Plant Mol Biol Rep (2012) 30:286–296
DOI 10.1007/s11105-011-0338-x

Retroelement Insertional Polymorphism and Genetic Diversity in *Medicago sativa* Populations Revealed by IRAP and REMAP Markers

Babak Abdollahi Mandoulakani · Yaser Piri · Reza Darvishzadeh · Iraj Bernoosi · Morad Jafari

Published online: 27 July 2011

Abstract Retrotransposons (RTNs) are common components of plant genomes, showing activity at transcription and integration levels. Their abundance, dispersion, ubiquity, and ability to transpose make them useful as molecular markers. In this study, we used multilocus PCR-based techniques, inter-retrotransposon amplified polymorphism (IRAP), and retrotransposon-microsatellite amplified polymorphism (REMAP) to study the integration events of native (Tms1Ret1) and non-native RTN (LORE1, LORE2, Tps12a, and Tps19) families in *M. sativa* genome. IRAP and REMAP markers derived from these RTNs were used to assess genetic diversity among and within 80 *M. sativa* genotypes belonging to eight populations. Results indicated the presence and transpositional activation of RTNs Tms1Ret1, LORE1, LORE2, and even Tps12a in *M. sativa* genome. REMAP analysis showed the insertion of studied RTN families near the microsatellites in *M. sativa* genome except for RTN Tps19. A total of 101 and 119 loci were amplified using 10 and 14 IRAP and REMAP primers, respectively. The number of polymorphic loci was 66 and 62 for IRAP and REMAP, respectively. Mantel test between IRAP and REMAP cophenetic matrices evidenced no significant correlation (r=0.15). Although populations could be relatively differentiated based on IRAP + REMAP data, overall genetic differentiation among populations was low (PhiPT= 0.082, P=0.0010). Analysis of molecular variance based on IRAP + REMAP analysis revealed the higher level of genetic variation within populations (92%) compared to among populations (8%). IRAP + REMAP-based cluster analysis of 80 *M. sativa* genotypes using complete linkage algorithm identified five heterotic groups that could be applied as crossing parents in *M. sativa* alfalfa breeding programs.

Keywords *Medicago sativa* L. · Transpositional activity · LORE1 · Tms1Ret1 · Tps12a

B. Abdollahi Mandoulakani · R. Darvishzadeh · I. Bernoosi · M. Jafari
Department of Agronomy and Plant Breeding,
Faculty of Agriculture, Urmia University,
Urmia, Iran

B. Abdollahi Mandoulakani (✉) · R. Darvishzadeh · I. Bernoosi · M. Jafari
Department of Agricultural Biotechnology,
Institute of Biotechnology, Urmia University,
Urmia, Iran
e-mail: b.abdollahi@urmia.ac.ir

Y. Piri
Department of Agronomy and Plant Breeding,
Faculty of Agriculture, Islamic Azad University of Tabriz,
Tabriz, Iran

Introduction

Alfalfa (*Medicago sativa* L.) is the most cultivated forage legume, originated in the Caucasus region: northeastern Turkey, Turkmenistan, and northwestern Iran (Michaud et al. 1988). Tetrasomic inheritance associated with an allogamous breeding system and pronounced inbreeding depression complicates genetic analyses and characterization of alfalfa compared with other major crops (Robins et al. 2007). The use of heterosis in alfalfa as a means to improve forage yield and other important traits has been demonstrated. Advances in genome research have generated interest in predicting hybrid performance using molecular markers in crop breeding programs (Diers et al. 1996; Bourguiba et al. 2010). Different molecular marker systems have been used to assess genetic diversity and hybrid performance in alfalfa, including restriction fragment length

polymorphism (Kidwell et al. 1994b), random amplified polymorphic DNA (Mengoni et al. 2000), simple sequence repeats (SSR) (Diwan et al. 1997; Falahati-Anbaran et al. 2007), and amplified fragment length polymorphism (AFLP) (Zaccardelli et al. 2003). Investigations in maize, rice, and oilseed rape have shown that the molecular genetic diversity of parents was significantly correlated with hybrid performance (Riaz et al. 2001; Betran et al. 2003). Significant positive associations have been reported between the genetic diversity of parental alfalfa genotypes and the forage yield of their respective single-cross progenies (Kidwell et al. 1994a, 1999). Brummer (1999) proposed that the identification and maintenance of distinctly divergent populations may be necessary to develop heterotic groups in alfalfa. Riday et al. (2003) found negative correlation between AFLP genetic distance and heterosis in crosses of *M. sativa* subsp. *sativa* and *M. sativa* subsp. *falcata*.

Several molecular marker methods based on retrotransposons (RTNs) have been developed (Waugh et al. 1997; Flavell et al. 1998; Kalendar et al. 1999). The dispersion, ubiquity, and prevalence of RTNs in plant genomes provide an excellent basis for the development of marker systems. Direct comparisons of RTN methods with other marker systems indicate that the RTN markers are more informative and polymorphic in a variety of crops (Abdollahi Mandoulakani et al. 2008; Queen et al. 2004; Tam et al. 2005). A major disadvantage of RTN-based molecular markers is the need for RTN sequence information to design family-specific primers. However, related species have similar RTN sequences, meaning that primers designed based on long terminal repeat (LTR) sequences of an RTN can be readily used across species lines, among closely related genera, and even sometimes between plant families (Lou and Chen 2007; Kalendar et al. 2011). Porceddu et al. (2002) used Tms1Ret1 LTR-based sequence-specific amplification polymorphism (S-SAP) to assess genetic diversity in genus *Medicago* and reported high marker index for S-SAP compared to AFLP and selective amplification of microsatellite polymorphic loci (SAMPL). Nevertheless, S-SAP relies on restriction digestion and the sensitivity of commonly used enzymes such as *Pst*I and *Eco*RI to DNA methylation, combined with the high and potential variable degree of CG and CXG methylation in plant DNA (Fiuk et al. 2010; Wei et al. 2010), which means that some apparent polymorphism may neither be sequence-based nor heritable. Inter-retrotransposon amplified polymorphism (IRAP) and retrotransposon-microsatellite amplified polymorphism (REMAP), two RTN-based markers, require no DNA digestion and have been applied as molecular markers for genetic diversity studies as well as study of the integration events of RTNs in plant genomes (Kalendar et al. 1999; Abdollahi Mandoulakani et al. 2008, 2009). To our knowledge, no IRAP- and REMAP-based study has been conducted in alfalfa. Therefore, the objective of the present study was to use IRAP and REMAP markers for detecting integration events and activity of native (Tms1Ret1) and non-native (LORE1, LORE2, Tps12a, and Tps19) RTN families in *M. sativa* genome. IRAP and RAMAP markers derived from these RTNs were also used to assess genetic diversity among and within *M. sativa* populations and genotypes in order to identify populations with large genetic distance as well as determine the heterotic groups for potential application in cultivated alfalfa breeding programs.

