Genome-Wide Association Study of Multi-Traits in Common Bean (Phaseolus Vulgaris L.)

Papias Hongera Binagwa (hongera1984@yahoo.com)  
Tuskegee University  https://orcid.org/0000-0001-6735-6325

Conrad Bonsi  
Tuskegee University

Guohao He  
Tuskegee University

Sy M Traore  
Tuskegee University

Xiangyu Chen  
Fujian Academy of Agriculture Sciences: Fujian Academy of Agricultural Sciences

Eunice Bonsi  
Tuskegee University

Binita Subedi  
Tuskegee University

Jesse Jaynes  
Tuskegee University

Research article

Keywords: GWAS, Zinc, Iron, micronutrient deficiencies, genomic region, SNPs.

DOI: https://doi.org/10.21203/rs.3.rs-78845/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

**Background:** Micronutrient deficiencies caused by lack of Iron (Fe), Zinc (Zn) and Vitamin A have negative effects on human health worldwide. Even though common bean meets Fe and Zn deficiencies based on recommended dietary allowances, the cooking time remains an important challenge. Understanding the genome organization, the loci and genes localization controlling multiple traits is crucial for developing enhanced micronutrient content and improved common bean with short cooking time.

**Results:** In this study, GWAS method was used to determine SNPs associated with four traits studied, including water uptake, cooking time, Fe and Zn concentration level, using ~5,000 SNPs for a panel of 206 genotypes. By the analysis of population structure, 206 common bean genotypes were formed in three groups, the Mesoamerican gene pool, the Andean gene pool, and the admixture gene pool using the Bayesian method. With the significance levels between $p = 2.94 \times 10^{-08}$ to $p = 8.33 \times 10^{-03}$ and phenotypic variation of 8.26% to 18.19%, total of 10 SNPs was found to be significantly associated with traits on chromosomes Pv03, Pv04 and Pv10. Among of these SNPs, four SNPs were associated with water uptake, another four SNPs with the cooking time, one SNP with Zn concentration and one of less significant SNP associated with Fe concentration. Within the range or nearby the peak of significant SNP, putative loci encoding pectin degradation, cell structure and cellular exchange/transport were identified, suggesting these loci may play important roles in modulating expression of traits.

**Conclusion:** The resultant data in this study contributed on deciphering the molecular mechanisms associated with water uptake, cooking time and micronutrient contents in common bean. The investigation of association between phenotypic and genotypic data has showed that GWAS approach is a powerful tool to dissect genetic architecture of complex traits and cooperate molecular breeding for advanced genotypes with enhanced micronutrients and short cooking time.

**Background**

Common bean (*Phaseolus vulgaris* L.) was domesticated in two centers, the Andean Center of Domestication (ACD) or andean gene pool and Mesoamerican Center of Domestication (MCD) or Mesoamerican gene pool from Andes mountains of Peru and Lerma-Santiago basin of Mexico respectively [1]. Common bean is one of the most important legume crops that contributes to human nutrition [2]. It provides an important source of nutrients for more than 300 million people in different parts of the world, including Eastern Africa and Latin America where it represents 65% of total protein consumed, and 32% of energy [3], [4]. It is also the major source of micronutrients including Iron (Fe) and Zinc (Zn) [5] that are essential for the normal growth, development, reproduction and other physiological functions in the human body [6]. Fe and Zn bioavailability may be impeded by phytate content in common bean even though they provide a rich source of Fe and Zn. Some studies showed that the iron bioavailability, i.e. the amount that a person would absorb, is higher in the quicker cooking beans from different bean market classes [7] as micronutrients gradually decreased with the increasing of cooking time [8]. In order to improve micronutrient bioavailability, pre-soaking, germination and or fermentation of
seeds in water for at least 12 hours are adopted to reduce the phytic acids [9] for accelerating the cooking time. It is believed that fast cooking with enhanced micronutrient bioavailability has impact on bean consumption.

More than 2 billion people around the world are suffering from vitamin and micronutrient deficiencies [10]. Imbalances of these nutrients lead to a significant risk of illness and mortality among children under five years of age, pregnant women and lactating mothers [11]. Common beans provide a rich source of Fe and Zn based on recommended dietary allowances; however, they often take a long time to cook [12]. Many communities around the world have limited fuels for cooking. They mainly rely on burning wood, charcoal or other available source of biofuels. Cooking of common bean is time consuming; therefore, leading to deforestation in areas using burning wood as main source of energy [13]. Traits of cooking time and micronutrient contents in common bean varies as changes in physical and chemical property (i.e. seed size, thickness and seed coat color) and environment (relative humidity and storage temperature) [14], which is genetically regulated. Understanding the genomic regions related to the variation of cooking time and micronutrient contents allows to genetically manipulate these traits and helps to improve this crop with enhanced micronutrient content and short cooking time, which may lead to higher consumption.

Genome-wide association study (GWAS) has been used to determine genomic variants associated with traits of interest using recombinant inbred line population or landraces population [15], [16]. Literatures have demonstrated that some genomic regions have been identified to be associated with traits, such as bruchid resistance [17], cooking time of andean diversity panel [18], agronomic traits [19], drought tolerance [20], anthracnose and angular leaf spot diseases [21] and symbiotic nitrogen fixation [22] in common bean. However, there is a limited information on grain micronutrient contents of Fe and Zn as well as seed water imbibition during pre-soaking and cooking time. Therefore, the aim of this study was to dissect these complex traits, identify genomic regions associated with variations of traits, and reveal putative genes mediating traits using the GWAS technology in a diverse population of common bean.

