Identification of QTLs/ Candidate Genes for Seed Mineral Contents in Common Bean (*Phaseolus vulgaris* L.) Through Genotyping-by-Sequencing

**Muslima Nazir**¹, Reetika Mahajan¹, Sheikh Mansoor², Sheezan Rasool¹, Rakeeb Ahmad Mir³, Ravinder Singh⁴, Vandana Thakral⁵, Virender Kumar⁵, Parvaze A. Sofi⁶, Hamed A. El-Serehy⁷, Daniel Ingo Hefft⁸ and Sajad Majeed Zargar¹*

¹Proteomics Laboratory, Division of Plant Biotechnology, Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir, Srinagar, India, ²Division of Biochemistry, FBSc, Sher-e Kashmir University of Agricultural Sciences and Technology of Jammu, Jammu, India, ³Department of Biotechnology, BGSB University, Rajouri, India, ⁴School of Biotechnology, Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu, Jammu, India, ⁵National Agr Food Biotechnology Institute (NABI), Mohali, India, ⁶Division of Genetics and Plant Breeding, Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir, Srinagar, India, ⁷Department of Zoology, College of Science, King Saud University, Riyadh, Saudi Arabia, ⁸University Centre Reaseheath, Reaseheath College, Nantwich, United Kingdom

Throughout the ages, the common bean has been consumed by humanity as an important food staple crop and source of nutrition on a global scale. Since its domestication, a wide spectrum of phenotypic and genotypic investigations have been carried out to unravel the potential of this crop and to understand the process of nutrient accumulation along with other desirable characteristics. The common bean is one of the essential legume crops due to its high protein and micronutrient content. The balance in micronutrients is critical for the growth and development of plants as well as humans. Iron (Fe), Zinc (Zn), Copper (Cu), Manganese (Mn), Magnesium (Mg), Calcium (Ca), and Molybdenum (Mo) are some of the important micronutrients present in legumes. Thus, we aimed to investigate the quantitative trait loci’s (QTLs)/single nucleotide polymorphisms (SNPs) to identify the candidate genes associated with micronutrients through genotyping by sequencing (GBS). In our investigation, through GBS we identified SNPs linked with traits and assessed seven micronutrients in 96 selected common bean genotypes for screening nutritionally rich genotypes. Among 96399 SNPs total identified through GBS, 113 SNPs showed significant phenotypic variance, ranging from 13.50 to 21.74%. SNPs associated with most of the seed micronutrients (Mg, Mn, Fe, Ca, Cu) were found on chr3 & chr11 (Mg, Mn, Mo, Ca, Zn). The findings from this study could be used for haplotype-based selection of nutritionally rich genotypes and for marker-assisted genetic enhancement of the common bean. Further, the identified SNPs for candidate genes/transporters associated with micronutrient content may pave the way for the enrichment of seeds by employing genomics-assisted breeding programs.

**Keywords:** common bean, single nucleotide polymorphism (SNP), genome wide association studies (GWAS), ionome, transporters, population structure
1 INTRODUCTION

Common bean (Phaseolus vulgaris L.), an important food legume, constitutes 50% of the grain legumes consumed as a human food source (Câmara et al., 2013; FAOSTAT, 2017). Common bean has huge genetic variation and based on domestication it is distinguished into two main gene pools, Andean and Mesoamerican. Cultivars in the mesoamerican gene pool have small to medium seed size and with “S” or “B” phaseolin patterns, while Andean cultivars have a large seed size with phaseolin patterns “Y”/“C”/“H”/“A” (Bitocchi et al., 2012; Bellucci et al., 2014).

Being a great source of carbohydrates, dietary proteins, soluble and insoluble fibers, vitamins, and essential micronutrients such as minerals including Iron (Fe), Zinc (Zn), Copper (Cu), Manganese (Mn), Magnesium (Mg), Calcium (Ca) and Molybdenum (Mo), beans have often been considered a “poor man’s meat” (Fennema owen, 2000; Hayat et al., 2014; Yeken, 2018). These micronutrients play a pivotal role in the proper growth and development of plants and animals. The deficiency of any essential micronutrients hinders the proper functioning of biological processes leading to several metabolic and physiological implications. In many low- and lower-middle income countries, especially those in Asia, Africa, and Latin America, micronutrients Fe and Zn are the main components of hidden hunger (Darnton-Hill et al., 2006). Similarly, Fe and Zn deficiency cause severe yield loss in crops and metabolic disturbances in humans (Zargar et al., 2015). In addition, Cu is one of the important trace elements that play a vital role in maintaining metabolic activities. In humans, Cu deficiency leads to anemia, cardiac dysfunction, myeloneurropathy, and myelopathy, whereas in plants, its deficiency leads to lignification dysfunctioning (Geir, 2013; Lehmann and Rillig, 2015; Papamargaritis et al., 2015).

Food crops are the major source of essential minerals (Hardiman et al., 1984; Ali et al., 2014). Biofortification of nutrient-rich plants is important in addressing malnutrition-related issues. Thus, given today’s population explosion and food shortage issues, there is a need to introduce smart food crops into our diet. So far, research has been carried out using both conventional and modern breeding approaches to increase the nutrient content in cultivated plants. It is now important to identify genetic loci that regulate the uptake of essential minerals, as each genetic loci is an essential factor in the success of the biofortification breeding effort. Association mapping is one of the modern breeding approaches used to identify genetic loci that determine desired traits (Lewontin and Kojima, 1960). Association mapping has several advantages over the biparental QTL mapping approach (Yu and Buckler, 2006; Tian et al., 2011; Wen et al., 2014; Nadeem et al., 2021). For example, association mapping explores the allelic diversity that exists in the diverse germplasm, while QTL mapping can examine the allele variation present in only two parental lines. Undoubtedly, both approaches are indispensable and have their own advantages and disadvantages (Sonah et al., 2013). We believe an integrated approach involving both association mapping as well as biparental mapping can lead to a breakthrough in crop improvement.

