Plant Omics Journal POJ

POJ 5(4):359-364 (2012) ISSN:1836-3644

R-RAP: a retrotransposon-based DNA fingerprinting technique in plants

Ali Aalami[1], Samira Safiyar[1], Babak Abdollahi Mandoulakani[2, 3]*

[1]Department of Agronomy and Plant Breeding, Faculty of Agriculture, University of Guilan, Rasht, Iran
[2]Department of Agronomy and Plant Breeding, Faculty of Agriculture, Urmia University, Urmia, Iran
[3]Department of Agricultural Biotechnology, Institute of Biotechnology, Urmia University, Urmia, Iran

***Corresponding author: b.abdollahi@urmia.ac.ir**

Abstract

The present research is aimed at testing a new retrotransposon-based marker, RAPD-retrotransposon amplified polymorphism (R-RAP), which used the combination of RAPD and LTR retrotransposon primers. Banding patterns which are obtained by R-RAP primer combinations were different from amplicons by IRAP and RAPD, demonstrating detection of different genomic regions that are not covered by the other molecular marker systems. The used LTR rertrotransposons had been previously isolated from barley and wheat, hence, twelve accessions of three species of *Triticum aestivum*, *Hordeum vulgare,* and *Aegilops tauschii* were used to evaluate the efficiency of this method. Ten polymorphic R-RAP primer combinations generated 141 loci, of which 114 were polymorphic. The studied species were divided with 66 % similarity into 3 clusters according to their species and even some sub-clusters within species. High polymorphism was observed in both between and within species. Polymorphic information content (PIC) ranged from 0.28 to 0.40. Reproducibility of each primer combination, tested 4 times in different conditions, was approximately 100%. Agarose gel was used to separate PCR products and showed good resolving ability. We showed that R-RAP can be used as an efficient marker as with other retrotransposon-based markers and can be applied in a similar way for DNA fingerprinting, genetic diversity, genome mapping, and gene tagging in plants. This method also represents a departure for IRAP and RAPD limitations.

Keywords: IRAP; LTR retrotransposon; PIC; R-RAP; RAPD.
Abbreviations: LTR-long terminal repeat; R-RAP- RAPD retrotransposon amplified polymorphism; RAPD-random amplified polymorphic DNA; AFLP- amplified fragment length polymorphism; SSR- simple sequence repeat; IRAP-inter-retrotransposon amplified polymorphism; REMAP-retrotransposon microsatellite amplified polymorphism; S-SAP- sequence specific amplification polymorphism; RBIP-retrotransposon-based insertional polymorphism; PIC- polymorphic information content.

Introduction

The emergence of marker systems has closely followed development in biochemistry and molecular biology for the last 40 years (Kalendar et al., 2010). Some of these techniques are basic and their limitations and advantages have been studied since that time. RAPD is one of the first developed methods that have been studied widely in all genetic programs. Using short primer (almost 10 nucleotides), no extra information was needed to know about the genome sequence, high speed, low cost and low technical requirement led RAPD to be a desirable marker for many years (Williams et al., 1990; Sawalha et al., 2008). On the other hand, presenting many potential priming sites for these sequences, low annealing temperature that causes low reproducibility, and inventing new methods such as AFLP (Amplified fragment length polymorphism) and SSR (Simple sequence repeat) which solved RAPD problems, thus eliminating this system from molecular markers today (Kalendar et al., 2011). In 1956 Barbara McClintock found some mutagenesis factors, called transposable elements, that can be divided into two main classes: DNA transposons that move through their intermediate DNA using a cut-and-paste mechanism and retrotransposons that increase their copy number through a cycle of transcription and integration back to the genome on condition that the older copy still persists and causes increase of genome size (Bennetzen, 2000; Kalendar et al., 2004; Rodriguez et al., 2006). In many crop plants, between 40 to 70 % of the total DNA comprise retrotransposons (Bennetzen, 2000). Due to the features, such as: integration activity, persistence, dispersion, conserved structure, sequence motifs and high copy number; retroelements can widely be used as molecular markers today. The advanced markers also utilize these retrotransposable elements (Agarwal et al., 2008). In general, retroelement-based molecular markers are based on PCR and one primer is designed to match a segment of LTR that is conserved within a given family of element, but is different in other families. The second primer is designed to match some other sequence of the genome (Kalendar and Schulman, 2006). According to the identity of the second primer, some retroelement-based techniques have been developed: IRAP (Inter-retrotransposon amplified polymorphism) is based on the fact that retrotransposons tend to cluster together in the genome. This method uses two LTR primers that can be from the same or from a different family. IRAP is experimentally simple (Kalendar et al., 2011), but it has some disadvantages such as: producing a huge product, making the low resolution or target sites too far a part to produce product (Mansour, 2008). REMAP (Retrotransposon-microsatellite amplified polymorphism), to some extent, is similar to IRAP, one of the primer matches to a microsatelite motif (Kalendar et al., 1999).RBIP (Retrotransposon-based insertional polymorphism) is based on flanking regions of a LTR to detect

