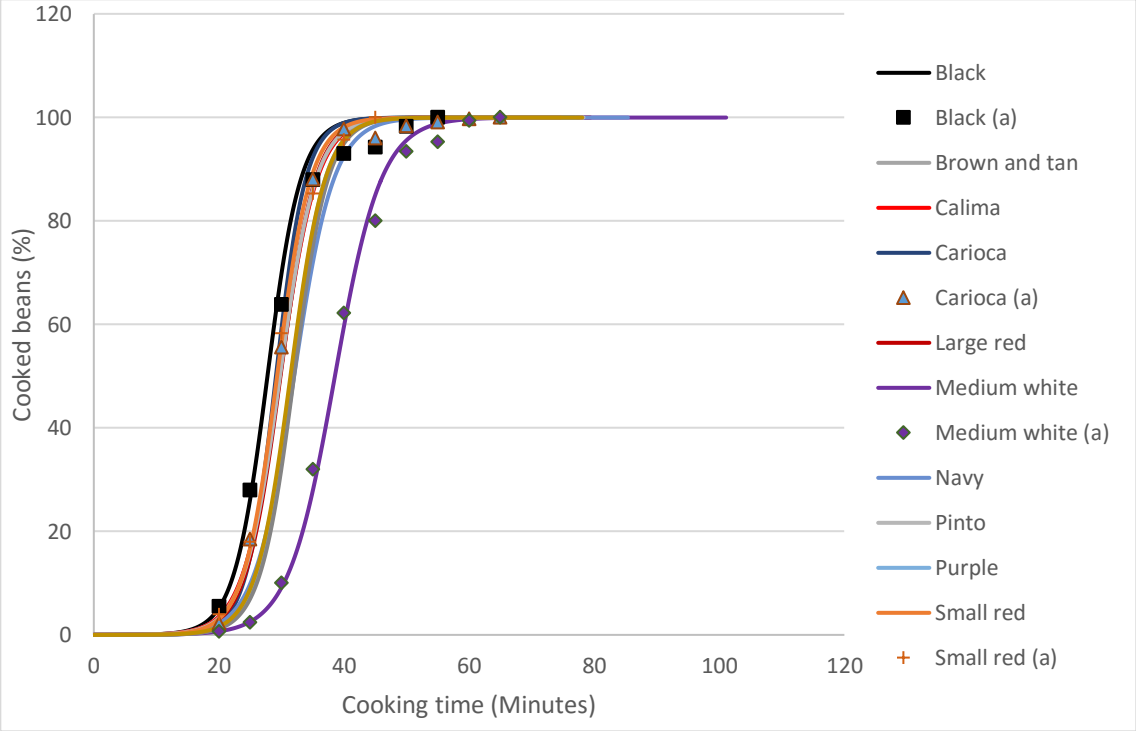


**Figure 4.3: Cooking time profile of fresh common bean seeds of different market classes**

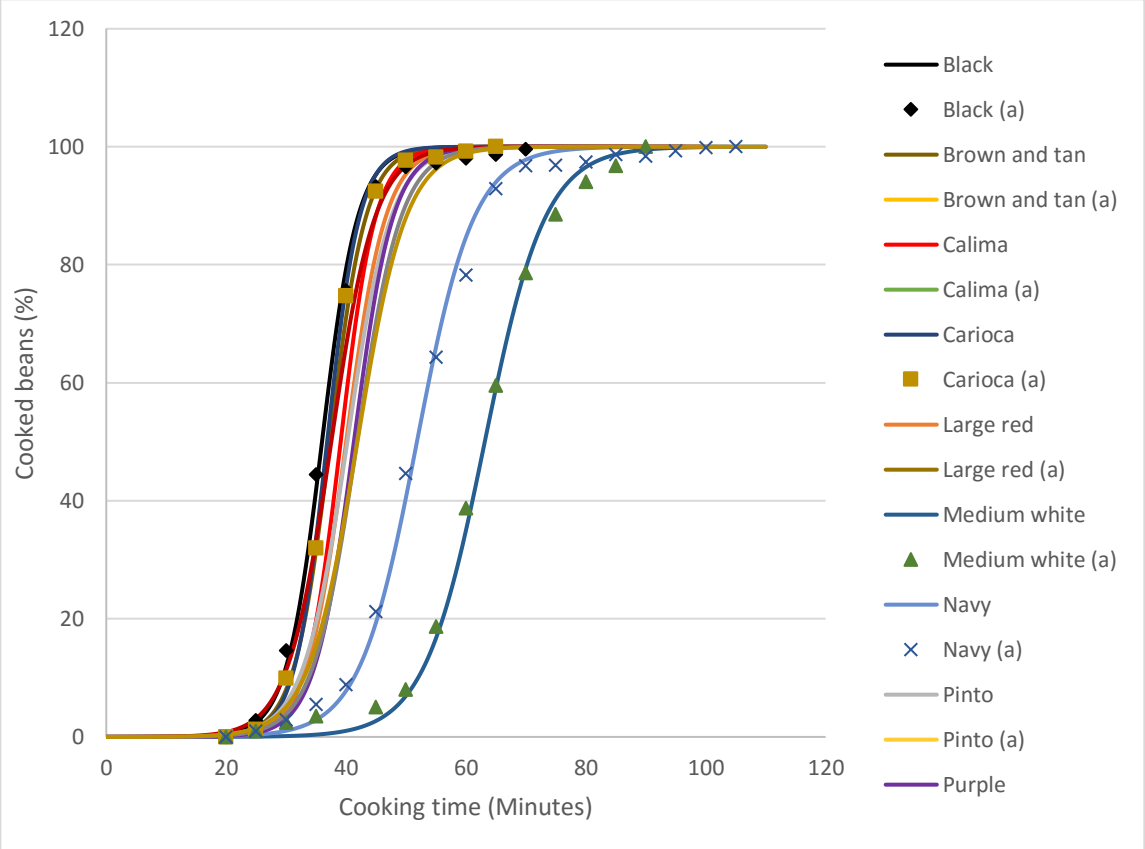


**Figure 4.4: Cooking time profile of aged common bean seeds of different market classes**

### 4.4.2 Marker data

The DArTseq produced 26945 SNPs markers of which 24878 were successfully assigned chromosomes, after filtering out SNPs with minor allele frequency (MAF) of <5% and missing data >20%, a final total of 19188 markers remained which were subjected to maker-trait association analysis. The average number of SNPs markers per chromosome was 1744.4, which ranged from 1387 markers on chromosome 4 to 2325 markers on chromosome 11. In the PC analysis, the first principal components accounted for 53.8% of the variation while the second and third principal components accounted for only 6.9% of the variation (Figure 4.6).

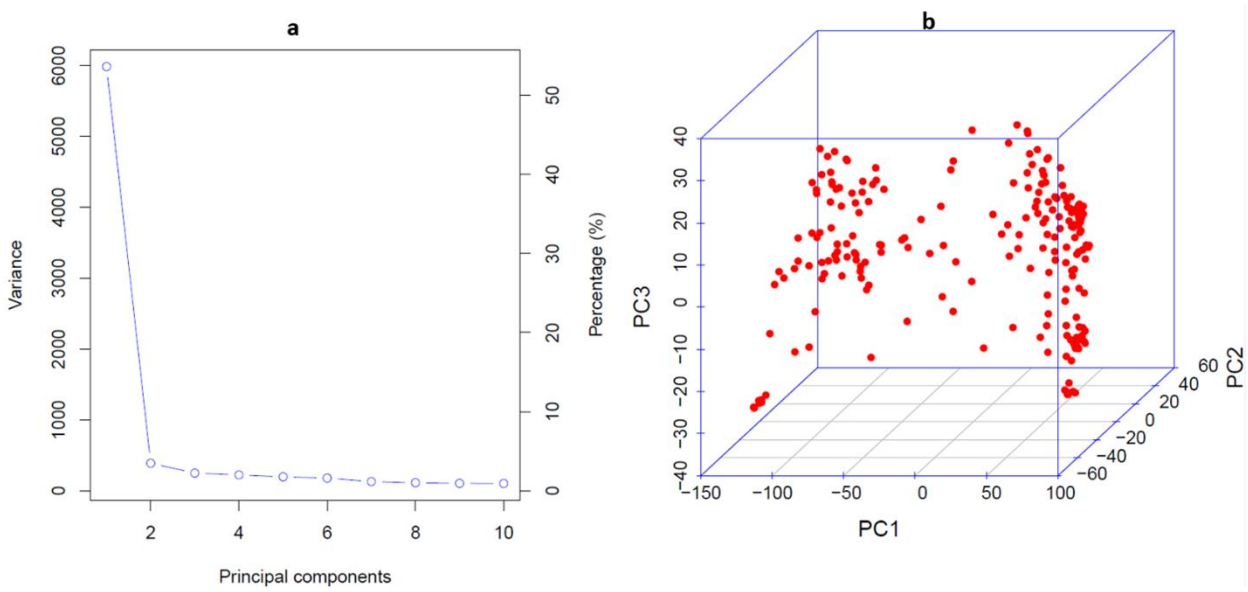


**Figure 4.5: Population structure of common bean accession, (a) amount of variation accounted for by principal components, (b) three-dimensional plot from principal component analysis for 222 common bean accessions and 1988 SNPs markers**

### 4.4.3 Linkage disequilibrium

Linkage disequilibrium (LD) analysis conducted using 1988 loci pairs showed pairwise 917276 loci with an average $r^2$ of 0.43 and ranged from 0.36 on chromosome 11 to 0.53 on chromosome 1 and extended to an average distance of 4406.2kb. About 2.9% of SNPs were in complete LD ($r^2$= 1). The average D' was 0.88 ranging from 0.84 on chromosome 11 to 0.91 on chromosome 1 (Fig 4.7).

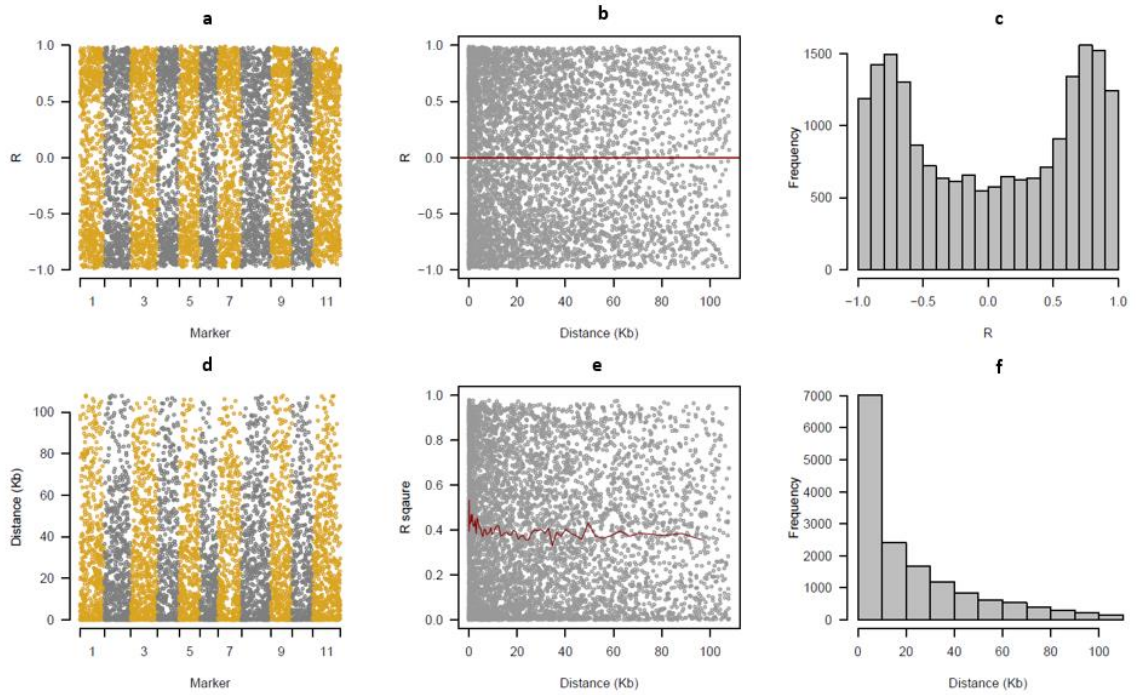


**Figure 4.6: (a) Distribution of makers by correlation (r) in each of the 11 chromosomes, (b) distribution of markers by correlation (r) as a function of genetic distance (kb), (c) distribution of SNPs markers by correlation (r), (d) distribution of markers along the 11 chromosomes, (e) Linkage disequilibrium (LD, $r^2$) decay plot for pairwise markers as a function of genetic distance (kb), (f) distribution of 1988 SNPs markers by genetic distance (kb) for the 222 common bean accessions**

### 4.4.4 Marker-trait association

Genome-wide analysis revealed two significant (P≤0.05) SNP markers associated with the cooking time of aged seeds on chromosome 10 (Figure 4.8). However, no significant SNP markers were identified associated with the cooking time of fresh seeds. The most significant SNP maker was 100096770|F|0-21:G>A-21:G>A on chromosome 10 at location 5600323 with a P-value of $6.9x10^{-7}$ which explained 36% of the phenotypic variation. The second SNP marker that was significantly associated with the cooking time of aged seeds was 3377419|F|0-24:A>T-24:A>T on chromosome 10 at location 4468450

with a P-value of $9.8x10^{-6}$ which explained 34% of the phenotypic variation. The two significant SNPs 100096770|F|0-21:G>A-21:G>A and 3377419|F|0-24:A>T-24:A>T had an allelic effect of 13.21% and 10.8% respectively were in strong linkage disequilibrium linkage disequilibrium with $r^2$ of 0.79, D prime of 0.95 (Table 4.4). The region around these SNPs markers also had several other SNP markers near significance ($P \leq 0.05$) threshold with a P-value ranging from $3.02x10^{-5}$ to $6.52x10^{-5}$ due to linkage disequilibrium (LD) (Figure 4.7).

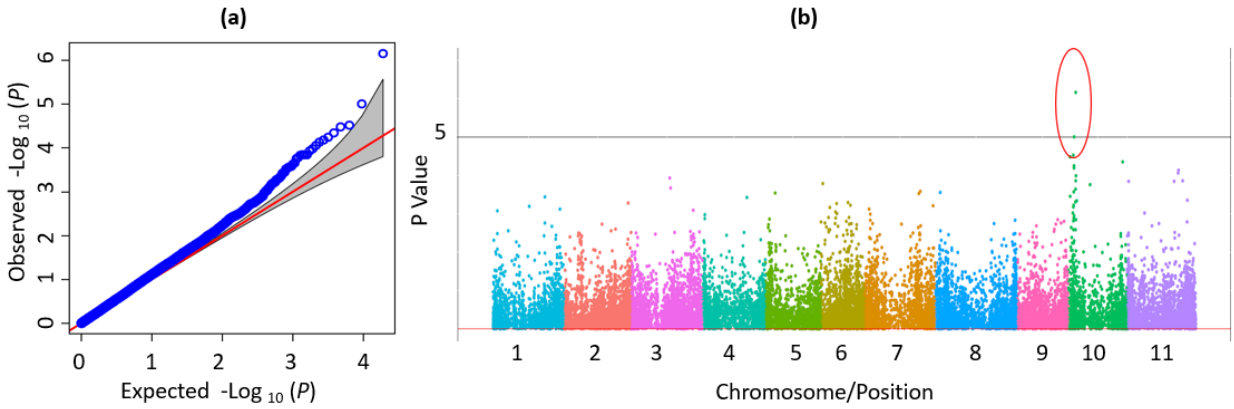


**Figure 4.7: (a) Quantile-quantile plot of estimated -log (P) for association analysis of cooking time. (b) Manhattan plot showing significant SNPs and their P-values for cooking time of aged and soaked common bean accessions**

**Table 4.4: Significant SNPs associated with cooking time of aged common bean accessions**

| SNP | Chr | Position | P-value | MAF | $R^2$ | AE (%) |
|---|---|---|---|---|---|---|
| 100096770\|F\|0-21:G>A-21:G>A | 10 | 5600323 | $6.9x10^{-7}$ | 0.1 | 0.36 | 13.2 |
| 3377419\|F\|0-24:A>T-24:A>T | 10 | 4468450 | $9.8x10^{-6}$ | 0.15 | 0.34 | 10.8 |

SNP=Single Nucleotide Polymorphism, Chr=Chromosome, MAF=Minor allele frequency, $R^2$=Phenotypic variation explained by the SNP, AE=Allelic effect.

### 4.4.5 Identification of potential candidate gene

When potential candidate genes were explored in the Phytozome database, three potential candidate genes were identified near the location of the most significant ($P \leq 0.05$) SNP marker and eight potential candidate genes were identified around the second most significant SNP marker at around 100kb of the location of the SNP markers. The nearest gene to the location of the most significant SNP at Phytozome database on chromosome 10 were Phvul.010G038000 at 5575958 bp, Phvul.010G038100.1 at 5627743 bp and Phvul.010G038200.1 at location 5644933 bp (Table 4.5).

A BLASTn search of the most significant SNP 100096770|F|0-21:G>A-21:G>A sequence (TGCAGTACCAGAAAAACAATCGGTTGTTTTACAAAAACAATCGGTTGTTTT ACAAAAACAATCGGTTGT) at NCBI database revealed Sequence ID AC254323.*1* and AC254327.1 of *Phaseolus vulgaris* L. with a 90.5% and 89.1% match, respectively. The functional annotation for transcript *Phvul.010G038000* indicated galacturan, 1, 4-alpha galacturonidase enzyme, polygalacturonase/pectinase enzyme for *Phvul.010G038100.1*, and NADPH-dependant alkenal reductase enzyme for *Phvul.010G038200.1*. There were no defined functions for the genes identified in the NCBI database.

**Table 4.5: Potential candidate genes identified in Phytozome data base around SNPs significantly associated with cooking time of common bean accessions**

| | | Phytozome map | |
|---|---|---|---|
| **Significant SNP** | **SNP position** | **Phytozome transcript (100kb)** | **Position in Chr 10** |
| 100096770\|F\|0-21:G>A-21:G>A | 5600172 | Phvul.010G038100.1 | 5627744 |
| | | Phvul.010G038000 | 5575959 |
| | | Phvul.010G038200.1 | 5644934 |
| 3377419\|F\|0-24:A>T-24:A>T | 4468484 | Phvul. 010G030800.1 | 4468349 |
| | | Phvul. 010G030900.1 | 4472539 |
| | | Phvul. 010G030700.1 | 4462259 |
| | | Phvul. 010G030600.2 | 4452797 |
| | | Phvul. 010G031000.2 | 4482915 |
| | | Phvul. 010G030500.2 | 4424827 |
| | | Phvul. 010G030600.1 | 4452790 |
| | | Phvul. 010G031100.1 | 4512862 |

SNP=Single Nucleotide Polymorphism, Chr=Chromosome

Most of the genes identified in the Phytozome database around the second significant SNP marker SNP 3377419|F|0-24:A>T-24:A>T did not have a defined function, the nearest transcript to the marker was *Phvul.010G030800.1* with functional annotation of asparagine tRNA ligase enzyme which belongs to class II family of tRNA synthetases localized in the cytoplasm where it plays a role in protein synthesis. A BASTN in search of the sequence of the second most significant SNP marker (TGCAGTTCAGGATCTGAAGAAAACAAATGACCTGGCATCACAATTTGAAGC AAGAGAAAACAGAAAGTT) in NCBI database revealed five genes with no defined function (Table 4.5).

### 4.5 Discussion

The variation in cooking time was higher for aged seeds when compared to fresh seeds. However, the cooking time for fresh seeds among bean accessions was also significantly ($P \leq 0.05$) different. This indicates that progress can be made when selecting common bean

accessions with a shorter cooking time using both fresh and aged seeds since high narrow sense heritability of 0.76 and 0.74 for cooking time has been reported in previous studies by Elia (2003) and Jacinto-Hernandez *et al.,* (2003) respectively. The significant difference observed between the cooking time of aged and fresh seeds indicates that the storage conditions significantly affected this trait. It is also evident that the extent of increase in cooking time due to storage varied among common bean accessions. This suggests that common bean susceptibility to hardening during storage varies among accessions. Accessions NUA Ciankui and commercial variety GLP2 (Rosecoco) were less susceptible to aging due to storage in adverse conditions. Previous studies have reported an increase in the cooking time of common beans stored in higher temperatures and relative humidity as demonstrated in this study (Nyakuni *et al.,* 2008; Ousman *et al.,* 2013). Nyakuni *et al.,* (2008) reported an increase in cooking time of four varieties stored in ambient temperatures and relative humidity of 63-74% and the percentage increase in the cooking time varied among varieties. Several other studies have reported that some varieties have a shorter cooking time while others have a prolonged cooking time. (Bressani and Chon 1996; Nyakuni *et al.,* 2008; Kinyanjui *et al.,* 2015).

The cooking time for commercial varieties observed in this study follows the same trend as reported in previous studies, where Rosecoco (GLP2) and Red Haricot (GLP24) have a relatively shorter cooking time in comparison to Pinto (GLPx92) and Canadian Wonder (GLP1127a) (Kinyanjui *et al.,* (2015). The genetic diversity in cooking time among the common bean accessions evaluated provides variability that can be utilized in breeding. The accessions with shorter cooking time can serve as genetic resources for selection or hybridization schemes to generate new common bean varieties that are easy-to-cook. Accessions with shorter cooking time can be improved through direct selection based on other desirable traits.

The initial lag phase of the cooking time curve creates much of the cooking time difference between fresh and aged seeds. The lag phase may be the stage at which the pectin is solubilized within the middle lamella to allow water to imbibe into the cells of the cotyledon. Njoroge *et al.,* (2014) reported that varieties that have a shorter cooking time

had a higher hot water-soluble pectin (8.44 mg/g) than slow cooking beans like pinto (5.51 mg/g). If this is the case, the modification of the composition of middle lamella may make beans easy to cook as proposed by Broughton *et al.,* (2003). The cooking time profile in this study agree with the findings of Kinyanjui *et al.,* (2015) who reported that cooking time is sigmoid with the lag and exponential phase being influenced by variety and storage.

Classification of common bean market classes is based on seed size and color. Seed color is determined by the presence and concentrations of flavanol glycosides, anthocyanins, and condensed tannins in the seed coat (Reynoso *et al.,* 2006). Lei *et al.,* (2020) classified those that weigh less than 25 g per 100 seeds as small-seeded while those that range from 25 to 40 g as medium-sized, and those that weigh more than 40 g are classified as large-seeded. The cluster analysis results in this study classified the accession into two groups mostly based on seed size. The results also indicate that large-seeded accessions had a relatively shorter cooking time, unlike the medium and small-seeded genotypes. Seed size depends on the genetic difference among varieties and can be traced back to the origin of the common bean, namely Mesoamerican and Andean gene pools (Angioi *et al.,* 2010). The Andean gene pool is generally large-seeded while the Mesoamerican gene pool is small-seeded and adapted to lower altitudes and higher temperatures (Beebe *et al.,* 2011). The significant results for the cooking time differences among the common bean seed classes used in this study confirm that the trait varies between and within seed classes as reported by Cichy *et al.,* (2015). However, the findings in this study contradict those of Cichy *et al.,* (2015), which found the white seed class to have the shortest cooking time in relative to other bean classes. This could be due to differences in the genetic background of the white bean accessions used in these two studies, the environment they were grown in, and the interaction between genotypes and the environment. In this study, the common bean accessions were sourced from Gene Bank of Kenya which had been collected mainly from Kenya while Cichy *et al.,* (2015) evaluated common bean accessions of Andean origin collected from Africa and North America.

The moderate (0.59) correlation in cooking time between fresh and aged seeds indicates that accessions with a higher cooking time of fresh seeds are more susceptible to hardening during storage in adverse conditions. This suggests that freshly harvested seed can be used to an extent for indirect selection for easy-to-cook accessions. The significant extremely weak negative association of cooking time with duration to maturity, pod length, and seed weight indicates that common bean accessions that have a longer duration to mature, have longer pods and larger seeds tend to have a slightly shorter cooking time. Similarly, common bean accession with a higher number of pods per plant will tend to have a slightly longer cooking time for aged seeds. This could be attributed to linkage disequilibrium where some genes tend to be inherited together or the presence of a third variable such as the prevailing temperature and humidity during growth. Similar negative correlation results between cooking time and duration to maturity (-0.44), and seed weight (-0.21) were reported in a previous study (Cichy *et al.*, 2019). A study conducted by Berry *et al.*, (2020) also revealed a negative correlation between cooking time and seed weight which ranged from -0.3 to -0.8 depending on the environment the common bean was grown in.

For a rapid and effective breeding program, genetic markers are used to identify QTLs for desirable traits. In this study, GWAS was used to identify the genomic regions that control cooking time in common beans. The lack of significant SNP associated with the cooking time of soaked freshly harvested seeds could be attributed to insufficient variation in the cooking time among fresh common bean accessions.

