

# SCIENTIFIC REPORTS



OPEN

## Identification of candidate genes and natural allelic variants for QTLs governing plant height in chickpea

Alice Kujur<sup>1</sup>, Hari D. Upadhyaya<sup>2</sup>, Deepak Bajaj<sup>1</sup>, C. L. L. Gowda<sup>2</sup>, Shivali Sharma<sup>2</sup>, Akhilesh K. Tyagi<sup>1</sup> & Swarup K. Parida<sup>1</sup>

Received: 28 October 2015

Accepted: 25 May 2016

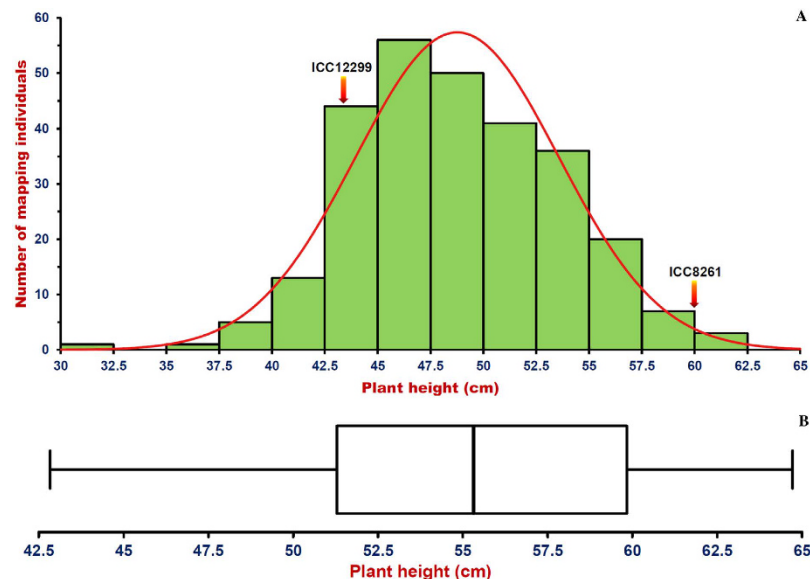
Published: 20 June 2016

In the present study, molecular mapping of high-resolution plant height QTLs was performed by integrating 3625 *desi* genome-derived GBS (genotyping-by-sequencing)-SNPs on an ultra-high resolution intra-specific chickpea genetic linkage map (dwarf/semi-dwarf *desi* cv. ICC12299 x tall *kabuli* cv. ICC8261). The identified six major genomic regions harboring six robust QTLs (11.5–21.3 PVE), associated with plant height, were mapped within <0.5 cM average marker intervals on six chromosomes. Five SNPs-containing genes tightly linked to the five plant height QTLs, were validated based upon their high potential for target trait association (12.9–20.8 PVE) in 65 *desi* and *kabuli* chickpea accessions. The vegetative tissue-specific expression, including higher differential up-regulation (>5-fold) of five genes especially in shoot, young leaf, shoot apical meristem of tall mapping parental accession (ICC8261) as compared to that of dwarf/semi-dwarf parent (ICC12299) was apparent. Overall, combining high-resolution QTL mapping with genetic association analysis and differential expression profiling, delineated natural allelic variants in five candidate genes (encoding cytochrome-c-biosynthesis protein, malic oxidoreductase, NADH dehydrogenase iron-sulfur protein, expressed protein and *bZIP* transcription factor) regulating plant height in chickpea. These molecular tags have potential to dissect complex plant height trait and accelerate marker-assisted genetic enhancement for developing cultivars with desirable plant height ideotypes in chickpea.

Plant height is a vital yield-component trait. It affects crop's performance, particularly lodging and subsequently grain yield and grain quality<sup>1,2</sup>. Crop plants of erect and a very tall habit tend to lodge as they approach maturity, whereas erect accessions with reduced height can avoid wind and rain damage, and are expected to be more lodging resistant. These desirable characteristics in turn enhance the potential of dwarf/semi-dwarf crop plant types to produce higher grain yield that are also more amenable to mechanical harvesting<sup>1,2</sup>. The reduction of plant height has therefore, been an important breeding and genomics research objective for many decades. Remarkably, the development of dwarf and semi-dwarf varieties of wheat and rice invoking "Green Revolution" during the 1960s and 1970s, is a famous historical milestone achieved in crop genetic improvement<sup>3,4</sup>. This has paved the way for plant breeders/molecular geneticists to substantially increase the grain yield through identification of variants (major dwarfing or semi-dwarfing genes) that reduce the plant height without adversely affecting the yield potential. However, tallness in plants has almost universally been shown to be dominant over dwarfness since Mendel's times and thus developing plants with reduced height is still a challenging task.

Chickpea (*Cicer arietinum* L.) belonging to family Fabaceae, is a self-pollinated and diploid legume crop species with a genome size of ~740 Mb<sup>5,6</sup>. It is an annual, much branched and herbaceous crop with plant height mostly ranging from 30 to 70 cm. The genomes of two most common types of chickpea cultivars - *desi* and *kabuli* have been sequenced<sup>7–9</sup>. These available genomic resources have accelerated the process of NGS (next-generation sequencing)-based whole genome and transcriptome sequencing of numerous cultivated (*desi* and *kabuli*) and wild accessions as well as NGS-/array-based large-scale discovery and high-throughput genotyping of informative microsatellite and SNP (single nucleotide polymorphism) markers in natural and mapping populations of chickpea<sup>10–18</sup>. Consequently, these efforts have significantly driven the process of constructing high-density intra- and inter-specific genetic linkage maps and genetic/association mapping to identify QTLs (quantitative trait loci)/genes governing useful agronomic traits in chickpea<sup>12,19–35</sup>. One such most promising outcome includes

<sup>1</sup>National Institute of Plant Genome Research (NIPGR), Aruna Asaf Ali Marg, New Delhi 110067, India. <sup>2</sup>International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502324, Telangana, India. Correspondence and requests for materials should be addressed to S.K.P. (email: swarup@nipgr.ac.in)



**Figure 1.** (A) Frequency distribution of plant height (PH cm) trait variation estimated in a 275 mapping population (ICC 12299 x ICC 8261) illustrated a goodness of fit to the normal distribution. (B) Boxplots illustrating the variation of PH trait among 65 *desi* and *kabuli* chickpea accessions. Box edges denote the lower and upper quartiles with median values in the middle of the box.

the use of GBS (genotyping-by-sequencing) assay in discovery/genotyping of genome-wide SNPs in natural germplasm lines and mapping populations for understanding domestication and LD (linkage disequilibrium) pattern and constructing ultra-high density genetic linkage maps. This will further assist in fine mapping and high-resolution genome-wide association study (GWAS) to identify genomic loci (candidate genes) harboring major/minor QTLs regulating stress tolerance and yield-contributing traits for genomics-assisted crop improvement of chickpea<sup>26,34,36–38</sup>.

