Discourse Journal of Agriculture and Food Sciences
www.resjournals.org/JAFS
ISSN: 2346-7002
Vol. 5(1): 1-11, April, 2017

Application of retrotransposon based molecular markers

Tadesse Abate

Department of Animal and Range Science, college of Agriculture and natural resources, Madda walabu University, P.O. Box 247, Bale Robe, Ethiopia.

Abstract

Molecular markers are indispensable tools for determining the genetic variation and biodiversity with high levels of accuracy and reproducibility. The dynamism and dispersion of the various groups of transposable elements have led to their widespread exploitation as molecular markers. With regard to their special utility as genetic markers for molecular ecologists and conservation biologists, retrotransposons offer certain advantages over microsatellites and sequence data which are currently in wide use. In recent years, several retrotransposons based molecular marker have been developed. The SSAP, IRAP, and REMAP methods are multiplex and generate anonymous marker bands; RBIP scores individual loci, much as microsatellite-based marker systems do. The methods are variously suited to marker detection on agarose and polyacrylamide slab gels, slab and capillary sequencing devices, and arrays on solid supports. This short review narrates the application of these molecular markers for genetic diversity, gene mapping and phylogenetic studies.

Keywords: Molecular markers, Transposable elements, Retrotransposon

INTRODUCTION

A genetic marker is any character that can be measured in an organism which provides information on the genotype of that organism. Markers are used for genetic diversity study, construction of linkage maps and in the tracking of individuals or lines carrying particular genes. Phenotypic and biochemical markers tend to have the disadvantages of a low degree of polymorphism, relatively few loci and environmentally variable expression. These marker types have been replaced with Molecular marker techniques which are based on naturally occurring polymorphisms in DNA sequences (Wetermeier, 1993). Random Amplified Polymorphic DNA (RAPD) is one of the first developed molecular markers that have been studied widely in all genetic programs. RAPD was desirable marker for many years (Williams *et al.,* 1990; Sawalha *et al*., 2008) because of it uses short primer, no need of priori genome sequence information, high speed, low cost and low technical requirement. On the other hand, presenting many potential priming sites for these sequences, low annealing temperature make the invention for new methods such as Amplified fragment length polymorphism (AFLP) and Simple sequence repeat (SSR) which solved RAPD problems, thus eliminating this system from molecular markers today (Kalendar *et al*., 2011).

Now a days, a vast number of mobile genetic elements (Transposons) have been detected and characterized, and they are regular and ubiquitous constituents of eukaryotic genomes (saedler *et al*., 1996). It was found that transposons have a great impact on genome structure and gene function in nearly all organisms (Kidwell and Lisch, 1997). The two major classes of eukaryotic mobile elements are defined by differences in their mechanism of transposition (Kleckner, 1989; Feschotte *et al*., 2002). DNA transposons (Class II elements), transpose directly via a DNA intermidate, usually present in a low copy number, probably as a consequence of their 'cut and paste' mechanism of transposition. While retrotransposons (Class I elements), transpose to new locations in the genome through RNA intermediate using replicative mechanisms (Kumar and Bennetzen, 1999; Sabot *et al*., 2004).

Recent insertion and excision of TEs have given rise to a series of transposon insertion polymorphisms in closely

related species, subspecies, and haplotypes and served as ongoing sources of genomic and genetic variation (Bennett *et al*., 2004). In particular, retro elements can widely be used as molecular markers today due to the features, such as: integration activity, persistence, dispersion, conserved structure, sequence motifs and high copy number. The advanced markers also utilize these retrotransposable elements (Agarwal *et al*., 2008). 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: sequence-specific amplification polymorphism (SSAP) approach (Waugh et al., 1997) is the most popular transposon-based marker method. It amplifies products between a retrotransposon integration site and a restriction site to which an adapter has been ligated. Inter-retrotransposon amplified polymorphism (IRAP) 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 apart to produce product (Mansour, 2008). Retrotransposon-microsatellite amplified polymorphism (REMAP), to some extent, is similar to IRAP, one of the primer matches to a microsatelite motif (Kalendar et al., 1999). Retrotransposon-based insertional polymorphism (RBIP) is based on flanking regions of a LTR to detect 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). These molecular markers have successfully been used to establish phylogenies, study biodiversity and generate linkage maps for agronomically important traits in several species such as barley, pea, rice and tobacco (Schulman et al 2004; Kalendar et al., 1999; Kenward et al., 1999).

Classification of transposable elements

Transposable elements (TEs) are DNA fragments that can insert into new chromosomal locations and often make copies of themselves. It was discovered in the 1940's by Barbara McClitock in maize as the genetic agents that are responsible for the sectors of altered pigmentation on mutant kernels, and subsequently identified in both prokaryotes and eukaryotes (Capy *et al*., 1998). With the advent of large-scale DNA sequencing, it has become apparent that, TEs are the single largest component of most genomes of eukaryotes. TEs account for at least 45% of the human genome (Lander *et al*., 2001) and 50–80% of some grass genomes (SanMiguel and Bennetze, 1998; Vicient *et al*., 1999; Meyers *et al*., 2001).

Transposable elements are classified into two major groups based on their mechanism of transposition and by comparison of their genomic structures and sequences (Kleckner, 1989; Saedler and Gierl 1996; Finnegan 1989). Class I elements (retrotransposon) do not excise during transposition but produce a new DNA copy from an RNA transcript using reverse transcriptase. This means that during transposition, a new copy of the transposon created, while the original copy remains intact at the donor site (Fig. 1). Class I elements therefore increase in copy number during the process of transposition and can become very abundant. Retrotransposons are further classified into two major subclasses that differ in their structure and transposition cycle. These are the long terminal repeat (LTR) retrotransposons and the non-LTR retrotransposons, which are distinguished by the respective presence or absence of LTRs at their ends. LTR retrotransposons are the most abundant class of retrotransposons (Schulman and Kalendar, 2005). LTR-retrotransposons contribute substantially to the structural diversity of plant genomes (Vitte and Panaud, 2005). In plant genomes, huge numbers of LTR retrotransposon insertions are found in many species and constitute more than half the entire genome in some cases (Kumar and Bennetzen, 1999; Sanmiguel *et al*., 1996). LTR elements are flanked by two identical long terminal repeats and usually have 1-3 open reading frames that encoded structural and enzymatic proteins involved in retrotransposition. Non-LTR retrotransposons lack LTRs and are transcribed from an internal promoter. LINEs, like LTR-retrotransposons, have *gag* and *pol* genes encoding structural and enzymatic activities, and it has been proposed that LINEs could be the precursors of LTR-retrotransposons (Xiong and Eickbush, 1990). On the contrary the small retrotransposons called SINEs are very different from the rest, and none of them encode their own transposition machinery and are thus retrotransposed in trans by enzymatic machinery encoded elsewhere.