Table1. List of primer sequences used in this study.

Primers name	Sequence(5' -3')
UBC3	CCTGGGCTTA
UBC77	GAGCACCAGG
OPA10	GTGATCGCAG
OPA02	TGCCGAGCTG
OPG10	AGGGCCGTCT
Wltr2105	ACTCCATAGATGGATCTTGGTGA
Sukkula	GATAGGGTCGCATCTTGGGCGTGAC
5'LTR	ATCATTGCCTCTAGGGCATAATTC

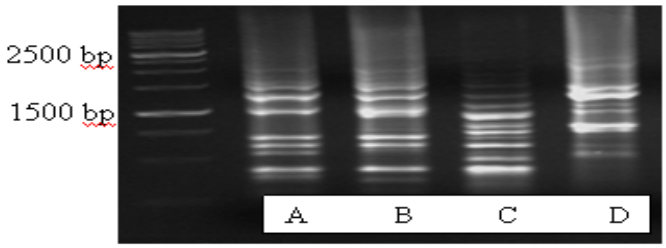

Fig 1. Banding pattern generated by primers Sukkula (A and B), OPA02 (C) and their combination (D) on *Triticum aestivum* accessions.

polymorphism for the integrating element at a particular locus (Kalendar et al., 1999). RBIP is more expensive and technically complicated than the other methods for detecting insertion, because it needs to know the complete sequence of the 3' and 5' of the flanking region of an insertion site of a retrotransposon and it is a co-dominant marker similar to SSR (Kalendar et al., 2011). We hypothesised a new method based on LTR retrotransposons that uses RAPD primers to match to a random sequence of the genome and LTR to another site, but because of the random feature of RAPD primer, it would be presented between two LTR and, therefore, make it possible to produce some new bands from a combination of RAPD and LTR primers. Our proposal is in representing a departure for IRAP and RAPD restrictions, because it overcomes not only low annealing temperature and, therefore, low reproducibility of RAPD, but also the problem with IRAP, that for far LTRs it would either produce no band or band with low resolution. In addition, this new combination may produce some new amplicons and, thus, gives a new kind of information about DNA sequences and plant evolution. Compared to other retrotransposon-based methods, R-RAP (RAPD-retrotransposon amplified polymorphism) neither needs restriction enzymes and silver staining like S-SAP, nor prior information about the SSR motifs like REMAP and solves the IRAP problems on condition that it is still as easy to perform as IRAP, and at the same time it can be as efficient as all of these methods. Simplicity, low technical and cost requirements and no primary information needed, makes possible to use it as a useful genetic tool in the near future in plant genetics.

