

Analysis of genetic structure and interrelationships in lentil species using morphological and SSR markers

Priyanka Mohan Koul¹ · Vikas Sharma² · Maneet Rana³ · Rakesh K. Chahota¹ · Shiv Kumar⁴ · Tilak R. Sharma¹

Received: 23 December 2016 / Accepted: 28 February 2017
© Springer-Verlag Berlin Heidelberg 2017

Abstract Genetic structure and relationships of 130 lentil accessions belonging to six taxa were analysed. For this purpose, seven morphological traits and 31 polymorphic simple sequence repeat (SSR) primers were used for this purpose. Morphological traits grouped lentil accessions into five main clusters. SSR primers collectively amplified 139 polymorphic alleles in a range of 2–10 with an average of 4.48 alleles. The size of amplified alleles varied from 50 to 650 bp. Polymorphism information content (PIC) ranged from 0.02 to 0.85 with an average of 0.46. Neighbour-joining tree grouped accessions broadly according to their taxonomic ranks, except *L. culinaris* ssp. *odemensis*. Analysis of molecular variance (AMOVA) revealed that a major portion (82.0%) of genetic variance resided within species, while only 18% resided among species. Bayesian model-based STRUCTURE analysis assigned all accessions into five clusters and showed some admixture within individuals. Cluster analysis showed that cultivated *Lens* accessions of Ethiopian origin clustered separately, from other cultivated accessions indicating its distinct lineage.

Among the analysed lentil species, *L. culinaris* ssp. *odemensis* seemed to have conserved genetic background and needs revision of its taxonomic status. Results of present study provide important information on genetic diversity and relationships among different wild and cultivated taxa of lentil. Thus, these results can be useful in designing breeding strategies for future improvement and taxonomic implications in lentil.

Keywords Lentil · Morphological traits · SSR markers · Genetic structure · Phylogenetic relationships

Introduction

Lens, commonly known as lentil, is a self-pollinated diploid ($2n = 2x = 14$) crop from Viciae tribe of Fabaceae family with a haploid genome size of 4063 Mbps (Arumuganathan and Earle 1991). It is comprised of the cultivated lentil (*Lens culinaris* Medik. ssp. *culinaris*) and six related wild taxa (Ferguson et al. 2000). The species *L. culinaris* Medikus has three wild subspecies: *L. culinaris* ssp. *orientalis*, *L. culinaris* ssp. *odemensis* and *L. culinaris* ssp. *tomentosus* (Ferguson et al. 2000); among these, *L. culinaris* ssp. *orientalis* is considered to be the progenitor of cultivated lentil. The other three wild species are *L. ervoides*, *L. nigricans* and *L. lamottei* (Van Oss et al. 1997). The cultivated lentil has been divided into two market classes, namely macrosperma having seed diameter range from 6 to 9 mm and microsperma having seed diameter range from 2 to 6 mm (Barulina 1930). The distribution and cultivation area of lentil ranges from Indian subcontinent to Northern Africa, Western Asia, Southern Europe, North and South America, and Australia (Rubeena and Taylor 2003). Thus, it is

Electronic supplementary material The online version of this article (doi:10.1007/s13205-017-0683-z) contains supplementary material, which is available to authorized users.

✉ Tilak R. Sharma
sharmat88@yahoo.com

- ¹ Department of Agricultural Biotechnology, CSK Himachal Pradesh Agricultural University, Palampur 176062, India
- ² Department of Botany, Punjabi University, Patiala, Punjab 147002, India
- ³ Division of Crop Improvement, ICAR-Indian Grassland and Fodder Research Institute, Jhansi 284128, India
- ⁴ ICARDA, B.P. 6299, Station Experiment, INRA-Quich, Rue Hafiane Cherkaoui, Agdal, Rabat-Institutes, Rabat, Morocco

one of the major pulse crops of the world and requires advanced breeding and improvement programmes to develop high-yielding varieties which should be resistant to various biotic and abiotic stresses. However, the knowledge of genetic variation and species relationships of lentil germplasm is not well understood yet which is important for its improvement in future. Genetic variations within populations of a crop species have been of major interest for plant breeders and geneticists for the introgression of desired genes from one species to other for improvement of a trait (Hayward and Breese 1993). Regarding lentil, cultivation and breeding programme in the past have relied mainly on local landraces of *L. culinaris* ssp. *culinaris*. Thus, unlike other food legumes, lentil has lagged behind in terms of genetic improvement for yield and other major traits (Gupta and Sharma 2006). Evaluation and screening of wild accessions for genetic diversity can help in tapping the unexplored variability and its probable use in lentil improvement. Several attempts have been made to examine the diversity and phylogeny of lentil using seed protein (Hoffman et al. 1986), restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD) and amplified fragment length polymorphism (AFLP) (Havey and Muehlbauer 1989; Abo-Elwafa et al. 1995; Sharma et al. 1995, 1996). However, few studies were reported where robust marker systems such as SSR were used for genetic analysis of lentil (Liu et al. 2008; Kushwaha et al. 2013, 2015). Thus, in the present study we utilized morphological and SSR markers to estimate the genetic structure and interrelationships in cultivated and wild lentil germplasm.

Materials and methods

Plant material

One hundred and thirty accessions of lentil were included in the present study. These accessions were procured from International Center for Agricultural Research in Dry Areas (ICARDA), Ethiopia, and National Bureau of Plant Genetic Resources (NBPGR), New Delhi. Among these, 87 were wild lentil accessions belonging to five species/sub-species (*L. culinaris* ssp. *odemensis*, *L. culinaris* ssp. *orientalis*, *L. nigricans*, *L. ervoides* and *L. lamottei*) and 43 were cultivated landraces/varieties belonging to *L. culinaris* ssp. *culinaris*. A detailed account of plant material is given in Supplementary Table 1. The present investigation was carried out in the Department of Agricultural Biotechnology, College of Agriculture, CSK Himachal

Pradesh Agricultural University, Palampur, Himachal Pradesh, India.

Morphological analysis

Field data were recorded on five plants of each accession. Seven quantitatively measured traits were recorded during cropping season 2013–2014. These seven traits were number of branches/plant, plant height, number of seeds per plant, number of pods per plant, number of seeds per pod, 100-seed weight and seed diameter.