In contrast, Class II consists of DNA transposons, which change their location in the genome by a 'cut and paste' mechanism (Figure1. Grzebelus 2006). This means that they excise themselves from the donor site and reintegrate themselves at the acceptor site. The transposase enzyme binds class II element DNA sequences at or near the inverted repeats and catalyzes transposition by cutting out the element sequence and inserting it at a new location.

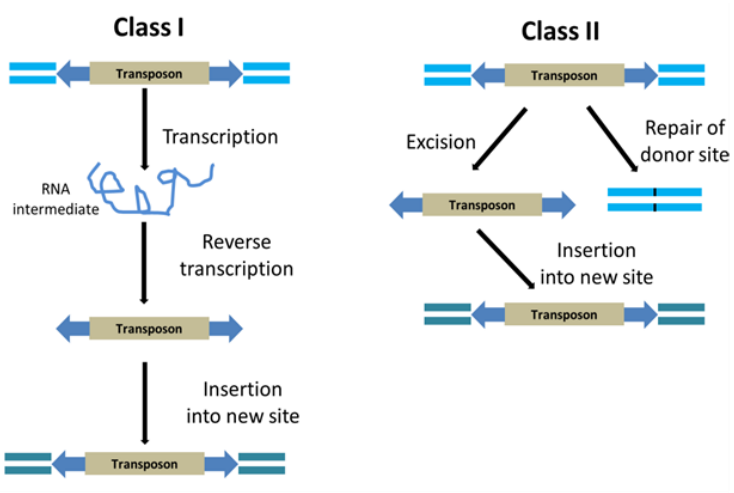

Figure 1. Class I and II transposons and mechanisms of integration

Retrotransposons as molecular markers

A role in generating diversity together with their ubiquitous nature is making TEs important tools for use as molecular markers (Kumar and Hirochika, 2001). TEs based markers have great potential as tools for investigating aspects of molecular ecology, including population structure, conservation genetics, the genetics of speciation, phylogeography and phylogeny (Ray, 2007). Inparticular, retrotransposons provides an excellent basis for the development of markers systems. It is because of they contain long, defined, conserved sequences which can be used for cloning of specific markers and flanking sequences. Moreover, replicationally active members of a retrotransposon family will produce new insertions in the genome leading to polymorphism. The new insertions may then be detected and used to order the insertion events in a lineage, thereby helping to establish phylogenies (Shimamura *et al*., 1997). However, mobile element based markers have the same drawback: difficultly of data interpretation and uncertainty about the true nature of the polymorphism. Specifically, the question may arise as to whether differences in banding patterns are due to the absence or presence of retrotransposons, or are caused by some other mechanism, e.g. indels or restriction site loss. Fortunately, advances in analytical methods and a number of successful studies indicate that these drawbacks can be overcome. In recent years, several molecular marker methods based on retrotransposons have been developed (Kumar and Hirochika, 2001) and they are presented in detail below. All rely on the principle that a joint is formed, during retrotransposon integration, between genomic DNA and the retrotransposon. These joints may be detected by amplification between a primer corresponding to the retrotransposon and a primer matching a nearby motif in the genome. The various methods for revealing insertion polymorphism of retrotransposons differ in the nature of the feature external to the TE that is used for primer design (Schulman *et al*., 2004). With regard to their special utility as genetic markers for molecular ecologists and conservation biologists, retrotransposons offer certain advantages over more commonly used genetic characters such as microsatellites and sequence data. First and foremost is the observation that these markers are essentially homoplasy free characters (Shedlock and Okada 2000; Okada *et al*., 2004; Shedlock *et al*., 2004). In other words, unlike many other genetic markers, they are not prone to exist as character states for reasons other than inheritance from a common ancestor. They are almost invariably identical by descent, not identical by state. Thus, they provide an extremely accurate picture of evolutionary relationships and have proven very successful in elucidating problematic phylogenies in several taxa (Murata *et al*., 1993; Zietkiewicz *et al*., 1999; Takahashi *et al*., 2001b; Salem *et al*., 2003; Sasaki *et al*., 2004; Ray *et al*., 2005a; Watanabe *et al*., 2006). In recent years, several molecular marker methods based on retrotransposons have been developed and they are presented in detail below.

Sequence Specific Amplification Polymorphism (SSAP)

SSAP is a multi-banding marker technology that differences in the insertion sites of transposable elements are targeted to generate polymorphic markers. It was first described by Waugh *et al*., 1997 to investigate the location of *BARE-1* retro transposons in the barley genome. This method is based on the shredding of genomic DNA with two different restriction

enzymes to generate a template for the specific primer PCR: amplification between retrotransposon and adaptors ligated at restriction sites using selective bases in the adaptor primer.

Primers are usually designed in the LTR region, but could also correspond to internal part of the element. Non selective primers can be used when enzymes used for digestion cut infrequently, or when the copy number of the TE is low. For high-copy-number families, the number of selective bases may be increased (Schulman *et al*., 2004). The use of two enzymes in SSAP correspondingly reduces genomic complexity, as does the use of selective bases on the primers associated with the adapters. TEs with low numbers of copies are not well suited to methods to reduce genomic complexity. It converts retrotransposon insertion sites into banding patterns using primers annealing to the junctions between the transposon and the host genome (Grzebelus, 2006). The sizes of the fragments are determined by the distance between the transposon insertion site and the adjacent restriction cut site, with differences in insertion sites between genomes easily visible as different banding patterns (Syed and Flavell, 2006).

Due to their conserved nature, different elements of the Class II retrotransposons can be used to perform an SSAP assay. It is a simple modification of the standard AFLP (Amplified Fragment Length Polymorphism) protocol (Vos *et al*., 1995) but SSAP has been reported to produce higher levels of polymorphism compared with AFLPs (Leigh *et al*., 2003; Queen *et al*., 2004; Tam *et al*., 2005) and requires fewer experimental procedures to generate the desired numbers of markers.