Results and discussion

Detection of new genomic regions by R-RAP method

IRAP method can be carried out with a single primer matching either 5´or 3´ end of the LTR but oriented away from the LTR itself, or with two LTR primers from the same or different family of retrotransposons (Kalendar and Schulman, 2006). Therefore, if a single primer was used in an IRAP PCR reaction, because of large distance between some LTR retrotransposons, it would either produce no product or products with low resolution (smear form). As a result, this sequence is combined with some other sequence of genome such as SSR motifs (REMAP) (Kalendar et al. 1999). Three LTR primers (WLTR2105, 5LTR and Sukkula) (Table 1) which had been used previously in IRAP and REMAP in aegilops and wheat (Saeidi et al., 2008), were combined with 5 RAPD primers (15 R-RAP combinations) and used in our study. 10 of 15 R-RAP primer combinations, produced clear banding patterns. Application of primer WLTR2105, solely in a PCR reaction on wheat and barley genotypes, led to production of a smear in gel electrophoresis. It seems that frequency of this sequence is not very high in the genome, consequently, large amplicons were produced which appear like a smear in the gel instead, a primer combination of RAPD and WLTR2105 in the PCR reaction produced very clear and distinguishable amplicons. Combinations of five different RAPD primers with WLTR2105 (Table 2) were tested and all of them amplified the product. Band sizes ranged from 250 to 3500 bp. All RAPD primers were also examined in PCR reaction solely and the amplified banding pattern compared with the combination of RAPD and retrotransposon primers clarified that the banding patterns are completely different in these two kinds of PCR products.

In some R-RAP reactions, banding pattern produced by a given combination of RAPD and retrotransposon primers led to larger amplicons to almost 3000 bp. It is likely because of the large distance between LTR regions and RAPD sequence which is lying in this region. Obtaining the pattern represents the result of competition between the targets and product in the reaction and as a result the product obtained with two primers does not represent the simple sum of the product obtained with the primer individually. Hence, it is logical that this combination demonstrates new loci in the genome. Sukkula, one of the abundant LTR retrotransposon in genome-produced clear bands acts as a single primer in IRAP reactions. As expected, banding patterns amplified by primers Sukkula, RAPD and RAPD+Sukkula were different. Some amplicons produced by single retrotransposon or RAPD primers vanished in combination and, at the same time, some new bands were produced (Fig. 1).

Efficiency of R-RAP for genetic diversity, gene tagging and mapping in plants

An efficiency of a molecular marker system is detection of polymorphism between and within the species. At this regard, 4 genotypes of each species *H. vulgare*, *A. tauschii* and *T. aestivum* were used in PCR reactions with 10 R-RAP combinations that we attained. Consequently, clear and distinguishable amplicons were produced (Fig. 2) and then the bands were scored as 0 for absence and 1 for presence of a particular locus. Cluster analysis was performed using UPGMA algorithm based on Dice similarity coefficient and 3 species were divided into 3 groups (Fig. 3) with an average similarity of 66% which indicates that these primers can be used efficiently for diversity and evolutionary programs in the mentioned species. Similar to the result obtained by cluster analysis, principle coordinate analysis divided these species in 3 distinguished groups as well (Fig. 4).

The number of R-RAP locus amplified by each primer combination ranged from 10 for Sukkula+UBC3 to 17 for WLTR2105+OPG10. A summary of diversity indices are given in Table 3. Primer combination WLTR2105+UBC3 was the highest in diversity indices. Polymorphic information

Table 2. R-RAP primer combinations used in this study and their comparison with corresponded single IRAP and RAPD primers.

