Range Mgmt. & Agroforestry 38 (1) :48-57, 2017
ISSN 0971-2070

Population studies and assessment of molecular genetic divergence among alfalfa (*Medicago* sp.) sub-species inhabiting cold arid province of Ladakh

Shahida Iqbal, Gulzar Zaffar, Asif B. Shikari*, Bilal A. Padder, Gazala H. Khan and Mushtaq A. Dar
Sher-e-Kashmir University of Agricultural Sciences & Technology, Kashmir-191121, India
*Corresponding author e-mail: asifshikari@gmail.com
Received: 2nd January, 2016 Accepted: 8th May, 2017

Abstract
Alfalfa (*Medicago sativa* L.) also called as queen of fodder or green gold, is the most important fodder crop grown in Ladakh region owing to its well adaptability to cold arid habitat (35000 m amsl). Three species of *Medicago* such as *M. falcata, M. sativa and M. varia* are found in Ladakh. Owing to little information available on population sub-structure and genetic diversity of alfalfa in the region, 45 accessions of *Medicago* sub species collected from Drass, Kargil and Leh regions of Ladakh were studied for molecular and morphological diversity. Observations recorded on 12 quantitative traits revealed presence of high genetic diversity among the sub-populations of alfalfa. The genotypes spread into 7 clusters based on D^2 statistic with maximum distance of 833.92 was observed between cluster VI and cluster V. Molecular characterization was carried out using 12 SSR markers. Analysis of molecular variance (AMOVA) revealed significant variation within population variance. Maximum PIC value was shown by the marker Afat-15 (0.74) and minimum by the marker MTIC-432 (0.54). The genotypes were grouped into five different clusters following UPGMA method. Small Fst values depicted minimum population sub-structure and greater sharing of gene pool among different sub-species. The kind of information we generated on the molecular divergence in alfalfa populations of the region shows that there is a need to introduce novel *Medicago* species in the region and cross them with already existing inter-sub-species derivatives. Besides, population improvement programs should be initiated to improve the frequency of desirable alleles among and within sub-species.

Keywords: Alleles, Genetic divergence, Markers, *Medicago* species, Population structure, SSR

Introduction

Alfalfa (*Medicago sativa* L., 2n = 4x = 32) is an important perennial leguminous forage crop with high nutritional quality and herbage yield. Alfalfa originated in Vavilov's "Near Eastern Centre" – Asia Minor, Transcaucasia, Iran and Turkistan. Alfalfa spread from its centre of origin into Europe, North Africa, Arabia and eventually South America with invading armies, explorers and missionaries who used it as feed for horses and other livestock. Alfalfa is grown over 32 million hectares worldwide (Michaud *et al*., 1988). Cultivated alfalfa has been improved from a complex taxonomic group known as the *Medicago sativa-falcata* complex, which includes both diploid (2n=2x=16) and tetraploid (2n= 4x=32) species. Flower colour, pod shape and pollen shape have traditionally been used to differentiate taxa in the complex. The diploid members of the complex include *M. sativa* ssp. *falcata*, with yellow flowers and sickle shaped pods, *M. sativa* ssp. *coerulea*, with purple flowers and pods having multiple coils, and their natural hybrid, *M. sativa* ssp. *hemicycla*, with variegated flower colour and partially coiled pods. The tetraploid subspecies is a complex that includes *M. sativa* ssp. *Sativa* (the direct analogue of diploid ssp. *coerulea)*, *M. sativa* ssp. *falcata* and the tetraploid hybrid *M. sativa* ssp. *varia* (Quiros and Bauchan, 1988). Hybridization among taxa is possible even across ploidy level by unreduced gametes (McCoy and Bingham, 1988). The original alfalfa of Ladakh, i.e. *Medicago falcata* was yellow flowered. The traders of the old silk route introduced *M. sativa* from Yarkand (Central Asia) to enhance the availability of forage for their horses and locally *M. falcata* is known as *Ole* while *M. sativa* as *Yarkandi Ole.* A lot of natural hybridization has taken place and now it is rare to find a true stand of either *M. sativa* or *M. falcata* and this has resulted in a wide range of variability in habit, leaf size, height, colour of the flower, shape of pods and the resistance of plants to cold and aridity. These characteristics of alfalfa in Ladakh attracted the attention of American scientists and they introduced the crop to USA during 1910 after naming it variety *Ladakh* (Bolton, 1962). As per Whyte *et al.* (1953) the present populations of *Medicago* in Ladakh are confined to cold arid habitats above 3500 m amsl and have been classified as *M. sativa, M. falcata* and *M. varia*. These species hold a rich source

of natural variation and are valuable genetic resources for developing better grazing legumes especially for dry and cold regions where green fodder is very scarce.

The information on genetic diversity in such populations would help to understand gene flow among and within natural populations and its likely effect on morphological traits including forage yield. Although, morphology, physiology and phenology of the three species has been studied previously in the region (Misri, 1986); their population genetic structures have not been understood yet. Molecular markers are powerful tools to identify the genetic diversity, population sub-structure and cultivar identification. Different molecular marker types that have been used to assess genetic diversity in alfalfa include RAPDs (Musial *et al.,* 2002), AFLP (Segovia-Lerma *et al.* 2003), RFLP (Maureira *et al.,* 2004), SSRs (Flajoulot *et al.,* 2005) and SRAP (Vandemark *et al.,* 2005). Among them, simple sequence repeat (SSR) markers are simple, quick and reliable assays for diversity analysis given their multi-allelic nature.With this background, the present study was a first time attempt aimed to characterize the genetic variability in *Medicago* species using morphological and DNA based markers to understand the genetic structure of *Medicago* populations in cold arid Ladakh province.