Moreover, SSAP-based markers appear to be better for estimating phylogenetic relationships in plants compared with the conventional AFLP-based makers because one of the primers is based on specific RTN sequences (Figure 1). Restriction digestion of genomic DNA, adapter ligation, and pre amplification PCR are steps similar to AFLP. However, only one restriction site-specific AFLP primer is employed in the final amplification step, whereas the second primer is complementary to a defined DNA sequence. In addition for a typical AFLP procedure no a priori sequence information is required, careful planning and prior transposon sequence knowledge is strongly recommended for SSAP.

SSAP has been carried out on sequencing gels due to the large number of products generated .In the SSAP technique, the selective bases added to the adapter primer or the LTR primer reduced the complexity of the amplified DNA, depending on the copy number of the retrotransposon targets. In the original S-SAP version, (Waugh *et al*., 1997) the specific primer targeted the LTR sequence of the barley retrotransposon *BARE*-1. *BARE*-1 elements are highly abundant in barley (70,000 to 100,000 copies per haploid genome (Manninen *et al*., 1993), but by using three selected bases at the 3′ end of the AFLP primer, banding complexity resulting from the abundance of BARE-1 retrotransposons in barley (*Hordeum vulgare*) was reduced to a useful level (Waugh *et al*., 1997).

SSAP is potentially applicable to any known sequence, given appropriate primer design (Waugh *et al*., 1997). But S-SAP also requires restriction digestion of genomic DNA to provide sites for adapter ligation as in AFLP method. Sensitivity of commonly used restriction enzymes to DNA methylation could provide false genotyping results. S-SAP markers exhibit important advantages for a number of application areas. First, the preference of many transposable elements to integrate into gene-rich regions is certainly beneficial for the purposes of gene isolation by transposon tagging, marker-assisted selection, and map-based cloning. Second, the relatively even distribution of S-SAP markers across genetic maps favorably contrasts with the behavior of AFLP markers, which are often clustered in certain genomic areas (Keim *et al*., 1997). Third, the high level of S-SAP polymorphism, exceeding that of most other PCR-based multilocus marker systems, may be helpful in discriminating closely related accessions. Fourth, as insertional polymorphisms are generated by transpositional events, S-SAP markers can be used to monitor the transpositional activity of the element at an evolutionary timescale.

The SSAP is now used in numerous species where terminal LTR sequences can be effectively cloned (Schulman *et al*., 2004). The use of S-SAP has been described for barley (Waugh *et al*., 1997), wheat (*Triticum aestivum* L.) and wild relatives (Gribbon *et al*., 1999, Queen *et al*., 2004), pea (*Pisum sativa* L.) (Ellis *et al*., 1998), alfalfa (*Medicago sativa* L.) (Porceddu *et al*., 2002). However, proper use of the SSAP technique requires either radioactivity or fluorescent labeling of primers and product detection. Both these drawbacks could be overcome by IRAP.

Inter-Retrotransposon Polymorphism (IRAP)

IRAP is a valuable retrotransposon-based marker. Although retrotransposons are dispersed, they can also be found clustered in the genome. It is the phenomenon of clustering that makes possible the IRAP method, which detects insertional polymorphisms by amplifying the portion of DNA between two nearby retrotransposons using two identical or different primers inserted sufficiently close to each other to allow for efficient amplification, and primers matching the outer segments of the LTRs are generally utilized (Fig.3; Kalender *et al*., 1999; Kalendar and Schulman, 2006). If all retrotransposons were dispersed equally throughout the genome, even for an abundant family such as BARE1, individual elements would be 50 kb apart, and could not yield IRAP amplification templates**.**

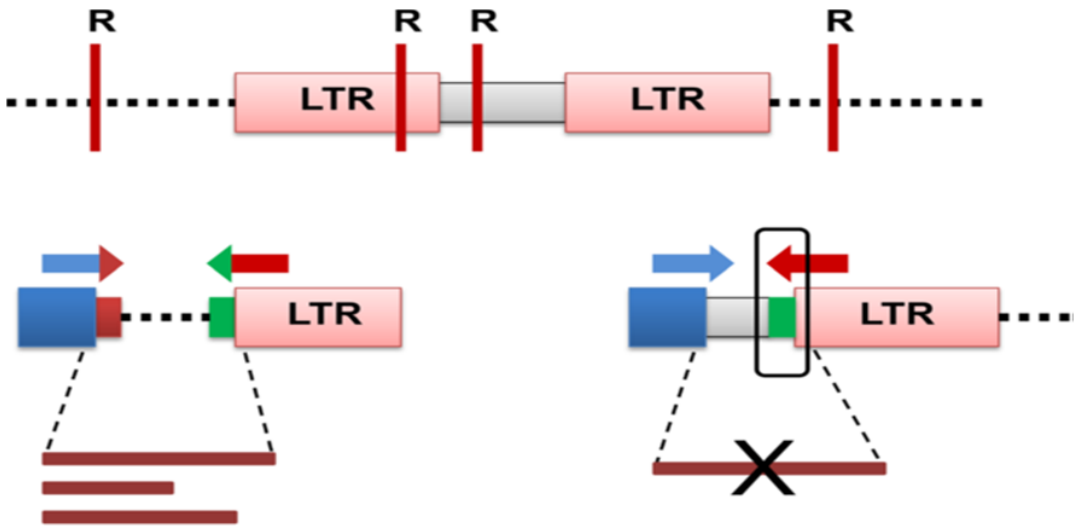

Figure 2. An outline of SSAP. DNA is digested with one frequently cutting and/or rarely cutting restriction enzyme (horizontal red lines marked 'R'). Adapters (blue boxes) are ligated to restricted ends and then a pre-selective amplification is carried out (not illustrated). Selective PCR amplification, shown below, is carried out with LTR (red arrow) and adapter specific (blue arrow) primers. Both primers contain selective nucleotides (colored heads of the arrows) to ensure specific amplification and reduce the number of generated bands to a manageable level. Transposon amplification can only be carried out from the construct shown on the left as the primer at the 3'end contains a selective nucleotide that is absent from the one shown on the right. Generated PCR products of variable length are indicated by brown bars. (**Source:** Waugh *et al*., 1997)

IRAP can be carried out with a single primer matching either the 5' or 3' end of the LTR or with two primers specifically designed for LTRs in the same PCR, but the results will be determined by the orientation of these regions. The targeted Class I elements use the 'copy-and-paste' method of transposition which can take place in either orientation (5' to 3' or 3' to 5'). Besides genomic abundance, this leads to differently oriented gene (copy) clusters found in head to- head, tail-to-tail or head-to-tail orientation (Fig.4). For head-to-head and tail-to-tail arrangements, only a single primer is necessary to generate IRAP products. For head-to-tail orientation, both 5' and 3' LTR primers are needed to amplify the intervening genomic DNA (Weining and Lngridge, 1991).

