frontiers in Plant Science

ORIGINAL RESEARCH
published: 01 May 2017
doi: 10.3389/fpls.2017.00412

Deciphering Genomic Regions for High Grain Iron and Zinc Content Using Association Mapping in Pearl Millet

N. Anuradha[1], C. Tara Satyavathi[1], C. Bharadwaj[1], T. Nepolean[1], S. Mukesh Sankar[1], Sumer P. Singh[1], Mahesh C. Meena[2], Tripti Singhal[1] and Rakesh K. Srivastava[3]**

[1] Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi, India, [2] Division of Soil Science and Agricultural Chemistry, ICAR-Indian Agricultural Research Institute, New Delhi, India, [3] International Crops Research Institute for the Semi-Arid Tropics, Patancheru, India

OPEN ACCESS

Edited by:
Manoj Prasad, National Institute of Plant Genome Research, India

Reviewed by:
Akila Chandra Sekhar, Yogi Vemana University, India
Prashant Vikram, International Maize and Wheat Improvement Center (CIMMYT), Mexico
Deepmala Sehgal, International Maize and Wheat Improvement Center (CIMMYT), Mexico

***Correspondence:**
C. Tara Satyavathi
csatyavathi@gmail.com
Rakesh K. Srivastava
r.k.srivastava@cgiar.org

Specialty section:
This article was submitted to Plant Genetics and Genomics, a section of the journal Frontiers in Plant Science

Received: *26 October 2016*
Accepted: *10 March 2017*
Published: *01 May 2017*

Citation:
Anuradha N, Satyavathi CT, Bharadwaj C, Nepolean T, Sankar SM, Singh SP, Meena MC, Singhal T and Srivastava RK (2017) Deciphering Genomic Regions for High Grain Iron and Zinc Content Using Association Mapping in Pearl Millet. Front. Plant Sci. 8:412. doi: 10.3389/fpls.2017.00412

Micronutrient malnutrition, especially deficiency of two mineral elements, iron [Fe] and zinc [Zn] in the developing world needs urgent attention. Pearl millet is one of the best crops with many nutritional properties and is accessible to the poor. We report findings of the first attempt to mine favorable alleles for grain iron and zinc content through association mapping in pearl millet. An association mapping panel of 130 diverse lines was evaluated at Delhi, Jodhpur and Dharwad, representing all the three pearl millet growing agro-climatic zones of India, during 2014 and 2015. Wide range of variation was observed for grain iron (32.3–111.9 ppm) and zinc (26.6–73.7 ppm) content. Genotyping with 114 representative polymorphic SSRs revealed 0.35 mean gene diversity. STRUCTURE analysis revealed presence of three sub-populations which was further supported by Neighbor-Joining method of clustering and principal coordinate analysis (PCoA). Marker-trait associations (MTAs) were analyzed with 267 markers (250 SSRs and 17 genic markers) in both general linear model (GLM) and mixed linear model (MLM), however, MTAs resulting from MLM were considered for more robustness of the associations. After appropriate Bonferroni correction, *Xpsmp* 2261 (13.34% R^2-value), *Xipes* 0180 (R^2-value of 11.40%) and *Xipes* 0096 (R^2-value of 11.38%) were consistently associated with grain iron and zinc content for all the three locations. Favorable alleles and promising lines were identified for across and specific environments. PPMI 1102 had highest number (7) of favorable alleles, followed by four each for PPMFeZMP 199 and PPMI 708 for across the environment performance for both grain Fe and Zn content, while PPMI 1104 had alleles specific to Dharwad for grain Fe and Zn content. When compared with the reference genome Tift $23D_2B_1$-P1-P5, *Xpsmp* 2261 amplicon was identified in intergenic region on pseudomolecule 5, while the other marker, *Xipes* 0810 was observed to be overlapping with aspartic proteinase (*Asp*) gene on pseudomolecule 3. Thus, this study can help in breeding new lines with enhanced micronutrient content using marker-assisted selection (MAS) in pearl millet leading to improved well-being especially for women and children.

Keywords: pearl millet, iron and zinc, association mapping, SSR, favorable alleles

INTRODUCTION

There is an increasing global attention towards addressing micronutrient malnutrition as the health impairment caused due to poor quality diet is more wide spread than low energy intake (Murgia et al., 2012). Among all mineral micronutrients, iron and zinc deficiency accounts for more than two billion people globally (WHO, 2012). One-third of its share is from India alone (Barthakur, 2010). It is more prominent in inhabitants of low income countries who depend on cereal based low quality staple food (Tiwari et al., 2016). Anemia, caused by iron deficiency is the most common disorder in such countries and people, who consume low quality diet are prone to possible risk of child mortality and other physiological disorders (Tako et al., 2015). Zinc is also an important micronutrient, which is required for proper growth, whose deficiency leads to stunting, increased susceptibility to many infectious diseases, morbidity and low mental ability (Deshpande et al., 2013). One among successful approaches to alleviate hidden hunger is biofortification, so that micronutrients can be delivered effectively to rural poor residents who majorly depend on staple food crops (Bouis and Welch, 2010; Saltzman et al., 2013). This can be easily achieved through molecular breeding. The biofortified products are cost-effective, sustainable, and remain within the purchasing power of rural poor (Pfeiffer and McClafferty, 2007; Bouis and Welch, 2010).

Pearl millet [*Pennisetum glaucum* (L.) R. Br.] is a climate resilient cereal which can be grown in harsh and adverse agricultural environments with scanty rainfall and less access to irrigation and fertilizers (Vadez et al., 2014). It exhibits greater level of tolerance to drought, heat stress and high temperature (Shivhare and Lata, 2016). It is widely cultivated as a source of food and fodder in arid and semi-arid regions of sub-Saharan Africa and India (Rao et al., 2006). It is gifted with many nutritional qualities compared to other cereals. It contains high amount of fiber, α-amylose (Nambiar et al., 2011), metabolizable energy, protein, essential amino acid, macro and micro nutrients like phosphorus, magnesium, iron and zinc and thus ensuring food and nutritional security. Pearl millet flour is used in fortification of food to improve nutritional quality of food products (Sonkar and Singh, 2015). Hence, consumption of these products were suggested by experts/physicians to all age groups, pregnant and nursing women for their proper mental and physical development. Consumption of pearl millet also forestalls cancer and Type-2 diabetes (T2D) and for relieving of celiac and several other non-communicable diseases (Nambiar et al., 2011).

Iron deficiency is prevalent in many parts of Africa and Asia, especially in India. Pearl millet provides 80–85% of total calorie intake per day (Tako et al., 2015) in Sahel region of Africa. The National Family Health Survey (NFHS) studies conducted in India indicate that more than 50% women (adolescent girls, lactating, and pregnant women) are anemic in states of Rajasthan, Haryana, Gujarat, and Maharashtra (Press Information Bureau, 2013) where pearl millet is largely grown (90% of Indian total production) and consumed (Rao et al., 2006; Finkelstein et al., 2015). Pearl millet can contribute 30–40% of these essential micronutrients and forms the cheapest source of staple food for iron and zinc in its growing regions (Rao et al., 2006). Hence development of micronutrient dense varieties or hybrids of pearl millet and their consumption would meet the required recommended dietary allowance (RDA) for these micronutrients and also would pave way to development of biofortified pearl millet.

Screening of germplasm lines for enhanced grain micronutrients and mapping the related QTLs will help in selection of parental lines or donors for developing iron/zinc rich hybrids or varieties. Despite having knowledge on different genes involved for iron and zinc uptake to final sequestration in grain (Grotz and Guerinot, 2006; Krohling et al., 2016), mapping for these traits is limited in most of the cereal crops (Kumar et al., 2010, 2016; Jin et al., 2015; Yu et al., 2015; Crespo-Herrera et al., 2016).

Accumulation of these micronutrients in the grain/edible portion is a complex mechanism involving many genes and moreover influenced by environment. One approach to dissect QTLs is through association mapping. It is a substitute to QTL mapping (Gómez et al., 2011) which utilizes the principle of linkage disequilibrium (LD) to find significant association of a molecular marker/QTL with a trait (Flint-Garcia et al., 2003; Gupta et al., 2005). It offers a great deal of advantage over linkage mapping in terms of higher mapping resolution since it accounts for historical mutations and recombinations in lineages leading to identification of markers nearby causative genes (Liu et al., 2016). More number of alleles can be considered, as broader population is included for investigation and allele mining can be attempted by exploitation of genetic diversity in a reference population (Flint-Garcia et al., 2003). SSR markers often are preferred in association mapping studies compared to other markers because of their co-dominant inheritance, multi-allelic nature, high reproducibility, and genome wide coverage (Varshney et al., 2005). Theoretically, SSRs are more powerful in detecting QTLs compared to SNPs (Ohashi and Tokunaga, 2003). With this background, the present study was undertaken to examine localization of genomic regions linked with enhanced grain iron and zinc content and its association with SSRs and genic markers using association mapping.

MATERIALS AND METHODS

Plant Material

The association mapping panel comprised of 130 diverse pearl millet lines including two checks, namely, ICTP 8203Fe and ICMB 98222 for grain iron and zinc content (**Table S1**). The association panel represents B- lines (seed parents), R- lines (restorers or pollen parents) and advanced breeding lines derived from different parts of India and some introduced selections from Africa.

Field Experiments and Evaluation

The experiment was conducted at three diverse geographical locations, representing all the three pearl millet growing agro-climatic zones of India, (i) ICAR-Indian Agricultural Research Institute Research farm, New Delhi (28° 382′ N, 77° 802′ E) representing Zone A receiving more than 400 mm annual rainfall (ii) ICAR-National Bureau of Plant Genetic Resources, Regional

Station farm, Jodhpur (26° 252′ N, 72° 992′ E) falling in zone A_1 with annual rainfall less than 400 mm and (iii) ICAR-IARI Regional Centre farm, Dharwad (15° 212′ N, 75° 052′ E) from zone B (covering the southern peninsular India). In all locations, the trial was taken for 2 consecutive years (Rainy Season, 2014 and 2015). The planting of genotypes was taken up in an alpha-lattice design (Patterson and Williams, 1976), with three replications. Each replication comprised of 13 incomplete blocks with 10 entries in each block. Each accession was represented in 2 rows of 4 m length with a spacing of 50 cm between rows and 15 cm from plant to plant. All the experiments were managed as per recommended agronomic practices across the locations and years to raise a normal healthy crop.

Estimation of Iron and Zinc Contents in Grain

Open-pollinated panicles from five representative plants were harvested for each accession at physiological maturity, threshed with a wooden mallet and then cleaned, while taking utmost care to avoid dust or metal contamination of the samples. Further, oven-dried samples were used on duplicate basis to estimate iron and zinc by di-acid digestion (Singh et al., 2005) followed by readings taken on an atomic absorption spectrometer (AAS, ZEE nit 700 tech Analytikjena). Genotypes, which recorded high iron content (>77 ppm, as this is the criterion given by Harvest Plus) were also tested for aluminum content in ICPMS (Nex ION 300X, Perkin Elmer USA) in order to determine, whether higher grain iron content could be due to contamination with soil, dust or is innate (Pfeiffer and McClafferty, 2007).

The data over locations and years were subjected to combined analysis of variance across the six environments using PROC GLM with random statement, considering environments as random and genotypes as fixed effect in SAS v9.2 (http://iasri.res.in/design). The adjusted means thus obtained after analysis were used to generate all the descriptive statistics using IBM SPSS v20. The broad-sense heritability was calculated as per the formula:

$$H^2 = \sigma^2_G/(\sigma^2_G + \sigma^2_e/r) \text{.....for single environment}$$
$$H^2 = \sigma^2_G/(\sigma^2_G + \sigma^2_{GE}/n + \sigma^2_e/nr) \text{.....for multi environment}$$

where, σ^2_G is genotypic variance, σ^2_{GE} is genotype × environment interaction variance, σ^2_e is error variance, *n* is number of environments and *r* is number of replications.

Datasets for analysis at Delhi during 2014 and 2015 were referred as Del-14 and Del-15. Similarly, data of Dharwad in 2014 and 2015 as DW-14 and DW-15. Likewise, Jodhpur data during 2014 and 2015 as Jod-14 and Jod-15. Mean data over years in a single location as Y14-M and Y15-M. Mean data across locations in a year were referred as Del-M, DW-M, and Jod-M. Mean data over all the six environments as grand mean (GM). Altogether, 12 datasets were included for analysis with TASSEL using both general liner model (GLM) and mixed linear model (MLM), but the results of only MLM are presented here. Further, though STRUCTURE and neighbor-joining (NJ) results were similar, we chose Q matrix files over PCA files due to greater relevance of the former with the pedigree of population for MLM analysis.

