Cytogenet Genome Res 110:598–605 (2005)
DOI: 10.1159/000084993

Cytogenetic and Genome Research

A movable feast: diverse retrotransposons and their contribution to barley genome dynamics

A.H. Schulman,[a,b] R. Kalendar[b]

[a] Plant Breeding Biotechnology, MTT Agrifood Research, Jokioinen
[b] MTT/BI Plant Genomics Laboratory, Institute of Biotechnology, Viikki Biocenter, University of Helsinki, Helsinki (Finland)

Manuscript received 10 December 2003; accepted in revised form for publication by J.-N. Volff 9 March 2004.

Abstract. Cellular genes comprise at most 5% of the barley genome; the rest is occupied primarily by retrotransposons. Retrotransposons move intracellularly by a replicative mechanism similar to that of retroviruses. We describe the major classes of retrotransposons in barley, including the two nonautonomous groups that were recently identified, and detail the evidence supporting our current understanding of their life cycle. Data from analyses of long contiguous segments of the barley genome, as well as surveys of the prevalence of full-length retrotransposons and their solo LTR derivatives in the genus *Hordeum*, indicate that integration and recombinational loss of retrotransposons are major factors shaping the genome. The sequence conservation and integrative capacity of barley retrotransposons have made them excellent sources for development of molecular marker systems.

An inkling of our current view of genome organization began to emerge from physical experiments with DNA hybridization carried out in the 1970s (e.g., Flavell et al., 1974; Hake and Walbot, 1980). The observations were that DNA from bacteria demonstrated a simple reannealing curve over time, whereas genomic DNA from higher organisms showed a compound curve indicating the presence of several components differing in their relative concentration, or repetitiveness. Shearing of the DNA to various lengths made it clear that the repetitive and non-repetitive fractions are interspersed. Subsequent analyses have shown that the repetitive DNA constitutes the majority of large genomes. In maize, 60–80% DNA is repetitive (Flavell et al., 1974; Springer et al., 1994; Bennetzen, 1996; Heslop-Harrison, 2000; Meyers et al., 2001), whereas in wheat, 83% of the genome is repetitive (Wicker et al., 2001).

Supported by funding from the European Commission under the MMEDV (QLK5-CT-1999-01499) and Gediflux (PROPOSAL: QLK5-CT-2001-00934) projects.

Request reprints from Alan H. Schulman, Plant Breeding Biotechnology
MTT Agrifood Research, Jokioinen, P.O. Box 56
Viikinkaari 4, FIN–00014 Helsinki (Finland)
telephone: +358 9 40 768 2242; fax: +358 9 191 58952
e-mail: alan.schulman@helsinki.fi

The measurements of the proportion of repetitive DNA in the genome are roughly in line with estimates on gene number from large-scale EST and genome sequencing projects. These indicate a total of 30,000–60,000 genes are present in plant genomes on average (Arabidopsis Genome Initiative, 2000; Goff et al., 2002; Kurata et al., 2002; Yu et al., 2002). A minimum total of 3×10^4 genes of 5 kb each require a gene space of 1.5×10^8 bp. Taking genome size into account, the minimum genome complement therefore represents 35% of the rice genome (4.3×10^8 bp; Kurata et al., 1997) but only 2.8% of the barley genome (5.3×10^9 bp, Kankaanpää et al., 1996). In contrast, the genome of *Arabidopsis thaliana* is remarkably compact (Schmuths et al., 2004). Its genome sequence reveals a total of 26,439 genes of 4,870 bp average length, occupying 40.4% of the 1.29×10^8 nucleotides in the genome (http://mips.gsf.de/proj/thal/db/tables/chrall_tables/exons.html).

Recent sequencing of long segments ("contigs") of cereal genomes confirms the interspersion of repetitive DNA with genes that are present as single copies or small families. The repetitive DNA is comprised of several components. These include transposable elements that move via a cut-and-paste mechanism, which are also called the Class II DNA transposons, and those that replicate via an RNA intermediate leading to cDNA copies that insert into new loci, which are called retrotransposons or Class I elements. The microsatellites or simple

1424–8581/05/1104–0598$22.00/0

Accessible online at:
www.karger.com/cgr

sequence repeats, as well as tandem repeats and special groups of repetitive sequences found in telomeres and centromeres, round out the genome. Comparisons of mapped genes and sequenced contigs show that, for the most part, cereal genomes are syntenic and collinear, having equivalent genes organized in the same order on homeologous chromosomes or chromosome segments (Smilde et al., 2001). The overall conclusion is that the bulk of cereal genomes is comprised of families of retrotransposons, with a major contributing factor to genome size differences being the abundance of these elements.