Variation in retrotransposon insertions into the genome means that the number of sites amplified and sizes of inter-retroelement fragments can be used as markers to detect genotypes poly-morphisms, which in turn, could be used to measure diversity or reconstruct phylogeny (Kumar and Hirochika 2001). Unlike AFLP or SSAP, the IRAP method requires only intact genomic DNA as the template and PCR reagents and apparatus for amplification. The amplification products are generally resolved by electrophoresis in wide-resolution agarose gels, but if labeled primers are used, sequencing gel systems may be employed. IRAP fingerprints with single primers often generate bands from 500 to 3000 bases, lengths that are not convenient for capillary electrophoresis. To reduce the size of the DNA products to be separated and visualized, fluorescent primers may be used in the PCR reaction and the amplicon DNA digested with a four-base-specific restriction enzyme such as *Tai* I or *Taq* I after the PCR reaction. The number of amplification products in IRAP can be reduced by adding selective bases either to the LTR primer. In this technique, polymorphism is detected by the presence or absence of the PCR product. The lack of amplification indicates the absence of the retrotransposon at the particular locus (Kalendar *et al*., 1999; Kalendar *et al*., 2011). The number of IRAP bands is affected both by the abundance of the retrotransposon family on which the primers are based on and on the distribution of the elements or, strictly speaking, of the LTRs (Vukich *et al*., 2009).

Of the various techniques available, IRAP can detect high levels of polymorphism without the need of DNA digestion, ligations or probe hybridization to generate marker data, thus increasing the reliability and robustness of the assay (Kalender *et al.,* 1999; Kalendar and Schul-man 2006). Due to dominance and possibly non- representative genome coverage, the IRAP method would not be suitable for applications in marker-assisted selection, even though it is used in cereals together with other methods (Smýkal 2006)

The IRAP method (Kalendar *et al*., 1999) has found applications in gene mapping in barley (Maninen *et al*., 2000) and wheat (Boyko *et al*., 2002) and in studies of genome evolution in the grasses (Vicient *et al*., 2001). Analyses of IRAP based polymorphism is a very efficient tool for describing of differences among plums (Senková *et al*., 2013) and for sunflower lines (Žiarovská *et al* .,2013).

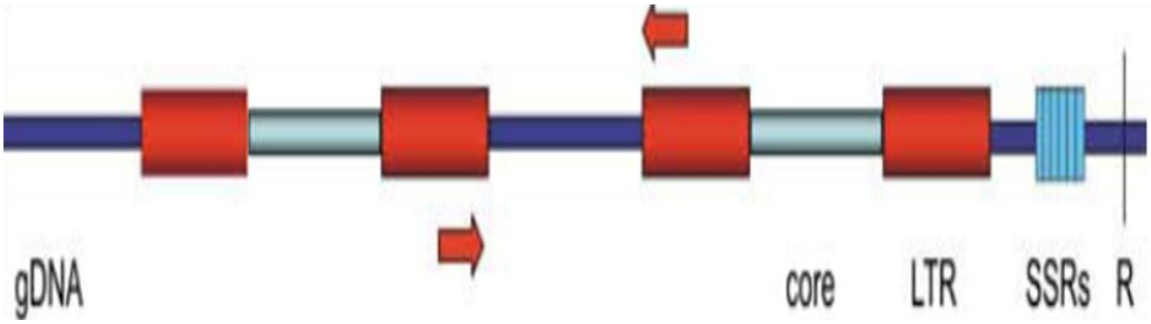

Figure 3. IRAP (Inter Retrotransposon Amplified Polymorphism) Amplification is carried out between the LTRs of two retrotransposons. Genomic DNA (gDNA) is shown as a solid blue line, primers as arrows above and below the genome segment, the retrotransposonas comprised of LTRs and a core domain (core). Other features may be present in the genome, such as microsatellites (SSRs) or restriction sites (R), but IRAP method does not take them into account.

Retrotransposon Microsatellite Amplified Polymorphism (REMAP)

Long terminal repeat regions of the retrotransposons are highly conserved and contain sequences that are essential for integration and expression of these elements (Kalendar *et al*.,1999). Primers based on such regions can be used to amplify dispersed members of a retrotransposon family. Abundant in most genomes, microsatellites or simple sequence repeats seem to be associated with retrotransposons. For instance, in cereals they appear to be associated with retrotransposons (Ramsay *et al*., 1999) and have high mutation rates due to polymerase slippage. Therefore, they may show much variation at individual loci within a species (Schulman *et al*., 2004). This property has been exploited to develop a molecular marker method, retrotransposon microsatellite amplified polymorphism (REMAP), which detects polymorphisms among amplicons produced between microsatellite sequences and RTNs (fig. 5 ; Kalendar *et al*., 1999, Provan *et al*., 1999). To achieve this, one specifically designed LTR primer is mixed with another arbitrarily chosen primer directly containing a simple repeat [e.g. (CA) n, (GA)n] plus one or more non repeat nucleotides at the 3' end to serve as an anchor. The anchor is necessary to provide specificity to the PCR amplification; otherwise, the repetitive structure of the primer might cause it to anneal in multiple positions in any given microsatellite. An anchored primer also prevents the detection of variation in repeat numbers within the microsatellite. The REMAP method is conceptually similar to IRAP, but it differs in that one of the two primers matches a microsatellite motif (Kalendar *et al*., 1999; Kalendar and Schulman, 2006). In addition ,this technique can be regarded as a modified or extended version of the inter- simple sequences repeat (ISSR) technique, since one of the primers in a REMAP reaction is an anchored ISSR primer combined with an IRAP primer.

This method requires neither digestion with restriction enzymes nor ligation to generate the marker bands. Amplification can be performed using single primer or with two primers. Gel electrophoresis of the product is performed using optimal electrophoresis buffers and conditions. This protocol can be completed in 1-2 days. For samples with many or large (41000 bp) bands, gel electrophoresis is performed at a constant voltage of 70 V overnight (20 h). A useful guide for gel architectures differing from 20-20 cm is to first calculate the total volt-hours required for 500-bp fragments to migrate to the bottom of the gel and then adjust the voltage for subsequent runs so that the run time to achieve that total is at least 12 h (Kalendar and Schulman, 2006).The presence or absence of PCR product detects polymorphism in this system. Lack of amplification, however, reveal that retrotransposons are absent at the particular locus

Banding patterns are completely different if REMAP primers are used individually or in combination, indicating that the majority of bands are derived from sequences bordered by a microsatellite on one side, and by an LTR on the other. Usually, the REMAP pattern is more variable than the corresponding inter-simple sequence repeat pattern; and often (but not always, depending on the LTR sequence) the IRAP pattern with primer combinations shows more variability than with a single primer (Leigh *et al*., 2003; Kalendar *et al.,* 1999, 2004; Kalendar and Schulman, 2006).