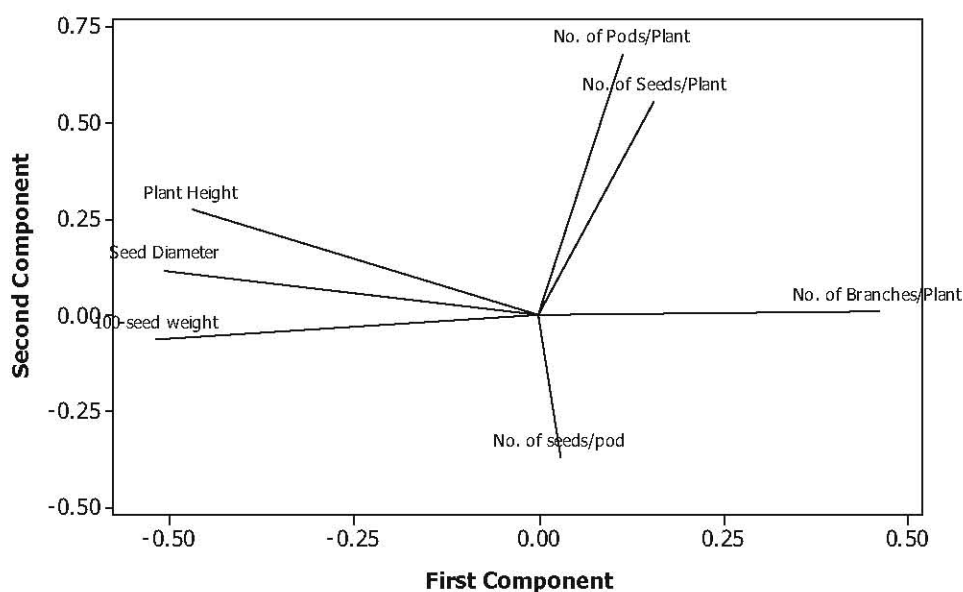
DNA extraction and SSR genotyping

Genomic DNA was isolated from fresh leaves using cetyl trimethyl ammonium bromide (CTAB) method (Doyle and Doyle 1990). After extraction, dried DNA pellet was dissolved in 1 ml TE buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0). The dissolved DNA was treated with RNase (10 mg/ml). The quantity and quality of DNA were estimated through electrophoresis using 0.8 percent agarose gel (Himedia). Further, quantified DNA was made into working stocks as per Sharma et al. 2015a, b. Thirty-one SSR primers showing polymorphic and reproducible bands were selected based on the published reports (Hamwieh et al. 2005; Phan et al. 2007; Saha et al. 2010). PCR reactions were comprised of 10 µl reaction volume as per Sharma et al. (2009). The total PCR reaction volume contained 20 ng of template DNA, 15 ng of each primer, 200 µM of each dNTP, 10 mM Tris-HCl (pH 8.3), 50 mM KCL, 1.5 mM MgCl₂, 0.01% gelatin and 0.5 U of Taq DNA polymerase. PCR amplifications were carried out in Gene Amp PCR System 9700[®] (Applied Biosystems, CA, USA) and 2720 Thermal Cycler (Applied Biosystems, CA, USA). The amplification products were electrophoresed in 3% agarose gel (Himedia) and stained with ethidium bromide. The gels were visualized and photographed using the Gel-Documentation Unit (Bio-Rad, Hercules, CA, USA).

Data analysis

Morphological data were analysed for mean, range, coefficient of variance, correlation coefficient and genetic diversity to deduce the genetic similarity/dissimilarity. Principal component analysis (PCA) was performed using the statistical software PLP Minitab, SYSTAT-12 and StatistiXL version 1.10. For molecular data, all SSR primers alleles were scored manually and the polymorphism information content (PIC) of each primer pair was calculated according to the formula given by Botstein et al. (1980) and implemented in Cervus version 3.0 (Kalinowski

Fig. 1 Biplot of different variables on PC1 and PC2 based on morphological traits data



et al. 2007). Other genetic diversity estimates such as expected heterozygosity (H_e), observed heterozygosity (H_o), Shannon information index (I), etc. were calculated with the help of POPGENE version 1.32 (Yeh and Boyle 1997). Neighbour-joining (N-J) tree was constructed using Dice coefficient with the help of DARwin (Perrier and Jacquemoud-Collet 2006). Bootstrapping with 1000 replicates was also performed with DARwin. Bayesian model-based clustering method implemented in STRUCTURE software, version: 2.3.3 (Pritchard et al. 2000; Falush et al. 2007) was utilized to assess the genetic structure at population level as well as to detect genetic stocks contributing to this germplasm collection. Ancestry model with admixture and correlated allele frequency model were set to get the estimates of posterior probability of data. STRUCTURE was run by setting the value of K from 1 to 10 with three iterations for each value of K. Length of burn-in period was set at 100,000 and number of Markov Chain Monte Carlo (MCMC) repeats after burn-in were set at 100,000. Evanno's method (Evanno et al. 2005) based programme, STRUCTURE HARVESTER, developed by Earl and Vonholdt (2011), was used to determine the value of estimated Ln probability of data-LnP(K) and to get the best fit value of K for the data. STRUCTURE was run for all genotypes collectively and separately for accessions of *L. culinaris* ssp. *culinaris* to assess their genetic structure at species level. Genetic differentiation (F_{st}) estimates for each cluster were also detected by STRUCTURE. Genetic relationships among the genotypes were analysed by Factorial analysis using DARwin (Perrier and Jacquemoud-Collet 2006). Analysis of molecular variance (AMOVA) was performed using GenAlex 6.4 program (Peakall and Smouse 2006).

Results

Morphological evaluation

The analysed lentil accessions revealed significant variations ($p = 0.05$) as was evident from analysis of variance, range, mean and coefficient of variation for all the quantitative traits studied (Supplementary Table 2; Fig. 1). Across all lentil species, the number of branches per plant was lowest (2.3) in Precoz (*L. culinaris* ssp. *culinaris*) and highest (64.6) in EC619872 (*L. nigricans*) with a mean value of 14.7. Plant height was observed minimum (7.3 cm) in EC619674 (*L. culinaris* ssp. *orientalis*) and maximum (66.6 cm) in Lcc13 (*L. culinaris* ssp. *culinaris*) with a mean value of 36.7. The number of seeds per plant was lowest (17.3) in EC619674 (*L. culinaris* ssp. *orientalis*) while highest value (261.3) for number of seeds per plant was observed in accession EC619872 (*L. nigricans*) with a mean value of 132.7. The number of pods per plant was lowest (17.3) in accession EC619674 (*L. culinaris* ssp. *orientalis*), whereas it was highest (197.6) in Lcc51 (*L. culinaris* ssp. *culinaris*) with a mean value of 98.3. Number of seeds per pod did not show much variation as it ranged from 1.00 to 2.00 with a mean value of 1.41. The trait, 100-seed weight showed significant variation which varied from 0.2 g in EC619654 (*L. culinaris* ssp. *orientalis*) to 10.4 g in Lcc187 (*L. culinaris* ssp. *culinaris*) with a mean value of 2.5. The lowest (1.9 mm) seed diameter was recorded in accession EC619860 (*L. ervoides*), while highest (5.3 mm) value of the trait was observed in Precoz (*L. culinaris* ssp. *culinaris*) with a mean value of 3.1 mm. Further, the bivariate correlation coefficients among seven quantitatively measured traits showed significant positive correlation between the traits (Supplementary

