



SSR markers in revealing extent of genetic diversity and phylogenetic relationships among chickpea core collection accessions for Western Himalayas

Asma Hamid Mir¹ · Mohd Ashraf Bhat¹ · Humara Fayaz^{1,2} · Aijaz A. Wani² · Sher A. Dar³ · Showkat Maqbool⁴ · Mohammad Yasin⁵ · Javid Iqbal Mir⁶ · Mohd Anwar Khan¹ · Parvaze A. Sofi¹ · Ahmed H. El-Sappah⁷ · Mahendar Thudi^{8,9} · Rajeev Kumar Varshney^{8,10} · Reyazul Rouf Mir¹

Received: 12 February 2022 / Accepted: 11 August 2022 / Published online: 25 August 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

Background The exploration of genetic diversity is the key source of germplasm conservation and potential to broaden its genetic base. The globally growing demand for chickpea suggests superior/climate-resilient varieties, which in turn necessitates the germplasm characterization to unravel underlying genetic variation.

Methodology and results A chickpea core collection comprising of diverse 192 accessions which include cultivated *Cicer arietinum*, and wild *C. reticulatum*, *C. echinospermum*, and *C. microphyllum* species were investigated to analyze their genetic diversity and relationship, by assaying 33 unlinked simple sequence repeat (SSR) markers. The results amplified a total of 323 alleles (Na), ranging from 2 to 8 with an average of 4.25 alleles per locus. Expected heterozygosity (He) differed from 0.46 to 0.86 with an average of 0.68. Polymorphic information content (PIC) ranged from 0.73 to 0.98 with an average of 0.89. Analysis of molecular variance (AMOVA) showed that most of the variation was among individuals (87%). Cluster analysis resulted in the formation of four distinct clusters. Cluster I represented all cultivated and clusters II, III, and IV comprised a heterogeneous group of cultivated and wild chickpea accessions.

Conclusion We report considerable diversity and greater resolving power of SSR markers for assessing variability and interrelationship among the chickpea accessions. The chickpea core is expected to be an efficient resource for breeders for broadening the chickpea genetic base and could be useful for selective breeding of desirable traits and in the identification of target genes for genomics-assisted breeding.

Keywords Chickpea · Core collection · Genetic diversity · Microsatellites · Cluster analysis

✉ Reyazul Rouf Mir
imrouf2006@gmail.com; rrmir@skuastkashmir.ac.in

¹ Division of Genetics and Plant Breeding, Faculty of Agriculture (FoA), SKUAST-Kashmir, Wadura Campus, Sopore, Kashmir, India

² Department of Botany, Cytogenetics and Reproductive Biology Laboratory, University of Kashmir, Srinagar, Kashmir, India

³ Dryland Agriculture Research Station (DARS), SKUAST-Kashmir, Budgam, Kashmir, India

⁴ Division of Animal Genetics and Breeding, SKUAST-Kashmir, FVSc & AH, Shuhama, Srinagar, Kashmir, India

⁵ Rafi Amhad Kidwai (RAK) College of Agriculture, Rajmata Vijayaraje Scindia Krishi Vishwa Vidyalyaya (RVSKV), Sehore, Madhya Pradesh, India

⁶ ICAR-Central Institute of Temperate Horticulture (CITH), Srinagar, Kashmir, India

⁷ School of Agriculture, Forestry and Food Engineering, Yibin University, Yibin, China

⁸ Center of Excellence in Genomics & Systems Biology (CEGSB), International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Hyderabad, Telangana, India

⁹ Department of Agricultural Biotechnology and Biotechnology, Rajendra Prasad Central Agricultural University, Pusa, Samastipur, Bihar, India

¹⁰ State Agricultural Biotechnology Centre, Crop & Food Innovation Centre, Food Futures Institute, Murdoch University, Murdoch, WA, Australia

Introduction

Chickpea (*Cicer arietinum* L.) remains as the second most important cool season legume crop in the world [1] offering a cheap source of dietary protein (23% in dry seed), and provides nutritional, and income security in countries like India, Pakistan, and Bangladesh among many other developing countries [2]. It contributes 14.78 million tons to global food production from 14.56 million hectares area across the world with average productivity of 0.96 tons per hectare [3]. In India, chickpea represents national pulse acreage, and production of 35 and 46%, respectively. India serves as the largest chickpea producer (69%), consumer, and importer of chickpea globally [4]. The genome sequence of both Kabuli [5], and Desi type [6], whose advanced version was later published by [49] are available thus making chickpea a candidate crop for legume genetic, and genomic research.

The genetic variability for the traits of economic importance, and the relationship between the cultivated, and its wild relatives are of paramount importance for advanced chickpea programs. Domestication, modern plant breeding, and agricultural systems have narrowed the genetic base of cultivated chickpea. This has promoted the need for new sources of variation arising within the germplasm that might be useful in plant breeding programs [5]. However, the large size of germplasm collections, and its inadequate data/information, have resulted in low use (<1%) of germplasm further narrowing the genetic base in many crops [7]. India is known for its largest diversity in cultivated chickpea [24]. Despite the availability of a diverse, and large number of chickpea germplasm collection at the Gene banks like the National Bureau of Plant Genetic Resources (NBPGR), and International Crop Research Institute of Semi-arid Tropics (ICRISAT), a very small proportion has been used in crop improvement/breeding programs mainly because the resources available for evaluation of such large collection is quite limited. The core approach which represents the mini crop collections capturing most of the available genetic diversity in the whole germplasm collection, is an efficacious methodology to advance crop improvement programs. Such core sets are cost-effective, easy to maintain by individual breeders, and help in the appropriate assessment of genetic diversity, population structure, association analysis, and targeted gene mining. Many national programs around the world are actively involved in the utilization of mini cores of crops like chickpea, pigeonpea, groundnut, foxtail millet sorghum, pearl millet, and finger millet, which have led to new and diverse sources of variation for, qualitative, and quantitative (biotic/abiotic stress resistance) traits in various crops.

The molecular characterization of the core collection will further enhance its use in genomics for crop

improvement programs. Molecular markers can explore the genetic variation in plants, and are useful in detecting the duplicates among the core collection, and thereby help in designing efficient strategies for optimal sampling from the existing variation. The utility of DNA-based markers for unbiased estimation of molecular diversity and establishing precise phylogenetic relationships among species as compared to morphological, cytological and biochemical markers have been well understood. In chickpea, diverse molecular markers including AFLPs [8], STMS [9], RAPDs [11], ISSRs [12], and SSRs [13–16] have been already used to unravel the genetic diversity available in chickpea. The desirable genetic attributes (co-dominant inheritance, reproducibility, bi-/multi-allelic nature, and abundant genomic distribution) of sequence-based robust microsatellite/SSR markers have encouraged their use for many applications of chickpea genetics, genomics, and breeding including cultivar identification, allele mining, genetic mapping, association studies, genetic diversity analysis, population structure studies, and establishing phylogenetic relationships [32, 50]. The draft genome sequence of chickpea has led to the identification of over 48,000 SSRs for use as genetic markers [5].