The RE-MAP method was first used in Hordeum (Kalendar *et al*., 1999) and later for fingerprinting, biodiversity, and genomic stability studies (Kalendar *et al*. 2000; Vicient *et al* ., 2001; Baumel *et al* ., 2002). This technique is used ever since to measure diversity, similarity and cladistic relationships in many genotypes such as (Manninen *et al*., 2000 and 2006; Brick *et al*., 2006), rice (*Oryza sativa.*) *(Branco et al*., *2007)*, Rice blast pathogen *(Magnaporthe grisea SP.*) (Chadha and Gopalakrishna, 2005)*,* spartina *(Spartina sp.)* (Baumel *et al*., 2002) and oat *(Avena sativa L.)*. (Tanhuanpää *et al*., 2007).

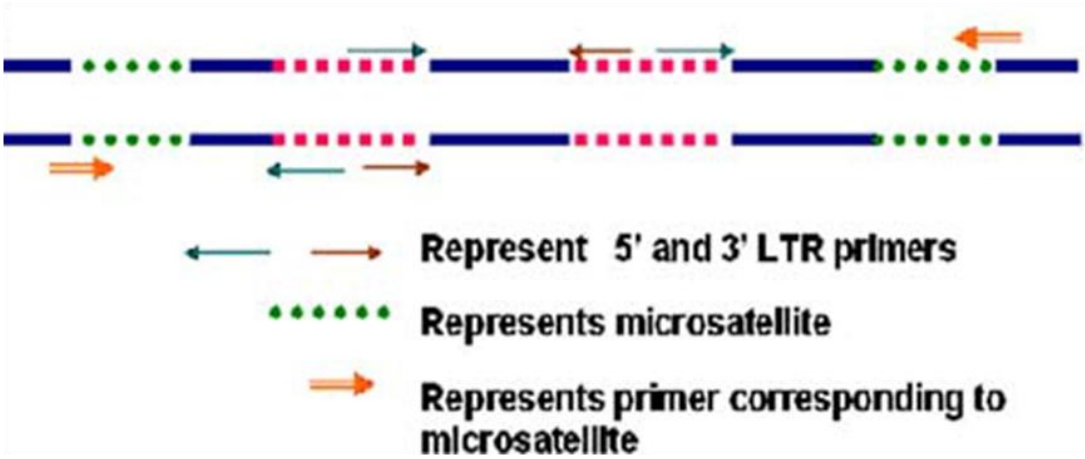

Figure 4. REMAP technique relies on amplification using one outward-facing LTR primer and a second primer from a microsatellites

Retrotransposon-Based Insertion Polymorphism (RBIP)

Retrotransposon Based Insertion Polymorphism (RBIP) is the sole retrotransposon method designed to detect polymorphism in the integration of an element at a particular locus and it has been developed using the *PDR1* retrotransposon in the pea (Flavell *et al.* 1998). RBIPs can detect individual insertions by PCR with flanking host sequence primers and a retrotransposon-specific primer (Flavell *et al*., 1998). The method has also been called insertion sequence based polymorphism (Paux *et al*., 2010).

RBIP differs from the methods described above because it requires the sequence of the 5' and 3' regions flanking the transposon insertions to be completely known. The RBIP method uses primers flanking retrotransposon insertions and scores the presence and absence of insertions at individual sites. Presence or absence of insertion is investigated by two PCRs, the first using one primer from the retro transposon and one from the flanking DNA, the second using primers designed from both flanking regions. When a primer specific to the transposon is used together with a primer designed to anneal to the flanking region, they generate a product from template DNA containing the insertion. On the other hand, primers specific to both flanking regions amplify a product if the insertion is absent. Polymorphisms are detected by simple agarose gelelectrophoresis or by dot hybridization assays. The dot assay for PCR product detection is fully automated for handling thousands of samples.

This method is more expensive and technically complicated than other methods for detecting transposon insertions. However, the particular feature of RBIP that distinguishes it from the other retrotransposon-based marker methods described in above is that it is a single locus, co-dominant technique, with presence and absence of the transposon insertion independently scorable**,** suitable for high-throughput analysis (Flavell *et al*., 1998, Jing *et al*.,2005) . RBIP produces less data per experiment than do multiplex approaches but is more accurate for studies of deeper phylogeny in wide germplasm (Jing *et al*., 2005). For instance, it can detect both the presence and absence of the insertion, whereas other multiplex approaches detect only the presence of insertion, while absence is inferred by band absence.

The main advantage of this approach, and the reason for its development, is the possibility of adapting it to high throughput situations. High throughput marker assays with automated data capture are particularly useful in situations where very large numbers of plant samples must be processed (Flavell *et al*., 1998). Retrotransposon-based SSAP, REMAP, and IRAP are well suited to deal with tens to hundreds of samples (Waugh *et al*., 1997**;** Ellis *et al*., 1998; Kalendar *et al*.,1999; Manninen *et al*.,2000). RBIP is more useful for far larger numbers of samples because it can, in principal, be completely automated. It does not necessarily require agel-based detection system but can easily be automated to gel-free procedures, such as TaqMan or DNA chip technology in order to increase sample throughput. RBIPs have been applied in the study of population genetics and phylogenetic analyses of both plants and animals (Stoneking *et al*., 1997; Batzer and Deininger, 2002; Vitte *et al*., 2004; Jing *et al*., 2005). A strategy akin to RBIP was used successfully to determine the distant phylogenetic relationships between whales and ungulates (Shimamura *et al*., 1997).