DNA Extraction and Genotyping

Pooled leaf samples from five plants for each genotype were collected at 2 weeks stage and genomic DNA of all genotypes in the association mapping panel was isolated using modified CTAB method (Murray and Thompson, 1980). The PCR was performed in a 10 μl volume consisting of 1.0 μl DNA (25 ng/μl), 1.0 μl 10 X Buffer (Banglore Genei), 0.60 μl dNTP (10 mM, Banglore Genei), 0.20 μl (25 mM $MgCl_2$, Banglore Genei), 1.0 μl each forward and reverse primer (10 mM), 0.13 μl Taq polymerase (3 U/μl, Banglore Genei) and the final volume is made up to 10 μl with autoclaved distilled water for every reaction. Thermo cycling was carried out in 384 well-blocked thermal cycler machine (GenePro, Bioer Technology Co. Ltd.) with a thermal cycling program having initial denaturation for 5 min at 94°C, followed by 40 cycles of 30 s at 94°C for denaturation, 30 s at temperatures as indicated in **Table S2**, extension at 70°C for 30 s and finally, 7 min at 72°C with final extension.

Analysis of Molecular Data

Population Structure and Genetic Relatedness

Initially, 114 polymorphic SSRs distributed in the entire genome was utilized for population structure and familial relatedness analysis. To realize the population structure, a Bayesian model-based program implemented in STRUCTURE v2.3.4 as suggested by Pritchard et al. (2000) was used with number of sub-populations (K) set from 1 to 12 with burn-in length and MCMC both set to 2,00,000. By using admixture model with correlated allele frequencies, K is replicated 10 times. The best *K*-value could not be determined easily by considering the log likelihood value [LnP(D)] of STRUCTURE output, hence, an *ad-hoc*, ΔK (Evanno et al., 2005) was determined to reveal number of sub-groups by importing structure output result zip file as input in the structure harvester program (http://taylor0.biology.ucla.edu/structureHarvester/index.php). Genotypes with $Q \geq 0.6$ were allotted to corresponding sub-groups A, B, or C, and those with values <0.6 were allotted to a mixed sub-population, admixture. Genetic relatedness or K matrix is generated from TASSEL *v*3 (Trait Analysis by Association Evolution and Linkage, Bradbury et al., 2007).

Population grouping pattern obtained from STRUCTURE was further supported with the neighbor-joining (NJ) tree based on Nei's genetic distance (Nei, 1972) among the genotypes and also supplemented by principal coordinate analysis (PCoA) using DARwin v6 (Perrier et al., 2003). AMOVA was attempted by Arlequin (Excoffier and Lischer, 2010) by considering A, B, C, and admixture sub populations obtained in STRUCTURE.

Linkage Disequilibrium (LD) and Marker-Trait Association (MTA)

For association mapping analysis, a total of 267 polymorphic markers were used of which 250 were SSR markers and 17 were genic markers. To target genes involved in the pathway of iron and zinc uptake from soil to final sequestration into grains by plants, 55 genic primers were screened for polymorphism. Twenty primer pairs were selected from previous studies on maize (Mondal et al., 2014) and rice (Sperotto et al., 2010). Further, orthologous genes were searched in

phytozome database of other related crops for cross amplification in pearl millet, since the genome sequencing in this crop (http://ceg.icrisat.org/ipmgsc/) is still underway and not available in public domain. After identification of orthologous genes in foxtail millet and sorghum genome, a total of 35 gene specific primers (foxtail millet: 22 and sorghum: 13) were designed with the help of Primer3Plus web application (http://primer3plus.com/cgi-bin/dev/primer3plus.cgi). Out of 55 genic markers used, 17 were observed to be polymorphic and they were further utilized in the study. TASSEL *v*3 (Bradbury et al., 2007) with 1,000 permutations was used to generate an LD plot with r^2 and *p*-values among all the 267 polymorphic markers. MTA was studied in TASSEL v3 (http://www.maizegenetics.net) with both general linear model (GLM) and mixed linear model (MLM) (Yu et al., 2006) as given below (Tadesse et al., 2015).

$$y = Xa + Qb + e \text{ .....GLM and}$$
$$y = Xa + Qb + Zu + e \text{ ..... MLM}$$

where,

y is phenotype vector, a is marker vector with fixed effects, b is a vector with fixed effects, u is a vector with random effects (kinship matrix), e is a residuals vector, X denotes the accessions/genotypes at the marker, Q is the Q-matrix, result of STRUCTURE software, Z is an identity matrix.

A widely adopted threshold for the significance, Bonferroni correction (Bland and Altman, 1995; Li C. Q. et al., 2016) was used uniformly for all traits to observe association between trait and markers. In this context, $p < 3.74 \times 10^{-3}$ (where, $p = 1/n$ and $n =$ number of total markers used (267); also $-\log10\ (1/267) = 2.43$). Quantile-quantile (QQ) plots, plotted against expected and observed *p*-values to assess the adequacy of Type I error were generated with TASSEL (**Figures S1**, **S2**).

The position of the marker trait associations were determined following the published consensus linkage map of pearl millet by Rajaram et al. (2013).

Further, the significantly associated markers were tested for associations at individual sub-population level using one way ANOVA (http://vassarstats.net/anova1u.html), assuming null hypothesis of no phenotypic difference between different alleles of a marker.

Favorable Allele Mining

Initially, Breseghello and Sorrells (2006) utilized null alleles (missing and rare alleles were also included) to calculate the phenotypic effect of an allele for identified MTAs. But all SSR markers may not have null alleles, sometimes rare alleles themselves are superior alleles. Hence, population mean was introduced in place of null alleles (Cai et al., 2014). Favorable allele for a marker loci associated with specific trait was identified using the formula (Cai et al., 2014; Li X. et al., 2016).

$$a_i = \frac{\Sigma x_{ij}}{n_i} - \frac{\Sigma N_k}{n_k}$$

where,

a_i is ith allele phenotypic affect, x_{ij} is phenotypic value of ith allele in jth genotype, n_i is number of accessions with ith allele, N_k is phenotypic value over all genotypes and n_k is total number of genotypes. When, $a_i > 0$, then this allele is said to have positive effect on the trait. When $a_i < 0$, the allele gives a negative effect (Zhang et al., 2013). Since, positive alleles give an increment, we considered superior positive alleles for planning crosses to pyramid maximum possible number of favorable alleles together for enhancement of grain iron and zinc content.

Validation of MTAs Obtained in Association Mapping

Four lines having higher mean iron and zinc content (PPMI 1102, PPMI 708, PPMI 683, PPMI 1107) and check (ICTP 8203Fe) along with two lines (5540B and PPMI 1155) with low mean iron and zinc content were selected for validation. Three consistently associated SSR marker primers, IPES 0096, IPES 0180, and PSMP 2261 were used to amplify DNA of all the seven selected genotypes. The amplified products were sequenced by standard Sanger's sequencing method by M/S Chromous Biotech, Bangalore. The sequenced reads were compared using BLAST against recently sequenced pearl millet reference genome,Tift23$D_2B_1^-$P1-P5 (unpublished) using MEGA6 (Tamura et al., 2013).

RESULTS

Molecular Genetic Diversity Analysis in Association Mapping Panel

Initially, 114 polymorphic SSR markers, covering the entire genome were used to characterize pearl millet association panel to assess the genetic relationship. A total of 294 alleles were detected with an average number of 2.65 alleles per locus (**Table S3**). The observed number of alleles (n_a) ranged from 2 to 5. Maximum number of five alleles were detected by genomic SSR loci, *Xpsmp* 2081. Four alleles per loci were observed in *Xipes* 200, *Xipes* 126 and *Xpgird* 50, and *Xpsmp* 2086. The mean effective allele number per loci (n_e) for 114 polymorphic markers studied was 1.64. This shows that there exists a difference of 37.89% in observed and expected number of alleles. The discrepancy observed between n_a and n_e indicates that alleles detected were having frequencies <5% (Sehgal et al., 2015). The average major allele frequency was 0.75 with a range of 0.32 (*Xpsmp* 2066) to 0.98 (*Xipes* 0079. Measure of gene diversity, Nei's gene diversity (Nei) was calculated across the population. *Xipes* 0079 had the least value of 0.03 while highest value (0.73) was recorded by *Xpsmp* 2066 with a mean of 0.35. A frequently used measure to know the polymorphism information of a marker, PIC ranged from 0.03 (*Xipes* 0079) to 0.68 (*Xpsmp* 2066). In this study, seven SSRs (*Xpsmp* 2066, *Xpsmp* 2081, *Xpsmp* 2263, *Xpsmp* 2086, *Xipes* 0200, *Xipes* 0163, and *Xicmp* 3066) had a PIC value more than 0.5 and in total 67 markers were showing moderate PIC values.

Population Structure and Cluster Analysis

The true association is known only after considering population structure (Yu et al., 2006). A Bayesian model based approach, STRUCTURE, assigns individuals to sub-population based on genotyping. Pritchard et al. (2000) was of the opinion that, actual number of sub-populations in a panel may not be easy to identify,

but one should consider the smallest K-value which captures the major composition in the population. Evanno et al. (2005) suggested the use of ΔK-value based on standard deviation to find the true value of K. The LnP(D) graph (**Figure 1A**) depicted with number of sub-populations(K) on x-axis and logarithmic of probability distribution on y-axis did not give a clear picture of the true value of K, so ΔK value was plotted against number of sub-population (**Figure 1B**) showed the highest peak at $K = 3$, which indicated that there are three sub-populations in the panel. The genotypes were assigned to individual sub-population keeping in view of highest membership likelihood criterion based on Q-values obtained from STRUCTURE software. The decision point taken here is, a genotype indicating a Q-value more than or equal to 0.6 is assigned to that particular sub group A or B or C and remaining genotypes which did not reach $Q = 0.6$ were allotted to admixture group (**Table S1**). In this manner, the first sub-population A is the biggest group containing 50 lines, B had 22, and C had 31, while remaining genotypes (30) were kept in admixture. These three sub populations were represented by three colors in the bar plot showing three sub-populations (**Figure 1C**). The sub population A is depicted by red, B by green, and C by blue color. Here, sub-population A is said to have fewer admixtures (0.19%) compared to B (0.29%) or C (0.21%) and in overall 23.08% of total population is under admixture. Sub-population A consisted of genotypes from different sources of pedigree, which included B lines while sub-population B comprised of few promising lines derived from a cross between, PPMI 683 and PPMI 627, few lines from 843B × 841B lines and some low iron Jamnagar lines and few other breeding lines. Sub-population C, majorly included genotypes selected for good agronomic score in a cross between PPMI 683 and PPMI 627, few lines derived from a cross between WGI 148 and WGI 52, and few others. Admixture consisted of thick panicle restorer lines, early maturing lines and dual purpose lines.

The dendrogram generated by DARwin using neighbor joining, sub-classified population into three groups. The maximum number of individuals was clustered in group 1 consisting 75 genotypes (**Figure 2**), whereas group 2 consisted of 29 and group 3 comprised of 26 individuals. Cluster 1 was only 64.0% similar to sub-population A, whereas Cluster 2 had 72.4% similarity with sub-population B while cluster 3 shared 88.5% similarity with sub-population C.

The principal coordinate analysis further supported this clustering pattern, where all the breeding lines were distributed in all four quadrants (**Figure 3**). Cluster I, similar to sub-population A of STRUCTURE analysis was observed in first quadrant. In same way, cluster II, similar to sub-population B fell in second quadrant and cluster III, similar to sub-population C, was mostly in third and a little of fourth quadrant. Remaining genotypes which belong to admixture in STRUCTURE were dispersed throughout all the quadrants.

Results from AMOVA showed that significant differences exist between sub populations which accounted for 11.8% of total variation. Within sub-population variation was much larger, contributing 82.2% to total variation and within individual variation was 6.0% (**Table 1**).

Variation for Grain Iron and Zinc Content in the Association Mapping Panel

Pearl millet grain iron and zinc content were having a wide range of variation at all six environments studied showing significant differences between genotypes in the analysis of variance (**Tables S4A–E**). Grain iron content (grand mean) ranged from 32.3 to 111.9 ppm whereas grain zinc content ranged from 26.6 to 73.7 ppm (**Tables 2A,B**, **Tables S5A,B**). The broad sense heritability estimates were high (>0.8) at all the individual environments (**Tables 2A,B**). It ranged from 0.84 (Jod-15) to 0.89 (DW-15) for grain zinc content while for grain iron content it was 0.85 (Del-14) to 0.92 (Jod-14). In pooled environments, the highest heritability value was recorded at Dharwad location (0.86) while the lowest during the year 2015 (0.60) across all the three locations for grain zinc content with 0.79 combined over all six environments. Grain iron content exhibited the highest heritability at Delhi location (0.8) while the lowest at Dharwad location (0.65). Hence, these results indicate that heritability was high for both grain iron and zinc content.