Considering the added-advantages of QTL and association mapping in quantitative dissection of complex traits, a number of genes underlying QTLs associated with plant height that rely on these approaches, have been identified in crop plants, including legumes<sup>39–54</sup>. Similar major efforts have been made in the past few years towards molecular mapping of low-resolution QTLs governing plant height on multiple intra- as well as inter-specific genetic linkage maps of chickpea<sup>6,29,34,37,55–58</sup>. However, none of these plant height QTLs could essentially be validated hitherto in diverse genetic backgrounds of chickpea accessions across multiple geographical locations (years). Consequently, none of the potential genes underlying plant height QTLs have been isolated and functionally validated/characterized through fine-mapping/map-based cloning and further functional genomic approaches to be effectively deployed in marker-assisted genetic enhancement of chickpea.

To expedite these useful genetic applications, the identification of functionally relevant genes harboring high-resolution QTLs at a genome-wide scale, by construction of SNP marker-based ultra-high resolution genetic linkage maps using GBS assay and their further validation through trait association analysis and differential expression profiling, at present is anticipated as an attractive strategy for chickpea genetic improvement. In this context, the present investigation employed aforesaid combinatorial approach to understand the complex genetic architecture of plant height trait for marker-assisted genetic enhancement with a definitive objective to develop high-yielding cultivars with desirable plant architecture (height) in chickpea.

## Results

**An ultra-high density intra-specific genetic linkage map accelerates molecular mapping of high-resolution PH QTLs in chickpea.** The sequencing of 288-plex *Ape*KI libraries constructed from 275 RIL mapping individuals and parents (ICC 12299 x ICC 8261) using the reference *desi* genome-based GBS assay discovered 26785 high-quality SNPs differentiating the two parental accessions (as per Kujur *et al.*<sup>26</sup>). The use of genotyping information (present in parental accessions and all 275 RIL mapping individuals) of 4448 high-quality SNPs in linkage analysis mapped 3625 SNPs on eight LGs of an intra-specific genetic map based on their physical positions (bp) on corresponding *desi* chromosomes (Table S1). As per our previous study<sup>26</sup>, this GBS SNP-based ultra-high density intra-specific genetic map covered a total map-length of 714.1 cM with an average inter-marker distance of 0.20 cM.

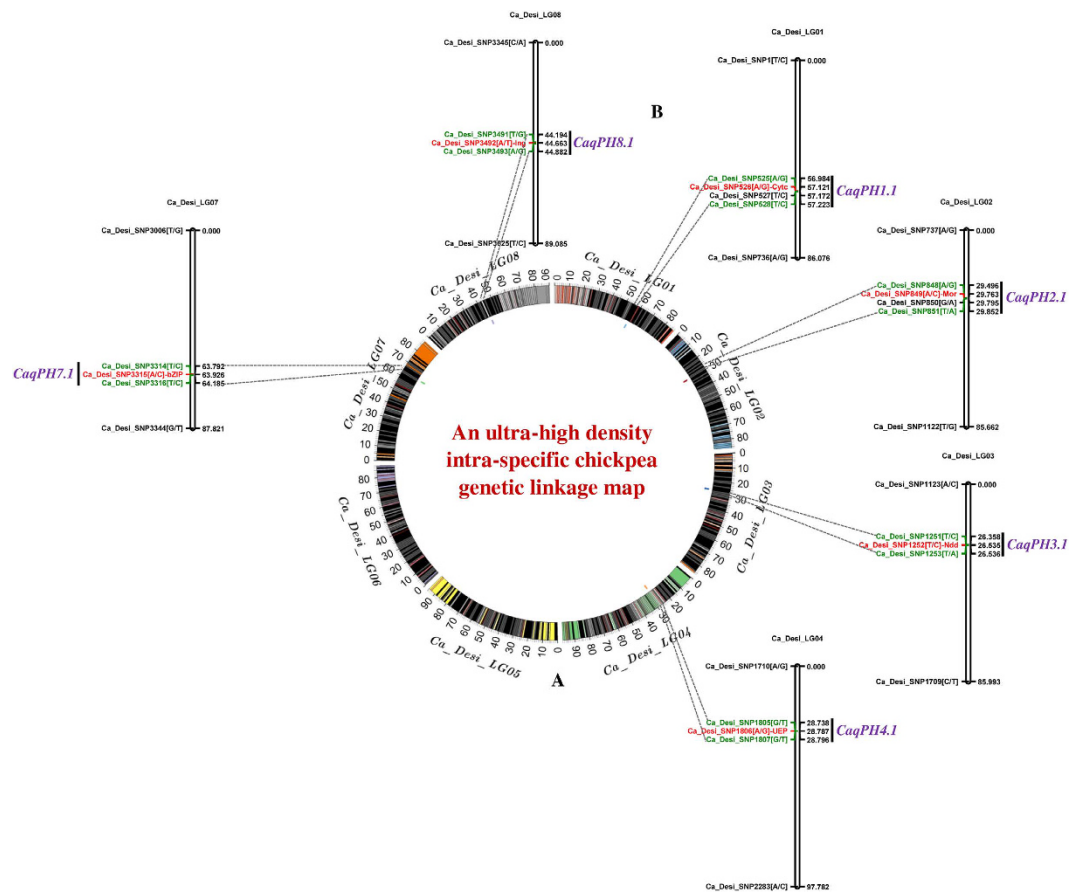
A significant difference ( $p \leq 0.001$ ) of PH trait [ranging 30.6–61.3 cm with mean  $\pm$  standard deviation (SD) of  $48.8 \pm 4.8$  and 82%  $H^2$  (broad-sense heritability)] in parental accessions and 275 RIL mapping individuals was observed across two years based on ANOVA. The coefficient of variation (CV) of PH trait was measured as 9.8%. A bi-directional transgressive segregation, including normal frequency distribution of PH trait in the generated RIL mapping population was evident (Fig. 1A). For molecular mapping of QTLs, the genotyping data of 3625 parental polymorphic *desi*-genome based SNPs mapped on an ultra-high density intra-specific genetic linkage