Since the whole genome sequence of the common bean is available in the public domain, as such there is an excellent opportunity to perform Genome-wide association studies (GWAS) to identify QTLs followed by candidate genes that govern the uptake and accumulation of minerals (Schmutz et al., 2014). Genome-wide association study is one of the modern breeding approaches for mapping genes associated with different traits (Lu et al., 2015; Liu et al., 2016). The introduction of Next-generation sequencing (NGS) technologies has sped up the identification of SNPs and subsequent genotyping (Verma et al., 2015). Genotyping-by-sequencing (GBS) is a robust and cost-effective method wherein selective small genome fragments obtained by restriction digestion are sequenced by NGS platform to identify SNPs (Narum et al., 2013; Peterson et al., 2014; Schilling et al., 2014). In the last two decades, numerous studies have been performed using the GBS approach in diverse plant species, including wheat, canola, barley, and soybean, which are known to have a complex genome. In this regard, GWAS in common bean will allow us to estimate population structure and linkage disequilibrium (LD), connecting the variation in the genome with the phenotypic variations in the population. The LD-based analysis is organized based on population structure and genetic relatedness among populations. Several association studies with respect to micronutrient contents have been carried out in food crops like rice (Shao et al., 2011; Norton et al., 2014), pea (Diapari et al., 2015), chickpea (Diapari et al., 2014; Upadhyaya et al., 2016; Ozkuru et al., 2019), common bean (Nemli et al., 2016; Mahajan et al., 2017a; Erdogmus et al., 2020).

In the present study, we examined the germplasm of common beans of the Jammu and Kashmir region of northwestern Himalayas, India, for various micronutrients, followed by the identification of QTLs associated with the accumulation of nutrients. To identify genes/QTLs that regulate micronutrients in beans, GBS-based SNPs were discovered from natural populations of beans. Based on the preliminary studies, a set of 96 different bean genotypes was created, which were collected from other regions of the northwestern Himalayas (Zargar et al., 2014). In the present study, the ionome profile, which contains Ca, Cu, Mg, Mn, and Mo, of the bean core set was deciphered, and the QTLs contributing to their accumulation by GWAS have been identified. The investigation led us to inventory the micronutrients of the 96 different types of beans to determine candidate lines (nutrient-rich for different micronutrient levels in the seed) that can be used in breeding high-nutrient, high yielding bean varieties. In addition, the identification of QTLs can serve as critical genomic resources for improving the micronutrient profile in beans. These studies can improve the understanding of possible correlations for the accumulation of different elements. To understand the function of the respective candidate genes that regulate micronutrient uptake in beans, further studies need to be performed that may include knock out or overexpression of responsive candidate genes.
2 EXPERIMENTAL PROCEDURES

2.1 Plant Material
A total of 96 common bean germplasm lines, mostly landraces and a few released varieties (SFB1, SR1, SR2, Arka Anoop, VLR-125) were used as plant material in the current study. Germplasm was collected from different geographical regions of Jammu and Kashmir (Supplementary Table S1) and maintained at the research fields at Division of Genetics and Plant Breeding, Faculty of Agriculture, Wadura, Sher-e-Kashmir University of Agricultural Sciences and Technology-Kashmir (SKUAST-K), Sopore. Most of the released varieties used in this study were developed through a single plant selection and have been used as checks.

2.2 Field Experiment and Micronutrient Profiling
Field experiments were conducted during 2016 and 2017 at the experimental field of the College of Agriculture, Wadura, SKUAST-Kashmir, India (34.34 North; 74.4 East; Altitude: 1,590.00 m). Clay loam textured soil with pH (7.2), organic carbon (.65%), electrical conductivity (.18 dS/m), CEC (16 meq/kg), and an available concentration of different elements in the soil i.e. P (4.91 mg kg$^{-1}$), K (5.55 mg kg$^{-1}$), Zn (.68 mg kg$^{-1}$), Fe (5.1 mg kg$^{-1}$), Cu (.29 mg kg$^{-1}$), Mn (6.2 mg kg$^{-1}$) was used for plantation of germplasm seeds in the experimental sites. The experiment was laid out according to augmented block design, which includes more than one released variety that is taken as replicated treatments, and these varieties are repeated in each block. Five released varieties (SFB1, SR1, SR2, Arka Anoop, VLR-125) were included in each block as checks. All the standard agronomical practices recommended were followed to raise healthy and disease-free crop plants. Harvesting was done at the time of 90% pod maturity. Further, the seed material of each genotype was powdered to analyze seven essential micronutrients i.e. Cu, Mn, Mg, Ca, Fe, Zn, and Mo. The elemental profiling of these genotypes was determined using a portable X-ray fluorescence spectrometer (pXRF). The pXRF instrument was calibrated as explained in Reidinger et al. (2012). In pXRF, a synthetic methylcellulose matrix was used to spike the known quantity of standard elements. Based on the methylcellulose pellet with known standards, an elemental composition standard curve was developed and subsequently used for sample evaluation. Similarly, samples were cross-verified with Energy Dispersive X-ray Fluorescence (ED-XRF).

2.3 Genotyping-by-Sequencing of Common Bean Genotypes
Genomic DNA was extracted from 15 day old leaves by using the CTAB method (Doyle and Doyle, 1990), and the quality, as well as quantity of DNA, was checked on both gel electrophoresis (8% Agarose) and nano-drop (mySPEC, Wilmington, USA). The extracted genomic DNA was purified for the preparation of multiplex GBS libraries via Illumina HiSeq 2000 (SciGenom Pvt. Ltd., India). ApeK1 (from Aeropyrum pernix K1) restriction enzyme was employed for restriction digestion of genomic DNA. After quality filtering, de-multiplexed high-quality sequences were mapped to the reference common bean genome (Phytozome v12.1 database), and SNPs were mined from the coding and non- coding regions of common bean genes and chromosomes. Subsequently, SNPs mined were structurally annotated on the diverse coding DNA sequence (CDS) and non-coding (upstream/downstream regulatory regions and introns) sequence components of genes and intergenic regions of the common bean genome.