LD and Association of Markers with Grain Iron and Zinc Content in Different Environments

The LD measured as squared value of Pearson correlation, R^2 ranged from 0 to 1 with a mean of 0.017. Highest R^2-value was observed between *Xpsmp* 2063 and *Xipes* 0145 on LG7, *Xipes* 0103 and *Xipes* 0146 of LG1 (**Figure 4**). Out of 35,511 pair-wise combinations obtained from 267 marker loci, 69 pairs had R^2 more than 40%. Only 9.53% of the 35,511 pairs were in extent of LD ($p < 0.05$) across all the genotypes accessed. Before performing the association mapping, correlations between grain iron and zinc content and population structure was computed (**Table S6**). Grain iron was significantly associated with structure at DW-15, while grain zinc content recorded its significance at DW-14, DW-15, and DW-M. In remaining datasets, there is no correlation of the trait with the structure. The Q+K model holds good in all the cases, but when there is correlation of trait with the structure, Q+K model is a better model which avoids spurious associations (Yu et al., 2006). The threshold for identifying significant associations with a trait kept here was $p < 0.0037$, which was a Bonferroni corrected value considered by many workers in association mapping studies. Significant associations with grain iron and zinc content were tested across six different environments and their six means in different combinations.

Environment-Wise Marker-Trait Associations—Iron

The MLM analysis (**Table 3**) identified *Xipes* 0180 and *Xsinramp* 6 to be significantly associated with the trait explaining 12.4 and 10.1% variation, respectively, in Delhi, 2014. During the year 2015 at Delhi, *Xpsmp* 2261 recorded significant association with grain iron content. Mean of Delhi over 2 years, indicated *Xpsmp* 2261, *Xipes* 0096 and *Xsinramp* 6 to be highly associated with grain iron with an R^2 more than 9.0%. Dharwad 2014 dataset analysis indicated significance of only one SSR marker

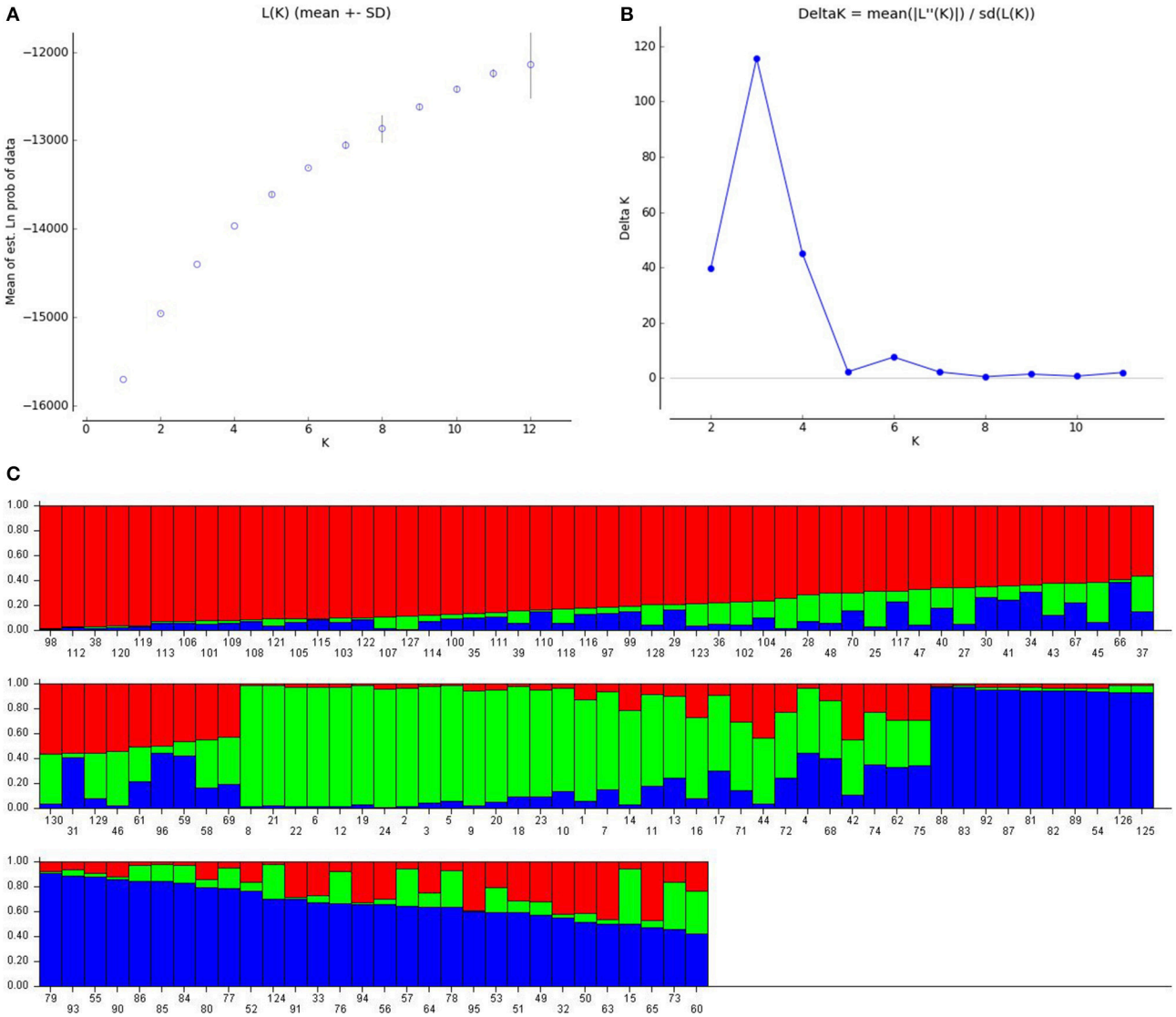

FIGURE 1 | Population stratification of pearl millet association mapping panel (A) LnP(D), the log likelihood values from $k = 1$ to 12. **(B)** Delta K, rate of change from 2 to 11. **(C)** Inferred population structure ($K = 3$, each color represents one sub-population).

(*Xicmp* 3092) with grain iron content, while during the year 2015 at Dharwad, *Xpsmp* 2209 was observed to be associated. No association was observed for grain iron content at Dharwad over mean of 2 years.

At Jodhpur during the year 2014, *Xpsmp* 2261 and *Xipes* 0096 were observed to be associated with the trait, where *Xpsmp* 2261 explained phenotypic variance of 15.7%, while another marker, *Xipes* 0096 recorded more than 8.0% variation. *Xpsmp* 2261 showed statistical significance by explaining more than 10% variation in the year 2015 at Jodhpur. Mean iron content of Jodhpur over 2 years revealed only one marker locus, *Xpsmp* 2261 to associate with the trait by explaining 18.1% variation.

During 2014, the data pooled over locations identified *Xpsmp* 2261, *Xipes* 0096, and *Xipes* 0180 to associate with grain iron content. Similarly, during year 2015, *Xpsmp* 2261 and *Xipes* 0096 were observed to be significantly associated with the trait. Likewise, three SSR loci, *Xpsmp* 2261, *Xipes* 0096, and *Xipes* 0180 were recorded to be associated with grain iron content in grand mean data set.

Association with Grain Zinc Content

During 2014 at Delhi, *Xpsmp* 2261 recorded significant association with grain zinc content explaining variance of more than 10% (**Table 3**). MLM analysis of the data during the year 2015 at Delhi, showed only one marker, *Xipes* 3016 to be significantly associated with grain zinc with more than 8.0% variation explained in the model, while *Xpsmp* 2261 was observed to be significantly associated with the mean over 2 years at Delhi. At Dharwad, during the year 2014, *Xipes* 0096 was observed to be associated with 12.7% variation, whereas, during the year

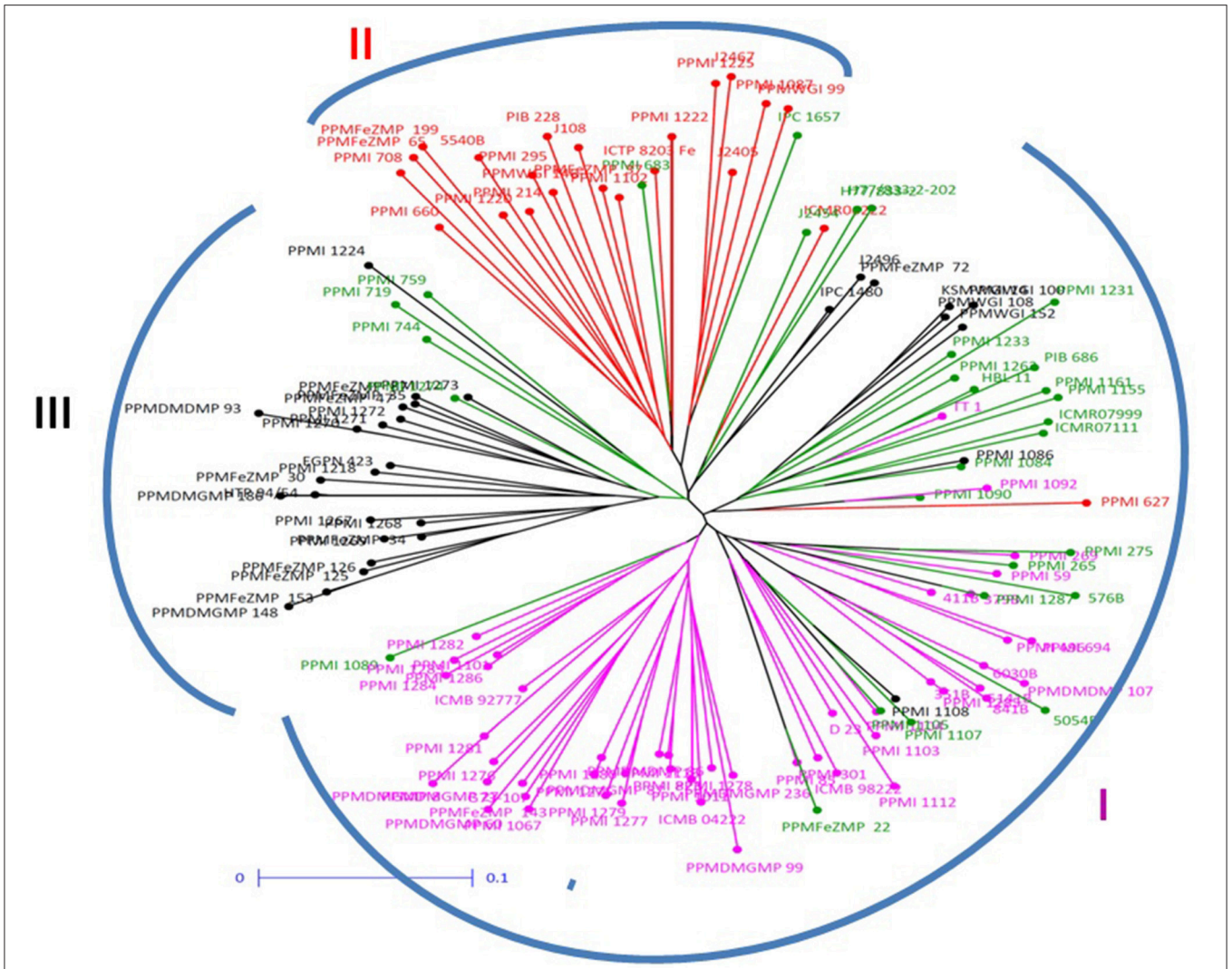

FIGURE 2 | Neighbor joining tree constructed for pearl millet association mapping panel (genotypes represented in different colors corresponding to the three sub-populations and admixture observed in STRUCTURE).

2015 at Dharwad, *Xicmp* 4006 revealed its significance with an R^2 of 10.4%, while *Xipes* 0096 and *Xsinramp* 6 were recorded as significant for Dharwad mean over 2 years.

Jodhpur, in the year 2014 recorded three SSRs (*Xicmp* 3004, *Xipes* 0224, and *Xpsmp* 2261) to be significantly associated with grain zinc content explaining more than 10% variation. In the year 2015 at Jodhpur, *Xpsmp* 2261 and *Xipes* 0180 were associated significantly, whereas the mean of Jodhpur over 2 years revealed three SSR markers (*Xicmp* 3004, *Xpsmp* 2213, and *Xpsmp* 2261) to be associated with grain zinc content. Mean zinc content of the years, 2014 and 2015 revealed *Xipes* 0096 and *Xpsmp* 2261 to be associated significantly with the trait explaining more than 10% variation each. When association were observed with the grand mean, over all locations and years, *Xipes* 0096 and *Xpsmp*2261 were found to be significantly associated with grain zinc content with R^2 more than 11.0% individually.