Two significant SNP markers were identified to be associated with the cooking time of soaked aged seeds. The study identified two positional potential candidate genes *Phvul.010G038000* and *Phvul.010G038100.1* that control galacturan 1,4-alpha galacturonidase and polygalacturonase enzymes, respectively, close to the position of the most significant SNP marker. The two enzymes are known to be involved in the breakdown of pectin in the plant cell wall. Therefore, this study supports the theory of the formation of insoluble pectin as the cause of the hard-to-cook trait. The difference in cooking time of stored bean accession is due to the activity of galacturan 1,4-alpha galacturonidase and polygalacturonase enzymes that hydrolyse pectin. The hydrolyses of

pectin could probably have happened during storage as proposed by Jones and Boulter, (1983) for pectin to translocate to middle lamella and combine with magnesium and calcium ions to produce an insoluble magnesium pectinate and calcium pectinate that cements cells together hardening the cell wall. Alternatively, the hydrolysis of pectin could have occurred during the pre-soaking treatment of aged seeds before cooking.

Pectin is made up of complex acid polysaccharides with a backbone of galacturonic acid residue with an alpha-1,4-glycosidic linkage. Homogalacturonan-rich pectin is commonly found in the middle lamellar region of plant cell walls where two cells border (Atkinson *et al.,* 2002). Galacturan 1,4-alpha galacturonidase enzyme is known to hydrolyze the first group of glycosidic bonds from the non-reducing end of the substrate, while polygalacturonase enzymes break down the pectin components found in the middle lamella of plant cells after PME makes the polymeric backbone accessible. The combined effect of both PME and pectinase enzymes has been reported to give softer fruits and vegetables at the end of maturation (Phutela *et al.,* 2005).

The second most significant SNP marker associated with the cooking time marker may be due to linkage disequilibrium ($r^2 = 0.74$, $D' = 0.95$) because of its proximity to the gene that controls HTC. The genes identified around this SNP allele variant marker may not be involved in the control of cooking time but based on linkage disequilibrium the marker can be used to determine the haplotype containing genes of interest. The identification of the two enzymes supports the theory of the formation of insoluble pectin in the middle lamella as the cause of the occurrence of HTC phenomena in common beans during storage in conditions of high temperature and relative humidity (Jones and Boulter, 1983).

Several studies have been carried out to map QTLs that control cooking time. A random amplified polymorphic DNA (RAPD) marker associated with cooking time was identified using 104 recombinant inbred lines, the identified marker explained 23% of the variation in cooking time (Jacinto-Hernandez *et al.,* 2003). Garcia *et al.,* (2012) mapped 6 QTLs that govern cooking time on chromosomes 1 and 9 using 105 polymorphic SSR markers and 140 $F_{2:4}$ recombinant inbred lines. The study found QTL CT1.1 on chromosome 1 which explained 21% of the phenotypic variation to be the most promising.

In a genome-wide association study using freshly harvested seeds, significant SNPs associated with cooking time were identified on chromosomes 2, 3, and 6 using 206 common bean accessions of Andean origin, the SNPs explained between 4 to 8.7% of the phenotypic variation (Cichy *et al.*, 2015). A recent study conducted by Berry *et al.*, (2020) using freshly harvested 146 recombinant inbred lines of common bean, identified 10 QTLs on chromosomes 1, 2, 3, 5, 6, 10, and 11, with the most robust QTLs being on chromosomes 3, 6, 10 and 11 that appeared in over two different environments. The variations observed in these studies could have resulted from different number of markers in some studies that affects markers saturation, pretreatments of seeds before the determination of cooking time such storage conditions and soaking, and the limitations and advantages different types mapping populations used in the studies.

### 4.6 Conclusion

This study assessed variation in cooking time among fresh and aged common bean accessions. In addition, genotyping of the common bean accessions was conducted to identify single nucleotide polymorphisms (SNPs) markers associated with cooking time. Storage significantly ($P \leq 0.05$) increased the cooking time of the common bean accessions, but the increase varied among the common bean accessions. Significant difference for cooking time among the common bean accession was observed. Genome-wide association analysis (GWAS) revealed two SNPs markers that were significant associated with the cooking time of the aged common bean accessions on chromosome 10. Consequently, two potential candidate genes co-localized with the most significant SNP marker. The two genes *Phvul.010G038000* and *Phvul.010G038100* encode galacturan 1, 4-alpha galacturonidase and polygalacturonase enzymes respectively. The two enzymes are involved in the hydrolysis of pectin in the plant cell wall. The easy-to-cook accessions and the SNP markers significantly associated with cooking time can be utilized in breeding programs to improve the cooking time of the common bean.

# CHAPTER FIVE

# GENOME WIDE ASSOCIATION STUDY OF VARIATION IN AGRONOMIC TRAITS AMONG COMMON BEAN (*PHASEOLUS VULGARIS* L.) ACCESSIONS USING DIVERSITY ARRRAYS TECHNOLOGY MARKERS

## 5.1 Abstract

Common bean is a major source of nourishment for a large population of people globally. Identifying genomic regions that control various critical agronomic traits of common bean is crucial to aid in the efforts to improve the crop through breeding. genome-wide association studies (GWAS) analysis is a popular method used to identify quantitative trait loci (QTLs) associated with phenotypic traits. The objective of the study was to identify genomic regions of common bean that control various agronomic traits using GWAS. A total of 234 common bean accessions obtained from selected farmers' fields and the National Gene Bank of Kenya were used in this study. Field experiments were conducted for four seasons at Jomo Kenyatta University of Agriculture and Technology (JKUAT), Kenya. Agronomic data recorded included days to flowering, days to maturity, number of pods per plant, pod length, number of seeds per pod, 100-seed weight, and seed yield. Genotyping By Sequencing (GBS) approach was used to generate single nucleotide polymorphism (SNP) markers using the Diversity Arrays Technology Sequencing (DArTseq) method. Based on 24879 SNP markers and seven agronomic traits, association analysis was carried out using the GWAS method to identify SNPs associated with various agronomic traits. The results revealed significant differences ($P \leq 0.05$) among accessions for all the traits evaluated, the seasonal and the interaction between accessions and seasons were also significant. A total of 32 SNPs associated with various agronomic traits were identified, and the association between trait and molecular marker was found to be influenced by the seasonal changes. Various possible potential candidate genes for various traits were identified around 100 kb, the position of the trait-associated SNP. The study managed to estimate the position and effects of genomic regions associated with various agronomic traits evaluated under different environmental conditions. The identified

markers can be used in developing varieties with desirable traits through markers assisted selection (MAS).

### 5.2 Introduction

Common bean (*Phaseolus vulgaris* L.) is an important crop for human nutrition. It is rich in proteins, dietary fibers, carbohydrates, vitamins, and minerals, playing a vital role in the development of plant-based diets, addressing malnutrition and promoting food security (Mecha *et al.,* 2018; Murube *et al.,* 2021). The crop is a major source of protein, carbohydrates, dietary fiber, and essential minerals to a large population globally (Broughton *et al.,* 2003;). A wide diversity of traits in common beans exist in duration to flowering and maturity, resistance to biotic and abiotic stresses, seed characteristics, nutritional quality, and yield (Fisseha *et al.,* 2018; Farrow and Andriatsitohaina, 2021).

Seed yield improvement is the most crucial breeding objective for the common bean. Understanding the genetic architecture of yield and its interaction with other yield components would lay a genetic foundation for improving seed yield (Vandemark *et al.,* 2014). Seed yield is a quantitative trait governed by multiple genes, and is conditioned primarily by three yield components namely, the number of pods per plant, the number of seeds per pod, and seed weight (Ghobary and Allah, 2010; Negahi *et al.,* 2014). The three yield components are quantitative and their interaction with seed yield is based on the physiological and morphological features of a plant (Burbano-Erazo *et al.,* 2021). The knowledge of the interaction between these yield attributes and seed yield may help identify a suitable donor for this trait.

Seed yield has been reported to be significantly positively associated with the number of pods per plant, number of seeds per pod, seed weight, and biomass yield (Ghobary and Allah, 2010; Negahi *et al.,* 2014). Ghobary and Allah, (2010) noted that breeders should be attentive to these traits as they are linked and exhibit a positive direct effect on seed yield. High-yielding varieties were also reported to flower early, mature late, taller, and have a higher number of primary branches per plant, pods per plant, and seeds per pod (Ashango and Alamerew, 2017). Genetic control for most plant traits has been reported to be predominantly due to genes with additive effects (Silva *et al.,* 2013). A more detailed

understanding of important traits at the molecular level is therefore critical to improving these traits in common bean.

The use of Deoxyribonucleic acid (DNA) analysis techniques in common bean breeding programs has improved our understanding of genetic factors controlling various traits. Several mapping studies in common bean have been conducted to understand genomic regions contributing to traits using different techniques and populations (Mukeshimana *et al.,* 2014; Kamfwa *et al.,* 2015; Briñez *et al.,* 2017; Sandhu *et al.,* 2018; Langat *et al.,* 2019).

Next-generation sequencing (NGS) using Single Nucleotide Polymorphism (SNPs) has become more practical in genotyping and discovery of markers. Marker-assisted selection (MAS) using SNP technology is much faster and inexpensive than the older generation markers that were gel-based (Gujaria-Verma *et al.,* 2016). The large number of SNPs that have been identified allow explorations of genetic diversity and population structure (Cichy *et al.,* 2015; Valdisser *et al.,* 2016). Genome-Wide Association (GWAS) is a popular method used to identify QTLs associated with bean traits. It involves the application of molecular markers to plant breeding using statistical methods to estimate the position and effects of genomic regions associated with variation in quantitative traits (Kafwa *et al.,* 2015). GWAs has been used to identify QTL related to biotic stress (Zuiderveen *et al.,* 2016; Perseguini *et al.,* 2016), abiotic stress (Villordo-Pineda *et al.,* 2015), agronomic traits (Kamfwa *et al.,* 2015; Rasende *et al.,* 2018), pod shattering (Ugwuanyi *et al.,* 2022) and grain quality (Cichy *et al.,* 2015). This method is important in breeding because it is dependent on genotype-environment interactions (GxE) (Beebe *et al.,* 2011), thereby giving an understanding of GxE at a molecular level, especially for a self-pollinating plant like the common bean that has been adapting to a constantly changing environment (Li *et al.,* 2003). The objective of this study is to identify genomic regions that control various agronomic traits of the common bean using GWAS.

## 5.3 Materials and Methods

### 5.3.1 Plant materials

A total of 234 of the 257 common bean accessions were used in this study sourced from selected farmers field and the National Gene Bank of Kenya. The number of accessions used, and their market classes is described in Table 5.1.

Field experiments were conducted for over four seasons at Jomo Kenyatta University of Agriculture and Technology (JKUAT) farm, Kenya. The site coordinates, seasonal rainfall and temperatures, and type of soil is described in section 3.3.1. The monthly rainfall and temperature that prevailed in the area in 2018 and 2019 is shown in Appendix 4. Data collected included days to flowering, days to maturity, number of pods per plant, number of seeds per pod, seed weight, and plot seed yield as described in section 3.3.4.

**Table 5.1: Common bean accessions used in this study and their characteristics**

| Seed class | Description | Number of accessions |
|---|---|---|
| Sugars | Cream, can be speckled | 38 |
| Carioca | Red and Red specks | 28 |
| Navy | Small whites | 26 |
| Calima | Rosecoco type | 24 |
| Pinto | Cream with brown specks-GLPx92 type | 21 |
| Black | Black coloured | 20 |
| Brown and tan | Brown and orange | 17 |
| Large Reds | Canadian wonder type | 16 |
| Small Reds | Red haricot type | 14 |
| Medium Whites | Medium and large whites | 12 |
| Purples | Mwezimoja type | 11 |
| Yellow | Yellow coloured | 7 |
| Total | | 234 |

### 5.3.2 Experimental design and trial management

The experimental design used and the management of the field experiment is as described in section 3.3.3.

### 5.3.3 DNA extraction and genotyping

DNA isolation and genotyping were conducted as described in section 4.3.3.

### 5.3.4 Data analysis

Data collected from field experiments was combined over seasons and subjected to a normality test and analysis of variance (ANOVA) using R software (version 4.0.2). The augmented block model described below was employed.

$$Y=Mx+Zg+S\beta+Tk+e$$

Where Y is the data vector, x is the vector of the assumed fixed effect (means of genotypes), g is the vector of the genotypic effect of the bean accessions, β is the vector of environmental (seasonal) effects, and k is the vector of the effects of the genotype and environmental interaction (GxE), and *e* is the vector of the residual effects. The capital letter M, Z, S, and T represents the incidence matrices for these effects.

Marker-trait association analysis and identification of potential candidate gene was conducted as described in section 4.3.6 and 4.3.7 respectively.

### 5.3.5 Linkage disequilibrium

Linkage disequilibrium analysis are described in section 4.3.5

### 5.3.6 Marker-trait association

Association analysis to identify SNP markers associated with phenotypic traits is described in section 4.3.6

### 5.3.7 Identification of potential candidate genes

The identification of potential candidate gene is described in section 4.3.7

## 5.4 Results

### 5.4.1 Phenotypic traits

Significant differences ($P \leq 0.05$) were observed among accessions for all the traits, the seasonal effect, and the interaction effect of seasons and accessions was also significant for all traits evaluated (Appendix 1).

The means, range, residual mean squares, standard error, and coefficient of variation for traits recorded are summarized in Table 5.2. The coefficient of variation ranged from 4.0% (days to maturity) to 36.5% (number of pods per plant). Overall mean for days to flowering, days to maturity, the number of seeds per pod, 100-seed weight, and grain yield were high during the long rain season than short rain season while the number of pods per plant and pod length was higher during the short rain season (Table 5.2).

**Table 5.2: Descriptive statistics for agronomic traits of common bean accessions**

| Trait | Season | Mean | Max | Min | MSE | SE | CV% |
|---|---|---|---|---|---|---|---|
| Days to flowering | LR | 38.4 | 48.0 | 32.5 | 2.41 | 0.14 | 4.0 |
| | SR | 37.4 | 44 | 32.25 | 2.6 | 0.10 | 4.3 |
| Days to maturity | LR | 82.3 | 99.0 | 72.0 | 16.0 | 0.18 | 4.9 |
| | SR | 81.7 | 89.8 | 70.0 | 7.8 | 0.17 | 3.4 |
| Pod length | LR | 9.9 | 14.8 | 7.0 | 1.6 | 0.06 | 12.8 |
| | SR | 9.9 | 17.3 | 6.6 | 2.3 | 0.07 | 15.1 |
| Seed per pod | LR | 4.9 | 7.5 | 3.3 | 0.6 | 0.03 | 15.1 |
| | SR | 4.3 | 6.3 | 2.8 | 0.6 | 0.03 | 18.0 |
| 100 seed weight | LR | 37.6 | 67.8 | 15.0 | 20.0 | 0.44 | 11.9 |
| | SR | 36.2 | 70.5 | 14.8 | 19.4 | 0.40 | 12.2 |
| Pods per plant | LR | 12.3 | 27.5 | 5.5 | 13.9 | 0.24 | 30.3 |
| | SR | 12.8 | 36.6 | 3.8 | 21.8 | 0.25 | 36.5 |
| Yield ($kgha^{-1}$) | LR | 1528.3 | 4926.3 | 248.8 | 212824 | 26.21 | 30.2 |
| | SR | 1005.7 | 4357.8 | 75 | 121967.0 | 22.59 | 34.7 |

n=234, LR=Long rains, SR=Short rains, CV=Coefficient of variation, MSE=Residual mean sum of squares, Max=Maximum, Min=Minimum, SE=Standard error.

### 5.4.2 Single Nucleotide Polymorphism markers

The number of SNPs significantly ($P \leq 0.05$) associated with each trait are displayed on Table 5.3. The total number of significant SNPs obtained ranged from one to 14, with

days to flowering and pod length having the highest number of significant SNPs of 14 and 12 respectively (Table 5.3).

**Table 5.3: Number of significant SNPs identified for various agronomic traits**

| Trait | SR season | LR season | Four seasons | Total |
|---|---|---|---|---|
| Days to flowering | 10 | 3 | 6 | 13 |
| Days to maturity | 0 | 0 | 2 | 2 |
| Pods per plant | 1 | 0 | 0 | 1 |
| Pod length | 3 | 10 | 7 | 12 |
| Seeds per pod | 1 | 0 | 1 | 2 |
| 100 seed weight | 0 | 0 | 1 | 1 |
| Grain yield | 0 | 0 | 1 | 1 |

SNP=Single Nucleotide Polymorphism, SR=Short rain season, LR=Long rain season.

Results show that the number of identified SNPs varied with trait and the season, some SNPs that were significantly associated with traits during long rain season were not identified during short rain season and vice versa. The results did not reveal any significant ($P \leq 0.05$) SNPs associated with days to maturity, number of seeds per pod, 100-seed weight, and grain yield during long and short rain seasons, but when the overall average of the four seasons for these traits was used in marker-trait association analyses we identified a few significant SNPs. Furthermore, SNPs significantly ($P \leq 0.05$) associated with the number of pods per plant and number of seeds per pod were only identified during the short rain season (Table 5.4).

### 5.4.3 Duration to flowering and maturity

In total, thirteen significant ($P \leq 0.05$) SNPs markers were identified for days to flowering, six SNPs were identified using average data from the four seasons (Figure 5.1) ten were also identified during the short rain season, and three during the long rain season. Eight of these SNPs were found on chromosome one all located around positions 44146828 to 44960059 bp. In this region SNP marker 8206455|F|0-42:G>A-42:G>A with a nucleotide sequence TGCAGAAATTCTATCAGTGGCACTGATGAGTTGGATATGCATGCTG

GGGTAGTTGATTCCGTCCGTTCC was the most significant with a p-value $1.46x10^{-10}$ and explained 50% phenotypic variation.

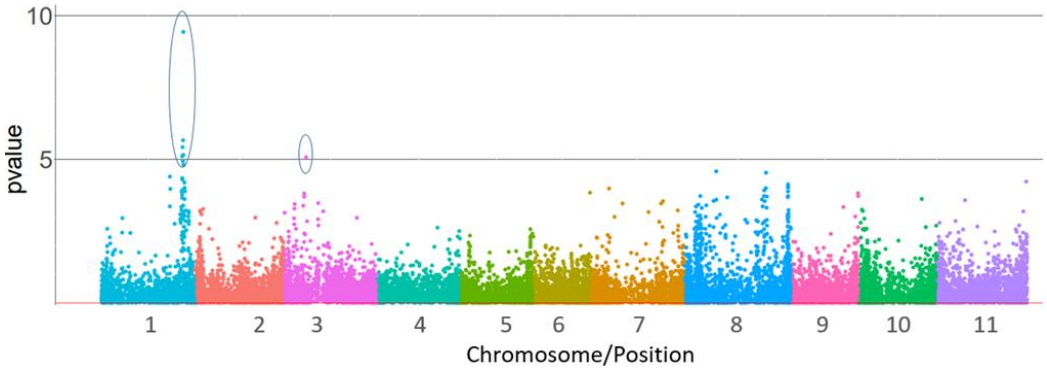


**Figure 5.1: Manhattan plot showing candidate SNPs and their P-values for days to flowering of common bean accessions grown in the year 2018 and 2019**

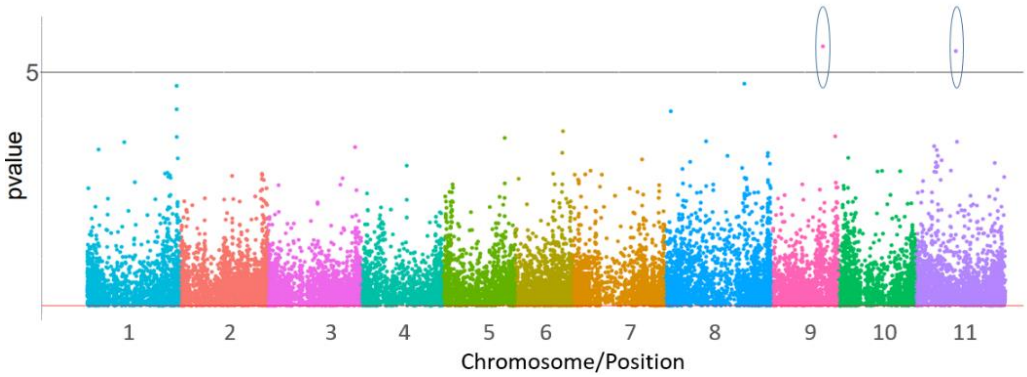


**Figure 5.2: Manhattan plot showing candidate SNPs and their P-values for days to maturity of common bean accessions grown in the year 2018 and 2019**

Other regions in the genome with SNPs marker significantly associated with days to flowering were chromosomes number three at positions 12798228 and 5975107 kb, chromosome eight at locations 19534456 and 48790667 kb, and lastly chromosome eleven in position 15684107 kb (Table 5.4).

**Table 5.4: Chromosome, chromosome position, P-values, minor allele frequency and proportion of phenotypic variation explained of the most significant Single Nucleotide Polymorphisms (SNPs) for days to flowering and maturity measured on 234 common bean accessions grown in the year 2018 and 2019**

| Trait | Season | SNP ID | Chr | SNP Position | P-value | MAF | $R^2$ |
|---|---|---|---|---|---|---|---|
| DF | Both | 8206455\|F\|0-42:G>A-42:G>A | 01 | 44775515 | $3.67x10^{-10}$ | 0.43 | 0.53 |
| DF | Both | 8215844\|F\|0-31:G>A-31:G>A | 01 | 44680698 | $2.17x10^{-06}$ | 0.42 | 0.48 |
| DF | Both | 8215758\|F\|0-28:T>C-28:T>C | 01 | 44386414 | $3.76x10^{-06}$ | 0.42 | 0.48 |
| DF | Both | 3378093\|F\|0-35:C>T-35:C>T | 01 | 44809453 | $7.00x10^{-06}$ | 0.42 | 0.48 |
| DF | Both | 3381041\|F\|0-55:T>C-55:T>C | 01 | 44146828 | $7.84x10^{-06}$ | 0.11 | 0.48 |
| DF | Both | 100095511\|F\|0-43:A>G-43:A>G | 03 | 12798228 | $8.47x10^{-06}$ | 0.06 | 0.48 |
| DF | LR | 8206455\|F\|0-42:G>A-42:G>A | 01 | 44775515 | $1.88x10^{-08}$ | 0.43 | 0.46 |
| DF | LR | 100112791\|F\|0-25:C>A-25:C>A | 08 | 48790667 | $4.93x10^{-06}$ | 0.31 | 0.42 |
| DF | LR | 100038445\|F\|0-43:G>C-43:G>C | 08 | 19534456 | $6.64x10^{-06}$ | 0.19 | 0.42 |
| DF | SR | 8206455\|F\|0-42:G>A-42:G>A | 01 | 44775515 | $1.46x10^{-10}$ | 0.43 | 0.50 |
| DF | SR | 100122138\|F\|0-19:A>G-19:A>G | 11 | 15324966 | $5.29x10^{-07}$ | 0.44 | 0.45 |
| DF | SR | 8215758\|F\|0-28:T>C-28:T>C | 01 | 44386414 | $5.67x\ 10^{-07}$ | 0.42 | 0.45 |
| DF | SR | 8215844\|F\|0-31:G>A-31:G>A | 01 | 44680698 | $1.40x10^{-06}$ | 0.42 | 0.44 |
| DF | SR | 3383101\|F\|0-8:C>A-8:C>A | 03 | 5975107 | $2.14x10^{-06}$ | 0.22 | 0.44 |
| DF | SR | 3378093\|F\|0-35:C>T-35:C>T | 01 | 44809453 | $2.63x10^{-06}$ | 0.42 | 0.44 |
| DF | SR | 3383103\|F\|0-22:C>G-22:C>G | 01 | 44604355 | $5.12x10^{-06}$ | 0.42 | 0.44 |
| DF | SR | 100095511\|F\|0-43:A>G-43:A>G | 03 | 12798228 | $5.71x10^{-06}$ | 0.06 | 0.44 |
| DF | SR | 100097512\|F\|0-55:T>C-55:T>C | 01 | 44960059 | $9.74x10^{-06}$ | 0.47 | 0.43 |
| DF | SR | 3377352\|F\|0-18:T>C-18:T>C | 01 | 44215472 | $1.05x10^{-05}$ | 0.41 | 0.43 |
| DM | Both | 100166226\|F\|0-13:G>A-13:G>A | 09 | 29625250 | $2.80x10^{-06}$ | 0.13 | 0.45 |
| DM | Both | 100157386\|F\|0-13:A>G-13:A>G | 11 | 27376763 | $3.53x10^{-06}$ | 0.34 | 0.44 |

SNP=Single Nucleotide Polymorphism, Chr=Chromosome, MAF=Minor allele frequency, $R^2$=Phenotypic variation explained by the SNP, DF=Days to flowering, DM=Days to maturity.

Association analysis using combined data for days to maturity from four seasons revealed two SNPs that were significantly associated with days to maturity (Figure 4.2). The two

SNPs were located at chromosomes nine (100166226|F|0-13:G>A-13:G>A) with a nucleotide sequence TGCAGTTAGCCCGGAGTGTGACAGATGACGGGGCAGTCT ACACGAGACGATCACGTAATGTTCATGGAT and eleven (100157386|F|0-13:A>G-13:A>G) with a nucleotide sequence TGCAGTATACATAACCAGTTACC GCCAAGATGAGTCATCGCCCAAGAGAAGGACATCGCCTAAGCAAGA and p-values of $2.80x10^{-06}$ and $3.53x10^{-06}$, respectively. The SNP markers on chromosomes nine and eleven accounted for 45% and 44% of the total phenotypic variation in days to maturity, respectively (Table 5.4). There were no SNPs identified to be significantly associated with days to maturity from the data recorded in separate seasons.

