Theoretical and Applied Genetics
https://doi.org/10.1007/s00122-019-03277-5

ORIGINAL ARTICLE

# Co-localization of genomic regions associated with seed morphology and composition in a desi chickpea (*Cicer arietinum* L.) population varying in seed protein concentration

**Runfeng Wang[1] · Manu P. Gangola[1] · Craig Irvine[1] · Pooran M. Gaur[2] · Monica Båga[1] · Ravindra N. Chibbar[1]**

Received: 20 June 2018 / Accepted: 7 January 2019


## Abstract

***Key message*** **Major QTL on LG 1 and 3 control seed filling and seed coat development, thereby affecting seed shape, size, color, composition and weight, key determinants of crop yield and quality.**

**Abstract** A chickpea (*Cicer arietinum* L.) population consisting of 189 recombinant inbred lines (RILs) derived from a cross between medium-protein ICC 995 and high-protein ICC 5912 genotypes of the desi market class was analyzed for seed properties. Seed from the parental lines and RILs was produced in four different environments for determination of seed shape (SS), 100-seed weight (100-SW), protein (PRO) and starch (STA) concentration. Polymorphic genetic markers for the population were identified by Genotyping by Sequencing and assembled into a 522.5 cM genetic map. Phenotype data from the different growth environments were analyzed by QTL mapping done by single and multi-environment analyses and in addition, single marker association mapping. The analyses identified in total 11 QTL, of which the most significant ($P < 0.05$) loci were located on LG 1 (q-1.1), LG 2 (q-2.1), LG 3 (q-3.2, q-3.3), LG 4 (q-4.2), and LG 5 (q-5.1). STA was mostly affected by q-1.1, which explained 19.0% of the phenotypic variance for the trait. The largest QTL effects were demonstrated by q-3.2 that explained 52.5% of the phenotypic variances for 100-SW, 44.3% for PRO, and 14.6% for SS. This locus was also highly associated with flower color (COL; 95.2% explained) and showed q-3.2 alleles from the ICC 5912 parent conferred the blue flower color and production of small, round seeds with relatively high protein concentration. Genes affecting seed filling at q-1.1 and seed coat development at q-3.2, respectively, were considered to underlie differences in seed composition and morphology in the RIL population.

## Introduction

Chickpea (*Cicer arietinum* L.) is an annual diploid ($2n = 2x = 16$) and self-pollinated legume crop, which ranks fourth among pulses in terms of production in the world (FAOSTAT 2016; http://www.fao.org/faostat/). Most of the chickpea cultivation and consumption is concentrated to the Indian subcontinent, the Middle East and the Mediterranean countries. Two market classes of chickpea denoted desi and kabuli are cultivated. Desi chickpea are mainly grown in semiarid tropical areas and typically produce small, slightly angular, dark-colored seeds with thick seed coat. A cooler temperate climate is preferred for the kabuli chickpea, which develop large, smooth and cream-colored seeds with thin seed coat. The carbohydrate storage component in desi seeds is large (51–65%) with a major contribution from starch (32–56%) followed by cellulose (4–13%), hemicellulose (3.5–8.8%) and pectin (1.5–3.8%) (Wood and Grusak 2007). Proteins constitute the second largest seed fraction

Communicated by Henry T. Nguyen.

Runfeng Wang and Manu P. Gangola have contributed equally to this work.

**Electronic supplementary material** The online version of this article (https://doi.org/10.1007/s00122-019-03277-5) contains supplementary material, which is available to authorized users.

✉ Ravindra N. Chibbar
ravi.chibbar@usask.ca

[1] Department of Plant Sciences, University of Saskatchewan, Saskatoon, SK S7N 5A8, Canada

[2] International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Hyderabad, Telangana 502 324, India

Published online: 19 January 2019

(16.7–30.6%), whereas lipids are present in smaller quantities (2.9–7.4%). The essential amino acids are well represented in chickpea except methionine and cysteine, which amount to undesirably low levels; nevertheless, chickpea provide valuable nutrients for both humans and livestock because of the high protein content and adequate levels of many vitamins and minerals (Wood and Grusak 2007).

Seed size is a component of crop yield and has been the target for selection since the beginning of crop domestication. An important factor for seed yield is the source–sink relationship, which determines nutrient allocation to seeds upon integration of cellular, hormonal and environmental signals (Yu et al. 2015; Ainsworth and Bush 2011; Ruan et al. 2012). During seed development, the source tissues provide a flux of sugars and amino acids to support growth of the three major seed compartments: triploid endosperm, zygotic embryo and surrounding maternal integuments. As seeds are formed within the ovule of the plant, the surrounding maternal tissues can greatly influence the development of three seed compartments (Alonso-Blanco et al. 1999).

The sequential enlargement of the different seed organs is genetically determined by the developmental program seeds undergo from the double fertilization step to maturation, but the process is influenced by environmental signals (Mizukami 2001). Endosperm development initiates soon after fertilization with a series of mitotic divisions that generate a multinucleate endosperm called coenocyte (Weber et al. 2005). This step is followed by endosperm cellularization and differentiation coinciding with shrinkage of the central vacuole storing imported nutrients (Morley-Smith et al. 2008). During this phase, auxin signals derived from the endosperm initiate growth and differentiation of maternal integuments that will form the seed coat (Figueiredo et al. 2016; Ingouff et al. 2006).

A second phase of seed development involves cell division in embryo, which depends on endosperm cellularization (Hehenberger et al. 2012), followed by a maturation phase when embryo grows by cell expansion (Weber et al. 2005). In dicots such as chickpea, almost all of the transiently stored nutrients in the early endosperm and seed coat are allocated to the embryo and the two large cotyledons during seed maturation. Thus, the final seed size is largely determined by the coordinated growth and development of the endosperm and seed coat as these events impact the subsequent differentiation and growth of the embryo (Garcia et al. 2005; Hehenberger et al. 2012). Many genes and pathways controlling the different phases of seed development in model crops such as Arabidopsis are known (Li and Li 2016; Sun et al. 2010; Berger and Chaudhury 2009), but they are less characterized in crop species. To increase productivity and seed quality in crops, it is imperative to understand the metabolic and regulatory networks that coordinate the different phases of seed development; this will allow selection of robust genetic markers to be used in marker-assisted selection for crop improvement.

The diploid and self-pollinated nature of chickpea facilitates genomic mapping of agronomic and seed quality traits. Recently, the progress in the Next-Generation Sequencing (NGS) technologies has aided high-throughput identification of single-nucleotide polymorphism (SNP) markers for chickpea genotypes and facilitated development of high-density genomic maps for genotypes ICC 4958 (Jain et al. 2013; Parween et al. 2015) and CDC Frontier (Varshney et al. 2013). The advancements of genomics technologies for plants have also led to the application of Genotyping-by-Sequencing (GBS) technologies for mapping of Quantitative Trait Loci (QTL) (Sonah et al. 2013). Most of the available QTL information for chickpea seed quality is derived from studies of yield, seed weight, seed color and seed coat thickness (Bajaj et al. 2015; Cobos et al. 2009; Das et al. 2015; Kujur et al. 2015; Saxena et al. 2014; S. Verma et al. 2015a). QTL reports on intrinsic seed quality traits are mainly focused on the content of protein (Upadhyaya et al. 2016b; Jadhav et al. 2015), minerals (Upadhyaya et al. 2016a) and carotenoids (Abbo et al. 2005). Traits affecting seed composition are generally controlled by numerous genes, each with small effect and often influenced by environmental conditions. Thus, consistent QTL can be identified by QTL studies across multiple populations and environments. To improve our understanding of the complex processes that govern seed development and seed filling in chickpea, we have analyzed the composition in seeds produced by a RIL population in four widely different environments. GBS was used for genotyping and allowed genomic regions associated with seed weight, concentration of protein and starch in seeds, seed shape, and flower color to be identified. Based on the current knowledge of seed development, candidate genes for identified QTLs were suggested.

## Materials and methods

### Plant material and growth of plants

A medium-protein genotype (ICC 995; female) of Indian origin was crossed with a high-protein Mexican genotype (ICC 5912; male) at the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) in Patancheru, India. Desi market class seeds are produced by both parents. The progenies were advanced to the F6 generation to generate a RIL population of 240 lines, of which 189 lines were studied in this report. To assess variation in seed quality traits in parents and RILs, seeds from the different genotypes were produced at three field locations and in one greenhouse environment (Table S1). The ICRISAT (ICR) trial was conducted in Patancheru, India, with relatively hot

climate and the Biggar (BIG) and Aberdeen (ABE) trials were done in Saskatchewan, Canada, with cooler climates. Seeding in the field was done in 7.5 m rows with 25 cm row spacing, 50 seeds/row and two randomly placed rows per genotype and trial. Fertilizer (10 kg/ha P; 40 kg/ha N) was applied before seeding, and weeds were controlled by manual weeding. In the greenhouse (GH) trial, the genotypes were grown in one-gallon pots and exposed to 18 h day length with 380 µM $m^{-2}$ $s^{-1}$ photosynthetic flux density and 20/18 °C day/night temperatures. Fully mature seeds from the trials were harvested and stored at room temperature until analysis.

## Determination of flower color, seed shape, and seed weight

Flower color (COL; pink or blue) and seed shape (SS; round = 1, slightly angular = 2, and angular = 3) were rated only for plants grown in the ICR trial. 100-seed weight (100-SW) for each genotype was determined for each of the four trials and determined from the average weights of three seed aliquots of 100 seeds each.

## Determination of total starch and protein concentrations in seeds

Chickpea seeds (about 10 g) were ground into a fine meal using a UDY cyclone mill (Udy Corporation, Fort Collins, CO, USA) equipped with a 0.5-mm sieve. The meal was stored at room temperature and assayed within 4 weeks. The total starch concentration in 100 ± 0.5 mg meal (STA) was determined in duplicate samples and using the Total Starch Assay Kit (Megazyme International Ireland Ltd, Wicklow, Ireland). Protein concentration in meal (PRO) was determined in flour samples using a combustion method as described (Wang et al. 2017).

## Statistical analyses

Statistical analysis including analysis of variance (general linear model), Pearson's correlation coefficients, and the Anderson–Darling normality test were performed using the MINITAB 17 statistical software (Minitab Inc., State College, PA, USA). Components of variance were derived from QTL data analysis using QTLNetwork software ver. 2.1 (Yang et al. 2008).

## DNA extraction and GBS

Leaf samples were collected from each genotype grown under optimal greenhouse conditions for 2 weeks, frozen in liquid nitrogen and stored at − 80 °C. Genomic DNA from collected samples was isolated using the Qiagen DNeasy Plant Mini Kit (Qiagen GmbH, Hilden, Germany) and quantified by the Quant-iT™ PicoGreen® dsDNA Assay Kit (Invitrogen, Molecular Probes, Eugene, OR, USA). DNA preparations were normalized to 10 ng $\mu L^{-1}$ and used for preparation of libraries for Ion Proton Genotyping by Sequencing (GBS), a service provided by Plateforme d'analyses génomiques at the Institut de Biologie Intégrative et des Systèmes (IBIS, Université Laval, Québec, Canada) and performed as described (Mascher et al. 2013) with the following exceptions: *ApeKI* with corresponding barcodes were used instead of the *PstI*/*MspI* combination, and the libraries were sized before PCR amplification using a BluePippin Size Selection System (Sage Science Inc., Beverly, MA, USA). Ion CHEF System, Ion PI Hi-Q Chef reagents and Ion P1 Chips v3 (Thermo Fisher Scientific, Waltham, MA, USA) were used for preparation of libraries and chips. The Ion Proton System (Thermo Fisher Scientific, Waltham, MA USA) was used for the DNA sequence analysis.