Materials and Methods

Plant Material and DNA Isolation

Plant materials consisted of 80 genotypes belonging to eight alfalfa populations (10 plants from each population) (Table 1) were kindly provided by the research center of agriculture, west Azarbayejan, Urmia, Iran. Sixty genotypes (six populations) are widely cultivated in different regions of Iran. The remaining 13 studied genotypes (three populations) are originally from Turkey and Azarbayejan and selected for our study because of their good field performance in west Azarbayejan of Iran.

Seeds were planted in small pots with 10 cm diameter containing mixture of garden soil and vermiculite in the greenhouse with an ambient temperature 25°C. Genomic DNA was extracted from young leaves of 25-day seedlings using the method described by Ausubel et al. (1995) with minor modifications. The quality and concentration of the DNA were measured using a spectrophotometer and electrophoresis in a 0.8% (*w*/*v*) agarose gel.

IRAP and REMAP Reactions

Five single and 10 IRAP primer combinations were used to study RTN activity and analyze genetic diversity in 80 alfalfa genotypes. Primers had been designed based on native (isolated from *M. sativa*) and non-native (isolated

Table 1 Studied alfalfa populations and their origin

Population	Number of plants	Origin
Gharayonjeh-Malekkandi	10	Iran-Khoy
Gharayonjeh-Urmia	10	Iran-Urmia
Mahalie-Esfahan	10	Iran-Esfahan
Baghdadi	10	Iran-Kerman
Hamedani	10	Iran-Hamedan
Ordubar-Azarbayejan	10	Azarbayejan
Turkey1	10	Turkey
Turkey-Sakuel	10	Turkey

Table 2 Sequences and numbers of primers used

Primer	Sequence (5′-3′)	Primer	Sequence (5′-3′)
Tms1Ret1	cggttttgtggggttgtgttaggccca	443	acacacacacacacact
LORE1	gagtcctgagtaaccaactaac	459	tgctgctgctgctgctgcc
LORE2	cagcttgaggacaagctgagtc	818	cacacacacacacag
Tps12a	gggctttgactaatggacctc	825	acacacacacacact
A7	agagagagagagagagagt	840	gagagagagagagagaYt
A12	gagagagagagacc	848	cacacacacacacacaRg
B1	tctctctctctctctctcc	849	gtgtgtgtgtgtgtgtcg
438	acacacacacacacacacg	857	acacacacacacacacYg

R purine (A/G), *Y* pyrimidine (C/T)

from *Lotus japonicus* and *Pisum sativum*) RTN families: Tms1Ret1 (Porceddu et al. 2002) from *M. sativa*; LORE1 (Madsen et al. 2005) and LORE2 (Fukai et al. 2008) from *L. japonicus*; and Tps12a and Tps19 (Pearce et al. 2000) from *P. sativum* (Tables 2 and 3). PCR amplifications were carried out in a Bio-Rad thermocycler (Bio-Rad Laboratories Inc., Hercules, CA, USA) in a total volume of 20 µl, containing 45 ng of genomic DNA, 1× PCR buffer (10 mM Tris–HCl, 50 mM KCl, pH=8.3), 1.5 mM $MgCl_2$, 0.2 µM of each dNTP, 0.5 U of Taq DNA polymerase (Cinagen Co., Iran), and 10 pmol of each primer. The amplification profile composed of an initial denaturation at 94°C for 5 min, followed by 35 cycles at 95°C for 45 s, 55°C to 63°C (Table 3) for 40 s, 72°C for 2 min, and a final extension of 5 min at 72°C. PCR products were separated by electrophoresis (Bio-Rad) using 1.8% Resolute™ line Biozyme agarose

Table 3 Used IRAP and REMAP primer combinations, annealing temperature, total loci, polymorphic loci, and size range of amplified loci

Primer name	Annealing temperature	Total loci	Polymorphic loci	Band size (bp)
IRAP				
Tms1Ret1	63	12	9	1,000–3,000
LORE1	58	11	8	700–1,800
LORE2	55	9	6	1,200–2,500
Tps12a	55	4	0	900–2,000
Tms1Ret1- LORE1	60	15	12	250–2,800
Tms1Ret1- LORE2	60	13	8	400–3,000
Tms1Ret1- Tps12a	57	11	10	700–2,500
LORE1-LORE2	58	13	8	300–1,600
LORE1-Tps12a	58	6	5	850–2,500
LORE2-Tps12a	57	7	0	420–1,500
Total		101	66	
REMAP				
Tms1Ret1-459	52	14	11	500–2,500
Tms1Ret1-A7	52	5	0	350–850
Tms1Ret1-A12	52	12	5	250–1,300
Tms1Ret1-B1	52	1	0	700
Tms1Ret1-438	52	8	2	200–1,000
Tms1Ret1-443	52	5	0	500–1,000
Tms1Ret1-818	52	9	3	75–2,000
Tms1Ret1-825	52	5	3	150–700
LORE1-840	52	7	3	200–1,000
LORE1-848	52	11	7	100–1,000
LORE2-849	52	7	6	350–1,000
LORE2-857	52	6	3	200–850
Tps12a-459	52	16	14	200–1,700
Tps12a-438	52	13	5	200–1,400
Total		119	62	

gel in 0.5× TBE buffer with constant voltage of 70 V for 3 h. Gels were stained by ethidium bromide. DNA fragments were then visualized under UV light and photographed using a gel documentation system.

Forty-eight REMAP primer combinations, derived from four single IRAP primers designed based on RTNs (Tms1Ret1, LORE2, LORE1, and Tps12a) with 12 ISSR primers A7, A12, B1, 438, 443, 459, 818, 825, 840, 848, 849, and 857, were tested on four alfalfa genotypes to select the primer combinations producing scorable and discernible banding patterns (Tables 2 and 3). PCR amplification reactions and temperature profile, electrophoresis, and visualization of REMAP markers were as for IRAP. Annealing temperature of REMAP primer combinations depended on the primers (Table 3).