**Results**

**Phenotyping multiple traits**

There were large variations in traits of water uptake and cooking time among genotypes studied. The initial moisture level across the samples ranged between 10-14% (data not shown). After soaking in water overnight, the maximum water uptake was 61.54% and minimum was 3.7% in this collection of genotypes (Figure 1a). The longest cooking time was 76 minutes and 15 seconds while the shortest was 14 minutes and 17 seconds (Figure 1b). Negative correlation of $r = -0.20$ at $p<0.001$ was observed between water uptake and cooking time among the common bean genotypes.

Traits for seed micronutrient contents of Fe and Zn showed significant difference ($p<0.001$) among the genotypes and environments. The concentration of Fe varied from 47.00 to 112.70 mg/kg (Figure 1c) while Zn concentration had a range of 21.73 to 42.50 mg/kg (Figure 1d). There was a positive correlation of $r = 0.51$ at $p<0.001$ between Fe and Zn contents. The positive correlation illustrated that Fe
concentration increased as Zn increased in both locations, suggesting micronutrients Fe and Zn might share a similar metabolic pathway, such as transport and translocation, for uptake and assimilation of Fe and Zn.

**Population structure**

A large number of genome-wide DNA markers could cover most recombination variation in a natural population. In this study, the SNP array containing 5052 SNP markers was used to assess population structure through the principal component analysis. The collection of 206 common bean genotypes is divided into three clusters (K=3) using the Bayesian method, in which the first cluster (K=1) consisted of 113 genotypes belonging to the Mesoamerican gene pool; the second cluster (K=2) consisted of 72 genotypes belonging to the Andean gene pool and the third cluster (K=3) consisted of 21 genotypes belonging to the admixture gene pool (Figures 2). Since this collection contained some breeding lines derived from different sources of parents, there was a possibility that offspring shared Mesoamerican and Andean gene pools leading to these “mosaic” chromosomes in the third cluster.

To evaluate the difference of micronutrient concentration levels among gene pools, the concentrations of Fe or Zn of genotypes in each group were calculated for their median concentration at different locations. The result demonstrated that the median concentration of Fe or Zn in Mesoamerican and Andean gene pools was not significantly different, except that the concentration of Zn was significant difference between two gene pools at SARI. However, the median concentration of Fe and Zn in the admixture group was significantly similar to Mesoamerican gene pool than to Andean gene pool depended on the location plants grow (Figure 3).

**Marker-trait associations**

**Water uptake and cooking time**

The water uptake of a dried mature seed represents the capability of seed imbibition that is related with the cooking time, i.e. higher water uptake, shorter cooking time. A total of 10 significant SNPs associated with water uptake was identified on the Pv03, Pv04, and Pv10 (Figure 4). The most significant SNP was ss715649433 ($p = 2.94 \times 10^{-08}$) located on Pv04, which accounted for 18.19% of phenotypic variability (Table 1). For cooking time trait, 11 SNPs were detected on the Pv03, Pv04, and Pv10 and explained about 9.31% - 10.06% phenotypic variations and the highest significant SNP was ss715643803 ($p = 2.14 \times 10^{-05}$) on Pv03.

**Micronutrient concentration**

Micronutrients are essential for human growth and play an important role in balanced human health. Common bean provides a rich source of Fe and Zn for human. In this study, Fe and Zn considered as significant invisible traits in common bean genotypes. Significant SNPs were identified for Zn on the Pv04 (Figure 4) and the most significant ($p = 5.98 \times 10^{-05}$) SNP ss715646134 explained 11.09% of phenotypic
variation (Table 1). Significant SNPs were identified for Fe on Pv10 with the most significant SNP (ss715646330) at $p = 8.33 \times 10^{-03}$ explaining 8.26% of phenotypic variation (Table 1).

### Putative causative loci

Traits evaluated in this study displayed a quantitative feature as continuous distribution of values in this collection. The GWAS approach was used to dissect genetic architectures of these complex traits and to uncover the putative causative loci for traits. Significant SNPs identified on chromosome regions associated with various traits facilitate the identification of putative causative loci nearby SNP peaks. In this study, several putative candidate genes related to traits studied were identified at the flanking sequence of significant SNPs. For instance, nuclear pore complex protein GP210 (PHVU_004G008000g), galactinol-sucrose galactosytransferase 5 (PHVU_004G0071000g), and glycosyltransferase BC10 (PHVU_004G007000g) were linked to significant SNPs located on Pv04 for the trait of water uptake; oligouidylate-binding protein 1-like (PHVU_003G129100g), probable polygalacturonase (PHVU_003G129300g), and ubiquinone biosynthesis protein COQ9 (PHVU_003G128500g) were near significant SNPs in chromosome 3 for cooking time; transmembrane 9 superfamily member 3-like (PHVU_004G145000g); transcriptional factor MYB52 (PHVU_004G144900g); and probable cyclic nucleotide-gated ion channel 20 (PHVU_004G144200g) were associated with Zn concentration within significant SNP peak on Pv04. Several putative causative loci underlying traits were listed in Table 2.