### 5.4.4 Number of pods per plant and pod length

One significant ($P \leq 0.05$) SNPs marker (100104488|F|0-33:C>T-33:C>T) with a nucleotide sequence TGCAGCAACAATAGCAACCACCACTTCATATGCCACCCC TGCTTCTAATATTAATATTAC on chromosome 11 at position 5531505bp was associated with the number of pods per plant, the marker had a p-value of $1.0x05^{-05}$ and accounted for about 54% of the total variability in the number of pods per plant during short rain seasons (Table 5.5 and Figure 5.3).

**Table 5.5: Chromosome, chromosome position, P-values, minor allele frequency and proportion of phenotypic variation explained of the most significant Single Nucleotide Polymorphisms (SNPs) for pod length and number of pods per plant measured on 234 common bean accessions grown in the year 2018 and 2019**

| Trait | Season | SNP ID | Chr | SNP Position | P-value | MAF | $R^2$ |
|---|---|---|---|---|---|---|---|
| PL | Both | 3383789\|F\|0-8:A>G-8:A>G | 02 | 47816492 | $1.29x10^{-07}$ | 0.27 | 0.50 |
| PL | Both | 3379907\|F\|0-17:G>T-17:G>T | 02 | 48172001 | $1.29x10^{-07}$ | 0.24 | 0.50 |
| PL | Both | 3374313\|F\|0-27:G>C-27:G>C | 08 | 60205731 | $6.03x10^{-07}$ | 0.38 | 0.49 |
| PL | Both | 100049685\|F\|0-55:T>A-55:T>A | 05 | 3857206 | $1.93x10^{-06}$ | 0.30 | 0.48 |
| PL | Both | 3368653\|F\|0-8:A>T-8:A>T | 02 | 47364334 | $2.96x10^{-06}$ | 0.25 | 0.48 |
| PL | Both | 3377440\|F\|0-53:A>G-53:A>G | 02 | 47505888 | $3.85x10^{-06}$ | 0.37 | 0.48 |
| PL | Both | 8668461\|F\|0-6:A>C-6:A>C | 08 | 59503639 | $1.00x10^{-05}$ | 0.24 | 0.48 |
| PL | LR | 3374313\|F\|0-27:G>C-27:G>C | 08 | 60205731 | $3.01x10^{-07}$ | 0.38 | 0.46 |
| PL | LR | 100088980\|F\|0-58:A>G-58:A>G | 08 | 59997265 | 4.76x10-$^{07}$ | 0.49 | 0.46 |
| PL | LR | 3383789\|F\|0-8:A>G-8:A>G | 02 | 47816492 | $5.96x10^{-07}$ | 0.27 | 0.46 |
| PL | LR | 3379907\|F\|0-17:G>T-17:G>T | 02 | 48172001 | $9.89x10^{-07}$ | 0.24 | 0.45 |
| PL | LR | 3377435\|F\|0-27:G>A-27:G>A | 03 | 38190877 | $1.95x10^{-06}$ | 0.30 | 0.45 |
| PL | LR | 100049685\|F\|0-55:T>A-55:T>A | 05 | 3857206 | $2.64x10^{-06}$ | 0.30 | 0.45 |
| PL | LR | 3368653\|F\|0-8:A>T-8:A>T | 02 | 47364334 | $3.11x10^{-06}$ | 0.25 | 0.45 |
| PL | LR | 3377440\|F\|0-53:A>G-53:A>G | 02 | 47505888 | $4.30x10^{-06}$ | 0.37 | 0.44 |
| PL | LR | 8212905\|F\|0-6:G>A-6:G>A | 02 | 48938155 | $5.49x10^{-06}$ | 0.38 | 0.44 |
| PL | LR | 3372743\|F\|0-25:G>A-25:G>A | 02 | 49029823 | $7.77x10^{-06}$ | 0.30 | 0.44 |
| PL | LR | 3370435\|F\|0-48:G>C-48:G>C | 04 | 44775515 | $9.50x10^{-06}$ | 0.33 | 0.44 |
| PL | SR | 3383789\|F\|0-8:A>G-8:A>G | 02 | 47816492 | $5.20x10^{-07}$ | 0.27 | 0.40 |
| PL | SR | 3379907\|F\|0-17:G>T-17:G>T | 02 | 48172001 | $1.58x10^{-06}$ | 0.24 | 0.39 |
| PL | SR | 100092861\|F\|0-24:G>A-24:G>A | 10 | 5758205 | $1.70x10^{-06}$ | 0.07 | 0.39 |
| PP | SR | 100104488\|F\|0-33:C>T-33:C>T | 11 | 5531505 | $1.00x10^{-05}$ | 0.09 | 0.54 |

SNP=Single Nucleotide Polymorphism, Chr=Chromosome, MAF=Minor allele frequency, $R^2$= Phenotypic variation explained by the SNP, PP=Number of pods per plant, PL=Pod length.

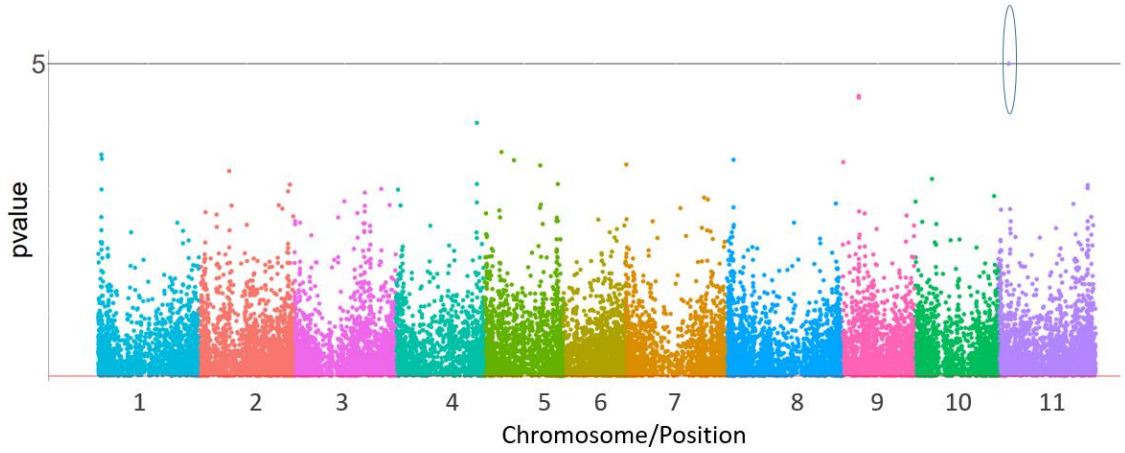


**Figure 5.3: Manhattan plot showing candidate SNPs and their P-values for number of pods per plant of common bean accessions grown during short rain seasons of the year 2018 and 2019**

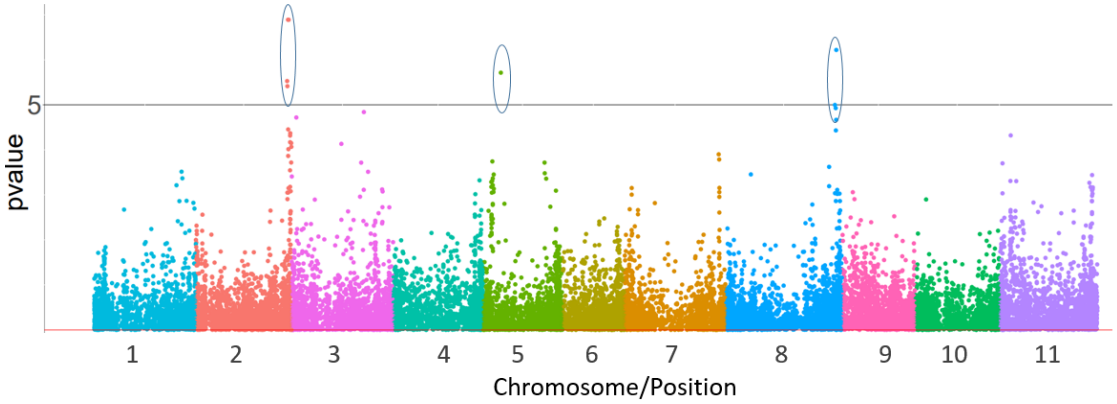


**Figure 5.4: Manhattan plot showing candidate SNPs and their P-values for pod length of common bean accessions grown in the year 2018 and 2019**

There were no significant (P≤0.05) SNP markers associated with the number of pods per plant identified during the long rain season or with the average data from both seasons. However, one significant SNP associated with number of pods per plant was identified during short rain seasons in chromosome eleven at position 5367046 with a P-value 1.0x05-05 and explained 54% of the total phenotypic variation.

Overall, thirteen significant (P≤0.05) SNPs associated with pod length were identified. Seven significant SNPs were identified using combined data of pod length for all seasons (Figure 5.4), eleven were identified during the long rain season, three during the short rain season, and one in both seasons (Table 5.5). Two SNPs 3383789|F|0-8:A>G-8:A>G with a nucleotide sequence of TGCAGGTTAGTGATCACAGAGGATTTTCCTGAGTGG GAGAATGAGTTGATGCCTTCAGGGAGAACGATG and 3379907|F|0-17:G>T-17:G>T (TGCAGACAGGAATACCAGTTTTCGTTCATCTCTGAACACAGATGCCC GAACATGTCAAGATGTAATAAC) on chromosome two were the most significant (P≤0.05) SNPs associated with pod length in both long rain and short rain seasons. These two markers were located at position 47816492 and 48172001 bp respectively, both with a p-value of $1.29x10^{-07}$ and explained 50% of the total variability. It was observed that the most significant SNPs markers were on chromosome two in the region around 47364334 to 49029823 bp. Other regions with significant SNPs included chromosome three at position 38190877 bp, chromosome four at position 46352340 bp, chromosome five at position 3857206 bp, chromosome eight at position 59503639 to 59997265 bp and 60205731 bp, and lastly chromosome ten in position 5758205 bp.

### 5.4.5 Number of seeds per pod, seed weight and grain yield

Two significant (P≤0.05) SNP markers associated with the number of seeds per pod were identified, one during short rain seasons and the other using the average data for both seasons (Figure 5.5). Both SNPs were on chromosome ten with the most significant SNP being 8669727|F|0-28:C>A-28:C>A with a nucleotide sequence TGCAGTGGGCAAG ACTCAAACCCAAGTCCTCGAGTGACCAAAGAAGCTTTAC at position 32425106 bp with a p-value of $2.04x10^{-06}$. The SNP explained 65% of the total phenotypic variation in the number of seeds per pod (Table 5.6). The second SNP significant (P≤0.05) SNP marker was 100112078|F|0-38:T>C-38:T>C with a sequence TGCAGTTGATTCTTG ATTGTGTGTGTTGCTTCCACGTGTGTTTTTTCCCATCGAGTGGTATCAAGAGC T and was located at 35888476 bp with a p-value of $2.04x10^{-06}$. When marker-trait association analysis was conducted using the average data for both seasons for this 100-

seed weight, one SNPs marker (3380104|F|0-61:T>A-61:T>A) was found to be significantly (P≤0.05) associated with 100-seed weight. This SNPs had a nucleotide sequence TGCAGCTATAGTCTCTGCAAATTTTGCCGGTAGAATTATATTACTG AGATAAAATTTCTATTGAAAAAA and was located on chromosome eleven at position 5893195bp with a p-value of $6.30x10^{-06}$ (Figure 6). This SNP explained 79% of phenotypic variation in seed weight (Table 5.6). There were no significant (P≤0.05) SNPs identified associated to 100-seed weight during long rain and short rain seasons.

**Table 5.6: Chromosome, chromosome position, P-values, minor allele frequency and proportion of phenotypic variation explained of the most significant Single Nucleotide Polymorphisms (SNPs) for number of seeds per pod, seed weight and yield measured on 234 common bean accessions grown in the year 2018 and 2019**

| Trait | Season | SNP ID | Chr | SNP Position | P-value | MAF | $R^2$ |
|---|---|---|---|---|---|---|---|
| SP | Both | 8669727\|F\|0-28:C>A-28:C>A | 10 | 32425106 | $2.04x10^{-06}$ | 0.09 | 0.65 |
| SP | SR | 100112078\|F\|0-38:T>C-38:T>C | 10 | 35888476 | $3.16x10^{-06}$ | 0.06 | 0.54 |
| SW | Both | 3380104\|F\|0-61:T>A-61:T>A | 11 | 5893195 | $6.30x10^{-06}$ | 0.09 | 0.79 |
| SY | Both | 3384204\|F\|0-17:T>C-17:T>C | 01 | 724994 | $8.34x10^{-06}$ | 0.11 | 0.28 |

SNP=Single Nucleotide Polymorphism, Chr=Chromosome, MAF=Minor allele frequency, $R^2$= Phenotypic variation explained by the SNP, SP=Seeds per pod, SW=Seed weight, SY=Seed yield.

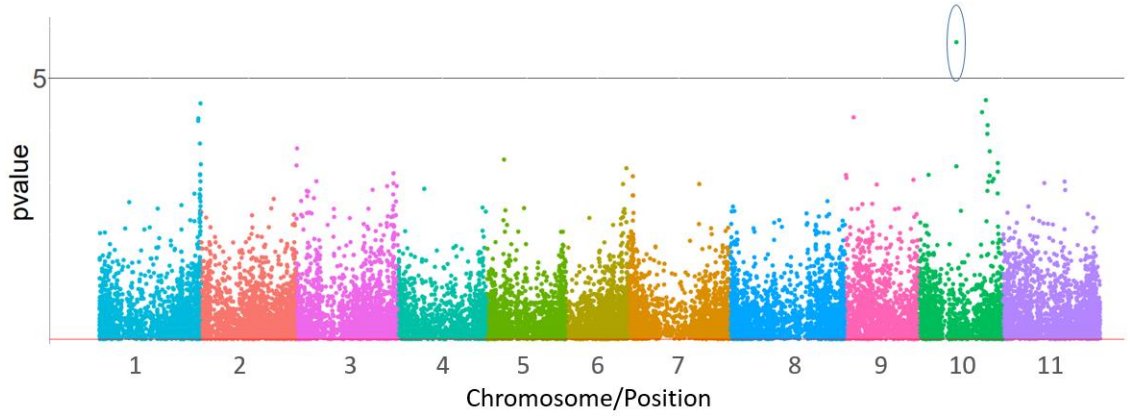


**Figure 5 5: Manhattan plot showing candidate SNPs and their P-values for number of seeds per pod of common bean accessions grown in the year 2018 and 2019**

SNP marker 3384204|F|0-17:T>C-17:T>C with nucleotide sequence TGCAGAGTGATGAAGTGTTGCCACGGTTTGATGGTGAGACGGAGGCTCATCTGGGGTCCTTCTCAACAG was significantly (P≤0.05) associated with grain yield per hectare on chromosome one at position 827668 bp and with a p-value of $8.34x10^{-06}$ (Figure 5.7). This marker was identified using combined data for grain yield of both seasons and accounted for 28% of phenotypic variation in grain yield (Table 5.6). No

significant (P≤0.05) SNP was identified during long rain and short rain seasons.

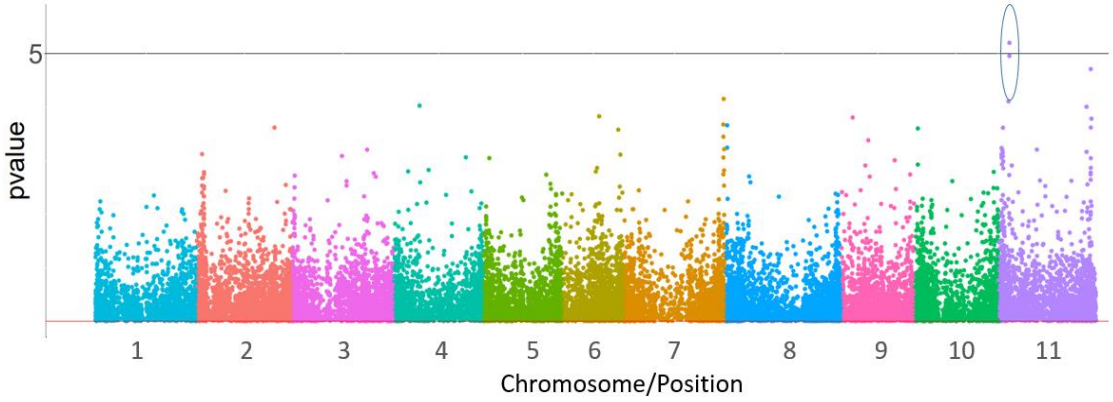


**Figure 5.6: Manhattan plot showing candidate SNPs and their P-values for number of 100-seed weight of common bean accessions grown in the year 2018 and 2019**

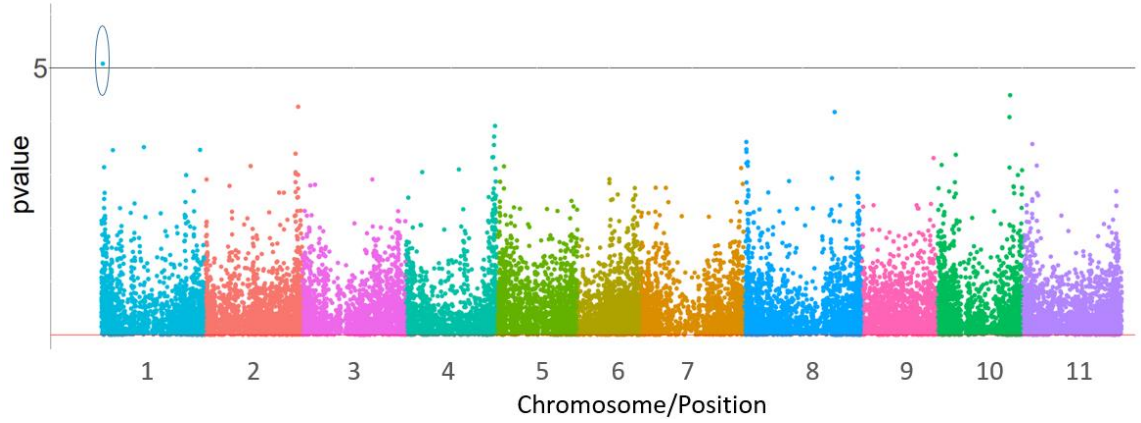


**Figure 5.7: Manhattan plot showing candidate SNPs and their P-values for grain yield of common bean accessions grown in the year 2018 and 2019**

### 5.5 Discussion

The results showed significant (P≤0.05) variability for various traits among the accessions in different seasons which is ideal for association mapping. Significant seasonal effects on traits could be attributed to changes in temperature and amount of rainfall that affected plant morphology and the expression of genes. The effect of seasonal variation on the expression of genetic loci was also revealed by trait-marker association analysis in that

the number of identified SNPs varied with trait and the seasons. The complex interactions of diverse environments and multiple genetic loci suggest an evaluation method that considers the interaction of genes and environment in analysis such as the trait-marker association studies. A GWAs approach involves rapidly scanning markers across the entire genome of many accessions to find genetic variations associated with phenological and yield component traits. This method allows simultaneous characterization of many accessions as well as examining thousands of genes under different environmental conditions to find a new genetic association with observable traits. Several studies have also reported the influence of the environment on the association of SNP marker to morphological traits (Kamfwa *et al.,* 2015) and cooking time (Berry *et al.,* 2020).

Optimal flowering time is an important agronomic trait that is a prerequisite to successful sexual reproduction and production of seed yield, it is an adaptive trait that is of interest to plant breeders (Muller, 2009). High-yielding varieties have been reported to flower early and mature late (Ashango and Alamerew, 2017), studies on flowering time for various plants have revealed that flowering time is controlled by many genes with diverse responses to seasonal cues (Andrés and Coupland, 2012). NCBI BLASTn search of the sequence of the most significant SNP marker (8206455|F|0-42:G>A-42:G>A) on chromosome one revealed a 100% sequence similarity match to sequence ID XM_007162819.1 (*PHAVU_001G188400g*).

QTL for flowering time on chromosome one had been reported in previous studies by Koinange *et al.,* (1996), Blair *et al.,* (2006), Perez-Vega *et al.,* (2010), Mukeshimana *et al.,* (2014), Kamfwa *et al.,* (2015) and (Langat *et al.,* 2019). Other QTLs have been reported on chromosome four by Mukeshimana *et al.,* (2014) and Langat *et al.,* 2019, on chromosome eight by Koinange *et al.,* (1996), Perez-Vega *et al.,* (2010), Kamfwa *et al.,* (2015), (Briñez *et al.,* 2017).

Duration to maturity is a critical trait for adaptation to geographical areas with varying seasonal duration and amounts of rainfall. Several common bean growing areas have irregular and marginal rainfall which makes them unreliable for production, this necessitates the use of early maturing and high-yielding common bean varieties. However,

earliness embodies an inherent loss of yield potential due to the short growth cycle and sub-optimal canopy (Mduruma *et al.,* 1998). Earliness has therefore been a breeding objective for many crops. In this study, two significant SNPs associated with duration to maturity were identified on chromosomes nine and eleven.

A BLASTn search of the nucleotide sequence of the two significant SNP markers 100166226|F|0-13:G>A-13:G>A on chromosome 9 and 100157386|F|0-13:A>G-13:A>G on chromosome 11 which were significantly ($P \leq 0.05$) associated to duration to maturity at NCBI database did not reveal any significant match similar to the nucleotide sequences. In previous studies, QTL for the duration to maturity has been reported on chromosome one (Langat *et al.,* 2019), chromosomes four (Mukeshimana *et al.,* 2014; Langat *et al.,* 2019), chromosome seven (Mukeshimana *et al.,* 2014), and chromosome nine (Mukeshimana *et al.,* 2014; Kamfwa *et al.,* 2015).

MADS-box loci also known as MICK-type genes have been reported to control various plant development processes like vegetative growth and reproductive organ development (Adamzyk and Fernandez, 2009). Phytochrome-interacting factors are basic helix-loop-helix transcription factors that play critical roles in the germination of seeds, photomorphogenesis, responses to shading, flowering time, and leaf senescence (Sakuraba *et al.,* 2014; Shi *et al.,* 2018). However, in this study, the MADS-box and phytochrome loci were not identified close to the position of the most significant SNPs associated with duration to flowering and maturity on the Phytozome genetic map.

The number of pods per plant is a primary yield component and part of the accumulated aerial biomass that is later partitioned to seed yield (Negahi *et al.,* 2014). A BLASTn search of the sequence of the identified significant SNP marker (100104488|F|0-33:C>T-33:C>T) associated with the number of pods per plant at NCBI revealed a sequence ID XM_007131979.1 (*PHAVU_011G061900g*) with a 100% match.

QTLs for the number of pods per plant have been identified in previous studies using different methods and populations. QTL for the number of pods per plant has been reported on chromosome one (Koinange *et al.,* 1996, Langat *et al.,* 2019), on chromosome two (Beattie, 2003; Mukeshimana, *et al.,* 2014; Langat *et al.,* 2019), on chromosome three

(Beattie, 2003; Langat *et al.,* 2019), on chromosome 5 (Beattie, 2003; Kamfwa *et al.,* 2015), on chromosome seven (Blair, 2006; Kamfwa *et al.*, 2015; Briñez *et al.,* 2017), on chromosome eight (Mukeshimana, *et al.*, 2014) and chromosome nine and eleven (Blair, 2006).

Pod length is a measure of the size of the harvested part of a snap bean, it is one of the quality traits that appeal to the consumer in green beans (Wahome *et al.,* 2013). Pod length is negatively associated with the number of pods per plant and is positively correlated with seed weight (Ghobary and Allah, 2010). A BLASTn search of the nucleotide sequence for the most significant (P≤0.05) SNP marker 3383789|F|0-8:A>G-8:A>G associated to pod length at NCBI revealed sequence ID XM_007160253.1 (*PHAVU_002G311300g*) with a 100% match. When the second most significant SNP (3379907|F|0-17:G>T-17:G>T ) sequence on the same chromosome two was searched it revealed sequence ID XM_007160298.1 (*PHAVU_002G315000g*) with also a 100% match.

The number of seeds per pod is a primary yield component for seed yield (Ghobary and Allah, 2010; Negahi *et al.,* 2014). Higher-yielding varieties have been reported to have more seeds per pod (Ashango and Alamerew, 2017). A BLASTn search of the two SNP markers 8669727|F|0-28:C>A-28:C>A and 100112078|F|0-38:T>C-38:T>C significantly associated with the number of seeds per pod at the NCBI database did not identify significant (P≤0.05) similarity with any known sequence. QTLs for the number of seeds per pod have been reported in previous studies on chromosomes two (Langat *et al.*, 2019) on chromosomes 6 and 7 (Briñez *et al.*, 2017) and on chromosome 8 (Briñez *et al.*, 2017; Langat *et al.*, 2019).