Table 3). The number of seeds per plant showed significant positive correlation ($r = 0.23$; $p < 0.01$) with number of branches per plant while number of pods per plant showed positive correlation ($r = 0.80$; $p < 0.01$) with number of seeds per plant. 100-seed weight was associated ($r = 0.66$; $p < 0.01$) with plant height while seed diameter showed significant positive correlation ($r = 0.72$; $p < 0.01$) with 100-seed weight. The degree of association was maximum between seed diameter and plant height followed by 100-seed weight. Hierarchical cluster analysis based on quantitative traits grouped the lentil accessions into five groups. Euclidean dissimilarity matrices varied from 3.59 between EC 619700 (*L. ervoides*, Turkey) and EC619857 (*L. ervoides*, Syrian Arab Republic) to 20,479.53 between Lcc170 (*L. culinaris* ssp. *culinaris*, Ethiopia) and Lcc187 (*L. culinaris* ssp. *culinaris*, Ethiopia). PCA showed that the first principal component (PC1) explained 42.40% of the total variance, followed by PC2 (29.20%) and PC3 (14.80%) as shown in Supplementary Table 4.

SSR polymorphism

All 31 lentil SSRs utilized in this study amplified unambiguous and reliable alleles. In total, 31 loci amplified 139 alleles with an average of 4.48 alleles per primer (Table 1). Size of alleles ranged from 50 to 650 bp. Primer pair SSR-156 amplified maximum of ten alleles, whereas SSR-204, CALTL, LG111 and GLLC-607 amplified minimum of two alleles each. The polymorphic information content (PIC) value ranged from 0.02 in primer CALTL to 0.85 in primer SSR-156 with an average of 0.46. Highest (0.87) and lowest (0.02) expected heterozygosity (H_e) values were obtained with SSR-156 and CALTL, respectively, with an average of 0.50. Highest (0.97) observed heterozygosity (H_o) was obtained with primer pair LSSR1b. Shannon's Information index (I) was recorded maximum (2.13) for SSR-156, whereas the lowest value (0.05) was observed for CALTL with an average of 0.96.

Bayesian genetic structure

STRUCTURE grouped all the accessions into five clusters (Fig. 2). Accessions belonging to *L. culinaris* ssp. *culinaris* were grouped into two separate clusters, i.e. cluster-II and cluster-IV; one of them contained accessions of only Ethiopian origin (cluster-II). All other accessions of cultivated lentil were clustered in a single group (cluster-IV) with some extent of admixtures in few accessions. Majority of accessions belonging to *L. ervoides* were grouped into a single cluster (cluster-III). Few accessions of this species were also dispersed with accessions in other clusters (cluster-I and cluster-IV). *L. culinaris* ssp. *odemensis* formed cluster-V. A very low admixture recorded in the accessions

of this cluster indicated their pure ancestry, which could be attributed to a conservative genetic background. Cluster-I mainly comprised accessions of *L. culinaris* ssp. *orientalis* with some mixed accessions of other species. As accessions of *L. culinaris* ssp. *culinaris* were found distributed in two clusters, with few accessions dispersed across other clusters, further assessment of the population structure within *L. culinaris* ssp. *culinaris* was performed to understand the species complexity. STRUCTURE analysis showed the existence of three groups in the species, with Ethiopian accessions forming a separate gene pool (Supplementary Fig. 1). Wright's fixation index F_{st} , which is a measure of genetic differentiation, was found to be 0.16 overall indicating moderate levels of genetic differentiation. It was recorded highest ($F_{st} = 0.24$) between *L. culinaris* ssp. *odemensis* and *L. ervoides*. Genetic differentiation between *L. ervoides* and *L. culinaris* ssp. *culinaris* was also found to be high ($F_{st} = 0.23$). The lowest genetic differentiation ($F_{st} = 0.11$) was recorded between *L. culinaris* ssp. *orientalis* and *L. culinaris* ssp. *culinaris*, which are taxonomically closer to each other and the former designated as progenitor species of cultivated lentil.

Phylogenetic analysis

The phylogenetic analysis using N-J tree indicated six clusters for wild and cultivated lentil accessions (Fig. 3). High bootstrap values were shown for all the nodes formed in the neighbour-joining tree. Of the 43 cultivated *L. culinaris* ssp. *culinaris* accessions, Ethiopian accessions clustered together in a cluster, named *L. culinaris* ssp. *culinaris* (E), whereas all other accessions of this species grouped into two different clusters. Accessions of *L. culinaris* ssp. *orientalis* largely grouped in a single cluster with few exceptions and showed integrity as a species complex. *L. culinaris* ssp. *odemensis* formed a distinct group with majority of accession while six accessions deviated from this cluster and were dispersed into three different clusters. Two accessions of *L. nigricans* clustered separately in two clusters of wild lentil. Single accession of *L. lamottei* grouped in parent cluster of *L. culinaris* ssp. *odemensis*. The most conservative grouping was shown by accessions of *L. ervoides* species while only two accessions of this species went out of the parent cluster. Overall, majority of the accessions clustered according to existing taxonomic hierarchy except few deviations. While recording distantness on the basis of similarities between accessions we found that wild lentil accessions viz., EC619700 (*L. ervoides*, Turkey) and EC619654 (*L. culinaris* ssp. *orientalis*, Syrian Arab Republic) were most distant with a dissimilarity value of 0.96. Factorial analysis confirmed the results of N-J clustering as it resulted into almost similar clustering pattern as shown in Fig. 4. In factorial analysis,