Data Analysis

The amplified fragments were scored independently as 1 or 0 for their presence or absence at each position, and the obtained binary data were used for analysis. Three genetic distance matrices for IRAP, REMAP, and combined data (IRAP + REMAP) were established between populations using Nei genetic distance coefficient (Nei 1978) in GenAlEx 6 (Peakall and Smouse 2006) and subsequently used to construct the three dendrograms based on the unweighted pair group method using an arithmetic average (UPGMA) clustering algorithm implemented in NTSYSpc 2.1 (Rohlf 2000). To verify the adjustment between genetic distance matrices and respective dendrogram-derived matrices (cophenetic matrix), the cophenetic correlation coefficient (r) was estimated. To estimate the degree of correlation among the three cophenetic matrices for IRAP, REMAP, and IRAP + REMAP, a Mantel's test of matrix comparison with 1,000 permutations was performed by using the NTSYSpc 2.1 software (Rohlf 2000). Three Nei genetic distance matrices (Nei 1978) were used to establish dcenter matrices and subsequently used to accomplish principal coordinate analysis (PCO) in NTSYSpc 2.1 (Rohlf 2000) to resolve patterns of genetic relationships of alfalfa populations. Genetic similarity matrix between individual pairs of genotype was calculated for IRAP + REMAP data using Dice coefficient (Nei and Li 1979) in NTSYSpc 2.1 (Rohlf 2000). This matrix was used to construct dendrogram using complete linkage algorithm to identify heterotic groups among 80 genotypes. In order to partition the total genetic variation among and within populations, analysis of molecular variance (AMOVA) was carried out based on IRAP, REMAP, and IRAP + REMAP data using GenAlEx 6 (Peakall and Smouse 2006). Number of loci, percentage of polymorphic loci, number of alleles or loci with a frequency higher or equal to 5%, number of private loci or alleles, number of less common loci with frequency lower or equal to 25% and 50%, mean of heterozygosity (Lynch and Milligan 1994), Nm (Number of migrants between populations), and standard error of mean heterozygosity were also calculated for each population using GenAlEx 6 for IRAP, REMAP, and IRAP + REMAP data.

Result

RTN Activity in *M. Sativa* Genome and IRAP Analysis

Five single and 10 IRAP primer combinations (Table 3) from native and non-native RTN families were used to analyze genetic diversity in 80 alfalfa genotypes. All single IRAP primers designed based on Tms1Ret1 and non-native RTNs LORE1 and LORE2 produced distinguishable and polymorphic banding patterns. Single IRAP primer designed based on non-native RTN Tps12a from *P. sativum* also amplified scorable but not polymorphic banding pattern. IRAP amplification of another non-native RTN Tps19 (from *P. sativum*) generated unscorable banding pattern in low annealing temperature (less than 40°C). IRAP primer combinations from native and non-native used RTN families produced polymorphic and scorable banding pattern except for primer combinations that Tps19 was involved. Ten out of 15 used IRAP primers produced 101 loci. Out of 101 amplified loci, 66 were polymorphic (65%). Primer combination Tms1Ret1-LORE1 generated the most amplified and polymorphic loci (Fig. 1). The average of polymorphic loci was 6.6 per primer. The size of

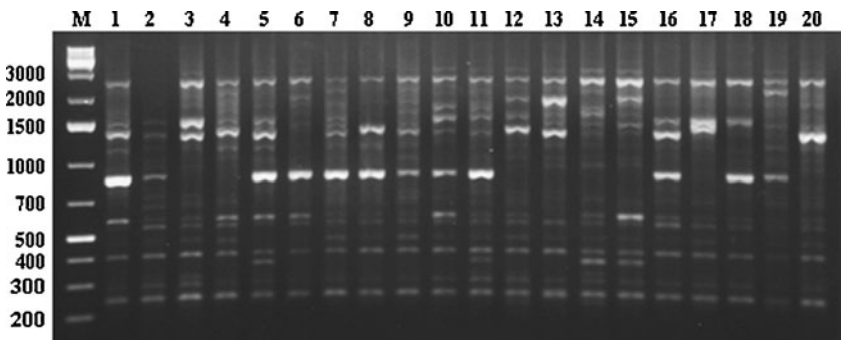

Fig. 1 Polymorphism detected by IRAP primer combination Tms1Ret1-LORE1. *Lanes from left to right*: *M*: 1 kb O'GeneRuler™ DNA ladder (Fermentas) in base pairs; *1 to 10*: individuals from population Azarbayejan-Ordubar; *11 to 20*: individuals from population Mahalie-Esfahan

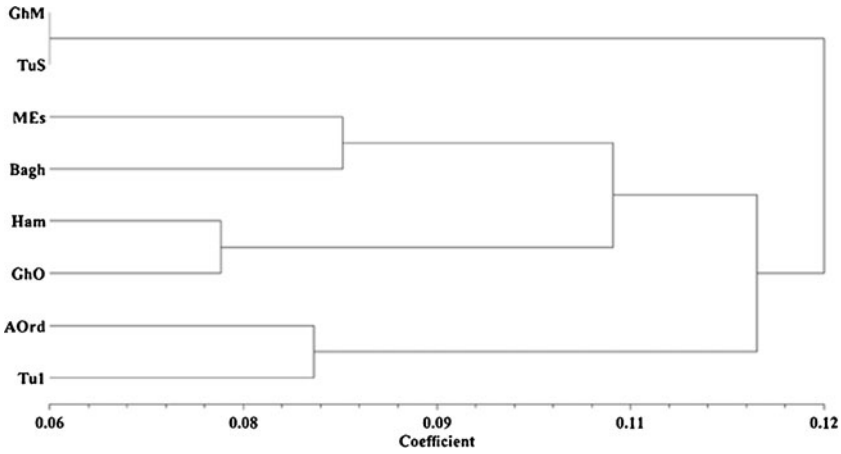

Fig. 2 UPGMA dendrogram of eight alfalfa populations using 101 IRAP loci based on Nei genetic distance coefficient. *GhO* Gharayonjeh-Orumieh, *GhM* Gharayonjeh-Malekkandi, *MEs* Mahalie-Esfahan, *Bagh* Baghdadi, *Ham* Hamedani, *AOrd* Azarbayejan-Ordubar, *Tu1* Turkey1, *TuS* Turkey-Sakuel

the amplified IRAP loci ranged from 250 to 3,000 bp (Table 3).

Genetic distance between populations based on IRAP markers ranged from 0.063 (GhM and TuS) to 0.151 (TuS and AOrd) with a mean value of 0.11. IRAP-based UPGMA dendrogram clustered populations into three main groups (Fig. 2). Group I included populations GhM and TuS. The second group constituted of two subgroups: MEs and Bagh in subgroup I and Ham and GhO in subgroup II. Populations AOrd and Tu1 located in the third group. Similar to the results obtained from cluster analysis, populations GhM and TuS separated from the rest of the populations on PCO biplot derived from IRAP markers (data not shown).

