The GLDC Innovation Fund-Final report

Activity Title	Association mapping of nutritional quality traits (protein, iron and zinc) in chickpea
Activity period	1 Aug 2018 to 30 June 2019
Lead Centre	ICRISAT
Key contact from Lead centre	
Name	Srinivasan Samineni (Collaborating scientists: Mahendar Thudi, Rajeev Gupta, Pooran Gaur)
Designation	Senior Scientist (Chickpea Breeding)
Country	India
Name of Partner organisation	Indian Institute of Pulses Research (ICAR-IIPR), Kanpur
Key contact from Partner organisation	
Name	Biswajit Mondal
Designation	Scientist
Telephone	-
Mobile	+91-7505423161
Fax	-
Email	biswagpb@gmail.com
Country	India

*Project funds could not be utilized to conduct the proposed activities due to issues related to finance department in ICAR.

Rationale:

Chickpea (*Cicer arietinum* L.) is the world's second-largest smallholder-cultivated food legume crop. It has historically been an important daily staple in the diet of millions of people's lives in the developing countries. Chickpea grain is an excellent source of high-quality proteins and micronutrients such as Iron (Fe) and Zinc (Zn). A large proportion of the Indian population, particularly children suffer from protein and micronutrient deficiency. Genetically biofortified crops offer a sustainable and low-cost solution to provide essential micronutrients to people. Chickpea is the most important pulse crop of India accounting for about 45% of pulse production and providing a range of health benefits to people. Large genetic variations exist for contents of protein and micronutrients in chickpea. These variations present a viable opportunity for genetic biofortification. Precision and efficiency of breeding programs for improving these traits can be enhanced by using marker-assisted selections. Marker-trait associations can be identified through association mapping. In addition, efficient phenotyping methods need to be established and standardized for phenotyping the large number of

accessions. High throughput machines XRF and NIRS are available for the quality analysis of grain samples after calibration of best curves. This will enable us to screen large number of samples at very low cost and in a short time.

Summary

The present study was designed to develop micronutrient-dense chickpea lines, by evaluating the genetic variability and identifying marker trait association for grain protein, Fe and Zn contents. The experiment was conducted using 140 diverse chickpea accessions in two different locations (ICRISAT, Patancheru and ICAR-IIPR, Kanpur) under different environment conditions (normal and late planting). The trial at ICAR-IIPR was failed due to seed damage by bruchids, hence seed could not be analysed for nutrients traits. Phenotypic data from the trial conducted at ICRISAT revealed a large variation for phenology, plant height, grain yield and 100 seed weight traits. Wide range of variation was observed for grain Fe (55.6-80.0 & 45.6-66.7 ppm) and zinc (32.3-48.7 & 32.1-52.5 ppm) contents under normal and heat stress conditions, respectively. Under normal season protein content varied from 21.8 to 29.5%. A single marker (S3_16891399) was associated with grain Fe content on Chr. 3. Two markers were associated with grain Zn content on Chr. 3 (S3_16891399) and Chr.6 (S6_10479236). A single co-localized marker (S3_16891399) for Fe and Zn was found on Chr. 3. For protein content a single SNP (S6_13523529) with $P \leq 2.40 \times 10^{-3}$ was significantly associated on Chr.6. High throughput phenotyping machines were calibrated for estimation of protein, Fe and Zn contents in seeds. XRF machine was calibrated for Fe and Zn contents where R2 value was found to be 0.74 (SE=6.9) and 0.58 (SE=5.8), respectively. To measure the protein content, NIRS machine was calibrated and the R^2 value was estimated to be 0.62 (SE=0.2).

Methodology

A set of 140 diverse chickpea (*Cicer arietinum* L.) accessions were used for this study. Field experiments were conducted in two locations at ICRISAT, Patancheru and IIPR, Kanpur under normal sown and late sown conditions. These experiments were planted in randomized complete block design in three replications on 12th Oct 2018 and 4th Nov 2018 (normal conditions) 17th Jan 2019 and 10th Dec 2018 (late sown conditions) at ICRISAT Patancheru and IIPR, Kanpur respectively. All the genotypes were planted in single row with four-meter row length with inter row distance of 60cm. During the crop growth period, observations were recorded on days to 50% Flowering (DF), days to maturity (DM), plant height (cm) and plant stand (PLSD). Currently harvesting of the trial is in progress where grain yield per plot (Yldg) and 100 seed weight are being recorded at IIPR location. Post-harvest observations are being taken at Patancheru location. Grain nutrient contents were estimated using XRF (Grain Fe and Zn) and NIRS (protein) machines and the traits correlations determined by statistical software.