IRAP	Tm	Resolution of bands	Reproducibility
WLTR2105	63.4	very low	-
Sukkula	75	High	High
5LTR	64.9	very low	-
RAPD			
UBC3	35	High	Low
UBC77	33	High	Low
OPA10	32	High	low
OPA02	42	High	low
OPG10	42	high	low
R-RAP			
WLTR2105+OPA02	62	high	high
WLTR2105+OPA10	58	high	high
WLTR2105+OPG10	62	high	high
WLTR2105+UBC3	63	high	high
WLTR2105+UBC77	63	high	high
Sukkula+OPA10	62	high	high
Sukkula+UBC3	65	high	high
Sukkula+UBC77	64	high	high
5LTR +OPA10	62	high	high
5LTR+UBC3	63	high	high

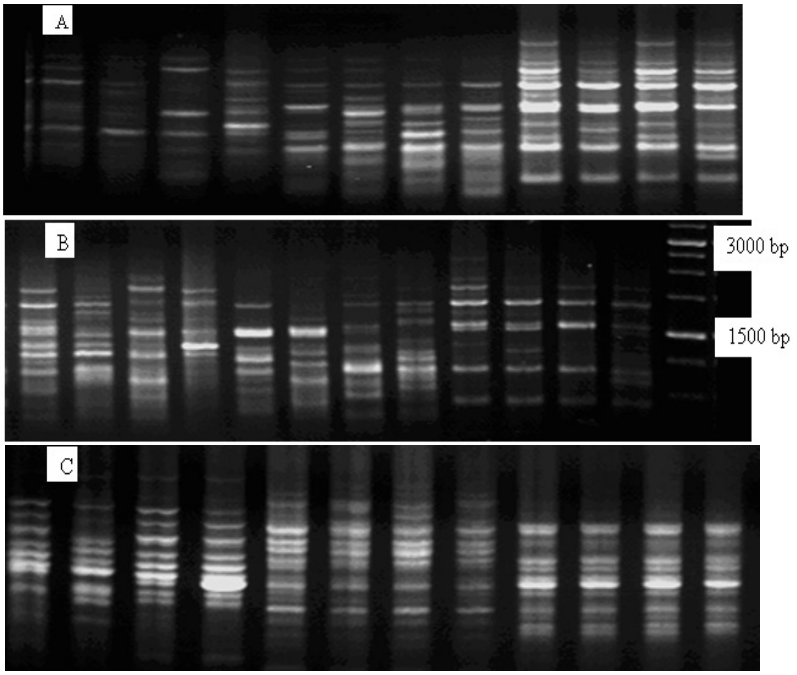

Fig 2. Banding pattern generated by R-RAP primer (A) WLTR2105+UBC3 (B) WLTR2105+UBC77 and (C) Sukkula+OPG10. Lanes from right to left: four accessions of *Triticum aestivum*, *Hordeum vulgare* and *Aegilops tauschii*

content (PIC) ranged from 0.28 (Sukkula+OPA10) to 0. 40 (WLTR2105+UBC3 and OPA02), averaging 0.35 (Table 3).

In another experiment, the reproducibility of the RAPD, IRAP and R-RAP primer combinations were measured. In this case, different types of PCR machine with the same PCR reactions in 4 repeats was applied and, as a result, the obtained products were similar in all replications of R-RAP reactions, which indicates good reproducibility of this new method (Table 2). Ubiquity, abundance and activity of the retrotransposons in plant genomes make them appealing as molecular marker systems. Frequency of a retrotransposon in the genome is different and depends on retrotransposon transposition and activity. Some retrotransposon families tend to cluster together in repeat seas surrounding gene islands (Kalendar and Schulman, 2006). Retrotransposon-based molecular markers in comparison to the other methods, detect large changes in the genome. By contrast, molecular marker technology based on single nucleotide polymorphism (SNP), simple sequence repeat (SSR) which is reversible and their use in determining parental lineage data in any study of the phylogenetic relationship is limited, retrotransposons are irrevisible (Kumar and Hirochika. 2001) and the characteristics such as ubiquity, abundance and super distribution in euchromatin and heterochromatin regions of the plant genome make them the powerful marker system which detects polymorphism, both between and within the plant species. Recently, due to related simplicity and informativity, research on retrotransposon marker systems in

Table 3. Genetic diversity estimates for R-RAP primer combinations in 12 accessions of *Aegilops tauschii*, *Hordeum vulgare* and *Triticum aestivum* .