2.4 Statistical Analysis

2.4.1 Micronutrient Profiling Analysis
All the observations were recorded in replicates of three, and values were then averaged. One-way ANOVA was applied to evaluate the variance of seven micronutrients among the genotype and Pearson’s pairwise correlation coefficient was calculated for all elements using the SPSS program (ver. 16).

2.4.2 Population Genetic Analysis
Population structure was estimated using a Bayesian Markov Chain Monte Carlo model (MCMC) implemented in STRUCTURE (Pritchard et al., 2000). Five runs were performed for each number of populations (K) set from 01 to 12. Burn-in time and MCMC replication number were both set to 100,000 for each run. The most likely K value was determined by the log probability of the data lnP(D) and delta K, based on the rate of change in lnP(D) between successive K values. These analyses were performed using Structure Harvester (Earl and vonHoldt, 2012). The neighbor-joining tree was built using Phylip and MEGA5 (Felsenstein, 1989; Tamura et al., 2011).

2.4.3 Genome-wide Association Analysis
All the analyses were performed using TASSEL3.0 and the Genomic Association and Prediction Integrated Tool (GAPIT) (Bradbury et al., 2007; Lipka et al., 2012). Mixed linear models (MLM) were used for the identification of SNPs associated with these traits. Van Raden method (K) was used to calculate the kinship matrix (Loiselle et al., 1995; Hyun et al., 2008). Covariate like P from principal component analysis and Q from STRUCTURE along with kinship matrix (K) were used for mixed linear models (MLM). The negative log(1/n) was used to establish a significance threshold (Wang et al., 2012; Yang et al., 2013).

2.4.4 LD Plots and Haplotype Blocks
The SNP matrix for all the samples was converted to HapMap format and TASSEL was used for the filtering out of SNP with major allele frequency (MAF) of less than .05. The LD plots for individual chromosomes were created in Tassel > Analysis > Diversity > Linkage Disequilibrium with LD Type as “Sliding Window” and LD Window size set to 50. The heterozygous sites were treated as missing and the R$^2$ accumulated was calculated for 100 intervals with a size of .01. The R$^2$ data obtained following LD analysis was used for plotting chromosome-wise LD information using MS Excel.

The PLINK v1.90b6.24 64-bit was used to determine the block size of SNP haplotypes. The individual chromosome Tassel files
were saved to Plink format (.map and .ped) using the option available in Tassel. The haplotype blocks were calculated using—blocks no-pheno-req--blocks-max-kb 100--make-founders command in Plink. The Haplotype Block analysis revealed the size of Haplo Blocks, including the number of SNPs on each block. The data was plotted using MS Excel.

2.5 Candidate Gene Identification
For candidate genes identification, the reference genome of *P. vulgaris* (V2.1) was used. The candidate genes were identified in .1 Mb both flanking regions of significant SNPs using the BioMart tool (Smedley et al., 2009), and the information related to gene function, Pfam ID, Panther ID, KOG ID, gene ontology ID, and their description were downloaded.

3 RESULTS
3.1 Genetic Variation in Common Bean Seed Micronutrients
Micronutrient profiling of 96 genotypes was conducted to study the distribution and correlation among each other. Normal distribution was observed for Cu, Mn, Mg, Ca, Fe, Zn, and Mo (Figure 1; Supplementary Tables S2, S3). Seed micronutrient content showed continuous variation for Fe ranging from 67.35–133.02 mg kg⁻¹ (with average (av) of 94.21 mg kg⁻¹), 21.16–49.77 mg kg⁻¹ (av 35.98 mg kg⁻¹) for Zn, 1,293.21–2,657.46 mg kg⁻¹ (av 1893.56 mg kg⁻¹) for Ca, 2.02–28.02 mg kg⁻¹ (av 13.11 mg kg⁻¹) for Cu, 1,057.55 mg kg⁻¹ 208–2,492.26 mg kg⁻¹ (av 1827.71 mg kg⁻¹) for Mg, 22.34–93.36 mg kg⁻¹ (av 58.26 mg kg⁻¹) for Mn and 2.09–7.80 mg kg⁻¹ (av 4.63 mg kg⁻¹) for Mo (Table 1). Very low coefficients of variation (CVs) were observed for all these micronutrients.

3.2 Correlation Analysis
Pearson’s correlation analysis showed a highly significant positive correlation between Fe and Zn (r = .61**), Ca and Mg (r = .37**), Cu and Fe (r = .28**), and Cu and Zn (r = .26*) (Table 2). All the other micronutrients showed a non-significant correlation with each other. Correlation studies between the micronutrient contents of the seeds in the present study showed that Mn is the only micronutrient that correlates negatively with all other micronutrients such as Fe, Zn, Cu, Ca, Mg, and Mo. However, the Zn content was found positively correlated with Fe and Cu and negatively correlated with Ca, Mo, and Mn. However, the Zn content was found positively correlated with Fe and Cu and negatively correlated with Ca, Mo, and Mg. Mo was found to correlate positively with Ca, Mg, Cu and negatively with Fe, Zn, and Mn. The Fe content was found to be positively correlated with Ca, Mg, and Cu, while it was negatively correlated with Mo and Mn.

3.3 Characterization and Distribution of SNPs in the Common Bean
A total of 96,399 SNPs were found among a whole set of 96 diverse genotypes. The highest number of SNPs (10,978) were observed on chr. 3, whereas the lowest number of SNPs (6524) were observed on chr. 6 (Figure 2).
3.4 Genetic Diversity and Population Structure

In the present study, all paired genetic distances between the 96 bean lines were determined from SNP-based genotypic data. A neighboring tree based on these genetic distances showed that the genotypes were divided into five main and five subgroups (Figure 3A). The Dendrogram revealed that of the minor groups, one included only one genotype, WB6; another group included only two genotypes (N13, WB1137); while another two groups were found to have only three genotypes each (KD17, WB1664, KD5 in one group and WB877, KD11, WB864 in the second group) and yet another group comprised of four genotypes (WB1282, KD13, R9, WB1436). The released lines like SR2 and ARKAANOOP were grouped in one cluster with other local lines, whereas VLR-125 and SFB1 clustered in another group.