Nutrient estimation:

The grains were harvested at maturity and stored in glassine bags. Precautions were taken to avoid any metallic or dust contamination of grains while harvesting and analyzing. For each accession the grains were divided into three parts and analyzed as three replicates for Fe and Zn content using simultaneous multi-element inductively coupled plasma–optical emission

spectrometer (ICP-OES, Perkin Elmer). Briefly, the whole grain samples were quickly washed with distilled water to remove any surface contamination and dried in hot air oven at 50°C for 24 h. The samples (0.5 g) along with operational blanks and standard solution of known concentrations were digested in 5 ml of distilled nitric acid (Analytical Reagent Grade, Merck) at 140°C for 45 min in a Microwave Digestion System (Perkin Elmer) to obtain clear digests. Following digestion, the volume of each sample was made up to 25 ml using Milli-Q water and elemental determination was performed by ICP-OES. For calculating the grain micronutrient concentration, the mean of element specific blank concentration was subtracted from each data point. The data were then multiplied by initial sample volume, divided by initial weight of grains, and expressed as µg element g-1 dry grain material (ppm) (Khokhar et al., 2018).

Under generation challenge program (GCP), the chickpea accessions were genotyped using the KASPar assays method as described in Hiremath et al., 2012. Phenotypic best linear unbiased predictor (BLUP) was estimated for each accession and trait using GenStat v16 software (VSN International, Hemel Hempstead, Hertfordshire, United Kingdom).

Association Mapping

Association mapping based on a GLM, MLM, FarmCPU, SUPER and BLINK were conducted primarily using the Genome Association and Prediction Integrated Tool (GAPIT; Lipka et al. 2012) in R: a language and environment for statistical computing. It includes the phenotypic data with SNP markers coming from the high-density arrays. Significant association between a marker and phenotypic traits was declared at $P \leq 0.01$ presented in Manhattan plots. For significant marker-trait associations (MTAs) detection, we set a threshold P-value of -log10 (P) ≥ 2.5. Quantile-Quantile (Q-Q plots) of the observed and expected P values were plotted to assess the adequacy of a fitted normal straight line to the markers. Q-Q plots were drawn based on the observed and expected -log10 (P) values. Explained phenotypic variance (R^2) and marker effects (positive/negative) were extracted from GWAS results.

Results

Phenotyping

At ICRISAT, Patancheru location data was recorded under normal sown conditions (Table 1). In two genotypes (ICCV 96030, ICC 16641) flowering was observed at 26 days after flowering, first podding at 34 days and maturity at 68, 75 days, respectively. Maximum plant height was recorded in ICC 8923 (68 cm) followed by ICC 12955 and ICCV 92219. High grain yield (339-376 g) was observed in following genotypes ICC 11321, ICCV 97126, ICC 11102 and ICC 10136. The genotypes with 100 seed weight more than 30 g was observed in ICCV 05107, ICCV 97306, ICCV 8305, ICC 15833, ICCV 95418 and ICC 4936, whereas it was less than 8 g in ICC 5362 and ICC 5590. Post-harvest observations are being taken on the experiment conducted under late planting conditions.

At ICAR-IIPR Kanpur, ICCV 96030 initiated flowering at 40 and 36 days under normal and late sown conditions, respectively. The entries ICCV 96030, ICCV 8305, ICC 5362, ICC 8584, ICC 16216, ICC 5434, ICC 16641 and ICC 16343 flowered early under both sowing

conditions. Final data on field experiment could not be recorded as the experiment was failed. The same experiment has been planted again during 2019 crop season.

Table 1: Characterization of diverse chickpea genotypes for various agronomic traits grown under normal season at ICRISAT, Patancheru during 2018-19.

Trait	**Minimum value**	**Maximum value**	**Average of all genotypes**
Days to First Flowering	23.0 (ICCV 96030)	85.0 (ICC 12366)	47.0
Days to 50%Flowering	26.0 (ICCV 96030)	89.0 (ICC 12366)	51.0
Days to First Podding	34.0 (ICCV 96030)	93.0 (ICC 12366)	58.0
Days to Maturity	68.0 (ICCV 96030)	132.0 (ICC 7509)	103.0
Plant Height (cm)	20.2 (ICC 16948)	68.0 (ICC 8923)	42.0
Grain Yield (g)	66.3 (ICC 16641)	375.6 (ICC 11321)	221.0
100 Seed Weight (g)	6.3 (ICC 5362)	35.0 (ICCV 05107)	17.6