Seed weight is also a primary yield component that exhibits higher heritability than the number of pods per plant and the number of seeds per pod (Vandemark *et al.*, 2014). Large-seeded (Andean bean type) has been reported as the most popular in Africa compared to the small-seeded type (Beebe, 2012). A BLASTn of the SNP marker 3380104|F|0-61:T>A-61:T>A significantly associated with seed weight at NCBI revealed sequence ID CP039347.1, XM_028053922.1, XM_028053921.1, XM_028053920.1 and XM_028053918.1 of cowpea with 92.75% match. Several QTLs for seed weight have

been mapped in previous studies on chromosome one (Koinange *et al.,* 1996; Broughton *et al.,* 2003; Briñez *et al.,* 2017), chromosome five (Blair *et al.,* 2012; Briñez *et al.,* 2017), on chromosome seven (Koinange *et al.,* 1996; Mukeshimana *et al.,* 2014), on chromosome eight (Langat *et al.,* 2019) and chromosome eleven (Koinange *et al.,* 1996).

Seed yield is a polygenic trait determined primarily by the number of pods per plant, number of seeds per pod, and seed weight (Negahi *et al.,* 2014). Identifying QTLs that control the three yield components and their interaction with seed yield would assist breeding efforts aimed at improving the seed yield of common bean (Burbano-Erazo *et al.,* 2021). In this study SNPs marker significantly associated with grain yield was identified on chromosome one, around this SNP twelve positional candidate genes were revealed in the phytozome database. A BLASTn search of the SNP marker 3384204|F|0-17:T>C-17:T>C sequence which was significantly associated with seed yield at NCBI revealed a sequence ID XM_007160638.1 encoding for a hypothetical protein *PHAVU_001G009700g* in common bean with 100% match.

QTLs for seed yield had been reported in various studies using different methods and populations, QTL for seed yield has been identified on chromosome one (Briñez *et al.,* 2017; Resende *et al.,* 2018; Langat *et al.,* 2019), on chromosome two (Langat *et al.,* 2019), on chromosome three (Mukeshimana *et al.,* 2014, Kamfwa *et al.,* 2015; Sandhu *et al.,* 2018; Resende *et al.,* 2018), on chromosome four (Resende *et al.,* 2018), on chromosome seven (Sandhu *et al.,* 2018), on chromosome eight (Sandhu *et al.,* 2018; Resende *et al.,* 2018), on chromosome nine (Mukeshimana *et al.,* 2014, Kamfwa *et al.,* 2015) and chromosome eleven (Briñez *et al.,* 2017).

### 5.6 Conclusion

The study estimated the position and effects of genomic regions associated with various agronomic traits evaluated in different seasons. The association between trait and marker was influenced by the seasonal changes which validate the effectiveness of GWAs in identifying QTLs under different environmental conditions. A total of 33 SNPs were significantly ($P \leq 0.05$) associated with various agronomic traits were identified, and 296 potential candidate genes for various traits were identified around 100kb the position of

the trait-associated SNPs. These potential candidate genes require further investigations to ascertain whether they are in linkage disequilibrium with the loci or play a role in the expression of the trait. The findings expand the current knowledge on genomic regions known to control various traits in common beans, and the identified SNPs markers can be used in developing varieties with desirable traits through markers-assisted selection (MAS).

# CHAPTER SIX

# QUANTITATIVE TRAIT LOCI FOR COOKING TIME AND MORPHOLOGICAL TRAITS OF COMMON BEAN (*PHASEOLUS VULGARIS* L.) USING RECOMBINANT INBRED LINES AND DIVERSITY ARRAYS TECHNOLOGY MARKERS

## 6.1 Abstract

Common bean is a critical source of nourishment globally, second only to cereals. Common bean seeds stored in adverse storage conditions develop a hard-to-cook trait that prolongs cooking time leading to a high cost of preparation. The aim of this study was to map Quantitative Trait Loci (QTLs) associated with the hard-to-cook trait and morphological traits of agronomic importance using $F_{2:6}$ recombinant inbred lines (RILs) of common beans derived from two biparental crosses. Two common bean varieties GLP2 (easy-to-cook), GLPx92 (hard-to-cook) sourced from Kenya Seed Company, and accession GBK035420 (easy-to-cook) from the National Gene Bank of Kenya were used in this study to develop $F_{2:6}$ recombinant inbred lines. The populations were advanced using the single seed descent method in a greenhouse and later multiplied in the field at the Jomo Kenyatta University of Agriculture and Technology (JKUAT) farm. The RILs were genotyped using Diversity Arrays Technology Sequencing (DArTseqTM) to identify variation in single nucleotide polymorphism markers (SNP). Seed samples of RILs were stored at a temperature of 35$^{o}$C and 50% RH for four months and later used for a cooking experiment to determine their cooking time using a subjective finger-pressing method. A total of 35246 SNP markers were scored, out of which 11277 markers were found polymorphic. A linkage genetic map was constructed with a distance and marker density that ranged from 2513 to 3352 cM and 0.5 to 3.5, respectively. This study identified QTLs responsible for cooking time and some morphological traits. QTLs influencing cooking duration were detected on chromosome 1, 2, 3, 5, 6, 9, 10, and 11. QTLs on chromosomes 3 and 10 had the highest additive effect of 27.2 towards longer cooking time and both explained 37.7% of the phenotypic variance. The study identified genes known to control

polygalacturonase/pectinase, pectin methylesterase, pectinesterase inhibitor, and galacturan 1, 4 alpha-galacturonidase enzymes to co-localize with the QTLs detected.

## 6.2 Introduction

The common bean crop is an important contributor to food security worldwide, it is a good source of nourishment second only to cereals. Common bean seeds increase the protein content of a meal, improving the quality of the diet by a factor of 50% to 70% when served with cereals (Bressani *et al.,* 1988; Taptue, 2018). Common bean is also a source of vitamin C, vitamin B components such as thiamine, folate, and niacin, and essential minerals such as iron, zinc, magnesium, and potassium (Mitova *et al.,* 2008; Mitchell *et al.,* 2009).

Cooking beans is a fundamental practice of bean preparation, as it enhances their digestibility, nutrient availability, and taste while reducing the levels of antinutrients that could interfere with nutrient absorption (Costa *et al.,* 2006). Cooking beans is a high-energy demanding process due to their prolonged cooking time. The cooking time is influenced by genetic differences, growth environment, post-harvest handling, storage conditions such as temperature and humidity, storage time, and treatments before cooking (Arruda *et al.,* 2012; Kinyanjui *et al.,* 2015). Common bean seeds that have been stored for a while develop hard-to-cook (HTC) characteristics, which affect cooking time and digestibility (Kinyanjui *et al.,* 2015). It is proposed that storage of common beans at high temperatures and relative humidity leads to the formation of insoluble pectin in the cell wall and middle lamella that cements cells together rendering the cotyledon cells fail to separate easily upon cooking (Jones and Boulter, 1983a; Kinyanjui *et al.,* 2015).

DNA analysis techniques in common bean breeding programs have improved our understanding of genetic factors controlling various traits. The discovery of next-generation sequencing (NGS) Single Nucleotide Polymorphism (SNPs) has become more practical in genotyping and the discovery of molecular markers (Gujaria-Verma *et al.,* 2016). The identified single nucleotide polymorphism (SNPs) markers for common bean allow explorations of genetic diversity and population structure in common beans (Cichy *et al.,* 2015; Valdisser *et al.,* 2016).

Several studies have mapped QTLs that control cooking time. A random amplified polymorphic DNA (RAPD) marker associated with cooking time was identified using 104 recombinant inbred lines, the marker explained 23% of the variation in cooking time (Jancinto-Hernandez *et al.*, 2003). Garcia *et al.*, (2012) mapped 6 QTLs that govern cooking time on chromosomes 1 and 9 using 105 polymorphic SSR markers and 140 $F_{2:4}$ recombinant inbred lines. A study conducted by Berry *et al.*, (2020) using 146 recombinant inbred lines of common bean, 10 QTLs on chromosomes 1, 2, 3, 5, 6, 10, and 11 were identified, with the most robust QTLs being on chromosome 3, 6, 10 and 11 that appeared in over two different environments. In a recent study, QTLs for cooking time were identified on Pv8 and Pv10 using 242 recombinant inbred lines population developed from a cross between Ervilha (Manteca) and PI527538 (Njano) (Bassett *et al.*, 2021). This study aimed at mapping Quantitative Trait Loci (QTLs) associated with the hard-to-cook trait using $F_{2:6}$ recombinant inbred lines of common bean derived from two biparental crosses.

## 6.3 Materials and Methods

### 6.3.1 Plant material and field multiplication

Two common bean varieties GLP2 (Rosecoco type), GLPx92 (Pinto type) sourced from Kenya Seed Company, and GBK035420 (Black type) from the National Gene Bank of Kenya were used to develop $F_{2:6}$ RILs (Table 3.1). GLPx92 contributed the hard-to-cook (HTC) characteristic, while GLP2 and GBK035420 provided the easy-to-cook (ETC) trait. HTC parent (pinto) was intercrossed with ETC parents (rosecoco and black) to generate an $F_1$ populations which were self-pollinated to produce $F_2$ population. A random sample of 300 $F_2$ segregating populations were advanced using Single Seed Descent (SSD) method in a greenhouse to produce $F_{2:6}$ Recombinants Inbred Lines (RILs) for each cross (Figure 6.1, 6.2 and 6.3).

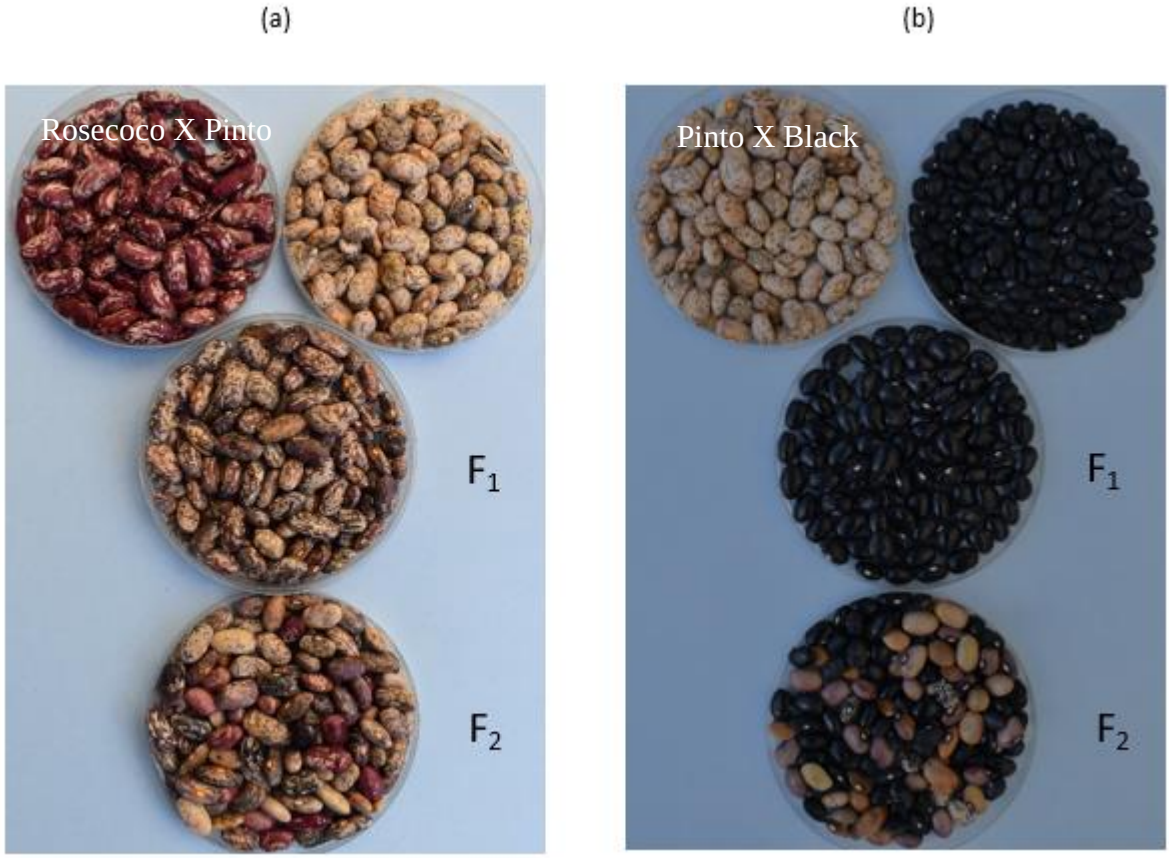


**Figure 6.1: (a) GLP2 (Rosecoco) and GLPx92 (Pinto) their $F_1$ and $F_2$'s seeds, (b) GLPx92 (pinto) and GBK035420 (black) and their $F_1$'s and $F_2$ seeds**


**Figure 6.2: Development of RILs in a greenhouse at Jomo Kenyatta University of Technology**


**Figure 6.3: Seeds from various single plants at $F_3$ generation from a cross between (a) GLP2 (Rosecoco) X GLPx92 (Pinto) and (b) between GLPx92 X GBK035420 (Black)**

At $F_6$ generation the lines were multiplied in the field at the Jomo Kenyatta University of Agriculture and Technology (JKUAT) farm. Morphological traits recorded included days to flowering, days to maturity, number of pods per plant, pod length, number of seeds per pod, and seed weight during field multiplication as described in section 3.3.4.

### 6.3.2 Incubation of seed and determination of cooking time

The storage and cooking experiment were conducted as described in section 4.3.2; and cooking time was defined as time taken for 95% of a 100 seed sample to cook.

### 6.3.3 DNA extraction and genotyping

DNA isolation and genotyping were conducted as described in section 4.3.3.

### 6.3.4 Data analysis

The relationship between cooking time and the percentage of cooked bean seeds was demonstrated with Logistic regression modelling as described by Wafula *et al.,* (2020). Cooking time was defined as the time corresponding to the probability of having 95% of the bean seeds cooked (Table 6.1). Analysis of variance was performed on the obtained cooking time using R software (version 4.0.2) at a 95% confidence interval.

Linkage mapping and Quantitative Trait Loci (QTL) analysis were conducted using QTL IciMapping software version 4.0.6.0 in its default settings. SNPs markers with minor allele frequency (MAF) < 0.05 and integrity < 0.2 were filtered out from QTLs mapping analysis. The remaining SNPs markers were grouped based on the LOD score of 3.0. Most SNPs markers were assigned their respective chromosomes by anchoring based on the *Phaseolus vulgaris* L. genetic map (*Phaseolus vulgaris* 442 version 2.1) (Schmutz *et al.*, 2014; Goodstein *et al.*, 2012). The loci order of each linkage group was established using the nnTwoOpt algorithm. Rippling criteria to fine-tune the order of the chromosomes was performed using the sum of Adjacent Recombination Frequencies (SARF) (Meng *et al.*, 2015). Inclusive composite interval mapping for the Additive and Dominance QTL (ICIM-ADD) method was used in QTL analysis with map distances in (centiMorgans, cM) calculated using the kosambi mapping function. A walking speed of 1.0 cM was used to determine the experiment-wise threshold at P≤0.05 and a LOD score threshold of 2.5 (Meng *et al.*, 2015).

The sequences of flanking SNPs markers of each QTL were used to identify potential candidate genes using the BLASTn search tool in the phytozome database (www.phytozome.net) to reveal the position of QTL in the physical map of *Phaseolus vulgaris* v2.1. The genes within the QTL region were selected based on their function in the formation of pectin in the cell wall.

## 6.4 Results

### 6.4.1 Phenotypic data

There were significant (P≤0.05) differences among RILs for all traits measured (Table 6.1, and appendix 3). Results revealed a significant (P≤0.05) moderate and positive correlation between days to flowering and days to maturity (0.5) and between pod length and seed weight (0.47) (Table 6.2).

**Table 6.1: Descriptive statistics for the parents and RILs for various phenotypic traits**

| | Parent | | | Pinto x Rosecoco RILs | | | | Pinto x Black RILs | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Trait | Pinto | Rosecoco | Black | Mean | P-value | Min | Max | Mean | P-value | Min | Max |
| CT (min) | 58.0 | 50.3 | 43.9 | 49.4 | ** | 32.0 | 103.2 | 52.9 | ** | 30.7 | 90.0 |
| DF | 40.0 | 42.0 | 45.0 | 41.4 | ** | 34.0 | 51.0 | 40.0 | ** | 31.0 | 54.0 |
| DM | 92.0 | 84.5 | 90.0 | 87.4 | ** | 78.0 | 100.0 | 87.4 | ** | 78.0 | 103.0 |
| PP | 11.0 | 6.2 | 12.8 | 14.5 | * | 3.3 | 41.7 | 9.5 | ** | 2.5 | 28.7 |
| PL | 7.4 | 12.8 | 7.4 | 8.3 | ** | 6.4 | 10.3 | 9.6 | * | 5.7 | 15.5 |
| SP | 3.7 | 4.0 | 5.5 | 4.9 | * | 3.0 | 6.7 | 4.3 | ** | 2.3 | 8.5 |
| SW | 41.5 | 56.5 | 19.0 | 25.5 | ** | 8.7 | 54.0 | 36.1 | ** | 19.0 | 65.0 |

*, **= significant at P≤0.05 and P≤0.01 probability level, CT=Cooking time, DF=Days to flowering, Days to maturity, PL=Pod length, SP=Seeds per pod, SW=Seed weight.

The results also revealed significant (P≤0.05) weak positive associations between cooking time and days to flowering (0.16), days to flowering and number of seeds per pod (0.2), days to maturity and pod length (0.12), days to maturity and number of seeds per pod (0.11), number of pods per plant and number of seeds per pod (0.21) and between pod length and number of seed per pod (0.27) (Table 6.2). A significant (P≤0.05) weak negative relationship was observed between cooking time and number of pods per plant (-0.13), number of pods per plant and pod length (-0.19), number of pods per plant and seed weight (-0.27), and between the number of seeds pod and seed weight (-0.2) (Table 6.2).

**Table 6.2: Pearson phenotypic correlation coefficient between various traits of RILs**

| | CT | DF | DM | PP | PL | SP | SW |
|---|---|---|---|---|---|---|---|
| CT (min) | 1.00 | | | | | | |
| DF (days) | 0.16** | 1.00 | | | | | |
| DM (days) | 0.08 | 0.52** | 1.00 | | | | |
| PP (no.) | -0.13* | 0.04 | -0.04 | 1.00 | | | |
| PL (cm) | 0.04 | 0.05 | 0.12* | -0.19** | 1.00 | | |
| SP (no.) | -0.08 | 0.20* | 0.11* | 0.21** | 0.27* | 1.00 | |
| SW (gms) | 0.05 | -0.04 | 0.24 | -0.27** | 0.47** | -0.20** | 1.00 |

*, **= significant at P≤0.05 and P≤0.01 probability level, CT=Cooking time, DF=Days to flowering, Days to maturity, PL=Pod length, SP=Seeds per pod, SW=Seed weight

### 6.4.2 QTL mapping

A total of 11277 markers were found to be polymorphic out of the 35246 markers scored, the number of polymorphic markers for each population are presented in Table 6.3. The polymorphic markers were used to construct a linkage genetic map for the four populations with a distance that ranged from 2513 to 3352 cM and marker density that ranged from 0.5 to 3.5 (Table 6.4). Various QTLs were identified for days to flowering, days to maturity, number of pods per plant, pod length, number of seeds per pod, seed weight, and cooking time (Tables 6.5 and 6.6).

**Table 6.3: Population size and number of SNP markers sequenced**

| Cross | Population size | SNP Markers sequenced | Polymorphic SNP markers |
|---|---|---|---|
| PintoXRosecoco | 137 | 9127 | 3507 |
| RosecocoXPinto | 129 | 8913 | 4156 |
| BlackxPinto | 79 | 7102 | 1329 |
| PintoxBlack | 131 | 10104 | 2285 |

SNP=Single nucleotide polymorphism

**Table 6.4: Linkage map of $F_{2:6}$ RILs of common bean**

| Chr | Pinto x Rosecoco | | | Rosecoco x Pinto | | | Black x Pinto | | | Pinto x Black | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | No. of markers | Chr size (cM) | Marker density | No. of markers | Chr size (cM) | Marker density | No. of markers | Chr size (cM) | Marker density | No. of markers | Chr size (cM) | Marker density |
| 1 | 371.0 | 256.4 | 0.7 | 399.0 | 298.1 | 0.7 | 134.0 | 175.1 | 1.3 | 179.0 | 312.6 | 1.7 |
| 2 | 401.0 | 321.4 | 0.8 | 503.0 | 429.0 | 0.9 | 152.0 | 330.0 | 2.2 | 272.0 | 233.1 | 0.9 |
| 3 | 404.0 | 417.5 | 1.0 | 467.0 | 368.0 | 0.8 | 161.0 | 285.7 | 1.8 | 275.0 | 273.3 | 1.0 |
| 4 | 176.0 | 338.1 | 1.9 | 223.0 | 424.0 | 1.9 | 55.0 | 192.4 | 3.5 | 91.0 | 151.2 | 1.7 |
| 5 | 243.0 | 193.2 | 0.8 | 311.0 | 187.2 | 0.6 | 86.0 | 226.8 | 2.6 | 170.0 | 223.3 | 1.3 |
| 6 | 338.0 | 192.9 | 0.6 | 307.0 | 157.2 | 0.5 | 104.0 | 180.4 | 1.7 | 174.0 | 185.8 | 1.1 |
| 7 | 333.0 | 325.2 | 1.0 | 423.0 | 321.6 | 0.8 | 101.0 | 268.5 | 2.7 | 204.0 | 241.4 | 1.2 |
| 8 | 412.0 | 334.5 | 0.8 | 452.0 | 283.2 | 0.6 | 160.0 | 300.3 | 1.9 | 296.0 | 286.6 | 1.0 |
| 9 | 356.0 | 338.1 | 0.9 | 423.0 | 203.1 | 0.5 | 90.0 | 164.4 | 1.8 | 165.0 | 331.9 | 2.0 |
| 10 | 234.0 | 221.8 | 0.9 | 246.0 | 294.3 | 1.2 | 154.0 | 127.0 | 0.8 | 233.0 | 220.8 | 0.9 |
| 11 | 239.0 | 304.4 | 1.3 | 402.0 | 386.2 | 1.0 | 131.0 | 262.4 | 2.0 | 226.0 | 369.6 | 1.6 |

Chr=Chromosome

### 6.4.3 Cooking time

From the two crosses, QTLs affecting cooking time were detected on chromosomes 1, 2, 3, 5, 6, 9, 10, and 11. Three QTLs were on chromosome 1, and two QTLs were detected on chromosomes 3 and 10, while chromosomes 2, 5, 6, 9, and 11 had one QTL each affecting cooking time (Figures 6.3 and 6.4).