### Discussion

Genetic diversity of common bean provides a source of variations in various agronomic traits for improvement of varieties. The genetic variation characterized by different molecular markers can be utilized in plant breeding via marker-assisted selection to pyramid favorable loci into a cultivar. Two gene pools are available in common bean and most of the genotypes collected in this study belongs to Mesoamerican gene pool. However, a separated group of genotypes (admixture) was observed between the two gene pools after the analysis of population structure. This admixture group consists of breeding lines that stemmed from the hybridization of genotypes of the two gene pools. Similar results were observed from previous studies [23]–[25]. Breeding for crop improvement is depending on the enhanced of single or multiple traits that is desirable for food and nutrition systems [26]. In this study, the median concentration of Fe and Zn from genotypes in the admixture group was higher than the median concentration of genotypes in the two gene pools, suggesting that the Fe or Zn concentration in hybridization bean is mediated by minor additive effect and influenced by environment. Micronutrient uptake and accumulation traits in plant are heritable [27]; few additive genes controlling Zn concentration were also reported in rice, maize and soybean [28]–[30]. Trait controlled by additive effect has been beneficial to effectively breeding for micronutrient improvement.

Common bean contains high levels of important micronutrients especially Fe and Zn compared to the low level of Fe and Zn in many other staple food crops [31]. A previous study demonstrated that QTLs related to the concentration levels of Fe and Zn were identified on Pv02, Pv06, Pv07 and Pv08 by utilizing
microsatellite and RAPD markers [32]. Another study reported that SSRs markers associated to the Fe concentration level on Pv02, Pv05, Pv06, Pv07, Pv09 and Pv10, while markers associated with Zn on Pv01, Pv03, Pv05, Pv07 and Pv10 with one marker in Pv04 [33]. In this study, probable cyclic nucleotide-gated ion channel 20 (CNGC20) locus was observed nearby the significant SNP associated with the Zn concentration on Pv04. Zinc is micronutrient essential for all living organisms with key roles related to growth, development, and defense. Zinc ions are involved in oxidation/reduction processes, thus have both beneficial and toxic effects on plant cells. CNGCs are believed to be involved in the uptake of both essential and toxic cations [34]. The identified CNGC20 in common bean may play a role for maintaining Zn homeostasis. Although Zn seems to affect the capacity for water uptake and transport in plants [35], the correlation ($r = 0.06$) between Zn concentration and water uptake obtained in this study did not support this conclusion.

For the trait of Fe concentration, although we did not identify highly significant SNPs associated with Fe concentration, a putative causative locus was detected very closed to a less significant SSR ($8.33 \times 10^{-03}$) located on Pv10. This gene is annotated as cation/calcium exchanger 4 (CCX4) and identified as cation transporter. In Arabidopsis, $AtCCX4$ is expressed throughout the plant [36]. Although the mechanisms underlying the trafficking of iron are not clearly understood, chloroplast is the iron-richest system in plant cell [37]. We also identified another locus cytochrome P450 (CYP) (PHVU_010G126900g) within same peak of significant SNP. The CYPs are enzymes that oxidize substance using iron [38]. Because iron is highly reactive and need to shield Fe ions interaction with other molecules causing oxidative damage to the cell [37]. The CCX4 identified in this study may play a key role for cellular iron homeostasis as CNGC20. The correlation ($r = 0.13$) between Fe concentration and water uptake may suggest that a higher concentration of Fe may have a synergized effect with water uptake.

Interestingly, nuclear pore complex protein GP210 locus was detected on Pv04 and the most significant SNP was exactly resided within the sequence range of this locus. The nuclear pore complex (NPC) controls all trafficking of molecules in and out of the nucleus [39]. GP210 is a membrane-spanning glycoprotein and a major component of the NPC, mediating molecular transport between nucleoplasm and cytoplasm [40], [41]. Whether $GP210$ gene affects water uptake in common bean, needs further detail investigation because GP210, not only plays a role in molecules transit; but also, is required for nuclear pore assembly. Knockdown research study of the $GP210$ gene in the $C. elegans$ embryos, prevented the accumulation of aberrant structures on the nuclear membranes [41]. In addition, a gene encoding subtilisin protease SBT1.4 was observed in the flanking region of significant SNP on Pv10. Subtilisin proteases (SBTs) are serine proteases controlling diverse developmental processes with six distinct subfamilies [42]. For instance, $MtSBT1.1$ in $M. truncatula$ controls seed size in legume while the gene $AtSBT1.7$ is essential for mucilage release from seed coats, triggering the accumulation and/or activation of cell wall modifying enzymes necessary for the loosening of seeds in Arabidopsis [40]. It is believed that the SBT1.4 was associated with leaf senescence in Arabidopsis [43], [44]. The function of SBT1.4 in common bean is not fully understood, analysis of loss-of-function mutants would provide evidence and elucidate the function of SBT1.4 in seeds. A study in common bean has revealed a correlation between water uptake and the seed coat luster [45]. Because shiny seeds have firmer texture controlled by the
Asper (Asp) gene, this gene is the major determinant of water uptake and is located on Pv07. However, Asp gene is not detected within the regions of significant SNPs on Pv07 in this study, suggesting that this gene may not be directly related to water uptake.

The cooking time in common bean has become an important trait due to its association with the volume of fuel, water and time spent on food preparation. Among traits interested by consumers, cooking time was prioritized in farmers with proportional of 63% followed by taste (51%) and flatulence (50%) as stated in consumer acceptance report [38]. Cooking time is accompanied with water uptake trait. In this study, water uptake and cooking time were significantly associated with negative correlation $r=-0.20$. These two variables indicated that a decrease in cooking time predictably relies on the increase in water uptake or vice versa.

Reports from previous studies showed that genomic regions associated with cooking time were identified on different chromosomes, such as Pv06 [18]; Pv02, Pv03 and Pv06 [46]. In this study, in the upstream of the significant SNP on Pv03, a probable polygalacturonase homolog was detected as a candidate causal locus. Polygalacturonase is a major component of the pectin remodeling and disassembly network [47]. It is involved in catalyzing the degradation of pectin in the cell wall via the hydrolytic cleavage of glycosidic bonds [48], probably making seeds soften and shorten the cooking time.