Frequency of Significant Markers in MTAs with Grain Iron and Zinc Content

Table S7A shows the number of times a marker was observed as significantly associated with grain iron or zinc content *via* MLM analysis. Out of total 6 markers identified to be associated with grain iron content, *Xpsmp* 2261 recorded significance in highest number of times (Del-15, Jod-14, Jod-15, Del-M, Jod-M, Y14-M, Y15-M, and GM). Further, higher number of associations were observed with *Xipes* 0096 which recorded as significant in five datasets (Del-M, Jod-14, Y14-M, Y15-M, and GM) followed by *Xipes* 0180 which recorded significant associations in four datasets (Del-14, Del-M, Y14-M, and GM). Genic marker, *Xsinramp* 6 identified to be associated with iron content in two datasets (Del-14 and Del-M) while *Xicmp* 3092 and *Xpsmp* 2209 were significant at Dharwad location during the year 2014 and 2015, respectively.

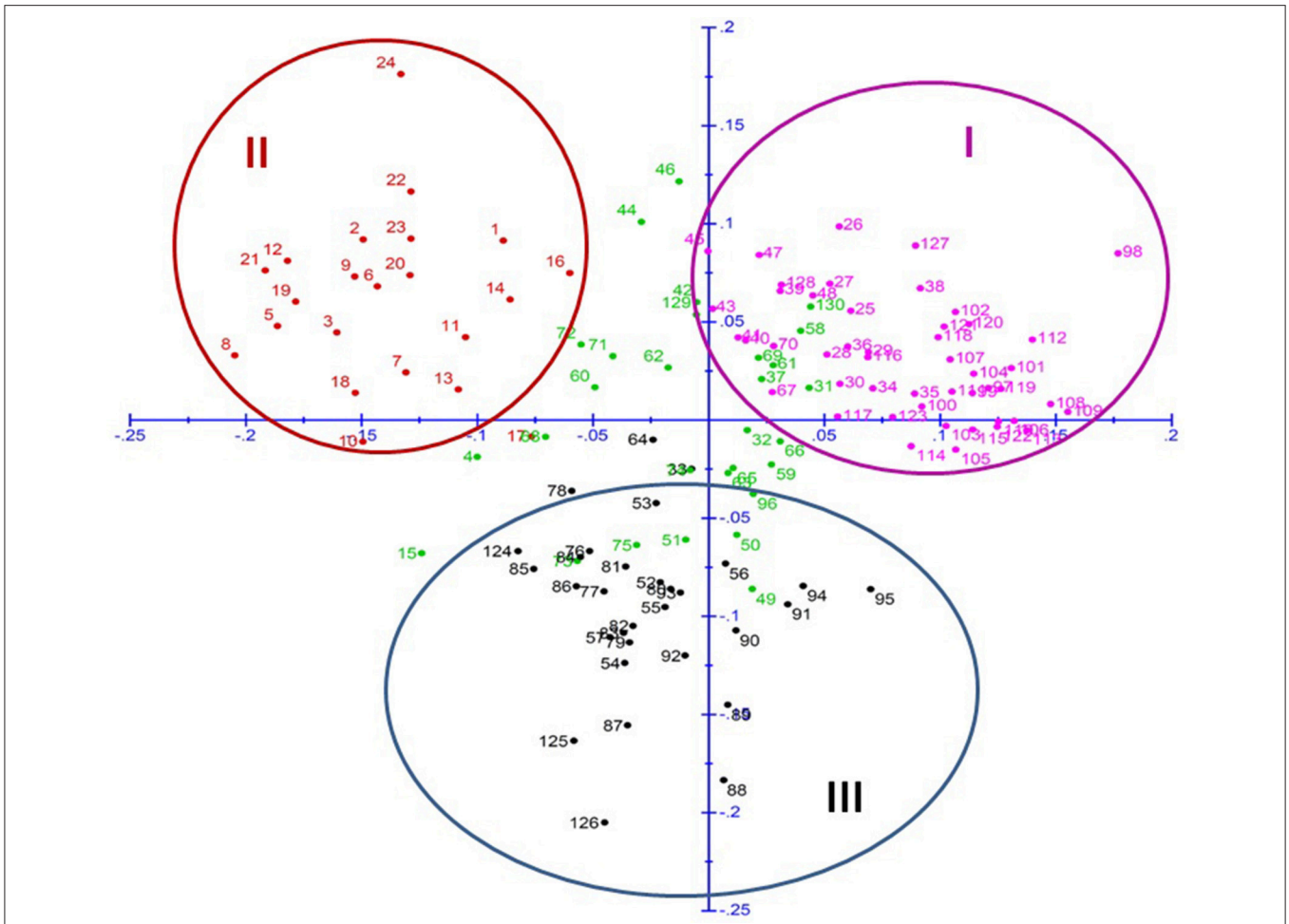

FIGURE 3 | Principal Coordinate Analysis (PCoA) of pearl millet association mapping panel (genotypes represented in different colors corresponding to the three sub-populations plus admixture observed in STRUCTURE).

For grain zinc content (**Table S7B**), among 10 markers identified as significantly associated with the trait, *Xpsmp* 2261 showed association in eight datasets (Del-14, Jod-14, Jod-15, Del-M, Jod-M, Y14-M, Y15-M, and GM). Next, *Xipes* 0096 was observed as significant for five times (DW-14, DW-M, Y14-M, Y15-M, and GM). Genic marker, *Xsinramp* 6 was associated with the trait in three datasets (DW-14, DW-15 and DW-M) while, seven SSR markers (*Xipes* 0180, *Xpsmp* 2086, *Xipes* 0224, *Xpsmp* 2213, *Xicmp* 3004, *Xicmp* 3016, and *Xicmp* 4006) showed significance only at one location.

For grain iron content alone, six markers were identified to be statistically significant at Bonferroni correction ($p < 0.0037$) by MLM analysis in all the combinations of environments studied (**Table S8**, **Figure 5**). Two SSR markers were from LG7 and one each marker belonging to LG3 and LG5 in the published pearl millet consensus map of Rajaram et al. (2013) were associated with iron content. Remaining two markers (*Xpsmp* 2209 and *Xsinramp* 6) were unmapped to any linkage group of the published consensus map of pearl millet.

For grain zinc content (**Table S8**, **Figure 5**), total 10 markers were observed to be associated in all the environments taken together. The maximum number of significant markers (4) were from unmapped markers (from Rajaram et al., 2013) followed by LG6, where *Xipes* 0224 and *Xpsmp* 2213 were observed to be significant. One each marker from LG3, LG4, LG5, and LG7 were observed to associate with grain zinc content in MLM.

TABLE 1 | AMOVA between sub-populations and genotypes.

Source of variation	Degrees of freedom	Sum of squares	Variance components	Percentage of variation
Among sub-populations	3	523.86	2.24 Va	11.77
Among individuals within sub-populations	126	4,095.14	15.68 Vb	82.20
Within individuals	130	149.50	1.15 Vc	6.03
Total	259	4,768.49	19.07	

Va, Vb, and Vc all are significant at $p < 0.01$.

TABLE 2A | Descriptive statistics and heritability of grain iron content at six individual environments and their six pooled environments (ppm).

	Del-14	Del-15	Jod-14	Jod-15	DW-14	DW-15	Y14-M	Y15-M	Del-M	Jod-M	DW-M	GM
Minimum	27.29	29.99	23.27	24.64	31.22	34.36	29.00	34.73	29.27	29.65	36.02	32.30
Maximum	125.01	117.27	121.63	111.40	139.12	123.78	114.60	112.10	116.40	109.62	122.34	111.90
Mean	57.52	60.44	55.73	57.81	60.71	61.49	57.99	59.92	58.98	56.77	61.10	58.95
SEm	1.54	1.43	1.80	1.43	1.99	1.51	1.38	1.13	1.34	1.38	1.48	1.13
H^2	0.85	0.87	0.92	0.91	0.94	0.88	0.67	0.68	0.80	0.72	0.65	0.79

Where, Del-14, Del-15, Jod-14, Jod-15, DW-14, DW-15, Y14-M, Y15-M, Del-M, Jod-M, DW-M, and GM are Delhi during 2014, Delhi during 2015, Jodhpur during 2014, Jodhpur during 2015, Dharwad during 2014, Dharwad during 2015, Year 2014 mean, Year 2015 mean, Delhi mean, Jodhpur mean, Dharwad mean, and grand mean, respectively. SEm and H^2 are standard error mean and heritability, respectively.

TABLE 2B | Descriptive statistics and heritability of grain zinc content at six individual environments and their six pooled environments (ppm).

	Del-14	Del-15	Jod-14	Jod-15	DW-14	DW-15	Y14-M	Y15-M	Del-M	Jod-M	DW-M	GM
Minimum	27.90	23.40	20.30	22.29	21.70	19.12	26.40	26.64	28.50	24.85	21.01	26.62
Maximum	86.20	78.38	86.00	84.36	73.27	74.22	77.20	74.19	81.08	81.71	70.65	73.68
Mean	46.61	49.06	42.75	45.16	36.40	35.74	41.92	43.32	47.83	43.96	36.07	42.62
SEm	1.04	1.11	1.14	1.05	0.95	1.16	0.78	0.83	1.00	1.01	0.99	0.76
H^2	0.85	0.86	0.87	0.84	0.87	0.89	0.68	0.60	0.83	0.80	0.86	0.79

Where, Del-14, Del-15, Jod-14, Jod-15, DW-14, DW-15, Y14-M, Y15-M, Del-M, Jod-M, DW-M, and GM are Delhi during 2014, Delhi during 2015, Jodhpur during 2014, Jodhpur during 2015, Dharwad during 2014, Dharwad during 2015, Year 2014 mean, Year 2015 mean, Delhi mean, Jodhpur mean, Dharwad mean, and grand mean, respectively. SEm and H^2 are standard error mean and heritability, respectively.

In total, 12 markers (4 genomic SSRs, 7 EST-SSRs, and 1 genic marker) accounting for16 MTAs were observed to be significantly associated with either grain iron or zinc content at significant threshold of Bonferroni correction ($p < 0.0037$) using MLM analysis of data from different environments and combination of their means (**Table S8**, **Figure 5**). The maximum contribution was from LG6 and LG7, which included two SSRs each. Both of the markers from LG6 were associated with zinc content only, whereas one SSR, *Xipes* 0096 of LG7 was highly associated with both iron and zinc content, while another SSR, *Xicmp* 3092 was associated with iron content only. One SSR each from LG3, and LG5 were associated with both the traits. One genic marker, *Xsinramp* 6 also showed significance of association with both iron and zinc content. Remaining four unmapped markers in the published consensus map (Rajaram et al., 2013) were either associated with grain iron or zinc content.

Marker-Trait Associations in Sub-Populations

Twelve significant markers resulted from MLM analysis were used to study their associations at sub-population level. For grain iron content, 13, 26, and 23 MTAs were recorded in sub-population A, B and C respectively (**Table S9A**). Similarly, 19, 24, and 18 MTAs were recorded for grain zinc content in sub-population A, B and C respectively (**Table S9B**). Maximum number of associations were observed in sub-population B for both traits, where most of promising lines like PPMI 1102, PPMFeZMP 199, PPMI 708 etc. were included.

Xpsmp 2261 at Del-15 and *Xsinramp* 6 at Del-M showed significant association with grain iron content in all the three sub-populations (comparing **Table S9A** with **Table 3**). Significance in at least two sub-populations were observed by *Xicmp* 3092 at DW-14; *Xipes* 0180 in Del-14 and GM; *Xsinramp* 6 in Del-14 and *Xpsmp* 2261 in Del-M, Jod-M, Y14-M, Y15-M, and GM datasets. Few associations *viz., Xipes* 0096 at Del-M, Jod-14, Y14-M, Y15-M, and GM; *Xipes* 0180 at Del-M and Y14; *Xpsmp* 2209 at DW-15; *Xpsmp* 2261 at Jod-14 and Jod-15 were restricted only to one sub-population, but resulted in significant associations using MLM analysis of the whole population. Few MTAs were not observed while analyzing the whole population but were recorded in two out of three sub-populations for some datasets.

Similarly for grain zinc content (comparing **Table S9B** with **Table 3**), *Xpsmp* 2261 in Y14-M, Y15-M, and GM; *Xsinramp* 6 in DW-M; *Xsinramp* 6 in DW-14 recorded associations in two sub-populations out of three, whereas *Xpsmp* 2261 in Del-14, Del-15, Jod-14, and Jod-M; *Xipes* 0096 in DW-14, Y14-M, Y15-M, and GM; *Xicmp* 3016 in Del-15; *Xipes* 0096 in DW-M; *Xipes* 0180 in Jod-15; *Xicmp* 4006 in DW-15; *Xicmp* 3004 and *Xipes* 0224 in Jod-14 datasets recorded significance only in one sub-population. MTAs, *Xpsmp* 2086 at DW-14; *Xsinramp* 6 in DW-15; *Xpsmp* 2213 in Jod-M did not show significance at sub-population level, though they are associated in the whole population. *Xpsmp* 2261 in DW-14 and DW-M; *Xipes* 0180 in Del-14 and *Xsinramp* 6 in Del-15 appeared in significant association with grain zinc content in two sub-populations, though they are not associated at whole population level.