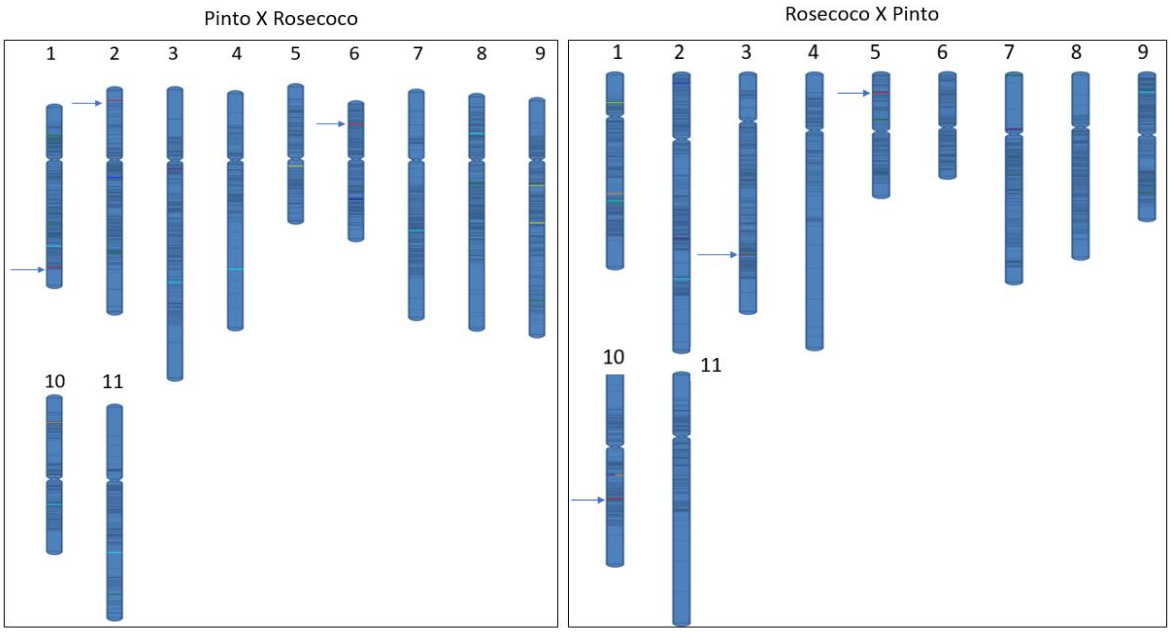


**Figure 6.4: Common bean linkage map constructed using Pinto x Rosecoco $F_{2:6}$ RILs, the arrows point to locations of QTL for cooking time on chromosome**

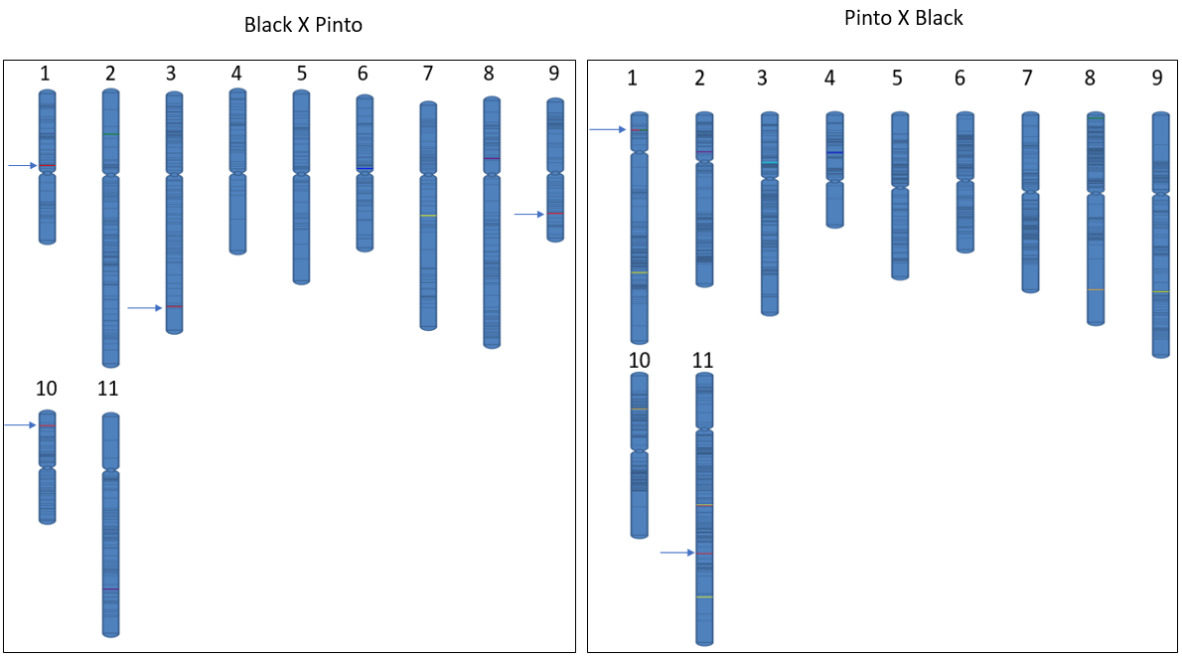


**Figure 6.5: Common bean linkage map constructed using Pinto x Black $F_{2:6}$ RILs, the arrows point to locations of QTL for cooking time on chromosome**

The additive effect of the QTLs detected ranged from -4.2 to 27.2. QTLs that contributed towards shorter cooking time were only detected in the cross of PintoXRosecoco. QTL on chromosome eight that contributed most towards shorter cooking time had an additive effect of -4.2, a LOD score of 6, and explained 16.8% of the phenotypic variance for cooking time. QTLs with the highest additive effect contributed to the hard-to-cook trait. QTLs on chromosomes three and ten had the highest additive effect towards longer cooking time with a LOD score of 5.0 and 3.8, respectively, and both explained 37.7% phenotypic variance for cooking time (Tables 6.5).

**Table 6.5: Significant QTLs for cooking time detected using inclusive composite interval mapping for Pinto (GLPx92) X Rosecoco (GLP2) and Pinto (GLPx92) X Black (GBK035420) $F_{2:6}$ RILs**

| Cross | Type | Chr | Position | Left marker | Right marker | LOD score | PVE (%) | Add |
|---|---|---|---|---|---|---|---|---|
| PintoXRosecoco | R | 1 | 230 | 3369772\|F\|0-15:G>A-15:G>A | 3380651\|F\|0-45:G>C-45:G>C | 3.2 | 8.0 | -3.0 |
| PintoXRosecoco | R | 2 | 16 | 3380232\|F\|0-6:T>C-6:T>C | 13121299\|F\|0-23:C>G-23:C>G | 3.3 | 21.4 | 10.0 |
| PintoXRosecoco | R | 6 | 30 | 3381145\|F\|0-24:T>A-24:T>A | 3378882\|F\|0-16:T>C-16:T>C | 3.9 | 9.4 | 3.2 |
| RosecocoXPinto | R | 3 | 277 | 3373798\|F\|0-56:C>G-56:C>G | 8198235\|F\|0-24:T>C-24:T>C | 6.0 | 13.8 | 3.8 |
| RosecocoXPinto | R | 5 | 29 | 3377170\|F\|0-7:C>T-7:C>T | 3365727\|F\|0-33:T>C-33:T>C | 3.8 | 17.5 | 4.4 |
| RosecocoXPinto | R | 10 | 189 | 3365755\|F\|0-41:C>A-41:C>A | 3377779\|F\|0-55:G>C-55:G>C | 6.5 | 14.1 | 4.1 |
| BlackXPinto | R | 1 | 92 | 3379442\|F\|0-64:A>G-64:A>G | 8202872\|F\|0-22:G>A-22:G>A | 3.9 | 38.6 | 19.6 |
| BlackXPinto | R | 3 | 256 | 3377474\|F\|0-24:C>T-24:C>T | 13122417\|F\|0-6:T>A-6:T>A | 5.0 | 37.7 | 27.2 |
| BlackXPinto | R | 9 | 133 | 3382186\|F\|0-32:G>A-32:G>A | 3377625\|F\|0-15:T>C-15:T>C | 5.7 | 38.8 | 17.8 |
| PintoXBlack | R | 1 | 21 | 8214403\|F\|0-30:C>T-30:C>T | 8198338\|F\|0-64:A>T-64:A>T | 4.0 | 28.3 | 12.0 |
| PintoXBlack | R | 11 | 243 | 100052374\|F\|0-6:T>C-6:T>C | 100052136\|F\|0-16:C>T-16:C>T | 4.4 | 33.0 | 6.9 |
| BlackXPinto | R | 10 | 15 | 3382815\|F\|0-17:G>A-17:G>A | 3383861\|F\|0-67:G>A-67:G>A | 3.8 | 37.7 | 27.1 |
| PintoXRosecoco | C | 2 | 218 | 3383226\|F\|0-59:T>C-59:T>C | 3378069\|F\|0-44:T>G-44:T>G | 2.7 | 3.0 | -1.8 |
| PintoXRosecoco | C | 3 | 236 | 3382182\|F\|0-44:T>C-44:T>C | 3382093\|F\|0-58:T>A-58:T>A | 3.0 | 3.4 | 1.9 |
| PintoXRosecoco | C | 5 | 125 | 3381095\|F\|0-11:A>G-11:A>G | 3377582\|F\|0-21:A>G-21:A>G | 5.8 | 14.2 | -3.9 |
| PintoXRosecoco | C | 8 | 43 | 3378448\|F\|0-17:A>T-17:A>T | 3381609\|F\|0-25:C>T-25:C>T | 6.0 | 16.8 | -4.2 |
| PintoXBlack | C | 1 | 126 | 8669561\|F\|0-16:C>T-16:C>T | 3379475\|F\|0-52:T>C-52:T>C | 4.6 | 8.9 | 6.2 |
| PintoXBlack | C | 8 | 87 | 3380849\|F\|0-46:A>G-46:A>G | 3384254\|F\|0-57:A>T-57:A>T | 3.1 | 5.6 | 2.3 |
| PintoXBlack | C | 9 | 18 | 3381479\|F\|0-62:C>T-62:C>T | 8200928\|F\|0-43:C>A-43:C>A | 3.1 | 5.8 | 2.4 |
| PintoXBlack | C | 9 | 35 | 8215841\|F\|0-30:A>G-30:A>G | 3382186\|F\|0-32:G>A-32:G>A | 3.3 | 26.2 | 12.1 |
| PintoXBlack | C | 9 | 39 | 3382186\|F\|0-32:G>A-32:G>A | 3377625\|F\|0-15:T>C-15:T>C | 4.1 | 26.2 | 12.3 |

Chr=Chromosome, PVE=Phenotypic variance explained, Add=Additive effect, LOD=Log of odds, CT=Cooking time, R= RILs from population of individual reciprocal cross, C=RILs where reciprocals were combined.

### 6.4.4 Days to flowering and maturity

In total, seventeen QTLs affecting days to flowering were detected from the two crosses, and most of them (15) were detected in RILs from the cross-involving pinto and rosecoco. QTLs for days to flowering were on chromosomes 1, 2, 3, 5, 7, 8, 9, and 11, with chromosome one having the highest number (6) of QTLs for days to flowering. QTLs detected had a LOD score ranging from 2.9 to 30.6 and explained phenotypic variance in the range of 4.3 to 40.5, the additive effect of these QTLs ranged from -2.3 to 4.2. Two QTLs with the highest contribution towards early and late flowering were on chromosome one. QTL with an additive effect (-2.3) towards early flowering had a LOD score of 18.0 and explained 37.5% of the phenotypic variance. On the other hand, QTLs with the highest positive additive effect had a LOD score of 6.7 and explained 24.4% of the phenotypic variance (Tables 6.6). The region with QTL with the highest additive effect on late-flowering was also detected in the genome-wide association study (GWAS) in section 4.4.5.

We detected eleven QTLs that influenced duration to maturity, the majority (7) of them were in RILs derived from the cross of pinto x rosecoco. The QTLs were detected on chromosomes 1, 2, 3, 4, 7, 8, 10, and 11. Seven of these QTLs contributed toward early maturity with an additive effect that ranged from -1.1 to -2.2. The QTL that contributed most to early maturing (-2.2) was found on chromosome one with a LOD score of 12.1 and explained 17.6% of the phenotypic variance. QTLs with the highest (3.3) positive additive effect was detected on chromosome nine with a LOD score of 3.3 and explained 15.3% of the phenotypic variance (Table 6.6).

**Table 6.6: Significant QTLs for days to flowering and days to flowering detected using inclusive composite interval mapping for Pinto (GLPx92) X Rosecoco (GLP2) and Pinto (GLPx92) X Black (GBK035420) $F_{2:6}$ RILs**

| Cross | Trait | Chr | Position | Left marker | Right marker | LOD score | PVE (%) | Add |
|---|---|---|---|---|---|---|---|---|
| PintoXRosecoco | DF | 1 | 42 | 8214729\|F\|0-29:A>G-29:A>G | 100042247\|F\|0-52:G>T-52:G>T | 9.6 | 8.2 | 1.2 |
| PintoXRosecoco | DF | 1 | 48 | 8208744\|F\|0-19:C>T-19:C>T | 3380734\|F\|0-38:G>A-38:G>A | 6.7 | 5.2 | 1.0 |
| PintoXRosecoco | DF | 1 | 165 | 3377876\|F\|0-30:A>C-30:A>C | 3382650\|F\|0-43:G>A-43:G>A | 30.6 | 40.5 | 2.7 |
| PintoXRosecoco | DF | 2 | 234 | 8198597\|F\|0-15:T>G-15:T>G | 3377263\|F\|0-22:G>C-22:G>C | 4.3 | 4.3 | -0.9 |
| PintoXRosecoco | DF | 5 | 186 | 3365766\|F\|0-14:T>C-14:T>C | 3367185\|F\|0-53:G>A-53:G>A | 2.9 | 6.2 | -1.0 |
| PintoXRosecoco | DF | 8 | 118 | 3378175\|F\|0-46:T>G-46:T>G | 3384226\|F\|0-45:G>A-45:G>A | 6.3 | 11.4 | 1.4 |
| PintoXRosecoco | DF | 9 | 288 | 3377566\|F\|0-12:C>G-12:C>G | 3366375\|F\|0-37:T>C-37:T>C | 8.1 | 8.6 | 1.2 |
| PintoXRosecoco | DF | 11 | 270 | 16646762\|F\|0-26:T>C-26:T>C | 8210375\|F\|0-19:G>A-19:G>A | 12.7 | 11.8 | -1.5 |
| RosecocoXpinto | DF | 1 | 46 | 3378970\|F\|0-9:C>T-9:C>T | 3381677\|F\|0-26:C>G-26:C>G | 3.3 | 3.7 | -0.7 |
| RosecocoXpinto | DF | 1 | 194 | 3383103\|F\|0-22:C>G-22:C>G | 8213730\|F\|0-21:G>A-21:G>A | 18.0 | 37.5 | -2.3 |
| RosecocoXpinto | DF | 3 | 277 | 3373798\|F\|0-56:C>G-56:C>G | 8198235\|F\|0-24:T>C-24:T>C | 3.2 | 3.5 | 0.7 |
| RosecocoXpinto | DF | 5 | 74 | 8211088\|F\|0-29:A>G-29:A>G | 3381994\|F\|0-45:G>C-45:G>C | 3.9 | 6.5 | 2.7 |
| RosecocoXpinto | DF | 7 | 1 | 13122512\|F\|0-38:G>C-38:G>C | 100051674\|F\|0-17:G>A-17:G>A | 16.2 | 22.2 | 1.8 |
| RosecocoXpinto | DF | 7 | 150 | 16646999\|F\|0-5:T>C-5:T>C | 3378944\|F\|0-45:C>T-45:C>T | 3.2 | 7.9 | 1.1 |
| RosecocoXpinto | DF | 9 | 181 | 8207343\|F\|0-60:T>A-60:T>A | 8209966\|F\|0-27:A>G-27:A>G | 3.2 | 3.8 | -0.7 |
| BlackXPinto | DF | 2 | 52 | 100051993\|F\|0-12:G>A-12:G>A | 100045876\|F\|0-40:A>G-40:A>G | 2.9 | 20.6 | -1.4 |
| PintoXBlack | DF | 8 | 4 | 100045132\|F\|0-46:A>T-46:A>T | 3370680\|F\|0-21:C>T-21:C>T | 3.0 | 7.7 | 1.0 |
| PintoXBlack | DF | 1 | 21 | 8214403\|F\|0-30:C>T-30:C>T | 8198338\|F\|0-64:A>T-64:A>T | 6.7 | 24.4 | 4.2 |
| PintoXRosecoco | DM | 1 | 198 | 3381013\|F\|0-65:T>C-65:T>C | 8200984\|F\|0-48:A>T-48:A>T | 12.1 | 17.6 | -2.2 |
| PintoXRosecoco | DM | 3 | 276 | 8198780\|F\|0-17:C>G-17:C>G | 100051248\|F\|0-28:C>T-28:C>T | 3.2 | 4.5 | -1.1 |
| PintoXRosecoco | DM | 4 | 251 | 100027868\|F\|0-27:C>T-27:C>T | 13122574\|F\|0-38:A>G-38:A>G | 2.9 | 3.3 | 0.9 |
| PintoXRosecoco | DM | 7 | 196 | 3378670\|F\|0-18:T>A-18:T>A | 3381830\|F\|0-48:C>T-48:C>T | 6.8 | 9.8 | 1.6 |
| PintoXRosecoco | DM | 8 | 55 | 3378594\|F\|0-55:T>C-55:T>C | 3371535\|F\|0-27:C>T-27:C>T | 4.2 | 5.4 | -1.2 |
| PintoXRosecoco | DM | 10 | 151 | 8215179\|F\|0-27:T>G-27:T>G | 8213716\|F\|0-42:A>G-42:A>G | 5.4 | 6.6 | 1.4 |
| PintoXRosecoco | DM | 11 | 207 | 100042431\|F\|0-24:C>T-24:C>T | 100075102\|F\|0-9:T>C-9:T>C | 8.5 | 11.1 | -1.8 |
| RosecocoXpinto | DM | 1 | 193 | 3378093\|F\|0-35:C>T-35:C>T | 8215758\|F\|0-28:T>C-28:T>C | 7.1 | 23.9 | 2.6 |
| RosecocoXpinto | DM | 2 | 315 | 8212471\|F\|0-56:G>A-56:G>A | 3384276\|F\|0-35:C>G-35:C>G | 3.5 | 10.3 | -1.7 |

| Cross | Trait | Chr | Position | Left marker | Right marker | LOD score | PVE (%) | Add |
|---|---|---|---|---|---|---|---|---|
| RosecocoXpinto | DM | 9 | 28 | 3379459\|F\|0-65:G>A-65:G>A | 3379459\|F\|0-44:T>G-44:T>G | 3.3 | 15.3 | 3.3 |
| PintoXBlack | DM | 3 | 68 | 8215537\|F\|0-49:G>T-49:G>T | 3380638\|F\|0-9:A>G-9:A>G | 3.0 | 13.6 | -1.6 |

Chr=Chromosome, PVE=Phenotypic variance explained, Add=Additive effect, LOD=Log of odds, DF=Days to flowering, Days to maturity.

### 6.4.5 Number of pods per plant and pod length

QTLs detected to control the number of pods per plant were five in total located on chromosomes 2, 4, and 6. Four of these QTLs contributed toward a higher number of pods per plant in the range of 1.8 to 8.7. QTLs on chromosome six had a negative additive effect of -3.4 and -1.3. QTLs with the highest positive additive effect of 8.7 was found on chromosome four with a LOD score of 4.8 and explained 25.6% of the phenotypic variance (Table 6.7).

A total of eight QTLs affecting pod length were detected on chromosomes 2, 3, 7, 8, 10, and 11. Six of these QTLs had a positive additive effect ranging from 0.3 to 0.6, while the rest had a negative additive effect of -0.4 to -0.5. The QTLs with the highest contribution towards longer pods were found on chromosome seven, with a LOD score of 6.9, and explained 19.4% of the phenotypic variance. On the other hand, the QTL that contributed most to shorter pods was detected on chromosome ten with a LOD score of 2.8 and explained 11.2% of the phenotypic variance of pod length (Tables 6.7).

**Table 6.7: Significant QTLs for number of pods per plant and pod length detected using inclusive composite interval mapping for Pinto (GLPx92) X Rosecoco (GLP2) and Pinto (GLPx92) X Black (GBK035420) $F_{2:6}$ RILs**

| Cross | Trait | Chr | Position | Left marker | Right marker | LOD score | PVE (%) | Add |
|---|---|---|---|---|---|---|---|---|
| PintoXRosecoco | PP | 2 | 121 | 100027117\|F\|0-24:G>A-24:G>A | 13122120\|F\|0-29:T>C-29:T>C | 2.6 | 18.8 | 2.5 |
| PintoXRosecoco | PP | 6 | 134 | 3384331\|F\|0-44:T>A-44:T>A | 3370745\|F\|0-5:T>A-5:T>A | 2.6 | 8.4 | -1.3 |
| RosecocoXpinto | PP | 2 | 13 | 3379962\|F\|0-20:T>G-20:T>G | 3366452\|F\|0-28:A>G-28:A>G | 5.5 | 18.9 | 1.8 |
| BlackXPinto | PP | 6 | 89 | 8200710\|F\|0-22:A>G-22:A>G | 3383514\|F\|0-39:A>G-39:A>G | 2.8 | 40.9 | -3.4 |
| PintoXBlack | PP | 4 | 55 | 100050314\|F\|0-30:A>G-30:A>G | 100047639\|F\|0-26:T>A-26:T>A | 4.8 | 25.6 | 8.7 |
| PintoXRosecoco | PL | 3 | 108 | 3378047\|F\|0-31:G>A-31:G>A | 8211763\|F\|0-5:A>G-5:A>G | 4.3 | 8.8 | 0.5 |
| RosecocoXpinto | PL | 2 | 251 | 100035053\|F\|0-45:A>G-45:A>G | 3383135\|F\|0-59:G>C-59:G>C | 3.3 | 8.0 | -0.4 |
| RosecocoXpinto | PL | 7 | 86 | 100036259\|F\|0-29:T>G-29:T>G | 3379840\|F\|0-28:C>A-28:C>A | 6.9 | 19.4 | 0.6 |
| RosecocoXpinto | PL | 10 | 151 | 100084370\|F\|0-68:T>C-68:T>C | 8216613\|F\|0-45:G>C-45:G>C | 2.8 | 11.2 | -0.5 |
| BlackXPinto | PL | 8 | 73 | 3377109\|F\|0-54:C>T-54:C>T | 3380605\|F\|0-10:A>G-10:A>G | 2.5 | 29.1 | 0.5 |
| BlackXPinto | PL | 11 | 207 | 3380363\|F\|0-29:G>A-29:G>A | 8209950\|F\|0-19:T>A-19:T>A | 2.7 | 22.4 | 0.4 |
| PintoXBlack | PL | 2 | 53 | 3377989\|F\|0-43:T>A-43:T>A | 8206626\|F\|0-65:A>G-65:A>G | 9.4 | 27.7 | 0.4 |
| PintoXBlack | PL | 11 | 175 | 3382513\|F\|0-19:T>A-19:T>A | 8215619\|F\|0-50:A>T-50:A>T | 4.7 | 16.3 | 0.3 |

PVE=Phenotypic variance explained, Add=Additive effect, LOD=Log of odds, CT=Cooking time, PP=number of pods per plant, PL=Pod length.

### 6.4.6 Number of seeds per pod and seed weight

QTLs detected to affect the number of seeds per pod were five in total and were on chromosomes 1, 8, and 10. Three QTLs found on chromosomes 1, 8, and 10 contributed towards more seeds per pod with an additive effect of 0.2 to 1.3, while two QTLs found on chromosome 10 contributed toward fewer seeds per pod with both having an additive effect of (-0.4). QTLs on chromosome one had the highest additive effect (1.3) with a LOD score of 2.5 and explained 17.3% of the phenotypic variance of the number of seeds per pod (Tables 6.8).

In total, 10 QTLs affecting seed weight were detected on chromosomes 1, 5, 6, 7, 9, and 11. The majority (13) of the QTLs for seed weight were detected in RILs from the cross-involving pinto and black (Tables 6.8). Six of these QTLs contributed toward low seed weight with a negative additive effect ranging from -2.5 to -0.4. The rest had a positive additive effect that ranged from 1.3 to 7.2. The QTLs which contributed most to more seed weight were on chromosome nine with a LOD score of 3.8 and explained 8.8% of the phenotypic variance. QTLs that had the highest effect towards less seed weight were on chromosome one with a LOD score of 4.4 and explained 9.2% of the phenotypic variation in seed weight (Tables 6.8).

**Table 6.8: Significant QTLs for number of seeds per pod and seed weight detected using inclusive composite interval mapping for Pinto (GLPx92) X Rosecoco (GLP2) and Pinto (GLPx92) X Black (GBK035420) $F_{2:6}$ RILs**

| Cross | Trait | Chr | Position | Left marker | Right marker | LOD score | PVE (%) | Add |
|---|---|---|---|---|---|---|---|---|
| PintoXRosecoco | SP | 10 | 37 | 3375452\|F\|0-32:C>A-32:C>A | 3377795\|F\|0-13:G>A-13:G>A | 3.0 | 17.3 | -0.4 |
| RosecocoXpinto | SP | 1 | 181 | 100052110\|F\|0-16:G>A-16:G>A | 3378280\|F\|0-49:A>G-49:A>G | 2.5 | 17.3 | 1.3 |
| RosecocoXpinto | SP | 10 | 151 | 100084370\|F\|0-68:T>C-68:T>C | 8216613\|F\|0-45:G>C-45:G>C | 3.5 | 18.1 | -0.4 |
| PintoXBlack | SP | 8 | 240 | 100033278\|F\|0-24:C>A-24:C>A | 100046319\|F\|0-27:T>C-27:T>C | 5.3 | 17.8 | 0.3 |
| PintoXBlack | SP | 10 | 48 | 3382802\|F\|0-10:A>T-10:A>T | 100031864\|F\|0-56:G>T-56:G>T | 2.8 | 9.9 | 0.2 |
| PintoXRosecoco | SW | 5 | 110 | 3365598\|F\|0-34:T>G-34:T>G | 3378671\|F\|0-39:C>T-39:C>T | 3.4 | 3.9 | 2.0 |
| PintoXRosecoco | SW | 9 | 116 | 3382317\|F\|0-31:A>G-31:A>G | 3384285\|F\|0-55:A>C-55:A>C | 2.8 | 6.5 | -2.1 |
| PintoXRosecoco | SW | 9 | 172 | 3381643\|F\|0-47:C>T-47:C>T | 3381371\|F\|0-8:A>C-8:A>C | 2.6 | 3.1 | 1.4 |
| BlackXPinto | SW | 7 | 129 | 3384098\|F\|0-23:C>T-23:C>T | 3384303\|F\|0-57:C>G-57:C>G | 2.5 | 15.4 | -1.7 |
| PintoXBlack | SW | 1 | 215 | 8214289\|F\|0-18:G>A-18:G>A | 3382934\|F\|0-49:A>G-49:A>G | 4.7 | 10.6 | -1.6 |
| PintoXBlack | SW | 6 | 185 | 100031585\|F\|0-22:G>T-22:G>T | 100031452\|F\|0-6:C>A-6:C>A | 2.7 | 6.2 | -1.2 |
| PintoXBlack | SW | 9 | 242 | 3377682\|F\|0-11:G>A-11:G>A | 8196251\|F\|0-26:A>G-26:A>G | 3.8 | 8.8 | 7.2 |
| PintoXBlack | SW | 11 | 174 | 3369798\|F\|0-10:G>A-10:G>A | 3382513\|F\|0-18:C>T-18:C>T | 6.3 | 16.2 | -1.9 |
| PintoXBlack | SW | 11 | 305 | 16649721\|F\|0-25:T>A-25:T>A | 3383217\|F\|0-7:T>C-7:T>C | 5.5 | 15.7 | 1.9 |
| RosecocoXpinto | SW | 1 | 45 | 100058942\|F\|0-48:T>C-48:T>C | 3383292\|F\|0-60:T>A-60:T>A | 4.4 | 9.2 | -2.5 |

PVE=Phenotypic variance explained, Add=Additive effect, LOD=Log of odds, SP=Seeds per pod, SW=Seed weight.