**Conclusion**

GWAS technology is a powerful tool to determine the chromosome regions associated with various traits and facilitate the identification of candidate causative genes within these chromosome regions. We have demonstrated that several putative causal loci mediating four traits were identified through this GWAS analysis. Further validated loci could be used in common bean breeding to develop varieties with higher concentration level of micronutrients, efficient water uptake for shorter cooking time.

**Methods**

**Plant materials and field experiment**

A natural population of 206 common bean genotypes collected from International Centre for Tropical Agriculture, CIAT – Uganda (179), Ethiopia (12), Kenya (10), Tanzania (3) and Rwanda (2) were used in this study [49]. The experimental materials were planted in the northern Tanzania at medium altitudes with 1407 m above sea level (a.s.l) of S03°21.690’ and E36°37.879’ (Selian Agricultural Research Institute, abbreviated SARI) and at low altitudes with 992 m a.s.l of S03°19.905’and E037°14.067’ (Tanzania Coffee Research Institute, commonly called Lyamungo) in 2017 and 2018. The soil characteristics of these areas were Eutrophic Brown Soils on volcanic and Alluvial sediments – medium texture (loamy soils) [50]. The range of Fe was 29.85 mg/kg to 39.24 mg/kg and Zn was 0.33 mg/kg to 0.60 mg/kg. In both seasons and agro-ecologies, the experiments were laid-down in Incomplete Block Design with two replications for 206 genotypes. The plot size was 3.2 m length and 1.5 m width with 4 rows per plot and
0.5 m inter row spacing and 0.2 m intra row spacing under rainfed. Middle two rows were harvested and used for data collection.

**Sampling and sample preparation for Fe and Zn determination**

At full maturity, 30 well filled pods of each genotype were randomly harvested from two centered rows in each plot. Total of 100 seeds/genotype/location were sampled and dried in an Oven (Binder drying chamber, Model: ED 115, Tuttlingen, German) for 12 hours at 60°C. The moisture of dried seeds was measured by Grain Moisture Tester – Soybeans (Baton cooperation. Model: 8500, Michigan 48084) and seeds were ground by Retsch SK 100 (Retsch GmbH, Model: SK 1001C, 42781, Haan, Germany) with a stainless-steel grinding ring to produce bean fine powder. About 5 - 10 g flour of each genotype was scanned for the content of Fe and Zn (mg/kg) from different locations using X-ray Fluorescence (XRF) (Bruker AXS GmbH, Model: Tracer 5i, Östliche Rheinbrückenstr, 76187 Karlsruhe Germany).

**Sampling and sample preparation for water uptake and cooking time determination**

The rest of harvested seeds from each plot was bulked, sun dried in a paper bag sized 6.5 cm ×14 cm and stored at room temperature. When seed moisture reached between 10-14%, 30 seeds from each genotype were weighted (g) using electronic scale (SONASH ®, Model: SKS-006, Frances). The dried seeds were soaked in a jar with 150 ml water overnight. Thirty-seed weight was measured again to quantify water uptake (ml) using the formula (i) below. Then, 25 seeds were placed on the top of each well in a Matson Cooker (Customized Machining and Hydraulics Co., Winnipeg, Canada), which consists of a plate with 25 wells for individual seeds and 25 metal pins [51]. Cooking time in minutes were recorded when the 20th pin drop down, which represents the time of 80% seeds to be completely pierced with 85 g stainless steel rod with a 2 mm pin [51], [52]. The average cooking time was determined for each genotype based on two locations.

\[
\text{water uptake (\%)} = \frac{\text{seed weight after soaking (g)} - \text{seed weight before soaking (g)}}{\text{seed weight after soaking (g)}} \times 100 \quad (i)
\]

**Phenotypic data analysis**

Analysis of Variance (ANOVA) from GenStat software was used to determine the significant effects (p<0.05) among the genotypes and locations used and Fisher’s protected least significant differences (LSD) was applied to identify the populations whose means differs statistically [53], [54]. The formula in excel analysis software and Pearson correlation were used for plotting graph, correlation, regression, scatter plots and pie charts.

**Genotyping data collection**

Total of 15 seeds per genotype were shipped from Tanzania to Tuskegee University through USDA-Animal and Plant Health Inspection Services (APHIS) with permit number P587-180801-005 and phytosanitary certificate number 00310248. Four seeds were planted in the pot with diameter of 15 cm in the
greenhouse at George Washington Carver Agricultural Experiment Station in Tuskegee University Research Center, AL. Total genomic DNAs of 206 genotypes were extracted from trifoliate leaves using Wizard® Genomic DNA Purification Kit from Promega Corporation (Madison, WI). The concentration of each DNA sample was determined using NanoDrop Spectrophotometer (NanoDrop 2000) and the quality of DNA samples was also observed on 0.8% agarose gel. These DNA samples were sent to Heflin Center Genomic Core Lab at the University of Alabama (Birmingham, AL) for genotyping using the Next generation Sequencing (NGS) Method of Illumina BARCBEAN6K_3 BeadChip with 5398 SNP markers [55]. All markers were distributed across 11 pairs of common bean chromosomes (2n=22) and the Illumina BeadStation 500G was used to scan the BeadChips. The SNP calling was conducted with the genotyping modules of GenomeStudio v2.0.4 of 2018 available as a free download software.