A drawback of the method is that sequence data of the flanking regions is required for primer design. This situation is closely analogous to the collection of new flanking sequences for microsatellite or SSR markers. RBIP markers have the added disadvantage with respect to SSR markers that the retrotransposon itself needs to be characterized before

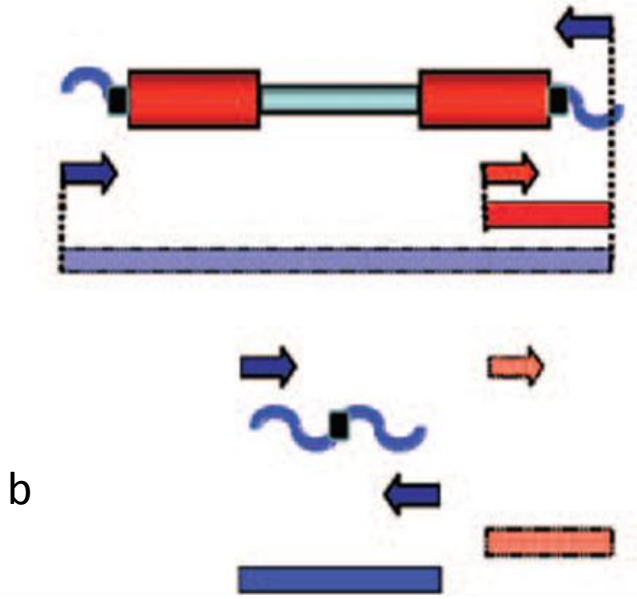

Figure 5. RBIP. (a) Full sites are scored by amplification between a primer in the flanking genomic DNA (shown as a blue wavy line) and a retrotransposon primer. The single product is shown as a red bar beneath the diagram. The alternative reaction between the primers for the left and right flanks (light blue bar beneath the diagram) is inhibited in the occupied site by the length of the retrotransposon. (b) The flanking RBIP primers are able to amplify the empty site, depicted as a deep blue bar beneath the diagram, but amplification from the retrotransposon primer does not occur (missing product shown as a light red bar) because the TE insert is missing.

marker development can begin. This is a smaller problem than it seems at first sight because retrotransposon families span genus boundaries (Flavell *et al*., 1992; Moore *et al*., 1991). Therefore, characterization of the retrotransposons in a species is a one-off investment that can be applied to the entire tribe or family containing them. Sequence data for new RBIP markers may be obtained from sequence analysis of genomic clones. Alternatively, SSAP markers can be converted into RBIP markers.

Conclusion

Retrotransposons are major, dispersed components of most eukaryotic genomes. They replicate by a cycle of transcription, reverse transcription, and integration of new copies, without excising from the genome in the process. Because they represent a major share of the genome, cause easily detectable genetic changes having known ancestral and derived states, and contain conserved regions for which polymerase chain reaction (PCR) primers may be designed, retrotransposon insertions can be exploited as powerful molecular marker systems. The Sequence-Specific Amplified Polymorphism (SSAP) method, the first retrotransposon-based method to be described, amplifies products between a retrotransposon integration site and a restriction site to which an adapter has been ligated. In Inter-Retrotransposon Amplified Polymorphism (IRAP), segments between two nearby retrotransposons or LTRs are amplified. The Retrotransposon-Microsatellite Amplified Polymorphism (REMAP) technique detects retrotransposons integrated near a microsatellite or stretch of SSRs. The Retrotransposon-Based Insertional Polymorphism (RBIP) marker system, in contrast to the others, detects a given locus in both alternative states, namely, empty and occupied b y a retrotransposon, b y using both flanking primers and a retrotransposon primer. These molecular markers revealed high degree of heterogeneity and insertional polymorphism both within and between species, Retrotransposon insertions are irreversible and hence considered particularly useful in phylogenetic studies. They are wide spread throughout the genome and often observed in regions adjacent to known plant genes which revealed their potential to be used in gene mapping study.

REFERENCES

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

Batzer, M., Deininger, P. (2002) Alu repeats and human genomic diversity. *Nat Rev Genet* 3: 370–379.

Baumel A, Ainouche M, Kalendar R, Schulman,A. (2002). Inter-retrotransposon amplified polymor- 24 phism (IRAP), and retotransposonmicrosatellite amplified polymorphism (REMAP) in populations of the young allopolyploid species Spartina (*Spartina SP*.) angelica Hubbard (Poaceae). *Mol Biol Evol* 19: 1218- 1227

Bennett, E.,Coleman, L., Tsui, C., Pittard, W., Devine, S. (2004) Natural genetic variation caused by transposable elements in humans. *Genetics* 168:933–951.

Boyko, E., Kalendar, R., Korzun, V., (2002). A high density cytogenetic map of the *Aegilops tauschii* genome incorporating retrotransposons and defense related genes: insight into cereal chromosome structure and function. *Plant Mol. Biol.* 48 : 767–790.

Branco, C., Vieira, E., Malone, G., Kopp, M., Malone, E., Bernardes, A., Mistura C., Carvalho, F., Oliveira, C. (2007). IRAP and REMAP assessments of genetic similarity in Rice (Oryza *Sativa*). *J. Appl Genet.* 48: 107–113.

Brik, A., Kalendar, R., Stratula, O., Sivolap, M. (2006). IRAP and REMAP analyses of Barley (Hordeum vulgare) varieties of Odessa breeding. *Cytology and Genetics* 3: 24-33.

Capy, P, Bazin ,C., Higuet, D., Langin, T. (1998). Dynamics and evolution of transposable elements. Austin: Landes Bioscience.

Chadha,S. and Gopalakrishna,T. (2005). Retrotransposon-nmicrosatellite amplifie polymorphism (REMAP) markers for genetic diversity assessment of the Rice (*Oryza Sativa* .) blast pathogen (*Magnaporthe grisea*). *Genome* 48: 943-945.

Ellis, T., Poyser, S., Knox, M.,Vershinin, A. and Ambrose, M. (1998) Polymorphism of insertion sites of Ty*1-copia* class retrotransposons and its use for linkage and diversity analysis in pea. *Mol. Gen. Genet.* 260: 9–19.

Feschotte, C., Jiang and Wessler,S. (2002). Plant transposable elements: where genetics meets genomics. *Nature Reviews Genetics* 3: 329-341

Finnegan, D. (1989) Eukaryotic transposable elements and genome evolution. *Trends Genet*. 5: 103–107

Flavell, A., Knox, M., Pearce, S. and Ellis, T. (1998) Retrotransposon- based insertion polymorphisms (RBIP) for high throughput marker analysis. *Plant J.* 16: 643–650.

Flavell, A., Dunbar, E., Anderson, R., Pearce, S., Hartley, R. and Kumar, A. (1992). Ty1 copia group retrotransposons are ubiquitous and heterogeneous in higher plants. *Nucleic Acids Res.* 20 : 3639–3644.