Our recently published information of intensive screening and genetic evaluation of the composite set of 384 chickpea germplasm has led to the identification of chickpea core collection, which would enhance the utilization of chickpea genetic resources in crop improvement programs [10]. The present investigation was aimed to further evaluate the level of genetic diversity and allelic richness in the above selected chickpea core set of 192 accessions, using a larger number of thirty-three genome-wide SSR markers. The results derived from this study could be used for more efficient utilization of the entire collection. The microsatellite markers used in the present study were also tested for their ability to detect the extent of genetic diversity by using polymorphism information content values.

Methods

Germplasm collection

Experimental material for the study comprised of a core set of 192 chickpea accessions, consisting of landraces, elite cultivars, advanced breeding lines, and wild species from the primary, secondary, and tertiary gene pools. These lines include annual cultivated 178 *Cicer arietinum* lines from the three seed types (kabuli, desi, and intermediate pea-shaped) and the fourteen (14) wild relatives belonging to *Cicer reticulatum* Ladiz. (8), *Cicer echinospermum* P. H. Davis (5), and a perennial *Cicer microphyllum* Benth (1) species. The seed material of accessions from primary and

secondary gene pools has been collected from different hot spots in Turkey under USAID project ‘‘Chickpea Innovation Lab’’, and the same has been procured by us from ICRISAT, Hyderabad, India. And tertiary gene pool has been collected from its wild natural habitat in Leh/Ladakh region of state Jammu and Kashmir, India. The seed material of cultivated chickpea has been attained from various ethnic/ global institutes. (Supplementary Table EMS_1). Details of some of the accessions like their cold tolerance/ nutrient density are available elsewhere [17, 18]. The passport data of the species is presented in Table 1.

DNA isolation

Young leaf tissues of each accession were lyophilized to extract the total genomic DNA following CTAB-DNA extraction protocol [19] with some minor modifications. The extracted DNA was purified by treating with RNase to remove the RNA contamination. The purified DNA was quantified, and its quality was assessed by 0.8% agarose gel electrophoresis). Based on the intensity of the bands, DNA was diluted for marker genotyping.

SSR markers

A total of 33 SSR markers spanning the entire chickpea genome, one marker at least covering each arm of the 8 linkage groups, were selected to characterize the set of 192 core chickpea accessions (Fig. 1). These SSR primer pairs were selected from the composite collections of the chickpea database (CicArMiSatDB: the chickpea microsatellite database (<https://cegresources.icrisat.org/CicArMiSatDB/>) developed by ICRISAT. The database contains SSRs which are reported by [5]. The details of SSR primer-pairs are provided in Supplementary Table EMS_2.

Polymerase chain reaction (PCR)

To study polymorphism in the chickpea core collection, PCR amplification was carried out in a Universal Gradient Thermal Cycler with 96 wells using a 20- μ L reaction mixture: 2 μ L of template DNA (25 ng/ μ L), 10 \times PCR buffer (1.8 mM, MgCl₂, 10 mM Tris-HCl, 50 mM KCl), 2.5 mM dNTPs (Molecular Biology for Life Science, Fermentas, Lithuania, USA), 5 μ M each of forward and reverse primer and 5U of Taq DNA polymerase (Sigma Aldrich USA. The thermal cycler (Peqlab) was programmed as follows: initial denaturation for 5 min at 94 °C followed by 40 cycles of denaturation for 30 s (s) at 94 °C, annealing at temperature specific for each primer pair for 30 s and extension at 72 °C for 30 s. The final extension was allowed for 10 min at 72 °C and storage at 4 °C until further use. The resulting PCR products were run on 10% Poly-acrylamide Gel Electrophoresis (PAGE) using a Dual Gel Vertical Electrophoresis System (Peqlab) and silver stained for manual visualization of bands on gel documentation system.

Data analysis

To calculate genetic diversity parameters GenALex software version 6.5b3 [20] was used. The overall diversity, number of different alleles (N_a), the effective number of alleles (N_e), number of private alleles, and gene diversity measured as expected heterozygosity (H_e), was measured along with other parameters. Total genetic variation was partitioned among populations, among individuals, and within individuals, by performing an analysis of molecular variance (AMOVA). Allelic frequency was calculated using Polymorphism Information Content (PIC) values which indicates the ability of markers to differentiate the genotypes was calculated as $1 - \sum (f_{ij})^2$ where f_{ij} is the frequency of the j th allele for i th locus summed across all alleles for the locus [21]. For the calculation of the pairwise genetic distances, DARwin

Table 1 Some passport information on *Cicer* accessions used in the present study

S. No	Species	Biological status	Breeding cycle	Gene pool	Inbreeding	Origin	Geographical distribution	Latitude range (°N)
1	<i>C. arietinum</i>	Cultivated	Annual	1°	Inbreeder	India	North, Central and Peninsular India	8.7–40.0
2	<i>C. reticulatum</i> Ladizinsky	Wild	Annual	1°	Inbreeder	Turkey	East Turkey (South East Anatolia)	37.3–38.1
3	<i>C. echinospermum</i> P.H.Davis	Wild	Annual	2°	Inbreeder	Turkey and Syria	East Turkey (East Anatolia), North Iraq	36.9–37.9
4	<i>C. microphyllum</i> Benth	Wild	Perennial	3°	Outbreeder	India and Turkey	North India (hot spots of Leh/ Ladakh)	34.1–35.0