Materials and Methods

Plant materials: Hot spots of alfalfa cultivation namely, Kargil, Leh and Drass of Ladakh Province of J&K were explored. Sub-species designation was done on the basis of flower colour and pod traits and the populations were marked as those belonging to *Medicago sativa* ssp. *falcata*, *M. sativa* ssp. *sativa* and *M. sativa* ssp. *varia*. In total, fifty plants were selected for each sub-population. Hence forth, the term sub-populations will be used synonymously to sub-species for the present discussion. During the subsequent visits to the identified locations inhabited by sub-populations of alfalfa, morphological data was recorded on individual plants. Plant populations comprising of the sub-population from diverse habitats were selected at random. Fifty plants were selected from each sub-population. The morphological parameters were recorded *in vivo* on these selected plants from each sub-population and also their seed and leaf samples were obtained from respective locations.

Marker analysis: The leaf tissue was collected from sampled plants across populations from all the locations. Genomic DNA was extracted from leaves using CTAB procedure (Cetyl Trimethyl Ammonium Bromide) as modified by Maroof *et al.* (1984). Heat treated RNase (Fermentas, Lithuania, USA) was then added to a final concentration of 1 µg/ml. DNA samples were stored at -80°C for further use. Total of 12 SSR markers were used to discriminate *Medicago* populations. The primer sequences were retrieved from www.medicago.org (Table 1). The PCR amplification was carried out in 25 µl reaction volume using thermal cycler (Whatman Biometra, T-Gradient, Goettingen, Germany) after it was programmed for initial denaturation for 5 min at 94°C, denaturation for 40 s at 94°C, for 40 s at varying annealing temperatures and primer extension for 40 s at 72°C. A final elongation step at 72°C for 10 min was performed after 30 cycles. Prior to loading, the PCR product was denatured by adding one volume of denaturing gel loading buffer (98% formamide; 0.025% bromophenol blue, 0.025% xylene cyanol and 10 mM EDTA pH 8) and was heated at 94°C for 4 min. Microsatellite alleles were resolved by running the PCR product on a 6 % denaturing PAGE (Urea:420 g; 10X TBE: 100 ml; acrylamide mix: 40% (19:1) in 1x TBE buffer on Sequencing Unit (Genetic Asia Biotech Pvt. Ltd.) using 10 bp DNA ladder as size standard. The alleles were visualized after silver staining procedure.

Statistical analysis: Descriptive statistics for all morphological traits was computed to estimate variability of all traits. Morphological data was subjected to cluster analysis as per Singh and Chaudhary (1985). ARLEQUIN 3 (Excoffiel *et al*., 2005) was used to workout AMOVA. Genalex 6.1 (Peakall and Smause, 2001) was used to compute heterozygosity measures. DARwin 5.0 software was used to estimate molecular diversity. Polymorphism information content (PIC) was calculated as described by Botstein *et al.* (1980).

Results and Discussion

Variability for agro-morphological traits: The observations were recorded on 45 single plants for twelve agro-morphological traits *viz.,* days to 50% flowering, plant height, green fodder yield, dry weight per plant, leaf fresh weight, leaf dry weight, number of flowers on mainraceme, number of pods on mainraceme, number of seeds per pod, 1000-seed weight, seed yield per plant and crude protein content of leaves (Table 2). After employing Mahalanobis D^2 analysis, the 45 genotypes grouped into seven different clusters as per Toucher's clustering method (Rao, 1952). Cluster I comprised maximum number (11) of genotypes followed by cluster IV with 10 genotypes (Table 3). The cluster VI had highest intra cluster distance of 37.65. The Clusters V and VI

showed highest inter-cluster distance of 833.92; whereas, minimum inter-cluster distance (30.52) was observed between cluster V and II. Cluster VI expressed maximum inter cluster distance with most of the other clusters. The utility of multivariate analysis in quantifying the degree of divergence between populations so as to understand the trend of their evolutionary pattern and assess the relative contribution of different components to total divergence together with the nature of forces operating at intra- and inter-cluster level are greatly emphasized (Anand and Murty, 1968; Mishra *et al.,* 1994). Large inter-cluster distances signify that the accessions grouped in a cluster are different from the accessions of other clusters for one or more characters, which made them so divergent from others (Yadav *et al.,* 2010). The pattern of group constellation was based on species irrespective of their geographical origin and therefore, could be attributed to factors like heterogeneity, genetic architecture of the population, past history of selection development trait and degree of general combining ability (Murty and Aranachalam, 1966). Cluster-I recorded highest mean value for most of the traits such as Plant height (106.12 cm), Green fodder yield 33.73 (g/plant), 1000-seed weight (2.44 g) and Leaf crude protein content. The trait green fodder yield happened to be the major trait with 24% contribution to the total divergence followed by seed yield plant^{-1} (18.50%). The information helps in selecting parents for the trait which explains total divergence. Selection of the parents for hybridization should be done from different clusters having wide inter-cluster distance and those selected parents should have high *per se* performance for the traits contributing maximum towards divergence (Singh *et al*., 1996). Thus, from the present study crosses between the genotypes of cluster I with IV, genotypes of cluster V with VI and that of I with II are likely to exhibit right heterosis.