Table 1 Diversity statistics for 31 SSR loci studied in 130 lentil accessions

S. no.	Locus	T_a (°C)	N_a	Size range (bp)	N_e	I	H_o	H_e	PIC
1.	SSR-66	47	5	205–400	2.24	1.10	0.06	0.56	0.52
2.	SSR-80	51	6	50–300	3.57	1.47	0.71	0.72	0.67
3.	SSR-99	51	4	195–500	1.29	0.47	0.21	0.23	0.21
4.	SSR-124	49	4	190–480	1.43	0.61	0.00	0.31	0.28
5.	SSR-156	53	10	180–500	7.58	2.13	0.02	0.87	0.85
6.	SSR-184	55	5	200–490	2.68	1.18	0.09	0.63	0.57
7.	SSR-204	47	2	180–200	1.40	0.46	0.33	0.28	0.25
8.	SSR-212-1	43	4	190–380	1.44	0.61	0.29	0.31	0.28
9.	SSR-213	47	5	120–300	2.52	1.14	0.42	0.61	0.55
10.	SSR-302	49	6	170–600	4.23	1.57	0.86	0.77	0.73
11.	SSR-317-2	53	6	110–150	3.02	1.32	0.09	0.67	0.62
12.	CALTL	51	2	330–490	1.01	0.05	0.00	0.02	0.02
13.	FENR	51	3	470–550	2.58	1.02	0.27	0.61	0.54
14.	HPB2	51	6	400–620	3.52	1.40	0.67	0.72	0.67
15.	Lup183	47	3	260–380	1.24	0.41	0.12	0.19	0.18
16.	Lup242	47	4	160–650	2.67	1.07	0.73	0.63	0.55
17.	Lup336	47	4	205–300	2.22	0.98	0.15	0.55	0.49
18.	LG91	50	3	200–650	1.72	0.72	0.37	0.42	0.37
19.	LG111	51	2	210–400	1.84	0.64	0.71	0.46	0.35
20.	LSSR1b	49	3	170–510	2.08	0.78	0.97	0.52	0.40
21.	LSSR11b	49	5	140–510	2.33	0.97	0.91	0.57	0.48
22.	MsU141	49	4	150–600	1.73	0.77	0.44	0.42	0.38
23.	GLLC-106	51	4	220–530	1.25	0.43	0.05	0.20	0.19
24.	GLLC-108	50	6	190–550	4.47	1.60	0.78	0.78	0.74
25.	GLLC-511	50	4	140–500	1.97	0.75	0.51	0.50	0.38
26.	GLLC-527	48	5	210–380	2.72	1.22	0.03	0.64	0.58
27.	GLLC-548	50	4	80–270	1.50	0.65	0.31	0.33	0.31
28.	GLLC-598	48	9	120–600	6.66	2.01	0.84	0.85	0.83
29.	GLLC-607	47	2	110–130	1.11	0.21	0.04	0.10	0.09
30.	GLLC-609	47	5	190–500	3.45	1.38	0.42	0.72	0.66
31.	GLLC-614	50	4	170–400	1.73	0.81	0.16	0.43	0.38
Mean			4.48		2.55	0.96	0.37	0.50	0.46

T_a annealing temperature, bp base pair, N_a observed number of alleles, N_e effective number of alleles, I Shannon's information index, H_e expected heterozygosity, H_o observed heterozygosity, PIC polymorphism information content

first two principal coordinates explained 40.64% cumulative variations. AMOVA revealed that a major portion (82%) of genetic variance resided within species, while a lesser part (18%) of genetic variance resided among species (Table 2).

Discussion

Morphological evaluation

Lentil accessions analysed in the present study have shown high morphological diversity for the traits studied and,

therefore, accessions from analysed germplasm can be a valuable resource for lentil improvement programme. The highest number of seeds and maximum number of branches were observed in *L. nigricans*, a wild species of lentil. This important characteristic of *L. nigricans* can make it a potential donor for breeding experiments aiming to increase the number of seeds per plant. Earlier workers have also advocated the need to give more emphasis on such traits for improving the lentil productivity (Roopam and Singh 2014). The significant variation for morphological traits observed in lentil could be due to the diverse origin of these species as also inferred by Singh et al. (2014). The number of branches per plant and number of

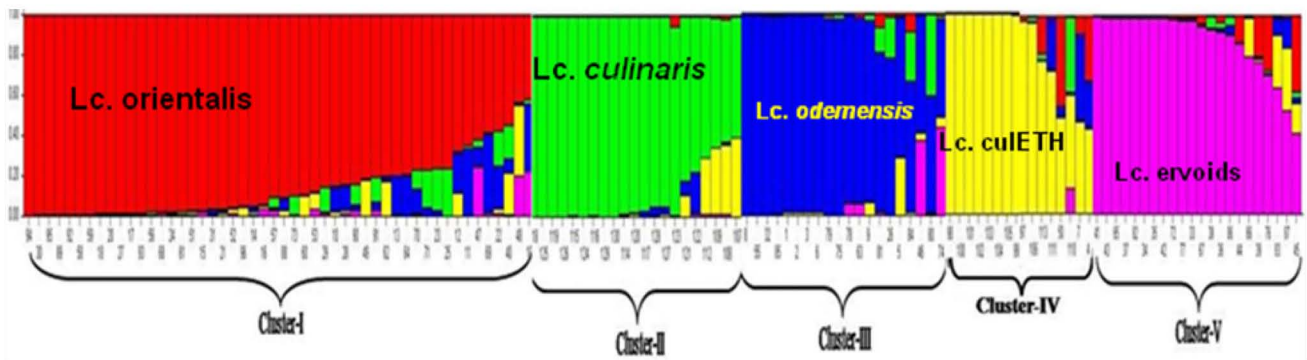


Fig. 2 Genetic structure of 130 lentil accessions as inferred by STRUCTURE v2.3.3 with SSR markers data set. The five inferred clusters are shown with *different colours*