**Population structure analysis**

When a nature population possesses several sub-groups with different genetic background, i.e. population structure, it may result in a confusing effect of genetic background and lead to false association. In order to determine number of subgroups in this collection, the Bayesian model-based clustering method was applied using STRUCTURE 2.3.4 software [56]. The admixture model with independent allele frequencies without prior population information was used for simulation. The STRUCTURE software was set at burn-in period length of 50,000 and after burn-in 50,000 Markov chain Monte Carlo (MMC) repetition were set five times. For joint inference of population substructure, the kinship (K) set at the range of 1 to 10 with five number of iterations runs for each (Kamfwa et al., 2015a; Tigist et al., 2019). The ideal number of sub-populations was determined using K method [57] implemented in the HARVESTER software [58].

**Marker-trait association analysis**

Filtering the SNPs with monomorphic marker was set at <2% of minor allele frequency (MAF) and 5052 SNP markers were retained and used in principal component analysis (PCA) and population structure analysis. The obtained SNP markers in respective distribution of chromosomes were employed in association analyses with TASSEL 5.0 software [59]. The following mixed linear model (MLM) formula (ii) was used:

$$Y = \gamma \alpha + \rho \beta + k\mu + \varepsilon$$

Where, Y is Phenotype of each genotype; $\gamma$ is the fixed effects of the SNP; $\rho$ is the fixed effect of population structure from PCA results; k is the random effect of kinship relative; $\varepsilon$ is the term error under normally distribution with mean =0 and variance $\delta^2$. The statistical model was used to test for trait-marker association studies [60].

**Putative candidate loci linked to traits**

By GWAS analysis, significant SNPs associated with traits can be detected in specific regions of chromosomes. To identify putative candidate loci influencing on traits, coding genes were explored within or nearby the significant SNP region based on the common bean reference genome.
(https://plants.ensembl.org/Phaseolus_vulgaris/Location/Genome?fbclid=IwAR1rLGjP0bqPtzMoJTGKoeU7KCRGYmXtcUIRfYgbUtmf-KfHLhLgjKxNk8k) and annotated by BLAST in the NCBI (https://www.ncbi.nlm.nih.gov).

Declarations

Acknowledgements

Authors are grateful to George Washington carver agricultural experiment station, Tanzania Agricultural Research Institute and Plant genomic lab Tuskegee University for greenhouse experimentation, phenotyping and genotyping respectively.

Author contributions

PHB, GH and CB designed the research study and developed the manuscript. SMT and BS analyzed data, XC, EB and JJ commented and corrected the manuscript. All authors proofread and approved the final version.

Funding

This research work was supported by USDA-NIFA-Evans-Allen through Tuskegee University, Government of Tanzania and CIAT/PABRA through Tropical Legume III project.

Availability of data and materials

All data generated and analyzed during this study are included in this published research article.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interest

Authors details

1Integrative Biosciences (IBS). Ph.D. Program, Tuskegee University, AL 36088. 2Department of Agricultural and Environmental Sciences, Tuskegee University, AL 36088, 3Department of Food and Nutritional Sciences, Tuskegee University, AL, 36088. 4Corp Research Institute, Fujian Academy of Agricultural Sciences, Fuzhou, China.
References

1. E. Bitocchi et al., “Mesoamerican origin of the common bean (Phaseolus vulgaris L.) is revealed by sequence data,” *Proc. Natl. Acad. Sci.*, vol. 109, no. 14, pp. E788–E796, Apr. 2012, doi: 10.1073/pnas.1108973109.

2. HarvestPlus, “Iron beans,” 2018. [Online]. Available: www.harvestplus.org.

3. M. Blair, L. Gonzales, P. Kiman, and L. Butare, “Genetic diversity, intergene pool introgression and nutritional quality of common beans (Phaseolus vulgaris L.) from central Africa,” *Theory Appl. Genet.*, no. 121, pp. 237–248, 2010.

4. R. M. Welch, W. A. House, S. Beebe, and Z. Cheng, “Genetic Selection for Enhanced Bioavailable Levels of Iron in Bean (Phaseolus vulgaris L.) Seeds,” *J. Agric. Food Chem.*, vol. 48, no. 8, pp. 3576–3580, Aug. 2000, doi: 10.1021/jf0000981.

5. W. Broughton, G. Hernandez, M. Blair, S. Beebe, P. Gepts, and J. Vanderleyden, “Beans (Phaseolus vulgaris spp.)-model food legumes,” *Plant Soil*, no. 252, pp. 55–128, 2003.

6. A. Fallah, A. Mohammad-Hasani, and A. Colagar, “Zinc is an Essential Element for Male Fertility: A Review of Zn Roles in Men’s Health, Germination, Sperm Quality, and Fertilization,” *J. Reprod. Infertil.*, vol. 19, no. 2, pp. 69–81, 2018.

7. J. A. Wiesinger et al., “Demonstrating a Nutritional Advantage to the Fast-Cooking Dry Bean (Phaseolus vulgaris L.),” *J. Agric. Food Chem.*, vol. 64, no. 45, pp. 8592–8603, Nov. 2016, doi: 10.1021/acs.jafc.6b03100.

8. W. Yong, L. Amin, and C. Dongpo, “Status and prospects of nutritional cooking,” *Food Qual. Saf.*, vol. 3, no. 3, pp. 137–143, Nov. 2019, doi: 10.1093/fqsafe/fyz019.

9. R. S. Gibson, F. Yeudall, N. Drost, B. Mtitimuni, and T. Cullinan, “Dietary interventions to prevent zinc deficiency,” *Am. J. Clin. Nutr.*, vol. 68, no. 2 SUPPL., pp. 484–487, 1998, doi: 10.1093/ajcn/68.2.484S.