Table 1. The primer sequence and linkage information of microsatellite markers used in assessment of genetic diversity

Name of marker	Primer forward (5-3)	Primer Reverse (5-3)	Linkage group	Annealing temperature
AFct45	TAAAAAACGGAAAGAGTTGGTTAG	GCCATCTTTTCTTTTGCTTC	7	52
AFat15	TTACGGGTCTAGATTAGAGAGTATAG	CAAAATGAGTATAGGGAGTGG	4	51
MTIC432	TGGAATTTGGGATATAGGAAG	GCCATAAGAACTTCCACTT	7	55
MTIC93	AGCAGGATTTGGGACAGTTGT	ACCGTAGCTCCCTTTTCCA	6	55
MTIC 297	CTAAGCTTTGGCCATGTATC	TGAAATGAGTTTGACTGAGG	2	50
MTIC249	TAGGTCATGGCTATTGCTTC	GTGGGTGAGGATGTGTGTAT	8	55
MTIC234	GGATGATTCCCTAAATTCAA	AAAAGGAGAATTTATTCATTTCA	7	55
MTIC430	GCGTCTTTTCTTCATTTCAC	TGATAGCCATAACTCCGAAT	5	55
AFca1	CGTATCAATATCTCGGGCAG	TGTTATCAGAGAGAGAAAGCG	4	52
AFct11	GGACAGAGCAAAAGAACAAT	TTGTGTGGAAAGAATAGGAA	6	52
MTLEC2A	CGGAAAGATTCTTGAATAGATG	TGGTTCGCTGTTCTCATG	3	51
AFca16	GGTCGAACCAAGCATGT	TAAAAAACATTACATGACCTCAAA	8	52

Table 2. Agro-morphological traits (mean values) recorded on *Medicago* populations in Ladakh

Species	Days to 50 % flowering	Plant height (cm)	Leaf fresh weight (g)	Leaf dry weight (g)	Number of flowers/ raceme	Number of pods/ raceme
Medicago falcata	118.84	48.47	9.86	3.52	20.44	11.97
Medicago sativa	91.63	82.15	15.02	4.98	23.96	16.48
Medicago varia	99.65	104.85	16.9	5.60	23.07	16.21

Species	Number of flowers on main raceme	Number of pods on main raceme	1000-seed weight	Seed yield per plant	Green fodder yield per plant	Leaf crude protein content (%)
Medicago falcata	20.44	11.97	1.96	1.56	19.73	20.22
Medicago sativa	23.96	16.48	1.94	2.22	30.07	18.05
Medicago varia	23.07	16.21	2.40	1.96	33.81	23.79

Table 3. Average inter cluster (above diagonal) and intra cluster (diagonal) distances based on Mahalanobis D^2 statistic

No. of genotypes	Name of the genotype	Cluster	I	II	III	IV	V	VI	VII
11	MV-5, MV-14, MV-12, MV-10, MV-3, MV-4, MV-7, MV-6, MV-13, MV-8, MV-2	I	**10.43**	663.66	412.29	144.96	854.02	30.72	306.48
9	MF-3, MF-15, MF-13, MF-18, MF-9, MF-4, MF-5, MF-12, MF-19	II	-	**19.74**	50.22	415.08	30.52	647.31	381.35
9	MF-1, MF-6, MF-20, MF-16, MF-7, MF-17, MF-14, MF-8, MF-10	III	-	-	**15.82**	265.03	109.60	408.57	283.26
10	MS-5, MS-9, MS-4, MS-10, MS-6, MS-8, MS-3, MS-7, MS-2, MS-11	IV	-	-	-	**24.60**	541.15	97.95	49.85
2	MF-2, MF-11	V	-	-	-	-	**6.08**	833.92	477.84
3	MV-9, MV-11, MV-1	VI	-	-	-	-	-	**37.65**	221.76
1	MS-1	VII	-	-	-	-	-	-	**0.00**

Cluster means of agronomic traits: Cluster I: Plant height (106.12 cm), Green fodder yield 33.73 (g/plant), 1000-seed weight (2.44 g) and Leaf crude protein content (23.93%); Cluster IV: Number of pods on main raceme (15.93); Seed yield per plant (2.27 g); Cluster VII: Seeds per pod (5.00)

Analysis of molecular variance (AMOVA) and diversity: Significantly high microsatellite diversity was recorded across *Medicago* populations. Here 20 genotypes belonged to *Medicago falcate* (included collections MF-1 to MF-20); 11 genotypes belonged to *Medicago sativa,* designated as MS1 to MS11 and the rest 14 genotypes belonged to *Medicago varia* designated as MV-1 to MV-14. Significant variability existed at marker loci across alfalfa populations and in major part this was contributed by variability among individuals across populations (96.76%), while little within group variability was recorded (Table 4). Thus, AMOVA revealed that enough of allele sharing has taken place among species. The variability studies using RAPD and SSR alleles were carried out among Italian ecotypes and cultivars of alfalfa (Mengoni *et al*., 2000). Kolliker *et al.* (1999) in *Dactylis glomerata* found about 85.1 per cent of variation among genotypes within cultivars. leira *et al.* (2004) reported that 98 per cent of total divergence in annual ryegrass was attributable to that of intra-population diversity. Similarly, AMOVA in diverse red clover breeding population and cultivars revealed high variability (80.4%) within population (Ulloa *et al*., 2003).

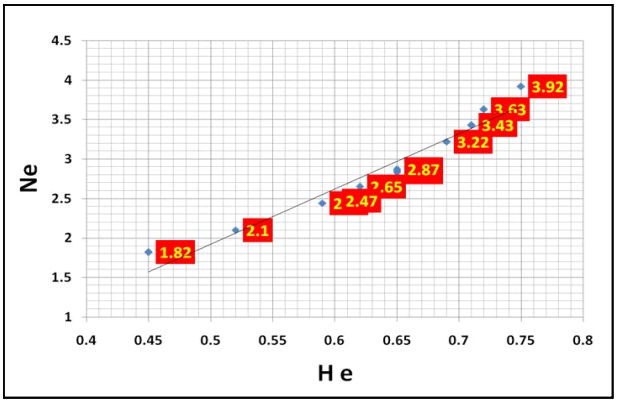
Table 4. Analysis of molecular variance (AMOVA)

Source of variation	Sum of squares	Variance component	Percentage variant	P-value
Among population	7.79	0.0064	3.23	
Within individuals	188.21	2.7559	96.76	0.0889
Total	196.00	2.7623	99.99	