QTL mapping						Association mapping		SNPs tightly linked to QTLs (trait-associated SNPs)	Structural annotation	Functional annotation
Linkage groups (LGs)	QTLs identities	Marker intervals with genetic positions (cM)	LOD	Phenotypic variation explained (PVE%) by QTLs	Additive effects	P	R <sup>2</sup> (%)			
Ca_Desi_LG01	<i>qPH1.1</i>	0.239: Ca_Desi_SNP525 (56.984) to Ca_Desi_SNP528 (57.223)	10.7	21.3	4.5	$2.1 \times 10^{-6}$	20.8	Ca_Desi_SNP526 (A/G)	Ca19207 (URR)	Cytochrome C biosynthesis protein
Ca_Desi_LG02	<i>qPH2.1</i>	0.356: Ca_Desi_SNP848 (29.496) to Ca_Desi_SNP851 (29.852)	8.6	15.9	3.9	$4.2 \times 10^{-6}$	17.9	Ca_Desi_SNP849 (A/C)	Ca10330 (CDS-Syn)	Mitochondrial NAD-dependent malic enzyme (Malic oxidoreductase)
Ca_Desi_LG03	<i>qPH3.1</i>	0.178: Ca_Desi_SNP1251 (26.358) to Ca_Desi_SNP1253 (26.536)	8.1	16.2	2.8	$1.3 \times 10^{-5}$	16.7	Ca_Desi_SNP1252 (T/C)	Ca20411 (DRR)	NADH dehydrogenase [ubiquinone] iron-sulfur protein 3
Ca_Desi_LG04	<i>qPH4.1</i>	0.058: Ca_Desi_SNP1805 (28.738) to Ca_Desi_SNP1807 (28.796)	7.8	12.4	2.5	$1.9 \times 10^{-5}$	12.9	Ca_Desi_SNP1806 (A/G)	Ca04741 (Intron)	Unknown expressed protein
Ca_Desi_LG07	<i>qPH7.1</i>	0.393: Ca_Desi_SNP3314 (63.792) to Ca_Desi_SNP3316 (64.185)	9.5	20.5	2.1	$3.9 \times 10^{-6}$	20.3	Ca_Desi_SNP3315 (A/C)	Ca09193 (Intron)	Basic leucine zipper ( <i>bZIP</i> ) transcription factor 60
Ca_Desi_LG08	<i>qPH8.1</i>	0.688: Ca_Desi_SNP3491 (44.194) to Ca_Desi_SNP3493 (44.882)	6.1	11.5	1.7	$4.8 \times 10^{-5}$	–	Ca_Desi_SNP3492 (A/T)	Intergenic	–

**Table 1. Candidate genes harbouring major QTLs regulating PH identified by QTL mapping and genetic association analysis in chickpea.**

map was correlated with phenotyping information of 275 RIL mapping individuals and parental accessions of chickpea. This identified six major genomic regions harboring six robust (6.1–10.7 LOD) QTLs (*qPH1.1*, *qPH2.1*, *qPH3.1*, *qPH4.1*, *qPH7.1* and *qPH8.1*) governing PH, which were mapped on six LGs (except LG5 and LG6) of chickpea (Table 1, Fig. 2A,B). The proportion of phenotypic variation explained (PVE) for PH trait by individual QTL ranged from 11.5–21.3%. These six QTLs in combination gave 35.8% PVE for PH trait. All these QTLs, validated across two geographical locations and years/seasons in the field and revealed consistent phenotypic expression (PVE: >10%) and major effects on PH traits, were thus considered as ‘robust PH QTLs’ in chickpea. The major genomic regions (from 0.058 cM on LG4 to 0.688 cM on LG8) underlying the robust QTLs covered with 20 SNPs, were mapped on six LGs (Table 1, Fig. 2B). These QTLs exhibited mostly the additive gene effects (1.7 to 4.5) signifying the effective contributions of ICC 12299 alleles at these loci for dwarf/semi-dwarf characteristics in chickpea. The identified major genomic regions harboring six robust PH QTLs as well as SNPs flanking/tightly linked to these PH QTLs mapped on an ultra-high density intra-specific genetic linkage maps were annotated based upon their structural and functional annotation information as documented in our previous study<sup>26</sup>. This analysis identified diverse coding (synonymous and non-synonymous) and non-coding regulatory and intronic SNPs in the five genes [encoding cytochrome c biosynthesis protein, malic oxidoreductase, NADH dehydrogenase iron-sulfur protein, unknown expressed protein and basic leucine zipper (*bZIP*) transcription factor] and one intergenic SNP tightly linked to the QTLs controlling PH trait in chickpea (Table 1, Fig. 2B).

**Trait association mapping to delineate candidate gene(s) regulating PH trait in chickpea.** For genetic association mapping, the SNPs (flanking/linked to PH QTLs) mapped at the six major PH QTL regions of an ultra-high density intra-specific genetic linkage map were genotyped in 65 *desi* and *kabuli* chickpea accessions (Table S2). The SNP genotyping information was integrated with PH field phenotyping data (varying 42.8–64.7 cm with mean  $\pm$  standard deviation of  $55.1 \pm 5.7$  and 80% H<sup>2</sup>) of 65 accessions (Fig. 1B). The association analysis based on CMLM and false discovery rate (FDR) correction (minimizing the confounding effect of population structure) detected five SNPs-containing genes revealing significant association ( $1.3 \times 10^{-5}$  to  $2.1 \times 10^{-6}$  P with 12.9 to 20.8% R<sup>2</sup>) with PH trait in chickpea (Table 1). These PH-associated SNPs-carrying genes localized at the major genomic regions harboring five robust QTLs were also found to be tightly linked with PH QTLs based on high-resolution QTL mapping (Table 1). Henceforth, five functionally relevant SNP loci-carrying genes (encoding cytochrome c biosynthesis protein, malic oxidoreductase, NADH dehydrogenase iron-sulfur protein, unknown expressed protein and *bZIP* transcription factor) validated both by high-resolution QTL and association mapping, were considered as candidates for plant height regulation in chickpea. Interestingly, two upstream regulatory (URR) and intronic SNPs (A/G and A/C)-containing cytochrome c biosynthesis protein-coding gene and *bZIP* transcription factor gene, respectively, tightly linked to the PH QTLs (*qPH1.1* and *qPH7.1*) exhibited strong association (average 20% PVE) with plant height based on QTL mapping and trait association analysis in chickpea.