An understanding of mobile genetic elements that have an RNA phase in their life cycle emerged about the same time the repetitive and interspersed nature of plant genomes was revealed. While research on pathogenic retroelements began with the work of Rous (1911) on the avian leucosis virus and sarcoma virus, only the provirus hypothesis of Temin (1964), its confirmation through the discovery of reverse transcriptase (Baltimore, 1970; Temin and Mizutani, 1970), and the details of the first retrovirus sequence (Shinnick et al., 1981) clarified their nature. Non-pathogenic but related elements, retrotransposons, where identified in *Drosophila* and *Saccharomyces*; related elements in the plant genome were demonstrated in the mid and late 1980s (Shepherd et al., 1984; Voytas and Ausubel 1988; Grandbastien et al., 1989).

The retrotransposons, or Class I elements, constitute three groups, the LINEs, SINEs, and LTR retrotransposons, members of this last group being bounded by long terminal repeats (LTRs). The LTR retrotransposons may be divided into two large categories both by sequence similarity and organization, those that resemble the *Drosophila* type element *copia* and those that resemble *gypsy* elements. Members of each group display a greater similarity to members of the same group from different species and Kingdoms than they do to elements from the same species, supporting their ancient divergence into independent lineages (Xiong and Eickbush, 1990). The LTR retrotransposons together appear to derive from non-LTR retroelements (Malik and Eickbush, 2001). The structure and life cycle of LTR retrotransposons (reviewed in Kumar and Bennetzen, 1999) is similar to that of the intracellular phase of mammalian retroviruses. The 5′ LTR functions as the promoter, whereas the 3′ LTR provides terminator and polyadenylation signals.

The LTRs of cereal retrotransposons can vary in length from a few 100 bp (*Tos17*; Hirochika et al., 1992) to over 5 kb (*Sukkkula*, Shirasu et al., 2000; *Grande*, García-Martínez and Martínez-Izquierdo, 2003) and have at their termini small inverted repeats, which give the LTRs a universal 5′ TG....CA 3′ structure. Immediately internal to the 5′ and 3′ LTRs are the priming sites for reverse transcription, respectively for generating the (–) strand and (+) strand. An internal domain specifying one or more open reading frames lies in between the primer binding sites. In *copia*-like elements this encodes in the sense orientation a capsid protein (GAG), aspartic proteinase (AP), which cleaves the expressed polyprotein into functional components, integrase (IN), which carries out the insertion of the cDNA into the genome, reverse transcriptase (RT), which makes the cDNA daughter copies, and RNase H (RH), which is required in replication (Turner and Summers, 1999). In *gypsy*-like elements, the integrase is at the end of the open reading frame. As discussed below in more detail, a recently recognized subset of plant *gypsy*-like retrotransposons also specifies a coding domain *(env)* for a putative envelope protein (Vicient et al., 2001a; Wright and Voytas, 2002). The *env* domains have also been found in *copia*-like elements (Laten et al., 1998; Kapitonov and Jurka, 1999; Peterson-Burch et al., 2000).

LINEs and SINEs are distinct from the LTR retrotransposons in their structure and mode of replication. These elements are the major class of transposable elements in the human genome and important players in its dynamics (Kazazian and Goodier, 2002). In plants, by contrast, generally the LTR retrotransposons appear to have that role. LINE elements contain an internal promoter and encode proteins respectively with RNA-binding and encapsidation, endonuclease, and reverse transcriptase functions. Replication takes place at the integration site itself, where the endonuclease nicks the host DNA to generate a free 3′ hydroxyl that serves as the primer for reverse transcription of the LINE. The SINEs, unlike the LINEs, do not encode proteins for their replication, and likely utilize those of the LINEs for their replication and propagation (Boeke, 1997; Lenoir et al., 2001; Szafranski et al., 2004). The SINEs are generally tRNA derivatives and, as such, are transcribed by RNA polymerase III; they contain internal promoters that render them transcriptionally active (Deragon and Capy, 2000).