Name of markers from top to bottom: Afca1; Afc1 15; MTIC 93; Lanes 1-45: Genotypes; Lane M: 100 bp size standard (Fermentas, Luthuania, U S) (The brighter band is 300 bp)

Fig 1. Microsatellite allelic profile of individuals representing alfalfa populations in Ladakh

Number of alleles identified and polymorphism of SSR primers: A varied allelic pattern was generated on a set of populations within each of the three sub-species. The 12 SSR markers detected a total of 61 alleles in 45 accessions of alfalfa. The number of alleles per locus ranged from three (MTIC-234, MTIC-430) to 8 (MTIC-432) with an average of 5.8 alleles per locus (Fig 1; Table 5). The polymorphic information content (PIC) value ranged from 0.538 (MTIC-432) to 0.742 (Afat-15) with an average of 0.637. Major allelic frequency ranged from 0.28 (Afat-15) to 0.86 (MTIC-297). Analysis of few individual markers revealed distinct allelic profile across three species. For instance, the frequency of 155 bp fragment at marker locus Afat15 was 0.60 in *M. sativa* compared to only 0.30 and 0.10 in *M. varia* and *M. falcata*, respectively. Similarly, 160 bp allele amplified only in *M. varia* and was absent in other two species. The 150 bp band was amplified in 30% of *M. varia* and 20% of *M. falcata* population while, it was altogether absent in *M. sativa*. With respect to marker MTIC-432, 175 bp allele was amplified across all the three sub-populations. The 179 bp allele was present in about 20% *M. falcata* and *M. sativa* species but was absent in *M. varia*. The marker AFCT-45 amplified four alleles out of which 145 base pair fragment was missing in *M. sativa* and rest were common. The allele 115 bp at marker loci MTIC-297 showed 80% distribution in *M. falcata*, 73% distribution in *M. sativa* and 50% distribution in *M. varia*. Thus, varied allelic diversity appeared among populations at some of the marker loci. The loss of allele or very low or high allelic frequency in one population compared to others may be sometimes due to sampling or chance error. The major allelic frequency ranged from 0.28 for marker Afat-15 to 0.86 for marker MTIC-297 with minor allelic frequency as low as 0.02 for marker AFCa-1. As per Hale *et al.* (2012), the accuracy of sampling frequencies of rare alleles (frequency between 0.01 and 0.05) happens to be slightly affected by sample size, while the accuracy of sampling frequencies of very rare alleles (real frequency, 0.01) was found to be unaffected by sample size. They argue that very rare alleles were unlikely to be sampled. Therefore, one more dimension adds to total genetic diversity which could be investigated more appropriately is the level of heterozygosity rather presence or absence of allele(s) *per se*.



Ne or Effective no of alleles = [1/ 1 - He]. Markers beyond Ne = 3.0 have high discriminatory power where number of alleles tend to equal number of effective alleles.

Fig 2. Trend line of Number of effective alleles (Ne) versus Expected heterozygosity

Table 5. Information on Number of alleles and Polymorphic Information Content at SSR Loci

Locus	SSR motif (repeat sequence)	Alleles per locus	PIC value				Major allelic frequency	Minor allelic frequency
			M. falcata	*M. sativa*	*M. varia*	Mean		
AFCT-45	$(CT)_8 AT (CT)_3$	4	0.589	0.531	0.640	0.597	0.72	0.03
Afat-15	$(AT)_{23}$	6	0.724	0.500	0.734	0.747	0.28	0.03
MTIC-432	$(AG)_6$	8	0.523	0.611	0.438	0.538	0.41	0.03
MTIC-93	$(TTC)_6$	7	0.689	0.574	0.510	0.629	0.85	0.03
MTIC-297	$(TAC)_5$	5	0.451	0.568	0.616	0.574	0.86	0.03
MTIC-249	$(TCA)_5$	5	0.708	0.375	0.681	0.634	0.53	0.06
MTIC-234	$(ATT)_7$	3	0.622	0.594	0.500	0.591	0.56	0.21
MTIC-430	$(AG)_5$	3	0.646	0.599	0.580	0.613	0.63	0.30
AFCA-1	$(CT)_4 (CA)_{10}$	4	0.649	0.660	0.597	0.677	0.76	0.02
AFCT-11	$(CT)_{12}$	6	0.595	0.611	0.622	0.693	0.48	0.03
MTLEC-2A	$(AT)_{19}$	5	0.651	0.622	0.628	0.644	0.64	0.03
AFCA-16	$(CA)_{12}$	5	0.745	0.586	0.727	0.706	0.48	0.13
Total number of alleles		**61**						
Alleles per locus		**5.08**						

Table 6. The relative measures of observed (H_o) and expected (H_e) heterozygosity and inbreeding coefficient (F) for *Medicago* populations