To assess and partition total genetic variation among and within populations, AMOVA was performed based on the eight populations using IRAP data. Significant differences ($P<0.05$) was detected within populations. The level of genetic variation was higher within populations (90%) compared to among populations (10%). The characteristics of amplified IRAP loci using 10 primers were reported in Table 4. The percentage of IRAP polymorphic loci in population varied from 74.24 (Ham) to 89.39 (Bagh), averaging 82.01. Frequency of the all IRAP amplified loci was more than 5%. A genotype-specific locus was produced in population GhM by IRAP primer combination Tms1Ret1-LORE1. Primer LORE2 amplified a private locus in a genotype of population AOrd. Mean of heterozygosity varied from 0.208 (Ham) to 0.261 (GhM), averaging 0.238 (Table 4).

REMAP Analysis

Of the tested 48 REMAP primer combinations, 14 generated 119 scorable loci (Table 3). Primer combination Tms1Ret1-A7 produced monomorphic banding pattern. No bands were amplified using REMAP primer combinations with non-native RTN Tps19. Out of 119 amplified loci, 62 were polymorphic (52%). Polymorphism detected by REMAP primer combination LORE1-848 in populations Ham and MEs are presented in Fig. 3. The average of polymorphic loci was 4.7 per primer. The size of the amplified loci ranged from 150 to 2,500 bp (Table 3).

REMAP-derived genetic distance between populations ranged from 0.048 (GhM and Bagh) to 0.124 (AOrd and Bagh), averaging 0.089. Three main clusters were identified

Table 4 Characteristics of amplified IRAP loci in studied alfalfa populations

Population	GhM	TuS	MEs	Bagh	Ham	GhO	Aord	Tu1
Number of loci	61	57	56	60	51	53	56	52
Percentage of polymorphic loci (%)	87.88	78.79	84.85	89.39	74.24	80.30	83.33	77.27
Number of loci with frequency of ≥5%	61	57	56	60	51	53	56	52
Number of private loci	1	0	0	0	0	0	1	0
Number of less common loci (≤25%)	0	1	0	1	0	0	0	0
Number of less common loci (≤50%)	5	5	3	3	0	3	1	1
Mean of heterozygosity	0.261	0.227	0.256	0.254	0.208	0.219	0.244	0.236
Standard error of mean heterozygosity	0.019	0.019	0.019	0.020	0.021	0.022	0.020	0.021

GhO Gharayonjeh-Orumieh, *GhM* Gharayonjeh-Malekkandi, *MEs* Mahalie-Esfahan, *Bagh* Baghdadi, *Ham* Hamedani, *AOrd* Azarbayejan-Ordubar, *Tu1* Turkey1, *TuS* Turkey-Sakuel

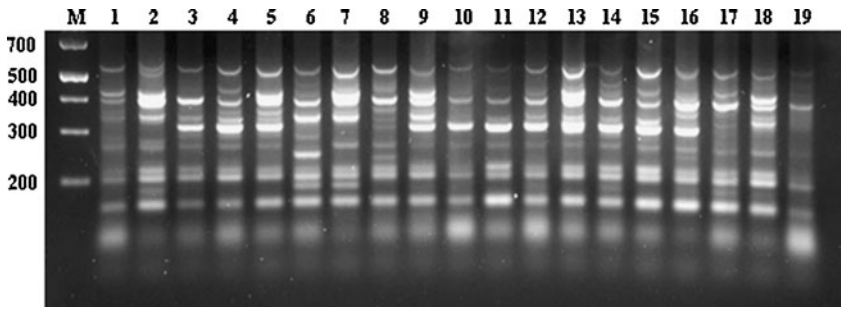

Fig. 3 Polymorphism detected by REMAP primer combination LORE1-848. *Lanes from left to right*: *M*: 1 kb O'GeneRuler™ DNA ladder (Fermentas) in base pairs; *1 to 10*: individuals from population Mahalie-Esfahan; *11 to 19*: individuals from population Hamedani

based on REMAP data (Figs. 4 and 5). Group I consisted of three subgroups: GhM and Bagh in subgroup I, GhO and Tu1 in subgroup II, and MEs constituted subgroup III. TuS clustered in group II. The third group constituted of populations Ham and AOrd. Associations among eight populations using REMAP markers were also confirmed through PCO (data not shown). REMAP-based AMOVA was carried out using the eight populations as the basis for the analysis. Similar to the results obtained by IRAP method, the level of genetic variation was higher within populations (94%) compared to among populations (6%), indicating no clear differentiation based on the studied populations. The characteristics of amplified loci using 14 REMAP primers are presented in Table 5.

The percentage of polymorphic loci in population based on REMAP data varied from 79.03 (MEs) to 88.71 (GhO), with a mean value of 83.67. Four loci with a frequency less than 25% were detected (Table 4), suggesting that most of the amplified loci are prevalent in the studied populations. Three genotype-specific REMAP loci were detected in populations Ham, GhO, and AOrd. Mean of heterozygosity varied from 0.249 (TuS) to 0.290 (GhO), averaging 0.267 (Table 4).

Combined Data Analysis

Combined data of IRAP and REMAP markers were used to build a dendrogram to evaluate the power of both techniques when accumulated. Genetic distance between populations based on IRAP + REMAP analysis ranged from 0.08 (GhM and TuS) to 0.134 (AOrd and TuS) with a mean value of 0.102. Populations were grouped into three main clusters (Fig. 5). Group I included GhM and TuS in subgroup I and MEs, and Bagh and Tu1 in subgroup II. Populations Ham and GhO clustered in group II and AOrd located in group III. PCO for IRAP + REMAP data was carried out to confirm associations among eight populations. The first (PCO1) and second (PCO2) principle coordinates accounted for 33% and 21% of the total variation, respectively. Relationships among populations on the biplot were similar to the results obtained from cluster analysis (Fig. 6).

AMOVA was performed to divide total genetic variation among and within populations. Similar to what was recorded by IRAP and REMAP markers, the level of genetic variation was higher within populations (92%) compared to among populations (8%). The percentage of polymorphic loci in population varied from 79.69 (Ham) to 85.16 (GhM), averaging 82.81. Mean of heterozygosity varied from 0.238 (TuS) to 0.257 (MEs and Tu1), averaging 0.252 (Table 6). Because of high genetic diversity within populations, IRAP + REMAP-based cluster analysis was performed using complete linkage algorithm based on Dice similarity coefficient to identify heterotic groups among all 80 genotypes (Fig. 7). Five heterotic groups were identified and summarized in Table 7.