Calibration of XRF machine for Fe Zn contents: X-ray Fluorescence analysis (XRF) is a high throughput machine ideal for screening large numbers of samples for Fe and Zn content in whole grains of rice, pearl millet and now has been calibrated for chickpea. The machine can perform non-destructive analysis and no hazardous chemicals used to estimate Fe and Zn from the grains. About 100 diverse chickpea genotypes were analysed from standard accredited chemistry lab in Adelaide, Australia for grain Fe and Zn contents. For calibration of machine, we have selected 32 chickpea accessions at different intervals from a range of 41.0 to 84.0 ppm for Fe and 24 to 53 ppm for Zn. The regression coefficient of the calibration curve for Fe and Zn contents was 0.74 and 0.58, respectively. However, the standard errors (SE) were found high (6.9 for Fe and 5.8 for Zn) for these curves. Efforts are in progress towards reducing the SE of the calibration curves by using more diverse genetic backgrounds.

Calibration of NIRS machine for protein content: Near infrared reflectance spectroscopy (NIRS) is a robust, cost-effective and non-destructive method of estimating oil, protein and fatty acids. The technique of NIRS was utilized at ICRISAT to predict the performance of genotypes including oil and protein content in station trials. Calibration equations are being developed using the readouts of wet chemistry methods. Different mathematical treatments were tested to identify the best calibration equation based on their coefficient of determination in calibration (R^2) and coefficient of determination in cross-validation (1-VR) values. Preliminary analysis could identify best curve with R^2 value 0.62 (SE=0.2). More diverse genotypes are being used now to improve the regression coefficient.

Nutrient analysis and association studies: The phenotypic values for the 2 seasons were converted into BLUP values to get unbiased mean estimates. Under normal season protein content varied from 21.8 to 29.5%. Wide range of variation was observed for grain Fe (55.6-

80.0 & 45.6-66.7 ppm) and zinc (32.3-48.7 & 32.1-52.5 ppm) contents under normal and heat stress conditions, respectively. A strong positive linear relationship was found between BLUPs and mean values with the shrinkage of BLUPs toward the population average. The BLUP values depicted a normal distribution for grain Fe and Zn content.

In normal season significant positive correlations between grain Fe and Zn content were observed, as micronutrient concentrations are highly influenced by the environment. No correlation was observed under heat stress conditions between grain Fe and Zn content (Fig. 1a,1b).