Identification of Favorable Alleles

Alleles having positive effect were considered as favorable for both grain iron and zinc content. Details of these alleles along with top three genotypes carrying them are given in **Tables S10A,B**. The favorable allele is expressed in terms of

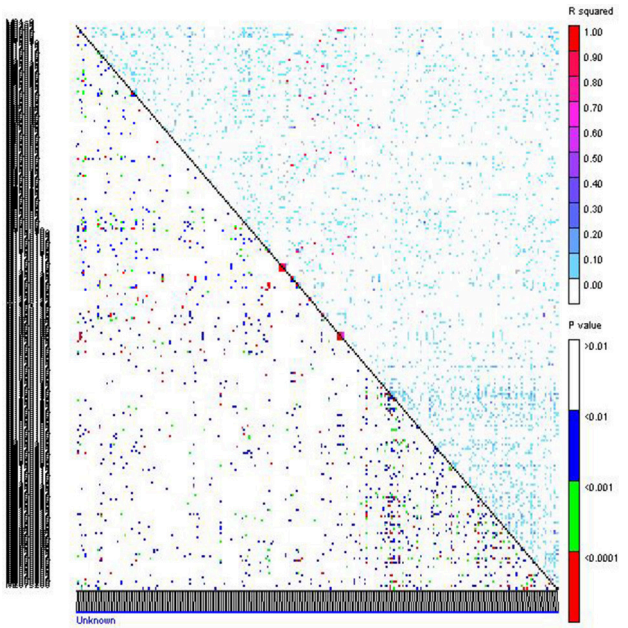

FIGURE 4 | Linkage disequilibrium pattern of association mapping panel genotyped with 267 markers. Upper triangle represents R^2 of markers and in lower triangle corresponding *p*-values are given.

amplicon size of the SSR marker. A total of six alleles were detected for grain iron content, and five out of 10 alleles detected for grain zinc content had phenotypic effect of more than four. Highest phenotypic effect of alleles for grain iron and zinc content was observed by *Xpsmp* 2261-180. Among the top three performing genotypes, PPMI 1102 had highest number (7) of favorable alleles (**Tables S11A,B**), followed by PPMI 1104 (5), PPMFeZMP 199 (4), and PPMI 708 (4). PPMI 1102, PPMFeZMP 199, PPMI 708 and PPMI 683 had alleles which expressed across environments, whereas PPMI 1104 had *Xicmp* 3092-220 allele specific for Fe content, and *Xicmp* 4006-280 allele specific for Zn content at Dharwad.

Preferred Crosses to Improve Target Traits

Considering the average phenotypic effect of an allele approximately equal to four or greater, nine better crosses (**Table 4**) were suggested to pyramid maximum number of alleles in a single genotype. To account for maximum pyramiding of alleles even having minor effect, one single cross (PPMI 1102 × PPMI 1104) can capture as many as 10 alleles favorable for both grain iron and zinc content.

Validation of Marker-Trait Associations Using Bioinformatics Tools

Out of three SSR markers selected for amplifying seven genotypes, *Xipes* 0096 was dropped for further analysis, because of its poor read quality. The amplicons of *Xpsmp* 2261 and *Xipes* 0810 for seven genotypes were compared with reference genome Tift23$D_2B_1^-$ P1-P5. Among two markers studied, *Xpsmp* 2261 amplicon landed in intergenic region on pseudomolecule 5 and was having >50% GC content, while the other marker, *Xipes* 0810 was observed to be overlapping with gene on pseudomolecule 3. This flanking gene was annotated to be aspartic proteinase (*Asp1*) gene. On comparison with the

TABLE 3 | Markers associated with grain iron and zinc content along with phenotypic variance explained at $p < 0.0037$ (Bonferroni corrected *p*-value) in MLM.

Data set	Marker	Fe		Zn	
		marker_p	Marker R^2	marker_p	Marker R^2
Del-14	*Xipes* 0180	0.0017	0.124	ns	ns
	Xpsmp 2261	ns	ns	0.0021	0.106
	Xsinramp 6	0.0036	0.101	ns	ns
Del-15	*Xicmp* 3016	ns	ns	0.0032	0.082
	Xpsmp 2261	6.00×10^{-4}	0.122	ns	ns
Del-M	*Xipes* 0096	0.0012	0.11	ns	ns
	Xipes 0180	0.0036	0.102	ns	ns
	Xpsmp 2261	0.0034	0.92	0.0031	0.099
	Xsinramp 6	0.0033	0.098	ns	ns
DW-14	*Xicmp* 3092	0.0036	0.091	ns	ns
	Xipes 0096	ns	ns	4.56×10^{-4}	0.127
	Xpsmp 2086	ns	ns	0.0015	0.108
	Xsinramp 6	ns	ns	0.002	0.107
DW-15	*Xicmp* 4006	ns	ns	0.0017	0.104
	Xpsmp 2209	0.0034	0.07	ns	ns
	Xsinramp 6	ns	ns	0.0012	0.119
DW-M	*Xipes* 0096	ns	ns	0.003	0.094
	Xsinramp 6	ns	ns	6.17×10^{-4}	0.13
Jod-14	*Xicmp* 3004	ns	ns	0.0023	0.103
	Xipes 0096	0.0024	0.098	ns	ns
	Xipes 0224	ns	ns	0.0017	0.105
	Xpsmp 2261	8.1×10^{-5}	0.157	3.61×10^{-4}	0.131
Jod-15	*Xipes* 0180	ns	ns	0.0032	0.095
	Xpsmp 2261	6.44×10^{-4}	0.121	6.14×10^{-5}	0.162
Jod_M	*Xicmp* 3004	ns	ns	0.0047	0.091
	Xpsmp 2213	ns	ns	2.09×10^{-4}	0.141
	Xpsmp 2261	2.22×10^{-5}	0.181	5.35×10^{-5}	0.165
Y14-M	*Xipes* 0096	7.42×10^{-4}	0.118	6.43×10^{-4}	0.121
	Xipes 0180	0.0024	0.118	ns	ns
	Xpsmp 2261	3.28×10^{-4}	0.133	2.11×10^{-4}	0.14
Y15-M	*Xipes* 0096	0.0017	0.104	0.0017	0.104
	Xpsmp 2261	2.17×10^{-4}	0.14	0.0014	0.107
GM	*Xipes* 0096	1.36×10^{-4}	0.148	9.41×10^{-4}	0.114
	Xipes 0180	9.89×10^{-4}	0.131	ns	ns
	Xpsmp 2261	1.7×10^{-4}	0.144	3.00×10^{-4}	0.134

Where Del-14, Del-15, Jod-14, Jod-15, DW-14, DW-15, Y14-M, Y15-M, Del-M, Jod-M, DW-M, and GM are Delhi during 2014, Delhi during 2015, Jodhpur during 2014, Jodhpur during 2015, Dharwad during 2014, Dharwad during 2015, Year 2014 mean, Year 2015 mean, Delhi mean, Jodhpur mean, Dharwad mean, and grand mean, respectively. ns, non-significant.

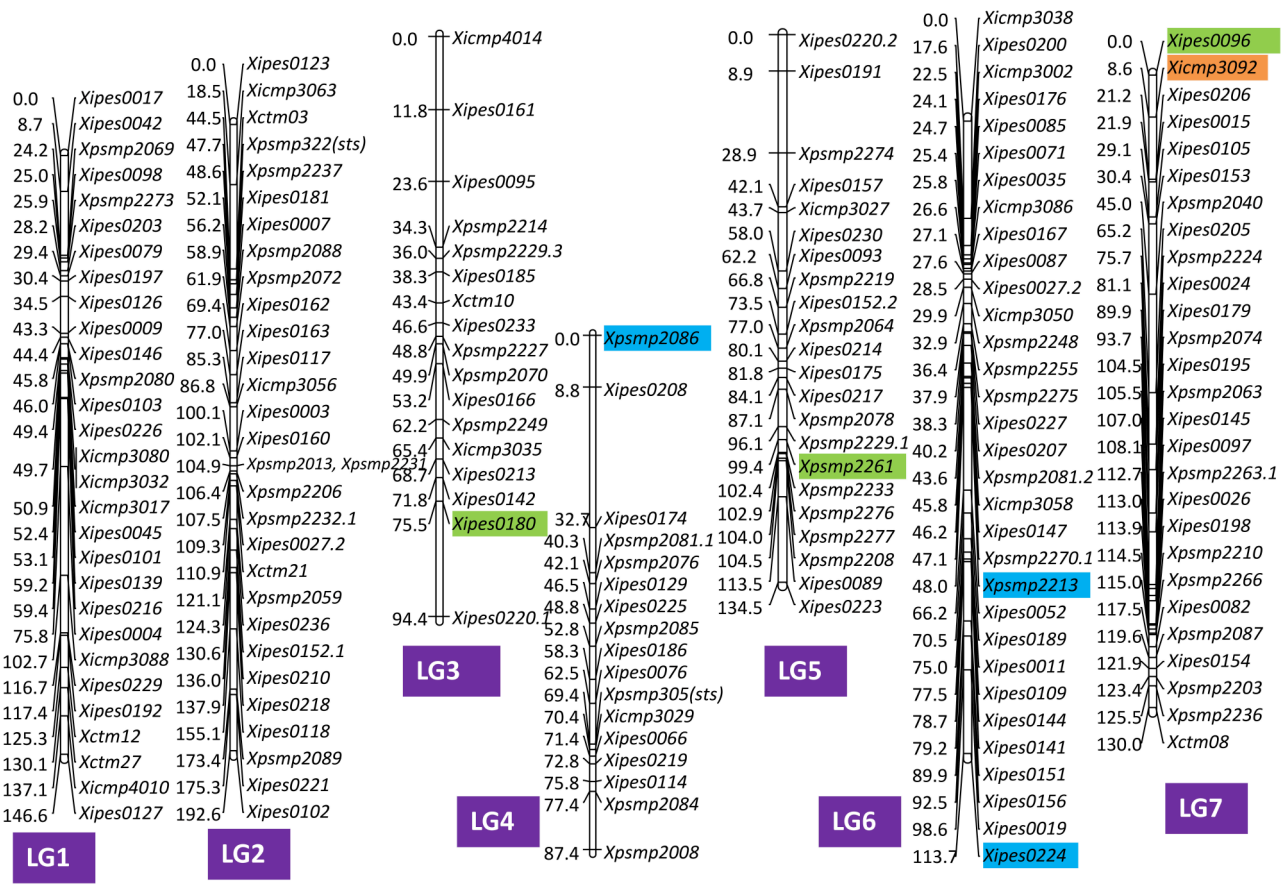

FIGURE 5 | Genomic positions of significantly associated SSR markers with grain iron and zinc content in the consensus map of Rajaram et al. (2013). Color code: Orange for iron; Blue for zinc; and Green for both iron and zinc.

reference sequence, two putative SNPs (**Figure 6**) and a 15 bp InDel were observed.

DISCUSSION

Lack of adequate quantities of vitamin A, iron and zinc in diet are main factors for micronutrient malnutrition. Biofortification of crop plants paved an economic and sustainable way to alleviate hidden hunger. Micronutrient deposition in grains are controlled by many genes and influenced by environmental conditions. In the process of biofortifying pearl millet with higher grain iron content, one open pollinated variety ICTP 8203Fe (Dhanshakti) and one hybrid ICMH 1201 (Shakti 1201) were released for cultivation by ICRISAT in collaboration with Harvest Plus, which had iron and zinc content of 71, 41 ppm and 75, 39 ppm, respectively (Rai et al., 2013) through conventional breeding. Further process is underway in this crop by many institutes to reach a target level of >77 ppm kept by Harvest Plus (http://www.harvestplus.org). This process will hasten up when QTLs related to grain iron and zinc content are identified and incorporated through marker-assisted breeding programs. Hence, the present study was attempted to decipher genomic regions for high grain iron and zinc content in pearl millet using association mapping.

Constitution of a panel with diverse lines is of prime requirement for association mapping (Flint-Garcia et al., 2005). The present association mapping panel comprised of 130 diverse lines of pearl millet with diverse pedigrees originating from different parts of India and Africa. The panel was phenotypically diverse with a wide range grain micronutrient content observed in different environments. Also, it was earlier reported that population consisting of <100 genotypes was treated as sub-optimal for association studies, in particularly with markers having low minor allele frequency (Zhu et al., 2008). However, some studies (Zhao et al., 2011) showed significant associations with a small panel of 57–97 genotypes. Hence, present size of the panel (130 genotypes) is considered substantial for association studies.

The average number of alleles per locus (2.65) detected in the present study is less than previous studies on pearl millet (Nepolean et al., 2012; Sehgal et al., 2015). According to Pasam et al. (2014), number of accessions taken, number and type of markers used for characterization play a role in detection of total number of alleles per locus. Same panel can result in varied detection of alleles depending upon the detection

TABLE 4 | Crosses proposed to accumulate favorable alleles for enhancing both grain iron and zinc content.