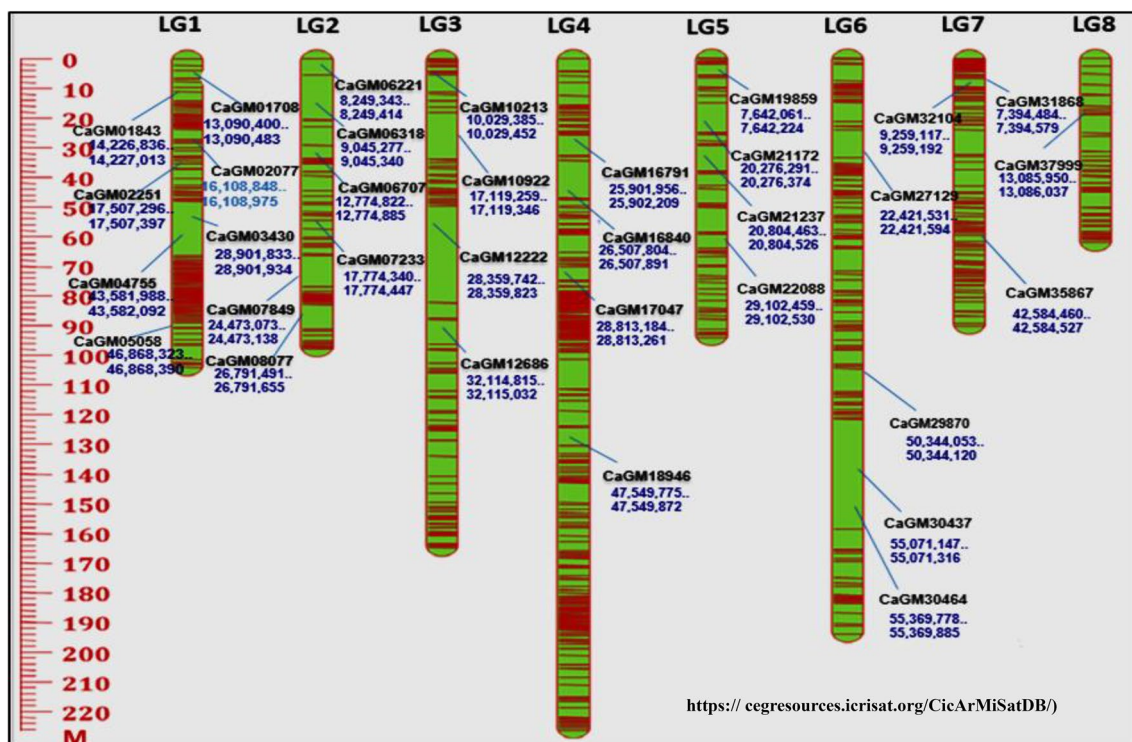


Fig. 1 The distribution of 33 SSR markers on different chickpea chromosomes used during the present study

version 5.0 was used [22]. The dissimilarity matrix obtained was subjected to cluster analysis, using the Unweighted Pair Group Method with Arithmetic Average (UPGMA) based hierarchical clustering for the construction of dendrograms.

Results

A set of 33 primer pairs physically mapped across eight chickpea chromosomes showing reproducible amplification, and 100% polymorphism were genotyped across the core collection of 192 cultivated and wild *Cicer* accessions using the polyacrylamide gel-based assay. The SSR markers used in this study were found informative enough to explain genetic variation among core chickpea accessions.

Allelic and gene diversity

A total of 323 alleles (N_a) were identified at 76 marker loci generated by 33 SSR markers giving an average of 4.25 alleles per marker locus. The microsatellite marker locus CaGM06707A (mapped on LG2) was the most variable with maximum (eight) 8 alleles and microsatellite marker locus CaGM06221B was least variable with only (two) 2 alleles. The total number of effective alleles (N_e) was 263.96 with an average of 3.47 alleles/locus. The number of effective alleles was lower than the total number of alleles, showing

that only a few alleles contributed to the variation. The average number of private alleles i.e., the number of alleles unique to single population was 4.25.

Gene diversity (expected heterozygosity; H_e) was variable across loci reflecting the different number and frequencies of the alleles found. The average gene diversity of the core collection was found moderate (0.68), showing that two randomly chosen genes will carry different alleles roughly half of the time. The lowest H_e (0.46) was recorded for SSR marker locus CaGM12686E and the highest H_e (0.86) for SSR marker locus CaGM06707A. For each population, the observed heterozygosities (H_o) were less than the expected heterozygosities (H_e) (Table 2, Fig. 2).

Allelic/gene diversity across sub-populations (cultivated and wild)

The allelic/genetic patterns studied across the populations are presented in Table 2. The allelic diversity, when studied within two sub-populations (cultivated v/s wild), revealed a total number of 345 alleles with an average of 4.25 alleles/locus observed in the cultivated sub-population, as against a total number of 201 alleles with an average of 3.28 alleles/locus in wild sub-population. The average number of effective alleles (N_e) was observed to be 3.46 in cultivated chickpea, and that of 2.60 in wild chickpea. The number of alleles with frequency ≥ 5 percent

Table 2 Comparison of diversity estimates in the whole population and two sub-populations (cultivated and wild) of core chickpea accessions

	Whole population	Sub-population (Cultivated)	Sub-population (Wild)
Diversity parameters	Mean		
N	191.605	177.605	13.052
Na	4.250	4.250	3.289
Na Freq. > =5%	4.026	4.026	3.263
Ne	3.473	3.468	2.602
Number of private alleles	4.250	0.961	0.000
He	0.687	0.686	0.572
UHe	0.688	0.688	0.592

*N = sample size, Na = no. of different alleles, Na = Freq. > =5% number of different alleles with a frequency > =5%, Ne = no. of effective alleles, He = expected heterozygosity, UHe = unbiased expected heterozygosity

(Na Freq. \geq 5%), (i.e., measurement taken to alleviate the sampling error associated with the sampling of distinct alleles) was for cultivated and wild sub-populations was 4.02 and 3.26, respectively. The average number of private alleles (alleles unique to a single population) was 0.96 in the cultivated sub-population. However, no private allele was found among the wild sub-population. Since the sample size between the two sub-populations was not equal, the resampling technique was therefore used, to make valid comparisons of diversity, wherein a small sample size i.e., 14 for wilds was drawn randomly 3000 times from the 178

cultivated group and average mean alleles were recorded. The mean alleles/locus for the cultivated group based on a sample size of 14 genotypes was found to be 3.65. The more diversity apparent in cultivated sub-population may be attributed to the presence of private alleles in the group (mean private alleles = 0.96), while no private alleles were present in the wild sub-population. Similarly, gene diversity (expected heterozygosity; He) was observed to be higher in cultivated sub-population and varied from 0.14 to 0.80 with an average of 0.68, and from 0.09 to 0.75 with an average of 0.57 in wilds (Fig. 2).

Polymorphic information content (PIC)

The PIC value provides an idea about the polymorphism detected by the marker system and its suitability for the study of genetic diversity. The higher the PIC value of a locus, the higher the number of alleles detected. During the present study, the PIC value of all the 76 polymorphic SSR marker loci for all 192 chickpea core set varied from 0.73 to 0.98 with an average of 0.89. Interestingly, all the marker loci were highly informative with the PIC value of > 0.50. Markers *CaGM06707A*, *CaGM12686E*, and *CaGM19859B* were most informative with high PIC value of 0.98, 0.98, and 0.97, respectively. SSR marker *CaGM12686E* and *CaGM06221B* showed the least PIC value of 0.72 and 0.73, respectively. (Table 3). The correlation worked out between the PIC value and the number of alleles was significantly positive ($r = 0.67$, $p < 0.01$).