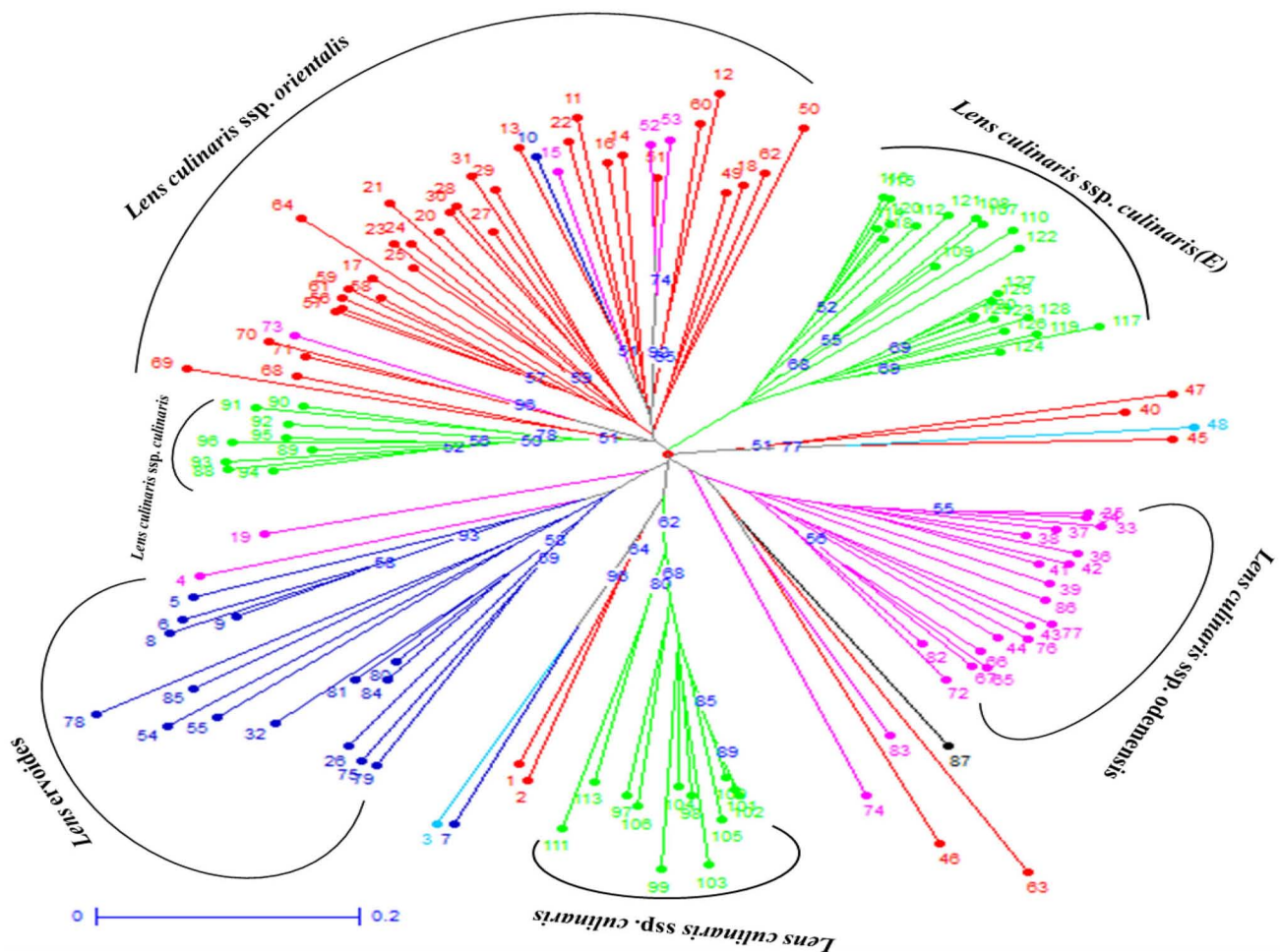


Fig. 3 Radial neighbour-joining tree based on 139 alleles from 31 SSR loci among 130 lentil accessions belonging to six species. Bootstrap values $\geq 50\%$ are also shown

seeds per plant were higher in the wild *L. nigricans* accession EC619872 collected/originated from Turkey. The maximum plant height was recorded in the accessions of *L. culinaris* ssp. *culinaris* and lowest in *L. ervoides*. However, the wild annual *L. culinaris* ssp. *orientalis* accession EC619662 collected from Turkey also showed

increased plant height. *L. nigricans* accession EC619872 collected from Turkey was superior for number of seeds per plant and number of pods per plant. Range and mean for 100-seed weight and seed diameter were found maximum in accessions of *L. culinaris* ssp. *odemensis* and *L. lamottei* collected from Turkey and Spain, respectively. It

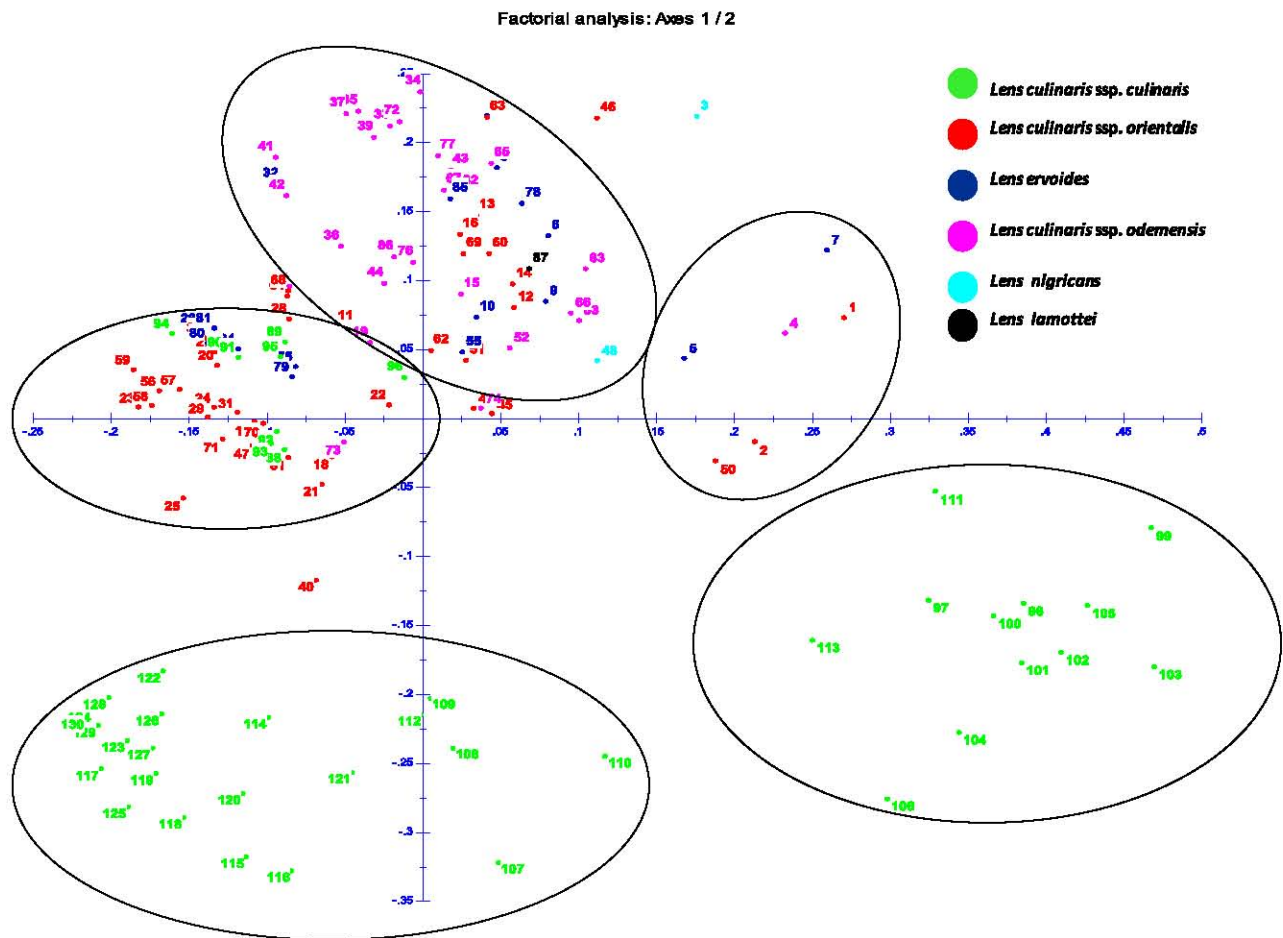


Fig. 4 Factorial analysis of *Lens* accessions based on molecular data generated by 31 SSRs