Crosses	Alleles predicted		
	Fe	Zn	Total
PPMI 1102 × PPMI 1104	6	4	7
PPMI 1102 × PPMI 1116	5	5	6
PPMI 1102 × PPMI 1105	5	5	6
PPMI 1102 × PPMI 1231	6	3	6
PPMI 1102 × PPMDMGPM 186	6	3	6
PPMI 1102 × PPMI 265	6	3	6
PPMI 1102 × PPMI 1285	6	3	6
PPMFeZMP 199 × PPMI 1104	5	4	6
PPMI 1104 × PPMI 683	4	3	6

techniques (Gupta et al., 2010), for example, PAGE or capillary electrophoresis with DNA fragment analyzer are capable enough to resolve even two base pairs differences in the amplicons whereas 3.5% agarose gels used here do not have the same resolution. Hence, even though diverse breeding lines from different pedigrees were included in the present investigation and previously studied SSR markers were utilized, the difference in allele detection methodology used in this study from earlier studies led to this distinction. However, the number of alleles per loci detected here is more than detected by Tara et al. (2009) (2.0) in pearl millet, Gupta et al. (2012) (2.2), Pandey et al. (2013) (2.1) in foxtail millet. The present panel was moderately polymorphic according to Vaiman et al. (1994) since the average PIC value falls in the range of 0.25–0.5 values. The value obtained for average gene diversity (0.35) also indicates moderate diversity in the present lines chosen for study. This value is less than that of reported in earlier studies (Nepolean et al., 2012; Tara et al., 2013; Sehgal et al., 2015). Though the highest value of gene diversity (0.73) obtained here is similar to that reported by Mariac et al. (2006) (0.75) who characterized wild pearl millet genotypes for diversity studies, the average value in present study is lesser than that recorded by them (0.49 in cultivated and 0.67 in wild lines). This may be because of the differences in the genotypes and the allele detection methods used.

Before performing association analysis, genetic differentiation of the population was observed by STRUCTURE and compared with other clustering approaches. STRUCTURE, clustering analysis and PCoA showed the division of population into three sub-populations. All the three models exhibited similarity in the grouping pattern which is akin to earlier studies (Sehgal et al., 2015). But, there was no 100% match between clustering and STRUCTURE grouping pattern. Resemblance between two was more than 60%. In STRUCTURE grouping was based on highest percentage of membership of an individual to a sub-population and in cluster analysis each genotype is assigned to a fixed branch position (Gupta et al., 2014). Also, admixture was separated from groups and placed in separate group depending on their *Q*-value, but, in clustering those genotypes were also involved in sub groups leading to reduced similarity in the grouping pattern obtained by two different analysis. Totally, there was no structural pattern observed based on high and low grain iron and zinc content either in STRUCTURE or in clustering. Grouping

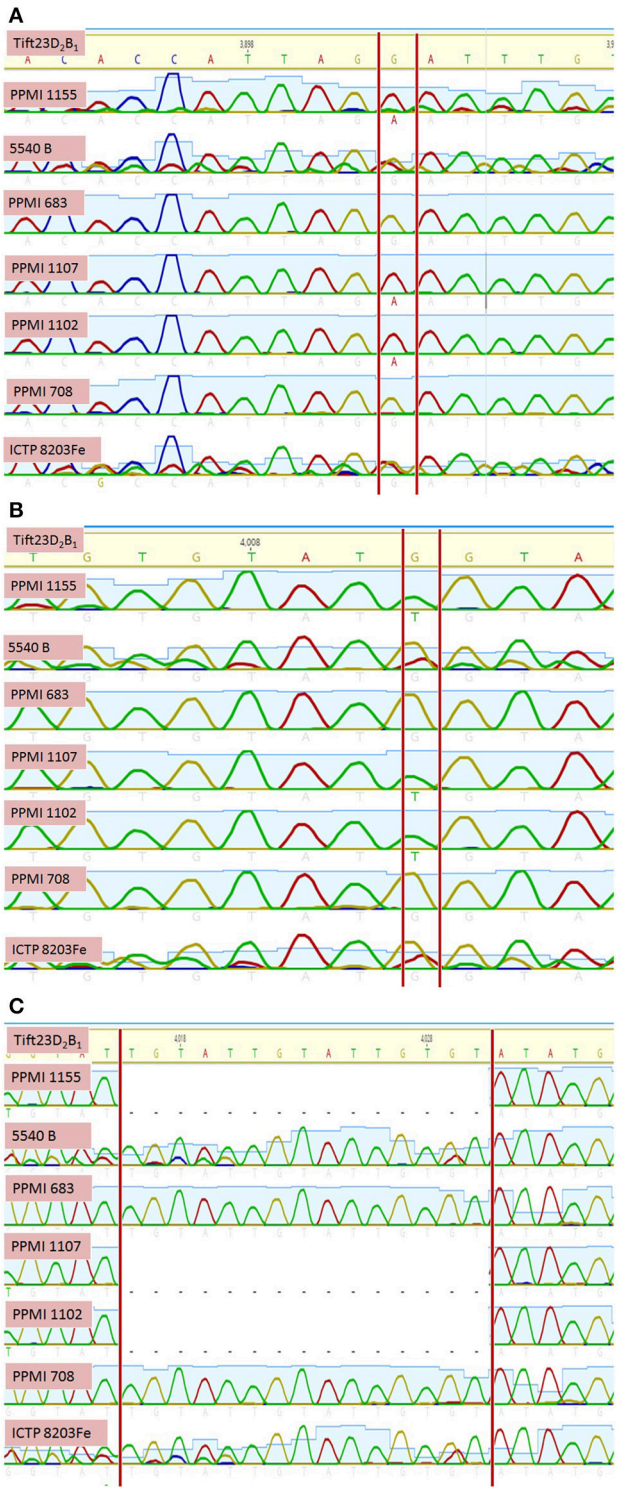

FIGURE 6 | Sequence homology of high and low grain iron and zinc genotypes with pearl millet reference sequence identifies putative (A) SNP1, **(B)** SNP2, and **(C)** InDel.

pattern, was in accordance with their pedigree. These results were similar to those obtained earlier (Sehgal et al., 2015).

Genetic differentiation between sub populations was tested here using AMOVA, which presents general skeleton for the study of population structure (Michalakis and Excoffier, 1996). In this study, all pair wise F_{ST}s were significantly different from one another, thus all sub-populations may be considered as significantly different from each other (Nachimuthu et al., 2015).Variation within sub-population was more compared to between sub-populations. Only 11% out of total variation was between sub-populations indicating lesser genetic structure of the population having free gene flow between sub-groups. This may be because of constant exchange of genetic material among breeding programs (Würschum et al., 2013). According to Wright (1978), when F_{ST} is between 0.05 and 0.15 values, it implies population differentiation is moderate. However, he defined it for biallelic markers or allozyme markers. Multi-allelic markers have low F_{ST} values compared to biallelic markers, because of inherent properties in calculation (Jakobsson et al., 2013). Since, the markers used here were multi-allelic, the value obtained (0.11) may indicate that genetic differentiation obtained here was moderate to high. Many diversity studies in pearl millet (Lewis, 2010; Nepolean et al., 2012), the maximum variation was contributed by within sub-population indicating that genotypes considered in this study were different from each other. In cross-pollinated species gene flow is likely to be greater within population than between population (Hamrick et al., 1990). We also observed some amount of within individual variance, which is indicative of presence of heterozygosity in the population. One reason is peri-centric regions have high degree of heterozygosity because of reduced recombination (McMullen et al., 2009 in Maize). In cross-pollinated crops like pearl millet, some amount of heterozygosity is expected. Nachimuthu et al. (2015) also observed some amount of individual variance in rice. The same reason i.e. cross-pollinated nature of pearl millet holds good for the low value of significant linkage disequilibrium (LD) obtained here. High values cannot be sustained because of frequent recombination as observed in maize (Zhu et al., 2008). LD is influenced by different factors like mode of reproduction, selection, rate of recombination, rate of mutation, genetic drift, and population structure. The low value of LD may ensure high resolution mapping, but more number of markers are needed to take this advantage (Gupta et al., 2005). Similarly, Zhang et al. (2013) observed only 2.95% locus pairs to be in significant LD in cotton crop. Broad-based population was reported to have low LD compared to narrow-based population. Hence the present panel may be broad-based, which supports the usefulness of panel for association mapping studies. A low LD demands more number of markers compared to high LD of same size population (Gupta et al., 2005), but there are some instances where less number of markers were quite enough to get MTAs (Li C. Q. et al., 2016). To capture the resolution in low LD crops, thousands of polymorphic markers are needed but it is possible only when those many markers are available, or when genome sequence is already known and all types of markers are easily predicted (Gupta et al., 2014).

Most widely used statistical models for marker-trait association studies are GLM (Pritchard et al., 2000) and MLM (Yu et al., 2006; Price et al., 2010). While GLM accounts for the population structure, MLM considers both population structure and familial relatedness. In GLM, there are chances of spurious associations (Type I error) because it only considers population structure and not kinship (Zhao et al., 2007). In MLM, sometimes, over compensation with both Q and K may lead to false negatives (Type II error; Zhao et al., 2007, 2011). This leads to detection of some MTAs only with a particular model (Liu et al., 2016). Hence, both models were tested for grain micronutrient association, but only the MTAs revealed by MLM are presented here, as MLM results were found to be more reliable.

The associations of the trait with markers were tested at three locations in the years 2014 and 2015. The data of individual environments (Del-14, Del-15, DW-14, DW-15, Jod-14, Jod-15) and their means pooled over environments in different combinations (Del-M, DW-M, Jod-M, Y14-M, Y15-M, and GM) were used to find significant associations of molecular markers with the individual trait. In this study, Bonferroni corrected *p*-value ($p = 0.0037$) was considered as a threshold for significance for declaring MTAs (Liu et al., 2016). Except for two markers (*Xipes* 0096 and *Xpsmp* 2261), there was inconsistency in marker-trait association by GLM and MLM which is expected and it may be because of the above mentioned Type I and Type II errors in one or the other models. These two loci which were detected by both approaches can be considered as the best MTAs. MLM analysis detected one SSR marker, *Xipes* 0180 to be associated consistently with grain iron content explaining more than 10% variation. It was also observed to be associated with grain zinc content at Jodhpur during 2015. A genic marker, *Xsinramp* 6, was also observed to be associated with both grain iron and zinc content. It recorded more than 9.5% R^2-value for grain iron content at Delhi location, while for grain zinc content R^2-value was more than 10.0% at Dharwad location.

A total of 16 MTAs for both grain iron and zinc content were identified in this study by using MLM analysis (**Figure 5**), out of which six MTAs were for grain iron and 10 for grain zinc content. *Xipes* 0180, belonging to LG3 was detected consistently for grain iron content. This marker was associated at Jod-15 (Jodhpur, 2015) for grain zinc content which is in accordance with the previous findings of bi-parental mapping, where LG2, LG3, LG5, and LG7 were reported to house QTLs for both iron and zinc content (Kumar et al., 2010, 2016). In present study, *Xpsmp* 2261 and *Xipes* 0096 belonging to LG5 and LG7, respectively, identified strong and consistent associations (not only with grain iron content, but also with grain zinc content in accordance with Kumar et al., 2010, 2016). However, in this study we did not find any MTA on LG2. Linkag group of five more markers (*Xicmp* 3004, *Xicmp* 3016, *Xicmp* 4006, *Xpsmp* 2209, and *Xsinramp* 6) associated with traits could not be ascertained. Hence, there is possibility that any of those five markers may map to LG2. Consistently identified markers namely, *Xpsmp* 2261 (LG5), *Xipes* 0096 (LG7), *Xipes* 0180 (LG3) and *Xsinramp* 6 (unmapped) were identified for both high grain iron and zinc content in the current study (**Figure 5**). Co-localization of high grain iron and zinc

content alleles/QTLs in pearl millet has also been reported earlier by Kumar et al. (2016). This may be because, starting from uptake to final deposition into grain of iron and zinc may share some common pathways (Grotz and Guerinot, 2006).

This study identified 16 MTAs for grain iron and zinc content. Some were consistent across locations and years, while some were specific to certain locations or years. This may be because, these traits are governed by many genes and exhibit considerable genotype-by-environment (G × E) interactions. Some associations were specific to certain sub-populations and some though not observed at whole population level, were significant in more than one sub-population. Few associations were not detected at sub-population level but they were recorded when the whole population was analyzed. This variation is because of the structure present in the population. Even though few differences existed at sub-population levels, *Xpsmp* 2261, *Xipes* 0096, *Xipes* 0180, and *Xsinramp* 6 were consistently associated in all sub-populations. Different populations of varying sizes and structure also affect the detection of associations. Earlier workers (Zhao et al., 2007; Liu et al., 2016) reported different associations with different set of populations. Hence, true associations obtained here must be validated by testing in another panel having different genotypes.