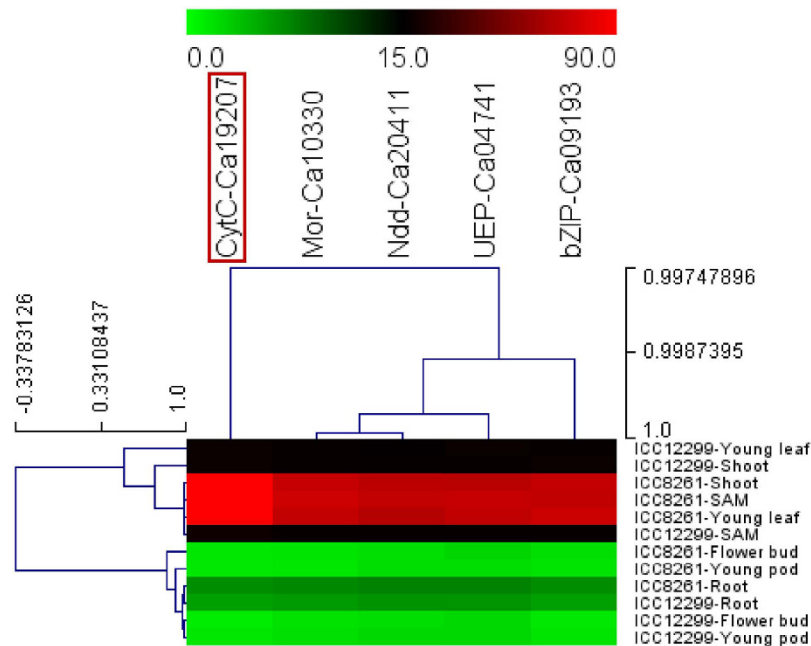


**Figure 2.** (A) An ultra-high resolution intra-specific genetic map (ICC 12299 x ICC 8261), constructed by integrating 3625 SNPs on eight LGs of chickpea, is illustrated by a Circos circular ideogram. The circles signifies the different genetic map length (cM) (spanning 10 cM uniform genetic distance intervals between bins) of eight LGs coded with diverse colours. (B) Molecular mapping of six major genomic regions underlying robust PH QTLs on six *desi* LGs/chromosomes (1, 2, 3, 4, 7 and 8). The SNPs exhibiting strong linkage with PH QTLs are highlighted with red colour. The SNPs flanking the PH QTL intervals are marked with green colour. The genetic (cM) distance and identities of the SNPs mapped on the LGs are denoted on the right and left side of LGs, respectively. The detail information regarding SNPs and QTLs are provided in the Table S1 and Table 1.

**Expression profiling of PH-associated genes.** To determine the differential gene regulatory expression patterns, five PH-associated SNPs-containing genes delineated at the major genomic regions harboring robust PH QTLs (by high-resolution QTL and association mapping) were utilized in expression analysis. The differential expression profiles of these genes in diverse vegetative and reproductive tissues (root, shoot, young leaf, shoot apical meristem, flower bud and young pod) of tall (ICC 8261) and dwarf/semi-dwarf (ICC 12299) parental accessions of an intra-specific mapping population (ICC 12299 x ICC 8261) were determined (Fig. 3). This revealed vegetative tissue-specific expression (as compared to reproductive tissues) as well as up-regulation (>5-folds) of five SNPs-carrying genes especially in shoot, young leaves and shoot apical meristem of tall parental accession (ICC 8261) than that of dwarf/semi-dwarf parent (ICC 12299). Notably, two strong PH-associated URR and intronic SNPs-containing vegetative tissue-specific cytochrome c biosynthesis protein-coding gene and *bZIP* transcription factor tightly linked to the PH QTLs (*qPH1.1* and *qPH7.1*) exhibited pronounced up-regulation (>6–7 folds) in shoots, young leaves and shoot apical meristems of tall parental accession (ICC 8261) as compared to that of dwarf/semi-dwarf parent (ICC 12299) (Fig. 3). Therefore, natural allelic variants-containing five potential candidate genes validated by an integrated genomic approach encompassing QTL and association mapping and expression profiling, were selected as target candidates for quantitative dissection of complex plant height trait in chickpea (Table 1).

## Discussion

The implication of GBS-assay in simultaneous genome-wide mining and high-throughput genotyping of high-quality and non-erroneous SNPs in diverse advanced generation intra- and inter-specific mapping populations for construction of ultra-high resolution genetic linkage maps have been demonstrated in chickpea<sup>26,36,37</sup>. A 3265 genome-wide SNP markers-led ultra-high density intra-specific genetic linkage map constructed by us using a 275 chickpea RIL mapping population (ICC 12299 x ICC 8261) exhibited a much higher mean map-density (with a low average inter-marker distance 0.20 cM). This is highly saturated and/or



**Figure 3. Hierarchical cluster display depicting the differential expression profiles of five PH-associated genes (validated by QTL mapping and association analysis) in vegetative and reproductive tissues (root, shoot, young leaf, shoot apical meristem, flower bud and young pod) of tall (ICC 8261) and dwarf/semi-dwarf (ICC 12299) parental accessions of an intra-specific mapping population (ICC 12299 x ICC 8261) using quantitative RT-PCR assay.** The colour scale at the top represents the average log signal expression values of genes in different tissues; in which green, black and red color signify low, medium and high level of expression, respectively. The gene revealing pronounced differential up-regulation especially in shoot, young leaf and shoot apical meristem of tall accessions is highlighted with red box. The details regarding genes are provided in the Table 1. The tissues and genes utilized for expression profiling are illustrated on the right and top side of expression map, respectively. To normalize the expression values across diverse tissues of accessions, an endogenous control *elongation factor-1 alpha* was utilized. *CytC*: Cytochrome C, *Mor*: Malic oxidoreductase, *Ndd*: NADH dehydrogenase, *UEP*: Unknown expressed protein and *bZIP*: basic leucine zipper.

comparable to the latest available high-resolution inter- and intra-specific genetic linkage maps (0.16–3.68 cM) in chickpea<sup>14,22,26,29,35–37,59–61</sup>. Therefore, an ultra-high density intra-specific genetic linkage map constructed in the present study has efficacy in identification and fine-mapping (map-based cloning) of major/minor QTLs governing traits of agricultural importance in chickpea.

The use of ultra-high density intra-specific genetic linkage map generated in our study as a reference identified and mapped six major genomic regions underlying six robust QTLs (11.5–21.3% PVE with 35.8% combined PVE) governing PH traits in chickpea. To ascertain the robustness and novelty of the identified PH QTLs, the major genomic regions underlying six robust PH QTLs were compared/correlated with that of previous QTL mapping studies based on correspondence of physical positions of SNPs covering (flanked/linked) the target QTL intervals<sup>37,58</sup>. None of these six QTLs were found to be congruent with previous QTL mapping studies involving diverse intra- and inter-specific mapping populations of chickpea. Henceforth, novelty and population-specific characteristics of the identified six major and robust QTLs governing PH trait was apparent in chickpea. A number of QTL mapping studies involving high resolution intra- and inter-specific genetic linkage maps as references/anchors have been undertaken for rapid molecular mapping as well as fine mapping/map-based cloning of major genes/QTLs governing multiple stress tolerance- and yield-component traits to expedite marker-assisted genetic enhancement in chickpea<sup>26,28,29,35–37,59</sup>. Interestingly, with the use of an ultra-high density genetic linkage map as a reference in the present investigation, the major PH QTLs significantly scaled-down into the marker intervals of 0.058–0.688 cM (average <0.5 cM) in chickpea. This is much lower than that documented (1.05–3.6 cM) till date particularly for PH QTLs in different QTL mapping studies of chickpea<sup>37,59</sup>. The comparative analysis between past and present QTL mapping study suggests the efficacy of these identified major and high-resolution robust PH QTLs for understanding the genetic inheritance pattern and in dissection of complex plant height trait of chickpea. Therefore, the SNPs derived from diverse coding and non-coding sequence components of genes and/or intergenic regions tightly linked to the six major PH QTLs can be deployed in marker-assisted breeding to select cultivars with desirable plant height for genetic improvement in chickpea. The subsequent validation of the PH QTLs in multiple genetic backgrounds/environments along with their fine mapping/map-based cloning will be helpful to select favorable and superior molecular tags for accelerating genomics-assisted crop improvement in chickpea.