10. M. Asfaw, M. Wondaferash, M. Taha, and L. Dube, “Prevalence of undernutrition and associated factors among children aged between six to fifty nine months in Bule Hora district, South Ethiopia,” *BMC Public Health*, vol. 15, no. 1, pp. 1–9, 2015, doi: 10.1186/s12889-015-1370-9.

11. WHO, “Global health Estimates: Deaths by age, sex and cause.,” 2017. [Online]. Available: www.who.int/health.

12. L. Smith, S. Ng, and B. Popkin, “Trends in US home food preparation and consumption: analysis of national nutrition surveys and time use studies from 1965–1966 to 2007-2008,” *Nutr. J.*, no. 12, pp. 45–49, 2008.

13. A. Menendez and M. Curt, “Energy and socioeconomic profile of a small rural community in the highlands of central Tanzania: A case study,” *Energy Sustain. Dev.*, no. 17, pp. 201–209, 2013.

14. R. Nakalema, “Relationships Between Physicochemical Properties and Hard-To-Cook Phenomenon of Dry Common Beans Abstract Relationships Between Physicochemical Properties and Hard-To-Cook Phenomenon of Dry Beans,” 2015.
15. S. Challa and N. Neelapu, “Genome-Wide Association Studies (GWAS) for Abiotic Stress Tolerance in Plants,” in *Biochemical, Physiological and Molecular Avenues for Combating Abiotic Stress Tolerance in Plants.* Visakhapatnam, 2018, pp. 135–150.

16. H. J. Liu and J. Yan, “Crop genome-wide association study: a harvest of biological relevance,” *Plant J.*, vol. 97, no. 1, pp. 8–18, 2019, doi: 10.1111/tpj.14139.

17. S. G. Tigist, R. Melis, J. Sibiya, A. B. Amelework, G. Keneni, and A. Tegene, “Population Structure and Genome-Wide Association Analysis of Bruchid Resistance in Ethiopian Common Bean Genotypes,” *Crop Sci.*, vol. 59, no. 4, pp. 1504–1515, Jul. 2019, doi: 10.2135/cropsci2018.09.0559.

18. K. Cichy and J. M. Wiesinger, “Genetic diversity and genome-wide association analysis of cooking time in common bean (Phaseolus vulgaris L.),” *Theory Appl. Genet.*, no. 128, pp. 1555–1567, 2015.

19. K. Kamfwa, K. A. Cichy, and J. D. Kelly, “Genome-Wide Association Study of Agronomic Traits in Common Bean,” *Plant Genome*, vol. 8, no. 2, p. plantgenome2014.09.0059, Jul. 2015, doi: 10.3835/plantgenome2014.09.0059.

20. V. Hoyos-Villegas, Q. Song, and J. D. Kelly, “Genome-wide Association Analysis for Drought Tolerance and Associated Traits in Common Bean,” *Plant Genome*, vol. 10, no. 1, p. plantgenome2015.12.0122, 2017, doi: 10.3835/plantgenome2015.12.0122.

21. J. Morin, K. Cardoso, P. Oblessuc, B. Rosa, K. Gomes, and A. Chiorato, “Genome-wide association studies of anthracnose and angular leaf spot resistance in common bean (Phaseolus vulgaris L.).” *PLOS One*, 11(3):e0150506. doi.org/10.1371/journal.pone.0150506.," *PLoS One*, vol. 3, no. 11, 2016, doi: org/10.1371/journal.pone.0150506.

22. K. Kamfwa, K. A. Cichy, and J. D. Kelly, “Genome-wide association analysis of symbiotic nitrogen fixation in common bean,” *Theor. Appl. Genet.*, vol. 128, no. 10, pp. 1999–2017, 2015, doi: 10.1007/s00122-015-2562-5.

23. K. A. Cichy *et al.*, “A Phaseolus vulgaris diversity panel for andean bean improvement,” *Crop Sci.*, vol. 55, no. 5, pp. 2149–2160, 2015, doi: 10.2135/cropsci2014.09.0653.

24. Kwak M and P. Gepts, “Structure of genetic diversity in the two major gene pools of common bean (Phaseolus vulgaris L., Fabaceae),” *Theor. Appl. Genet.*, vol. 118, no. 5, pp. 979–992, 2009.

25. M. M. Mensack, V. K. Fitzgerald, E. P. Ryan, M. R. Lewis, H. J. Thompson, and M. A. Brick, “Evaluation of diversity among common beans (Phaseolus vulgaris L.) from two centers of domestication using ‘omics’ technologies,” *BMC Genomics*, vol. 11, no. 1, p. 686, Dec. 2010, doi: 10.1186/1471-2164-11-686.

26. D. C. Lima, Â. de F. B. Abreu, R. A. D. C. Ferreira, and M. A. P. Ramalho, “Breeding common bean populations for traits using selection index,” *Sci. Agríc.*, vol. 72, no. 2, pp. 132–137, Feb. 2015, doi: 10.1590/0103-9016-2014-0130.

27. B. Sadeghzadeh, “A review of zinc nutrition and plant breeding,” *J. soil Sci. plant Nutr.*, no. ahead, pp. 0–0, 2013, doi: 10.4067/S0718-95162013005000072.

28. N. D. Majumder, S. C. Rakshit, and D. N. Borthakur, “Genetic effect on uptake of selected nutrients in some rice (O. sativa L.) varieties in phosphorus deficient soils,” *Plant Soil*, vol. 123, no. 1, pp. 117–
29. A. A. El-Bendary, M. M. El-Fouly, F. A. Rakha, A. A. Omar, and A. Y. Abou-Youssef, “Mode of inheritance of zinc accumulation in maize,” *J. Plant Nutr.*, vol. 16, no. 10, pp. 2043–2053, Oct. 1993, doi: 10.1080/01904169309364673.