Similarly, several genomic regions were detected for grain iron and zinc content in other crops (Upadhyaya et al., 2016 in chickpea, Nawaz et al., 2015 in rice) through association mapping. Likewise, many QTLs have been identified for other agronomically important traits in pearl millet through association mapping (Saïdou et al., 2014; Sehgal et al., 2015). Tadesse et al. (2015) also observed consistent and specifically adapted QTLs in wheat for grain quality traits. Hence, different genomic regions in this study can be introgressed for trait improvement in pearl millet based on the targeted environment depending upon common and location specific MTAs. *Xipes* 0180 amplicon sequence was found matching with a segment of pearl millet reference genome, and was annotated as aspartic proteinase (*Asp1*) gene. While *Asp1* may not have a direct role in grain iron and zinc metabolism, it might be indirectly involved through other gene networks and pathways in different metabolic pathways. In addition, the two putative SNPs and 15bp InDel (**Figures 6A–C**) identified in the present study do not seem to be associated with a particular trait. These may call for further investigations.

Mining down to the effect of alleles of significant MTAs will make us understand favorable alleles and the genotypes carrying them. In this regard, we calculated effect of an allele using population mean instead of null alleles (Cai et al., 2014; Liu et al., 2016) while, on the other hand, computation of effect of an allele with a null allele is still followed by many workers (Breseghello and Sorrells, 2006; Dang et al., 2016). We are of the opinion that mean is a representative value of the whole data, while missing alleles which are also included as null alleles are error prone due to differential handling of experiment will lead to different missing data which in turn changes the resultant effect of an allele. In many instances, it was noticed that average effect of an allele will be higher when the allele is rare (Cai et al., 2014). Stably expressing alleles linked to markers, *Xpsmp* 2261, *Xipes* 0180 *Xipes* 0096, and *Xsinramp* 6 recorded greater phenotypic effect which offer advantage for Fe and Zn enhancement and higher trait expression in the recipient lines. These linked alleles may be promising targets for marker-assisted selection (MAS). On the basis of occurrence of favorable alleles for high grain iron and zinc content in the panel, we have suggested a set of lines (**Table 4**) which can be used in the crossing programs to accumulate favorable alleles together in the segregating generations, resulting in higher Fe and Zn content. This is in line with other studies (Dang et al., 2016).

CONCLUSION

The results of the current association mapping demonstrates usefulness of association mapping using SSRs and genic markers, and the current association mapping panel for identification of superior alleles for grain nutritional traits like iron and zinc content. The favorable alleles and the associated markers identified in the present study need to be validated in more diverse genetic backgrounds. Such validated markers might be useful in marker-assisted back-crossing (MABC), marker-assisted recurrent selection (MARS), and forward breeding programs. The promising lines with favorable alleles identified in this study may be used for generating new cultivars which accumulate all or most of the favorable alleles for high grain iron and zinc content.

AUTHOR CONTRIBUTIONS

CS and RS planned and designed the research; NA, SMS, and SPS performed the field experiment; MM helped in Fe and Zn estimation; NA, CB, and TS conducted genotyping of the association mapping panel; NA and TN conducted STRUCTURE and TASSEL analysis; RS organized bioinformatics analysis; NA, CS, and RS prepared the manuscript. CS and RS edited the manuscript for publication. All authors have read and approved the final manuscript.

ACKNOWLEDGMENTS

This research work has been funded by CRP on biofortification (ICAR funding), Division of Genetics, IARI, New Delhi. All the facilities availed from Division of Genetics to carry out genotyping and phenotyping, and Division of Soil Science and Agricultural Chemistry for Fe and Zn estimation at IARI, New Delhi are gratefully acknowledged. NA also thank PG school for accepting her as a Ph.D. scholar, and ANGRAU, Andhra Pradesh for according deputation to pursue doctoral degree program at IARI, New Delhi. The authors thank Dr. Rajeev K. Varshney for granting access to the pearl millet reference genome sequence, and Prasad Bajaj for carrying out the bioinformatics work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2017.00412/full#supplementary-material

Table S1 | List of pearl millet genotypes, their pedigree and inferred sub-population.

Table S2 | List of 267 (250 SSRs and 17 genic) primer pairs used to characterize the association mapping panel.

Table S3 | Summary statistics of 114 SSR markers used in the study.

Table S4A | One way ANOVA for grain iron content (ppm) for six environments.

Table S4B | One way ANOVA for grain zinc content (ppm) for six environments.

Table S4C | Pooled ANOVA of years 2014 and 2015 for grain iron and zinc content for different locations.

Table S4D | Pooled ANOVA across three locations for grain iron and zinc content for the years 2014 and 2015.

Table S4E | Pooled ANOVA across 6 environments (2 years and three locations) for grain iron and zinc content.

Table S5A | Mean values of grain iron content across six environments and their six pooled environments (ppm).

Table S5B | Mean values of grain zinc content across six environments and their six pooled environments (ppm).

Table S6 | Correlation between grain iron and zinc content and population structure using multiple regression analysis.

Table S7A | Marker wise associations observed across all datasets for grain iron content in MLM at $p < 0.0037$.

Table S7B | Marker wise associations observed across all datasets for grain zinc content in MLM at $p < 0.0037$.

Table S8 | Linkage group wise number of significant markers associated with grain iron and zinc content.

Table S9A | Markers associated significantly ($p = 0.05$) with grain iron content at sub-population level.

Table S9B | Markers associated significantly ($p = 0.05$) with grain zinc content at sub-population level.

Table S10A | Phenotypic effect of favorable alleles for grain iron content and top three genotypes carrying them.

Table S10B | Phenotypic effect of favorable alleles for grain zinc content and top three genotypes carrying them.

Table S11A | Top most genotypes carrying favorable alleles for grain iron content.

Table S11B | Top most genotypes carrying favorable alleles for grain zinc content.

Figure S1 | Quantile-Quantile (QQ) plots for grain iron content in MLM showing distribution of marker-trait association.

Figure S2 | Quantile-Quantile (QQ) plots for grain zinc content in MLM showing distribution of marker-trait association.

REFERENCES

Barthakur, S. (2010). *Harnessing Hidden Hunger. Science Reporter.* Available online at: http://nopr.niscair.res.in

Bland, J. M., and Altman, D. G. (1995). Multiple significance tests: the Bonferroni method. *BMJ* 310:170. doi: 10.1136/bmj.310.6973.170

Bouis, H. E., and Welch, R. M. (2010). A sustainable agricultural strategy for reducing micronutrient malnutrition in the global south. *Crop Sci.* 50, 20–32. doi: 10.2135/cropsci2009.09.0531

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308

Breseghello, F., and Sorrells, M. E. (2006). Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics* 172, 1165–1177. doi: 10.1534/genetics.105.044586

Cai, C., Ye, W., Zhang, T., and Guo, W. (2014). Association analysis of fiber quality traits and exploration of elite alleles in Upland cotton cultivars/accessions (*Gossypium hirsutum* L.). *J. Integr. Plant Biol.* 56, 51–62. doi: 10.1111/jipb.12124

Crespo-Herrera, L. A., Velu, G., and Singh, R. P. (2016). Quantitative trait loci mapping reveals pleiotropic effect for grain iron and zinc concentrations in wheat. *Ann. Appl. Biol.* 169, 27–35. doi: 10.1111/aab.12276

Dang, X., Liu, E., Liang, Y., Liu, Q., Breria, C. M., and Hong, D. (2016). QTL detection and elite alleles mining for stigma traits in *Oryza sativa* by association mapping. *Front. Plant. Sci.* 7:1188. doi: 10.3389/fpls.2016.01188

Deshpande, J. D., Joshi, M. M., and Giri, P. A. (2013). Zinc: the trace element of major importance in human nutrition and health. *Int. J. Med. Sci. Public Health* 2, 1–6. doi: 10.5455/ijmsph.2013.2.1-6

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x

Excoffier, L., and Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Finkelstein, J. L., Mehta, S., Udipi, S. A., Ghugre, P. S., Luna, S. V., Wenger, M. J., et al. (2015). A randomized trial of iron-biofortified pearl millet in school children in India. *J. Nutr.* 145, 1576–1581. doi: 10.3945/jn.114.208009

Flint-Garcia, S. A., Thornsberry, J. M., and Buckler, E. S. IV. (2003). Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* 54, 357–374. doi: 10.1146/annurev.arplant.54.031902.134907

Flint-Garcia, S. A., Thuillet, A.-C., Yu, J., Pressoir, G., Romero, S. M., Mitchell, S. E., et al. (2005). Maize association population: a high-resolution platform for quantitative trait locus dissection. *Plant J.* 44, 1054–1064. doi: 10.1111/j.1365-313X.2005.02591.x

Gómez, G., Álvarez, M. F., and Mosquera, T. (2011). Association mapping, a method to detect quantitative trait loci: statistical bases. *Agron. Colomb.* 29, 367–376.

Grotz, N., and Guerinot, M. L. (2006). Molecular aspects of Cu, Fe and Zn homeostasis in plants. *Biochim. Biophys. Acta* 1763, 595–608. doi: 10.1016/j.bbamcr.2006.05.014

Gupta, P. K., Rustgi, S., and Kulwal, P. L. (2005). Linkage disequilibrium and association studies in higher plants: present status and future prospects. *Plant Mol. Biol.* 57, 461–485. doi: 10.1007/s11103-005-0257-z

Gupta, S., Kumari, K., Muthamilarasan, M., Parida, S. K., and Prasad, M. (2014). Population structure and association mapping of yield contributing agronomic traits in foxtail millet. *Plant Cell Rep.* 33, 881–893. doi: 10.1007/s00299-014-1564-0

Gupta, S., Kumari, K., Sahu, P. P., Vidapu, S., and Prasad, M. (2012). Sequence-based novel genomic microsatellite markers for robust genotyping purposes in foxtail millet [*Setaria italica* (L.) P. Beauv.]. *Plant Cell Rep.* 31, 323–337. doi: 10.1007/s00299-011-1168-x

Gupta, V., Dorsey, G., Hubbard, A. E., Rosenthal, P. J., and Greenhouse, B. (2010). Gel versus capillary electrophoresis genotyping for categorizing treatment outcomes in two anti-malarial trials in Uganda. *Malaria J.* 9:19. doi: 10.1186/1475-2875-9-19

Hamrick, J. L., Godt, M. W., Brown, A. H. D., Clegg, M. T., Kahler, A. L., and Weir, B. S. (1990). "Plant population genetics, breeding, and genetic resources,"

in *Allozyme Diversity in Plant Species*, eds A. H. D. Brown, M. T. Clegg, A. L. Kahler, and B. S. Weir (Sunderland, MA: Sinauer Associates Inc.), 43–63.

Jakobsson, M., Edge, M. D., and Rosenberg, N. A. (2013). The relationship between FST and the frequency of the most frequent allele. *Genetics* 193, 515–528. doi: 10.1534/genetics.112.144758

Jin, T., Chen, J., Zhu, L., Zhao, Y., Guo, J., and Huang, Y. (2015). Comparative mapping combined with homology-based cloning of the rice genome reveals candidate genes for grain zinc and iron concentration in maize. *BMC Genet.* 16:17. doi: 10.1186/s12863-015-0176-1

Krohling, C. A., Eutrópio, F. J., Bertolazi, A. A., Dobbss, L. B., Campostrini, E., Dias, T., et al. (2016). Ecophysiology of iron homeostasis in plants. *Soil Sci. Plant Nutr.* 62, 39–47. doi: 10.1080/00380768.2015.1123116

Kumar, S., Hash, C. T., Thirunavukkarasu, N., Singh, G., Rajaram, R., Rathore, A., et al. (2016). Mapping quantitative trait loci controlling high Fe and Zn density in self and open pollinated grains of pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Front. Plant. Sci.* 7:1636. doi: 10.3389/fpls.2016.01636

Kumar, S., Nepolean, T., Rai, K. N., Rajaram, V., Velu, G., Sahrawat, K. L., et al. (2010). "Mapping pearl millet [*Pennisetum glaucum* (L.) R. Br.]. QTLs for Fe and Zn grain density," in *International Symposium on Genomics of Plant Genetic Resources, II* (Bologna), 2427.