Interestingly, the integration of high-resolution QTL and association mapping validated informative SNPs-carrying five candidate genes harboring five major PH QTLs reflecting the robustness and reliability of

identified molecular tags for their immense deployment in marker-assisted genetic improvement of chickpea. The efficacy of these genes in regulating PH trait was further ascertained by their vegetative tissue-specific expression and pronounced up-regulation in tall than that of dwarf/semi-dwarf mapping parental accessions. The added-advantage of such integrated genomics-assisted breeding strategy for delineation of potential candidate genes regulating important yield-contributing traits have been demonstrated in chickpea<sup>23–26,31–33,35,38</sup>. The SNPs-carrying genes (encoding cytochrome c biosynthesis protein, NADH dehydrogenase iron-sulfur protein and malic oxidoreductase) regulating PH trait scaled-down by the integrated genomic approach, are evolutionarily conserved key components of mitochondrial respiratory complex and known to play a central role in classical electron transport chain (ETC) and tricarboxylic acid (TCA) cycle for regulating mitochondrial organization and respiration in crop plants<sup>62–67</sup>. The meristem-specific expression, including differential expression and accumulation of transcripts of these mitochondrial genes affect the cellular events related to mitochondrial respiration and photosynthetic activity by providing energy for biosynthesis and biomass storage during plant growth and development<sup>64,66,68</sup>. The regulation of these genes at transcriptional and post-transcriptional levels in different tissues is known to stimulate the mitochondrial respiratory apparatus of plants by assisting in shaping the organ development<sup>69</sup>. Another PH-associated basic leucine zipper functional domain-containing *bZIP* transcription factor gene regulates plant morphology by modulating the endogenous contents of gibberellins (GAs) and GA homeostasis. Interestingly, the reduction of GAs by down-regulation of genes (specifically *GA20-oxidase* and *GA3*) encoding GA biosynthetic enzymes severely inhibit the process of cell elongation/expansion and growth of stem internodes resulting in dwarf phenotype in multiple monocot and dicot plant species. Such GA response modulators, including green revolution genes- *Rht* (reduced height), *d8* (dwarf 8), *GAI* (gibberellin insensitive) and *Sd1* (semi-dwarf) governing GA signaling as well as plant height (dwarfism) have been identified and functionally characterized in wheat, maize, *Arabidopsis* and rice, respectively<sup>3,4,70,71</sup>. The correlation and interaction between mitochondrial ETC-mediated cytochrome c genes and GA synthesis-controlling *bZIP* transcription factors basing upon their similar expression characteristics in diverse tissues/developmental stages of crop plants is well deciphered. This indicates the involvement of these genes in an analogous functional regulatory pathway underlying both GA biosynthesis and mitochondrial respiration for maintaining growth and development of organs, including plant height<sup>72,73</sup>. Collectively, the PH-associated natural allelic variants-containing five genes delineated in the present study by integrating QTL and association mapping with differential expression profiling apparently play a crucial role in GA biosynthesis and mitochondrial respiration in crop plants. Henceforth, these molecular tags have got functional significance for understanding the regulation and quantitative dissection of complex plant height (dwarfism) trait in chickpea. The PH-associated superior natural allelic variants identified in the potential candidate genes once validated by translational genomic approaches (marker-assisted breeding and transgenics) can be deployed in crop genetic enhancement for developing genetically tailored improved chickpea cultivars with desirable plant height and architecture.

## Methods

**Generation and phenotyping of mapping population for plant height.** An intra-specific F<sub>7</sub> RIL mapping population (ICC 12299 x ICC 8261) consisting of 275 individuals was generated using a single seed descent method. The mapping parents ICC 12299 and ICC 8261 are dwarf/semi-dwarf and tall *desi* and *kabuli* chickpea (*C. arietinum*) accessions with plant height of 43 and 60 cm, respectively. These RIL mapping individuals along with parental accessions were grown in the field (following alpha-lattice design with two replications) for two consecutive years (2012–2013) at two diverse geographical locations of India. Subsequently, the RILs and parents were phenotyped individually for plant height (PH). The PH was documented by measuring the mean canopy height (cm) of 10 representative plants (selected from each RILs) from soil surface at the time of their flower ending/pod setting initiation stage. The diverse statistical parameters, including broad-sense heritability (H<sup>2</sup>%), mean, standard deviation, coefficient of variation (CV%), frequency distribution and analysis of variance (ANOVA) of PH trait in a RIL mapping population were determined using SPSS v17.0 and following Saxena *et al.*<sup>31</sup> and Bajaj *et al.*<sup>33</sup>.

**Mining and genotyping of genome-wide GBS-SNPs.** The genomic DNA isolated from 275 RIL mapping individuals and two parental accessions (as biological replicates) was used to construct the 288-plex (3 × 96-plex) GBS libraries as per Elshire *et al.*<sup>74</sup> and Kujur *et al.*<sup>25,26</sup>. The sequencing of these pooled libraries were performed by Illumina HiSeq2000 NGS platform. The sequencing data generated were submitted to a publicly accessible NCBI-SRA (sequence read archive) database (<http://www.ncbi.nlm.nih.gov/sra>) with accession number under BioProject PRJNA320682. The raw FASTA 100-bp single end sequence reads were processed for quality assessment, and subsequently high-quality sequences were filtered and mapped onto the reference genome using TASSEL 5.0 GBS Pipeline v2. The sequence reads with ≥10 *phred* Q-score across first 72-bp nucleotide sequence were considered high-quality, which were further ascertained through FASTQC v0.10.1. The retained good-quality sequences were de-multiplexed using unique barcodes tagged with each mapping parents and RIL individuals. The high-quality de-multiplexed reads of each mapping parents and RIL individuals were mapped against *desi* chickpea genome<sup>7</sup> using Burrows-Wheeler Alignment (BWA) tool with default parameters following Kujur *et al.*<sup>25,26</sup>. The SNPs were discovered among mapping parents and RIL individuals using SAMtools. To discover the high-quality SNPs, the imputation of missing SNP allele data was performed using haplotype probability method of fastPHASE<sup>26</sup>. The SNPs with base-quality ≥20 supported by minimum sequence read depth of 10 (< 10% missing data per individual and 100% reproducibility) in RIL mapping individuals and parental accessions were considered as high-quality SNPs.