Table 2 Analysis of molecular variance (AMOVA) among and within *Lens* species

Source	Df	SS	MS	Est. var.	Percent (%)
Among species	5	378.976	75.795	3.157	18
Within species	125	1830.612	14.645	14.645	82
Total	130	2209.588	17.802	17.802	100

df degree of freedom, *SS* sum of squares, *MS* mean squares, *Est.Var.* estimated variance

is important to note that wild accessions collected from Turkey exhibited substantial variations for useful agromorphological traits. These promising accessions of *Lens culinaris* ssp. *orientalis*, *L. culinaris* ssp. *odemensis*, *L. nigricans* and *L. lamottei* can be exploited for broadening the genetic base of cultivated lentil. Correlation analysis also showed some significant associations between few traits of importance. The number of seeds per plant showed significant positive correlation with number of branches per plant while number of pods per plant showed positive correlation with number of seeds per plant. 100-seed weight was associated with plant height while seed

diameter showed significant positive correlation with 100-seed weight. These estimates of correlation will be helpful in planning breeding strategies in the species (Singh et al. 2008; Tyagi and Khan 2011; Singh et al. 2012). Moreover, they have suggested that seed yield in lentil can be increased by selecting ideotypes having more number of branches per plant, more seeds per plant, more pods per plant, increased plant height, increased 100-seed weight and seed diameter (Younis et al. 2008). Clustering based on morphological traits and Euclidian distance depicted five groups for entire germplasm but it did not show any correlation with geographic occurrence of

accessions, which was in accordance with a study in lentil by Singh et al. (2014). Further, the lack of robustness in the dendrogram may be attributed to less numbers of phenotypic traits studied. PCA showed that PC1 was the most important and explained 42.4% of the total variance, which was mainly contributed by number of branches per plant, number of seeds per plant, number of pods per plant and number of seeds per pod. PC2 accounted for 29.2% variation and was attributed to number of pods per plant, number of seeds per plant, plant height, seed diameter and number of branches per plant. PC3 contributed 14.8% of the total variance through number of seeds per pod. The loading plot drawn for PC1 and PC2 (Fig. 1) also indicated the importance of different morphological traits for explaining the variance among accessions. The high morphological variability observed here is in consonance with observations of Bacchi et al. (2010) and Kumar et al. (2014).

SSR polymorphism

Genetic diversity revealed by SSR primers was high which was supported by the recorded high values of mean allele numbers, heterozygosity, PIC and other diversity indices. Average number of 4.48 alleles recorded in present work was higher than that obtained in the previous study by Kumar et al. (2014), whereas it was lower than the mean number of alleles recorded by Bacchi et al. (2010). However, mean heterozygosity and PIC of 0.50 and 0.46, respectively, were lower than those observed in other studies carried out in lentil (Liu et al. 2008; Andeden et al., 2013; Kumar et al. 2014). Average Shannon's Information Index (I) value of 0.96 was also lower than that observed by Liu et al. (2008). Further, species-wise analysis also revealed high genetic diversity within the accessions of *L. culinaris* spp. *orientalis*, of which most of the accessions were from Syrian Arab Republic, Uzbekistan, Turkey, Iran, Cyprus and Russian Federation. Ferguson and Robertson (1996) found that *L. culinaris* spp. *orientalis* possessed the high diversity based on morphological characters compared to isozymes markers. They also reported high genetic diversity based on RAPD analysis in *L. culinaris* spp. *orientalis* accessions from Jordan, Palestine, and southern Syria, with slightly lower diversity in those accessions from northwest Syria and southeast Turkey (Ferguson 1998a). Hamwiah et al. (2009) also observed high genetic diversity within *L. culinaris* spp. *orientalis* accessions originating from the Fertile Crescent (the north and south of Syria, Turkey, Iran, Jordan, Lebanon, Cyprus, and Palestine). Similarly, accessions of *L. culinaris* spp. *odemensis* have shown diverse genetic background in our study that is in agreement with the findings of Hamwiah et al. (2009). Moreover, we observed lowest genetic relatedness (0.04) between *L. ervoides* from Turkey and *L.*

culinaris spp. *orientalis* from Syrian Arab Republic which might be due to their diverse geographical origin and distant phylogenies.

Bayesian genetic structure

STRUCTURE assigned all the accessions into five ($K = 5$) clusters. Unexpectedly, accessions of *L. culinaris* spp. *culinaris* were divided into two clusters: cluster-II representing mixed accessions and cluster-IV representing accessions from Ethiopia. This indicated that cultivated lentil has differentiated and formed unique allelic combinations as a result of geographical isolation. The study also revealed that cultivated lentil species has two gene pools. Study by Alo et al. (2011) supports this fact as they have also observed two different groups in STRUCTURE analysis for domesticated lentil accessions. The diversified nature of cultivated lentil was further supported by STRUCTURE analysis when it was performed only on cultivated accessions, which showed presence of three clusters in it (Supplementary Fig. 1). Further, wild relatives of *L. culinaris*, namely *L. culinaris* spp. *orientalis* and *L. culinaris* spp. *odemensis*, formed two different clusters, i.e. cluster-I and cluster-III, respectively. Wild species *L. ervoides* clustered separately in cluster-V, which also included one accession of *L. nigricans*. The separate clustering of *L. culinaris* spp. *culinaris* and *L. culinaris* spp. *orientalis* was concordant with Dikshit et al. (2015) who also reported structuring of accessions according to their cultivated and wild nature. Based on STRUCTURE analysis Wong et al. (2015) divided cultivated and wild lentil species into four different gene pools, i.e. primary, secondary, tertiary and quaternary gene pools. However, STRUCTURE analysis in present study assigned accessions into three gene pools defined by them, which included *L. culinaris* spp. *culinaris*, and *L. culinaris* spp. *orientalis* in primary, *L. odemensis* and *L. lamottie* in secondary and *L. ervoides* and *L. nigricans* in tertiary gene pool. Further, Wright's F_{st} values between different species suggested that *L. culinaris odemensis* and *L. ervoides* were the most divergent species, while *L. culinaris* spp. *culinaris*, and *L. culinaris* spp. *orientalis* were found to be most similar among analysed accessions. This fact was in consonance with earlier studies by Ahmad and McNeil (1996) and Ladizinsky (1999) who reported poor or negligible crossability between *L. ervoids* and cultivated species and close relationship between *L. culinaris* spp. *culinaris*, and *L. culinaris* spp. *orientalis*.