## 6.5 Discussion

The results showed significant (P≤0.05) differences among RILs for all the traits recorded, which indicates that there existed genetic variability for the traits in the population. The study ended up with an unequal number of RILs due to root rot diseases during the final multiplication of seeds in the field (Table 6.3). The correlation analysis result shows a significant (P≤0.05) moderate positive correlation between duration to flowering and duration to maturity indicating that the two traits might be under the influence of the same genes. This suggests that duration to flowering can be used to predict duration to maturity and both can be selected simultaneously (Lobo, 2008). Duration to flowering is under the control of a lower dominance and a higher additive gene effect (Mendes *et al.,* 2008). When the dominance effect is present it reduces the number of days to flowering. A high correlation of 0.7 between days to flowering and days to maturity has been reported in a previous study conducted by Kamfwa *et al.,* (2015). Days to flowering has been reported in various studies to be highly associated with days to maturity (Okii *et al.,* 2014; Kamfwa *et al.,* 2015). Seed weight is used as a proxy of seed size, a significant correlation of 0.47 between pod length and 100-seed weight suggests that large-seeded genotypes tend to have longer pods. Okii *et al.,* (2014) reported a similar correlation result (0.48) between a 100-seed weight and pod length.

The formation of insoluble pectates at the cell wall and middle lamella is believed to be the cause of HTC in common bean (Shomer *et al.,* 1990; Hentges *et al.,* 1990). Pectin comprises complex acid polysaccharides with a backbone of galacturonic acid residue with an alpha-1,4-glycosidic linkage (Atkinson *et al.,* 2002). Homogalacturonan-rich pectin is commonly found in the middle lamella region of plant cell walls where two cells border (Atkinson *et al.,* 2002). At high temperatures and relative humidity, pectin methylesterase (PME) hydrolyses pectin molecules forming pectic acid and methanol. The magnesium and calcium released in the cells migrate to the middle lamella and produce an insoluble magnesium pectinate and calcium pectinate that cements cells together hardening the cell wall (Jones and Boulter, 1983a).

This study revealed two major QTLs on chromosomes 3 and 10 with the highest additive effect and explained the highest phenotypic variation in cooking time. The

region on chromosome 10 with the QTL for cooking time, was also detected in the genome wide study discussed in chapter 4. The QTL was at 15 cM on the linkage map and around 3968311 to 11971124 bp on the physical map. The QTL co-localized with 12 genes related to the formation of pectin at the cell wall, six genes encoded enzyme polygalacturonase/pectinase, three for pectin methylesterase, two genes for pectinesterase inhibitor, and one gene for galacturan 1, 4 alpha-galacturonidase. The co-localization of these loci with the identified QTLs supports the theory of insoluble pectin at the cell wall and middle lamella as the cause of hard-to-cook trait. However, there were no candidate genes related to the formation of pectin in the cell walls found within the QTL region on chromosome three. QTL for the cooking time on chromosome nine at position 133 cM on the linkage map had an additive effect of 17.8 and was located around 23833662 to 25258872 bp in the physical map. The region contained six genes encoding pectinesterase inhibitor, one gene for pectinesterase, and one gene related to the pectate lyase family.

Another region with QTL for the cooking time detected on chromosome one at 92 cM on the linkage map had an additive effect of 19.6. The QTL region was at 34267322 to 33304000 bp on the physical map and co-localized with two genes for polygalacturonase, one gene for polygalacturonase inhibitor, and one gene for pectinesterase.

Using the flanking markers to locate the QTL on the physical map for the cooking time on chromosome one at 21 cM, the region was located at 42975691 to 51047749 bp. The region had nine genes for pectin methylesterase inhibitor, three genes for pectin methylesterase, three genes for pectate lyase, one gene for polygalacturonase, and one gene for alpha-galactosidase.

In total, 52 candidate genes that play a role in the formation of pectin co-localized with QTLs for cooking time with an additive effect of ten and above. Pectinase and polygalacturonase enzymes are used to break down the pectin compound found in plant cell walls particularly, middle lamella to extract cell sap (Phutela *et al.*, 2005). This study supports the theory of the formation of insoluble pectin at the cell wall and middle lamella as the cause of HTC.

Several studies have reported QTLs that control cooking time. Jacinto-Hernandez *et al.*, (2003) reported a random amplified polymorphic DNA (RAPD) marker associated

with the cooking time that explained 23% of the variation in cooking time using 104 RILs. Six QTLs that govern cooking time were reported on chromosomes 1 and 9 using 105 polymorphic SSR markers and 140 $F_{2:4}$ RILs (Garcia *et al.,* 2012). Significant SNPs markers associated with cooking time were identified on chromosomes 2, 3, and 6 using GWAS on 206 common bean accessions, the SNPs explained between 4 to 8.7% of the phenotypic variation (Cichy *et al.,* 2015). Berry *et al.,* (2020) identified 10 QTLs on chromosomes 1, 2, 3, 5, 6, 10, and 11, with the most robust QTLs being on chromosomes 3, 6, 10, and 11 detected in over two different environments using 146 RILs of common bean.

A transcript locus for phytochrome interacting factor and agamous-like MADS-box loci have been reported to control plant development (Li *et al.,* 2017). Agamous-like MADS-box transcript *Phvul.001G186400.1* was located within the area of the identified QTL for days to flowering on chromosome one at 21 cM that had the highest negative additive effect (-2.3) and was in the region 43921483 to 44604321 bp on the physical genetic map. Results show a cluster of QTLs with a positive additive effect for days to flowering on chromosome one. Four loci that may play a role in flowering were also found within the QTL region with the highest positive additive defect (4.2) on chromosome one at 194 cM, located around 42975691 to 51047749 bp on the physical map. Loci found in this region include two genes encoding for phytochrome interacting transcription factor, one for growth regulator factor and one for agamous-like box protein.

MADS-box loci also known as MICK-type genes have been reported to control various plant development processes like vegetative growth and reproductive organ development (Adamczyk and Fernandez, 2009). Phytochrome-interacting factors are basic helix-loop-helix transcription factors that play critical roles in the germination of seeds, photomorphogenesis, responses to shading, flowering time, and leaf senescence (Sakuraba *et al.,* 2014; Shi *et al.,* 2018). However, Phytochrome-interacting loci were not found in the QTL regions detected in this study. QTLs for the duration to flowering on chromosome one had been reported in previous studies by Koinange *et al.,* (1996), Blair *et al.,* (2006b), Perez-Vega *et al.,* (2010), Mukeshimana *et al.,* (2014), Kamfwa *et al.,* (2015) and Langat *et al.,* (2019). Other QTLs have been reported on chromosome four by Mukeshimana *et al.,* (2014) and Langat *et al.,* (2019), on

chromosome eight by Koinange *et al.,* (1996), Perez-Vega *et al.,* (2010), Kamfwa *et al.,* (2015), and Briñez *et al.,* (2017).

The search for candidate genes within the QTL with high negative (-2.2) and positive (3.3) additive effects on duration to maturity located on chromosomes one and nine, respectively, showed no gene of interest. Agamous-like transcript *Phvul.001G186400.1* co-localized with the QTL with an additive effect of 2.6 on chromosome one for the duration to maturity located at 193 cM on the linkage map and 44386397 to 44809486 bp on the physical map. Langat *et al.,* (2019) reported a QTL for the duration to maturity that co-localized with QTL for the duration to flowering on chromosome one. QTLs for the duration to maturity have been reported on chromosome four (Mukeshimana *et al.,* 2014; Langat *et al.,* 2019), chromosome seven (Mukeshimana *et al.,* 2014), and chromosome nine (Mukeshimana *et al.,* 2014; Kamfwa *et al.,* (2015).

The number of pods per plant is a primary yield component and part of the accumulated aerial biomass partitioned to seed yield (Negahi *et al.,* 2014). A significant ($P \leq 0.05$) strong and positive correlation between grain yield and the number of pods per plant has been reported (Langat *et al.,* 2019). Five QTLs for the number of pods per plant were identified in this study, QTLs with the highest positive additive effect of 8.7 was found on chromosome four. QTLs for the number of pods per plant have been identified in previous studies using different methods and populations. QTL controlling the number of pods per plant was reported by Koinange *et al.,* (1996) on Pv01 and Pv08 in a population of 65 $F_8$ RILs developed from a cross of Mildas and G12873. Blair *et al.,* (2006b) reported a QTL of the same trait on Pv07, Pv09, and Pv11 in an inbred backcross population of 157 $BC_2$ $F_{3:5}$ from a cross between ICA Cerinza and G24404. Tar'an *et al.,* (2002) mapped a QTL for the number of pods per plant on Pv02 using 145 $F_{4:5}$ RILs from a cross of OAC Seaforth and OAC 95-4 navy bean. Kamfwa *et al.,* (2015) identified QTL for the number of pods per plant on Pv03 and Pv09 using 237 genotypes.

Pod length is a measure of the size of the harvested part of French beans, consumers prefer straight, rounded pods with a length ranging from 10 to 16 cm depending on the grade of the harvested pods (Wahome *et al.,* 2013). A total of eight QTLs affecting pod length were detected, six of these QTLs had a positive additive effect ranging from

0.3 to 0.6. The QTLs with the highest contribution towards longer pods were found on chromosome seven. This study and others have shown that pod length is also positively correlated with seed weight (Okii *et al.,* 2014), suggesting that the genes controlling these two traits are linked.

The number of seeds per pod is one of the primary seed yield components (Ghobary and Allah, 2010; Negahi *et al.,* 2014), high yielding varieties have a higher number of seeds per pod (Ashango and Alamerew, 2017). A total of five QTLs on chromosomes 1, 8, and 10 were detected in this study. Three of these QTLs contributed toward more seeds per pod with QTLs on chromosome one having the highest additive effect. QTLs for the number of seeds per pod have been reported on chromosome 2 (Langat *et al.,* 2019), chromosomes 6 and 7 (Briñez *et al.,* 2017), and chromosome 8 (Briñez *et al.,* 2017; Langat *et al.,* 2019).

Seed weight quantifies the size of the seeds, and it is one of the first-order yield components (Negahi *et al.,* 2014). The Andean gene pool is generally large-seeded and adapted to relatively higher altitudes and lower temperatures, on the other hand, the Mesoamerican gene pool is small-seeded and adapted to lower altitudes and higher temperatures (Beebe *et al.,* 2011). Ten QTLs affecting seed weight were detected on chromosomes in this study. Six of these QTLs contributed toward low seed weight, while the rest had a positive additive effect. The QTL which contributed most to higher seed weight was detected on chromosome nine while QTL with the highest effect towards less seed weight was on chromosome one. Several QTLs for seed weight have been mapped in previous studies on chromosome one (Koinange *et al.,* 1996; Broughton *et al.,* 2003; Briñez *et al.,* 2017), chromosome five (Briñez *et al.,* 2017), on chromosome seven (Koinange *et al.,* 1996; Mukeshimana *et al.,* 2014), on chromosome eight (Langat *et al.,* 2019) and chromosome eleven (Koinange *et al.,* 1996).

### 6.6 Conclusion

The study identified QTLs affecting cooking time and various morphological traits of common bean using $F_{2:6}$ recombinant inbred lines of common bean derived from two biparental crosses. QTLs associated with days to flowering, days to maturity, number of pods per plant, pod length, number of seed per pod and seed weight were detected.

Agamous-like MADS-box transcripts like *Phvul.001G186400.1* locus co-localized with QTLs for days to flowering and maturity. QTLs controlling cooking duration were detected on chromosomes 1, 2, 3, 5, 6, 9, 10, and 11, with chromosomes one and two having more than one QTL. QTLs on chromosomes three and ten had the highest additive effect of 27.2 towards longer cooking time and both explained 37.7% of the phenotypic variance. The study identified gene transcripts in the QTLs regions in the genome known to control enzymes involved in the formation and breakdown of pectin in plant cell walls. The genes found to co-localize with the detected QTLs for cooking time encodes for polygalacturonase/pectinase, pectin methylesterase, pectinesterase inhibitor, and galacturan 1, 4 alpha-galacturonidase enzymes. Therefore, this study points towards the theory of the formation of insoluble pectin as the cause of the HTC trait. The identified QTLs could be useful in the introgression of cooking time traits and implementation of MAS in common bean breeding.

# CHAPTER SEVEN
# CONCLUSION AND RECOMMENDATIONS

## 7.1 General conclusion

Common bean plays a critical role in the nutrition security of a large population as a vital source of protein in third-world countries and the diet of vegetarians. Identifying genomic regions that control the cooking time of grains and traits of agronomic importance of common bean is crucial to aid plant breeding efforts to improve the crop. Further, understanding the inheritance of these traits in common bean would assist breeders to choose an appropriate breeding method. This study evaluated the cooking time of fresh and aged seeds, duration to flowering and maturity, number of pods per plant, pod length, number of seeds per pod, seed weight, and seed yield of common bean through phenotyping and genotyping.

This study characterized and genotyped a population of common bean accessions and $F_{2:6}$ recombinant inbred lines (RILs) and identified easy-to-cook genotypes and quantitative trait loci (QTLs) associated with the cooking time, duration to flowering and maturity, number of pods per plant, pod length, number of seeds per pod, seed weight and seed yield. Significant variation existed among the common bean accessions and RILs evaluated for all the traits recorded. Traits that showed high broad sense heritability ($H^2$) included days to flowering, 100-seed weight and grain yield. Large-seeded accessions, climbing accessions, and popular seed classes (pinto, calima, small reds, and purples) had higher yields. The study found that storage of common grains at temperature of 35°C and relative humidity of 50% significantly increased cooking time by an average of 14.1 minutes.

Genome wide association study (GWAS) identified a total of 33 SNPs markers significantly associated with days to flowering, days to maturity, number of pods per plant, pod length, number of seed per pod, seed weight and yield. Two SNPs markers were also found to be significantly ($P \leq 0.05$) associated with cooking time of aged seeds. The association between trait and marker was found to be influenced by seasonal changes.

QTL analysis study identified various QTLs associated with days to flowering, days to maturity, number of pods per plant, pod length, number of seed per pod and seed weight. The genomic regions on chromosome one with QTLs for the days to flowering

and chromosome 10 for the cooking time were detected on both GWAS and QTL analyses studies.

The study found that agamous-like MADS-box transcript (*Phvul.001G186400.1*) loci co-localized with QTL for days to flowering and maturity, while galacturan 1,4-alpha galacturonidase (*Phvul.010G038000*) and polygalacturonase *(Phvul.010G038100*) loci co-localized with the QTL for cooking time. Other loci found to co-localize with the detected QTLs for cooking time include pectin methylesterase, pectinesterase inhibitor, and galacturan. These enzymes are involved in the formation and breakdown of pectin in the plant cell wall responsible for the development of the hard-to-cook trait. The findings of the GWAS and QTL analysis support the theory of the formation of insoluble pectin in the cell wall and middle lamella as the cause of the HTC trait. Common bean accessions evaluated in this study showed heritable variation that can be exploited to improve common bean in breeding programs. QTLs identified could be useful by enabling marker-assisted selection (MAS) in breeding of common bean breeding.

### 7.2 Recommendations

Common bean accessions evaluated in this study showed heritable variation that can be exploited to improve common beans in breeding programs. The identified accession with shorter cooking time and higher yields can be evaluated in different locations to determine their adaptability and stability of their performance in yield.

Higher yields were recorded for large-seeded accessions, climbing accessions, and popular (pinto, calima, small reds, and purples) seed classes. Farmers with small pieces of land could be encouraged to grow popular climbing common bean varieties to increase the productivity.

The identified quantitative trait loci (QTLs) for cooking time require further investigations to identify their robustness under different environment, storage conditions and with common bean of different genetic backgrounds. A replication of the GWAS study is recommended using cooking time of accessions grown in different locations to evaluate the robustness of the QTL regions identified. Further investigation using linkage study using recombinant inbred lines (RILs) developed from Easy-to-cook (ETC) and hard-to-cook (HTC) accessions identified in this study

as the parents. A combination of differential expression-based study and QTL mapping to identify the candidate gene would improve the precision in pursuit of candidate genes.

The identified molecular markers can be used in developing varieties with desirable traits through markers-assisted selection (MAS), and the identified potential candidate genes can be utilized in breeding programs to improve the cooking quality of the common bean.

## REFERENCES


Adamczyk, B., & Fernandez, D. E. (2009). MIKC* MADS domain heterodimers are required for pollen maturation and tube growth in Arabidopsis. *Plant Physiology*, 149:1713–1723. https://doi.org/10.1104/pp.109.135806

Agriculture and Food authority (2022). *Year book of statistics* 2022. 141 pages. Retrieved from the link https://www.agricultureauthority.go.ke › index.php

Akinyele, I. O., Onigbinde, A. O., Hussain, M. A., & Omololu, A. (1986). Physicochemical characteristics of 18 cultivars of Nigerian cowpeas (*V. unguiculata*) and their cooking properties. *Journal of Food Science,* 51, 1483-1485. https://doi.org/10.1111/j.1365-2621.1986.tb13840.x

Alexandre, A., & Oliveira, S. (2012). Response to temperature stress in rhizobia. *Critical reviews in microbiology*, 39. DOI. 10.3109/1040841X.2012.702097

Amanullah, Khan, A. A., Nawab, K., & Sohail, Q. (2006). Performance of promising common bean (*Phaseolus vulgaris* L.) germplasm at Kala-Swat. *Pakistan Journal of Biological Sciences,* 9(14):2642-2646. https://doi.org/10.3923/pjbs.2006. 2642.2646

Andrés, F., & Coupland, G. (2012). The genetic basis of flowering responses to seasonal cues. *Nature Reviews Genetics,* 13(9):627–639. https://doi.org/10.1038/nrg3291

Angioi, S. A., Rau, D., Nanni, l., Attene, G., & Bellucci, E. (2010). Beans in Europe: origin and structure of the European landraces of *Phaseolus vulgaris* L. *Theoretical and Applied Genetics*, 121:829–843. doi 10.1007/s00122-010-1353-2

Anonimo (1982). *Morphology of the common bean plant study guide to be used as a complement to the tutorial unit*. CIAT, Cali Colombia, 50 pages.

Anunda, H. N., Nyaboga, E.N., & Amugune, N. O. (2019). Evaluation of genetic variability, heritability, genetic advance and correlation for agronomic and yield components in common bean landraces from southwestern Kenya. *Journal of Plant Breeding and Crop Science,* 11(5):144-147. https://doi.org/10.5897/jpbcs 2018.0800

Arruda, B., Altamir, F., Jerfferson, L., & Jaqueline, B. (2012). Environment is crucial to the cooking time of beans. *Ciencia e Tecnologia de Alimentos,* 110:573-578 https://doi.org/10.1590/s0101-20612012005000078

Ashango, Z., & Alamerew, S. (2017). Seed yield and its related traits performance of common bean (*Phaseolus vulgaris* L.) varieties in Dawro zone, southwest Ethiopia. *Advances in Life Science and Technology*, 53:24-35. Retrieved from https://api.semanticscholar.org/CorpusID:56325410

Atkinson, R., Schroder, R., Hallett, C., Cohen, D., & Macke, E. (2002). Overexpression of polygalacturonase in transgenic apples tress leads to a range of novel phenotypes involving changes in adhesion. *Plant Physiology*, 129(1):122-133. doi: 10.1104/pp.010986

Bassett, A., Katuuramu, D. N., Song, Q., & Cichy, K. (2021). QTL mapping of seed quality traits including cooking time, flavor, and texture in a yellow dry bean (*Phaseolus vulgaris* l.) population. *Frontier in Plant Science*, 12:670284. doi: 10.3389/fpls.2021.670284

Beattie, A. D., Larsen, J., Michaels, T. E., & Pauls, K. P. (2003). Mapping quantitative trait loci for a common bean (*Phaseolus vulgaris* L.) ideo-type. *Genome*, 46(3):411-422. https://doi.org/10.1139/g03-015

Beebe, S. (2012). Common bean breeding in tropics. *Plant Breeding Reviews*, Pages 357-426. https://doi.org/10.1002/9781118358566.ch5

Beebe, S., Ramirez, J., Jarvis, A., Rao, I., & Mosquera, G. (2011). Genetic improvement of common beans and the challenge of climate change. In: Yadav s, Redden J, Hatfield L, Lotze-Campen, Hall E (ed) *Crop adaptation to climate change*. Wiley-Blackwell, Oxford, pages 356-369. https://doi.org/10.1002/9780470960929.ch25

Beebe, S., Rao, I., & Blair, M. (2009). Breeding for abiotic stress tolerance in common bean: Present and future challenges. SABRAO *Journal of Breeding and Genetics*, 41. Retrieved from the link https://www.researchgate.net/publication/228612601

Berry, M., Izquierdo, P., Jeffery, H., Shaw, S., Nchimbi-Msolla, S., & Cichy, K. (2020). QTL analysis of cooking time and quality traits in dry bean (*Phaseolus vulgaris* L.). *Theoretical and Applied Genetics*, 133(7):2291–2305. https://doi.org/10.1007/s00122-020-03598-w

Bitocchi, E., Nanni, L., Bellucci, E., Rossi, M., Giardini, A., Zeuli, P. S., Logozzo, G., Stougaard, J., McClean, P., Attene, G., & Papa, R. (2012). Mesoamerican origin of the common bean (*Phaseolus vulgaris* L.) is revealed by sequence data.

*Proceedings of the National Academy of Sciences*, 109(14), E788-E796. https://doi.org/10.1073/pnas.1108973109

Blair, M. W., Galeano, C. H., Tovar, E., Torres, M. C., Castrillón A. V., Beebe, S. E., & Rao, I. M. (2012). Development of a Mesoamerican intra-genepool genetic map for quantitative trait loci detection in a drought tolerant × susceptible common bean (*Phaseolus vulgaris* L.) cross. *Molecular Breeding*, 29(1):71-88. https://doi.org/10.1007/s11032-010-9527-9

Blair, M., Diaz, J., Hildago, R., Diaz, L., & Duque, M. (2007) Microsatellite characterization of Andean races of common bean (*Phaseolus vulgaris* L.). *Theoretical and Applied Genetics*, 116(1):29-43. https://doi.org/10.1007/s00122-007-0644-8

Blair, W., Iriate, G., & Beebe, S. (2006). QTL analysis of yield traits in an advanced backcross population derived from in an Andean X wild common bean (*Phaseolus vulgaris* L.) cross. *Theoretical and Applied Genetics*, 112(6):1149-1163. https://doi.org/10.1007/s00122-006-0217-2

Bressani, R., & Chon, C. (1996). Effects of altitude above sea level on the cooking time and nutritional value of common beans. *Plants Food for Human Nutrition*, 49:53-61. https://doi.org/10.1007/bf01092522

Bressani, R., Navarrete, D. A., & Garcia-Soto, A. (1988). Culinary practices and consumption characteristics of common beans at the rural home level. *Archvos Latina Americans de Nutricio*, 38: 925-934.

Briñez, B., Perseguini, J., Rosa, J. S., Bassi, D., Gonçalves, J.G., Almeida, C., Paulino, J. F., Blair, M. W., Chioratto, A. F., Carbonell, S. A., Valdisser, P. A., Vianello, R. P., & Benchimol-Reis, L. B. (2017). Mapping QTLs for drought tolerance in a SEA 5 x AND 277 common bean cross with SSRs and SNP markers. Genetics and *Molecular Biology*, 40(4) 813-823. https://doi.org/10.1590/1678-4685-gmb-2016-0222

Broughton, J. W., Hernandez, G., Blair, M., Beebe, S., & Vanderleyde, J. (2003). Beans (*Phaseolus spp.)* model food legumes. *Plant and Soil*, 252(1):55-128. https://doi.org/10.1023/a:1024146710611

Burbano-Erazo, E., León-Pacheco, R. I., Cordero-Cordero, C. C., López-Hernández, F., Cortés, A. J., & Tofiño-Rivera, A. P. (2021). Multi-Environment Yield

Components in Advanced Common Bean (Phaseolus vulgaris L.) × Tepary Bean (P. acutifolius A. Gray) Interspecific Lines for Heat and Drought Tolerance. *Agronomy*, 11(10):1978. https://doi.org/10.3390/agronomy11101978

Burle, M., Fonseca, J., Kami, J., & Gepts, P. (2010). Microsatellite diversity and genetic structure among common bean (*Phaseolus vulgaris* L.) landraces in Brazil, a secondary center of diversity. *Theoretical and Applied Genetics,* 121:801-813. https://doi.org/10.1007/s00122-010-1350-5

Burton, G. W., & De vane, E. H. (1953). Estimating heritability in tall fescue (*Festuca arundinacea*) from replicated clonal material. *Agronomy Journal,* 45:478-481. https://doi.org/10.2134/agronj1953.00021962004500100005x

CIAT (International Center for Tropical Agriculture) (1986). *Morphology of the common bean (Phaseolus vulgaris L.)*, CIAT, Cali, Colombia CIAT, 56 Pages.