**Construction of an intra-specific genetic linkage map.** The genotyping data of reference *desi* genome-based high-quality GBS-SNPs revealing polymorphism between parental accessions (ICC 12299 and ICC

8261) and 275 RIL mapping individuals were analyzed using JoinMap 4.1 (as per Bajaj *et al.*<sup>33</sup> and Kujur *et al.*<sup>26</sup>). At a higher LOD (logarithm of odds) threshold of >10.0, the SNPs were integrated [based on centiMorgan (cM) genetic distances] into eight linkage groups (LGs) of an intra-specific genetic map. The RECORD algorithm of RECORD\_WIN was employed to select the best marker-ordered genetic linkage maps with shortest map distances (cM). The LGs with genetically mapped SNPs were designated (LG1 to LG8) according to their corresponding physical positions (bp) on the chromosomes of *desi* chickpea genome.

**QTL mapping.** The reference genome-based GBS-SNP genotyping information was integrated with the genetic linkage map data of SNPs mapped on eight LGs of an intra-specific genetic map and PH field phenotyping data of 275 RIL mapping individuals and parental accessions using composite interval mapping functions of QTL Cartographer v2.5 and MapQTL v6.0. The novel genomic regions underlying the major QTLs associated with PH were identified and mapped on LGs at a significant LOD threshold score of  $\geq 5.0$  (1000 permutation at a  $p < 0.05$ ). The identified QTLs that validated across multiple environments (locations)/years were considered as 'robust QTLs' governing PH trait<sup>31</sup>. The phenotypic variation explained (PVE %) by QTLs and their additive effects (estimated by favorable alleles with parental origin) on PH trait were determined. The structural and functional annotation of genome-based SNPs, mapped on the eight LGs of the intra-specific genetic maps at target PH QTL regions, were performed according to the genome annotation (CGAP\_v1.0) information (GFF) of *desi* chickpea<sup>7</sup> following Kujur *et al.*<sup>25,26</sup>.

**Association analysis.** The SNPs (flanking/linked to PH QTLs) exhibiting differentiation between dwarf (ICC 12299) and tall (ICC 8261)-PH types parental accessions, mapped at the major QTL intervals, were genotyped in 65 *desi* and *kabuli* chickpea accessions (selected from the study of Kujur *et al.*<sup>25,38</sup>) using the Illumina GoldenGate SNP genotyping assay (as per Bajaj *et al.*<sup>33</sup>). The high-resolution LD (linkage disequilibrium)-based association analysis (compressed mixed model approach of EMMA) was performed following Saxena *et al.*<sup>31</sup>, Thudi *et al.*<sup>34</sup> and Kujur *et al.*<sup>38</sup>. The SNPs exhibiting strong association ( $R^2$ : correlation potential of significant SNPs with traits) with PH trait at a significant cut-off  $P \leq 10^{-5}$  were selected. This was performed by eliminating the confounding effect of population structure and implementing the Bonferroni correction of P-value (based on multiple comparisons) for each PH-associated SNPs at the false discovery rate (FDR) of 5% significance.

**Differential gene expression profiling.** The high PH-associated genes with SNPs delineated at the major genomic regions harboring robust PH QTLs (validated by QTL and association mapping) were selected to determine their regulation by differential expression profiling. The RNA isolated from diverse vegetative and reproductive tissues (root, shoot, young leaf, shoot apical meristem, flower bud and young pod) of tall (ICC 8261) and dwarf/semi-dwarf (ICC 12299) parental accessions was amplified using the gene-specific primers through quantitative RT-PCR assay (as per Bajaj *et al.*<sup>32</sup>).

## References

- Griffiths, S. *et al.* Meta-QTL analysis of the genetic control of ear emergence in elite European winter wheat germplasm. *Theor. Appl. Genet.* **119**, 383–395 (2009).
- Podolska, G. Plant lodging, effects, and control. In *Encyclopedia of Agrophysics, Part of the series Encyclopedia of Earth Sciences Series* (ed. Gliński, J., Horabik, J. & Lipiec, J.), 609–610 (Springer, Netherlands, 2014).
- Peng, J. *et al.* 'Green revolution' genes encode mutant gibberellin response modulators. *Nature* **400**, 256–261 (1999).
- Spielmeier, W., Ellis, M. H. & Chandler, P. M. Semidwarf (*sd-1*), "green revolution" rice, contains a defective *gibberellin 20-oxidase* gene. *Proc. Natl. Acad. Sci. USA* **99**, 9043–9048 (2002).
- Kumar, J., Choudhary, A. K., Solanki, R. K. & Pratap, A. Towards marker-assisted selection in pulses: a review. *Plant Breed.* **130**, 297–313 (2011).
- Varshney, R. K. *et al.* Achievements and prospects of genomics-assisted breeding in three legume crops of the semi-arid tropics. *Biotechnol. Adv.* **31**, 1120–1134 (2013).
- Jain, M. *et al.* A draft genome sequence of the pulse crop chickpea (*Cicer arietinum* L.). *Plant J.* **74**, 715–729 (2013).
- Varshney, R. K. *et al.* Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. *Nat. Biotechnol.* **31**, 240–246 (2013).
- Parween, S. *et al.* An advanced draft genome assembly of a *desi* type chickpea (*Cicer arietinum* L.). *Sci. Rep.* **5**, 12806 (2015).
- Garg, R., Patel, R. K., Tyagi, A. K. & Jain, M. *De novo* assembly of chickpea transcriptome using short reads for gene discovery and marker identification. *DNA Res.* **18**, 53–63 (2011).
- Garg, R., Bhattacharjee, A. & Jain, M. Genome-scale transcriptomic insights into molecular aspects of abiotic stress responses in chickpea. *Plant Mol. Biol. Rep.* **33**, 388–400 (2014).
- Hiremath, P. J. *et al.* Large-scale development of cost-effective SNP marker assays for diversity assessment and genetic mapping in chickpea and comparative mapping in legumes. *Plant Biotechnol. J.* **10**, 716–732 (2012).
- Jhanwar, S. *et al.* Transcriptome sequencing of wild chickpea as a rich resource for marker development. *Plant Biotechnol. J.* **10**, 690–702 (2012).
- Agarwal, G. *et al.* Comparative analysis of *kabuli* chickpea transcriptome with *desi* and wild chickpea provides a rich resource for development of functional markers. *PLoS One* **7**, e52443 (2012).
- Singh, V. K., Garg, R. & Jain, M. A global view of transcriptome dynamics during flower development in chickpea by deep sequencing. *Plant Biotechnol. J.* **11**, 691–701 (2013).
- Kudapa, H. *et al.* Comprehensive transcriptome assembly of chickpea (*Cicer arietinum* L.) using Sanger and next generation sequencing platforms: development and applications. *PLoS One* **9**, e86039 (2014).
- Pradhan, S. *et al.* Global transcriptome analysis of developing chickpea (*Cicer arietinum* L.) seeds. *Front. Plant Sci.* **5**, 698 (2014).
- Parida, S. K. *et al.* Development of genome-wide informative simple sequence repeat markers for large-scale genotyping applications in chickpea and development of web resource. *Front. Plant Sci.* **6**, 645 (2015).
- Nayak, S. N. *et al.* Integration of novel SSR and gene-based SNP marker loci in the chickpea genetic map and establishment of new anchor points with *Medicago truncatula* genome. *Theor. Appl. Genet.* **120**, 1415–1441 (2010).
- Gujaria, N. *et al.* Development and use of genic molecular markers (GMMs) for construction of a transcript map of chickpea (*Cicer arietinum* L.). *Theor. Appl. Genet.* **122**, 1577–1589 (2011).
- Thudi, M. *et al.* Novel SSR markers from BAC-end sequences, DArT arrays and a comprehensive genetic map with 1,291 marker loci for chickpea (*Cicer arietinum* L.). *PLoS One* **6**, e27275 (2011).