R-RAP	Polymorphic band	PIC	Ne	He	I
5LTR+OPA10	73%	0.36	1.47	0.23	0.42
5LTR+UBC3	69%	0.34	1.40	0.22	0.41
WLTR2105+OPA02	78%	0.40	1.51	0.29	0.44
WLTR2105+OPA10	92%	0.39	1.54	0.32	0.49
WLTR2105+OPG10	94%	0.37	1.46	0.29	0.44
WLTR2105+UBC3	100%	0.40	1.60	0.42	0.50
WLTR2105+UBC77	92%	0.34	1.53	0.25	0.48
Sukkula+OPA10	69%	0.28	1.40	0.23	0.36
Sukkula+UBC3	60%	0.30	1.42	0.23	0.33
Sukkula+UBC77	70%	0.35	1.49	0.27	0.40
Average	80%	0.35	1.48	0.27	0.42

PIC: polymorphic information content, Ne: effective number of alleles, He: Expected heterozygosity, I: Shanon's information index.

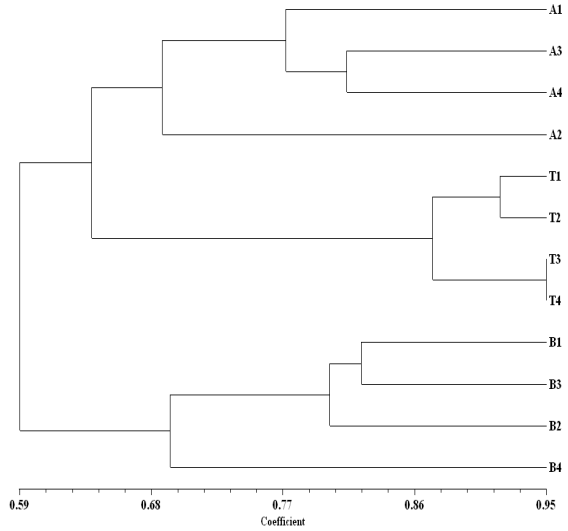

Fig 3. Dendrogram of 12 genotypes using the Dice similarity index and UPGMA clustering method based on R-RAP data. B1, B2, B3 and B4: *Hordeum vulgare*, T1, T2, T3 and T4: *Triticum aestivum* and A1, A2, A3 and A4: *Aegilops tauschii.*

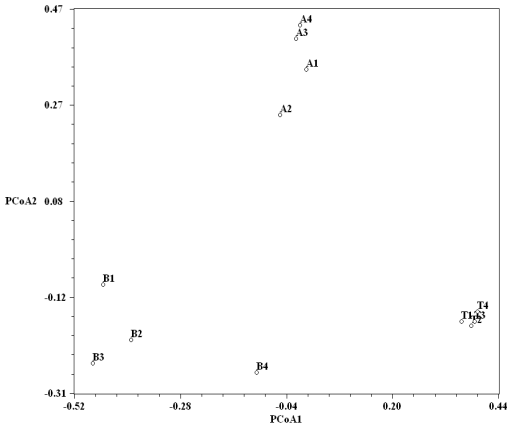

Fig 4. 2D plot of principle coordinate analysis based on 10 R-RAP polymorphic primers divided 3 species to 3 groups. B1, B2, B3 and B4: *Hordeum vulgare*, T1, T2, T3 and T4: *Triticum aestivum* and A1, A2, A3 and A4: *Aegilops tauschii.*