30. E. E. Hartwig, W. F. Jones, and T. C. Kilen, “Identification and Inheritance of Inefficient Zinc Absorption in Soybean,” *Crop Sci.*, vol. 31, no. 1, pp. 61–63, Jan. 1991, doi: 10.2135/cropsci1991.0011183X003100010015x.

31. H. E. Bouis and R. M. Welch, “Biofortification—a sustainable agricultural strategy for reducing micronutrient malnutrition in the global south,” *Crop Sci.*, vol. 50, no. April, p. S-20-S-32, 2010, doi: 10.2135/cropsci2009.09.0531.

32. M. Blair, C. Astudio, J. Rengifo, and S. Beebe, “QTL for seed iron and zinc concentrations in a recombinant inbred lines population of andean common beans (Phaseolus vulgaris L.),” *Theor. Appl. Genet.*, vol. 122, no. 3, pp. 511–521, 2011.

33. R. Mahajan et al., “Linkage disequilibrium-based association mapping of micronutrients in common bean (Phaseolus vulgaris L.): a collection of Jammu & Kashmir, India. 3,” *Biotechnology*, vol. 7, no. 5, p. 295, 2017, doi: org/10.1007/s13205-017-0928-x.

34. J. Moon, C. Belloeil, M. Ianna, and R. Shin, “Arabidopsis CNGC Family Members Contribute to Heavy Metal Ion Uptake in Plants,” *Int. J. Mol. Sci.*, vol. 20, no. 2, p. 413, Jan. 2019, doi: 10.3390/ijms20020413.

35. B. Hafeez, “Role of Zinc in Plant Nutrition- A Review,” *Am. J. Exp. Agric.*, vol. 3, no. 2, pp. 374–391, Jan. 2013, doi: 10.9734/AJEA/2013/2746.

36. J. Morris, H. Tian, S. Park, C. S. Sreevidya, J. M. Ward, and K. D. Hirschi, “AtCCX3 Is an Arabidopsis Endomembrane H + -Dependent K + Transporter,” *Plant Physiol.*, vol. 148, no. 3, pp. 1474–1486, Nov. 2008, doi: 10.1104/pp.108.118810.

37. S. B. Schmidt, M. Eisenhut, and A. Schneider, “Chloroplast Transition Metal Regulation for Efficient Photosynthesis,” *Trends Plant Sci.*, vol. 25, no. 8, pp. 817–828, Aug. 2020, doi: 10.1016/j.tpls.2020.03.003.

38. E. Stavropoulou, G. G. Pircalabioru, and E. Bezirtzoglou, “The Role of Cytochromes P450 in Infection,” *Front. Immunol.*, vol. 9, Jan. 2018, doi: 10.3389/fimmu.2018.00089.

39. A. Loschberger et al., “Super-resolution imaging visualizes the eightfold symmetry of gp210 proteins around the nuclear pore complex and resolves the central channel with nanometer resolution,” *J. Cell Sci.*, vol. 125, no. 3, pp. 570–575, Feb. 2012, doi: 10.1242/jcs.098822.

40. E. Lelièvre et al., “Signaling Pathways Recruited by the Cardiotrophin-like Cytokine/Cytokine-like Factor-1 Composite Cytokine,” *J. Biol. Chem.*, vol. 276, no. 25, pp. 22476–22484, Jun. 2001, doi: 10.1074/jbc.M101681200.

41. M. Cohen, N. Feinstein, K. L. Wilson, and Y. Gruenbaum, “Nuclear Pore Protein gp210 Is Essential for Viability in HeLa Cells and Caenorhabditis elegans,” *Mol. Biol. Cell*, vol. 14, no. 10, pp. 4230–4237, Oct. 2003, doi: 10.1091/mbc.e03-04-0260.
42. I. D’Erfurth et al., “A role for an endosperm-localized subtilase in the control of seed size in legumes,” New Phytol., vol. 196, no. 3, pp. 738–751, Nov. 2012, doi: 10.1111/j.1469-8137.2012.04296.x.

43. D. E. Martinez, M. L. Borniego, N. Battchikova, E.-M. Aro, E. Tyystjärvi, and J. J. Guiamét, “SASP, a Senescence-Associated Subtilisin Protease, is involved in reproductive development and determination of silique number in Arabidopsis,” J. Exp. Bot., vol. 66, no. 1, pp. 161–174, Jan. 2015, doi: 10.1093/jxb/eru409.

44. Y. O. Zubo et al., “Cytokinin induces genome-wide binding of the type-B response regulator ARR10 to regulate growth and development in Arabidopsis,” Proc. Natl. Acad. Sci., vol. 114, no. 29, pp. E5995–E6004, Jul. 2017, doi: 10.1073/pnas.1620491114.

45. K. Cichy et al., “QTL analysis of canning quality and color retention in black beans (Phaseolus vulgaris L.),” Mol. Breed., vol. 33, no. 1, pp. 139–154, 2014, doi: 10.1007/s11032-013-9940-y.

46. E. Perez-Vega, A. Paneda, C. Rodriguez-Suarez, A. Campa, and R. Giraldez, “Mapping of QTLs for morpho-agronomic and seed quality traits in a RIL population of common bean (Phaseolus vulgaris L.),” Theor. Appl. Genet., no. 120, pp. 1367–1380, 2014.