Marker	*M. falcata*					*M. sativa*					*M. varia*				
	H_o	H_e	F	χ^2	Prob	H_o	H_e	F	χ^2	Prob	H_o	H_e	F	χ^2	Prob
AFcl45	0.57	0.59	0.03	4.06	0.67	0.25	0.53	0.53	8.72*	0.03	0.50	0.64	0.22	19.92**	0.00
AFat15	0.14	0.72	0.8	31.05**	0.00	0.57	0.50	-0.14	3.11	0.38	0.25	0.73	0.66	26.88**	0.00
MTIC432	0.69	0.52	-0.31	4.39	0.93	0.44	0.61	0.27	10.44	0.11	0.54	0.44	-0.23	1.77	1.00
MTIC93	0.75	0.69	-0.09	9.22	0.16	0.33	0.57	0.42	12.58	0.05	0.50	0.51	0.02	15.90	0.39
MTIC297	0.56	0.45	-0.25	2.45	0.87	0.56	0.51	0.02	4.77	0.57	0.36	0.62	0.41	13.25	0.04
MTIC249	0.00	0.71	1.00	36.00**	0.00	0.00	0.38	1.00	8.00**	0.01	0.00	0.68	1.00	48.00**	0.00
MTIC234	0.00	0.62	1.00	28.00**	0.00	0.00	0.59	1.00	16.00**	0.00	0.00	0.50	1.00	24.00**	0.00
MTIC430	0.60	0.65	0.07	0.33	0.95	0.18	0.60	0.70	10.33**	0.02	0.33	0.58	0.43	6.51	0.09
AFCa1	0.55	0.65	0.15	7.90	0.25	0.78	0.66	-0.18	7.88	0.25	0.75	0.60	-0.26	2.85	0.83
ATCt11	0.24	0.60	0.61	20.67**	0.00	0.44	0.61	0.27	4.69	0.58	0.36	0.62	0.43	19.90*	0.03
MTLEC2A	0.33	0.65	0.49	15.03*	0.02	0.57	0.62	0.08	3.11	0.80	0.55	0.63	0.13	6.44	0.38
AFCa16	0.50	0.75	0.33	19.76*	0.03	0.38	0.59	0.36	3.48	0.32	0.18	0.73	0.75	23.91**	0.00
Mean	**0.41**	**0.63**	**0.32**			**0.38**	**0.56**	**0.36**			**0.36**	**0.61**	**0.38**		

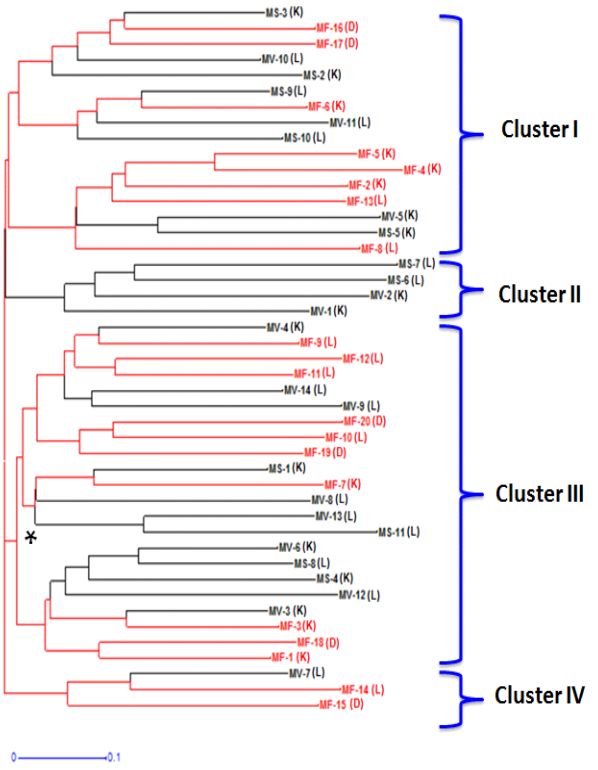
Overall measures of heterozygosity across populations: Ho: 0.39; He: 0.62 and F: 0.35

Estimation of heterozygosity observed (H_o) and expected (H_e), inbreeding coefficient (F) and Hardy-Weinberg equilibrium for the three Medicagosub-populations: H_o ranged from 0% for markers MTIC249 and MTIC234 to 0.75% for marker MTIC93 followed by MTIC432 (0.69%) and MTIC430 (0.60%), (Table 6). The H_e ranged from 0.45% for marker MTIC297 to 0.75% for AFca16. The highest inbreeding coefficient of 1.00 was shown by the markers MTIC249 and MTIC234 and lowest inbreeding coefficient of 0.03 by AFct45. The markers AFat15, MTIC249, MTIC234, ATCt11, MTLEC2A and AFCa16 showed significant deviation from Hardy Weinberg Equilibrium (HWE). The inbreeding coefficient (F) in *Medicago sativa* ranged from 0.02 to 1.00. The markers AFCt45, MTIC249, MTIC234 and MTIC430 showed significant deviation from HWE. The markers AFCt45, AFat15, MTIC297, MTIC249, MTIC234, ATct11, AFca16 showed significant deviation from HWE. The comparison of coefficient of inbreeding among three sub-populations revealed that *M. varia* showed maximum inbreeding coefficient (0.38) as compared to other two sub-populations *M. falcata and M. sativa.*

In the present study, the HWE was observed at the loci AFcl45, MTIC432, MTIC93, MTIC297, MTIC430, AFCa1 in case of *M. falcata*. In *M. sativa* the HWE was recorded at Afat15, MTIC432, MTIC93, MTIC297, AFCa1, ATCt11, MTLEC2A, AFCa16. Yet in *M. varia,* the markers MTIC432, MTIC93, MTLEC2A, AFCa1, MTIC430 were in HWE. Of these MTIC432, MTIC93 and AFCa1 were common to all the three species. The lack of commonality for other loci indicates the intra-population gene sharing within the respective species. One of the important consequences of population substructure is decrease in the level of heterozygosity. The cases where H_o was less than H_e indicated considerable degree of inbreeding. Here the two loci MTIC249 and MTIC234 exhibited complete fixation (F=1) across all three species. Interestingly, the number of alleles per locus for these loci are five and three, respectively, which may not be called very low and a reason to effect fixation. In addition, PIC value of MTIC249 is very low in *M. sativa* as compared to other two. Overall, the inbreeding coefficient was comparably lower in *M. falcata.*

Population genetic differentiation across three Medicago sub-populations based on SSR assay: The F_{ST} values which define fixation index with respect to 12 SSR loci were calculated for 45 genotypes of alfalfa and were highly significant ($p<0.001$), (Table 7). The SSR markers AFat15, MTIC129, AFca1 and ATCt-11 showed