Cichy, K., Wiesinger, J., & Mendoza, F. (2015). Genetic diversity and genome-wide association analysis of cooking time in dry bean (*Phaseolus vulgaris* L.). *Theoretical and Applied Genetics*, 128:1555–1567. DOI:10.1007/s00122-015-2531-z

Cichy, K., Wiesinger, J., Berry, M., Nchimbi-Msolla, S., Fourie, D., Porch, T., Ambechew, D., & Miklas, P. (2019). The role of genotype and production environment in determining the cooking time of dry beans (*Phaseolus vulgaris* L.). *Legume Science*, 1:1–15 https://doi.org/10.1002/leg3.13

Corporate Communications Office (2023). LCEFoNS phase ii to enhance legume value chain in kenya. Retrieved 13th August 2023 from the link https://www.jkuat.ac.ke/lcefons-phase-ii-to-enhance-legume-value-chain-in-kenya/

Corrêaa, M. M., Carvalhoa, L. M., Nuttib, M. R., Carvalhob, J. L., Netob, A. R., & Ribeiro, E. M. (2010). Water Absorption, Hard Shell and Cooking Time of Common Beans (Phaseolus vulgaris L.). *African Journal of Food Science*, 1, 13-20. Retrieved from https://api.semanticscholar.org/CorpusID:5504320

Cortes, A. C., Blair, W., & Chavaro, C. (2011). SNP marker diversity in common bean (*Phaseolus vulgaris* L.). *Theoretical and Applied Genetics,* 123(5):827-845. https://doi.org/10.1007/s00122-011-1630-8

Costa, A., Quieroz-Monici, K., & Oliveira, S. (2006). Chemical composition, dietary

fiber and resistant starch contents of raw and cooked pea, common bean, chickpea and lentil legumes. *Food Chemistry*, 94: 327-330. https://doi.org/10.1016/j.foodchem.2004.11.020

Cruz, G., Oliveira, M., Pires, C., Gomes, M., Costa, N., Brumano, M., & Moreira, M. (2003). Protein quality and in vivo digestibility of different varieties of bean (*Phaseolus vulgaris* L.). *Brazil Journal of Food Technology*, 6: 157-162. Retrieved from https://www.researchgate.net/publication/313662870

Diaz, S., Ariza-Suarez, D., Ramdeen, R., Aparicio, J., Arunachalam, N., Hernandez, C., Diaz, H., Ruiz, H., Piepho, H. P. & Raatz, B. (2021). Genetic architecture and genomic prediction of cooking time in common bean (*Phaseolus vulgaris* L.). *Frontier in Plant Science,* 11:622213. doi: 10.3389/fpls.2020.622213

Dudley, J. W., Moll, R. H. (1969). Interpretation and use of estimates of heritability and genetic variances in plant breeding. *Crop Science*, 9(3):257-262. https://doi.org/10.2135/cropsci1969.0011183x000900030001x

Elena, P., Cristina, R., Astrid, P., & Ana, C. (2010). Mapping QTLs for morph-agronomic and seed quality traits in RILs population of common bean (*Phaseolus vulgaris L.)*. *Theoretical and Applied Genetics*, 120: 1367-1380. https://doi.org/10.1007/s00122-010-1261-5

Elia, F. M. (2003). Heritability of cooking time and water absorption traits in common bean (*Phaseolus vulgaris* L.) using a North Carolina design II mating scheme. *Tanzania Journal of Science*, 29(1). https://doi.org/10.4314/tjs.v29i1.18364

Fabbri, A., Crosby, G. (2016). A review of the impact of preparation and cooking on the nutritional quality of vegetables and legumes. *International Journal of Gastronomy and Food Science*, 3:2-11. https://doi.org/10.1016/j.ijgfs.2015.11.001

FAOSTAT (Food and Agriculture Organisation Statistics) (2023). Food and Agriculture Organisation of the United Nations. Statistic division. Retrieved 23/06/2022 from the link https://www.fao.org/faostat/en/#data/QCL

Farrow, A. & Andriatsitohaina, R. (2021). *Atlas of Common Bean Production in Africa*. Pan-Africa Bean Research Alliance (PABRA); International Center for Tropical Agriculture (CIAT), ISBN: 978958694227-0. Retrieved from https://www.researchgate.net/publication/348860644

Fisseha, Z. & Kyallo, M., T., Kassahun, H., Jagger D., K. & Opiyo, S. (2018).

Integrating phenotypic evaluations with a molecular diversity assessment of an Ethiopian collection of common bean landraces. *African Crop Science Journal*, 26. DOI 10.4314/acsj.v26i2.12.

Freyre, R., Skroch, P. W., Geffroy, V., Adam-Blondon, A., Shirmohamadali A., Johnson W. C, Llaca, V., Nodari, R. O., Pereira, P. A., Tsai, S. M., Tohme, J., Dron, ., Nienhuis, J., Vallejos, C. E., & Gepts, P. (1998). Towards an integrated linkage map of common bean. Development of a core map and alignment of RFLP maps. *Theoretical Applied Genetics*, 97(5-6):847-956. https://doi.org/10.1007/s001220050964

Freytag, G., & Debouck, D. (2002). Review of taxonomy, distribution and ecology of the genus *Phaseolus* (*Lefuminosae-papilionoidae)* in North Americca, Mexico and Central America. *Botanical Miscellany,* 23: 1-300. Retrieved from the link https://www.researchgate.net/publication/44360006

Galisteo, M., Duarte, J., & Zarzuelo, A. (2008). Effects of dietary fibers on disturbances clustered in the metabolic syndrome. *The Journal of Nutritional Biochemistry,* 19(2):71-84. https://doi.org/10.1016/j.jnutbio.2007.02.009

Garcia, E., Filisetti, C., Udaeta, M., Lajolo, M. (1998). Hard-to-cook Beans (*Phaseolus vulgaris* L.): Involvement of phenolic compounds and pectates. *Journal of Agricultural and Food Chemistry*, 46(2):2110-2116. https://doi.org/10.1021/jf970848f

Garcia, R., Rangel, P., Bassinello, P., Brondani, C., Melo, L., Sibov, S., Vianello-Brondani, R. (2012). QTL mapping for the cooking time of common beans. *Euphytica,* 186:779–792 https://doi.org/10.1007/s10681-011-0587-7

Garret, A., Nita, M., De Wolf, D., Gomez, L., & Sparks, A. H. (2009). Plant pathogens as indicators of climate change. Climate change: Observed impacts on plant earth. Elsevier BV, Amsterdam pages 425-437. https://doi.org/10.1016/B978-0-444-63524-2.00021-X

Gepts, P. (1998). Origin and evolution of common bean: Past events and recent trends. *American Society for Horticultural Science*, 33(7):1124-1130. https://doi.org/10.21273/hortsci.33.7.1124

Gepts, P. (2014). Beans: Origins and development. *Encyclopedia of Global Archaeology*, Springer, pages 822-827. https://doi.org/10.1007/978-1-4419-0465-

2_2169

Gepts, P., Aragão, F. J., de Barros, E., Blair, M. W., Brondani, R., Broughton, W., Galasso, I., Hernández, G., Kami, J., Lariguet, P., McClean, P., Melotto, M., Miklas, P., Pauls, P., Pedrosa-Harand, A., Porch, T., Sánchez, F., Sparvoli, F., Yu, k. (2008). Genome of Phaseolus beans, a major source of dietary protein and micronutrients in the tropics. In: Moore, P., Ming, R. (ed) *Plant Genetics and Genomics: Crops and Models*, Springer New York, pages 1:113-143. https://doi.org/10.1007/978-0-387-71219-2_5

Ghobary, H., & Abd Allah, S. (2010). Correlation and path- coefficient studies in common bean (*Phaseolus vulgaris* L.). *Journal of Plant Production*, 1(9), 1233–1239. https://doi.org/10.21608/jpp.2010.86576

Goodstein, M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., & Rokhsar, D. S. (2012). Phytozome: a comparative platform for green plants genomics. *Nucleic Acid Research* 40(D1): D1178-D1186. https://doi.org/10.1093/nar/gkr944

Greenlife (2023). French beans. Retrieved 14th August 2023 from the link https://www.greenlife.co.ke/french-beans/

Gujaria-Verma, N., Ramsay. L., Sharpe, A. G., Sanderson, L., Debouck, D. G., Tar'an, B., & Bett, K. E. (2016). Gene-based SNP discovery in tepary bean (*Phaseolus acutifolius*) and common bean (P. vulgaris) for diversity analysis and comparative mapping. *BMC Genomics*, 17, 239. https://doi.org/10.1186/s12864-016-2499-3

Güzel, D., & Sayar, S. (2012). Effect of cooking methods on selected physiochemical and nutritional properties of barlotto bean, chickpea, faba bean and white kidney bean. *Journal of Food Science and Technology*, 49(1):89-95. https://doi.org/10.1007/s13197-011-0260-0

Gyori, Z., Nemeskeri, E., & Szilagyi, S. (1998). Legumes grown under nonirrigated conditions. *Journal of Agricultural and Food Chemistry*, 46(8):3087-3091. https://doi.org/10.1021/jf980060c

Haddad, E. H., & Tanzman, J. S. (2003). What do vegetarians in the United States eat? The *American Journal of Clinical Nutrition*, 78(3):626S-632S. https://doi.org/10.1093/ajcn/78.3.626s

Henry, A. N., Evans, N. N., & Nelson, A. O. (2019). Evaluation of genetic variability, heritability, genetic advance and correlation for agronomic and yield components in common bean landraces from Southwestern Kenya. *Journal of Plant Breeding and Crop Science*, 11(5):44-147. https://doi.org/10.5897/jpbcs2018.0800

Hentges, L., Weaver, M., & Nielsen, S. (1990). Reversibility of the hard-to-cook defect in dry beans (*Phaseolus vulgaris* L.) and cowpeas (*Vigna unguiculata*). *Journal of Food Science*, 55(5):1474-1474. https://doi.org/10.1111/j.1365-2621.1990.tb03965.x

Hoyos-Villegas, V., Song, Q., & Kelly, J. (2017). Genome-wide association analysis for drought tolerance and associated traits in common bean. *The Plant Genome*, 10(1). https://doi.org/10.3835/plantgenome2015.12.0122

Hutchin, A.M., Winham, D.M., & Thompson, S.V. (2012). *Phaseolus* beans: Impact on glycaemic response and chronic disease risk in human subjects. *British Journal of Nutrition* 108(S1):52-65. https://doi.org/10.1017/s0007114512000761

Irving, G. C. (1980). *Inositol phosphates : their chemistry, biochemistry, and physiology*. (Ed) D.J. Cosgrove. D.J. Elsevier, pages 85-96. Retrieved from the link https://api.semanticscholar.org/CorpusID:82565617

Jacinto-Hernandez, C., Azpiroz-Rivero, S., Acosta-Gallegos, A., Hernandez-Sanchez, H., & Bernal-Lugo, I. (2003). Genetic analysis and random amplified polymorphic DNA markers associated with cooking time in common bean. *Crop Science*, 43(1):329. https://doi.org/10.2135/cropsci2003.3290

Jaetzold, R., & Schmidt, H. (1983). *Farm Management Handbook of Kenya*. Ministry of Agriculture, Kenya, in cooperation with the German Agricultural Team (GAT), German Agency for Technical Cooperation (GTZ) 2, 245-285. Retrieved from the link https://edepot.wur.nl/487562

Johnson, H. W., Robinson, H., & Comstock, R. (1955). Estimates of Genetic and Environmental Variability in Soybeans. *Agronomy Journal*, 47(7):314–318. https://doi.org/10.2134/agronj1955.00021962004700070009x

Jones, B., & Boulter, D. (1983). The analysis of development of hard bean during storage of black beans (*Phaseolus vulgaris* L.*). Qualitas Plantarum Plant Foods for Human Nutrition*, 33(1):77-85. https://doi.org/10.1007/bf01093740

Kamfwa, k., Cichy, K., & Kelly, J. (2015). Genome wide association study for agronomic traits in common bean. *The Plant Genome*, 8(2). https://doi.org/10.3835/plantgenome2014.09.0059

Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., Blois, H., Caig, V., Heller-Uszynska, K., Jaccoud, D., Hopper, C., Aschenbrenner-Kilian, M., Evers, M., Peng, K., Cayla, C., Hok, P., & Uszynski, G. (2012). Diversity Arrays Technology: a generic genome profiling technology on open platforms. *Data Production and Analysis in Population Genomics*, 67–89. https://doi.org/10.1007/978-1-61779-870-2_5

Kinyanjui, P., Njoroge, D., Makokha, A., Christiaens, S., Ndaka, D., & Hendricky, M. (2015). Hydration properties and texture fingerprints of easy- and hard-to-cook bean varieties. *Food Science and Nutrition*, 3(1):39-47. https://doi.org/10.1002/fsn3.188

Kinyuru, J., Kinyanjui, P., Muchui, M., & Mungai, H. (2011). Influence of Post-harvest Handling on the Quality of Snap Bean (Phaseolus vulgaris L.). Journal of Agriculture and Food Technology 1. Retrieved from the link https://www.researchgate.net/publication/266871777

Koinange, K., Singh, P., & Gepts, P. (1996). Genetic control of the domestication syndrome in common bean. *Crop Science*, 36(4):1037-1045. https://doi.org/10.2135/cropsci1996.0011183x003600040037x

Kon, S., & Sanshuck, D. (2007). Phytate content and its effect on cooking quality of beans. *Journal of Food Processing and Preservation*, 5. 169 - 178. 10.1111/j.1745-4549.1981.tb00632.x.

Langat, C., Omwoyo-Ombori, Cheruiyot. R., Gathaara, M., Karanja, D., Leley, P. (2019). Mapping of Quantitative Trait Loci (QTL) related to drought tolerance in common bean (*Phaseolus vulgaris* L.) using $f_2$ population from (KATB1 X GLP2). *International Annals of Science*, 8(1):75-86. https://doi.org/10.21467/ias.8.1.75-86

Lei, L., Wang, L., Wang, S., & Wu, J. (2020) Marker-trait association analysis of seed traits in accessions of common bean (*Phaseolus vulgaris* L.) in China. *Frontiers in Genetics*, 11:698. doi: 10.3389/fgene.2020.00698

Li, Y., Yang, K., Yang, W., Chu, L., Chen, C., Zhao, B., Li, Y., Jian, J., Yin, Z., Wang, T., & Wan, P. (2017). Identification of QTL and qualitative trait loci for agronomic

traits using SNP markers in the adzuki bean. *Frontiers in Plant Science*, 8: https://doi.org/10.3389/fpls.2017.00840

Li, Z., Yu, S., Lafitte, H., Huang, N., Courtois, B., Hittalmani, S., Vijayakumar, C., Liu, G., Wang, G., Shashidhar, H., Zhuang, J., Zheng, K., Singh, V., Sidhu, J., Srivantaneeyakul, S., & Khush, G. (2003). QTL × environment interactions in rice. I. Heading date and plant height. *Theoretical and Applied Genetics* 108(1):141-153. https://doi.org/10.1007/s00122-003-1401-2

Lobo, I. (2008) One gene can affect multiple traits. *Nature education* 1(1)10. Retrieved from https://www.nature.com/scitable/topicpage/pleiotropy-one-gene-can-affect-multiple-traits-569/

Maryange, M., Msolla, N., Sawargonkar, S., Hudge, B., & Thanki, H. (2010). Screening of 39 advanced beans (*Phaseolus vulgaris* L.) lines for short cooking time using two different methods. *Electronic Journal of Plant Breeding*, 4:505-511. Retrieved from https://www.researchgate.net/publication/46179223

Marzooghian, A., Moghaddam, M., Valizadeh, M., Kooshki, M. H. (2014). Genetic diversity of common bean genotypes as revealed by seed storage proteins and some agronomic traits. *Plant Breeding and Seed Science*, 67:125-137. https://doi.org/10.2478/v10129-011-0075-1

Mashi, D. S. (2006). Genetic studies on seed coat texture and cooking time in some varieties of cowpea (*Vigna unguiculata* L.). PhD Thesis, University of Jos, Ibadan.

Mattson, S. (1946). *The cookability of yellow peas. Acta Agriculturae Suecana*. The Royal Swedish Academy of Agriculture Stockholm, Pages 185-231. https://doi.org/10.1097/00010694-194807000-00009

Mbithi-Mwikya, S., Ooghe, W., Van Camp, J., Ngundi, D., & Huyghebaert, A. (2000). Amino acid profiles after sprouting, autoclaving, and lactic acid fermentation of finger millet (*Eleusine Coracan*) and kidney beans (*Phaseolus Vulgaris* L.). *Food Science and Technology*, 48(8):3081-3085. https://doi.org/10.1021/jf0002140

Mduruma, Z., Nchimbi-Msolla, S., & Reuberi, M. R. (1998). Evaluation of maturity characteristics and of yield components of high protein bean (*Phaseolus vulgaris* L.) varieties in Morogoro, Tanzania. *Tanzania Journal of Agricultural Sciences*, 1(2):131-140. Retrieved from the link https://www.ajol.info/index.php/tjags/article/view/116323

Mecha, E., Figueira, M., Vaz Patto, M., & Bronze, M. (2018). Two sides of the same coin: The impact of grain legumes on human health: Common bean (*Phaseolus vulgaris* L.) as a case study. *IntechOpen*, DOI:10.5772/intechopen.78737

Mendes, M., Botello, F., Ramalho, M., Abreu, A., & Furtini, I. (2008). Genetic control of the number of days to flowering in common bean. *Crop Breeding and Applied Biotechnology*, 8(4):279-282. https://doi.org/10.12702/1984-7033.v08n04a04

Mendoza, M., Karen, C., Sparague, C., Goffnet, A., Lu, R., & Kelly, J. (2018). Prediction of canned black bean texture (*Phaseolus vulgaris* L.) from intact dry seeds using visible/near infrared spectroscopy and hyper-spectral imaging data. *Journal of the Science of Food and Agriculture*, 98(1):283-290. https://doi.org/10.1002/jsfa.8469

Meng, L., Li, H., Zhang, L., & Wang, J. (2015). QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *The Crop Journal*, 3(3): 269-283. https://doi.org/10.1016/j.cj.2015.01.001

Milkas, N., Kelly, D., Beebe, E., & Blair, W. (2006). Common bean breeding for resistance against biotic and abiotic stresses: From classical to MAS breeding. *Euphytica* ,147(1-2):105-131. https://doi.org/10.1007/s10681-006-4600-5

Mitchell. D., Lawrence, F., Hartman, J., & Curran, J. (2009). Consumption of dry beans, peas and lentils could improve diet quality in the US Population. *Journal of American Diet Association*, 109(5): 909-913. https://doi.org/10.1016/j.jada.2009.02.029

Mitova, I., Dimitrov, I., Atanasova, E., Stancheva, I. (2008). Effects of fore-crop fertilization on the yield and quality of kidney beans under vegetable crop rotation conditions. *Acta Agronomica Hangarica*, 56(4): 449-454. https://doi.org/10.1556/aagr.56.2008.4.11

Molina, M., Baten, M., Gomez-Brenes, R., King, R., & Bressani, R. (1976). Heat treatment: Process to control the development of the hard-to-cook phenomenon in black beans (*Phaseolus vulgaris* L.). *Journal of Food Science*, 41(3):661-666. https://doi.org/10.1111/j.1365-2621.1976.tb00694.x

Moscoso, W., Bourne, C., & Hood, F. (1984). Relationships between the hard-to-cook Phenomenon in red kidney beans and water absorption, puncture force, pectin,

phytic acid, and minerals. *Journal of Food Science*, 49:1577-1583. https://doi.org/10.1111/j.1365-2621.1984.tb12848.x

Mughi, I. M. (2017). Inheritance and characterization of cooking time, seed iron and zinc content in selected African common bean germplasm. Dissertation, Makerere University.

Mukeshimana, G., Butare, L., Cregan, P., Blair, M., & Kelly, J. (2014). Quantitative trait loci associated with drought tolerance in common bean. *Crop Science*, 54(3):923-938. https://doi.org/10.2135/cropsci2013.06.0427

Muller, A., & Jung, C. (2009). Flowering time control and applications in plant breeding. *Trends in Plant Science*, 14(10):563–573. https://doi.org/10.1016/j.tplants.2009.07.005

Murube, E., Beleggia, R., Pacetti, D., Nartea, A., Frascarelli, G., & Lanzavecchia, G. (2021). Characterization of nutritional quality traits of a common bean germplasm collection. *Foods*, 10(7):1-26. https://doi.org/10.3390/foods10071572

Nadeem, M. Y., Mehmet, S., Muhammad, Q.H., Ephrem, Y., Hilal, A., Ahmad, H., Rustu, C., Yeter, K., Khalid, M.L., Ndiko, A., Muhammad, K., Tolga, B., & Faheem (2021). Common bean as a potential crop for future food security: an overview of past, current and future contributions in genomics, transcriptomics, transgenics and proteomics. *Biotechnology & Biotechnological Equipment*, 35. 758-786. DOI.10.1080/13102818.2021.1920462.

Nakedde, T., Ibara-Perez, F., Mukankusi, C., Waines, J., & Kelly, J. (2016). Mapping of QTL associated with fusarium root rot resistance and root architecture traits in black beans. *Euphytica*, 212(1):51-63. https://doi.org/10.1007/s10681-016-1755-6

Nchanji, E. B., & Ageyo, O. C. (2021). Do common beans (*Phaseolus vulgaris* L.) promote good health in humans? A systematic review and meta-analysis of clinical and randomized controlled trials. *Nutrients*, 13(11):3701. doi: 10.3390/nu13113701

Negahi, A., Bihamta, M. R., Negahi, Z., & Alidoust, M. (2014). Evaluation of genetic variation of some agronomical and morphological traits in Iranian and exotic common bean (Phaseolus vulgaris L.). *Agricultural Communications*, 2(3): 22-26. Retrieved from the link https://api.semanticscholar.org/CorpusID:55632633

Nemeskéri, E. (2012). *Relationship between nutritional quality of beans and growing conditions*. Beans: Nutrition, Consumption and Health. 249-264. Retrieved from the link https://www.researchgate.net/publication/287320796

Njoroge, M., Kinyanjui, K., Makokha, O., Christiane, S., Shpigelman, A., Sila, N., & Hendrickx, E. (2014). Extraction and characterisation of pectic polysaccharides from easy and hard-to-cook common beans (*Phaseolus vulgaris*). *Food Research International*, 64:314-322. https://doi.org/10.1016/j.foodres.2014.06.044

Nodari, R., Tsail, M., Gilbertson, R., & Gepts, P. (1993). Towards an integrated linkage map of common bean 2. Development of an RFLP-based linkage map. *Theoretical and Applied Genetics*, 85(5):513-520. https://doi.org/10.1007/bf00220907

Nyakuni, G. A., Kikafunda, J. K., Muyonga, J. H., Kyamuhangire, W. M., Nakimbugwe, D., & Ugen, M. (2008). Chemical and nutritional changes associated with the development of the hard-to-cook defect in common beans. *International Journal of Food Sciences and Nutrition*, 59(7-8):652–659. https://doi.org/10.1080/09637480701602886

Nyamboki, L. (2012). Environmental implications of household energy use. M.Sc. Thesis, Institute of Development Studies, University of Nairobi.

Oki, D, Tukamuhabwa, P., Odong, T., Namayanja, A., Mukabaranga, J., Paparu, P., & Gepts, P. (2014). Morphological diversity of tropical common bean germplasm. *African Crop Science Journal*, 22: 59-67. Retrieved form the link https://hdl.handle.net/1807/47678

Onyango, P. (2023). *Common bean breeding in Africa takes new and more rapid innovative techniques*. Retrieved 13th August 2023 from https://www.pabra-africa.org/common-bean-breeding-in-africa-takes-new-and-more-rapid-innovative-techniques/

Ousman, B. (2013). Non-destructive determination of bean water absorption capacity using CFA images analysis for hard-to-cook evaluation. *International Journal of Electrical and Computer Engineering*, 3(3). https://doi.org/10.11591/ijece.v3i2.2149

Perera, D., Devkota, L., Garnier, G., Panozzo, J., & Dhital, S. 2023. Hard-to-cook phenomenon in common legumes: Chemistry, mechanisms and utilization. *Food Chemistry*, 415(2):135743. DOI:10.1016/j.foodchem.2023.135743

Pérez-Vega, E., Pañeda, A., Rodríguez-Suárez, C., Campa, A., Giraldez, R., & Ferreira, J. (2010). Mapping of QTLs for morpho-agronomic and seed quality traits in a RIL population of common bean (*Phaseolus vulgaris* L.). *Theoretical and Applied Genetics*, 120(7):1367–1380. https://doi.org/10.1007/s00122-010-1261-5

Perseguini, J., Oblessuc, P., Rosa, J., Gomes, K., & Chiorato, A. (2016). Genome-wide association studies of anthracnose and angular leaf spot resistance in common bean (*Phaseolus vulgaris* L.). *PLoS One*, 11(1): e0150506. https://doi.org/10.1371/journal.pone.0150506

Petry, N., Erick, D., James, W., & Richard, H. (2015). Review: The potential of the common bean (*Phaseolus vulgaris*) as a vehicle for iron biofortification. *Nutrients*, 7(2):1144-1173. https://doi.org/10.3390/nu7021144

Phutela, U., Dhuna, V., Sandhu, S., & Chadha, B. (2005). Pectinase and polygalacturonase production by thermophilic aspergillus fumigatus isolated from decomposing orange peels. Brazilian. *Journal of Microbilogy*, 36:63-69. https://doi.org/10.1590/s1517-83822005000100013

Reddy, R., Pierson, D., Sathe, K., & Salunke, K. (1984). Chemical, nutritional and physiological aspects of dry bean carbohydrates: A review. *Food Chemistry*, 13(1):25-68. https://doi.org/10.1016/0308-8146(84)90026-8

Resende, R. T., Resende, M. D., Azevedo, C. F., Silva, F. F., Melo, L. C., Pereira, H. S., Souza, T., Valdisser, P.C., & Vianello, R. P. (2018). Genome-Wide Association and regional heritability mapping of plant architecture, lodging and productivity in *Phaseolus vulgaris*. *G3 Genes|Genomes|Genetics*, 8(8):2841-2854. https://doi.org/10.1534/g3.118.200493

Reynoso, C., Gomez, R., & Pina, L. (2006). Bioactive components in common beans (*Phaseolus vulgaris* L.). *Advances in Agriculture and Food Biotechnology*, 34: 217-236.