22. Gaur, R. *et al.* High-throughput SNP discovery and genotyping for constructing a saturated linkage map of chickpea (*Cicer arietinum* L.). *DNA Res.* **19**, 357–373 (2012).
23. Kujur, A. *et al.* Functionally relevant microsatellite markers from chickpea transcription factor genes for efficient genotyping applications and trait association mapping. *DNA Res.* **20**, 355–374 (2013).
24. Kujur, A. *et al.* An efficient and cost-effective approach for genic microsatellite marker-based large-scale trait association mapping: identification of candidate genes for seed weight in chickpea. *Mol. Breed.* **34**, 241–265 (2014).
25. Kujur, A. *et al.* Employing genome-wide SNP discovery and genotyping strategy to extrapolate the natural allelic diversity and domestication patterns in chickpea. *Front. Plant Sci.* **6**, 162 (2015).
26. Kujur, A. *et al.* 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 (2015).
27. Sabbavarapu, M. M. *et al.* Molecular mapping of QTLs for resistance to *Fusarium* wilt (race 1) and *Ascochyta* blight in chickpea (*Cicer arietinum* L.). *Euphytica* **193**, 121–133 (2013).
28. Varshney, R. K. *et al.* Fast-track introgression of “QTL-hotspot” for root traits and other drought tolerance traits in JG 11, an elite and leading variety of chickpea. *Plant Genome* **6**, 1–9 (2013).
29. Varshney, R. K. *et al.* Genetic dissection of drought tolerance in chickpea (*Cicer arietinum* L.). *Theor. Appl. Genet.* **127**, 445–462 (2014).
30. Varshney, R. K. *et al.* Marker-assisted backcrossing to introgress resistance to *Fusarium* wilt race 1 and *Ascochyta* blight in C 214, an elite cultivar of chickpea. *Plant Genome* **7**, 1–11 (2014).
31. Saxena, M. S. *et al.* An integrated genomic approach for rapid delineation of candidate genes regulating agro-morphological traits in chickpea. *DNA Res.* **21**, 695–710 (2014).
32. Bajaj, D. *et al.* Genome-wide conserved non-coding microsatellite (CNMS) marker-based integrative genomics for quantitative dissection of seed weight in chickpea. *J. Exp. Bot.* **66**, 1271–1290 (2015).
33. Bajaj, D. *et al.* A combinatorial approach of comprehensive QTL-based comparative genome mapping and transcript profiling identified a seed weight-regulating candidate gene in chickpea. *Sci. Rep.* **5**, 9264 (2015).
34. Thudi, M. *et al.* Genetic dissection of drought and heat tolerance in chickpea through genome-wide and candidate gene-based association mapping approaches. *PLoS One* **9**, e96758 (2014).
35. Das, S. *et al.* Deploying QTL-seq for rapid delineation of a potential candidate gene underlying major trait-associated QTL in chickpea. *DNA Res.* **22**, 193–203 (2015).
36. Deokar, A. A. *et al.* Genome-wide SNP identification in chickpea for use in development of a high density genetic map and improvement of chickpea reference genome assembly. *BMC Genomics* **15**, 708 (2014).
37. Jaganathan, D. *et al.* Genotyping-by-sequencing based intra-specific genetic map refines a “QTL-hotspot” region for drought tolerance in chickpea. *Mol. Genet. Genomics* **290**, 559–571 (2015).
38. Kujur, A. *et al.* A genome-wide SNP scan accelerates trait-regulatory genomic loci identification in chickpea. *Sci. Rep.* **5**, 11166 (2015).
39. Tar'an, B. *et al.* Quantitative trait loci for lodging resistance, plant height and partial resistance to *Mycosphaerella* blight in field pea (*Pisum sativum* L.). *Theor. Appl. Genet.* **107**, 1482–1491 (2003).
40. Blair, M. W., Iriarte, G. & Beebe, S. QTL analysis of yield traits in an advanced backcross population derived from a cultivated Andean × wild common bean (*Phaseolus vulgaris* L.) cross. *Theor. Appl. Genet.* **112**, 1149–1163 (2006).
41. Josie, J., Alcivar, A., Rainho, J. & Kassem, M. A. Genomic regions containing QTL for plant height, internodes length, and flower color in soybean [*Glycine max* (L.) Merr.]. *Bios* **78**, 119–126 (2007).
42. Robins, J. G., Bauchan, G. R. & Brummer, E. C. Genetic mapping forage yield, plant height, and regrowth at multiple harvests in tetraploid alfalfa (*Medicago sativa* L.). *Crop Sci.* **47**, 11–18 (2007).
43. Checa, O. E. & Blair, M. W. Mapping QTL for climbing ability and component traits in common bean (*Phaseolus vulgaris* L.). *Mol. Breed.* **22**, 201–215 (2008).
44. Tullu, A., Tar'an, B., Warkentin, T. & Vandenberg, A. Construction of an intraspecific linkage map and QTL analysis for earliness and plant height in lentil. *Crop Sci.* **48**, 2254–2264 (2008).
45. Palomeque, L. *et al.* QTL in mega-environments: II. Agronomic trait QTL co-localized with seed yield QTL detected in a population derived from a cross of high-yielding adapted × high-yielding exotic soybean lines. *Theor. Appl. Genet.* **119**, 429–436 (2009).
46. Kumawat, G. *et al.* Molecular mapping of QTLs for plant type and earliness traits in pigeonpea (*Cajanus cajan* L. Millsp.). *BMC Genet.* **13**, 84 (2012).
47. Rossi, M. E., Orf, J. H., Liu, L. J., Dong, Z. & Rajcan, I. Genetic basis of soybean adaptation to North American vs. Asian mega-environments in two independent populations from Canadian × Chinese crosses. *Theor. Appl. Genet.* **126**, 1809–1823 (2013).
48. Sun, Y. *et al.* Mapping and meta-analysis of height QTLs in soybean. *Legume Genomics Genet.* **3**, 1–7 (2012).
49. Gao, L.-F., Guo, Y., Hao, Z.-B. & Qiu, L. J. Integration and “overview” analysis of QTLs related to plant height in soybean. *Yi Chuan* **35**, 215–224 (2013).
50. Liu, Y.-L. *et al.* Identification of quantitative trait loci underlying plant height and seed weight in soybean. *Plant Genome* **6**, 1–11 (2013).
51. Stanton-Geddes, J. *et al.* Candidate genes and genetic architecture of symbiotic and agronomic traits revealed by whole-genome, sequence-based association genetics in *Medicago truncatula*. *PLoS One* **8**, e65688 (2013).
52. Kumar, B. *et al.* Population structure and association mapping studies for important agronomic traits in soybean. *J. Genet.* **93**, 775–784 (2014).
53. Mir, R. R. *et al.* Candidate gene analysis for determinacy in pigeonpea (*Cajanus* spp.). *Theor. Appl. Genet.* **127**, 2663–2678 (2014).
54. Li, F. *et al.* A genome-wide association study of plant height and primary branch number in rapeseed (*Brassica napus*). *Plant Sci.* **242**, 169–177 (2015).
55. Gowda, S. J. M. *et al.* Mapping of QTLs governing agronomic and field traits in chickpea. *J. Appl. Genet.* **52**, 9–21 (2011).
56. Hamwih, A., Intiaz, M. & Malhotra, R. S. Multi-environment QTL analyses for drought-related traits in a recombinant inbred population of chickpea (*Cicer arietinum* L.). *Theor. Appl. Genet.* **12**, 1025–1038 (2013).
57. Jamalabadi, J. G., Saidi, A., Karami, E., Kharkesh, M. & Talebi, R. Molecular mapping and characterization of genes governing time to flowering, seed weight, and plant height in an intraspecific genetic linkage map of chickpea (*Cicer arietinum*). *Biochem. Genet.* **51**, 387–397 (2013).
58. Karami, E., Talebi, R., Kharkesh, M. & Saidi, A. A linkage map of chickpea (*Cicer arietinum* L.) based on population from ILC3279 × ILC588 crosses: location of genes for time to flowering, seed size and plant height. *Genetika* **47**, 253–263 (2015).
59. Stephens, A. *et al.* Genetic marker discovery, intraspecific linkage map construction and quantitative trait locus analysis of *Ascochyta* blight resistance in chickpea (*Cicer arietinum* L.). *Mol. Breed.* **33**, 297–313 (2014).
60. Gaur, R. *et al.* High density linkage mapping of genomic and transcriptomic SNPs for synteny analysis and anchoring the genome sequence of chickpea. *Sci. Rep.* **5**, 13387 (2015).
61. Khajuria, Y. P. *et al.* Development and integration of genome-wide polymorphic microsatellite markers onto a reference linkage map for constructing a high-density genetic map of chickpea. *PLoS One* **10**, e0125583 (2015).
62. Marienfeld, J. R. & Newton, K. J. The maize NCS2 abnormal growth mutant has a chimeric *nad4-nad7* mitochondrial gene and is associated with reduced complex I function. *Genetics* **138**, 855–863 (1994).