Principal component analysis (PCA) also showed diversity among the common bean genotypes (Figure 3B). The released bean lines and local lines have been pooled with no clear separation of local and released lines (Figure 3B). In addition, population structure analysis provides a robust analysis for understanding the genotypic origins of a particular crop. The population structure was scored for K values ranging from 1 to 12 across the panel using high-quality SNPs. The peak of delta K was found to be the highest at K = 5 and thus groups the 96 genotypes of the common bean into 5 populations (Figure 3C). Furthermore, this was in accordance with the neighbor-joining tree with little deviation (Figure 3A).

3.5 Genome-Wide Association Analysis

GWAS was performed for common bean seed micronutrient contents (Fe, Zn, Cu, Ca, Mn, Mg, and Mo). Out of 96,399 SNPs, included only two genotypes (N13, WB1137); while another two groups were found to have only three genotypes each (KD17, WB1664, KD5 in one group and WB877, KD11, WB864 in the second group) and yet another group comprised of four genotypes (WB1282, KD13, R9, WB1436). The released lines like SR2 and ARKAANOOP were grouped in one cluster with other local lines, whereas VLR-125 and SFB1 clustered in another group.

Principal component analysis (PCA) also showed diversity among the common bean genotypes (Figure 3B). The released bean lines and local lines have been pooled with no clear separation of local and released lines (Figure 3B). In addition, population structure analysis provides a robust analysis for understanding the genotypic origins of a particular crop. The population structure was scored for K values ranging from 1 to 12 across the panel using high-quality SNPs. The peak of delta K was found to be the highest at K = 5 and thus groups the 96 genotypes of the common bean into 5 populations (Figure 3C). Furthermore, this was in accordance with the neighbor-joining tree with little deviation (Figure 3A).

**TABLE 1** | Range, highest, lowest genotype and Coefficient of Variance (CV) of different micronutrient content of common bean seeds.

| S.No | Micronutrient       | Range (mg kg$^{-1}$) | Lowest | Highest  | Average (mg kg$^{-1}±$S.E) | CV  |
|------|---------------------|----------------------|-------|----------|-----------------------------|-----|
| 1    | Molybdenum (Mo)     | 2.09–7.80            | KD11  | WB1680   | 4.63 ± 0.07                 | 3.54|
| 2    | Zinc (Zn)           | 21.16–49.77          | WB352 | WB1190   | 35.98 ± 0.05                | .30 |
| 3    | Iron (Fe)           | 67.35–133.02         | K13   | WB1679   | 94.21 ± 0.04                | .10 |
| 4    | Calcium (Ca)        | 1,293.21–2,657.46    | WB1643| KD7      | 1893.56 ± 0.03              | .03 |
| 5    | Magnesium (Mg)      | 1,057.55–2,492.26    | WB371 | K16      | 1827.71 ± 0.03              | .004|
| 6    | Manganese (Mn)      | 22.34–93.36          | R2    | N15      | 58.26 ± 0.02                | .27 |
| 7    | Copper (Cu)         | 2.02–28.02           | WB1136| UJ       | 13.11 ± 0.02                | .07 |

**TABLE 2** | Correlation among seed micronutrient content in common bean.

| Correlation   | Mo     | Zn     | Fe     | Ca     | Mg     | Mn     | Cu     |
|---------------|--------|--------|--------|--------|--------|--------|--------|
| Mo            | 1      | .102   | .032   | .052   | .010   | .121   | .258   |
| Zn            |        | 1      | .032   | .052   | .010   | .121   | .258   |
| Fe            |        |        | 1      | .052   | .010   | .121   | .258   |
| Ca            |        |        |        | 1      | .052   | .010   | .121   |
| Mg            |        |        |        |        | 1      | .052   | .010   |
| Mn            |        |        |        |        |        | 1      | .052   |
| Cu            |        |        |        |        |        |        | 1      |

*aCorrelation is significant at the .01 level.
*bCorrelation is significant at the .05 level.

**FIGURE 2** | Frequency distribution of SNPs identified using GWAS on different chromosomes of common bean.
113 SNPs were found to be significantly associated with different seed micronutrient contents with 13.50–21.74% phenotypic variance (Figures 4A–G; Table 3). A total of 32 SNPs across all chromosomes except chr. 8 were found significantly associated with seed Mg content. The highest number of SNPs (7) associated with seed Mg content was found on chr.2, while only one associated SNP was found on chr.1 and chr.7. One SNP on chr.9 positioned at 31752041 (p-value = 1.61E-04) contributed 18.42% phenotypic variation. For seed Mo content, 29 SNPs were found significantly associated that are positioned on chr.4, chr.5, chr.6, chr.10, and chr.11 with the highest number of SNPs (13) on chr.10 and the lowest (1) on chr.6 and chr.7. One SNP on chr.11 contributed to 21.42% phenotypic variation. About 15 SNPs were found associated with seed Ca content on each of chr.3, chr.4, chr.5, chr.7, chr.9, and chr.11 and seed Cu content on chr.2, chr.3, chr.5, chr.6, and chr.8. For seed Ca content, SNP contributing to the highest phenotypic variation (17.10%) was found on chr.3 whereas, SNP associated to seed Cu content located on chr.2 and chr.3 contributed 21.74% and 21.71% phenotypic variation. Around 13 SNPs were found associated with seed Mn content on chr.1, chr.3, chr.8, chr.10 and chr.11 whereas, only 5 SNPs were found associated with seed Zn content on chr.2, chr.10, and chr.11 and 4 SNPs were associated to seed Fe content on chr.1, chr.3, and chr.7.