different plant species is increasing (Abdollahi Mandoulakani et al., 2012; Branco et al., 2007; Carvalho et al., 2010). According to our result, R-RAP can be applied as efficiently as other retrotransposon-based molecular markers in plant genetics. An ideal molecular marker technique should have the following criteria: 1) be polymorphic and evenly distributed throughout the genome, 2) provide adequate resolution of genetic differences, 3) generate multiple, independent and reliable markers, 4) be simple, quick and inexpensive, 5) need small amount of DNA sample, 6) require no prior information about the genome for an organism and 7) have linkage to distinct phenotype (Agrawal et al., 2008). As yet, no molecular marker has been known to have all of these aspects together, but scientists should select a suitable marker according to the aim of the research. Retrotransposon transcriptional activation will lead to an increase in copy number and genome size if the newly transposed copies survive selection. For that reason, all retrotransposon-based methods create high polymorphism between and within species and can meet most of the above-mentioned criteria (Kalendar et al., 2010; Abdollahi Mandoulakani et al. 2012). Retrotransposons are long and produce large genetic change at the point of insertion (Schulman, 2006). Over 50% of plant genome comprises of LTR retrotransposons, and the random feature of RAPD sequence present between two LTRs, enhanced the possibility for production of new bands with combination of these two primers. Presenting LTR and RAPD in almost all regions of the genome, R-RAP method covers all the genome. Consequently as demonstrated, it can be used widely in genetic diversity and evolutionary analysis according to high polymorphism generated by this method, both between and within species. Because of the ubiquity of this sequence in all of the genome, it might be used in gene tagging and mapping programs in plants. R-RAP produces a dominant banding pattern like IRAP, REMAP and AFLP. R-RAP is a simple method, no restriction enzymes and no high prior information about the genome are needed; also, the products amplified by this method can be separated into agarose gel with ethidiom bromide staining. Of course, R-RAP should be tested on more plant species with known retrotransposons and even in other organisms to check its utility and applicability. We are applying this method in *Medicago sativa*, *Cucumis melo* and *Linum usitatissimum*.

Materials and methods

Plant material

Plant materials consist of 12 accessions from three species; wheat (*Triticum aestivum* L.), barley (*Hordeum vulgare* L.) and *Aegilops tauschii*, provided by seed and plant improvement institute, Karaj, Iran. Fresh leaves were used for DNA extraction according to Saghai-Maroof's method (Saghai-Maroof et al., 1984) with slight modifications.

PCR program

Three IRAP primers were designed based on retrotransposons WLTR2105, 5'LTR and Sukkula (Saeidi et al., 2008), 5 RAPD and 10 R-RAP primer combinations (Tables 1 & 2) were used in IRAP, RAPD and R-RAP reactions, respectively. Because of the great difference between the Tm of RAPD and LTR primers, the annealing temperature for R-RAP reactions was optimized using gradient PCR. R-RAP PCR reactions were carried out in 10 µl reaction mixtures containing 50 ng of template DNA, 2 mM $MgCl_2$, 0.4mM of dNTPs, 1U of Taq polymerase and 0.4 µM of each primer in 1x PCR reaction buffer. The amplification reaction was performed in the Eppendorf Master cycler (Eppendorf North America, New York, United States) with an initial denaturation for 5 minutes at 94°C, then 35 cycles: 40 seconds denaturation at 94°C; 40 seconds annealing at 58 to 65°C (Table 2); 2 minutes extension at 72°C. Final extension was carried out at 72°C for 5 minutes. Amplified products were electrophoresed on 1.8 % agarose gel containing 1x TBE (Tris Borate EDTA) in constant voltage of 100V for two to three hours. Gels were stained using 0.5µg/ml ethidium bromide and a photograph was taken under UV light using Bio-Rad Gel Documentation system. PCR conditions and electrophoresis for RAPD and IRAP were the same as R-RAP, but the annealing temperatures for RAPD and IRAP were 32 to 42°C and 63.4 to 75 °C, respectively (Table 2). All PCRs were performed 3 to 4 times for each primer combination to measure the reproducibility of the three methods.