Fig. 2 Comparison of allelic and gene diversity detected in the whole chickpea population and two sub-populations (cultivated vs wild). The figure shows comparison for number of alleles detected (Na), Number of alleles with frequency more than or equal to 5%, Number of effective alleles (Ne) and Number of private alleles, etc.

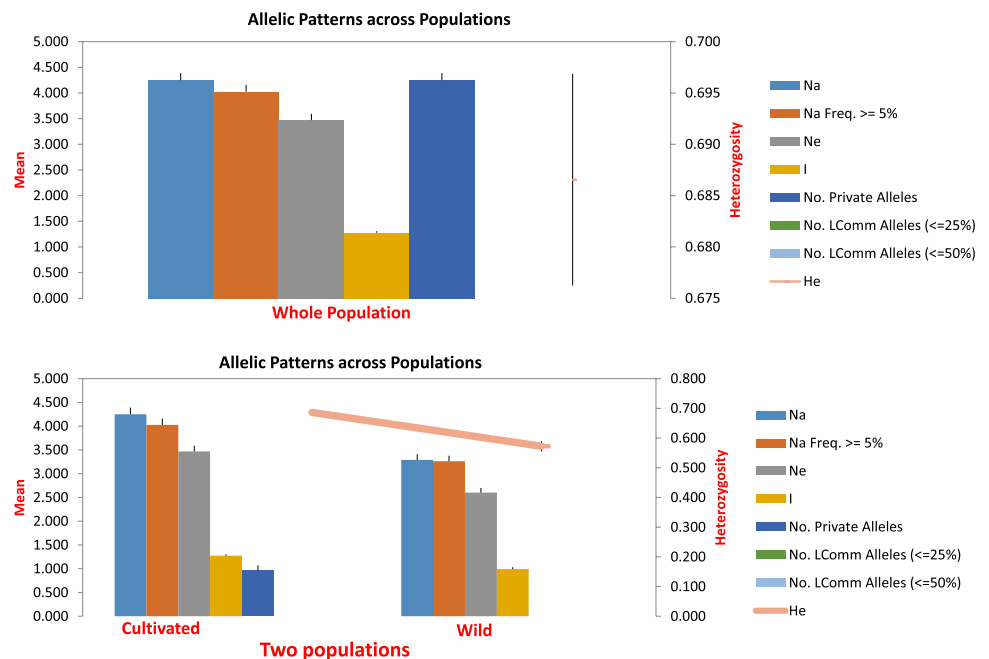


Table 3 Profile of 76 SSR marker loci detected during the present study across the core chickpea accessions

Marker Locus	Na	He	PIC	Marker Locus	Na	He	PIC
<i>CaGM01708A</i>	3	0.65	0.88	<i>CaGM17047A</i>	3	0.65	0.88
<i>CaGM01708B</i>	6	0.77	0.95	<i>CaGM17047B</i>	5	0.78	0.95
<i>CaGM01708C</i>	4	0.70	0.91	<i>CaGM18946A</i>	4	0.74	0.93
<i>CaGM01843A</i>	5	0.53	0.78	<i>CaGM18946B</i>	4	0.66	0.89
<i>CaGM01843B</i>	3	0.58	0.83	<i>CaGM19859A</i>	4	0.66	0.89
<i>CaGM02077A</i>	4	0.64	0.87	<i>CaGM19859B</i>	7	0.83	0.97
<i>CaGM02077B</i>	3	0.62	0.86	<i>CaGM19859C</i>	4	0.74	0.93
<i>CaGM02077C</i>	4	0.70	0.91	<i>CaGM21172A</i>	4	0.72	0.92
<i>CaGM05058A</i>	4	0.68	0.90	<i>CaGM21172B</i>	7	0.80	0.96
<i>CaGM05058B</i>	3	0.57	0.82	<i>CaGM21237A</i>	4	0.65	0.88
<i>CaGM06221A</i>	4	0.68	0.90	<i>CaGM21237B</i>	3	0.65	0.88
<i>CaGM06221B</i>	2	0.47	0.73	<i>CaGM21237C</i>	3	0.62	0.86
<i>CaGM06707A</i>	8	0.86	0.98	<i>CaGM22088A</i>	4	0.74	0.93
<i>CaGM06707B</i>	5	0.77	0.95	<i>CaGM22088B</i>	6	0.77	0.95
<i>CaGM06707C</i>	4	0.73	0.93	<i>CaGM2251A</i>	4	0.73	0.93
<i>CaGM06707D</i>	3	0.65	0.88	<i>CaGM25867A</i>	5	0.73	0.93
<i>CaGM07849A</i>	4	0.74	0.94	<i>CaGM25867B</i>	4	0.74	0.94
<i>CaGM07849B</i>	4	0.73	0.93	<i>CaGM27129A</i>	3	0.62	0.86
<i>CaGM07849C</i>	5	0.78	0.95	<i>CaGM27129B</i>	4	0.66	0.89
<i>CaGM08077A</i>	6	0.82	0.97	<i>CaGM30464A</i>	3	0.52	0.77
<i>CaGM08077B</i>	6	0.70	0.97	<i>CaGM30464B</i>	5	0.71	0.92
<i>CaGM08077C</i>	5	0.59	0.91	<i>CaGM30464C</i>	4	0.69	0.90
<i>CaGM08077D</i>	3	0.66	0.83	<i>CaGM31868A</i>	4	0.64	0.88
<i>CaGM10213A</i>	3	0.65	0.88	<i>CaGM31868B</i>	6	0.74	0.94
<i>CaGM10213B</i>	4	0.74	0.94	<i>CaGM32104A</i>	3	0.58	0.83
<i>CaGM10213C</i>	6	0.82	0.97	<i>CaGM32104B</i>	3	0.58	0.83
<i>CaGM10922A</i>	3	0.55	0.80	<i>CaGM32104C</i>	5	0.74	0.93
<i>CaGM10922B</i>	3	0.61	0.85	<i>CaGM3430A</i>	3	0.65	0.88
<i>CaGM10922C</i>	7	0.78	0.96	<i>CaGM37999A</i>	5	0.74	0.94
<i>CaGM12686A</i>	7	0.80	0.96	<i>CaGM37999B</i>	4	0.71	0.92
<i>CaGM12686B</i>	3	0.54	0.79	<i>CaGM4755A</i>	3	0.66	0.89
<i>CaGM12686C</i>	3	0.48	0.74	<i>CaGM6470A</i>	3	0.64	0.87
<i>CaGM12686D</i>	4	0.65	0.88	<i>CaGM7233A</i>	4	0.74	0.96
<i>CaGM12686E</i>	4	0.46	0.72	<i>CaGM30437A</i>	6	0.78	0.96
<i>CaGM16791A</i>	4	0.71	0.92	<i>CaGM29870A</i>	3	0.56	0.81
<i>CaGM16791B</i>	4	0.64	0.87	<i>CaGM29870B</i>	4	0.70	0.91
<i>CaGM16791C</i>	4	0.70	0.91	<i>CaGM29870C</i>	4	0.66	0.89
<i>CaGM16840A</i>	4	0.69	0.90	<i>CaGM27129A</i>	5	0.55	0.80