3.6 LD Plot and Haplotype Blocks

Overall LD measured as R² was correlated in all chromosomes (Figure 5; Supplementary Table S4). The maximum number of SNPs (125,580) correlated with other chromosomes were found on chr.10, whereas, a minimum number of SNPs (50,003) were found on chr.1 at R² = .01. With the increase in R² value, decreasing pattern in the number of associated SNPs across the 11 chromosomes was observed. At R² = .93 the lowest number of SNPs correlation was found in all the chromosomes. The lowest (33) and highest (232) were found on chr. 6 and 1, respectively. Moreover, a total of 1879 SNPs on chr.1 were found to correlate with other chromosomes at R² = .96, whereas 421 SNPs were found on chr.10. No SNPs were found associated across 11 chromosomes at R² = .97–.99. Through haplotype analysis, a set of 7107 haploblocks representative of the 11 chromosomes, ranging from 1244 (chr. 1) to 503 (chr. 4) were identified. A total of 22,090 SNPs were distributed in these blocks, with an average of ~3 SNPs per block. Chr. 1 (17) and chr. 4 and 10 (7 each) had the highest and lowest number of SNPs within their haploblocks, respectively (Figures 6A–K; Supplementary Table S5).
3.7 Candidate Gene Analysis
A total of 840 genes were identified in the .1 Mb flanking region of significant SNPs related to different traits; however, five SNPs were not co-localized with any gene in the .1 Mb flanking region (Supplementary Table S6). Out of these 840 genes, 16 transporter genes were identified (Table 4), some being metal transporters. We have also identified genes for 24 metal-binding proteins (Table 4) like zinc finger, calcium, and iron-binding proteins.

4 DISCUSSION
4.1 Micronutrients Variation and Correlation Among Micronutrients
Micronutrients play an indispensable role in the growth and development of eukaryotic organisms. A deficiency in these essential micro-and macronutrients leads to abnormal growth in living systems. Humans get most of their micronutrients from plant and animal sources. Therefore, the biofortification of important food crops is necessary nowadays. In the present study, an initiative was taken to investigate different micronutrient concentrations in the germplasm of common beans. The micronutrient content of the seeds has been varied widely in common bean seeds (Supplementary Table S2). Previous studies have shown that mineral variation has been observed in almost all major legumes, including the common bean. The different mineral content in beans has been studied in different parts of the world including India (O Abidemi et al., 2012; Martinez Meyer et al., 2013; Zaccardelli et al., 2013; Kumar and Chopra, 2014; Mahajan R et al., 2015; Erdogmus et al., 2020). Earlier reports also suggested that seed mineral content showed huge variation in common bean germplasm.

FIGURE 4 | (A–G): Summary of Genome-wide association result: Manhattan plots depicting association of 113 SNP markers with seed. (A) Mg, (B) Mo, (C) Ca, (D) Cu, (E) Mn, (F) Zn, (G) Fe content in common bean.
Seed mineral concentrations such as Mg, Ca, Fe, Zn, and Cu of 60 common bean genotypes collected from the Western Himalayas varied from 1,220.5 to 2,737.5 ppm, 300–5,350 ppm, 80.5–180.6 ppm for Fe, 14.64–104.08 ppm, and 9.3–14.4 ppm, respectively (Jan et al., 2021). The average seed Ca concentration was recorded as 1.37 and 1.41 g kg\(^{-1}\), whereas, the average Fe seed content was recorded as 79.57 and

| Trait | Marker | Chr. No | Position | \(p\)-value | \(R^2\) (%) |
|-------|--------|---------|----------|-------------|------------|
| Cu    | Pv45224 | 11      | 31094041  | 2.07E-04    | 17.10      |
|       | Pv18893 | 11      | 6191264   | 2.90E-04    | 16.25      |

(Continued in next column)
85.95 mg kg\(^{-1}\) from common bean seeds obtained through pedigree and single seed descent methods respectively (Ribeiro et al., 2014). Also, 88.14 mg kg\(^{-1}\) of Fe, 49.24 mg kg\(^{-1}\) of Zn, 0.25 g 100 g\(^{-1}\) of Mg, 11.30 mg kg\(^{-1}\) of Cu, and 22.71 mg kg\(^{-1}\) of Mn was found in common bean genotypes from Universidade Federal de Lavras (UFLA) (Silva et al., 2010) and 74.6 ppm Fe, 39.9 ppm Zn content in Ugandan common bean germplasm (Mukamuhirwa et al., 2012). The difference in micronutrient content in bean genotypes from different parts of the world may be due to the different number of samples taken for evaluation, environmental conditions such as climate and soil composition, and agricultural techniques. This diversity in the germplasm can help us identify potential candidate lines that can be used in the development of Multi-Parent Advanced Generation Intercross (MAGIC) or bi-parental mapping populations, the breeding of micronutrient-rich, high-yielding varieties of beans, can be used for investigating different levels of gene expression for different nutrients in common beans. The higher micronutrient bean lines could be used for biofortification programs.

4.2 GWAS for Micronutrients

Genetic diversity is an important parameter for studying variability in any crop and identifying superior alleles controlling qualitative and quantitative traits through association mapping (Nachimuthu et al., 2015). Molecular markers such as SSRs and SNPs have an important role in studying genetic diversity in most crops (Nachimuthu et al., 2015; Zargar et al., 2016). Insights into the genomic diversity and population structure of common bean germplasm can expedite the genetic gains in common bean-breeding programs (Blair et al., 2012; Blair et al., 2013). The diversity based on the dissimilarity coefficient divided selected germplasm into five main groups and subgroups. The results of the clustering showed that the local germplasm of the common bean of Jammu & Kashmir (J&K) is very diverse and could be used as advanced lines for the genetic enhancement of the common bean. Further, intermixing of released and local lines collected from
different regions of J&K indicated that the selected germplasm is diverse. Some of the previously conducted diversity studies on local landraces of the common beans collected from the J&K region have also divided the germplasm into different groups (Zargar et al., 2016; Mahajan et al., 2017a; Mahajan et al., 2017b). In order to have knowledge about the sub-populations in a particular crop structure analysis was performed.