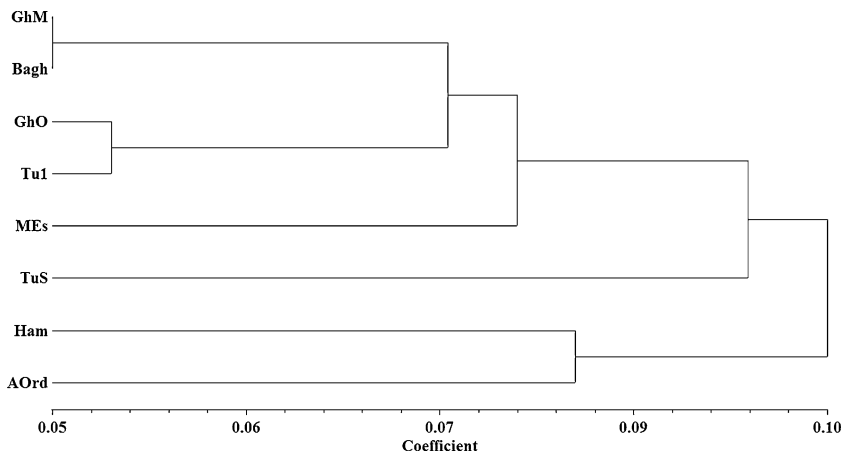

Fig. 4 UPGMA dendrogram of eight alfalfa populations using 119 REMAP loci based on Nei genetic distance coefficient

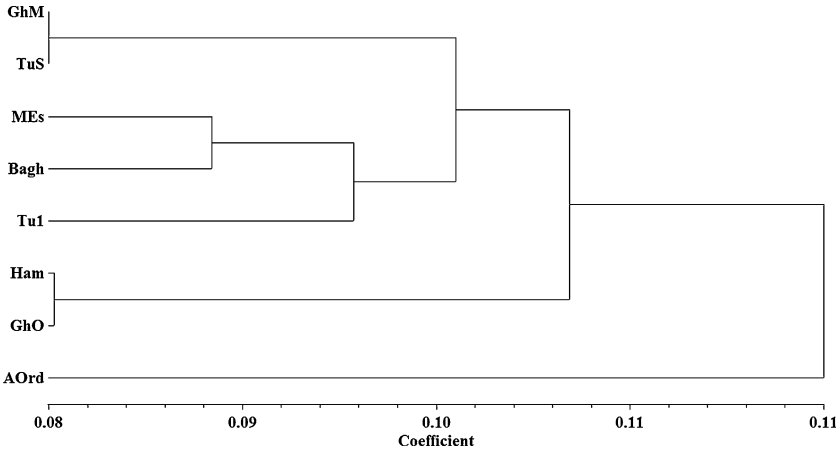

Fig. 5 UPGMA dendrogram of eight alfalfa populations using IRAP + REMAP loci based on Nei genetic distance coefficient

Discussions

RTN Activity and Insertional Polymorphism in Alfalfa Populations

This is the first report of IRAP- and REMAP-based assessment of native and non-native RTN activity and genetic diversity in alfalfa. Out of tested primers, 10 IRAP and 14 REMAP were shown to amplify discernible banding pattern and applied to study the used RTN activity and genetic diversity among 80 alfalfa genotypes. RTNs may integrate in principle in either orientation into the genome, and hence, any two members of one or different RTN families may be found head to head, tail to tail, or head to tail (Kalendar et al. 1999; Abdollahi Mandoulakani et al. 2009). Moreover, different RTN families may integrate into each other. Therefore, in order to increase the probability of finding bands, we also combined primers from LTR ends of different RTN families. Several investigations have demonstrated that primers designed based on LTR sequences of RTN families can be readily used across species lines, among closely related genera and even sometimes between plant families (Lou and Chen 2007; Kalendar et al. 2011). In this study, single IRAP primers Tms1Ret1 (native) and LORE1 and LORE2 (non-native) produced polymorphic banding patterns in studied alfalfa populations, showing the presence and activity of the mentioned retroelements in *M. sativa* genome. Tms1Ret1 elements represent about 16×10^3 copies in the *M. sativa* genome and are transpositionally active among and within different alfalfa species (also shown in the current study). High polymorphism of Tms1Ret1-based S-SAP (compared to SAMPL and AFLP markers) and no significant cluster of Tms1Ret1-based markers in *Medicago* genomes have been previously demonstrated (Porceddu et al. 2002); hence, these markers can provide good genome coverage in diversity studies in different *Medicago* species. Madsen et al. (2005) reported that LORE1, a low-copy-number TY3-gypsy RTN family in the model legume *L. japonicus*, is transpositionally active. LORE2A is an aged element, estimated as 600,000 years old, yet active in *L. japonicus* (Fukai et al. 2008). Tps12 may be transpositionally inactive in *M. sativa* genome because of produced monomorphic banding pattern, but it generated much more polymorphism in a combination with Tms1Ret1, indicating the insertion of these two RTNs near or into each other in alfalfa genome. Primer designed based on Tps19 amplified no bands, demonstrating its absence or much and fast divergence in *M. sativa* genome. Pearce et al. (2000) isolated a large heterogenous population of previously uncharacterized Ty1-copia RTNs from pea (Tps) and showed that each element group in pea has relatives in

Table 5 Characteristics of amplified REMAP loci in studied alfalfa populations

Population	GhM	TuS	MEs	Bagh	Ham	GhO	Aord	Tu1
Number of loci	51	54	50	50	54	56	52	54
Percentage of polymorphic loci (%)	82.26	85.48	79.03	80.65	85.48	88.71	82.26	85.48
Number of loci with Frequency of ≥5%	51	54	50	50	54	56	52	54
Number of private loci	0	0	0	1	0	1	1	0
Number of less common loci (≤25%)	0	2	1	1	0	0	0	0
Number of less common loci (≤50%)	0	3	1	1	1	1	1	2
Mean of heterozygosity	0.259	0.249	0.259	0.256	0.275	0.290	0.269	0.280
Standard error of mean heterozygosity	0.022	0.020	0.023	0.022	0.022	0.021	0.022	0.022