Data analysis

Bands were scored as 1 for their presence or 0 for their absence to generate a matrix. A genetic similarity (GS) matrix was computed based on Dice coefficient of similarity, subsequently used to carried out cluster analysis using the unweighted pair group method of arithmetical average (UPGMA) algorithm. Principle coordinate analysis (PCoA) was used. All of these computations were implemented in NTSYS 2.02 software (Rohlf, 2000). Polymorphic information content (PIC), as a marker discrimination power, was computed as $1-\sum P_i^2$, where P_i is the frequency of *i*th allele at a given locus. To estimate genetic variation, effective number of alleles (Ne), Nei's gene diversity (He) and Shanon's information index (I) were calculated using Popgen 1.32 program (Yeh and Boyle, 1997).

Conclusion

The retrotransposones are ubiquitous, abundant and dispersed components of eukaryotic genomes and, on the other hand, random primers of RAPD providing accessibility to whole genome. Thus, combination of them makes a high throughput marker system that is cheap, generic and easy to use. Our results showed that R-RAP can be used as an alternative and complementary method for IRAP and RAPD, although different aspects, ability and limitation of it are unknown. We can propose and imagine the usefulness of this protocol for all genetic programs. Hence, complementary studies should be performed.

Acknowledgements
The authors are grateful to seed and plant improvement institute, Karaj, Iran for providing the seeds. This work was financially supported by University of Guilan and Urmia University.

References

Abdollahi Mandoulakani B, Piri Y, Darvishzadeh R, Bernoosi I, Jafari M (2012) Retroelement insertional polymorphism and genetic diversity in *Medicago sativa* populations revealed by IRAP and REMAP markers. Plant Mol Biol Rep 30:286-296

Agarwal M,Shrivastava N, Padh H (2008) Advances in molecular marker techniques and their application in plant sciences. Plant cell Rep 27: 617-631

Bennetzen JL (2000) Transposable element contributions to plant gene and genome evolution. Plant Mol Biol 42: 251-269

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

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

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, Vicient CM, Ofer Peleg V, Anamthawat-jonsson K, Bolshoy A, Schulaman AH (2004) Large Retrotransposon Derivatives: Abundant, Conserved but Nonautonomous Retroelements of Barely and related Genomes. Genetics 166: 1437-1450

Kalendar R, Schulman A (2006) IRAP and REMAP for retrotransposon-based genotyping and finterprinting. Nature Protoc 1: 2478-2484

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

Kumar A, Hirochika K (2001) Applications of retrotransposons as genetic tools in plant. Trends in Plant Sci 6 (3): 127-134

Mansour A (2008) Utilization of genomic retrotransposons as cladistic markers. J cell Mol Biol 7: 17-28

McClintock B (1956) Intranuclear systems controlling gene action and mutation. Brookhaven Symp Biol. 8: 58-74

Rodriguez M, O'Sullivan D, Donini P, Papa R, Chiapparino E, Leigh F, Attene G (2006) Integration of retrotransposon-based markers in a linkage map of barely. Mol Breed 17: 173-184.

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

Saeidi H, Rahimnejad MR, Heslop-Harrison JS (2008) Retroelement insertional polymorphisms, diversity and phylogeography within diploid D-genome *Aegilops tauschii* (Triticeae, Poaceae) sub-taxa in Iran. Ann Bot 101: 855-861.

Saghai-Maroof MA, Soliman KM, Jorgesen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location and population dynamics. Proc Natl Acad Sci 81: 8014-8018.

Sawalha K, Eideh H, Laham S, Hasaneh H, and Mezeid B (2008) Genetic diversity studies on wheat Landraces in Palestine using RAPD markers in comparision to phenotypic classification. J Appl Biol Sci 2: 29-34

Schulman AH (2007) Molecular markers to assess genetic diversity. Euphytica 158: 313-321.

Williams JGK, Kubelic AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18: 6531-6535

Yeh FC, Boyle TJB (1997) Population genetic analysis of co-dominant and dominant markers and quantitative traits. Belg J Bot 129-157