F_{ST} =0.10 compared to average of 0.045 across all the markers. In the present study, the markers AFat15, MTIC297, ATCt11 recorded relatively higher F_{ST} values 0.10, 0.10 and 0.12, respectively. This suggests that only 10-12 per cent of differentiation at corresponding loci and explains population structure (species in present case). Individuals within species contributed 88-90% of this differentiation. At locus MTIC430, the three species are completely identical. Overall 4.5% genetic variation was attributed to species grouping and 95.5 per cent exists within species. The F_{IT} and F_{IS} values at locus MTIC432 were negative and for loci MTIC249 and MTIC234 both the values were equal to unity. The negative values suggest the higher frequencies of H_e than expected and F_{IT} or F_{IS} equal to one means the two loci are fixed for different alleles. F_{ST} of 0.10 means that of the total genetic differentiation based on allelic frequency only 10 per cent is explained by sub-population structure i.e. sub-species in our case and 90 per cent is attributed to heterozygosity among individuals throughout. This explains that sub-populations have very recent history of separation or enough gene sharing is taking place among three sub-population.



*denotes the node at a genetic distance of 0.87

Fig 3. UPGMA based molecular diversity analysis of 45 alfalfa genotypes across three sub-species of *M. falcata, M. varia and M. sativa*

Number of effective alleles (N_e): The number of effective alleles (N_e) in *M. falcata* ranged from 1.82 to 3.92 (Table 7). The highest N_e was shown by the marker AFCa-16 (3.92) and lowest by the marker MTIC-297 (1.82). In *M. sativa* N_e ranged from 1.60 (MTIC-249) to 2.95 (AFca-1). The N_e in *M. varia* ranged from 1.78 (MTIC-432) to 3.77 (AFat-45). Among the three sub-species of alfalfa, *M. falcate* showed maximum mean value for N_e (2.85) followed by *M. varia* (2.66) and lowest by *M. sativa* (2.38). The trend line of number of effective alleles (N_e) versus expected heterozygosity (H_e), indicated that the markers beyond N_e=3.0, have high discriminatory power where the number of actual alleles tend to equal the number of effective alleles (Fig 2). From the data of mean N_e and alleles per locus, the two had very poor correlation (-0.08^{ns}). The allele number per locus will have certainly low effect on population differentiation when compared to effective number of alleles. For example, the marker MTIC432 has highest number of alleles (8) but N_e only equal 2.15, surprisingly the lowest among all the twelve markers. Contrary to this MTIC430 had N_e = 2.57 for only three alleles. High N_e was recorded for markers Afat-15 (3.13) and AFCA-16 (3.31). This shows the balanced distribution of the alleles at the locus which should result in high PIC, as certainly was the case for these two loci with only having PIC > 70%. Otherwise, if only one allele predominates, that would result in low N_e ~1 and therefore, low PIC. N_e is a nonlinear function of the gene diversity (H_{exp}), which brings into play Jensen's inequality i.e. N_e will be highest between H_e=0.5 to 0.9. In present study, highest N_e values of 3.43, 3.63 and 3.92 were recorded against H_e of 0.71, 0.72 and 0.75. Markers beyond N_e = 3.0 have high discriminatory power. In such a case the number of alleles tends to equal number of effective alleles. In *M. falcata* populations, markers AFat15, MTIC249, AFCa16 showed N_e> 3.0. Similarly, in *M. varia* Afat15, MTIC249 and AFCa16 had N_e> 3.0. None of the markers had N_e>3.0 in *M. sativa* which is because of relatively greater contribution of few of the alleles over others.

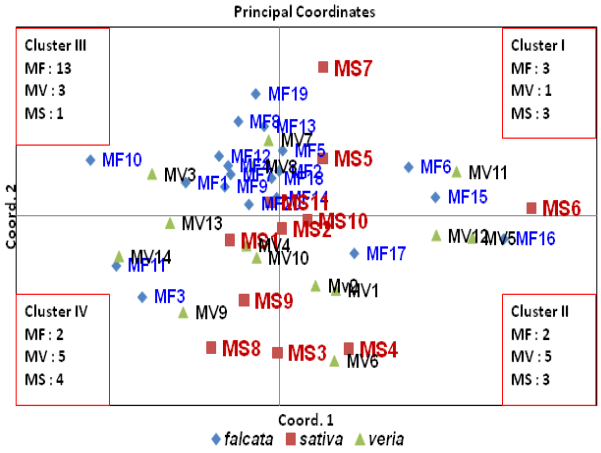
Analysis of molecular diversity: The dissimilarity matrix based on 12 SSR markers ranged from 0.33 to 0.95 with a mean of 0.64. Of the pair wise combinations generated, MF3 and MS7 showed highest dissimilarity index (0.95). The cluster analysis based on UPGMA across 45 genotypes generated four major clusters with three sub clusters in cluster-I having five, four and seven genotypes each. Cluster-II had four genotypes, Cluster-III had three sub clusters, each having nine, five and eight genotypes. Cluster-IV possessed only three genotypes (Fig 3). The

Table 7. Population genetic differentiation across all the three sub-populations of *Medicago* based on SSR analysis