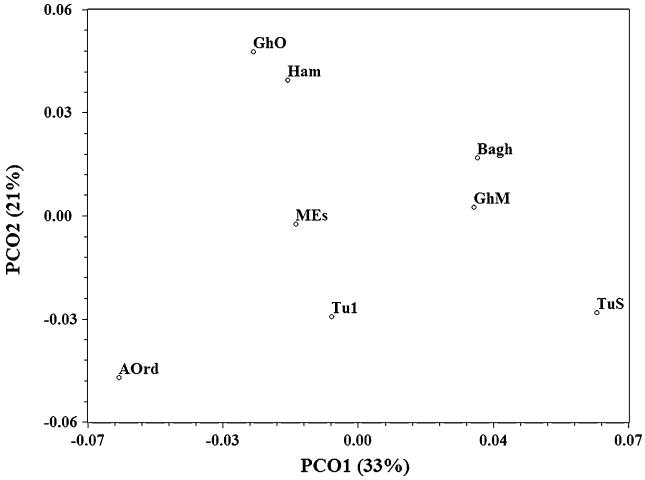

Fig. 6 Two-dimensional plot of the genetic relationship among eight alfalfa populations as revealed by principle coordinate analysis using IRAP + REMAP data

the more distantly related *Vicia* species indicating that heterogenous populations of these elements have been present throughout the evolution of the *Pisum* and *Vicia* genera from their common ancestor. They showed that Tps12a and Tps19 have a high level of insertional polymorphism in pea and have been active during *Pisum* species evolution. Our study shows that Tps elements also have relatives in *M. sativa* and probably come from the common ancestors before the divergence of *Medicago* and *Pisum* in evolutionary processes.

The generation of discernible and polymorphic banding pattern with used IRAP primer combinations (except for primer combinations that Tps19 was involved) showed that most of the studied RTN families integrate near or into each other in alfalfa genome. The high activity and polymorphism of different RTN families isolated from barley, oat, brachypodium, rice, and maize have been demonstrated in wheat (Abdollahi Mandoulakani et al. 2008, 2009). Kalendar et al. (2011) showed the activity and insertional polymorphism of Bare1-based IRAP markers in *Hordeum*, *Triticum*, and *Aegilops* species. Tam et al. (2005) in a study of comparative analyses of genetic diversity within tomato stated that RTN sequences isolated from one species can be used in related *Solanaceae* genera. REMAP amplification of the used RTN families indicated the insertion of the Tms1Ret1, as a native RTN, near the different SSRs in alfalfa genome. Non-native studied RTNs amplified bands in combination with a few SSR motifs, probably suggesting their low copy number, divergence, and preferential insertion within SSRs in *M. sativa* genome. The insertions of the RTN families in the vicinity of microsatellites have been previously reported in barley and wheat (Vicient et al. 1999, 2005; Carvalho et al. 2010). Polymorphisms detected by markers based on non-native RTNs from *P. sativum* were low compared to those developed based on non-native RTNs from *L. japonicus*. This could be expected because of the close polygenetic relationship between *M. sativa* and *L. japonicus*.

Genetic Relationship and Characterization of Alfalfa Populations

The estimated correlations between the three generated cophenetic matrices from IRAP, REMAP, and IRAP + REMAP dendrogram evidenced a relatively high and significant correlation of IRAP with IRAP + REMAP (r=0.7). However, the matrices estimated by the techniques individually revealed a low and non-significant correlation (r=0.15). These results suggest that similarly to what is found in barley (Kalendar et al. 1999) and rice (Branco et al. 2007), REMAP primers amplified DNA regions that could not be covered by IRAP. Therefore, we used combined analysis of IRAP + REMAP markers to reveal the association between populations and identification of heterotic groups among 80 alfalfa genotypes.

Using 220 amplified IRAP + REMAP loci, high levels of genetic variation were detected among individual plants within populations, compared to variation among populations, suggesting high amount of gene flow between studied populations (Falahati-Anbaran et al. 2007, Oliviera et al. 2010). The high level of intra- and inter-population variation detected in this work could be related to the outcrossing and

Table 6 Characteristics of amplified IRAP + REMAP loci in studied alfalfa populations

Population	GhM	TuS	MEs	Bagh	Ham	GhO	Aord	Tu1
Number of loci	112	111	106	110	105	109	108	106
Percentage of polymorphic loci (%)	85.16	82.03	82.03	85.16	79.69	84.38	82.81	81.25
Number of loci with frequency of ≥5%	112	111	106	110	105	109	108	106
Number of private loci	1	0	0	1	0	1	2	0
Number of less common loci (≤25%)	0	3	1	2	0	0	0	0
Number of less common loci (≤50%)	5	8	4	4	1	4	2	3
Mean of heterozygosity	0.260	0.238	0.257	0.255	0.240	0.253	0.256	0.257
Standard error of mean heterozygosity	0.014	0.014	0.015	0.015	0.015	0.015	0.015	0.015

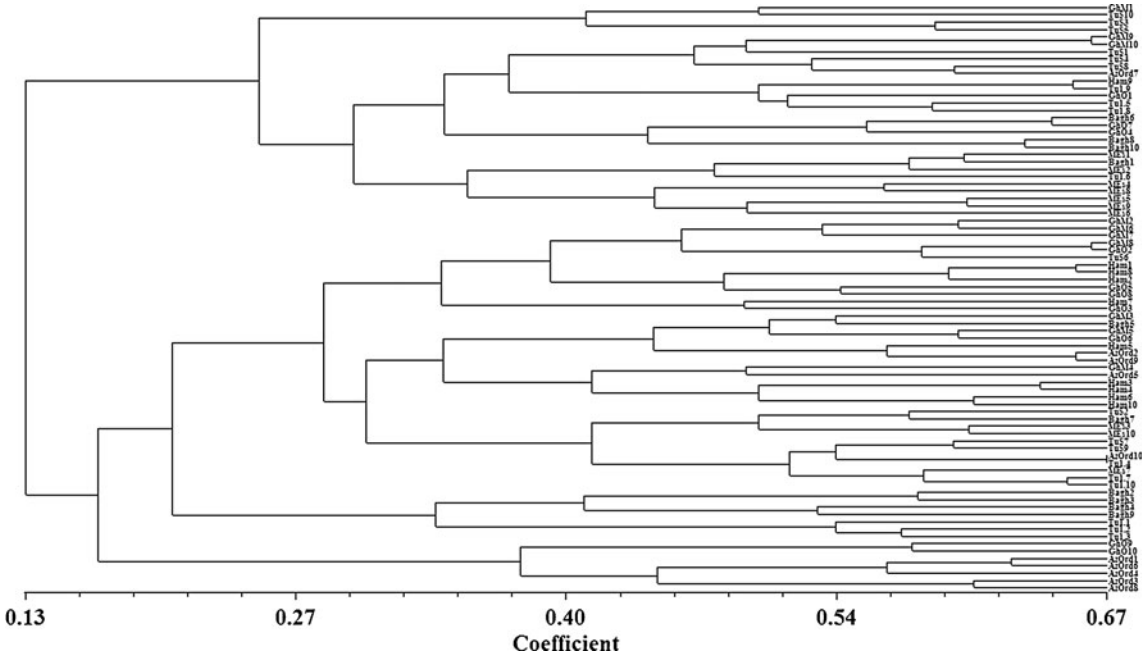

Fig. 7 Complete linkage dendrogram of 80 alfalfa genotypes using IRAP + REMAP loci based on Dice similarity coefficient