Richardson, C., & Stanley, W. (1991). Relationship of loss of membrane functionality and hard-to-cook defect in aged beans. *Journal of Food Science*, 56(2):590-591. https://doi.org/10.1111/j.1365-2621.1991.tb05334.x

Saha, S., Singh, G., Mahajan, V., & Gupta, H. (2009). Variability of nutritional and cooking quality in bean (Phaseolus vulgaris L) as a function of genotype. *Plant Foods for Human Nutrition*, 64 (2):174-180. https://doi.org/10.1007/s11130-009-0121-4

Sakuraba, Y., Jeong, J., Kang, M. Y., Kim, J., Paek, N. C., & Choi G (2014). Phytochrome interacting transcription factors PIF4 and PIF5 induce leaf senescence in Arabidopsis. *Nature Communications*, 5(1). https://doi.org/10.1038/ncomms5636

Sandhu, S., You, M., Conner, L., Balasubramanian, M., & Hou, A. (2018). Genetic analysis and QTL mapping of the seed hardness trait in a black common bean (*Phaseolus vulgaris*) recombinant inbred line (RIL) population. *Molecular Breeding*, 38(3). https://doi.org/10.1007/s11032-018-0789-y

Schmutz, J., McClean, P. E., Mamidi, S., Wu, G., Cannon, S., Grimwood, J., Jenkins, J., Shu, S., Song, Q., Chavarro, C., Torres-Torres, M., Geffroy, V., Moghaddam, S., Gao, D., Abernathy, B., Barry, K., Blair, M., Brick, M. A., Chovatia, M., Gepts, P., Goodstein, D. M., Gonzales, M., Hellsten, U., Hyten, D. L., Jia, G., Kelly, J. D., Kudrna, D., Lee, R., Richard, M., Miklas, P., Osorno, J. M., Rodrigues, J., Thareau, V., Urrea, C. A., Wang, M., Yu, Y., Zhang, M., Wing, R.A., Cregan, P. B., Rokhsar, D. S., & Jackson, S. A. (2014). A reference genome for common bean and genome-wide analysis of dual domestications. *Nature Genetics*, 46(7):707-713. https://doi.org/10.1038/ng.3008

Sgarbieri, C., & Whitaker, R. (1982). Physical, chemical, and nutritional properties of common bean (*Phaseolus*) proteins. *Advances in Food Research*, 28(3):93-166. https://doi.org/10.1016/s0065-2628(08)60111-1

Shehata, E. A. (1992). Hard-to-cook phenomenon in legumes. *Food Reviews International*, 8(2):191-221. https://doi.org/10.1080/87559129209540938

Shi, Q., Zhang, H., Song, X., Jiang, Y., Liang, R., & Li, G. (2018). Functional characterization of the maize phytochrome-interacting factors PIF4 and PIF5. *Frontiers in Plant Science*, 8. https://doi.org/10.3389/fpls.2017.02273

Shomer, I., Paster, H., Linder, P., & Vasiliver, R. (1990). The role of cell wall structure in the hard-to-cook phenomenon in beans (*Phaseolus vulgaris* L.) *Journal of Food Structure*, 9(2):10. Retrieved from the link https://digitalcommons.usu.edu/

foodmicrostructure/vol9/iss2/10

Silva, V., Menezes-Junior, J., Carneiro, P., Carneiro, J., Cruz, C. (2013). Genetic improvement of plant architecture in the common bean. *Genetics and Molecular Research*, 12(AOP). https://doi.org/10.4238/2013.january.30.8

Singh, R. K., & Chaudhary, B. D. (1985). *Biometrical methods in quantitative genetic analysis*. Kalyani Publishers, New Delhi, India pages 78-101.

Slupski, J. (2010). Effect of cooking and sterilisation on the composition of amino acids in immature seeds of flageolet bean (*Phaseolus vulgaris* L.) cultivars. *Food Chemistry*, 121(4):1171-1176. https://doi.org/10.1016/j.foodchem.2010.01.066

Swema, E., & Mwinuka, L. (2021). Understanding consumer preference for common beans from manifold viewpoints of attributes in Dar es Salaam, Tanzania. *Theoretical Economics Letters*, 11, 1038-1054. https://doi.org/10.4236/tel.2021.115066

Taptue, C. (2018). Nutritional properties and nutrients chemical analysis of common beans seed. *MedCrave Online Journal of Biology and Medicine*, 3. 10.15406/mojbm.2018.03.00074.

Tar'an, B., Michaels, T., & Pauls, K. P. (2002). Genetic mapping of agronomic traits in common bean. *Crop Science*, 42(2):544–556. https://doi.org/10.2135/cropsci2002.5440

Tar'an, B., Michaels, T. E., & Pauls, K. P. (2002). Genetic mapping of agronomic traits in common bean. *Crop Science*, 42: 544–556. doi: 10.2135/cropsci2002.0544

Thorwarth, P., Yourself, A., & Schmid, K. (2017). Genomic prediction and association mapping of curd-related traits in gene bank accessions of cauliflower. *G3: Genes, Genomes, Genetics* 8. https://doi.org/10.1534/g3.117.300199

Tosh, S., & Yada, S. (2010). Dietary fibres in pulse seeds and fractions: Characterization, functional attributes, and applications. *Food Research International*, 43(2):450-460. https://doi.org/10.1016/j.foodres.2009.09.005

Ugwuanyi, S., Udengwu, O. S., Snowdon, R. J., & Obermeier, C. (2022). Novel candidate loci for morpho-agronomic and seed quality traits detected by targeted genotyping-by-sequencing in common bean. *Frontier in Plant Science*, 13:1014282. doi: 10.3389/fpls.2022.1014282

Valdisser, R., Pappas, G., Menezes, I., Müller, B., Pereira, W., Narciso, M., Brondani, C., Souza, T., Borba, T., & Vianello, R. (2016). SNP discovery in common bean by restriction-associated DNA (RAD) sequencing for genetic diversity and population structure analysis. *Genetics and Genomics*, 291(3):1277-1291. https://doi.org/10.1007/s00438-016-1182-3

Van der Poel, A. F. B. (1990). Effect of processing on antinutritional factors and proteins nutritional value of dry beans (*Phaseolus vulgaris* L.). *Animal Feed Science and Technology*, 29(3-4):179-208. https://doi.org/10.1016/0377-8401(90)90027-6

Van Schoonhoven, A., & Pastor-Corrales, M. A. (1987). *Standard system for the evaluation of bean germplasm*. Cali, Colombia. 54pages

Vandemark, G. J., Brick, M. A., Orsono, J. M., Kelly, J. D., & Urrea, C. A. (2014). Edible Grain Legumes. In: Smith S, Diers B, Specht J, Carver B (ed) *Yield Gains in Major U.S. Field Crops.* Madison pp 87-124. https://doi.org/10.2135/cssaspecpub33.c5

Villordo-Pineda, E., González-Chavira, M. M., Giraldo-Carbajo, P., Acosta-Gallegos, J., & Caballero-Pérez, J. (2015) Identification of novel drought-tolerant-associated SNPs in common bean (*Phaseolus vulgaris*). *Frontiers in Plant Science*, 6. https://doi.org/10.3389/fpls.2015.00546

Vindiola, L., Seib, A., & Hoseney, C. (1986). Accelerated development of hard-to-cook state in beans. *Cereal Foods World*, 31(8):608. Retrieved from the link https://eurekamag.com/research/004/650/004650915.php

Virginia, M. (2014). Nutritional and health benefits of dry beans. *The American Journal of Clinical Nutrition*, 100:437-442. DOI: 10.3945/ajcn.113.071472

Viteri, D.M., Terán, H., Asensio-S.-Manzanera, M.C., Asensio, C., Porch, T.G., Miklas, P.N., & Singh S.P. (2014). Progress in Breeding Andean Common Bean for Resistance to Common Bacterial Blight. *Crop Science*, 54:2084–2092. doi: 10.2135/cropsci2014.03.0177

Wacu, V. K. (2016). Effect of storage conditions and pre-processing treatments on physical, cooking and nutritional properties of two common beans varieties in Kenya. M.Sc. Thesis, Food science and Nutrition JKUAT.

Wafula, E. N., Wainaina, I. N., Buve, C., Nguyen, N., Kinyanjui, P. K., Saeys, W., Sila, D. N., & Hendrickx, M. (2020). Application of near-infrared spectroscopy to predict the cooking times of common beans (*Phaseolus vulgaris* L.). *Journal of Food Engineering*, 284:110056. https://doi.org/10.1016/j.jfoodeng.2020.110056

Wahome, S., Kimani, P., Muthomi, J., & Narla, R. (2013). Quality and yield of snap bean lines locally developed in Kenya. *International Journal of Agronomy and Agricultural Research*, 3(7):1-10. Retrieved from the link https://innspub.net/quality-and-yield-of-snap-bean-lines-locally-developed-in-kenya/

Wairimu, N. G. (2015). Analysis of consumer preference in product attributes: a case study of common beans in Kiambu county, Kenya. M.Sc. Thesis, Agribusiness management and Trade, Kenyatta University.

Wang, N., & Daun, K. (2005). Determination of cooking times of pulses using an automated Mattson cooker apparatus. *Journal of the Science of Food and Agriculture*, 85(10):1631-1635. https://doi.org/10.1002/jsfa.2134

Wang, S., Chen, L., Yang, H., Gu, J., Wang, J., & Ren, F. (2020). Regular intake of white kidney beans extract (*Phaseolus vulgaris* L.) induces weight loss compared to placebo in obese human subjects. *Food Science and Nutrition*, 8:1315–1324. doi: 10.1002/fsn3.1299

Wortmann, C. S., Kirkby, R. A., Eledu, C. A., & Allen, D. J. (1998). Atlas of common bean (*Phaseolus vulgaris* L.) production in Africa. *Centro Internacional de Agricultura Tropical*, (CIAT), Cali, Colombia 297:131 pages. Retrieved from the link https://www.researchgate.net/publication/348860644

Yohannes, S., Loha, G., & Gessese, M. K. (2020). Performance evaluation of common bean (*Phaseolus vulgaris* L.) genotypes for yield and related traits at Areka, southern Ethiopia. *Advances in Agriculture*, 2020:1-8. https://doi.org/10.1155/2020/1497530

Zaminder, N., Baghekhandan, S., Nasirpour, A., & Sheikhzeinoddin, M. (2013). Effect of lime, soaking and cooking on water absorption texture and splitting of red kidney beans. *Journal of Food Science and Technology*, 50(1):108-114. https://doi.org/10.1007/s13197-011-0234-2

Zhang, X., Blair, M. W., & Wang, S. (2008). Genetic diversity of Chinese common

bean (*Phaseolus vulgaris* L.) landraces assessed with simple sequence repeat markers. *Theoretical and Applied Genetics*, 117(4):629-640. https://doi.org/10.1007/s00122-008-0807-2

Zilio, M., Souza, A. C., Coelho, M. C., & Jose, D. (2014). The genotype and crop environment effect on the technological quality of common beans grains. *American-Eurasian Journal of Agriculture and Environmental Science*, 14:212-220. DOI: 10.5829/idosi.aejaes.2014.14.03.12315

Zuiderveen, G., Padder, B., Kamfwa, K., Song, Q., & Kelly, J. (2016). Genome-wide association study of anthracnose resistance in Andean beans (*Phaseolus vulgaris* L.). *PLoS One*, 11(6): e015639. https://doi.org/10.1371/journal.pone.0156391

## APPENDICES

**Appendix I: Analysis of variance for agronomic traits and cooking time of common bean accession grown during year 2019/2020**

| Sources of Variation | DF | SS | MS | F-value | Pr (>F) | |
|---|---|---|---|---|---|---|
| **(a) Days to flowering** | | | | | | |
| Block | 1 | 97.9 | 97.9 | 25.5 | 5.35E-07 | *** |
| Seasons | 3 | 3741.4 | 1247.1 | 324.2 | < 2.2e-16 | *** |
| Accessions | 256 | 17172.3 | 67.1 | 17.4 | < 2.2e-16 | *** |
| Season X Accession | 768 | 8689.2 | 11.3 | 2.9 | < 2.2e-16 | *** |
| Residuals | 1027 | 3950.2 | 3.9 | | | |
| **(b) Number of number pods per plant** | | | | | | |
| Block | 1 | 3170.8 | 3170.8 | 157.1 | < 2.2e-16 | *** |
| Seasons | 2 | 28416.1 | 14208 | 704.1 | < 2.2e-16 | *** |
| Accessions | 256 | 24841.4 | 97 | 4.8 | < 2.2e-16 | *** |
| Season X Accession | 512 | 15160.5 | 29.6 | 1.5 | 7.73E-07 | *** |
| Residuals | 770 | 15538.8 | 20.2 | | | |
| **(c) Days to maturity** | | | | | | |
| Block | 1 | 33.6 | 33.6 | 2.8 | 0.09 | |
| Seasons | 3 | 11008.9 | 3669.6 | 307.6 | < 2e-16 | *** |
| Accessions | 256 | 28478.5 | 111.2 | 9.3 | < 2e-16 | *** |
| Season X Accession | 768 | 16235.9 | 21.1 | 1.8 | < 2e-16 | *** |
| Residuals | 1027 | 12240.5 | 11.9 | | | |
| **(d) Pods length (cm)** | | | | | | |
| Block | 1 | 44.7 | 44.7 | 23.2 | 1.70e-06 | *** |
| Seasons | 3 | 194.7 | 64.9 | 33.7 | < 2.2e-16 | *** |
| Accessions | 256 | 4849.7 | 18.9 | 9.8 | < 2.2e-16 | *** |
| Season X Accession | 768 | 1920.4 | 2.5 | 1.3 | 5.43E-05 | *** |
| Residuals | 1027 | 1978.5 | 1.9 | | | |
| **(e) Number of seeds per pod** | | | | | | |
| Block | 1 | 5.2 | 5.2 | 9.1 | 2.66E-03 | ** |
| Seasons | 3 | 226.4 | 75.5 | 131.6 | < 2.2e-16 | *** |
| Accessions | 256 | 885.1 | 3.5 | 6.0 | < 2.2e-16 | *** |
| Season X Accession | 768 | 630.2 | 0.8 | 1.4 | 4.45E-08 | *** |
| Residuals | 1027 | 588.2 | 0.6 | | | |
| **(f) Number of 100-seed weight** | | | | | | |
| Block | 1 | 56 | 56.2 | 2.8 | 9.31e-02 | |
| Seasons | 3 | 1272 | 424.1 | 21.3 | 2.12e-13 | *** |
| Accessions | 256 | 333195 | 1301.5 | 65.4 | < 2.2e-16 | *** |
| Season X Accession | 768 | 22658 | 29.5 | 1.5 | 2.04e-09 | *** |
| Residuals | 1027 | 20429 | 19.9 | | | |
| **(g) Seed yield ($Kgha^{-1}$)** | | | | | | |
| Block | 1 | 17025456 | 17025456 | 101.5 | < 2.2e-16 | *** |
| Seasons | 3 | 2.15E+08 | 71647938 | 427.2 | < 2.2e-16 | *** |
| Accessions | 256 | 5.69E+08 | 2222240 | 13.3 | < 2.2e-16 | *** |
| Season X Accession | 768 | 4.21E+08 | 548375 | 3.3 | < 2.2e-16 | *** |

| Sources of Variation | DF | SS | MS | F-value | Pr (>F) | |
|---|---|---|---|---|---|---|
| Residuals | 1027 | 1.72E+08 | 167714 | | | |
| **(h) Cooking time of fresh and aged seeds (min)** | | | | | | |
| Block | 1 | 8.0 | 7.7 | 0.5 | 4.65E-01 | |
| Storage | 1 | 26823.0 | 26822.9 | 1860.8 | <2e-16 | *** |
| Accessions | 230 | 76213.0 | 331.4 | 23.0 | <2e-16 | *** |
| Accession X Storage | 230 | 31164.0 | 135.5 | 9.4 | <2e-16 | *** |
| Residuals | 461 | 6645.0 | 14.4 | | | |

DF=Degree of freedom, SS=Sum of squares, MS=Mean sum squares, , '*' '**'***= Significant at P≤0.05, P≤0.01 and P≤0.001 respectively

**Appendix II: Analysis of variance for agronomic traits and cooking time of recombinant lines derived from a cross of GLPx92 (pinto) X GLP2 (rosecoco)**

| Sources of Variation | DF | SS | MS | F-value | Pr (>F) | |
|---|---|---|---|---|---|---|
| **(a) Days to flowering** | | | | | | |
| Block | 1 | 1.6 | 1.6 | 1.2 | 2.79e-01 | |
| Maternal effect | 1 | 108.7 | 108.7 | 80.7 | 3.19e-16 | *** |
| RIL | 189 | 3859.6 | 20.4 | 15.2 | < 2.2e-16 | *** |
| Residuals | 183 | 246.5 | 246.5 | 1.347 | | |
| **(b) Days to maturity** | | | | | | |
| Block | 1 | 34.4 | 34.4 | 9.9 | 1.91e-03 | ** |
| Maternal effect | 1 | 66.9 | 66.9 | 19.3 | 1.89e-05 | *** |
| RIL | 189 | 6537.7 | 34.6 | 10.0 | < 2.2e-16 | *** |
| Residuals | 183 | 634 | 3.5 | | | |
| **(c) Number of pods per plant** | | | | | | |
| Block | 1 | 1162.5 | 1162.5 | 155.9 | < 2.2e-16 | *** |
| Maternal effect | 1 | 42.6 | 42.7 | 5.7 | 1.78e-02 | * |
| RIL | 189 | 2984.3 | 15.8 | 2.1 | 2.35e-07 | *** |
| Residuals | 183 | 1364.9 | 7.5 | | | |
| **(d) Pod length (cm)** | | | | | | |
| Block | 1 | 5.34 | 5.3 | 13.2 | 3.58e-04 | *** |
| Maternal effect | 1 | 0.59 | 0.6 | 1.5 | 2.28e-01 | |
| RIL | 189 | 671.2 | 3.6 | 8.8 | < 2.2e-16 | *** |
| Residuals | 183 | 73.84 | 0.4 | | | |
| **(e) Number of seeds per pod** | | | | | | |
| Block | 1 | 1.723 | 1.7 | 9.1 | 2.95E-03 | ** |
| Maternal effect | 1 | 0.428 | 0.4 | 2.3 | 1.35E-01 | |
| RIL | 189 | 194.3 | 1.0 | 5.4 | < 2.2e-16 | *** |
| Residuals | 183 | 34.7 | 0.2 | | | |
| **(f) 100 seed weight (g)** | | | | | | |
| Block | 1 | 39.4 | 39.4 | 15.8 | 0.0 | *** |
| Maternal effect | 1 | 20.9 | 20.9 | 8.4 | 0.0 | ** |
| RIL | 189 | 20311.4 | 107.5 | 43.0 | < 2.2e-16 | *** |
| Residuals | 183 | 457.1 | 2.5 | | | |

| Sources of Variation | DF | SS | MS | F-value | Pr (>F) | |
|---|---|---|---|---|---|---|
| **(g) Cooking time (min)** | | | | | | |
| Block | 1 | 5 | 5.4 | 0.8 | 3.71e-01 | |
| Maternal effect | 1 | 127 | 127.2 | 19.0 | 2.16e-05 | *** |
| RIL | 193 | 42572 | 220.6 | 32.9 | < 2.2e-16 | *** |
| Residuals | 193 | 1294 | 9.1 | | | |

DF=Degree of freedom, SS=Sum of squares, MS=Mean sum squares, '*' '**'***= Significant at P≤0.05, P≤0.01 and P≤0.001 respectively

**Appendix III: Analysis of variance for agronomic traits and cooking time of recombinant inbred lines derived from a cross of GLPx92 (pinto) X GBK035420 (black coloured)**

| Sources of Variation | DF | SS | MS | F-value | Pr (>F) | |
|---|---|---|---|---|---|---|
| **(a) Days to Flowering** | | | | | | |
| Block | 1 | 0.36 | 0.4 | 0.4 | 5.38E-01 | |
| Maternal effect | 1 | 9.41 | 9.4 | 9.9 | 1.98E-03 | ** |
| RIL | 166 | 2936.5 | 17.7 | 18.6 | < 2.2e-16 | *** |
| Residuals | 165 | 157.14 | 1.0 | | | |
| **(b) Days to maturity** | | | | | | |
| Block | 1 | 41.7 | 41.7 | 26.6 | 7.18E-07 | *** |
| Maternal effect | 1 | 39.2 | 39.2 | 25.0 | 1.48E-06 | *** |
| RIL | 166 | 3875.3 | 23.3 | 14.9 | < 2.2e-16 | *** |
| Residuals | 165 | 258.8 | 1.6 | | | |
| **(c) Number of pods/plant** | | | | | | |
| Block | 1 | 763.5 | 763.6 | 64.9 | 1.52E-13 | *** |
| Maternal effect | 1 | 88.3 | 88.3 | 7.5 | 6.84E-03 | ** |
| RIL | 166 | 4168.4 | 25.1 | 2.1 | 7.66E-07 | *** |
| Residuals | 165 | 1942.5 | 11.8 | | | |
| **(d) Pod length (cm)** | | | | | | |
| Block | 1 | 2.04 | 2.0 | 10.8 | 1.26E-03 | ** |
| Maternal effect | 1 | 0.165 | 0.2 | 0.9 | 3.52E-01 | |
| RIL | 166 | 140.674 | 0.8 | 4.5 | < 2.2e-16 | *** |
| Residuals | 165 | 31.249 | 0.2 | | | |
| **(e) Number of seeds per pod** | | | | | | |
| Block | 1 | 2.166 | 2.2 | 15.2 | 1.41E-04 | *** |
| Maternal effect | 1 | 0.955 | 1.0 | 6.7 | 1.05E-02 | * |
| RIL | 166 | 95.758 | 0.6 | 4.0 | < 2.2e-16 | *** |
| Residuals | 165 | 23.522 | 0.1 | | | |
| **(f) 100-seed weight (g)** | | | | | | |
| Block | 1 | 122.4 | 122.4 | 69.7 | 2.63E-14 | *** |
| Maternal effect | 1 | 256.3 | 256.3 | 146.0 | < 2.2e-16 | *** |
| RIL | 166 | 4650.6 | 28.0 | 16.0 | < 2.2e-16 | *** |
| Residuals | 165 | 289.6 | 1.8 | | | |

| Sources of Variation | DF | SS | MS | F-value | Pr (>F) | |
|---|---|---|---|---|---|---|
| **(g) Cooking time (min)** | | | | | | |
| Block | 1 | 0 | 0.0 | 0.0 | 9.48E-01 | |
| Maternal effect | 1 | 37.4 | 37.4 | 5.9 | 1.60E-02 | * |
| RIL | 198 | 29202.1 | 147.5 | 23.3 | < 2.2e-16 | *** |
| Residuals | 197 | 1247.2 | 6.3 | | | |

DF=Degree of freedom, SS=Sum of squares, MS=Mean sum squares, '*' '**'***= Significant at P≤0.05, P≤0.01 and P≤0.001 respectively, RIL=Recombinant inbred lines.

## Appendix IV: Average rainfall and temperature for Juja area in the year 2018/2019

| | Rainfall (mm) | | Temperature (ºC) | |
|---|---|---|---|---|
| **Month** | **2018** | **2019** | **2018** | **2019** |
| 1 | 38.0 | 34.0 | 21.7 | 21.1 |
| 2 | 27.0 | 17.8 | 21.1 | 19.7 |
| 3 | 33.0 | 36.0 | 20.1 | 20.2 |
| 4 | 60.0 | 63.9 | 20.5 | 20.3 |
| 5 | 45.0 | 48.6 | 19.5 | 20.0 |
| 6 | 11.0 | 6.8 | 18.9 | 17.9 |
| 7 | 8.0 | 10.0 | 16.5 | 17.0 |
| 8 | 7.0 | 8.0 | 17.0 | 18.0 |
| 9 | 4.0 | 5.0 | 19.0 | 19.4 |
| 10 | 30.0 | 27.0 | 19.5 | 20.5 |
| 11 | 43.0 | 62.2 | 19.4 | 20.0 |
| 12 | 27.0 | 24.5 | 19.2 | 19.5 |