47. W. Duan et al., “Comprehensive analysis of the polygalacturonase and pectin methylesterase genes in Brassica rapa shed light on their different evolutionary patterns,” Sci. Rep., vol. 6, no. 1, p. 25107, Jul. 2016, doi: 10.1038/srep25107.

48. Y. Babu and M. Bayer, “Plant Polygalacturonases Involved in Cell Elongation and Separation—The Same but Different?,” Plants, vol. 3, no. 4, pp. 613–623, Dec. 2014, doi: 10.3390/plants3040613.

49. P. H. Binagwa, G. He, M. Egnin, G. C. Bernard, D. Mortley, and C. K. Bonsi, “Evaluating natural infection of fungal, bacterial and viral pathogens to dry bean genotypes under field conditions,” vol. 12, no. March, pp. 70–90, 2020, doi: 10.5897/JPBCS2019.0866.

50. N. C. Brady and R. R. Weil, “Instructor’s Manual with Test Item File to accompany The Nature and Properties of Soils Fourteenth Edition,” Up. Saddle River, NJ Prentice Hall, vol. 13, 2008.

51. N. Wang and J. K. Daun, “Determination of cooking times of pulses using an automated Mattson cooker apparatus,” J. Sci. Food Agric., vol. 85, no. 10, pp. 1631–1635, 2005, doi: 10.1002/jsfa.2134.

52. S. Mattson, “The Cookability of yellow peas,” Acta Agric. Scand., no. 4, pp. 154–164, 1946.

53. A. G. Clewer and D. H. Scarisbrick, Practical statistics and experimental design for plant and crop science. Wiley. Chichester, New York: Wiley, 2001.

54. S. Liu, K. Yu, and S. Park, “Development of STS markers and QTL validation of common bacterial blight resistance in common bean,” Plant Breed., no. 127, pp. 62–68, 2008.

55. Q. Song et al., “SNP Assay Development for Linkage Map Construction, Anchoring Whole-Genome Sequence, and Other Genetic and Genomic Applications in Common Bean,” G3&amp;#58; Genes/Genomes/Genetics, vol. 5, no. 11, pp. 2285–2290, Nov. 2015, doi: 10.1534/g3.115.020594.

56. J. Pritchard, M. Stephens, and P. Donnelly, “Inference of population structure using multilocus genotype data,” Genetics, no. 155, pp. 945–959, 2000.

57. G. Evanno, S. Regnaut, and J. Goudet, “Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study,” Mol. Ecol., vol. 14, no. 8, pp. 2611–2620, 2005, doi:
58. D. Earl and B. Von Holdt, “STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method,” *Genet. Resour.*, no. 4, pp. 359–361, 2012.

59. P. J. Bradbury, Z. Zhang, D. E. Kroon, T. M. Casstevens, Y. Ramdoss, and E. S. Buckler, “TASSEL: Software for association mapping of complex traits in diverse samples,” *Bioinformatics*, vol. 23, no. 19, pp. 2633–2635, 2007, doi: 10.1093/bioinformatics/btm308.

60. S. Mamidi *et al.*, “Genome-Wide Association Analysis Identifies Candidate Genes Associated with Iron Deficiency Chlorosis in Soybean,” *Plant Genome*, vol. 4, no. 3, pp. 154–164, 2011, doi: 10.3835/plantgenome2011.04.0011.

**Tables**
Table 1
List of most significant SNPs associated with traits, their locations and account for the percentage of phenotypic variation

| Trait               | SNPs       | Chr. | Position  | P-value           | Phenotypic variation (R²) |
|---------------------|------------|------|-----------|-------------------|--------------------------|
| Water uptake (%)    | ss715647339 | 3    | 49,319,901| $9.42 	imes 10^{-05}$ | 10.23                    |
|                     | ss715649433 | 4    | 547,449   | $2.94 	imes 10^{-08}$ | 18.19                    |
|                     | ss715646915 | 4    | 1,058,178 | $4.57 	imes 10^{-06}$ | 13.31                    |
|                     | ss715639293 | 10   | 40,150,769| $7.90 	imes 10^{-06}$ | 12.76                    |
| Cooking time (min)  | ss715643803 | 3    | 31,752,330| $2.14 	imes 10^{-05}$ | 10.06                    |
|                     | ss715642813 | 3    | 31,787,315| $2.14 	imes 10^{-05}$ | 10.06                    |
|                     | ss715649337 | 4    | 43,577,378| $5.03 	imes 10^{-05}$ | 9.31                     |
|                     | ss715639842 | 10   | 1,821,163 | $6.01 	imes 10^{-05}$ | 9.13                     |
| Zn (mg/kg)          | ss715646134 | 4    | 42,529,945| $5.98 	imes 10^{-05}$ | 11.09                    |
| Fe (mg/kg)          | ss715646330 | 10   | 39,796,958| $8.33 	imes 10^{-03}$ | 8.26                     |
| Trait               | Coding gene    | Position (bp)       | Annotation                                      | Gene accession     |
|---------------------|----------------|---------------------|------------------------------------------------|--------------------|
| Water uptake        | PHVU_004G006800g | Pv4:528,718 - 556,900 | Nuclear pore complex protein GP210              | XM_028068273.1     |
| Cooking time        | PHVU_003G129300g | Pv3:31,716,125 - 31,724,390 | Probable polygalacturonase                      | XM_028064102.1     |
| Zn concentration    | PHVU_004G144200g | Pv4:42,438,787 - 42,446,049 | Probable cyclic nucleotide-gated ion channel 20 | XM_028068510.1     |
| Fe concentration    | PHVU_010G128500g | Pv10:39,817,553 - 39,820,115 | Cation/calcium exchanger 4                     | XM_028051654.1     |