Based on the sharp peak for the delta-K value and the results of the PCA, 96 common bean genotypes were classified into five major groups. A similar pattern of population structure K = 5 was also found in common bean germplasm from Jammu and Kashmir (Mahajan et al., 2017a; Mahajan et al., 2017b). However, earlier studies also classified the common bean germplasm into K = 2 (Gupta et al., 2020; Nkhata et al., 2020; Mir et al., 2021); K = 3 (Blair et al., 2012; Blair et al., 2013; Dennis et al., 2014; Nemli et al., 2014). The difference in cluster and structure analysis could be due to different algorithms used by the two approaches. The cluster analysis is based on evolutionary dissimilarity, while the population structure is based on a Bayesian algorithm. The deviations in the results can be attributed to the different germplasm, the different marker system, and the different geographical locations.

For the identification of genes associated with different traits in a large population, GWAS offers much higher mapping resolutions (Mamo et al., 2014; Norton et al., 2014). To the best of our knowledge, there are only a few association studies on different traits in common bean germplasm collected from the Himalayan region (Mahajan et al., 2017b; Choudhary et al., 2018; Gupta et al., 2020; Mir et al., 2021). However, in earlier studies, genic and genomic SSRs were used for GWAS. The present study is the first report on the association of SNPs related to seed micronutrient content in common bean germplasm from the Himalayan region. In the present study, it was found that SNPs associated with most of the seed micronutrients (Mg, Mn, Fe, Ca, Cu) were found on chr.3 and chr.11 (Mg, Mn, Mo, Ca, Zn) whereas, chr.8 have SNPs associated with Cu, Mg and Mn and chr.9 have SNPs associated with only Mg and Ca. Earlier studies revealed that QTLs linked with Fe content were found on chr. 2, 5, 6, 7, 9, and 10, whereas Zn content was found on chr.1, 2, 3, 5, 7, 8, and 10 (Blair et al., 2011; Mahajan et al., 2017a). In a recent study, a single QTL on chr. 9 and chr. 8 was found to be associated with seed Ca content and seed Mg content, respectively, whereas two QTLs on chr. 6 were associated with seed Zn content (Gunjača et al., 2021). In our study, we found that all of the 11 chromosomes have SNPs associated with more than one seed micronutrient content, which indicates the preciseness of using high throughput genotypic data in the present study.
4.3 LD Plot and Haploblock in Common Bean Germplasm

Knowledge about the LD properties in domesticated crops is important as it underlies all types of genetic mapping and may be used in the fine mapping of genes associated with complex traits in crop plants. It is said that in-depth LD in a crop is important for having more SNP-based associations to predict the average number of markers required for GWAS (Nordborg and Tavaré, 2002). Identification of QTLs in association mapping is based on Linkage Disequilibrium measurement. The population structure and genetic relatedness between the genotypes can lead to false-positive LD analysis. LD pattern in germplasm is affected by reducing genetic diversity by various factors like the type of selection, population admixture (Contreras-Soto et al., 2017). Our study is in accordance with the previous common bean LD estimation studies (Erdogmus et al., 2020; Gunjača et al., 2021). Hence, we can conclude that as the number of SNPs increases, there will be more $R^2$ and the higher the likelihood of association of markers with traits of interest. This also indicated that a significant association would be possible with LD block having a higher number of SNPs compared to those windows having lower SNPs. The present study suggests that GBS is an advanced approach to analyzing genetic diversity and population structure in the common bean. The haplotype-based analysis showed that more haplocks were found in the centromeric region than in the telomeric region. In Chr.10 (26,295 kb) higher haplotypes per kb were found than in the rest of the chromosomes, which suggests that the LD decay in Chr.10 is stronger than in other chromosomes.

5 CANDIDATE GENES

GWAS is often used for comparative genome analysis and helps explicitly in dissection as well as in understanding the complex quantitative feature analysis. The GWAS helps identify significant SNPs associated with a trait that is not the function of the region of interest, and sometimes these SNPs are present in non-coding or non-regulatory regions of chromosomes (Bararyenya et al., 2020). Thus, it is important to identify the candidate genes in the vicinity of the significant SNPs, and in our study, the significant SNP regions were examined for the identification of putative protein-coding genes using the *P. vulgaris* genome. The genes that are present in the .1 Mb flanking region of significant SNPs are given in Supplementary Table S6. Many previous studies also reported some markers linked with traits but no genes in the
and 11 has more number of genes in 0.1 MB. The number of genes in SNPs revealed that chromosome 2, 6, 9, and 11 has more number of genes in 0.1 MB flanking region of SNPs has a high density of genes or the hot spot of QTLs.