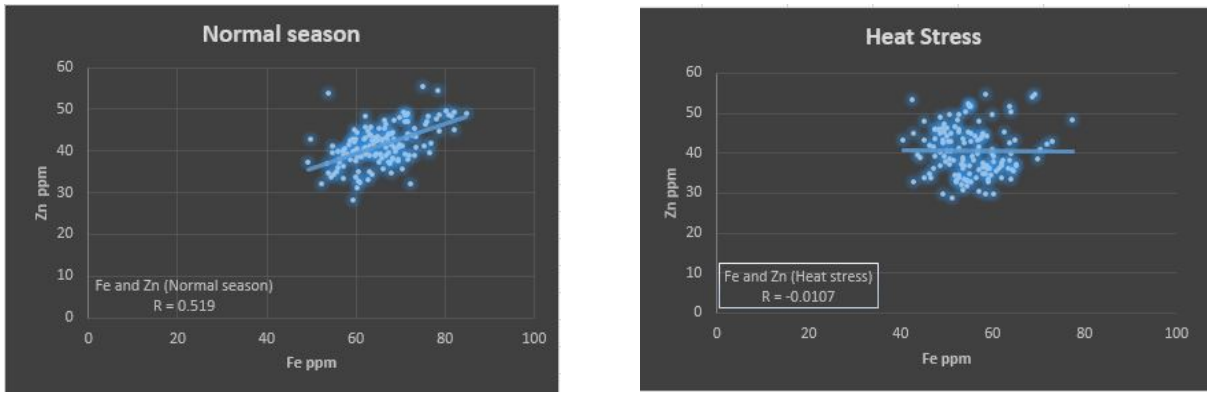

Fig 1. Correlation for Fe and Zn grain content (a) Normal season (b) Heat stress

KASPar assays based SNP markers were used to find the genetic identity between the accessions. From the group of accessions that had >99% genetic identity and high phenotypic similarity, only single accessions were retained for further marker trait association (MTA) studies. The association mapping was conducted for 140 non-redundant accessions using 333 SNP markers on the BLUP values of protein, Fe, Zn traits. For conducting MTA, an R-GWAS package called GAPIT. It includes the phenotypic data with SNP markers coming from the high-density arrays. GWAS analysis was done for 140 diverse genotypes with 333 SNPs to scan the association of the existing SNP polymorphism with the variations in grain protein, Fe, Zn content in the population. A single marker (S3_16891399) was associated with grain Fe content on Chr. 3. Two markers were associated with grain Zn content on Chr. 3 (S3_16891399) and Chr.6 (S6_10479236). In addition, on Chr. 3 a single marker (S3_16891399) was co-associated with grain Fe ($P \leq 1.99 \times 10^{-3}$) and Zn ($6.8 \times 10^{-5}$) content. For protein, a single marker (S6_13523529) was found to be associated ($P \leq 2.40 \times 10^{-3}$) on Chr.6. The results were presented in the Manhattan plot (Fig. 2 a, b and c). The threshold calculated by –log 10 value of P value for Manhattan plot. The Q-Q plot shows the associated SNPs based on the observed and expected -log10 (P) values and the results were matched with the Manhattan plotted SNP (Fig. 3 a, b and c).