Phylogenetic analysis

Phylogeny of *Lens* genus largely remained unambiguous. Inclusion of seven taxa in this genus was agreed upon by

most of the workers (Ferguson et al. 2000). However, many earlier and contemporary researchers have reported a major controversial aspect in this phylogenetic classification and that is distantness of *L. culinaris* ssp. *odemensis* from *L. culinaris* (Sonnante et al. 2003; Duran and Perez de la Vega 2004; Wong et al. 2015). According to Ferguson et al. (2000) genus *Lens* comprises seven taxa. These are *L. culinaris* with four of its sub species, namely *L. culinaris* ssp. *culinaris*, *L. culinaris* ssp. *orientalis*, *L. culinaris* ssp. *tomentosus* and *L. culinaris* ssp. *odemensis*, while other three species are *L. nigricans*, *L. lamottei* and *L. ervoides*. In our study, we have also followed this existing system for entire germplasm analysed. But, N-J clustering of the present study depicted that *L. culinaris* ssp. *odemensis* clustered differently from its parent species *L. culinaris* and seems to be an entirely different entity. Therefore, we propose revisiting of the taxonomic status of this taxon and advocate its existence to a rank of independent species. This fact is in agreement with a recent finding of Wong et al. (2015), who characterized *Lens* species and revisited its classification using genotyping-by-sequencing (GBS) approach. The separate clustering of *L. culinaris* ssp. *odemensis* from *L. culinaris* was in opposition to study conducted by Alo et al. (2011). However, they also suggested the revision of existing classification of *Lens*. The close relation between *L. culinaris* ssp. *culinaris* and *L. culinaris* ssp. *orientalis* was observed as also reported by other workers (Sharma et al. 1996; Mayer and Bagga 2002; Sonnante et al. 2003; Duran and Perez de la Vega 2004; Fikiru et al. 2010; Reddy et al. 2010; Alo et al. 2011; Singh et al. 2014; Dikshit et al. 2015; Wong et al. 2015). *L. ervoides* appeared as a conservative cluster and can be considered a distant relative of cultivated species and sub-species. Accession of *L. lamottei* clustered closely with *L. culinaris* ssp. *odemensis*. Similar findings were reported by Verma et al. (2014) and Wong et al. (2015) while two accessions of *L. nigricans* clustered separately from each other and one accession (LnigEC619872) showed affinity to *L. ervoides* which was in concurrence with earlier studies (Fikiru et al. 2010; Wong et al. 2015). Grouping of the other *L. nigricans* accession (LnigEC619873) with *L. culinaris* ssp. *orientalis* and deviation of few other accessions from their parent clusters need further investigations with larger sample size. The clustering pattern of N-J tree was supported by factorial analysis as the grouping of accessions was nearly identical to the phylogenetic tree. Furthermore, partitioning of genetic variance within and between species indicated that larger variance was harboured within species while a very small portion was harboured between the species which indicated that frequency of gene flow within species is very high than that between species.

In conclusion, the morphological traits and SSR data have revealed high genetic diversity among both cultivated and wild lentil genotypes analysed in this study. Genetic structuring showed five different genetic stocks for entire germplasm suggesting broad genetic background of the crop. Moreover, structure analysis of cultivated lentil accessions revealed Ethiopian accessions as a different gene pool. Clustering patterns of N-J tree can be useful for resolving phylogenetic and taxonomic issues in this genus. Phenotypic traits showing correlation with seed yield can be employed for the development of high-yielding varieties in lentil. Overall, the results of present work provide important genetic information which can be useful for lentil breeding programmes and taxonomic inferences in lentil species.

Acknowledgements Authors gratefully acknowledge the Department of Biotechnology (DBT), Government of India, New Delhi, for providing financial support to conduct this work under HRD programme.

Compliance with ethical standards

Conflict of interest Authors declare that they have no conflict of interest.

References

- Abo-elwafa A, Murai K, Shimada T (1995) Intra-and interspecific variations in *Lens* revealed by RAPD markers. *Theor Appl Genet* 90:335–340
- Ahmad M, McNeil DL (1996) Comparison of crossability, RAPD, SDS-PAGE and morphological markers for revealing genetic relationships within and among *Lens* species. *Theor Appl Genet* 93:788–793
- Alo F, Furman BJ, Akhunov E, Dvorak J, Gepts P (2011) Leveraging genomic resources of model species for the assessment of diversity and phylogeny in wild and domesticated lentil. *J Hered* 102:315–329
- Andeden EE, Derya M, Baloch FS, Kilian B and Ozkan H (2013) Development of SSR markers in lentil. In Proceedings of Plant and Animal Genome Conference XXI P0351, San Diego, CA, USA
- Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. *Plant Mol Biol Rep* 9:208–218
- Bacchi M, Leone M, Mercati F, Preiti G, Sunseri F, Monti M (2010) Agronomic evaluation and genetic characterization of different accessions in lentil (*Lens culinaris* Medik.). *Italian J Agron* 5:303–314
- Barulina H (1930) Lentil of the USSR and other countries. *Bull Appl Bot* 40:1–319
- Botstein D, White KL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Human Genet* 32:314–331
- Dikshit HK, Singh A, Singh D, Aski MS, Prakash P, Jain N et al (2015) Genetic diversity in *Lens* species revealed by EST and genomic simple sequence repeat analysis. *PLoS ONE* 10:e0138101
- Doyle JJ, Doyle JL (1990) A rapid total DNA preparation procedure for fresh plant tissue. *Focus* 12:13–15