Lewis, L. R. (2010). Biogeography and genetic diversity of pearl millet (*Pennisetum glaucum*) from Sahelian Africa. *Prof. Geogr.* 62, 377–394. doi: 10.1080/00330124.2010.483640

Li, C. Q., Ai, N. J., Zhu, Y. J., Wang, Y. Q., Chen, X. D., Li, F., et al. (2016). Association mapping and favourable allele exploration for plant architecture traits in upland cotton (*Gossypium hirsutum* L.) accessions. *J. Agric. Sci.* 154, 567–583. doi: 10.1017/S0021859615000428

Li, X., Zhou, Z., Ding, J., Wu, Y., Zhou, B., Wang, R., et al. (2016). Combined linkage and association mapping reveals QTL and candidate genes for plant and ear height in maize. *Front. Plant Sci.* 7:833. doi: 10.3389/fpls.2016.00833

Liu, N., Xue, Y., Guo, Z., Li, W., and Tang, J. (2016). Genome-wide association study identifies candidate genes for starch content regulation in maize kernels. *Front. Plant. Sci.* 7:1046. doi: 10.3389/fpls.2016.01046

Mariac, C., Luong, V., Kapran, I., Mamadou, A., Sagnard, F., Deu, M., et al. (2006). Diversity of wild and cultivated pearl millet accessions (*Pennisetum glaucum* [L.] R. Br.) in Niger assessed by microsatellite markers. *Theor. Appl. Genet.* 114, 49–58. doi: 10.1007/s00122-006-0409-9

McMullen, M. D., Kresovich, S., Villeda, H. S., Bradbury, P., Li, H., Sun, Q., et al. (2009). Genetic properties of the maize nested association mapping population. *Science* 325, 737–740. doi: 10.1126/science.1174320

Michalakis, Y., and Excoffier, L. (1996). A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* 142, 1061–1064.

Mondal, T. K., Ganie, S. A., Rana, M. R., and Sharma, T. R. (2014). Genome-wide analysis of zinc transporter genes of maize (*Zea mays*). *Plant Mol. Biol. Rep.* 32, 605–616. doi: 10.1007/s11105-013-0664-2

Murgia, I., Arosio, P., Tarantino, D., and Soave, C. (2012). Biofortification for combating 'hidden hunger' for iron. *Trends Plant Sci.* 17, 47–55. doi: 10.1016/j.tplants.2011.10.003

Murray, M. G., and Thompson, W. F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8, 4321–4326. doi: 10.1093/nar/8.19.4321

Nachimuthu, V. V., Muthurajan, R., Duraialaguraja, S., Sivakami, R., Pandian, B. A., Ponniah, G., et al. (2015). Analysis of population structure and genetic diversity in rice germplasm using SSR markers: an initiative towards association mapping of agronomic traits in *Oryza sativa*. *Rice* 8:1. doi: 10.1186/s12284-015-0062-5

Nambiar, V. S., Dhaduk, J. J., Sareen, N., Shahu, T., and Desai, R. (2011). Potential functional implications of pearl millet (*Pennisetum glaucum*) in health and disease. *JAPS* 1, 62–67

Nawaz, Z., Kakar, K. U., Li, X. B., Li, S., Zhang, B., Shou, H. X., et al. (2015). Genome-wide association mapping of quantitative trait loci (QTLs) for contents of eight elements in brown rice (*Oryza sativa* L.). *Agric. Food Chem.* 63, 8008–8016. doi: 10.1021/acs.jafc.5b01191

Nei, M. (1972). Genetic distance between populations. *Am. Nat.* 106, 283–292. doi: 10.1086/282771

Nepolean, T., Gupta, S. K., Dwivedi, S. L., Bhattacharjee, R., Rai, K. N., and Hash, C. T. (2012). Genetic diversity in maintainer and restorer lines of pearl millet. *Crop Sci.* 52, 2555–2563. doi: 10.2135/cropsci2011.11.0597

Ohashi, J., and Tokunaga, K. (2003). Power of genome-wide linkage disequilibrium testing by using microsatellite markers. *J. Hum. Genet.* 48, 487–491. doi: 10.1007/s10038-003-0058-7

Pandey, G., Misra, G., Kumari, K., Gupta, S., Parida, S. K., Chattopadhyay, D., et al. (2013). Genome-wide development and use of microsatellite markers for large-scale genotyping applications in foxtail millet [*Setaria italica* (L.)]. *DNA Res.* 20, 197–207. doi: 10.1093/dnares/dst002

Pasam, R. K., Sharma, R., Walther, A., Özkan, H., Graner, A., and Kilian, B. (2014). Genetic diversity and population structure in a legacy collection of spring barley landraces adapted to a wide range of climates. *PLoS ONE* 9:116164. doi: 10.1371/journal.pone.0116164

Patterson, H. D., and Williams, E. R. (1976). A new class of resolvable incomplete block designs. *Biometrika* 63, 83–92. doi: 10.1093/biomet/63.1.83

Perrier, X., Flori, A., and Bonnot, F. (2003). *Methods for Data Analysis. Genetic Diversity of Cultivated Tropical Plants.* Montpellier: Science Publishers, Inc. and CIRAD.

Pfeiffer, W. H., and McClafferty, B. (2007). HarvestPlus: breeding crops for better nutrition. *Crop Sci.* 47, 88–105. doi: 10.2135/cropsci2007.09.0020IPBS

Press Information Bureau (2013). *Anaemia, Iodine Deficiency and Micro Nutrient Disorders.* Press Information Bureau Government of India 2013. Ministry of Health and Family Welfare. Available online at: http://pib.nic.in

Price, A. L., Zaitlen, N. A., Reich, D., and Patterson, N. (2010). New approaches to population stratification in genome-wide association studies. *Nat. Rev. Genet.* 11, 459–463. doi: 10.1038/nrg2813

Pritchard, J. K., Stephens, M., Rosenberg, N. A., and Donnelly, P. (2000). Association mapping in structured populations. *Am. J. Hum. Genet.* 67, 170–181. doi: 10.1086/302959

Rai, K. N., Yadav, O. P., Rajpurohit, B. S., Patil, H. T., Govindaraj, M., Khairwal, I. S., et al. (2013). Breeding pearl millet cultivars for high iron density with zinc density as an associated trait. *J. SAT Agric. Res.* 11, 1–7. Available online at: http://oar.icrisat.org/id/eprint/7291

Rajaram, V., Nepolean, T., Senthilvel, S., Varshney, R. K., Vadez, V., Srivastava, R. K., et al. (2013). Pearl millet [*Pennisetum glaucum* (L.) R. Br.] consensus linkage map constructed using four *RIL mapping populations and newly developed EST-SSRs. BMC Genomics* 14:159. doi: 10.1186/1471-2164-14-159

Rao, P. P., Birthal, P. S., Reddy, B. V., Rai, K. N., and Ramesh, S. (2006). Diagnostics of sorghum and pearl millet grains-based nutrition in India. *Int. Sorghum Millets Newslett.* 47, 93–96. Available online at: http://oar.icrisat.org/id/eprint/1119

Saïdou, A. A., Clotault, J., Couderc, M., Mariac, C., Devos, K. M., Thuillet, A. C., et al. (2014). Association mapping, patterns of linkage disequilibrium and selection in the vicinity of the PHYTOCHROME C gene in pearl millet. *Theor. Appl. Genet.* 127, 19–32. doi: 10.1007/s00122-013-2197-3

Saltzman, A., Birol, E., Bouis, H. E., Boy, E., De Moura, F. F., Islam, Y., et al. (2013). Biofortification: progress toward a more nourishing future. *Glob. Food Sec.* 2, 9–17. doi: 10.1016/j.gfs.2012.12.003

Sehgal, D., Skot, L., Singh, R., Srivastava, R. K., Das, S. P., Taunk, J., et al. (2015). Exploring potential of pearl millet germplasm association panel for association mapping of drought tolerance traits. *PLoS ONE* 10:e0122165. doi: 10.1371/journal.pone.0122165

Shivhare, R., and Lata, C. (2016). Selection of suitable reference genes for assessing gene expression in pearl millet under different abiotic stresses and their combinations. *Sci. Rep.* 6:23036. doi: 10.1038/srep23036

Singh, D., Chhonkar, P. K., and Dwivedi, B. S. (2005). *Manual on Soil, Plant and Water Analysis*. New Delhi: Westville Publishing House.

Sonkar, S., and Singh, A. (2015). Improvement in nutritional properties of product through fortification with aloevera, mushroom and pearl millets. *Plant Arch.* 15, 795–800.

Sperotto, R. A., Boff, T., Duarte, G. L., Santos, L. S., Grusak, M. A., and Fett, J. P. (2010). Identification of putative target genes to manipulate Fe and Zn concentrations in rice grains. *J. Plant Physiol.* 167, 1500–1506. doi: 10.1016/j.jplph.2010.05.003

Tadesse, W., Ogbonnaya, F. C., Jighly, A., Sanchez-Garcia, M., Sohail, Q., Rajaram, S., et al. (2015). Genome-wide association mapping of yield and

grain quality traits in winter wheat genotypes. *PLoS ONE* 10:e0141339. doi: 10.1371/journal.pone.0141339

Tako, E., Reed, S. M., Budiman, J., Hart, J. J., and Glahn, R. P. (2015). Higher iron pearl millet (*Pennisetum glaucum* L.) provides more absorbable iron that is limited by increased polyphenolic content. *J. Nutr.* 14:11. doi: 10.1186/1475-2891-14-11

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Tara, C. S., Begum, S., Singh, B., Unnikrishnan, K., and Bharadwaj, C. (2009). Analysis of diversity among cytoplasmic male sterile sources and their utilization in developing F1 hybrids in Pearl millet [*Pennisetum glaucum* (R.) Br]. *Indian J. Genet. Plant Breed.* 69, 352–360.

Tara, C. S., Tiwari, S., Bharadwaj, C., Rao, A. R., Bhat, J., and Singh, S. P. (2013). Genetic diversity analysis in a novel set of restorer lines of pearl millet [*Pennisetum glaucum* (L.) R. Br] using SSR markers. *Int. J. Plant Res.* 26, 72–82. doi: 10.5958/j.2229-4473.26.1.011

Tiwari, C., Wallwork, H., Arun, B., Mishra, V. K., Velu, G., Stangoulis, J., et al. (2016). Molecular mapping of quantitative trait loci for zinc, iron and protein content in the grains of hexaploid wheat. *Euphytica* 207, 563–570. doi: 10.1007/s10681-015-1544-7

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*. 6:24050. doi: 10.1038/srep24050

Vadez, V., Hash, T., Bidinger, F. R., and Kholova, J. (2014). "Phenotyping pearl millet for adaptation to drought," *Drought Phenotyping in Crops: From Theory to Practice*, eds P. Monneveux and J. M. Ribaut (Frontiers E-books), 158.

Vaiman, D., Mercier, D., Moazami-Goudarzi, K., Eggen, A., Ciampolini, R., Lépingle, A., et al. (1994). A set of 99 cattle microsatellites: characterization, synteny mapping, and polymorphism. *Mamm. Genome* 5, 288–297. doi: 10.1007/BF00389543

Varshney, R. K., Graner, A., and Solrells, M. E. (2005). Genic microsatellite markers in plant: features and applications. *Trends Biotechnol.* 23, 48–55. doi: 10.1016/j.tibtech.2004.11.005

WHO (2012). *The World Health Report.* Geneva: World Health Organization.

Wright, S. (1978). *Variability Within and Among Natural Populations in Evolution and the Genetics of Populations, Vol. 4.* Chicago, IL: University of Chicago Press.

Würschum, T., Langer, S. M., Longin, C. F. H., Korzun, V., Akhunov, E., Ebmeyer, E., et al. (2013). Population structure, genetic diversity and linkage disequilibrium in elite winter wheat assessed with SNP and SSR markers. *Theor. Appl. Genet.* 126, 1477–1486. doi: 10.1007/s00122-013-2065-1

Yu, J., Pressoir, G., Briggs, W. H., Bi, I. V., Yamasaki, M., Doebley, J. F., et al. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38, 203–208. doi: 10.1038/ng1702

Yu, Y. H., Shao, Y. F., Liu, J., Fan, Y. Y., Sun, C. X., Cao, Z. Y., et al. (2015). Mapping of quantitative trait loci for contents of macro-and microelements in milled rice (*Oryza sativa* L.). *Agric. Food Chem.* 63, 7813–7818. doi: 10.1021/acs.jafc.5b02882

Zhang, T., Qian, N., Zhu, X., Chen, H., Wang, S., Mei, H., et al. (2013). Variations and transmission of QTL alleles for yield and fiber qualities in upland cotton cultivars developed in China. *PLoS ONE* 8:e57220. doi: 10.1371/journal.pone.0057220

Zhao, K., Aranzana, M. J., Kim, S., Lister, C., Shindo, C., Tang, C., et al. (2007). An Arabidopsis example of association mapping in structured samples. *PLoS Genet.* 3:e4. doi: 10.1371/journal.pgen.0030004

Zhao, K., Tung, C. W., Eizenga, G. C., Wright, M. H., Ali, M. L., Price, A. H., et al. (2011). Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Commun.* 2:467. doi: 10.1038/ncomms1467

Zhu, C., Gore, M., Buckler, E. S., and Yu, J. (2008). Status and prospects of association mapping in plants. *Plant Genome* 1, 5–20. doi: 10.3835/plantgenome2008.02.0089

Conflict of Interest Statement: 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.

Copyright © 2017 Anuradha, Satyavathi, Bharadwaj, Nepolean, Sankar, Singh, Meena, Singhal and Srivastava. 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) or licensor 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.