## Gene-specific and SSR markers

Primer used for PCR amplification of CaGM13632 and CaGM13641 simple sequence repeat (SSR) markers, ANGUSTIFOLIA3-like (AN3) and AHA10-like genes are described in Table S2. The PCR was done in 25 µL volumes containing 50 ng genomic DNA, 0.2 mM of each dNTP, 0.2 µM of each primer, 1 unit DreamTaq Polymerase (Thermo Fisher Scientific, Waltham, MA, USA) and supplied buffer adjusted to 2.5 mM $Mg^{2+}$. Amplifications were done for 32 cycles of 20 s at 94 °C, 20 s at 57 °C, and 60 s at 72 °C, followed by 1 cycle at 72 °C for 10 min. The SSR PCR products were analyzed by 1.8% agarose gel electrophoresis and the AN3 marker by the single-strand conformation polymorphism system as described (Tondelli et al. 2006). PCR products for the AN3 and AHA10 markers were analyzed by DNA sequencing performed by Eurofins Genomics (Louisville, KY, USA).

## Processing of GBS data and construction of genetic map

The raw "fastq" files obtained from the parent lines in duplicate and 189 RILs were processed using Trimmomatic version 0.36 software (Bolger et al. 2014) with a four-base window setting and a minimum average quality score of 16. Adapter sequences and read lengths less than 36 nucleotides were removed, and sequences greater than 170 nucleotides were trimmed. The processed sequences were aligned to the reference ICC 4958 chickpea genome, version 3.0 (Edwards 2016) using Bowtie2, version 2.2.6 software (Langmead et al. 2009). Duplicates were removed using the rmdup function, and SNPs including insertion/deletion sites (InDels) were called using the Samtools software suite, version 0.1.19

(Li et al. 2009). The markers were filtered using VCFtools version 0.1.15 (Danecek et al. 2011) with the following settings: minimum allele frequency of 1%, maximum missing data of 8% and minimum read depth of 6.0.

The JoinMap 3.0 software (van Ooijen and Voorips 2001) was used for construction of a genetic map for the ICC 995 × ICC 5912 population. The markers were grouped with a minimum logarithm of odds ratio (LOD) score of 3.0 and a recombination fraction of < 0.40. Conversion of recombination frequencies into centiMorgan (cM) genetic distances was done by the Kosambi mapping function.

### QTL mapping

QTL mapping was done using the QTLNetwork software (Yang et al. 2008). The genome-wide error rate was set to 0.05, and *F* value thresholds for QTL detection were determined by 1000 permutations. QTL and environmental interactions were identified by multi-environment analyses. One-dimensional genome scan to select single-locus QTL and their environmental interactions was done using the general linear model (GLM)-based interval mapping. QTL with epistatic interactions were identified by a two-dimensional genome scan. An experimental significance level of 0.05 was selected with 5 cM testing and filtration windows and walking speed of 1 cM. The proportion of observed phenotypic variance explained by each additive and epistatic QTL and the corresponding additive effects were estimated. The genetic markers were also tested individually for association with trait values by the general linear model (GLM) used by the TASSEL 3.0 software (Glaubitz et al. 2014). The nomenclature of identified QTL included four parts: trait (SW, PRO, STA, SS), location (ICR, GH, BIG, ABE), linkage group of sequenced ICC 4958 genome, and QTL order on linkage group (e.g., qSW-ICR-1.1 for locus identified at ICR location; qSW-1.1 for locus identified by multi-environment analysis). Genetic maps showing the identified QTL was graphically displayed using the MapChart 2.3 software (Voorrips 2002).

### Prediction of candidate genes for QTL

Predicted genes within the QTL confidence intervals on the ICC 4958 genome assembly uwa v3.0 (Edwards 2016) were examined using the online tools UniProt Knowledgebase [(Magrane and Consortium 2011); www.uniprot.org/], Quick GO (http://www.ebi.ac.uk/QuickGO/), the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (Szklarczyk et al. 2015), Chickpea Transcriptome Database (Verma et al. 2015b) and resources at the NCBI database (https://www.ncbi.nlm.nih.gov/).

## Results

### Analysis of seed quality traits in parental lines and RIL population

The 189 RILs analyzed were produced from a cross between genotypes ICC 995 and ICC 5912, which differ in flower color, seed morphology and composition. Pink flowers, slightly angular and light-colored seeds with about 17% protein concentration, are produced by ICC 995. The ICC 5912 genotype produces blue flowers, relatively small, round/pea-shaped and dark-colored seeds with high protein concentration (25–29%) (Gaur et al. 2016; Wang et al. 2017). To determine the genomic regions causing these phenotypic differences between the parental lines, the RIL population was investigated by QTL mapping.

Seeds produced in one greenhouse experiment in Saskatoon, Canada (GH), and three field trials (Table S1) in ICRISAT India (ICR), Biggar Canada (BIG) and Aberdeen Canada (ABE) were studied for 100-SW and the two major seed components, starch and protein (Table S4). The protein concentration in harvested ICC 5912 seeds from the four locations varied from 20.8 to 26.4%, and was 5.2–8.9 percentage units higher than values for the ICC 995 seeds (15.6–19.5%). However, ICC 5912 seeds produced in the different locations had lower 100-SW (7.2–19.7 g) than the ICC 995 seeds (10.1–25.2 g). The starch concentration in ICC 5912 seeds ranged from 37.3 to 48.6% and was 0.9 to 12.5 percentage units lower than values recorded for ICC 995 seeds (45.3–51.8%) in the BIG, GH and ICR trials, but slightly higher in the ABE trial (43.4% vs. 42.2%). Thus, the differences in STA values between parental seed samples from the different growth locations were not as consistent as the differences observed for PRO and 100-SW values.

The trait values determined for the RIL population demonstrated a continuous distribution of phenotypic values for the four locations (Fig. 2). However, only 100-SW in the BIG trial (SW-BIG) and STA in the ICR trial (STA-ICR) showed normal distributions as determined by the Anderson–Darling normality test ($P < 0.05$), whereas the remaining distributions appeared multi-modal. Transgressive segregation was observed for all traits, but none of the RILs showed a consistently higher protein concentration in seeds than the ICC 5912 parent.

Seeds harvested from the ABE trial had the highest 100-SW values (20.8 ± 4.2 g), followed by the GH (17.9 ± 6.7 g) and the ICR (17.5 ± 3.8) trials (Fig. 2, Table S4). In the BIG trial, 100-SW values were very low for parents and RILs (10.1 ± 3.0 g), probably due to suboptimal growing conditions in this environment. The BIG trial also showed the largest variation for STA (CoefVar = 15.1) and PRO (CoefVar = 13.3) (Table S4). Nevertheless, 100-SW for the RIL

**Table 1** Consistency of trait values between environments

| Trait | Location | ICR | GH | BIG | ABE |
|---|---|---|---|---|---|
| 100-SW | ICR | 1.00 | | | |
| | GH | 0.88 | 1.00 | | |
| | BIG | 0.66 | 0.67 | 1.00 | |
| | ABE | 0.87 | 0.84 | 0.69 | 1.00 |
| PRO | ICR | 1.00 | | | |
| | GH | 0.65 | 1.00 | | |
| | BIG | 0.63 | 0.57 | 1.00 | |
| | ABE | 0.74 | 0.59 | 0.61 | 1.00 |
| STA | ICR | 1.00 | | | |
| | GH | 0.36 | 1.00 | | |
| | BIG | 0.27 | 0.44 | 1.00 | |
| | ABE | 0.54 | 0.37 | 0.20 | 1.00 |

**Table 2** Correlations between traits in different trials

| Location | Trait | 100-SW | PRO | STA |
|---|---|---|---|---|
| ICR | 100-SW | 1.00 | | |
| | PRO | −0.61 | 1.00 | |
| | STA | 0.31 | −0.36 | 1.00 |
| GH | 100-SW | 1.00 | | |
| | PRO | −0.71 | 1.00 | |
| | STA | 0.51 | −0.49 | 1.00 |
| BIG | 100-SW | 1.00 | | |
| | PRO | −0.61 | 1.00 | |
| | STA | 0.59 | −0.48 | 1.00 |
| ABE | 100-SW | 1.00 | | |
| | PRO | −0.63 | 1.00 | |
| | STA | 0.33 | −0.28 | |

population was relatively stable between trials with correlation factors ranging from 0.66 to 0.88 (Table 1). A relatively good consistency was also noted for PRO ($r^2 = 0.57$–$0.74$), whereas STA was less consistent ($r^2 = 0.20$–$0.54$). As commonly found for seeds, the 100-SW and PRO values were negatively correlated ($r^2 = -0.61$ to $-0.71$; Table 2) and a positive correlation existed between 100-SW and STA values ($r^2 = 0.31$ to $0.59$). The PRO and STA values showed a negative correlation ($r^2 = -0.28$ to $-0.49$).

Seed shape segregated within the RIL population, for which round seeds were produced by 136 lines, slightly angular seeds by 23 lines and angular seeds by 30 lines (Table 3). Each seed shape class showed variation for 100-SW, and an inverse relationship between protein and starch content was noted for RILs producing round seeds. The slightly angular and angular seed classes had on average higher 100-SW, but protein concentration was lower than in the small round seeds. All 59 RILs with blue flowers produced round and small seeds (100-SW ≤ 14.5 g) that were dark colored like the ICC 5912 seeds, whereas 130 RILs with pink flowers produced larger round (100-SW > 17.5 g), slightly angular or angular seeds with seed colors similar to ICC 995 seeds (cream, light brown, reddish brown). Thus, a strong association between flower color, seed shape and seed color as previously observed for chickpea accessions (Gaur et al. 2016) was confirmed for the RIL population. In addition, a difference in seed coat attachment was observed for the seeds samples. Like the ICC 5912 seeds, round small seeds produced by the 59 RILs with blue flowers showed higher seed coat attachment, and thus were more difficult to dehull than the larger seeds. Strong seed coat attachment is a characteristic of kabuli chickpea and certain difficult-to-mill desi chickpea and is associated with differences in seed coat composition when compared to the easy-to-mill desi class (Wood et al. 2014).

## Construction of chickpea linkage map

The GBS analysis generated a total of 654.3 million reads, which were passed through processing and quality filters to produce 580.9 million sequences with Phred quality scores > 20 and an average read length of 103 bp. Alignment of the processed sequences to the ICC 4958 genome assembly uwa v3.0 (Edwards 2016) produced 6388 markers at 6× read depth. A total of 3588 markers (605 SNP and 2983 InDels) remained upon filtering for < 8% missing data and 1% minor allele frequency. Of these markers, 76.8% mapped to annotated genes and 42.4% were positioned within exons. Upon removal of markers showing redundancy between parents, heterozygosity in one or both parents, and InDels containing five or more single-nucleotide repeats, 967 markers remained and were applied for linkage mapping. A total of 16 linkage groups were obtained, but five groups were short and contained less than five markers each. The genetic maps of the 11 longest linkage groups containing 486 SNP and 14 InDels are shown in Fig. 1, where LG 3 has been supplemented with four additional markers (Table S2). Of the eight chickpea chromosomes, LG 4 was split into three linkage groups and LG 7 was composed of two groups, whereas the remaining chromosomes were represented by one group each. Based on marker coverage, low genetic diversity was indicated for LG 2, 3, and 5, whereas LG 4 with the largest number of markers (231) indicted a higher diversity between parents. The marker order along LG 2, 3, 4, 5 and 8 agreed relatively well with the ICC 4958 reference genome (Jain et al. 2013; Parween et al. 2015), but discrepancies in marker order were noted for the distal end of LG 1 and the entire length of LG 6. However, the marker order along LG 6 matched well to the corresponding positions of LG 6 on CDC Frontier genome assembly v. 1.0 (data not shown).