**Na* Number of different alleles, *He* Expected Heterozygosity, *PIC* Polymorphic information content

Analysis of molecular variance (AMOVA)

AMOVA was carried out among and within the two sub-populations (cultivated and wild). The results of the analysis showed that the observed variance was mostly partitioned among individuals (87%), whereas 7 percent of the variance was partitioned within individual accessions and the rest 6 percent was among populations (Table 4). The low level of genetic variation among populations is indicative of gene flow between populations. Since the maximum number (8)

of the wilds in the present study are from *C. reticulatum* which is the progenitor of cultivated species *C. arietinum* leading to smooth introgression of genes. There is a moderate amount of genetic differentiation between the groups ($F_{ST}=0.10$), indicating that the groups are less genetically distinct. The F_{IS} and F_{IT} values are 0.92 and 0.82, respectively indicating that the chickpea accessions making up the individual groups are inbred lines. Furthermore, the gene flow (Nm) between the groups was observed to be 4.26, suggesting a high level of gene flow.

Table 4 Analysis of molecular variance for variation among and within sub-population based on 33 SSRs

Source	df	SS	MS	Estimated Variance	% variance
Among Populations	1	134.50	134.50	1.52	6
Among Individuals	190	9517.43	50.09	24.04	87
Within Individuals	192	384.00	2.000	2.00	7
Total	383	10,035.93		27.57	100

* *df* degree of freedom, *SS* sum of squares, *MS* mean sum of squares

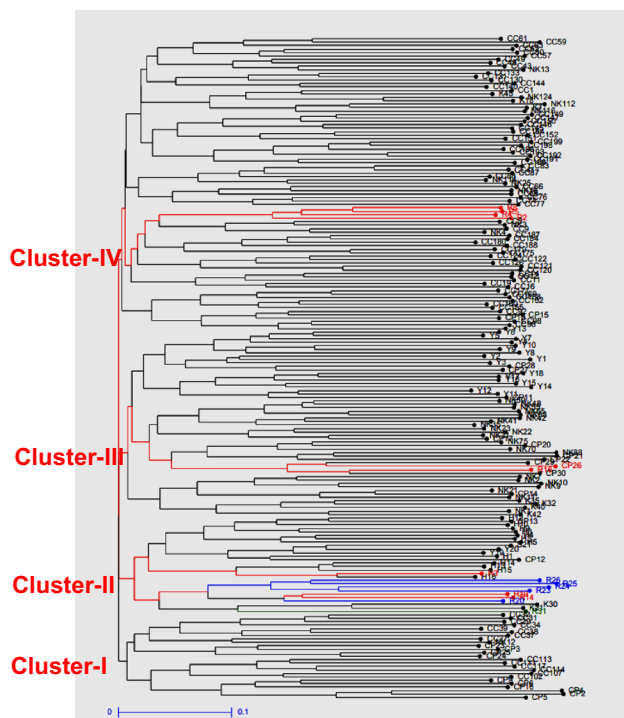


Fig. 3 Hierarchical clustering based on simple matching dissimilarity matrix genotyped across 192 core chickpea accessions using 33 SSR marker; Black lines indicate cultivated, Red ones are *C. reticulatum*, blue ones are *C. echinospermum* and green is *C. microphyllum*

Cluster analysis

Cluster analysis based on the dissimilarity index values delineated 192 chickpea accessions into 4 main clusters by employing all the 76 marker loci, the results of which are presented in Fig. 3. The cluster-I contained 25 accessions which were all cultivated species. Cluster-II possessed 29 accessions, which include wilds from the three gene pools (*C. reticulatum* (3), *C. echinospermum* (5), and *C. microphyllum* (1)) and the remaining 20 in the cluster were cultivated *C. arietinum*. A slight speciation pattern of structure was detected especially in cluster-II, where the wild accessions

belonging to three different gene pools were grouped in close vicinity. Cluster-III had a total of 53 accessions wherein all cultivated except 1 wild species (*C. reticulatum*) were included. Cluster-IV was the largest cluster with 85 accessions which included four *C. reticulatum* species and 81 cultivated species. These results infer that wild *C. reticulatum* belonging to primary gene pool was present with *C. arietinum* (cultivated) in three out of four clusters, owing to the origin of cultivated *C. arietinum* from the wild progenitor. Also, among all the four clusters, the cluster II was found to be the most heterogeneous.

Discussion

An extensive characterization of plant genetic resources, and an understanding of the genetic relationships in the germplasm collection is essential for effective conservation, management, and exploitation of genetic resources in varietal improvement programs. Also, the use of wild species provides a wider genetic base in otherwise, less diverse cultivated chickpea and is known as a potential source of resistance genes for various biotic and abiotic stresses [23]. Thus, the investigation of the nature and magnitude of genetic diversity and relatedness within and among the cultivated chickpea and its wild relatives is an obvious necessity to identify new sources of variation in the existing germplasm. A core set of chickpea acts as a working collection having an optimal and convenient size that can be evaluated for all important traits with minimum error rates. Intensive screening and evaluation of the core collections have led to the identification of diverse accessions leading to the advancement of plant breeding and thus paving the way towards booming crop improvement programs. To cite an example, the utility of the chickpea core set (300 accessions) capturing 78.1% alleles of the composite collection, developed at ICRISAT [25] has been extensively used in breeding as well as germplasm management and improvement. The molecular characterization of the core collection provides information related to unique/rare alleles from cultivated and wild species which could be used to select distinct genotypes for allele mining. Among the variety of molecular markers already developed and used in breeding programs, SSRs are considered markers of choice [26, 27]. Universal distribution and high density in a multitude of genomes along with other merits associated with SSR markers enhance their desirability, especially for germplasm characterization and have been extensively used for identification of variation in chickpea germplasm. Therefore, efforts were made during the present study to carefully select random SSR markers covering all the eight linkage groups of chickpea genome.