### TABLE 4 | Identified gene related to transporter/metal transporter and metal ion binding proteins.

| QTL name | Gene ID | Chr name | Start | End | Strand | Description |
|----------|---------|----------|-------|-----|--------|-------------|
| QTL_Zn_2_37819199 | Phvul.002G20107000 Chr02 | 37776912 | 37782162 | –1 | ABC-2 TYPE TRANSPORTER |
| QTL_Zn_2_37819199 | Phvul.002G20106000 Chr02 | 37786780 | 37792275 | –1 | ABC TRANSPORTER |
| QTL_Mn_1_1240566 | Phvul.001G079200 Chr01 | 12009849 | 12013454 | –1 | CALCIUM-BINDING TRANSPORTER-LIKE PROTEIN |
| QTL_Mg_1_13517317 | Phvul.001G087300 Chr01 | 13547691 | 13551468 | 1 | ZINC FINGER PROTEIN CONSTANS-LIKE 14-RELATED |
| QTL_Mg_1_13577463 | Phvul.001G087600 Chr01 | 13632923 | 13640935 | –1 | ZINC FINGER PROTEIN CONSTANS-LIKE 14-RELATED |
| QTL_Mg_2_39493707 | Phvul.002G223300 Chr02 | 39466477 | 39469833 | 1 | OXIDOREDUCTASE, 2OG-FE II OXYGENASE FAMILY PROTEIN |
| QTL_Ca_3_38092190 | Phvul.003G162400 Chr03 | 38151881 | 38153482 | –1 | ZINC FINGER PROTEIN |
| QTL_Ca_3_38092190 | Phvul.003G162500 Chr03 | 38160568 | 38162888 | 1 | ZINC FINGER PROTEIN |
| QTL_Mg_6_26644872 | Phvul.006G163300 Chr06 | 26634344 | 26644212 | –1 | ZINC FINGER COCH-DOMAIN-CONTAINING PROTEIN-RELATED |
| QTL_Ca_9_8733609 | Phvul.009G043700 Chr09 | 8779518 | 8780270 | –1 | CALCIUM-ACTIVATED CHLORIDE CHANNEL |
| QTL_Ca_9_8733609 | Phvul.009G042600 Chr09 | 8633736 | 8637094 | 1 | PROBABLE ZINC-RIBBON DOMAIN (ZINC_RIBBON_12) |
| QTL_Mn_11_5199874 | Phvul.011G068500 Chr11 | 5228636 | 5232078 | 1 | CALCIUM-DEPENDENT CHLORIDE CHANNEL REGULATOR |
| QTL_Mn_11_5168818/ | Phvul.011G068500 Chr11 | 5228636 | 5232078 | 1 | CALCIUM-DEPENDENT CHLORIDE CHANNEL REGULATOR |
| QTL_Ca_9_8733609 | Phvul.009G043700 Chr09 | 8779518 | 8780270 | –1 | CALCIUM-DEPENDENT CHLORIDE CHANNEL REGULATOR |
| QTL_Ca_9_8733609 | Phvul.009G042600 Chr09 | 8633736 | 8637094 | 1 | PROBABLE ZINC-RIBBON DOMAIN (ZINC_RIBBON_12) |
| QTL_Mn_11_5168818/ | Phvul.011G068500 Chr11 | 5228636 | 5232078 | 1 | CALCIUM-DEPENDENT CHLORIDE CHANNEL REGULATOR |
| QTL_Ca_9_8733609 | Phvul.009G043700 Chr09 | 8779518 | 8780270 | –1 | CALCIUM-DEPENDENT CHLORIDE CHANNEL REGULATOR |
| QTL_Ca_9_8733609 | Phvul.009G042600 Chr09 | 8633736 | 8637094 | 1 | PROBABLE ZINC-RIBBON DOMAIN (ZINC_RIBBON_12) |

**6 CONCLUSION**

Micronutrient deficiency is the leading cause of human health deterioration worldwide. An animal or plant-based diet alone...
cannot provide humans with excessive amounts of essential micronutrients. Biofortification of food crops is therefore important in order to provide humans with essential micronutrients. Modern breeding approaches such as QTL mapping and association mapping are important to identify QTLs that are related to micronutrient levels in plants. In the present study, a significant variation in the micronutrient content of the seeds in the germplasm of common beans was found. The present data on genetic loci, particularly the key SNPs associated with seven elements, will be helpful in identifying candidate genes, understanding molecular mechanisms, and developing molecular markers for breeding applications. We firmly believe that the results of the current studies will help accelerate bean biofortification efforts to overcome nutritional deficiencies.

DATA AVAILABILITY STATEMENT
The data sets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

AUTHOR CONTRIBUTIONS
MN did experimental work and wrote the manuscript; RM helped in data analysis and editing the manuscript; SM helped in writing the manuscript; SR helped in experimental work; RAM helped with writing the manuscript and analysis; RS helped in data analysis; VT and VK: molecular data analysis; PS helped in fieldwork; HE-S and DH help in editing the manuscript; SZ developed the concept and guided all the laboratory work and edited the manuscript.

ACKNOWLEDGMENTS
MN and SZ acknowledge the financial support of SERB, New Delhi (Project File No: PDF/2016/000486) for the award of the National Postdoc Fellowship (NPDF) to MN under the mentorship of SZ. SZ also acknowledges Dr. Swarup Parida (NIPGR, New Delhi, India) for his support in genotyping; Dr. Humaira Sonah and Dr. Rupesh Deshmukh (NABI, Mohali, India) for their support in data analysis and editing the manuscript. SZ also acknowledges the Hon’ble Vice-Chancellor SKUAST-K (Prof. Nazir A. Ganai) for his support. The authors would like to extend their sincere appreciation to the Researchers Supporting Project Number (RSP-2021/19), King Saud University, Riyadh, Saudi Arabia.

SUPPLEMENTARY MATERIAL
The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.750814/full#supplementary-material
Concentrations in the Seeds of C. Arietinum and C. reticulatum. Genomics 111. doi:10.1016/j.jgenet.2018.12.010

Papamargaritis, D., Aasheim, T., Sampson, B., and le Roux, C. W. (2015). Copper, Selenium and Zinc Levels after Bariatric Surgery in Patients Recommended to Take Multivitamin- mineral Supplementation. J. Trace Elem. Med. Biol. 31, 167–172. doi:10.1016/j.jtemb.2014.09.005

Peterson, G. W., Dong, Y., Horbach, C., and Fu, Y. R. (2014). Genotyping-by-sequencing for Plant Genetic Diversity Analysis: A Lab Guide for SNP Genotyping. Diversity 6. doi:10.3390/d6040065

Pfeiffer, W. H., and McClafferty, B. (2007). “HarvestPlus: Breeding Crops for Better Nutrition,” in Crop Science. doi:10.2135/cropsci2007.09.0021P88

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. Genetics 155, 945959.