**Table 3** Characteristics of seed shape classes within ICC 995 × ICC 5912 population

| Seed shape | Genotype | Trait | 100-SW (g) | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | ≤11.5 | 11.6–14.5 | 14.6–17.5 | 17.6–20.5 | 20.6–23.5 | 23.6–26.5 |
| Round | 136 RILs | Pink flowers[a] | 0/11 | 0/36 | 13/26 | 36/36 | 22/22 | 5/5 |
| | | Dark seed-coat[b] | 10/11 | 36/36 | 13/26 | 0/30 | 0/22 | 0/5 |
| | | Protein (%): | 22.4 | 22.3 | 19.2 | 18.1 | 17.4 | 17.3 |
| | | Starch (%): | 43.7 | 47.5 | 49.7 | 50.6 | 51.5 | 53.2 |
| | ICC 5912 | Protein (%): | 25.8 | | | | | |
| | | Starch (%): | 48.6 | | | | | |
| Slightly angular | 23 RILs | Pink flowers[a] | | 1/1 | 6/6 | 9/9 | 5/5 | 1/1 |
| | | Protein (%): | | 16.5 | 17.5 | 18.1 | 18.2 | 18.0 |
| | | Starch (%): | | 47.6 | 47.5 | 48.6 | 48.3 | 47.0 |
| | ICC 995 | Protein (%): | | | | | 17.4 | |
| | | Starch (%): | | | | | 49.5 | |
| Angular | 30 RILs | Pink flowers[a] | | | 7/7 | 18/18 | 4/4 | 2/2 |
| | | Protein (%): | | | 17.6 | 18.2 | 20.9 | 18.3 |
| | | Starch (%): | | | 44.7 | 44.6 | 47.4 | 48.3 |

[a]Number of lines producing pink flowers/total number of lines

[b]Seed-coat color similar to ICC 5912

## QTL analysis of seed traits

The initial QTL mapping on collected trait data was done using the QTLNetwork software ver. 2.1 (Yang et al. 2008) and employing the 11 larger linkage groups (Fig. 1). This analysis showed the genotypic variance for the population was largest for 100-SW (0.35) and PRO (0.33), but smaller for STA (0.18) (Table S5). Similarly, environmental influences on 100-SW (0.42) and PRO (0.42) were larger than on STA (0.25). The variances for genotype and environment interactions were small for all seed traits (0.02 to 0.06). Trait data from each trial (ICR, GH, BIG, ABE) were analyzed individually and as a multi-environment analysis (Table 4). From both analyses, the significant ($P < 0.05$) additive QTL identified for all traits were located to five regions (q-1.1, q-3.2, q-4.2, q-6.2 and q-8.1) on the developed maps (Table 4; Fig. 1). Four of these regions (q-1.1, q-3.2, q-4.2, and q-8.1) affected 100-SW, with qSW-1.1 and qSW-3.2 significant in all four environments. The qSW-3.2 locus explained a large part (33.9 to 71.1%) of the total phenotypic variance for 100-SW in the different trials. Weaker effects were demonstrated by qSW-1.1 (3.5 to 8.9% explained) and qSW-4.2; the latter was only significant in the ABE (13.4% explained) and ICR (7.3% explained) trials. In addition, a very weak QTL for 100-SW on LG 8 (qSW-8.1; 0.4% explained) was only indicated by the multi-environment analysis. For all 100-SW QTL, high values were consistently contributed by ICC 995 alleles and the total effect from the additive QTL explained 73.1 to 76.2% of the phenotypic variance for 100-SW in three trials. In the BIG trial producing the lowest 100-SW values (Fig. 2), the impact from identified QTL on 100-SW was less (37.6% explained).

The single and multi-environment analyses showed q-1.1 and q-3.2 affected PRO and STA in addition to 100-SW (Table 4; Fig. 2). The dominant QTL for PRO was located to q-3.2, which was significant in all four individual trials (34.8 to 57.0% explained). Lower impact on PRO was noted for qPRO-1.1, which was significant in three of the four trials (ABE, GH, ICR; 1.8 to 11.8% explained). In addition, two weak QTL, qPRO-3.1 and qPRO-6.2 explaining 0.5 and 1.3%, respectively, were suggested by the multi-environment analysis. For all four QTL for PRO, high protein content in seeds was favored by ICC 5912 alleles. In contrast to PRO and 100-SW, the STA values were more impacted by q-1.1 (14.3–31.6% explained) than by q-3.2 (9.2% explained in GH only). High STA values at qSTA-1.1 ($F = 53.6$; 18.4%) and qSTA-3.2 ($F = 13.6$; 3.9%) were favored by ICC 995 alleles. In the single-environment analyses, qSTA-1.1 was a stronger and more consistent locus than qSTA-3.2.

As observed for 100-SW, PRO and STA, the q-1.1 and q-3.2 regions also affected SS. The effects from qSS-ICR-1.1 (9.0%) and qSS-ICR-3.2 (14.6%) explained together 23.6% of the total variation in SS scores (Table 4; Fig. 1). Angular seed shape was promoted by ICC 5912 alleles at qSS-1.1 and ICC 995 alleles at qSS-3.2. Similar to QTL for 100-SW, PRO and STA, the peak for qSS-3.1 was mapped close to the CaGM13632 marker at Ca3:23.25 Mb (Table 4; Fig. 1) that was tightly associated with flower color (95.2% explained; Table 4). The strong co-localization of several traits observed for the RIL population is common in crops

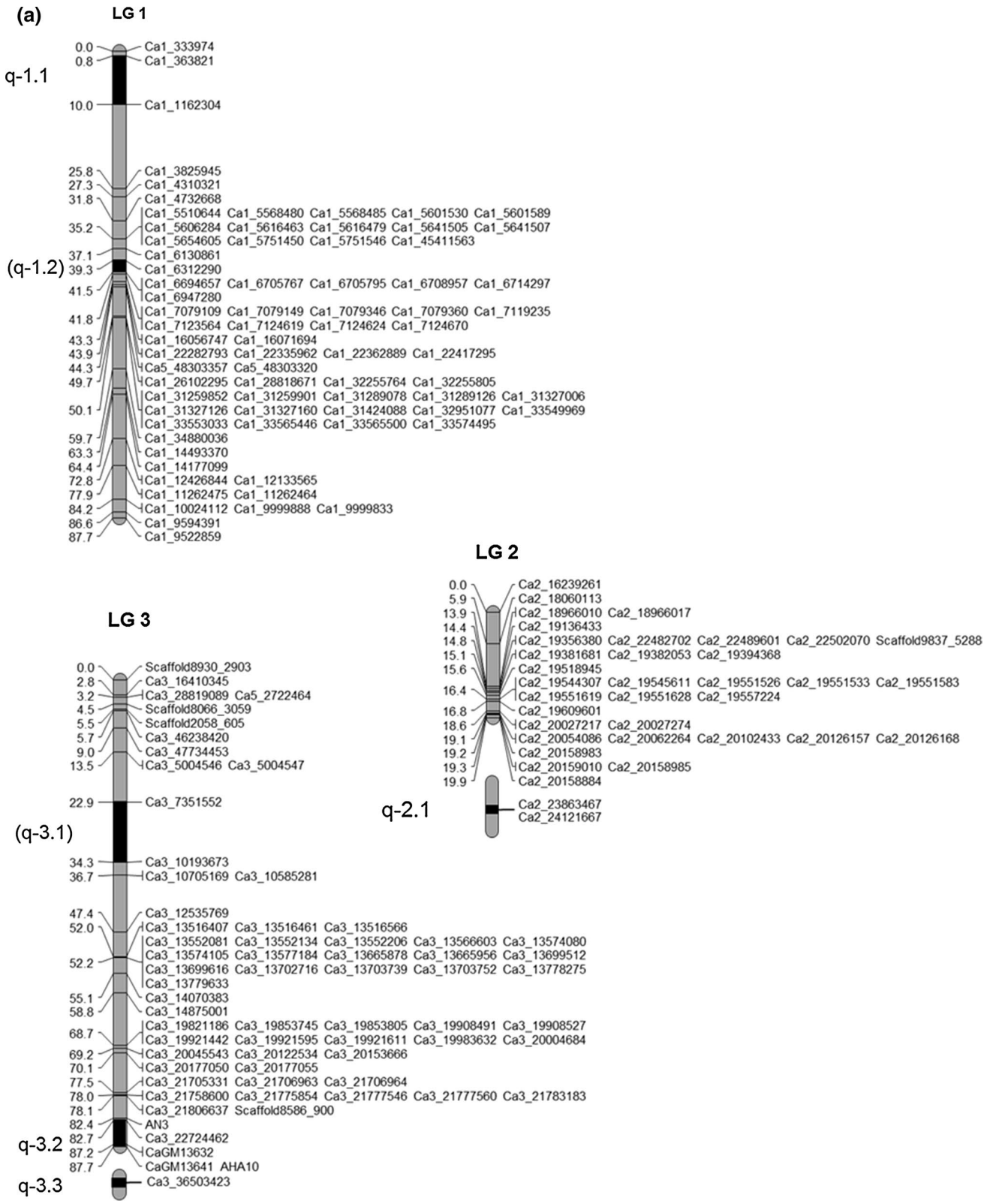


**Fig. 1** Genetic linkage map of chickpea genome derived from genotyping ICC 995 × ICC 5912 RIL population. Genetic distances are given in Kosambi centiMorgan (cM) units. Physical location of markers corresponds to chickpea genome assembly uwa v3.0 (Edwards 2016). The confidence intervals for additive and epistatic QTL identified for ICC 995 × ICC 5912 population (Tables 4, 5) are indicated by black areas on the chromosome bars. QTL within brackets are considered tentative