Gribbon, B., Pearce, S., Kalendar R., Schulman, A., Paulin L., Jack P., Kumar A., Flavell A. (1999). *Mol. Gen. Genet*. 261: 883–891

Grzebelus, D. (2006). Transposon insertion polymorphism as a new source of molecular markers. *J Fruit Ornam Plant Res.* 14:21–29

Jing, R., Knox, M., Lee, J., Vershinin ,A., Ambrose,M., Ellis, T., Flavell, A. (2005). Insertional polymorphism and antiquity of PDR1 retrotransposon insertions in Pisum species. *Genetics* 171: 741–752

Kalendar , R., Schulman, H. (2006). IRAP and REMAP for retrotransposon-based genotyping and fin-gerprinting. *Nature Protocols* 1: 2478 – 2484

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

Kalendar, R. (2011). The use of retrotransposon-based molecular markers to analyze genetic diversity. *Ratarstvo i Povrtarstvo.* 48: 261–274.

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

Kalendar, R., Tanskanen, J., Immonen, S., Nevo, E., Schulman, A. (2000) Genome evolution of wild barley (Hordeum spontaneum) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence. *Proc Natl Acad Sci.* 97:6603– 6607

Kalendar,R., Antonius, K., Smýkal, P., Schulman, A. (2010): iPBS: A universal method for DNA fingerprinting and retrotransposon isolation. *Theor. Appl. Genet*. 121: 1419- 1430

Keim, P., Schupp, J., Travis, S., Clayton, K., Zhu,T., Shi, L., Ferreria, A., and Webb, D. (1997) A high-density soybean genetic map based on AFLP markers, *Crop Sci*. 37: 537–543.

Kenward, K.,Bai,D.,Ban, M.,Brandle,J. (1999). Isolation and characterization of *Tnd-1*, a retrotransposon marker linked to black root resistance in tobacco. *Theor. Appl. Genet.* 98: 387–395

Kidwell, M., Lisch, D. (1997). Transposable elements as sources of variation in animals and plants. *Proc. Natl. Acad. Sci*. 94: 7704--7711.

Kleckner, N. 1989. Transposon Tnl0. In *Mobile DNA* (Berg,D. and Howe,M.eds), pp. 227-268. American Society for

Microbiology, Washington, DC.
Kumar A. and Bennetzen J. (1999). Plant Retrotransposon. *Annual Review of Genetics.* 33: 479- 532
Kumar, A. and Hirochika, H. (2001). Applications of retrotransposons as genetic tools in plant biology. *Trends in Plant Science* 6:127-134.
Lander, E. Linton,L.,Birren, B.,Nusbaum,C.,Zody,M. (2001) Initial sequencing and analysis of the human genome. *Nature* 409: 860–921
Leigh ,F., Kalendar, R., Lea, V., Lee, D., Donini, P., Schulman, A. (2003). Comparison of the utility of barley retrotransposon families for genetic analysis by molecular marker techniques. *Mol Genet Genomics* 269: 464-474
Manninen, I., and Schulman, A. (1993) BARE-1, a *copia*-like retroelement in barley (*Hordeum vulgare*), *Plant Mol. Biol*. 22: 829–864
Manninen, O., Jalli, M., Kalendar, R., Schulman, A., Afanasenko, O., and Robinson, J.(2006) Mapping of major spot-type and net-type netblotch resistance genes in the Ethiopian Barley *(Hordeum vulgare*) line CI 9819. *Genome* 49: 1564–1571.
Manninen, O., Kalendar, R., Robinson, J., and Schulman, A. (2000) Application of *BARE*-1 retrotransposon markers to map a major resistance gene for net blotch in barley. *Mol. Gen. Genet.* 264: 325–334.
Mansour,A. (2008). Utilization of genomic retrotransposons as cladistic markers. *Jurnal of Cell and Molecular Bioligy*. 7:17-28
Meyers, B., Tingey, S. and Morgante, M. (2001). Abundance, distribution, and transcriptional activity of repetitive elements in the maize genome. *Genome Res.* 11: 1660–1676).
Moore, G., Lucas, H., Batty, N. and Flavell, R. (1991) A family of retrotransposons and associated genomic variation in wheat. *Genomics* 10 : 461–468.
Murata, S., Takasaki, N., Saitoh, M., Okada, N. (1993). Determination of the phylogenetic relationships among Pacific salmonids by using short interspersed elements (SINEs) as temporal landmarks of evolution. *Proceedings of the National Academy of Sciences* 90: 6995–6999.
Okada, N, Shedlock, A., Nikaido, M. (2004) Retroposon mapping in molecular systematics. In: *Mobile Genetic Elements: Protocols and Genomic Applications* pp. 189–226. Humana Press, Totowa, NJ.
Paux, E., Faure, S., Choulet, F., Roger, D., Gauthier, V., Martinant, J., Sourdille, P., Balfourier, F., Le Paslier, M., Chauveau, A., Cakir, M., Gandon, B., Feuillet, C. (2010). Insertion site-based polymorphism markers open new perspectives for genome saturation and marker-assisted selection in wheat. *Plant Biotechnol J*. 8:196–210
Porceddu A., Albertini E., Barcaccia G., Marconi G., Bertoli, F. and Veronesi F. (2002) Development of S-SAP markers based on an LTR-like sequence from Medicago sativa L. *Mol. Genet*. *Genomics* 267: 107–114.
Provan, J. Thomas,W., Forster,B. and Powell,W. (1999). *Copia*-SSR: a simple marker technique which can be used on total genomic DNA. *Genome* 42: 363–366
Queen, R., Gribbon, B., James, C.,Jack, P., Flavell, A. (2004). Retrotransposon-based markers for linkage and genetic diversity analysis in wheat. *Molecular Genetics and Genomics* 271: 91-97.
Ramsay, L., Macaulay, M., Cardle, L., Morgante, M., Ivanissevich, S., Maestri, E., Powell,W. and Waugh, R.(1999). Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, UK, Intimate association of microsatellite repeats with retrotransposons and other dispersed repetitive elements in Barley (*Hordeum vulgare)*. *Plant J*. 17: 415–425.
Ray, D. (2007). SINEs of progress: mobile element applications to molecular ecology. *Molecular Ecology* 16:19–33
Ray,D., Hedges, D., Hall, M. (2005a) Alu Insertion Polymorphisms and Platyrrhine Primate Phylogenetic Relationships. *Molecular Phylogenetics and Evolution*. 35: 117–126.
Sabot, F., Simon, D., Bernard, M. (2004). Plant transposable elements, with an emphasis on grass species. *Euphytica* 139: 227-247.
Saedler, H., and Gierl, A. (1996). Transposable Elements, Springer, Berlin
Salem, A., Ray, D., Xing, J. (2003). Alu elements and hominid phylogenetics. *Proceedings of the National Academy of Sciences.* 100: 12787- 12791.
SanMiguel, P. and Bennetzen, J. (1998). Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. *Ann. Bot.* 81: 37–44
Sanmiguel,P.,Tikhonov,A., Jin,Y., Motchoulskaia,N.,Zakharov,D. (1996). Nested retrotransposons in the intergenic regions of the maize genome. *Science* 274: 765–768.
Sasaki T, Takahashi, K., Nikaido, M. (2004) First application of the SINE (short interspersed repetitive element) method to infer phylogenetic relationships in reptiles: an example from the turtle superfamily Testudinoidea. *Molecular Biology and Evolution* 21: 705–715.
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, A. and Kalendar, R. (2005). Amovable feast: diverse retrotransposons and theircontribution to Barley(*Hordeum vulgare*) genome dynamics. *Cytogenet. Genome Res*. 110: 598–605.