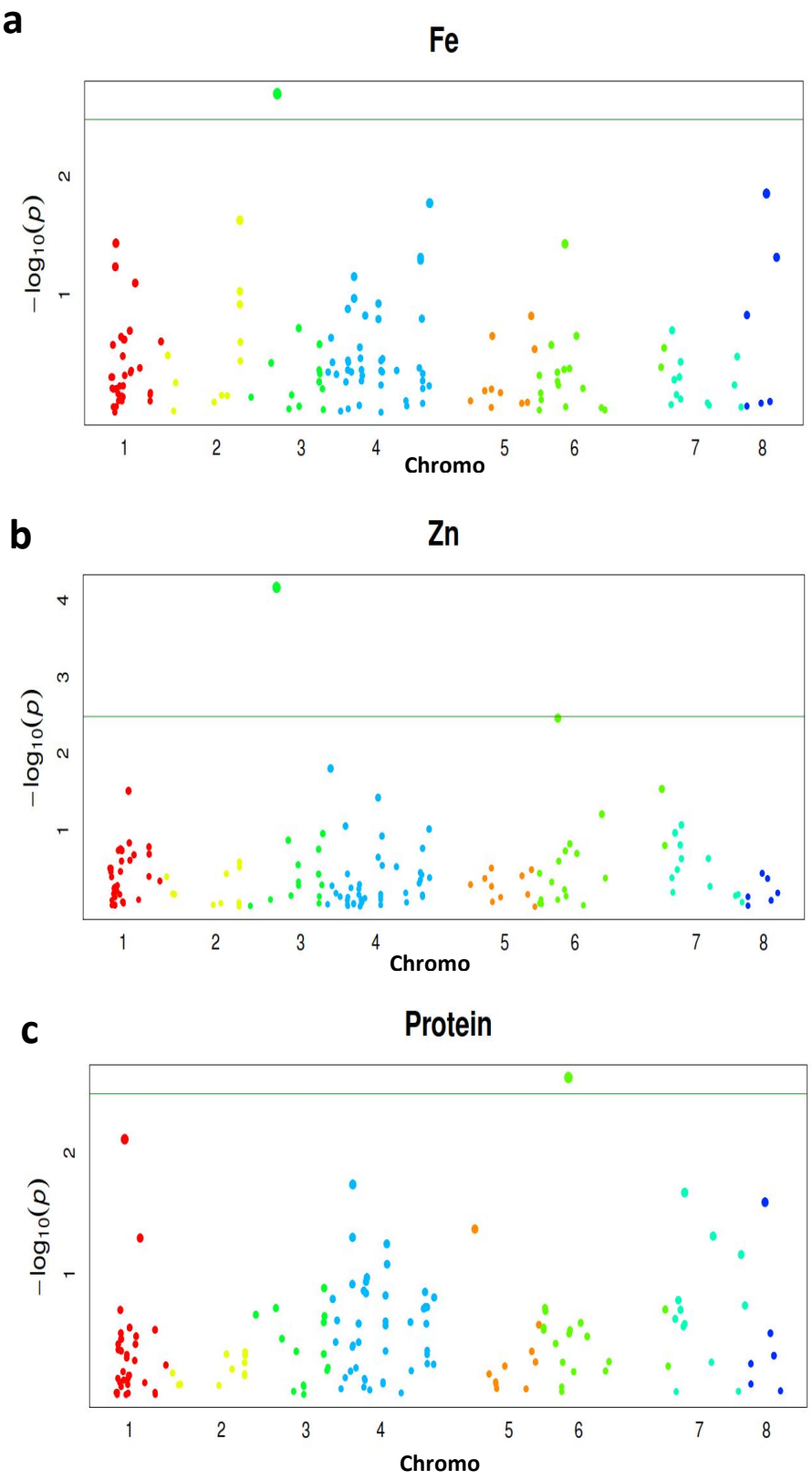

Fig 2. Manhattan plot GWAS model for **(a) Fe, (b) Zn and (c) Protein** plotted with the individual SNPs of all chromosomes on the X-axis and − log10 P value of each SNP in the Y-axis. The different colors indicate the 8 chromosomes of chickpea. The light green horizontal line shows the multiple testing threshold − *log10 P* value of 2.5 for this panel. [Fe and Zn values in ppm (parts per million) and Protein value in % (percentage)]

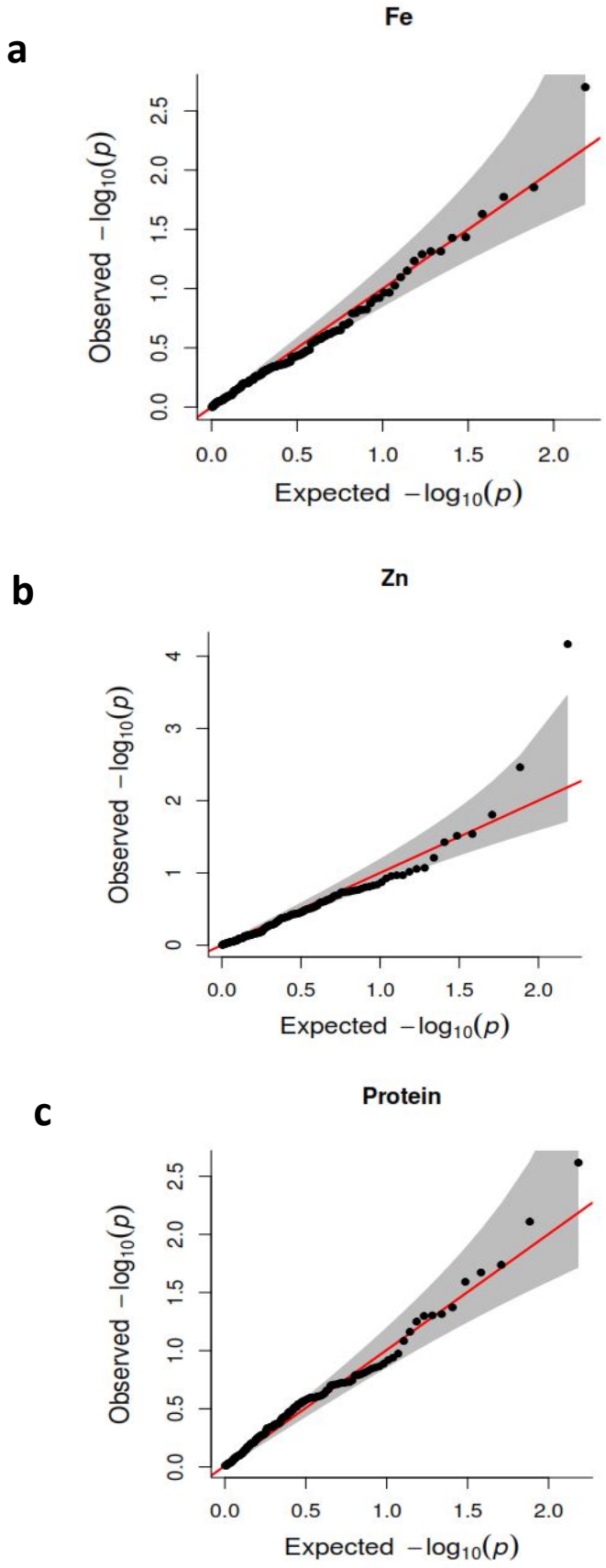

Fig 3. Quantile-quantile plot (QQ-plot) from GWAS model (a) Fe, (b) Zn, and (c) Protein illustrating the comparison between expected and observed -$\log_{10}$(P)-values with a FDR cut-off <0.05 to detect significant genomic loci (genes) associated with Grain-Fe Zn and Protein concentrations in chickpea. [Fe and Zn values in ppm (parts per million) and Protein value in % (percentage)]