the tetraploid nature of alfalfa (Mengoni et al. 2000; Falahati-Anbaran et al. 2007), as well as high activity and insertional polymorphism of the used RTNs in alfalfa genome. Although populations could be relatively differentiated in the current study, overall genetic differentiation among populations was low (PhiPT=0.082, P=0.0010). Low level of genetic differentiation between Iranian alfalfa populations has been previously reported (Falahati-Anbaran et al. 2007). Population TuS (from Turkey) was not differentiated from GhM (from Iran), indicating the occurrence of high amount of gene flow between these two populations (PhiPT=0.041; Nm, number of migrants=5.902). This could be expected as

Table 7 Five heterotic groups of alfalfa genotypes based on IRAP + REMAP data

Heterotic groups						
1	2		3		4	5
GhM1	GhM9	Bagh8	GhM2	AzOrd9	Bagh2	GhO9
TuS10	GhM10	Bagh10	GhM6	GhM4	Bagh3	GhO10
TuS3	TuS1	MEs1	GhM7	AzOrd5	Bagh4	AzOrd1
TuS5	TuS4	Bagh1	GhM8	Ham3	Bagh9	AzOrd6
	TuS8	MEs2	GhO2	Ham4	Tu1.1	AzOrd4
	AzOrd7	Tu1.6	TuS6	Ham6	Tu1.2	AzOrd3
	Ham9	MEs4	Ham1	Ham10	Tu1.3	AzOrd8
	Tu1.9	MEs8	Ham8	TuS2		
	GhO1	MEs5	Ham2	Bagh7		
	Tu1.5	MEs9	GhO5	MEs3		
	Tu1.8	MEs6	GhO8	MEs10		
	Bagh6		Ham7	TuS7		
	GhO7		GhO3	TuS9		
	GhO4		GhM3	AzOrd10		
			Bagh5	Tu1.4		
			GhM5	MEs7		
			GhO6	Tu1.7		
			Ham5	Tu1.10		
			AzOrd9			

ecotype TuS is cultivated in regions of west Azarbayejan in Iran, where population GhM is planted as well. All of these might facilitate gene flow between these two populations. Environment has a major effect on the amplification of retroelements (Carvalho et al. 2010), but no clear correlation was found in our study between the climate conditions and RTN insertional polymorphism and activity.

Populations TuS and AOrd had the most genetic distance. Populations were clustered in three main groups. Populations from different groups could be introduced as parents with enough genetic distance to produce hybrids or semi-hybrids in alfalfa. Because of high genetic diversity within populations and among genotypes, heterotic groups were identified using IRAP + REMAP analysis (Table 7). The genotypes from different heterotic groups could be also introduced as suitable parents in alfalfa breeding programs. Crossing between populations in different groups and genotypes in heterotic groups 4 and 5 is undergoing to test if IRAP and REMAP markers could be properly used to predict heterosis in alfalfa.

The use of RTN-based markers can be a valuable tool for alfalfa breeders, as it is for barley (Kalendar et al. 1999), wheat (Abdollahi Mandoulakani et al. 2008, 2009; Carvalho et al. 2010), and citrus (Biswas et al. 2010). The IRAP and REMAP techniques can be used combined for a more complete genome survey. The ubiquitous presence of LTR RTNs in plant genomes suggests that the use of these techniques in combination would allow breeders to obtain markers close to virtually any important agronomical trait and that the hypervariable nature of these elements should make them excellent sources of polymorphic markers. However, it is advisable to use elements with different insertional pattern or at least elements that do not insert preferentially to cover the complete genome. In conclusion, IRAP and REMAP markers derived from native and non-native RTNS from *L. japonicus* and *P. sativum* were polymorphic enough to allow the detection of populations and genotypes with large genetic distance in *M. sativa*.

Acknowledgments The authors are grateful to Institute of Biotechnology, Urmia University, Iran, for the financial support of the work.

References

Abdollahi Mandoulakani B, Bihamta MR, Zali AA, Yazdi-Samadi B, Naghavi MR, Schulman AH (2008) Fine mapping of stripe rust resistance gene *Yr15* in durum wheat. Seed and Plant J (In Persian) 24:371–387

Abdollahi Mandoulakani B, Bihamta MR, Schulman AH, Zali AB, Naghavi M (2009) Evaluation of retrotranposons as molecular markers in wheat. Modern Genet J (In Persian) 4:17–25

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K, Albright LM, Coen DM, Varki A (1995) Current protocols in molecular biology. Wiley, New York

Betran FJ, Ribaut JM, Beck D, Gonzalez de Leon D (2003) Genetic diversity, specific combining ability and heterosis in tropical maize under stress and non stress environments. Crop Sci 43:797–806

Biswas MK, Baig MNR, Cheng YJ, Deng XX (2010) Retrotransposon based genetic similarity within the genus citrus and its relatives. Genet Resour Crop Evol 7:963–972

Bourguiba H, Krichen L, Audergon JM, Khadari B, Tarifi-Farah N (2010) Impact of mapped SSR markers on the genetic diversity of apricot (*Prunus armeniaca* L.) in Tunisia. Plant Mol Biol Rep 28:578–587

Branco CJS, Vieira EA, Malone G, Kopp MM, Malone E, Bernardes A, Mistura CC, Carvalho FIF, Oliveira CA (2007) IRAP and REMAP assessments of genetic similarity in rice. J Appl Genet 2:107–113

Brummer EC (1999) Capturing heterosis in forage crop cultivar development. Crop Sci 39:943–954

Carvalho A, Guedes-Pinto H, Martins-Lopes P, Lima-Brito J (2010) Genetic variability of old Portuguese bread wheat cultivars assayed by IRAP and REMAP markers. Ann Appl Biol 3:337–345

Diers BW, Vetty PBE, Osborn TC (1996) Relationship between heterosis and genetic distance based on RFLP markers in oilseed rape (*Brassica napus* L.). Crop Sci 36:79–83