LTR retrotransposons in the barley genome

Moore and coworkers (1991) reported a structure in barley, referred to as BIS-1, that later was established to have the organizational features of a retrotransposon (Abbo et al., 1995). Similiarly, early work by McIntyre et al. (1988) revealed structures reminiscent, in their size and repetitiveness, of retrotransposon LTRs. However, the first complete retrotransposon described for barley was *BARE*-1 (Manninen and Schulman, 1993). The *BARE*-1 element was isolated serendipitously, aided by its high copy number. The *BARE*-1 retrotransposon is 8.9 kb in size and *copia*-like, comprised of LTRs of 1.8 kb, an unusually long untranslated leader of 2 kb, and coding domains for GAG, AP, IN, RT, and RH. The coding domains and replicational signals are well conserved in this family (Suoniemi et al., 1997, 1998a).

The *BARE*-1 family has proven to be a fruitful experimental system. The elements have been shown to be transcriptionally active in various barley tissues and in tissue culture (Suoniemi et al., 1996a) and their encoded polyprotein translated and processed not only in barley (Jääskeläinen et al., 1999), but also in other cereals (Vicient et al., 2001b) where members of this family are present. The *BARE*-1 family is present on average in 14,000 full-length copies, or about 2.9% of the total size of the genome (Vicient et al., 1999). Solo LTRs of *BARE*-1 account for at least ~64,000 additional elements in the genome, an equivalent share of its bulk. *BARE*-1 contributes both to size variation over the entire genome (Vicient et al., 1999) and to local genome diversification and size differences on a local scale (Kalendar et al., 2000), in which drought-related abiotic factors appear to play a role. As is discussed below, the *BARE*-1

Table 1. Barley LTR retrotransposons present in DNA databases

Retroelement[a]	Synonym or similarity [b]	Accession	Length	
			Element	LTR
Unknown class				
Ikeros		AY013246	7700	380
Inav		AF474982	5180	204
Leojyg		AY268139	4600	141
Usier	Nikita-1	AY268139	10300	1503
Nikita-1	Usier	AF254799, Y14573, AF474072	?	>1077
Copia-like				
BARE-1	Wis-2 (wheat)	Z17327, AF427791	8600	1800
Bianca		AF521177	7683	170
HORPIA-1		AF427791		
Horpia-2		AF521177	5142	323
Inga		AF474982	11700	1576
Katarina		AF521177	5140	323
Gypsy-like				
Bagy-1		Y14573	14500	2500
Cereba		AY040832	7800	931
Fatima		AF521177	9100	484
Kulkuri	Daniela-1 (wheat)?	AJ489246, AF474071	12000	893
Latidu		AF521177	9800	464
Laura	Bagy-1	AY188332	12403	4192
Lolaog		AY268139	10510	
Romani-Hv1, -Hv2, -Hv3		AJ002615, AJ002616, AJ002617, AJ002618, AJ002619	11000	2755, 3100
Gypsy-env[c]				
Egug	Bagy-2, Wilma (wheat)	AY268139	7300	1358
Bagy-2	Egug, Wilma (wheat)	AF254799, AF427791	8700	1500
Sabrina	R173 (rye), Wobi	AF474071	7924	1661
R173	Sabrina	AF474071, AF474072, Z17327	?	597
Wham	Claudia (wheat)-like LTR	AY268139	9560	1446
Wobi	R173, Sabrina	AY268139	?	?
LARD				
Sukkula-2		AF474072, AF254799	11200	5585
Sukkula-1(a,b,c)		AF474072, AF254799, Y14573, AY054377, AY054376, AF453665	11800 – 12200	5020
TRIM				
Pikku-1		AY164585	725	262

[a] Some names are present in the TREP database (http://wheat.pw.usda.gov/ITMI/ Repeats/ nrTREP_list.html).
[b] "?" indicates uncertain affinity.
[c] gypsy-env refers to gypsy elements containing *env* domains, also called Errantiviruses.

elements so prevalent in the genome, are frequently present as solo LTRs in extensive nests of elements integrated into each other. The activity of *BARE*-1 has made it suitable for exploitation as a molecular marker for genetic diversity and mapping (Kalendar et al., 1999; Manninen et al., 2000; Baumel et al., 2002; Boyko et al., 2002; Leigh et al., 2003).