63. Jenner, H. L. *et al.* NAD malic enzyme and the control of carbohydrate metabolism in potato tubers. *Plant Physiol.* **126**, 1139–1149 (2001).
64. Welchen, E. & Gonzalez, D. H. Differential expression of the *Arabidopsis* cytochrome *c* genes *Cytc-1* and *Cytc-2*, Evidence for the involvement of TCP-domain protein-binding elements in anther- and meristem-specific expression of the *Cytc-1* gene. *Plant Physiol.* **139**, 88–100 (2005).
65. Tronconi, M. A. *et al.* *Arabidopsis* NAD-malic enzyme functions as a homodimer and heterodimer and has a major impact on nocturnal metabolism. *Plant Physiol.* **146**, 1540–1552 (2008).
66. Welchen, E., Hildebrandt, T. M., Lewejohann, D., Gonzalez, D. H. & Braun, H. P. Lack of cytochrome *c* in *Arabidopsis* decreases stability of Complex IV and modifies redox metabolism without affecting Complexes I and III. *Biochem. Biophys. Acta* **1817**, 990–1001 (2012).
67. Wydro, M. M. *et al.* The evolutionarily conserved iron-sulfur protein INDH1 is required for complex I assembly and mitochondrial translation in *Arabidopsis*. *Plant Cell* **25**, 4014–4027 (2013).
68. Palmer, J. M. The organization and regulation of electron transport in plant mitochondria. *Ann. Rev. Plant Physiol.* **27**, 133–157 (1976).
69. Millar, A. H., Whelan, J., Soole, K. L. & Day, D. A. Organization and regulation of mitochondrial respiration in plants. *Annu. Rev. Plant Biol.* **62**, 79–104 (2011).
70. Fukazawa, J. *et al.* Repression of shoot growth, a *bZIP* transcriptional activator, regulates cell elongation by controlling the level of gibberellins. *Plant Cell* **12**, 901–915 (2000).
71. Ishida, S., Fukazawa, J., Yuasa, T. & Takahashi, Y. Involvement of 14-3-3 signaling protein binding in the functional regulation of the transcriptional activator REPRESSION OF SHOOT GROWTH by gibberellins. *Plant Cell* **16**, 2641–2651 (2004).
72. Bai, S. & Chaney, W. Gibberellin synthesis inhibitors affect electron transport in plant mitochondria. *Plant Growth Regul.* **35**, 257–262 (2001).
73. Welchen, E. *et al.* A segment containing a G-box and an ACGT motif confers differential expression characteristics and responses to the *Arabidopsis* *Cytc-2* gene, encoding an isoform of cytochrome *c*. *J. Exp. Bot.* **60**, 829–845 (2009).
74. Elshire, R. J. *et al.* A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* **6**, e19379 (2011).

### Acknowledgements

The authors gratefully acknowledge the financial support by the core grant of National Institute of Plant Genome Research (NIPGR), New Delhi, India. Alice Kujur acknowledges the CSIR (Council of Scientific and Industrial Research) for Senior Research Fellowship award.

### Author Contributions

A.K. and H.D.U. conducted all experiments and drafted the manuscript. A.K. and D.B. involved in genotyping-by-sequencing and data analysis. H.D.U., S.T., C.L.L.G. and S.S. helped in generation of mapping population and performed their phenotyping. S.K.P. and A.K.T. conceived and designed the study, guided data analysis and interpretation, participated in drafting and correcting the manuscript critically and gave the final approval of the version to be published. All authors have read and approved the final manuscript.

### Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Kujur, A. *et al.* Identification of candidate genes and natural allelic variants for QTLs governing plant height in chickpea. *Sci. Rep.* **6**, 27968; doi: 10.1038/srep27968 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>