Diwan N, Bhagwat AA, Bauchan GR, Cregan PB (1997) Simple sequence repeat DNA markers in alfalfa and perennial and annual *Medicago* species. Genome 40:887–895

Falahati-Anbaran M, Habashi AA, Esfahani M, Mohammadi SA, Gharayazi B (2007) Population genetic structure based on SSR markers in alfalfa (*Medicago sativa* L.) from various regions contiguous to the centers of origin of the species. J Genet 1:59–63

Fiuk A, Bednarek PT, Rybczynski JJ (2010) Flow cytometry, HPLC-RP, and metAFLP analysis to assess genetic variability in somatic embryo-derived plantlets of *Gentiana pannonica* scop. Plant Mol Biol Rep 28:413–420

Flavell AJ, Knox MR, Pearce SR, Ellis THN (1998) Retrotransposon-based insertion polymorphisms (RBIP) for high throughput marker analysis. Plant J 16:643–650

Fukai E, Dobrowolska AD, Madsen LH, Madsen EB, Umehara Y, Kouchi H, Hirochika H, Stougaard J (2008) Transposition of a 600 thousand-year-old LTR retrotransposon in the model legume *Lotus japonicus*. Plant Mol Biol 68:653–663

Kalendar R, Grob T, Regina M, Souniemi A, Schulman AH (1999) IRAP and REMAP: two new retrotransposon-based DNA fingerprinting techniques. Theor Appl Genet 98:704–711

Kalendar R, Flavell AJ, Ellis THN, Sjakste T, Moisy C, Schulman AH (2011) Analysis of plant diversity with retrotransposon-based molecular markers. Heredity 106:520–530

Kidwell KK, Austin DF, Osborn TC (1994a) RFLP evaluation of nine *Medicago* accessions representing the original germplasm sources for North American alfalfa cultivars. Crop Sci 34:230–236

Kidwell KK, Woodfield DR, Bingham ET, Osborn TC (1994b) Molecular marker diversity and yield of isogenic 2x and 4x single crosses of alfalfa. Crop Sci 34:784–788

Kidwell KK, Hartweck LM, Yandell BS, Crump PM, Brummer JE, Moutray J, Osborn TC (1999) Forage yields of alfalfa populations derived from parents selected on the basis of molecular marker diversity. Crop Sci 39:223–227

Lou Q, Chen J (2007) Ty1-copia retrotransposon-based SSAP marker development and its potential in the genetic study of cucurbits. Genome 50:802–810

Lynch M, Milligan BG (1994) Analysis of population genetic structure with RAPD markers. Mol Ecology 3:91–99

Madsen LH, Fukai E, Radutoiu S, Karl Yost C, Sandal N, Schauser L, Stougaard J (2005) LORE1, an active low-copy-number TY3-gypsy retrotransposon family in the model legume *Lotus japonicus*. Plant J 44:372–381

Mengoni A, Gori A, Bazzicalupo M (2000) Use of RAPD and microsatellite (SSR) variation to assess genetic relationships among populations of tetraploid alfalfa, *Medicago sativa*. Plant Breed 119:113–117

Michaud R, Lehman WF, Runbaugh MD (1988) World distribution and historical development. In: Hanson AA, Barnes DK, Hill RR (eds) Alfalfa and alfalfa improvement. Agronomy Monograph 29, Madison, pp 25–92

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. PNAS 76:5269–5273

Oliviera EJ, Amorim VBO, Matos ELS, Costa JL, Castellen MS, Padua JG, Dantas JLL (2010) Polymorphism of microsatellite markers in papaya (*Carica papaya* L.). Plant Mol Biol Rep 28:519–530

Peakall R, Smouse PE (2006) GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295

Pearce SR, Knox M, Ellis THN, Flavell AJ, Kumar A (2000) Pea Ty1-copia group retrotransposons: transpositional activity and use as markers to study genetic diversity in pisum. Mol Genet Genomics 263:898–907

Porceddu A, Albertini E, Barcaccia G, Marconi G, Bertoli FB, Veronesi F (2002) Development of S-SAP markers based on an LTR-like sequence from *Medicago sativa* L. Mol Genet Genomics 267:107–114

Queen RA, Gribbon BM, James C, Jack P, Falvell AJ (2004) Retrotransposon-based molecular markers for linkage and genetic diversity analysis in wheat. Mol Genet Genomics 271:91–97

Riaz A, Li G, Quresh Z, Swati MS, Quiros CF (2001) Genetic diversity of oilseed *Brassica napus* inbred lines based on sequence-related amplified polymorphism and its relation to hybrid performance. Plant Breed 120:411–415

Riday H, Brummer CE, Campbell TA, Luth D, Cazcarro PM (2003) Comparisons of genetic and morphological distance with heterosis between *Medicago sativa* subsp. *sativa* and subsp. *Falcate*. Euphytica 131:37–45

Robins JG, Luth D, Campbell TA, Bauchan GR, He C, Viands DR, Hansen JL, Brummer EC (2007) Genetic mapping of biomass production in tetraploid Alfalfa. Crop Sci 47:1–10

Rohlf FJ (2000) NTSYS-pc: numerical taxonomy and multivariate analysis system, version 2.1. Exeter Software, New York

Tam SM, Mhiri C, Vogelaar A, Kerkveld M, Pearce SR, Le Grandbastien MA (2005) Comparative analyses of genetic diversities within tomato and pepper collections detected by retrotransposon-based SSAP, AFLP and SSR. Theor Appl Genet 110:819–831

Vicient CM, Kalendar R, Anamthawat-Jonsson K, Suoniemi A, Schulman AH (1999) Structure, functionality, and evolution of the BARE-1 retrotransposon of barley. Genetica 107:53–63

Vicient CM, Kalendar R, Schulman AH (2005) Variability, recombination and mosaic evolution of the barley BARE-1 retrotransposon. J Mol Evol 61:275–291

Waugh R, McLean K, Flavell AJ, Pearce SR, Kumar A, Thomas BBT, Powell W (1997) Genetic distribution of Bare1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). Mol General Genomics 253:687–694

Wei Z, Zhang K, Yang C, Liu G, Liu G, Lian L, Zhang H (2010) Genetic linkage maps of *Betula platyphylla* suk based on ISSR and AFLP markers. Plant Mol Biol Rep 28:169–175

Zaccardelli M, Gnocchi S, Carelli M, Scotti C (2003) Variation among and within Italian alfalfa ecotypes by means of bio-agronomic characters and amplified fragment length polymorphism analyses. Plant Breed 122:61–65