The worldwide expansion of sequencing efforts, particularly of the long contiguous DNA segments represented in BAC and YAC clones, has greatly expanded the known repertoire of barley retrotransposons. Due to the prevalence of many retrotransposon families, even clones from gene-rich regions of the genome contain ample numbers of retrotransposons. Retroelements that can be gleaned from the database accessions of barley are presented in Table 1. An online database of repeats present in the Triticeae, TREP, which includes many barley retrotransposons, is now available (http://wheat.pw.usda.gov/ITMI/Repeats/nrTREP_list.html). It is now clear that barley and other plants contain all groups of retroelements present elsewhere among eukaryotes.

The *gypsy*-like elements, as described above, were originally described in the arthropods and in a few taxonomically scattered plants. We were able to establish, however, that *gypsy*-like retroelements are present in barley and are widespread in the plant kingdom (Suoniemi et al., 1998b). Vershinin and collaborators (2002) cloned RT domains in barley representing four families of *gypsy*-like elements. Hybridizing representatives of four *gypsy*-like families to gridded and blotted BAC clones, they found that together these families represented some 3% of the barley genome, the most abundant being present in almost 7,000 copies. If the barley genome resembles that of rice, where *gypsy*-like elements outnumber those that are *copia*-like by almost six to one on chromosome 1 (Sasaki et al., 2002) and

perhaps 2:1 overall (McCarthy et al., 2002), there are many more families of this class to be found, and they will prove to be more diverse than the *copia*-like elements as they are in rice (Vicient and Schulman, 2002).

One group of LTR retrotransposons that is especially interesting contains envelope domains. As described above, retrotransposons and retroviruses share the basic features of their intracellular life cycles as well as the major proteins that they encode. One crucial difference leading to distinct terms being used for these elements is that retrotransposons lack a domain, *env*, encoding an envelope glycoprotein. This glycoprotein enables the VLP to be sheathed in an envelope and also mediates fusion of the envelope with cellular membranes, facilitating cellular transit and infectivity. The retroviruses were previously thought to be restricted to the mammals, although a few invertebrate retrotransposons, including *gypsy* itself, contain *env* domains and can be infective (Malik et al., 2000). During the last several years, a clade of *gypsy*-like retrotransposons, *Gypsy-env*, that contains *env* domains were identified in *Arabidopsis thaliana* (Pélissier et al., 1995; Kapitonov and Jurka, 1999; Peterson-Burch et al., 2000) together with the related elements *Cyclops* and *Calypso* of the legumes (Chavanne et al., 1998; Peterson-Burch et al., 2000). A few *copia*-like retroelements containing *env* domains were also found (Laten et al., 1998; Wright and Voytas, 2002). The *env*-containing *gypsy*-like elements are not rare, but form part of a separate and widespread clade found throughout the flowering plants (Vicient et al., 2001a). In barley, the *Bagy*-2 member of this group is transcribed and the *env* domain spliced into a subgenomic RNA in a manner parallel to what is found in the retroviruses (Vicient et al., 2001a). The predicted envelope protein of *Bagy*-2 contains the N-glycosylation sites, leucine zipper, and transmembrane domains typical of the retroviral protein. The function of this protein in cellular or extracellular phases of the *Bagy*-2 life cycle remains, however, to be established.

SINEs and LINEs in barley

SINE elements are active in a variety of plants, and have been investigated in particular in *Brassica* (Gilbert et al., 1997; Arnaud et al., 2000; Tikhonov et al., 2001). They are useful phylogenetic markers in that genus (Tatout et al., 1999), in *Oryza* (Cheng et al., 2002), and in humans and animals (Shedlock and Okada, 2000; Watkins et al., 2003). However, they have been little investigated in barley and in cereals other than rice. The distribution of *copia*-like, *gypsy*-like, and LINE elements was examined in *H. spontaneum* and in *Aegilops speltoides* (Belyayev et al., 2001). LINE elements showed a dispersed, but nevertheless clustered, pattern. Vershinin et al. (2002) reported on LINE distribution in members of the genus *Hordeum*. They observed dispersion along most chromosomes, limited by the elements' low copy numbers in cultivated barley. The group employed degenerate RT primers to amplify a LINE RT segment from barley, and then used this to isolate a LINE element, named BLIN, from a BAC library. The isolated BLIN element contained the structural features and organization typical of other LINE elements, although the open reading frame had decayed. The BLIN is present in only 40 to 50 copies, in contrast to the *del2* LINE element of *Lilium*, which is present in 2.5×10^5 copies and comprises 4 % of that genome (Leeton and Smyth, 1993).