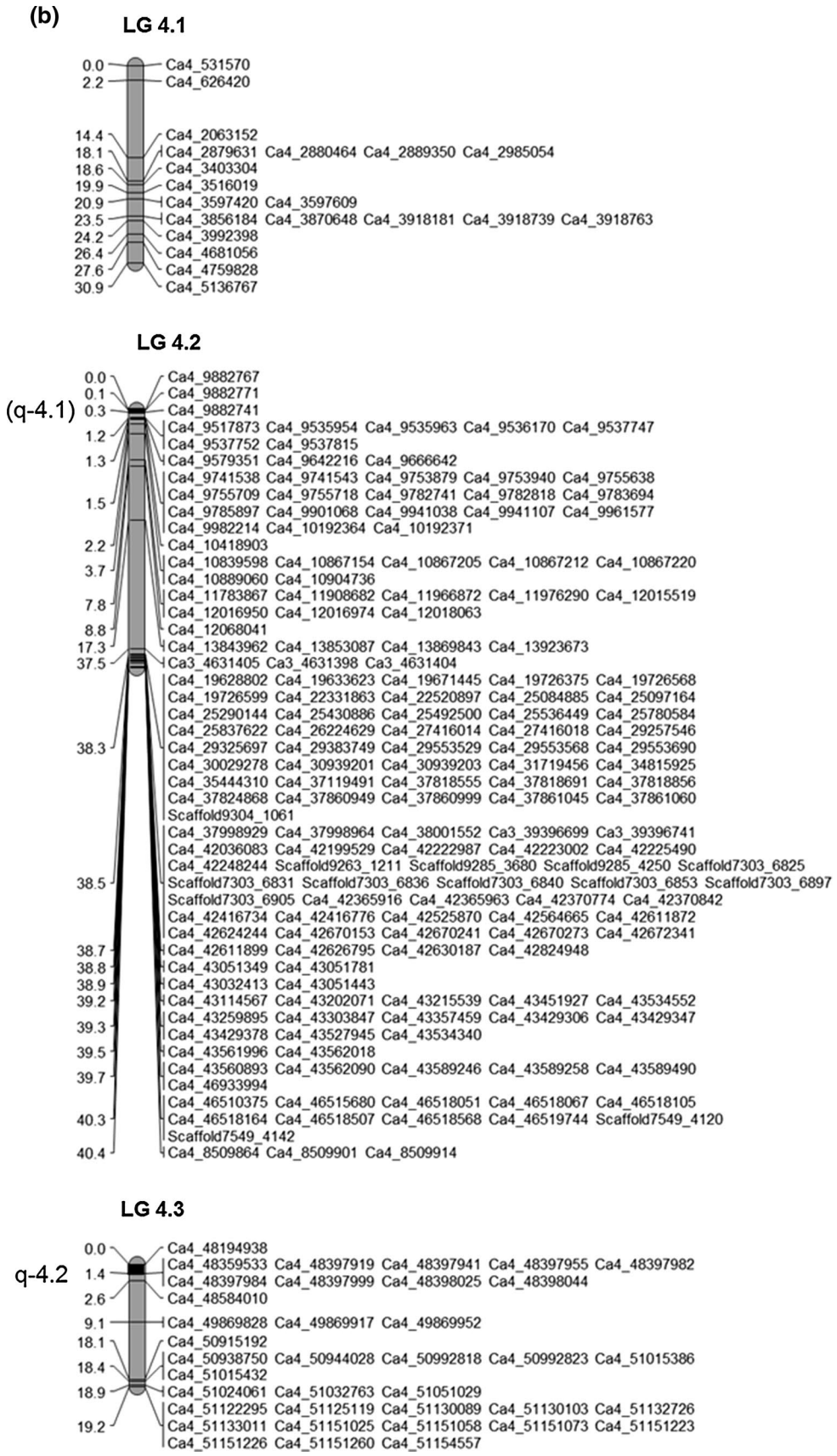


**Fig. 1** (continued)

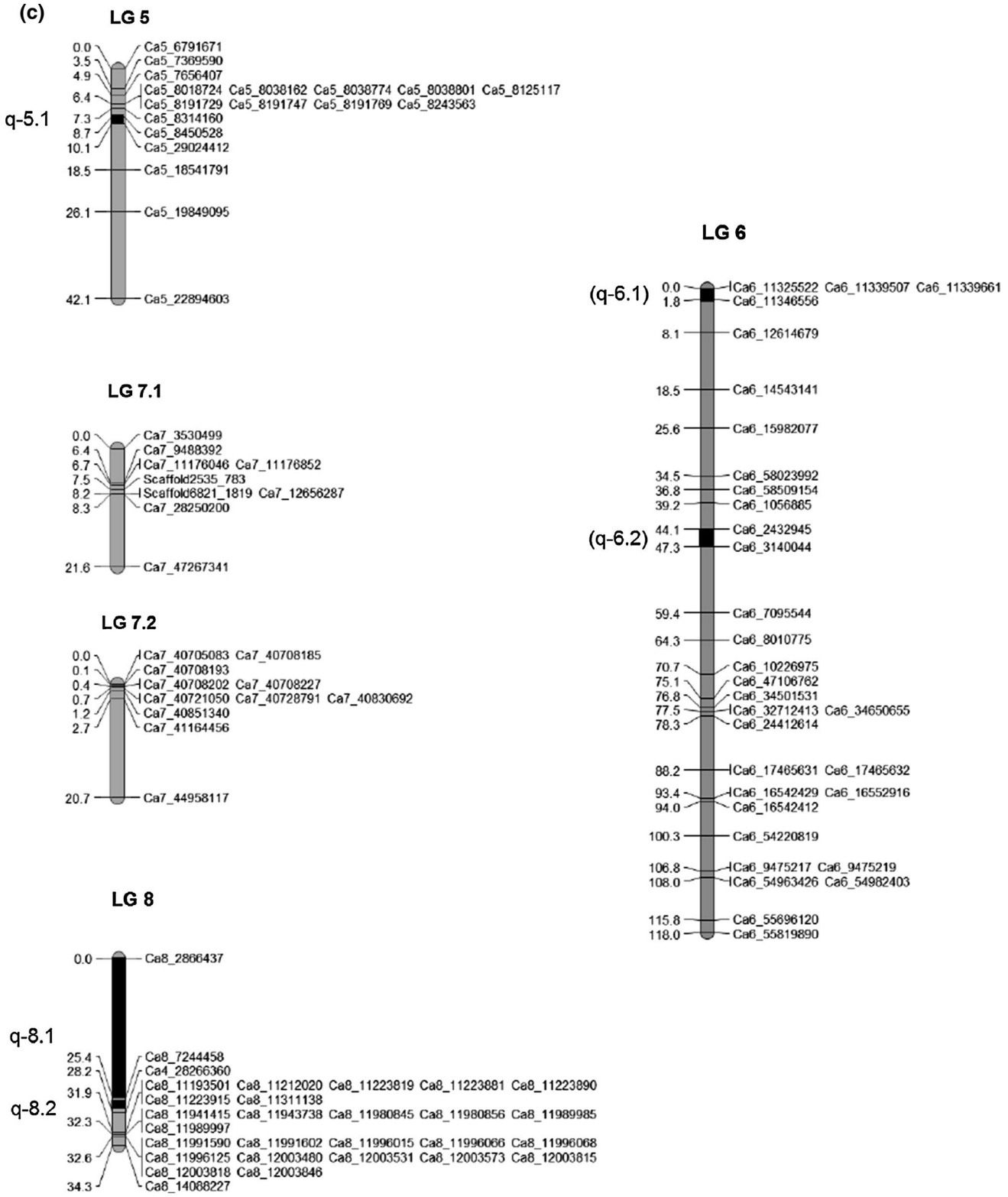


**Fig. 1** (continued)

**Table 4** Additive QTL for seed traits identified for ICC 995 × ICC 5912 population

| Trait | Crit. *F* value[a] | Trial[b] | Linkage group | QTL[c] | Marker interval | Range (cM) | Peak (cM) | *F* value | Additive effect[d] | Explained (%)[e] | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 100-SW | 12.1 | ABE | LG 1 | qSW-ABE-1.1* | Ca1:363821–Ca1:1162304 | 0.8–7.8 | 2.8 | 27.8 | 0.80 | 3.5 | 74.2 |
| | | | LG 3 | qSW-ABE-3.2* | Ca3:22724462–CaGM13632 | 84.7–87.2 | 86.7 | 195.7 | 3.43 | 57.3 | |
| | | | LG 4.3 | qSW-ABE-4.2 | Ca4:48194938–Ca4:48359533 | 0.0–1.4 | 1.0 | 21.4 | 1.07 | 13.4 | |
| | 11.8 | BIG | LG 1 | qSW-BIG-1.1 | Ca1:363821–Ca1:1162304 | 0.8–5.8 | 1.8 | 13.8 | 0.71 | 3.7 | 37.6 |
| | | | LG 3 | qSW-BIG-3.2 | Ca3:22724462–CaGM13632 | 83.7–87.2 | 86.7 | 72.8 | 1.92 | 33.9 | |
| | 12.4 | GH | LG 1 | qSW-GH-1.1* | Ca1:363821–Ca1:1162304 | 1.8–7.8 | 4.8 | 77.0 | 1.99 | 5.1 | 76.2 |
| | | | LG 3 | qSW-GH-3.2* | Ca3:22724462–CaGM13632 | 85.7–87.2 | 86.7 | 303.3 | 6.39 | 71.1 | |
| | 11.6 | ICR | LG 1 | qSW-ICR-1.1 | Ca1:363821–Ca1:1162304 | 0.8–6.8 | 3.8 | 75.3 | 1.41 | 8.9 | 73.1 |
| | | | LG 3 | qSW-ICR-3.2 | Ca3:22724462–CaGM13632 | 85.7–87.2 | 86.7 | 176.2 | 3.12 | 56.9 | |
| | | | LG 4.2 | qSW-ICR-4.2 | Ca4:48194938–Ca4:48359533 | 0.0–1.4 | 1.0 | 12.3 | 0.56 | 7.3 | |
| | 5.0 | Multi | LG 1 | qSW-1.1* | Ca1:363821–Ca1:1162304 | 1.8–5.8 | 3.8 | 51.6 | 1.34 | 4.5 | 57.4 |
| | | | LG 3 | qSW-3.2* | Ca3:22724462–CaGM13632 | 85.7–87.2 | 86.7 | 232.3 | 3.75 | 52.5 | |
| | | | LG 8 | qSW-8.1* | Ca8:2866437–Ca8:7244458 | 6.0–25.4 | 16.0 | 6.1 | 0.70 | 0.4 | |
| PRO | 12.1 | ABE | LG 1 | qPRO-ABE-1.1 | Ca1:363821–Ca1:1162304 | 0.8–9.8 | 2.8 | 14.8 | − 0.44 | 1.8 | 58.8 |
| | | | LG 3 | qPRO-ABE-3.2 | Ca3:22724462–CaGM13632 | 84.7–87.2 | 86.7 | 184.2 | − 1.82 | 57.0 | |
| | 12.3 | BIG | LG 3 | qPRO-BIG-3.2* | Ca3:22724462–CaGM13632 | 85.7–87.2 | 86.7 | 81.5 | − 1.47 | 34.8 | 39.1 |
| | | | LG 6 | qPRO-BIG_6.2* | Ca6:2432945–Ca6:3140044 | 44.1–47.3 | 45.1 | 16.5 | − 0.65 | 4.3 | |
| | 11.9 | GH | LG 1 | qPRO-GH-1.1 | Ca1:363821–Ca1:1162304 | 0.8–8.8 | 4.8 | 42.9 | − 0.77 | 8.7 | 48.4 |
| | | | LG 3 | qPRO-GH-3.2 | Ca3:22724462–CaGM13632 | 85.7–87.2 | 86.7 | 99.0 | − 1.44 | 39.7 | |
| | 12.2 | ICR | LG 1 | qPRO-ICR-1.1 | Ca1:363821–Ca1:1162304 | 0.8–6.8 | 2.8 | 71.2 | − 1.07 | 11.8 | 59.3 |
| | | | LG 3 | qPRO-ICR-3.2 | Ca3:22724462–CaGM13632 | 85.7–87.2 | 86.7 | 90.9 | − 1.97 | 47.5 | |
| | 5.1 | Multi | LG 1 | qPRO-1.1* | Ca1:363821–Ca1:1162304 | 2.8–7.8 | 4.8 | 25.7 | − 0.67 | 4.7 | 50.8 |
| | | | LG 3 | qPRO-3.1 | Ca3:7351552–Ca3:10193673 | 22.9–34.3 | 29.9 | 6.1 | − 0.21 | 0.5 | |
| | | | LG 3 | qPRO-3.2* | Ca3:22724462–CaGM13632 | 85.7–87.2 | 86.7 | 124.4 | − 0.71 | 44.3 | |
| | | | LG 6 | qPRO-6.2* | Ca6:2432945–Ca6:3140044 | 44.1–47.3 | 46.1 | 5.7 | − 0.38 | 1.3 | |
| STA | 11.6 | ABE | LG 1 | qSTA-ABE-1.1 | Ca1:363821–Ca1:1162304 | 0.8–6.8 | 1.8 | 33.9 | 1.66 | 15.8 | 15.8 |
| | 11.7 | BIG | LG 1 | qSTA-BIG-1.1 | Ca1:363821–Ca1:1162304 | 0.8–4.8 | 1.8 | 38.0 | 2.67 | 17.4 | 17.4 |
| | 12.4 | GH | LG 1 | qSTA-GH-1.1 | Ca1:363821–Ca1:1162304 | 1.8–6.8 | 3.8 | 108.0 | 4.16 | 31.6 | 40.8 |
| | | | LG 3 | qSTA-GH-3.2 | Ca3:22724462–CaGM13632 | 84.7–87.2 | 86.7 | 34.0 | 2.41 | 9.2 | |
| | 12.0 | ICR | LG 1 | qSTA-ICR-1.1 | Ca1:363821–Ca1:1162304 | 0.8–9.8 | 4.8 | 30.3 | 1.88 | 14.3 | 14.3 |
| | 4.8 | Multi | LG 1 | qSTA-1.1* | Ca1:363821–Ca1:1162304 | 0.8–4.8 | 2.8 | 56.8 | 2.69 | 19.0 | 23.2 |
| | | | LG 3 | qSTA-3.2* | Ca3:22724462–CaGM13632 | 85.7–87.2 | 86.7 | 15.1 | 1.32 | 4.2 | |