Schulman,A.,Flavell,A.,Ellis, T. (2004). The application of LTR retrotransposons as molecular markers in plants. *Methods Mol Biol.* 260:145-73

Senková,S., Žiarovská,J., Bežo,M. V. Štefúnová,V. and K. Ražná,K. (2013). Utilization of IRAP technique for plums genotypes differentiation. *Bioscience Research* 10: 01-07.

Shedlock, A.,Takahashi, K.,Okada, N. (2004). SINEs of speciation: tracking lineages with retroposons. *Trends in Ecology and Evolution*. 19: 545–553.

Shedlock,A., Okada,N. (2000) SINE insertions: powerful tools for molecular systematics. *Bioessays*, 22:148–160.

Shimamura, M., Yasue, H., Ohshima, K., Abe, H., Kato, H., Kishiro, T. (1997). Molecular evidence from retroposons that whales form a clade within eventoed ungulates. *Nature* 388: 666–670.

Smýkal,P. (2006). Development of an efficient retrotransposon-based fingerprinting method for rapid pea variety identification. *J Appl Genetic* 47: 221–230

Stoneking M, Fontius, J., Clifford, S., Soodyall, H., Arcot, S., Saha, N., Jenkins, T., Tahir, M., Deininger, P., Batzer, M. (1997) Alu insertion polymorphisms and human evolution: evidence for a larger population size in Africa. *Genome Res* 7: 1061–1071

Syed,N., Flavell, A. (2006). Sequence-specific amplification polymorphisms (SSAPs): a multi-locus approach for analyzing transposon insertions. *Nat Protoc* 1: 2746–2752.

Takahashi, K., Terai, Y., Nishida, M., Okada, N. (2001b). Phylogenetic relationships and ancient incomplete lineage sorting among cichlid fishes in Lake Tanganyika as revealed by analysis of the insertion of retroposons. *Molecular Biology and Evolution* 18: 2057–2066.

Tam,S., Mhiri, C., Vogelaar, A., Kerkveld, M., Pearce, S., Grandbastien, M. (2005) Comparative analyses of genetic diversities within tomato and pepper collections detected by retrotransposon-based SSAP, AFLP and SSR. *Theoretical and Applied Genetics* 110: 819-31.

Tanhuanpää, P, Kalendar, R, Schulman, A., Kiviharju, E . (2007). A major gene for grain cadmium accumulation in oat (*Avena sativa L*.). *Genome*. 50: 588- 94

Vicient, C., Suoniemi,A., Anamthawat-Jonsson,K., Tanskanen,J.,Beharav,A.,Nevo,E.,Schurman,A. (1999) Retrotransposon BARE-1 and its role in genome evolution in the genus *Hordeum*. *Plant Cell* 11: 1769–1784

Vicient,C., Jaaskelainen, M., Kalendar, R., and Schulman, A. (2001). Active retrotransposons are a common feature of grass genome. *Plant Physiol*. 125: 1283–1292.

Vitte, C., Ishii, T., Lamy, F., Brar, D., Panaud, O. (2004) Genomic paleontology provides evidence for two distinct origins of Asian rice (Oryza sativa L.). *Mol Genet Genomics* 272: 504–511

Vitte, C. and Panaud, O.(2005). LTR retrotransposons and flowering plant genome size: emergence of the increase/decrease model. *Cytogenet Genome Res* . 110: 91-107

Vos,P., Hogers,R. Bleeker,M, Reijans,M., Lee,T., Hornes,M Frijters,A. Pot,J., Peleman,J., Ellis T., Poyser, S., Knox M., Vershinin, A., Ambrose, M. (1995) *Mol. Gen. Genet*. 260: 9-19

Vukich, M., Schulman, A., Giordani, T., Natali, I., kalendar, R., Cavallini, A. (2009). Genetic variability in sunflower (*Helianthus annuus* L.) and in the *Helianthus* genus as assessed by retrotransposon-based molecular markers. *Theoretical and Applied Genetics* 119: 1027: 1038.

Watanabe, M., Nikaido, M., Tsuda, T. (2006) The rise and fall of the CR1 subfamily in the lineage leading to penguins. *Gene* .365 : 57–66.

Waugh, R., Mclean,K.,Flavell,A., Pearce,S., Kumar,A., Thomas,B. and Powell,W. (1997). Genetic distribution of Bare-1-like retrotransposable elements in barley genome revealed by sequence specific amplification polymorphisms (SSAP), *Molecular General Genetics* 253: 687-694.

Weining, S., Langridge, P. (1991). Identification and mapping of polymorphisms in cereals based on the polymerase chain reaction. *Theor Appl Genet*. 82: 209–216

Wetermeier, R. Electrophoresis in Practice, 1993; VCH, New York, USA.

Williams, J., Kubelik, A., Livak, K., Rafalski, J. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res*. 18: 6531 6535

Xiong, Y., Eickbush, T. (1990). Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J*. 9: 3353–3362.

Žiarovská,J., Milan Bežo,M. (2013). Inter hacre-1 retrotransposon length polymorphism is not conserved throught the sunflower lines. *Journal of Microbiology, Biotechnology and Food Sciences.* 2: 1629-1641

Zietkiewicz, E., Richer, C., Labuda, D. (1999) Phylogenetic affinities of tarsier in the context of primate Alu repeats. *Molecular Phylogenetics and Evolution* 11: 77–83.