In this study, every SSR primer pair successfully amplified the target DNA in a core set of 192 chickpea accessions.

The 33 random markers generated 76 polymorphic loci and were useful for detection of a total of 323 alleles (Na) at all marker loci. The average number of alleles/locus (4.25) found in this study are within the ranges reported by many earlier studies. A relatively higher diversity was observed in our core collection when compared to our previous study on genetic diversity in a composite set of 384 genotypes using 8 SSR markers which generated a total of 63 (Na) alleles with an average of 3.7/locus, indicating that the chickpea core maximized the diversity available in the entire composite germplasm set. Our observations are fairly comparable with earlier genetic diversity studies on the cultivated and wild *Cicer* accessions using random microsatellite markers with some deviations. For instance, the use of 478 microsatellites identified a total of 3703 alleles with an average of 4.26 alleles/marker locus in 94 cultivated and wild chickpea accession [28]. Similarly, 38 accessions of *Cicer arietinum* using 100 SSR markers generated an average of 4.8 amplicons per locus [17]. Likewise, a total of 59 alleles with an average of 4.2 alleles/locus in 60 chickpea accessions were reported using 14 SSR markers [15]. Similar results were also observed by [16] where he reported an average of 3.88 alleles/locus in 66 *Cicer arietinum* accessions. However, the level of diversity detected in the present study was less while comparing it with average number of alleles/locus (35, 17, 10.3, 7.4, 8.61, 7.75, 10.5) detected by [13, 30–35]. There are also reports where even less diversity than in our case has been reported. For example, 33 microsatellites were used to study genetic diversity in 155 chickpea accessions which resulted in a total of 111 bands with a range of 2–5 alleles/locus and on average 3.364 bands per marker [14]. In another study, using 27 SSR markers a total of 81 alleles with an average of 3.0 alleles per locus were detected in 50 chickpea genotypes [36]. The use of more cultivated chickpea accessions which possess narrow genetic base as reported by various researchers [37–40] may have led to observance of less diversity in our study. Likewise, higher polymorphic potential among the species/accessions originating from Fertile Crescent (79.5% polymorphism) in contrast to those from Central Asia- India (35.8%) was also reported by [28]. Since most of the germplasm used in the present study is indigenous which may have attributed to the presence of less diversity in our chickpea core collection. The inclusion of more chickpea wild species could produce a higher number of allele count/locus as suggested by [15, 30]. The overall genetic diversity (He) showed relatively moderate value of expected heterozygosity (He=0.68) which are comparable to previous studies [15, 33, 41, 42]. Differences in results for estimated genetic diversity between studies may be attributed to the different number of accessions, different number of loci examined, nature of markers, and perhaps the geographical origin of the accessions used in each study. The inter and intra- specific polymorphism,

when studied within two sub-populations (cultivated vs wild), did not differ much from the single population of 192 chickpea accessions, where the average number of alleles per locus for cultivated sub-population was 4.25, while that for wild was found to be 3.28. These results indicate that the number of accessions could still be reduced to study genetic diversity. The mean private alleles in cultivated was (0.96) and no private alleles in wilds were detected within the two sub-populations. Similar results were shown by [30], where no rare alleles were found in wild accessions. Among the two sub-populations, cultivated chickpea showed more diversity (possessing more number of private alleles and an average number of alleles/locus than wild population, which could be due to very less number of wild accessions involved in our study as compared to the large number of cultivated accessions used. The inclusion of more wild accessions particularly from secondary and tertiary gene pools could result in a higher level of diversity since most of the desirable gene complexes are present in them. For instance, a higher degree of polymorphism was detected among the accessions belonging to the secondary gene pool as compared to primary gene pool [28].

The results of PIC values indicate that all the markers used in the present study were informative with a PIC value greater than 0.50. The average PIC value for the 76 marker loci was 0.89 which suggests that the majority of markers enabled a high level of polymorphism and could be effective to determine the genetic differences among the chickpea accessions studied. Similar to our results the PIC value of 0.854 has been reported by [30] and PIC values of 0.77, 0.75, 0.75, 0.72, were observed by [13, 29, 35, 43], respectively. Comparatively, [17, 36, 44, 45] revealed an average PIC value of 0.53, 0.60, 0.68, and 0.43, respectively which is relatively less than that of the present study.

AMOVA is a method to study population differentiation utilizing molecular markers [46]. To evaluate the extent of population differentiation among/between the sub-populations (cultivated and wild) in the present study, the molecular variance based on 99 permutations was analyzed using 33 random SSR markers. The analysis of the two sub-populations showed that most of the variation was among individuals (87%), and 7% of the variation was partitioned within individuals. The relatively higher variation among individuals could be attributed to the differences among genotypes within sub-populations. In agreement with our results, [47] reported higher variance (61%) among the populations compared to within (39%) populations. Similarly, 71% of the allelic variation was documented among the domesticated and the wild gene pools and the remaining 29% of the allelic variation was observed within them [48].

To decipher the phylogenetic relationships between accessions and closest and distant genotypes from each other, UPGMA based Hierarchical clustering approach was

found best in the present study for interpretation of results. The 33 random SSR markers could cluster chickpea core collection of chickpea accessions into 4 main clusters, and several sub-clusters and sub-sub clusters, whereby the different members within a cluster are assumed to be more closely related to each other than with those members in different clusters. From the present investigation, it was found that accessions of different origins/species were grouped together in different clusters. The *C. reticulatum* accessions fell in 3 clusters out of 4 main clusters. All accessions belonging to *C. echinospermum* were clustered in a single cluster (cluster II). Another feature of the tree is the relative position of the single *C. microphyllum* accession (belonging to tertiary gene pool) in close vicinity of *C. echinospermum* accessions (secondary gene pool). This may suggest that both species are closely related to each other (Fig. 3). Results from the present study support the observations of several workers [10, 51] the grouping followed a definite pattern with all the cultivated accessions falling into all distinct clusters while the wild species grouped into distinct subclusters and in closer vicinity to each other. Similarly, [8] used Nei's pairwise distance calculations and identified clustering of cultivated species *Cicer arietinum*, *Cicer reticulatum* and *Cicer echinospermum* in the same group. It can be inferred from this study that SSR markers were effective in placing chickpea accessions in appropriate clusters based on relationships at an evolutionary level. The allelic/genetic diversity, PIC values, molecular variation, phylogenetic relationship information generated in this study for the set of 192 chickpea core collection, and with the availability of current genomic tools, could have implications for many genotyping applications including allele mining, gene tagging association analysis and QTL mapping targeting different qualitative and quantitative traits of agricultural importance in chickpea.