**Table 4** (continued)

| Trait | Crit. *F* value[a] | Trial[b] | Linkage group | QTL[c] | Marker interval | Range (cM) | Peak (cM) | *F* value | Additive effect[d] | Explained (%)[e] | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SS | 12.1 | ICR | LG 1 | qSS-ICR-1.1* | Ca1:363821–Ca1:1162304 | 0.8–10.0 | 3.8 | 17.4 | −0.13 | 9.0 | 23.6 |
| | | | LG 3 | qSS-ICR-3.2* | Ca3:22724462–CaGM13632 | 82.7–87.2 | 84.7 | 27.9 | 0.31 | 14.6 | |
| COL | 5.5 | ICR | LG 3 | qSS-ICR-3.2 | Ca3:22724462–CaGM13632 | 85.7–87.2 | 86.7 | 894.4 | 0.50 | 95.2 | 95.2 |

[a]Determined by 1000 permutation test

[b]Multi = data from multi-environment analyses

[c]Asterisk = interacting QTL (see Table 5)

[d]Positive value indicates higher trait value contributed by ICC 995 alleles, and negative value indicates higher values contributed by ICC 5912 alleles

[e]Additive effect of individual QTL and total additive effect are presented

and may be due to either pleiotropic effects of a major gene or tight linkage of several genes affecting the different traits.

## Identification of QTL interactions

In addition to the additive QTL, the QTLNetwork analysis identified six pairs of interactions from the single-environment analyses and five pairs of epistatic interactions from the multi-environment analysis (Table S6). Several of these QTL interactions involved the same loci, reducing the number of interactions to six. Four additional QTL regions (q-1.2, q-4.1, q-6.1 and q-8.2) without additive effects were predicted to be involved in epistatic and/or environmental interactions. The major additive QTL at q-3.2 showed epistatic interactions with several loci that affected 100-SW, PRO, STA and SS, respectively, but the additive and environmental effects were small (0.5–4.3%; Table 5). Parental alleles at q-3.2 and q-1.1 slightly increased 100-SW (0.5% explained), whereas alleles of different parental origins at the two loci increased STA values (1.8% explained) and promoted production of angular-shaped seeds (4.3% explained). The q-3.2 region was also suggested to interact with the weak additive loci q-8.1 and q-6.2 affecting 100-SW and PRO values, respectively. Alleles of different parental origins at q-3.2 and q-8.1 increased 100-SW values, whereas alleles from the same parent at q-3.2 and q-6.2 increased PRO values, but both interactions had small effects (0.5 and 1.2% explained, respectively).

One of the QTL interactions involved q-4.2 and q-1.2 on LG 1 (Table 5). This interaction was only revealed in the ABE trial, where alleles of different parental origins increased PRO values (2.7% explained). A locus on LG 8, q-8.2, was predicted to be involved in two interactions. Both interactions involved alleles of different parental origins, where the q-8.2/q-4.1 interaction weakly affected PRO (1.3%), and q-8.2/q-6.1 epistasis promoted development of angular-shaped seeds (5.8% explained). Upon combining all the effects of additive QTL, epistasis and environmental interactions from the multi-environment analyses, the identified QTL for 100-SW were estimated to explain 58.4% of trait variation as determined by the multi-environment analysis. Similarly, the total QTL effects for PRO were also large (53.3%), whereas smaller effects were identified for STA (25.0%). For SS, the total additive QTL effects determined from the ICR trial explained 33.7% of the total phenotypic variance.

## Association mapping indicates additional QTL

Due to low marker coverage of the genome, it was possible QTL located in regions with low marker density went undetected in the QTL analysis. Thus, all 3588 filtered markers and four additional markers developed (Table S2)

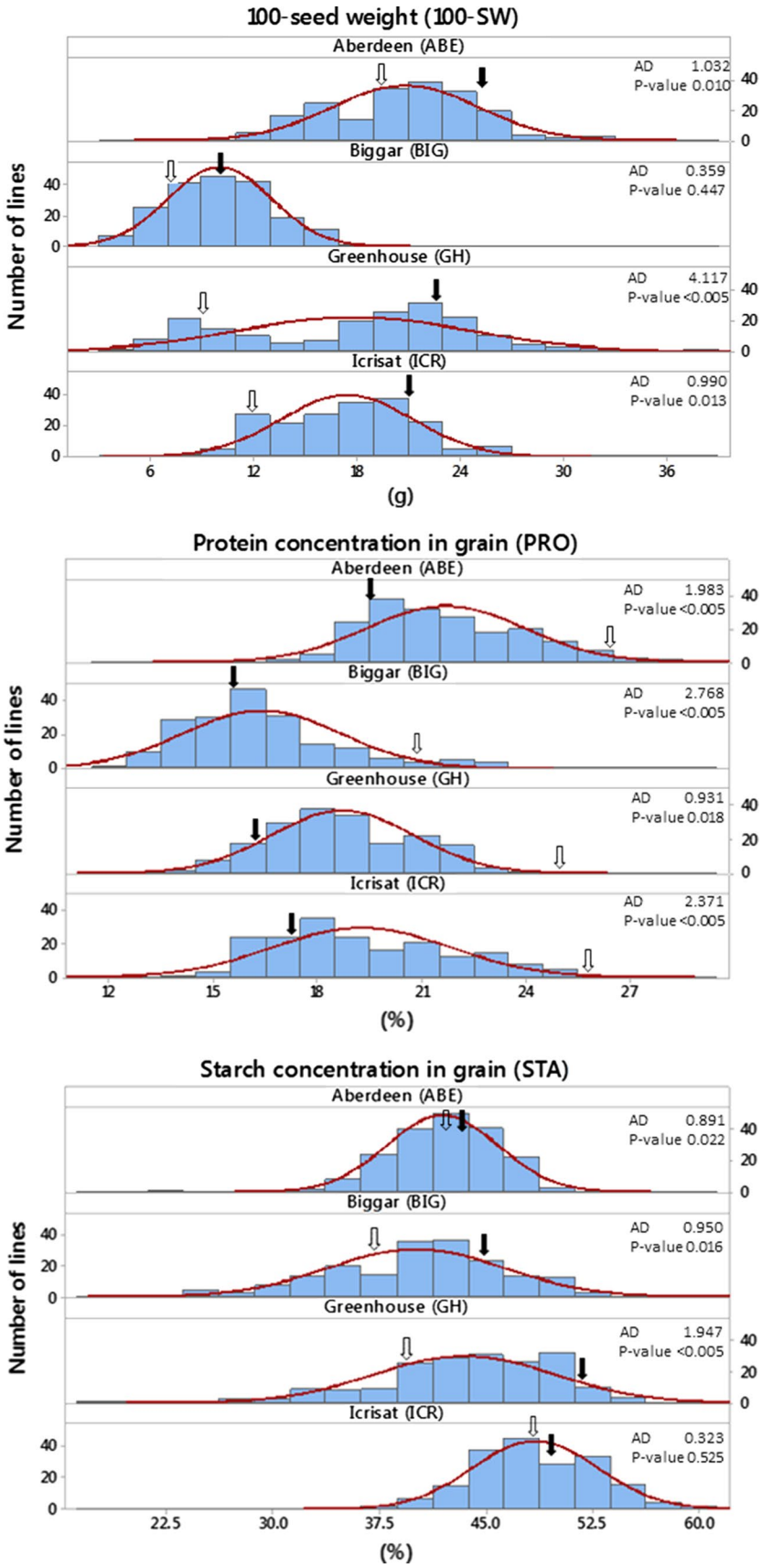


**Fig. 2** Frequency distributions of seed trait data for ICC 995 × ICC 5912 RIL population grown in four environments. Trait values for 100-seed weight (100-SW), protein (PRO) and starch (STA) concentrations in seeds harvested from trial conducted in ICRISAT, India (ICR), Biggar, Canada (BIG), greenhouse, Saskatoon, Canada (GH) and Aberdeen, Canada (ABE) are shown. Mean values for the parental genotypes ICC 995 and ICC 5912 are indicated by vertical arrows. AD scores obtained from normality test by the Anderson–Darling goodness of fit test are shown where scores with $P$ values $\leq 0.05$ were considered as significant for non-normal distribution

**Table 5** Genetic markers located within genes at major QTL on LG 1 and LG 3

| Genetic marker | Location | Predicted effect of variation | Annotated gene |
|---|---|---|---|
| Ca1_333974 | Intron | – | Uncharacterized; probable magnesium transporter |
| Ca1_363821 | Exon | – | Predicted COBRA-like protein 6 |
| Ca1_1162304 | Exon | Thr to Lys change | Predicted ubiquitin-like-specific protease 2B |
| Ca3_18996437 | Intron | | Subtilisin-like protease SBT4.15 |
| Ca3_19821186 | Intron | | Predicted glutamate carboxypeptidase 2 |
| Ca3_19853745 | Exon | No aa change | Predicted calcium homeostasis endoplasmic reticulum protein |
| Ca3_19853805 | Exon | Ala to Leu change | Predicted calcium homeostasis endoplasmic reticulum protein |
| Ca3_19908491 | Intron | | Predicted sister-chromatid cohesion protein 3 |
| Ca3_19908527 | Intron | | Predicted sister-chromatid cohesion protein 3 |
| Ca3_19921595 | Intron | | Predicted NPL4-like protein 1 |
| Ca3_19921611 | Intron | | Predicted NPL4-like protein 1 |
| Ca3_20045543 | Exon | No aa change | Predicted uncharacterized protein, LOC101503823 |
| Ca3_20122534 | Exon | Pro to Ser change | Predicted autophagy-related protein 18a-like |
| Ca3_20177050 | Intron | | Predicted uncharacterized protein OsI_027940-like |
| Ca3_20177055 | Intron | | Predicted uncharacterized protein OsI_027940-like |
| Ca3_21283920 | 5′-UTR | TACCA repeat affected | NHL repeat-containing protein 2 |
| Ca3_21706963 | Intron | | Predicted sucrose nonfermenting 4-like protein |
| Ca3_21706964 | Intron | | Predicted sucrose nonfermenting 4-like protein |
| Ca3_21758600 | Intron | | Predicted probable Xaa-Pro aminopeptidase P |
| Ca3_21775854 | Exon | No aa change | Predicted kinesin-like protein KIFC3 |
| Ca3_21777546 | Intron | | Predicted kinesin-like protein KIFC3 |
| Ca3_21777560 | Intron | | Predicted kinesin-like protein KIFC3 |
| Ca3_21783183 | Intron | | Predicted kinesin-like protein KIFC3 |
| Ca3_21806637 | 5′-UTR | Alternative start codon | Protein YIP1 domain family member (YIPF) |
| Scaffold8586_900 | 5′-UTR | Alternative start codon | Protein YIP1 domain family member (YIPF) |
| AN3_22374920 | 5′-UTR | CT-repeat affected | GRF1-interacting factor 1-like; ANGUSTIFOLIA3 (AN3)-like |
| Ca3_22724462 | Exon | Arg to Cys change | MUG1-like protein |
| AHA10_23262576 | Exon | Truncated protein | Membrane ATPase; AHA10-like |
| CaGM13641_~23312509 | Intron | | Predicted exportin-T |
| Ca3_36503423 | Exon | Truncated protein | Predicted ABC transporter B family member 26, chloroplastic-like |

were tested for association with trait values. The genotype/phenotype association was carried out by the GLM model using the TASSEL 3.0 software (Glaubitz et al. 2014). A total of 58 markers (45 mapped; 10 unmapped) showed strong ($P < 10^{-5}$) association with flower color and/or one or more of the seed traits (Table S7). Consistent with the QTL analysis, markers positioned within q-1.1, q-3.2, and q-4.2 showed the strongest associations with trait values, whereas markers associated with q-6.2 and q-8.2 were of low significance ($P > 10^{-4}$, data not shown). For q-3.2, four new unmapped markers were added to the locus. Of these new markers, Ca3:18996437 showing heterozygosity in ICC 5912, affected flower color and multiple seed traits ($F_{max}$ for COL = 13.8). The other three markers, also associated with multiple traits ($F_{max}$ for COL = 67.1–70.5), were clustered within the Ca3:19921595–19921611 region and could involve a gene duplication as the polymorphic alleles showed heterozygosity in the ICC 5912 parent. At q-4.2, all markers associated with trait values were included on the developed genetic map for the region.