Marker	He	Ho	HI	Hs	Ht	F_{IS}= (1-HI/Hs)	F_{IT}= (1-HI/HT)	F_{ST}= (1-Hs/HT)	$N_{e\ (falcata)}$	$N_{e\ (sativa)}$	$N_{e\ (varia)}$
AFcl45	0.59	0.44	0.47	0.59	0.60	0.20	0.21	0.01	2.44	2.13	2.78
AFat15	0.65	0.32	0.28	0.67	0.75	0.58	0.62	0.10	3.63	2.00	3.77
MTIC432	0.52	0.56	0.58	0.52	0.54	-0.12	-0.08	0.04	2.10	2.57	1.78
MTIC93	0.59	0.53	0.57	0.61	0.63	0.06	0.09	0.04	3.22	2.35	2.04
MTIC297	0.54	0.49	0.50	0.52	0.57	0.03	0.13	0.10	1.82	2.31	2.60
MTIC249	0.59	0.00	0.00	0.62	0.63	1.00	1.00	0.03	3.43	1.60	3.14
MTIC234	0.57	0.00	0.00	0.58	0.59	1.00	1.00	0.02	2.65	2.46	2.00
MTIC430	0.61	0.37	0.41	0.61	0.61	0.32	0.32	0.00	2.83	2.50	2.38
AFCa1	0.64	0.69	0.67	0.64	0.68	-0.05	0.01	0.06	2.85	2.95	2.48
ATCt11	0.61	0.35	0.32	0.61	0.69	0.47	0.53	0.12	2.47	2.57	2.65
MTLEC2A	0.63	0.48	0.46	0.64	0.64	0.28	0.29	0.01	2.87	2.65	2.69
AFCa16	0.69	0.35	0.37	0.70	0.71	0.47	0.48	0.01	3.92	2.42	3.61
Mean	**0.60**	**0.38**	**0.38**	**0.61**	**0.63**	**0.35**	**0.38**	**0.04**	**2.85**	**2.38**	**2.66**

F_{ST} scale: 0.0 to 0.15: Low genetic differentiation; 0.15 to 0.25: Moderate genetic differentiation > 0.25: High genetic differentiation; Ne or Effective no of alleles = [1/ 1 - He]

original alfalfa of Ladakh, *M. falcata* is yellow flowered. A lot of hybridization has taken place and now it is rare to find a true stand of either *M. falcata* or *M. sativa* in the backdrop of wide range of variability in habit, leaf shape, height, colour of flower, shape of pod and the resistance of plant to cold and aridity.



Percentage of variation explained by the first three Principal Coordinates: PC1: 20.65%; PC2: 18.86%; PC3: 17.50%

Fig 4. Principal coordinate analysis of 45 alfalfa genotypes across three sub-species of *M. falcata, M. varia and M. sativa*

Principle coordinate analysis: The PCA analysis was applied to explain the genetic diversity with the sub-population of alfalfa in Ladakh region (Fig 4). The PC-1and PC-2 accounted for 20.65% and 18.86% of total variability, respectively. The cluster analysis performed on the first three components grouped the alfalfa sub-population into four clusters. Cluster-I contain seven genotypes, cluster-II was composed of ten genotypes, cluster-III withthe highest 16 genotypes and cluster-IV carried 12 genotypes. The PCA reduces the dimensions of data matrix into few meaningful axis by singular value decomposition. PCA is comparable to PCOA when there is no missing data and number of individuals more than the number of clusters (Rohlf, 1972). The discordance in clustering through PCA based on morphological versus molecular data was as expected because of the obvious genetic differences among SSR loci, which are evolutionary less conserved than morphological trait loci. Secondly, the number of SSR data points against morphological traits, which amounts to number of alleles, differs between the two, which may have a relational bias in clustering pattern. The PCs with values >1 should be utilized to derive clustering (Lezzoni and Psitts, 1991). To extract maximum information from molecular data, PCA or PCoA can be used in combination with cluster analysis particularly when first three PCs explain more than 1 per cent variation (Messmor and Raines, 2000). The method was used in the present study for grouping of four clusters based on SSR data. In a similar study carried by Sood *et al.* (2016), PCA helped to classify twenty-five genotypes of *Trifolium* belonging to seventeen species into distinct groups and it matched to the analysis based on tochers method.

Conclusion

Significant variability was recorded within sub-populations for various morphological traits and at molecular level. High level of heterozygosity was recorded

at marker loci; however, this could not explain the population sub-structure based on three species. Number of effective alleles was found to be high in *M. falcate* that influenced the clustering pattern based on dice coefficient and PCA. Finally, the three sub-species did not maintain a definite population structure which indicates sufficient gene flow among these populations and therefore, high divergence among individuals within a group. Therefore, this warrants that hybridization followed by selection should be practiced based upon the utilization of useful genetic variation across the species and cannot be progressed by confinement to any individual geographical location or particular species. Besides, to carry out meaningful population improvement in alfalfa, detailed cytogenetics study is warranted in order to workout and validates the sub-species categorization of populations in the region.

Acknowledgement

The first author acknowledges UGC, Govt. of India for financial assistance in pursuing the research objectives.

References

Anand, I. J. and B. R. Murty. 1968. Genetic divergence and hybrid performance in linseed. *Indian Journal of Genetics* 28: 178-185.

Bolton, J. L. 1962. *Alfalfa*: *Botany, Cultivation and Utilization.* Inter-science Publishers, New York, USA.

Botstein, D., R. L. White, M. Skolnick and R. W. Davis. 1980. Constructions of genetic linkage map in non-using restriction fragment length polymorphism. *American Journal of Human Genetics* 32: 314-333.

Excoffier, L., P. E. Smouse and M. Quattra. 2005. Analysis of molecular variance inferred from metric distance among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479-491.

Flajoulat, S., J. Ronfort, P. B. Bandouin, T. Huguet, C. Huyghe and B. Juliea. 2005. Genetic diversity among alfalfa (*Medicago sativa*) cultivars coming from a breeding program, using SSR markers. *Theoretical and Applied Genetics* 111: 1420-1429.

Hale, M. L., T. M. Burg and T. E. Steeves. 2012. Sampling for Microsatellite Based Population Genetic Studies: 25 to 30 Individuals per Population Is Enough to Accurately Estimate Allele Frequencies. *PLOS ONE 7(9)*: e45170.