Conclusion

The present study shows the potential use of selected polymorphic microsatellite markers in discriminate among the core set chickpea accessions. Our results revealed considerable genetic diversity with most of the variation found among the individuals and provided information on phylogenetic relationships among the members of different gene pools. The core collection is regarded as the gateway to unravel the diversity estimates. The results obtained in the present investigation revealed that the core collection exhibited equal diversity when compared with the entire composite chickpea germplasm from which it was derived. Thus, shall provide greater confidence to the breeding community for assessment of distinctiveness and relationships among the various accessions. The grouping of wild species with cultivated *C. arietinum* indicates that they could possibly serve

as good sources of broadening chickpea genetic base, and for effective pre-breeding for breaking yield barriers. The clustering also indicated different genetic basis of different gene pools. The diverse genotypes could be used as genetic resources for developing mapping populations and in transcriptomics studies. The chickpea core collection is expected to ensure effective core deployment, especially for future chickpea varietal identification, conservation, and management, and the development of broad-based/ climate resilient cultivars that should meet the immediate and future breeding challenges.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11033-022-07858-4>.

Acknowledgements The authors are thankful to ICRISAT (Hyderabad), IIPR (Kanpur), NBPGR (New Delhi) for providing the germplasm for the present study.

Funding This study was not funded by any external funding agency.

Data availability The data analyzed and generated during the current study are available within the article [and/or] its supplementary materials.

Declarations

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

Ethical approval This article does not contain any studies/ research with human or animal subjects.

References

- Mallikarjuna BP, Samineni S, Thudi M, Sajja SB, Khan AW et al (2017) Molecular mapping of flowering time major genes and QTLs in chickpea (*Cicer arietinum* L.). *Front Plant Sci* 8:1140
- Archak S, Tyagi RK, Harer PN, Mahase LB, Singh N, Dahiya OP et al (2016) Characterization of chickpea germplasm conserved in the Indian national genebank and development of a core set using qualitative and quantitative trait data. *Crop J* 4(5):417–424
- FAOSTAT, Food and Agriculture Organization of the United Nation Statistics for (2017) Online database at <https://www.fao.org/faostat/en/#data>
- Annual Report DPD. 2017. Directorate of Pulses Development, Ministry of Agriculture and Farmers Welfare, Government of India. DPD/Pub./TR/29/2017–18
- Varshney RK, Song C, Saxena RK, Azam S, Yu S, Sharpe AG et al (2013) Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. *Nat Biotechnol* 31(3):240
- Jain M, Misra G, Patel RK, Priya P, Jhanwar S, Khan AW (2013) A draft genome sequence of the pulse crop chickpea (*Cicer arietinum* L.). *Plant J* 74:715–729
- Upadhyaya HD, Yadav D, Dronavalli N, Gowda CLL, Singh S (2010) Mini core germplasm collections for infusing genetic diversity in plant breeding programs. *Electron J Plant Breed* 1(4):1294–1309