The remaining significant associations involving unmapped markers were weaker than those for q-1.1, q-3.2 and q-4.2, but nevertheless suggested the presence of additional minor QTL that were not identified by QTL mapping. One new QTL locus was indicated for the Ca2:23863467–24121667 region (q-2.1) that was not included on LG 2 map (Fig. 1). Two unmapped SNP in this region were associated with seed shape ($F = 17.2$ and 18.8; Table S7). Another suggested QTL (q-3.3) was located outside of the constructed LG 3 map, 13.2 Mb distal of q-3.2 (Fig. 1). The significant marker at this locus was Ca3:36503423, which affected 100-SW in two trials ($F_{max} = 23.6$). A minor QTL for seed shape ($F = 17.6$) was associated with Ca5:10302503 on LG 5 (q-5.1). In summary, the association mapping supported the existence of six additive QTL (q-1.1, q-2.1, q-3.2, q-3.3, q-4.2, and q-5.1)

affecting seed traits and flower color. STA was primarily affected by q-1.1, PRO by q-1.1 and q-3.2, 100-SW by q-3.2, q-3.3 and q-4.2, SS by q-2.1, q-3.2 and q-5.1, and COL by 3.2. The additional and weaker QTL suggested by one of the mapping strategies (q-1.2, q-4.1, q-6.1, q-6.2, q-8-2) should be regarded as tentative QTL.

### Genes associated with major QTL on LG 1 and LG 3

The major QTL for STA was q-1.1, restricted by flanking markers to the 0.36–1.16 Mb region on LG 1 (Table 4; Fig. 1). The locus encompasses 115 annotated genes on the ICC 4958 genome assembly uwa v3.0 (Edwards 2016), but only three SNPs were associated with the locus in this study (Fig. 1; Table 5). The most significant marker was Ca1:333974 positioned within an intron of a probable magnesium transporter gene. The two other markers coincided with a COBRA6-like and an ubiquitin 2B-like gene, respectively. For the ubiquitin 2B-like gene, the ICC 5912 allele caused Thr to Lys substitution when compared to the ICC 4958 reference and ICC 995 alleles.

The q-3.2 interval determined by QTL mapping was delineated by the Ca3:22724462 marker and the end of LG 3 (~Ca3:23312509), a region containing 74 annotated genes. The CaGM13632, AHA10 and CaGM13541 markers positioned at the distal end of this region had the largest effects on all trait values. For these markers, the trait association was highest for COL ($F = 1874$-Inf.) followed by 100-SW ($F = 56$–277), PRO ($F = 56$–136), SS ($F = 17$) and lowest for STA ($F \leq 11$) as determined by association mapping (Table S7). The AHA10 marker at the QTL peak corresponded to a P type $H^+$-ATPase 10/TRANSPARENT TESTA 13 gene (Table 5), which in Petunia encodes a tonoplast $H^+$ pump involved in the regulation of flower color (Baxter et al. 2005; Verweij et al. 2008). Located in the tonoplast, AHA10 promotes formation of an electrochemical gradient across the membrane and thereby supports transfer of anthocyanin pigments by multidrug and toxic compound extrusion (MATE) transporters into the vacuoles of petal cells (Faraco et al. 2014; Verweij et al. 2008). Due to vacuolar acidification promoted by AHA10 activity, the anthocyanin pigments adopt a reddish/pink color in the presence of certain metal ions and/or co-pigments. Absence of AHA10 activity results in higher pH in the vacuolar lumen and a shift toward blue color for the anthocyanin compounds. Polymorphism for AHA10 marker for the parental alleles was caused by an out-of-frame mutation within an *AHA10* exon for the ICC 5912 allele, which was predicted to encode a non-functional transporter. Consistent with AHA10 deficiency, ICC 5912 produced blue flowers, whereas ICC 995 with wild-type *AHA10* alleles produced pink flowers. Among the RIL, 59 out of 60 lines with homozygous ICC 5912 alleles for *AHA10* had blue flowers, whereas flower color was pink for all 126 lines with homozygous ICC 995 alleles for *AHA10*. Three RIL carrying heterozygous *AHA10* alleles produced pink flowers. A transcript produced in young flowers (TC03272), matching the annotated *AHA10* coding region, supported AHA10 activity in chickpea flowers. Thus, the variation for *AHA10* alleles was likely to underlie the blue/pink variation in flower color observed within the RIL population.

Studies in Arabidopsis show AHA10 is produced during seed development and promotes accumulation of the color-less proanthocyanidin molecules into seed coats (Baxter et al. 2005). However, *AHA10* mutants have not been reported to produce seeds with altered size, shape or composition, and therefore, it was more likely that gene(s) located close to *AHA10* caused variations in seed traits at q-3.2. Within the confidence interval for q-3.2 identified by QTL mapping lies the Ca3:22724462 marker (Table 4) that is positioned within a region with no annotated genes on the ICC 4958 uwa v3.0 genome. However, a chickpea transcript (TC07742) showed good sequence alignment (2193/2194 nt) to the Ca3:22722713–22725182 region carrying the Ca3:22724462 marker. The protein predicted from the TC07742 sequence was 597 amino acids long and shared up to 96% sequence identity with numerous plant proteins annotated as uncharacterized or mutator-like transposase-like. However, among the best matches among Arabidopsis proteins were members belonging to a small family of "domesticated" and functional proteins denoted MUSTANG/MUG (Cowan et al. 2005; Joly-Lopez et al. 2012). Of the eight MUG proteins known in Arabidopsis, the TO07742-encoded polypeptide showed sequence highest identity to AtMUG1 (62%) and contained all the highly conserved residues for MUG proteins characterized from various plant species. A change from a conserved arginine to a cysteine residue for the MUG-like protein encoded from ICC 5912 allele was the result of polymorphism at the Ca3:22724462 marker. Thus, this sequence change may alter the function of the MUG-like protein produced by ICC 4912.

Considering the possibility that the strong q-3.2 effect is derived from several genes, genes located upstream or downstream of the identified QTL interval could also contribute to trait variation observed at q-3.2. Many markers located within the Ca3:21806637–22724462 region showed polymorphism that could affect gene expression or activity of encoded protein (Table 5). For example, marker Ca3:21806637 was positioned within sequence encoding the leader sequence of an YPT/RAB GTPase Interacting Protein (YIP1) domain transcript (TC03646) produced in young pods. The ATGTG/ATG allele variation for this marker could possibly involve an alternative translational start codon for the YIP1 transcript variant XP_004517227. Three markers for the CCA$(TACCA)_5$T sequence identified by association mapping were positioned within the leader sequence of

an NHL repeat-containing protein 2 gene (Tables S7 and 5) active in etiolated seedlings (XM_012714094.1) and young pods (TC22777). Differences in the number of $(TACCA)_{5–6}$ repeats, with possible regulatory function, were observed between the parental *YIP1* alleles. Allele variation within the 5′-UTR of a growth-regulating factor (GRF)-interacting factor 1 (GIF1) gene was also identified for parental alleles. The difference involved a $(CT)_{12–15}$ repeat, which could serve as a regulatory GAGA element. The number of CT-repeat was 12 for ICC 995 and CDC Frontier reference alleles, 15 for ICC 5912 allele and 13 for the ICC 4958 reference allele. For the region located distal to q-3.2, the Ca3:36503423 marker at q-3.3 was located within an exon of an ABCB26 gene, for which a frameshift caused by the ICC 5912 allele predicted a truncated ABCB26 protein for this parent (Table 5).

## Discussion

### Low genetic diversity within RIL population

In this study, the marker coverage for several of the assembled linkage groups was low (Fig. 1) as fewer-than-expected GBS markers were identified from the DNA sequencing data. This was mainly due to very low marker polymorphism between parental alleles, which resulted in a low number of usable markers for linkage mapping and poor marker coverage for several chromosomes. However, a relatively high fraction of the identified GBS markers were positioned within genes (76.8%), which was desired as it increased the chance of identifying the cause of trait variation at QTL. To compensate for the low-density genetic maps, the QTL mapping strategy involved both QTL mapping and a single marker/phenotype association study. Three major additive QTL (q-1.1, q-3.2, and q-4.2) were confirmed by both strategies (Tables 4 and S7), and three more additive QTL of lower significance were suggested by association mapping (q-2.1, q-3.3, q-5.1; Fig. 1). Five additional QTL were identified in the study (indicated within brackets in Fig. 1), but they were weak and will need further validation.

### Seed weight and protein content in seeds were mainly controlled by genetic factors

Only two regions (q-1.1 and q-3.2) had major effects on all seed traits (Table 4; Fig. 1). The power of these QTL caused several of the trait distributions curves for the RIL population to be multi-modal (Fig. 2) instead of showing a normal distribution that is expected for a quantitative trait. The low seed weights recorded for RIL population and parents in the BIG trial showed the environmental conditions at this site was especially poor for chickpea production (Fig. 2; Table S4). Abiotic and biotic stresses during seed development can cause precocious endosperm cellularization, which negatively affects endosperm to embryo signaling and growth of the embryo (Folsom et al. 2014; Lafon-Placette and Köhler 2014). Also, the photosynthesis reactions in green seed coats in legumes can be impacted by stress causing effects on energy and oxygen levels within the seed and resulting in reduced accumulation of storage products. These observations have led to the conclusion that low ATP levels stimulate protein accumulation, whereas high ATP levels improve seed filling by stimulation of starch biosynthesis (Borisjuk et al. 2003). Thus, the seed filling efficiency in the different trials may have been affected by stresses influencing the energy status within the developing seeds.