Ribeiro, N. D., Jost, E., Maziero, S. M., Storck, L., and Rosa, D. P. (2014). Selection of Common Bean Lines with High Grain Yield and High Grain Calcium and Iron Concentrations. Rev. Ceres 61. doi:10.1590/S0034-737X2014000100010

Reidinger, S., Ramsey, M. H., and Hartley, S. E. (2012). Rapid and Accurate Analyses of Silicon and Phosphorus in Plants Using a Portable X-ray Fluorescence Spectrometer. New Phyt. 195, 699706. doi:10.1111/j.1469-8137.2012.04179.x

Schilling, M. P., Wolf, P. G., Duffy, A. M., Rai, H. S., Rowe, C. A., Richardson, B. A., et al. (2014). Genotyping-by-sequencing for Populus Population Genomics: An Assessment of Genome Sampling Patterns and Filtering Approaches. PLoS ONE 9. doi:10.1371/journal.pone.0095292

Schmutz, J., McClean, P. E., Mamidi, S., Wu, G. A., Cannon, S. B., Grimwood, J., et al. (2014). A Reference Genome for Common Bean and Genome-wide Analysis of Dual Domestications. Nat. Genet. 46. doi:10.1038/ng.3008

Shao, Y., Jin, L., Zhang, G., Lu, Y., Shen, Y., and Bao, J. (2011). Association Mapping of Grain Color, Phenolic Content, Flavoround Content and Antioxidant Capacity in Dehulled rice. Theor. Appl. Genet. 122. doi:10.1007/s00122-010-1905-4

Silva, C. A., De Fátima Barbosa Abreu, A., Antonio, M., Ramalho, P., Duarte Correa, A., Contijo, L., et al. (2010). Genetic Variability for Protein and Minerals Content in Common Bean Lines. Phaseolus vulgaris, LEMBRAPA, Embrapa Scientific Information-Alice, Brazil.

Smedley, D., Haider, S., Ballester, H., Holland, R., London, D., Thorisson, G., and Kasprzyk, A. (2009). BioMart–Biological Queries Made Easy. BMC Genomics 10, 22. doi:10.1186/1471-2164-10-22

Sonah, H., Bastien, M., Iquira, E., Tardivel, A., Légaré, G., Boyle, B., et al. (2013). An Improved Genotyping by Sequencing (GWAS) Approach Offering Increased Versatility and Efficiency of SNP Discovery and Genotyping. PLoS ONE 8. doi:10.1371/journal.pone.0054603

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol. Biol. Evol. 28. doi:10.1093/molbev/msr121

Tian, F., Bradbury, P. J., Brown, P. J., Hung, H., Sun, Q., Flint-Garcia, S., et al. (2011). Genome-wide Association Study of Leaf Architecture in the maize Nested Association Mapping Population. Nat. Genet. 43. doi:10.1038/ng.746

Upadhyaya, H. D., Bajaj, D., Das, S., Kumar, V., Gowda, C. L. L., Sharma, S., et al. (2016). Genetic Dissection of Seed-Iron and Zinc Concentrations in Chickpea. Sci. Rep. 6. doi:10.1038/srep24850

Verma, S., Gupta, S., Bandhiwal, N., Kumar, T., Bharadwaj, C., and Bhatia, S. (2015). High- Density Linkage Map Construction and Mapping of Seed Trait (02.003) in Chickpea (Cicer Arietinum L.) Using Genotyping-By-Sequencing (GWAS). Sci. Rep. 5. doi:10.1038/srep17512

Wang, M., Yan, J., Zhao, J., Song, W., Zhang, X., Xiao, Y., et al. (2012). Genome-wide Association Study (GWAS) of Resistance to Head Smut in maize. Plant Sci. 196. doi:10.1016/j.plantsci.2012.08.004

Wen, W., Li, D., Li, X., Gao, Y., Li, W., Li, H., et al. (2014). Metabolome-based Genome-wide Association Study of maize Kernel Lends to Novel Biochemical Insights. Nat. Commun. 5. doi:10.1038/ncomms4438

Yang, Z., Li, Z., and Bickel, D. R. (2013). Empirical Bayes Estimation of Posterior Probabilities of Enrichment: A Comparative Study of Five Estimators of the Local False Discovery Rate. BMC Bioinformatics 14. doi:10.1186/1471-2105-14-87

Yeken, M. Z. Assessment of Mineral Content Variations for Biofortification of the Bean Seed UsalyeTohumunun Biyofortifikasyonu in MineralIeriKısılıVasıyanların Degerlendirilmesi.2018;4: 261–269. doi:10.24180/ijaws.455311

Yu, J., and Buckler, E. S. (2006). Genetic Association Mapping and Genome Organization of maize. Curr. Opin. Biotechnol. 17. doi:10.1016/j.copbio.2006.02.003

Zaccardelli, M., Pentangelo, A., and Tripodi, P. (2013). Characterization of bean (Phaseolus vulgaris L.) ecotype “Fagiolo occhio nero di Oliveto Citra” using agronomic, biochemical and molecular approaches. Pak. J. Biol. Sci. 16. doi:10.3923/pjbs.2013.901.910

Zargar, S. M., Farhat, S., Mahajan, R., Bakhri, A., and Sharma, A. (2016). Unraveling the Efficiency of RAPD and SSR Markers in Diversity Analysis and Population Structure Estimation in Common Bean. Saudij. J. Biol. Sci. 23. doi:10.1016/j.sjbs.2014.11.011

Zargar, S. M., Mahajan, R., Farhat, S., Nazir, M., Mir, R. A., Nazir, M., et al. (2015). Understanding the Role of Iron and Zinc in Animals and Crop Plants from Genomics Perspective. Curr. Trends Biotechnol. Pharm. 9.

Zargar, S. M., Sharma, A., Sadhu, A., Agrawal, G. K., and Rakwal, R. (2014). Exploring Genetic Diversity in Common Bean from Unexploited Regions of Jammu & Kashmir-India. Mol. Plant Breed. doi:10.5376/mbp.2014.05.0002

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Nazir, Mahajan, Massoor, Rasool, Mir, Singh, Thakral, Kumar, Sofi, El-Serehy, Hefti and Zargar. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.