Kolliker, R., F. J. Stadelmann, B. Reidy and J. Nosberger. 1999. Genetic variability of forage grass cultivar: A comparison of *Festuca pratensis*, *Lolium perenne* L. and *Dactylis glomerate* L. *Euphytica* 106: 261-270.

Leira, E. A., C. M. Castro, A. C. Oliveira, F. I. F. Carvalho, P. D. Zimmer and L. F. Martins. 2004. Genetic structure of annual ryegrass (*Lolium multiflorum*) populations estimated by RAPD. *Sci Agric* (Piracicaba, Braz) 4: 407–413.

Lezzoni, A. F. and M. P. Pritts. 1991. Applications of Principle component Analysis to Horticulture Research. *Horticultural Science* 26: 334-338.

Maroof, M. A., K. M. Saliman, R. A. Jorgensen and R. W. Anand. 1984. Ribosomal DNA spacer length polymorphisms in Barley: species diversity, chromosomal locations and population dynamics. *Proceeding of National Academy of Sciences (PNAS)* 91: 5466-5470.

Maureira, I. J., F. Ortega, H. Compos and T. C. Osborn. 2004. Population structure and combining ability of diverse *Medicago sativa* germplasm. *Theoretical and Applied Genetics* 109: 775-782.

McCoy, T. J. and E. T. Bingham. 1988. Cytology cytogenetics of alfalfa. In: A.A. Hanson, O.K. Barnes and R.R. Hill (eds.). *Alfalfa and Alfalfa Improvement.* American Society of Agronomy, Inc. Publishers, Madison, Wisconsin, USA. pp. 1084.

Mengoni, A., A. Gori and M. Bazzicalupo. 2000. Use of RAPD and microsatellite (SSR) variation to assess genetic relationships among populations of tetraploid alfalfa *Medicago sativa*. *Plant Breeding* 119: 311–317.

Messmore, J. M. and R. T. Raines. 2000. Pentavalent organo-vanadates as transition state analogues for phosphoryl transfer reactions. *Journal of the American Chemical Society* 122: 9911–9916.

Michaud, R., E. F. Lehman and M. D. Rambaugh. 1988. World distribution historical development. In: A. A. Hanson, D. K Barnes and Jr. R. R. Hill (eds.). *Alfalfa and Alfalfa Improvement*. American Society of Agronomy, Inc. Publishers, Madison, Wisconsin, USA, pp. 1084.

Mishra, S. B., U. K. Singh and P. K. Sadu. 1994. Studies on genetic divergence in five land races of rice in northern Bihar. *Annals of Agriculture Research* 15: 217-221.

Misri, B. 1987. Disjunct distribution of *Avena fatua* as a weed in Ladakh. *Oat News Letter.* pp. 37.

Murty, B. R. and V. Arunachalam. 1966. The nature of genetic divergence in relation to breeding systems in crop plants. *Indian Journal of Genetics* 26: 188-198.

Musial, J. M., K. E. Basford and J. A. G. Irwin. 2002. Analysis of genetic diversity within Australian Lucerne cultivars and implications for future genetic improvement. *Australian Journal of Agriculture Research* 53: 629–636.

Peakall, R. and P. E. Smouse. 2001. GenAlEx V5: Genetic Analysis in Excel. Population genetic software for teaching and research. Australian National\ University, Canberra, Australia. http:// www.anu.edu.au/BoZo/GenAlEx.

Quiros, C. F. and G. R. Bauchan. 1988. The genus Medicago and the origin of the *Medicago sativa* complex. In: A. A. Hanson, D. K Barnes and Jr. R. R. Hill (eds.). *Alfalfa and Alfalfa Improvement.* American Society of Agronomy, Inc. Publishers, Madison, Wisconsin, USA. pp. 1084.

Rao, C. R. 1952. *Advanced Statistical Method in Biometrical Research*. John Willey and Sons Inc. New York.

Rohlf, F. J. 1972. An empirical comparison of three ordination techniques in numerical taxonomy. *Systematic Zoology* 21: 271-280.

Segovia-Lerma, A., R. G. Cantrell, J. M. Conway and I. M. Ray. 2003. AFLP-based assessment of genetic diversity among nine alfalfa germplasms using bulk DNA templates. *Genome* 46: 51–58.

Singh, A. K., S. B. Singh and S. M. Singh. 1996. Genetic divergence in scented and five genotypes of rice (*Oryza sativa* L.). *Annals of Agricultural Research* 17: 163-166.

Singh, R. K. and B. D. Chaudhary. 1985. *Biometrical Methods in Quantitative Genetic Analysis.* Kalyani Publisher, India. pp. 318.

Sood, V. K., R. Dhiman, A. Pathania, J.C. Bhandaria and H.K. Chaudhary. 2016. Genetic diversity based on mutilvariate analyses for breeding strategies in *Trifolium. Range Management and Agroforestry* 37: 175-180.

Ulloa, O., F. Ortega, and H. Camposh, 2003. Analysis of genetic diversity in red clore (*Trifolium pretense* L.) breeding populations as revealed by RAPD genetic markers. *Genome* 46: 529-535.

Vandemark, G. J., T. J. Hughes, and R. C. Larsen. 2005. Genetic similarities between alfalfa cultivars based on sequence related amplified polymorphism (SRAP) DNA markers. In: Proc. *Plant & Animal Genomes XIII Conference* (January 15-19, 2005*)*, San Diego, CA, USA.

Whyte, R. O., L. Nilsson and H. C. Trumbee. 1953. *Legumes in Agriculture*. FAO, Rome.

Yadav, V.P., G.P. Shukla, R.V. Kumar and H.C. Pandey. 2010. Genetic divergence across the cutting management in alfalfa (*Medicago sativa* L.). *Range Management and Agroforestry* 31: 76-78.