### The role of seed coat development for seed shape

The ICC 995 alleles at q-1.1 and q-3.2 promoted development of slightly angular light-colored seeds with high STA and 100-SW values. Small round dark-colored seeds with high PRO values with tightly attached seed coats were favored by ICC 5912 alleles at q-3.2 only. Whether any of the two main QTL for SS corresponds to the previously described *Rd* locus conferring round or angular seed shape (Knights et al. 2011) is not known. The QTL at q-3.2 had the largest impact on SS (Table 4), and we speculate that maternal ICC 5912 alleles at this locus repress cell division and/or cell elongation in integument cells resulting in development of small seed coat. These under-developed seed coats put a physical constraint on endosperm and embryo growth, and seeds will be completely filled resulting in a round shape even when flux of nutrients to embryo is low. This was consistent with all RIL with ICC 5912 alleles at q-3.2 produced small round seeds irrespective of other alleles on the genome. Furthermore, we speculate angular-shaped seeds promoted by ICC 995 alleles at q-3.2 are developed with larger seed coats and only become fully filled and round when optimal transport of sugars and amino acids to the growing embryo and cotyledons is promoted by ICC 995 alleles at q-1.1. Angular (indented) seeds, a phenotype not seen for the parents and produced by RIL carrying ICC 5912 alleles at q-1.1 and ICC 995 alleles at q-3.2, resulted from incomplete filling of seed with large seed coat. As the growth of endosperm and integuments largely determine seed weight in chickpea (Garcia et al. 2005; Hehenberger et al. 2012; Weber et al. 2005), the variation in seed properties observed in the RIL population was expected to be mainly caused by the higher dose of maternal alleles active during the early stages of seed development.

## QTL on LG 1 (q-1.1) affected the efficiency of seed filling

The ICC 5912 alleles at q-1.1 had a negative influence on seed filling as RILs carrying these alleles showed lower starch accumulation in seeds and lower 100-SW. As a consequence of low starch accumulation, the protein concentration in seeds generally shows an increase (Hurkman and Wood 2011) as observed for RIL population studied here. The ICC 5912 line is reported to be relatively heat sensitive when compared to other chickpea accessions and shows a reduction in stomata conductance, invertase, Rubisco, sucrose phosphate synthase and sucrose synthase activities when heat-stressed (Kaushal et al. 2013). As activities associated with carbohydrate biosynthesis are highly affected by heat stress in ICC 5912, we propose gene(s) causing variation at q-1.1 have a role in this sensitivity.

Several of the early mapping studies of seed weight in chickpea identified QTL for the trait on LG 1 (Abbo et al. 2005; Cobos et al. 2009; Hossain et al. 2010), but due to lack of common markers, it is difficult to determine if these first reported QTL correspond to q-1.1 identified here. Lately, QTL with higher precision have been identified by employing the NGS-based whole-genome QTL-sequencing strategy for genotyping. This allowed one of the 100-SW QTL identified in an ICC 184 × ICC 15061 population to be associated with six annotated genes on LG 1 located at Ca1:0.52 Mb (Das et al. 2015), which lies within the confidence interval for q-1.1 reported here (Table 4). Based on expression profiles, a COP9 signalosome complex subunit 8 (CSN8) gene was suggested to cause seed weight variation in the ICC 184 × ICC 15061 population (Das et al. 2015). Approximately in the same region (Ca1:0.69 Mb), an ATP-dependent RNA helicase gene was associated with seed protein content in chickpea accessions (Upadhyaya et al. 2016b). The few markers mapped within q-1.1 did not allow us to pinpoint if q-1.1 coincided with CSN8 or the RNA helicase gene. However, the strongest association in our study was seen for the most proximal marker at Ca1:333974 mapped to a putative magnesium transporter gene (Tables 5, S7), suggesting q-1.1 did not coincide with QTL reported in this region on LG 1 in earlier studies.

## QTL on LG 3 (q-3.2) had a central role in the regulation of flower color, seed composition and seed morphology

All QTL peaks under q-3.2 mapped close to the CaGM13632, AHA10, and CaGM13641markers, for which homozygous ICC 5912 alleles showed very high association with blue flowers and production of small round seeds with high protein concentration (Table 4). The AHA10 gene at q-3.2 was considered to be a prime candidate gene for causing variation in flower color. In seeds, AHA10 has a role in the vacuolar import of proanthocyanins derived from the flavonoid pathway and affecting seed coat color when oxidized to condensed tannins (Baxter et al. 2005; Lepiniec et al. 2006). Although there are some suggestions that flavonoids influence endosperm cellularization and seed size (Doughty et al. 2014), there are no studies showing AHA10 mutants have a drastic effect on seed growth. For the RIL population investigated here, the most likely explanation for the tight linkage between blue flowers and small seeds, as observed in previous studies (Gaur et al. 2016), is the presence of closely linked genes affecting different traits. In a previous study of a RIL population derived from a *Cicer arietinum* ICC 4958 × *Cicer reticulatum* wild ICC 17160 cross, a QTL for 100-SW and number of pods per plant was mapped to LG 3 (Saxena et al. 2014). The locus identified on CDC Frontier v1 genome corresponds to Ca3:22271076–23242491 region on ICC 4958 v3 genome, thus largely coincides with the q-3.2 confidence interval in this study (Ca3:22724462–23312509). In the Saxena et al. (2014) study, four markers showing highest association with the traits were located within an NGA1 gene encoding a B3-domain transcription factor. However, the allele differences observed for *NGA1* were not present for ICC 995 and ICC 5912 parents (data not shown). In our analysis, the MUG-like gene appeared a more promising candidate gene for causing variation in seed traits at q-3.2. Disruption of MUG proteins is known to produce many different phenotypes, but generally results in lower fitness for the plant (Joly-Lopez et al. 2012). The mutation identified for the ABCB26 gene at q-3.3, which could contribute to the q-3.2 effect, is also interesting for further studies. This gene was in a recent study associated with grain yield in a GWAS study of chickpea accessions (Li et al. 2018). Several of the ABCB transporters have a role in auxin transport, but the exact function for the ABCB26 transporters in plants has not been determined (Lefèvre and Boutry 2018).

**Author contribution statement** RNC, MB, and PMG conceived and planned the project. PMG developed the RIL population and performed field experiments at ICRISAT. CI analyzed the GBS data and identified the SNPs in the population. RW did the phenotyping work. RW, MPG and MB carried out the field/greenhouse experiments in Canada, analyzed the data, did the QTL mapping and prepared the manuscript. All authors contributed to the review of the manuscript.

**Acknowledgements** This work was financially supported by Canada Research Chairs Program, Natural Sciences and Engineering Research Council, and Agriculture and Agri-Food Canada Internationalization program. The core research grant of International Crops Research Institute for Semi-Arid Tropics (ICRISAT, Patancheru, India) is

acknowledged for the development of the chickpea RIL population and field trial at ICRISAT. RW is a grateful recipient of the China Scholarship Council fellowship for Ph.D. We are very grateful to Mr. John Bennet (Biggar Saskatchewan) and Mr. Jeff Sopatyk (Aberdeen, Saskatchewan) who provided land for the field trials.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interests.

## References

Abbo S, Molina C, Jungmann R, Grusak MA, Berkovitch Z, Reifen R, Kahl G, Winter P, Reifen R (2005) Quantitative trait loci governing carotenoid concentration and weight in seeds of chickpea (*Cicer arietinum* L.). Theor Appl Genet 111:185–195. https://doi.org/10.1007/s00122-005-1930-y

Ainsworth EA, Bush DR (2011) Carbohydrate export from the leaf: a highly regulated process and target to enhance photosynthesis and productivity. Plant Physiol 155:64–69. https://doi.org/10.1104/pp.110.167684

Alonso-Blanco C, Blankestijn-de Vries H, Hanhart CJ, Koornneef M (1999) Natural allelic variation at seed size loci in relation to other life history traits of *Arabidopsis thaliana*. Proc Natl Acad Sci 96:4710–4717. https://doi.org/10.1073/pnas.96.8.4710

Bajaj D, Das S, Upadhyaya HD, Ranjan R, Badoni S, Kumar V, Tripathi S, Gowda CLL, Sharma S, Singh S, Tyagi AK, Parida SK (2015) A genome-wide combinatorial strategy dissects complex genetic architecture of seed coat color in chickpea. Front Plant Sci 6:979. https://doi.org/10.3389/fpls.2015.00979

Baxter IR, Young JC, Armstrong G, Foster N, Bogenschutz N, Cordova T, Peer WA, Hazen SP, Murphy AS, Harper JF (2005) A plasma membrane H + -ATPase is required for the formation of proanthocyanidins in the seed coat endothelium of *Arabidopsis thaliana*. Proc Natl Acad Sci 102:2649–2654. https://doi.org/10.1073/pnas.0406377102

Berger F, Chaudhury A (2009) Parental memories shape seeds. Trends Plant Sci 14:550–556. https://doi.org/10.1016/j.tplants.2009.08.003

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Borisjuk L, Rolletschek H, Walenta S, Panitz R, Wobus U, Weber H (2003) Energy status and its control on embryogenesis of legumes: ATP distribution within *Vicia faba* embryos is developmentally regulated and correlated with photosynthetic capacity. Plant J 36:318–329. https://doi.org/10.1046/j.1365-313X.2003.01879.x

Cobos MJ, Winter P, Kharrat M, Cubero JI, Gil J, Millan T, Rubio J (2009) Genetic analysis of agronomic traits in a wide cross of chickpea. Field Crop Res 111:130–136. https://doi.org/10.1016/j.fcr.2008.11.006

Cowan RK, Hoen DR, Schoen DJ, Bureau TE (2005) MUSTANG is a novel family of domesticated transposase genes found in diverse angiosperms. Mol Biol Evol 22:2084–2089. https://doi.org/10.1093/molbev/msi202

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R (2011) The variant call format and VCFtools. Bioinformatics 27:2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Das S, Upadhyaya HD, Bajaj D, Kujur A, Badoni S, Laxmi Kumar V, Tripathi S, Gowda CLL, Sharma S, Singh S, Tyagi AK, Parida SK (2015) Deploying QTL-seq for rapid delineation of a potential candidate gene underlying major trait-associated QTL in chickpea. DNA Res 22:193–203. https://doi.org/10.1093/dnares/dsv004

Doddamani D, Katta MAVSK, Khan AW, Agarwal G, Shah TM, Varshney RK (2014) CicArMiSatDB: the chickpea microsatellite database. BMC Bioinform 15:212. https://doi.org/10.1186/1471-2105-15-212

Doughty J, Aljabri M, Scott RJ (2014) Flavonoids and the regulation of seed size in *Arabidopsis*. Biochem Soc Trans 42:364–369. https://doi.org/10.1042/BST20140040

Edwards D (2016) Improved desi reference genome. CyVerse Data Commons. https://doi.org/10.7946/P2KW2Q

Faraco M, Spelt C, Bliek M, Verweij W, Hoshino A, Espen L, Prinsi B, Jaarsma R, Tarhan E, DeBoer AH, Di Sansebastiano G-P, Koes R, Quattrocchio FM (2014) Hyperacidification of vacuoles by the combined action of two different P-ATPases in the tonoplast determines flower color. Cell Rep 6:32–43. https://doi.org/10.1016/j.celrep.2013.12.009

Figueiredo DD, Batista RA, Roszak PJ, Hennig L, Köhler C (2016) Auxin production in the endosperm drives seed coat development in Arabidopsis. Elife 5:e20542. https://doi.org/10.7554/eLife.20542

Folsom JJ, Begcy K, Hao X, Wang D, Walia H (2014) Rice fertilization-independent endosperm1 regulates seed size under heat stress by controlling early endosperm development. Plant Physiol 165:238–248. https://doi.org/10.1104/pp.113.232413

Garcia D, Fitz Gerald JN, Berger F (2005) Maternal control of integument cell elongation and zygotic control of endosperm growth are coordinated to determine seed size in Arabidopsis. Plant Cell 17:52–60. https://doi.org/10.1105/tpc.104.027136