8. Nguyen TT, Taylor PWJ, Redden RJ, Ford R (2004) Genetic diversity estimates in *Cicer* using AFLP analysis. *Plant Breed* 123(2):173–179
9. Sethy NK, Shokeen B, Edwards KJ, Bhatia S (2006) Development of microsatellite markers and analysis of intra specific genetic variability in chickpea (*Cicer arietinum* L.). *Theor Appl Genet* 1(12):1416–1428
10. Fayaz H, Mir AH, Tyagi S, Wani AA, Jan N, Yasin M, Mir JI, Mondal B, Khan MA, Mir RR (2021) Assessment of molecular genetic diversity of 384 chickpea genotypes and development of core set of 192 genotypes for chickpea improvement programs. *Genet Resour Crop Evol.* <https://doi.org/10.1007/s10722-021-01296-0>
11. Talebi REZA, Fayaz F, Mardi M, Pirsyedi SM, Naji AM (2008) Genetic relationships among chickpea (*Cicer arietinum*) elite lines based on RAPD and agronomic markers. *Int J Agric Biol* 10(3):301–305
12. Jida Z, Alemu A, Mullualem D (2018) Genetic diversity among elite chickpea (*Cicer arietinum* L.) varieties of ethiopia based on inter simple sequence repeat markers. *Afr J Biotechnol* 17(34):1067–1075
13. Sefera T, Abebie B, Gaur PM, Assefa K, Varshney RK (2011) Characterisation and genetic diversity analysis of selected chickpea cultivars of nine countries using simple sequence repeat (SSR) markers. *Crop Pasture Sci* 62(2):177–187
14. Keneni G, Bekele E, Imtiaz M, Dagne K, Getu E, Assefa F (2012) Genetic diversity and population structure of ethiopian chickpea (*Cicer arietinum* L.) germplasm accessions from different geographical origins as revealed by microsatellite markers. *Plant Mol Biol Rep* 30(3):654–665
15. Ghaffari P, Talebi R, Keshavarzi F (2014) Genetic diversity and geographical differentiation of Iranian landrace, cultivars, and exotic chickpea lines as revealed by morphological and microsatellite markers. *Physiol Mol Biol Plants* 20(2):225–233
16. Aggarwal H, Choudhary SP, Rana MK, Choudhary R (2018) Assessment or evaluation of genetic diversity among 66 cultivars of chickpea (*Cicer arietinum* L.) of Indian origin Using SSR MARKERS. *Int J Curr Microbiol Appl Sci* 7(2):523–533
17. Choudhary P, Khanna SM, Jain PK (2012) Genetic structure and diversity analysis of the primary gene pool of chickpea using SSR markers. *Genet Mol Res* 11(2):891–905
18. Mir AH, Bhat MA, Dar SA, Sofi PA, Bhat NA, Mir RR (2021) Assessment of cold tolerance in chickpea (*Cicer* spp.) grown under cold/freezing weather conditions of North Western Himalayas of Jammu and Kashmir, India. *Physiol Mol Biol Plants.* <https://doi.org/10.1007/s12298-021-00997-1>
19. Fayaz H, Rather IA, Wani AA, Tyagi S, Pandey R, Mir RR (2019) Characterization of chickpea gene pools for nutrient concentrations under agro-climatic conditions of North- Western Himalayas. *Plant Genet Resour* 17:464–467
20. Saghai-Marouf MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer length polymorphism in barley: mendelian inheritance, chromosomal location and population dynamic. *Proc Natl Acad Sci USA* 81:8014–8019
21. Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. population genetic software for teaching and research. *Mol Ecol Notes* 6:288–295
22. Nei M (1973) Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci, USA* 70:3321–3323
23. Perrier X, Jacquemoud-Collet JP (2006) Darwin software. <http://darwin.cirad.fr/darwin>.
24. Singh R, Sharma P, Varshney RK, Sharma SK (2008) Chickpea improvement: role of wild species and genetic markers. *Biotechnol Genet Eng* 25:267–313
25. Maesen LJG (1972) *Cicer* L., A Monograph of the Genus, with Special Reference to the Chickpea (*Cicer arietinum* L.) Its Ecology and Distribution (Doctoral thesis) Wageningen University
26. Upadhyaya HD, Ortiz R (2001) A mini core subset for capturing diversity and promoting utilization of chickpea genetic resources in crop improvement. *Theor Appl Genet* 102:1292–1298
27. Mir RR, Varshney RK (2013) Future prospects of molecular markers in plants. In: Henry RJ (ed) *Molecular markers in plants*. Blackwell Publishing Ltd, Oxford, pp 170–190
28. Mir RR, Hiremath PJ, Riera-Lizarazu O, Varshney RK (2013) Evolving molecular marker technologies in plants: From RFLP to GBS. In: Lubberstedt T, Varshney RK (eds) *Diagnostics in plant breeding*. Springer, Dordrecht, pp 229–250
29. Saxena MS, Bajaj D, Kujur A, Das S, Badoni S, Kumar V et al (2014) Natural allelic diversity, genetic structure and linkage disequilibrium pattern in wild chickpea. *PLoS ONE* 9(9):e107484
30. Upadhyaya HD, Dwivedi SL, Baum M, Varshney RK, Udupa SM, Gowda CLL (2008) Genetic structure, diversity, and allelic richness in composite collection and reference set in chickpea (*Cicer arietinum* L.). *BMC Plant Biol* 8:106
31. Imtiaz M, Materne M, Hobson K, Van Ginkel M, Malhotra RS (2008) Molecular genetic diversity and linked resistance to ascochyta blight in Australian chickpea breeding materials and their wild relatives. *Aust J Agric Res* 59(6):554–560
32. Castro P, Millan T, Gil J, Merida J, Garcia ML, Rubio J, Fernandez-Romero MD (2011) Identification of chickpea cultivars by microsatellite markers. *J Agric Sci* 149(4):451–460
33. Khan R, Khan H, Harada F, Harada K (2010) Evaluation of microsatellite markers to discriminate induced mutation lines, hybrid lines and cultigens in chickpea (*Cicer arietinum* L.). *Aust J Crop Sci* 4(5):301
34. Shukla N, Babbar A, Prakash V, Tripathi N, Saini N (2011) Molecular diversity analysis in some promising lines of chickpea using SSR markers. *J Food Legumes* 24(4):320–323
35. Samyuktha SM, Kannan JR, Geethanjali S (2018) Molecular genetic diversity and population structure analysis in chickpea (*Cicer arietinum* L.) germplasm using SSR markers. *Int J Curr Microbiol Appl Sci* 7(2):639–651
36. Mohan S, Kalaimagal T (2020) Genetic diversity studies in chickpea (*Cicer arietinum* L.) genotypes using SSR markers. *J Food Legumes* 33(1):1–5
37. Udupa SM, Sharma A, Sharma RP, Pai RA (1993) Narrow genetic variability in *Cicer arietinum* L. as revealed by RFLP analysis. *J Plant Biochem Biotechnol* 2(2):83–86
38. Sonnante G, Marangi A, Venora G, Pignone D (1997) Using RAPD markers to investigate genetic variation in chickpea. *J Genet Breed* 51:303–307
39. Sant VJ, Patankar AG, Sarode ND (1999) Potential of DNA markers in detecting divergence and in analyzing heterosis in Indian elite chickpea cultivars. *Theor Appl Genet* 98:1217–1225
40. Singh R, Prasad CD, Singhal V, Randhawa GJ (2003) Assessment of genetic diversity in chickpea cultivars using RAPD, AFLP, and STMS markers. *J Breed Genet* 57:165–174
41. Hajibarat Z, Saidi A, Hajibarat Z, Talebi R (2014) Genetic diversity and population structure analysis of landrace and improved chickpea (*Cicer arietinum*) genotypes using morphological and microsatellite markers. *Environ Exp Biol* 12:161–166
42. Saeed A, Hovsepian H, Darvishzadeh R, Imtiaz M, Panguluri SK, Nazaryan R (2011) Genetic diversity of Iranian accessions, improved lines of chickpea (*Cicer arietinum* L.) and their wild relatives by using simple sequence repeats. *Plant Mol Biol Rep* 29(4):848–858
43. Khamassi K, Bettaiëb Ben Kaab L, Khoufi S, Chaabane R et al (2012) Morphological and molecular diversity of Tunisian chickpea. *Eur J Horticult Sci* 77(1):31
44. Bharadwaj C, Chauhan SK, Yadav S, Tara Satyavathi C, Singh R, Kumar J, Srivastava R, Rajguru G (2011) Molecular marker-based

- linkage map of chickpea (*Cicer arietinum*) developed from desi × kabuli cross. *Int J Agric Sci* 81(2):116–118
45. Sachdeva S, Bharadwaj C, Sharma V, Patil BS, Soren KR, Roorkiwal M, Varshney RK, Bhat KV (2018) Molecular and phenotypic diversity among chickpea (*Cicer arietinum* L.) genotypes as a function of drought tolerance. *Crop Pasture Sci* 69(2):142–153
 46. Excoffier L, Smouse PE, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131(2):479–491
 47. Roorkiwal M, Von Wettberg EJ, Upadhyaya HD, Warschefsky E, Rathore A, Varshney RK (2014) Exploring germplasm diversity to understand the domestication process in *Cicer* spp using SNP and DArT markers. *PLoS ONE* 9(7):e102016
 48. Van Oss R, Abbo S, Eshed R, Sherman A, Coyne CJ et al (2015) Genetic relationship in *Cicer* sp expose evidence for gene-flow between the cultigen and its wild progenitor. *PLoS ONE* 10(10):e0139789
 49. Parween S, Nawaz K, Roy R, Pole AK, Venkata S, Misra G, Jain M et al (2015) An advanced draft genome assembly of a desi type chickpea (*C. arietinum* L.). *Sci Rep* 5:12806
 50. Thudi M, Upadhyaya HD, Rathore A, Gaur PM et al (2014) Genetic dissection of drought and heat tolerance in chickpea through genome-wide and candidate gene-based association mapping approaches. *PLoS ONE* 9(5):e96758
 51. Bharadwaj C, Srivastava R, Chauhan SK, Satyavathi CT, Kumar J, Faruqui A et al (2011) Molecular diversity and phylogeny in geographical collection of chickpea (*Cicer* sp.) accessions. *J Genet* 1:7

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.