- Duran Y, Perez de la Vega M (2004) Assessment of genetic variation and species relationships in a collection of *Lens* using RAPD and ISSR. *Span J Agric Res* 2:538–544
- Earl DA, vonHoldt BM (2011) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Res* 4:359–361
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620
- Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol Ecol Notes* 7:574–578
- Ferguson M (1998) Studies of genetic variation within the genus *Lens*. Ph.D. Thesis. School of Biological Sciences, University of Birmingham, UK
- Ferguson M, Robertson LD (1996) Genetic diversity and taxonomic relationships within the genus *Lens* as revealed by allozyme polymorphism. *Euphytica* 91:163–172
- Ferguson ME, Maxted N, Van Slageren M, Robertson LD (2000) A re-assessment of the taxonomy of *Lens* Mill. (Leguminosae, Papilionoideae, Viciaeae). *Bot J Linn Soc* 133:41–59
- Fikiru E, Tesfaye K, Bekele E (2010) A comparative study of morphological and molecular diversity in Ethiopian lentil (*Lens culinaris* Medikus) landraces. *Afr J Plant Sci* 4:242–254
- Gupta D, Sharma SK (2006) Evaluation of wild *Lens* taxa for agromorphological traits, fungal diseases and moisture stress in North Western Indian Hills. *Genet Resour Crop Evol* 53:1233–1241
- Hamwiah A, Udupa S, Choumane W, Sarker A, Dreyer F, Jung C et al (2005) A genetic linkage map of *Lens* sp. based on microsatellite and AFLP markers and the localization of fusarium vascular wilt resistance. *Theor Appl Genet* 110:669–677
- Hamwiah A, Udupa SM, Sarker A, Jung C, Baum M (2009) Development of new microsatellite markers and their application in the analysis of genetic diversity in lentils. *Breed Sci* 59:77–86
- Havey MJ, Muehlbauer FJ (1989) Variability for restriction fragment lengths and phylogenies in lentil. *Theor Appl Genet* 77:839–843
- Hayward MD, Breese EL (1993) Population structure and variability. In: Hayward MD, Bosemark NO, Romayosa I (eds) *Plant breeding: principles and prospects*. Chapman and Hall, London, pp 7–29
- Hoffman DL, Soltis DE, Muehlbauer FJ, Ladizinsky G (1986) Isozyme polymorphism in *Lens* (Leguminosae). *Syst Bot* 11:392–402
- Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol Ecol* 16:1099–1106
- Kumar J, Srivastava E, Singh M, Kumar S, Nadarajan N, Sarker A (2014) Diversification of indigenous gene-pool by using exotic germplasm in lentil (*Lens culinaris* Medikus subsp. *culinaris*). *Physiol Mol Biol Plants* 20:125–132
- Kushwaha UKS, Ghimire SK, Yadav NK, Ojha BR (2013) Genetic relatedness of lentil (*Lens culinaris* L.) germplasm by using SSR markers. *Int J Appl Sci Biotech* 1:132–136
- Kushwaha UKS, Ghimire SK, Yadav NK, Ojha BR, Niroula RK (2015) Genetic characterization of lentil (*Lens culinaris* L.) germplasm by using SSR markers. *Agric Biol Sci J* 1:16–26
- Ladizinsky G (1999) Identification of the lentil's wild genetic stock. *Genet Resour Crop Evol* 46:115–118
- Liu J, Guan JP, Xu DX, Zhang XY, Gu J, Zong XX (2008) Genetic diversity and population structure in lentil (*Lens culinaris* Medik.) germplasm detected by SSR markers. *Acta Agron Sin* 34:1901–1909
- Mayer MS, Bagga SK (2002) The phylogeny of *Lens* (Leguminosae): new insight from ITS sequence analysis. *Plant Syst Evol* 232:145–154
- Peakall R, Smouse PE (2006) GenAlEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6:288–295
- Perrier X and Jacquemoud-Collet JP (2006) DARwin software <http://darwin.cirad.fr/darwin>
- Phan HT, Ellwood SR, Hane JK, Ford R, Materne M, Oliver RP (2007) Extensive macrosynteny between *Medicago truncatula* and *Lens culinaris* ssp. *culinaris*. *Theor Appl Genet* 114:549–558
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Reddy MRK, Rathour R, Kumar N, Katoch P, Sharma TR (2010) Cross-genera legume SSR markers for analysis of genetic diversity in *Lens* species. *Plant Breed* 129:514–518
- Roopam KM, Singh SK (2014) Genetic variability, heritability, expected genetic advance and character association in lentil (*Lens culinaris* Medic.). *Trends Biosci* 7:2017–2020
- Rubeena FR, Taylor PWJ (2003) Construction of an intraspecific linkage map of lentil (*Lens culinaris* ssp. *culinaris*). *Theor Appl Genet* 107:910–916
- Saha GC, Sarker A, Chen W, Vandemark GJ, Muehlbauer FJ (2010) Inheritance and linkage map positions of genes conferring resistance to stemphylium blight in lentil. *Crop Sci* 50:1831–1839
- Sharma SK, Dawson IK, Waugh R (1995) Relationships among cultivated and wild lentils revealed by RAPD analysis. *Theor Appl Genet* 91:647–654
- Sharma SK, Knox MR, Ellis THN (1996) AFLP analysis of the diversity and phylogeny of *Lens* and its comparison with RAPD analysis. *Theor Appl Genet* 93:751–758
- Sharma V, Bhardwaj P, Kumar R, Sharma RK, Sood A, Ahuja PS (2009) Identification and cross-species amplification of EST derived SSR markers in different bamboo species. *Conserv Genet* 10:721–724
- Sharma V, Rana M, Katoch M, Sharma PK, Ghani M, Rana JC, Sharma TR, Chahota RK (2015a) Development of SSR and ILP markers in horsegram (*Macrotyloma uniflorum*), their characterization, cross-transferability and relevance for mapping. *Mol Breed* 35:102
- Sharma V, Sharma TR, Rana JC, Chahota RK (2015b) Analysis of genetic diversity and population structure in horsegram (*Macrotyloma uniflorum*) using RAPD and ISSR markers. *Agric Res* 4:221–230
- Singh A, Singh SK, Sirohi A, Singh KV, Kumar R (2008) Genetic variability, heritability and character association yield and other characters in lentil (*Lens culinaris* Medik.). *Progress Agric* 8:42–44
- Singh RS, Sood BC, Kumar N (2012) Identification of yield attributes for selection through correlation and path analysis of lentil (*Lens culinaris* Medik.). *Crop Res* 43:171–177
- Singh M, Bisht IS, Kumar S, Dutta M, Bansal KC, Karale M, Sarker A, Amri A, Kumar S, Datta SK (2014) Global wild annual lens collection: a potential resource for lentil genetic base broadening and yield enhancement. *PLoS ONE* 9:e107781
- Sonnante G, Galasso I, Pignone D (2003) ITS sequence analysis and phylogenetic inference in the genus *Lens* Mill. *Ann Bot* 91:49–54
- Tyagi SD, Khan MH (2011) Correlation, path coefficient and genetic diversity in lentil (*Lens culinaris* Medik.) under rainfed conditions. *Int Res J Plant Sci* 2:191–200
- Van Oss H, Aron Y, Ladizinsky G (1997) Chloroplast DNA variation and evolution in the genus *Lens* Mill. *Theor Appl Genet* 94:452–457

- Verma P, Sharma TR, Srivastava PS, Abdin MZ, Bhatia S (2014) Exploring genetic variability within lentil (*Lens culinaris* Medik.) and across related legumes using a newly developed set of microsatellite markers. *Mol Biol Rep* 41:5607–5625
- Wong MML, Gujaria-Verma N, Ramsay L, Yuan HY, Caron C, Diapari M et al (2015) Classification and characterization of species within the genus *Lens* using genotyping-by-sequencing (GBS). *PLoS ONE* 10:e0122025
- Yeh FC, Boyle TJB (1997) Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belg J Bot* 129:157
- Younis N, Hanif M, Sadiq S, Abbas G, Asghar MJ, Haq MA (2008) Estimation of genetic parameters and path analysis in lentil. *Pak J Agri Sci* 45:44–48