Gaur PM, Singh MK, Samineni S, Sajja SB, Jukanti AK, Kamatam S, Varshney RK (2016) Inheritance of protein content and its relationships with seed size, grain yield and other traits in chickpea. Euphytica 209:253–260. https://doi.org/10.1007/s10681-016-1678-2

Glaubitz JC, Casstevens TM, Lu F, Harriman J, Elshire RJ, Sun Q, Buckler ES (2014) TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. PLoS ONE 9:e90346. https://doi.org/10.1371/journal.pone.0090346

Hehenberger E, Kradolfer D, Köhler C (2012) Endosperm cellularization defines an important developmental transition for embryo development. Development 139:2031–2039. https://doi.org/10.1242/dev.077057

Hossain S, Ford R, McNeil D, Pittock C, Panozzo JF (2010) Inheritance of seed size in chickpea (*Cicer arietinum* L.) and identification of QTL based on 100-seed weight and seed size index. Aust J Crop Sci 4:126–135

Hurkman WJ, Wood DF (2011) High temperature during grain fill alters the morphology of protein and starch deposits in the starchy endosperm cells of developing wheat (*Triticum aestivum* L.) grain. J Agric Food Chem 59:4938–4946. https://doi.org/10.1021/jf102962t

Ingouff M, Jullien PE, Berger F (2006) The female gametophyte and the endosperm control cell proliferation and differentiation of the seed coat in Arabidopsis. Plant Cell 18:3491–3501. https://doi.org/10.1105/tpc.106.047266

Jadhav AA, Rayate SJ, Mhase LB, Thudi M, Chitikineni A, Harer PN, Jadhav AS, Varshney RK, Kulwal PL (2015) Marker-trait association study for protein content in chickpea (*Cicer arietinum* L.). J Genet 94:279–286

Jain M, Misra G, Patel RK, Priya P, Jhanwar S, Khan AW, Shah N, Singh VK, Garg R, Jeena G, Yadav M, Kant C, Sharma P, Yadav G, Bhatia S, Tyagi AK, Chattopadhyay D (2013) A draft genome sequence of the pulse crop chickpea (*Cicer arietinum* L.). Plant J 74:715–729. https://doi.org/10.1111/tpj.12173

Joly-Lopez Z, Forczek E, Hoen DR, Juretic N, Bureau TE (2012) A gene family derived from transposable elements during early

angiosperm evolution has reproductive fitness benefits in *Arabidopsis thaliana*. PLoS Genet. https://doi.org/10.1371/journal.pgen.1002931

Kaushal N, Awasthi R, Gupta K, Gaur P, Siddique KHM, Nayyar H (2013) Heat-stress-induced reproductive failures in chickpea (*Cicer arietinum*) are associated with impaired sucrose metabolism in leaves and anthers. Funct Plant Biol 40:1334–1349. https://doi.org/10.1071/FP13082

Knights EJ, Wood JA, Harden S (2011) A gene influencing seed shape of desi type chickpea (*Cicer arietinum* L.). Plant Breed 130:278–280. https://doi.org/10.1111/j.1439-0523.2010.01810.x

Kujur A, Upadhyaya HD, Shree T, Bajaj D, Das S, Saxena MS, Badoni S, Kumar V, Tripathi S, Gowda CLL, Sharma S, Singh S, Tyagi AK, Parida SK (2015) Ultra-high density intra-specific genetic linkage maps accelerate identification of functionally relevant molecular tags governing important agronomic traits in chickpea. Sci Rep 5:9468. https://doi.org/10.1038/srep09468

Lafon-Placette C, Köhler C (2014) Embryo and endosperm, partners in seed development. Curr Opin Plant Biol 17:64–69. https://doi.org/10.1016/j.pbi.2013.11.008

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25. https://doi.org/10.1186/gb-2009-10-3-r25

Lefèvre F, Boutry M (2018) Towards identification of the substrates of ATP-binding cassette transporters. Plant Physiol 178:00325. https://doi.org/10.1104/pp.18.00325

Lepiniec L, Debeaujon I, Routaboul J-M, Baudry A, Pourcel L, Nesi N, Caboche M (2006) Genetics and biochemistry of seed flavonoids. Annu Rev Plant Biol 57:405–430. https://doi.org/10.1146/annurev.arplant.57.032905.105252

Li N, Li Y (2016) Signaling pathways of seed size control in plants. Curr Opin Plant Biol 33:23–32. https://doi.org/10.1016/j.pbi.2016.05.008

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li Y, Ruperao P, Batley J, Edwards D, Khan T, Colmer TD, Pang J, Siddique KHM, Sutton T (2018) Investigating drought tolerance in chickpea using genome-wide association mapping and genomic selection based on whole-genome resequencing data. Front Plant Sci 9:1–12. https://doi.org/10.3389/fpls.2018.00190

Magrane M, Consortium UP (2011) UniProt Knowledgebase: a hub of integrated protein data. Database 2011, bar009. https://doi.org/10.1093/database/bar009

Mascher M, Wu S, St. Amand P, Stein N, Poland J (2013) Application of genotyping-by-sequencing on semiconductor sequencing platforms: a comparison of genetic and reference-based marker ordering in barley. PLoS ONE 8:e76925. https://doi.org/10.1371/journal.pone.0076925

Mizukami Y (2001) A matter of size: developmental control of organ size in plants. Curr Opin Plant Biol 4:533–539. https://doi.org/10.1016/S1369-5266(00)00212-0

Morley-Smith ER, Pike MJ, Findlay K, Köckenberger W, Hill LM, Smith AM, Rawsthorne S (2008) The transport of sugars to developing embryos is not via the bulk endosperm in oilseed rape seeds. Plant Physiol 147:2121–2130. https://doi.org/10.1104/pp.108.124644

Parween S, Nawaz K, Roy R, Pole AK, Suresh BV, Misra G, Jain M, Yadav G, Parida SK, Tyagi AK, Bhatia S, Chattopadhyay D (2015) An advanced draft genome assembly of a desi type chickpea (*Cicer arietinum* L). Sci Rep 5:12806. https://doi.org/10.1038/srep12806

Ruan Y-L, Patrick JW, Bouzayen M, Osorio S, Fernie AR (2012) Molecular regulation of seed and fruit set. Trends Plant Sci 17:656–665. https://doi.org/10.1016/j.tplants.2012.06.005

Saxena MS, Bajaj D, Das S, Kujur A, Kumar V, Singh M, Bansal KC, Tyagi AK, Parida SK (2014) An integrated genomic approach for rapid delineation of candidate genes regulating agro-morphological traits in chickpea. DNA Res 21:695–710. https://doi.org/10.1093/dnares/dsu031

Sonah H, Bastien M, Iquira E, Tardivel A, Légaré G, Boyle B, Normandeau É, Laroche J, Larose S, Jean M, Belzile F (2013) An improved genotyping by sequencing (GBS) approach offering increased versatility and efficiency of SNP discovery and genotyping. PLoS ONE 8:e54603. https://doi.org/10.1371/journal.pone.0054603

Sun X, Shantharaj D, Kang X, Ni M (2010) Transcriptional and hormonal signaling control of Arabidopsis seed development. Curr Opin Plant Biol 13:611–620. https://doi.org/10.1016/j.pbi.2010.08.009

Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ, von Mering C (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 43:D447–D452. https://doi.org/10.1093/nar/gku1003

Tondelli A, Francia E, Barabaschi D, Aprile A, Skinner JS, Stockinger EJ, Stanca AM, Pecchioni N (2006) Mapping regulatory genes as candidates for cold and drought stress tolerance in barley. Theor Appl Genet 112:445–454. https://doi.org/10.1007/s00122-005-0144-7

Upadhyaya HD, Bajaj D, Das S, Kumar V, Gowda CLL, Sharma S, Tyagi AK, Parida SK (2016a) Genetic dissection of seed-iron and zinc concentrations in chickpea. Sci Rep 6:24050. https://doi.org/10.1038/srep24050

Upadhyaya HD, Bajaj D, Narnoliya L, Das S, Kumar V, Gowda CLL, Sharma S, Tyagi AK, Parida SK (2016b) Genome-wide scans for delineation of candidate genes regulating seed-protein content in chickpea. Front Plant Sci 7:302. https://doi.org/10.3389/fpls.2016.00302

van Ooijen JW, Voorips RE (2001) JoinMap 3.0, software for the calculation of genetic linkage maps

Varshney RK, Song C, Saxena RK, Azam S, Yu S, Sharpe AG, Cannon S, Baek J, Rosen BD, Tar'an B, Millan T, Zhang X, Ramsay LD, Iwata A, Wang Y, Nelson W, Farmer AD, Gaur PM, Soderlund C, Penmetsa RV, Xu C, Bharti AK, He W, Winter P, Zhao S, Hane JK, Carrasquilla-Garcia N, Condie JA, Upadhyaya HD, Luo M-C, Thudi M, Gowda CLL, Singh NP, Lichtenzveig J, Gali KK, Rubio J, Nadarajan N, Dolezel J, Bansal KC, Xu X, Edwards D, Zhang G, Kahl G, Gil J, Singh KB, Datta SK, Jackson SA, Wang J, Cook DR (2013) Draft genome sequence of chickpea (*Cicer arietinum)* provides a resource for trait improvement. Nat Biotechnol 31:240–246. https://doi.org/10.1038/nbt.2491

Verma M, Kumar V, Patel RK, Garg R, Jain M (2015a) CTDB: an integrated chickpea transcriptome database for functional and applied genomics. PLoS ONE 10:1371. https://doi.org/10.1371/journal.pone.0136880

Verma S, Gupta S, Bandhiwal N, Kumar T, Bharadwaj C, Bhatia S (2015b) High-density linkage map construction and mapping of seed trait QTLs in chickpea (*Cicer arietinum* L) using genotyping-by-sequencing (GBS). Sci Rep 5:17512. https://doi.org/10.1038/srep17512

Verweij W, Spelt C, Di Sansebastiano G-P, Vermeer J, Reale L, Ferranti F, Koes R, Quattrocchio F (2008) An H+P-ATPase on the tonoplast determines vacuolar pH and flower colour. Nat Cell Biol 10:1456–1462. https://doi.org/10.1038/ncb1805

Voorrips R (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered 93:77–78

Wang R, Gangola MP, Jaiswal S, Båga M, Gaur PM, Chibbar RN (2017) Variation in seed-quality traits of chickpea and their correlation to raffinose family oligosaccharides concentrations. Crop Sci 57:1594–1602. https://doi.org/10.2135/cropsci2016.08.0710

Weber H, Borisjuk L, Wobus U (2005) Molecular physiology of legume seed development. Annu Rev Plant Biol 56:253–279. https://doi.org/10.1146/annurev.arplant.56.032604.144201

Wood JA, Grusak MA (2007) Nutritional value of chickpea. In: Yadav SS, Redden R, Chen W, Sharma B (eds) Chickpea breeding and management. CAB International, Wallingford, pp 101–142

Wood JA, Knights EJ, Campbell GM, Choct M (2014) Differences between easy- and difficult-to-mill chickpea (*Cicer arietinum* L.) genotypes. Part I: broad chemical composition. J Sci Food Agric 94:1437–1445. https://doi.org/10.1002/jsfa.6437

Yang J, Hu C, Hu H, Yu R, Xia Z, Ye X, Zhu J (2008) QTLNetwork: mapping and visualizing genetic architecture of complex traits in experimental populations. Bioinformatics 24:721–723. https://doi.org/10.1093/bioinformatics/btm494

Yu S-M, Lo S-F, Ho T-HD (2015) Source-sink communication: regulated by hormone, nutrient, and stress cross-signaling. Trends Plant Sci 20:844–857. https://doi.org/10.1016/j.tplants.2015.10.009