Non-autonomous LTR retrotransposons in barley

Systems of non-autonomous and autonomous elements were found to be characteristic of Class II, DNA transposons since early in their analysis (McClintock, 1956; Burr and Burr, 1982; Fedoroff et al., 1983). The non-autonomous transposons depend for their mobilization on the transposase of their autonomous partner. Recently, MITE elements were shown as well to be a form of highly reduced non-autonomous Class II transposable elements (Jiang et al., 2003; Kikuchi et al., 2003; Nakazaki et al., 2003). For the retrotransposons, although fully functional retrotransposons could be isolated in plants by selection in tissue culture (Grandbastien et al., 1989), the majority of elements in large families of retrotransposons appeared to contain stop codons or small indels rendering their reading frames unable to code for full-length polyproteins (e.g., Suoniemi et al., 1998a). These retroelements are similar in that way to most nonautonomous Class II transposons, but differ from the highly reduced, perhaps minimum transposable Class II element represented by the MITEs.

Recently, the conceptual equivalent of MITE elements has, however, been found for retrotransposons. Witte et al. (2001) described a class of highly reduced non-autonomous retrotransposons, which were given the appellation "TRIMs". The TRIM elements are composed of terminal direct repeats (LTRs or TDRs) of 100 to 250 bp, flanking typical retrotransposon (–) strand PBS and (+) strand PPT primers and an internal "stuffer" region of 100 to 300 bp. The internal region is non-coding. Their small size prompted the acronym for Terminal Repeat retrotransposons In Miniature. The only cereals TRIMs heretofore found are in rice (16 copies, of which eight were intact) and maize (two copies, one intact). We now have isolated a family of highly repetitive, insertionally polymorphic TRIMs from barley, which we demonstrate to be actively transcribed as well. We also have shown that TRIM elements highly conserved in length and sequence can be found across the tribe Triticeae (Kalendar et al., in preparation).

We have recently found another class of retrotransposons, non-autonomous but at the opposite end of the size spectrum from TRIMs. These elements, named LARDs for LArge Retrotransposon Derivatives, contain LTRs of ~ 4.4 kb (Kalendar et al., 2004). The *Sukkula* element, a solo LTR derived from a LARD, was previously reported (Shirasu et al., 2000). These large LTRs flank primer-binding sites, as is typical for both TRIMs and autonomous elements. However, the internal domains of ~ 3.5 kb contain no open reading frames encoding typical retrotransposon proteins. The internal domains are sufficiently diverged from typical retrotransposon coding sequences, if that is their origin, that they cannot be satisfactorily aligned with them. However, the LARDs specify well-conserved RNA secondary structures. Although non-autonomous, the LARDs are also well conserved in sequence, polymorphic

by transposons display methods, and are highly abundant, being present in 10^4 copies. We have described them in 13 species and four genera of the grass tribe Triticeae, including barley (Kalendar et al., 2004).

The retrotransposon life cycle in barley

To date, no single copy of a barley retrotransposon has been demonstrated to integrate during the course of a laboratory experiment, as was the *Tnt1* of tobacco, which is present in moderate numbers (Grandbastien et al., 1989), or was *Tos*-17 of rice, which occurs only rarely in the genome (Hirochika et al., 1996; Miyao et al., 2003). However, progress has been made in demonstrating the various steps of the retrotransposon life cycle, including transcription, translation, and formation of virus-like particles (VLPs). For the cereals generally, retrotransposon transcripts appear to be present in many tissues of the plant, and can be found as about 0.1% of the clones in EST databases (Vicient et al., 2001b; Echenique, et al., 2002). In barley, 0.3% of the ESTs are derived from retrotransposons (Vicient et al., 2001b), the highest of the nine monocots sampled, and the highest for all species having more than 1×10^3 ESTs in the database. Furthermore, among the 27 monocot, one pine, and 10 dicot retrotransposon families investigated, which are about evenly split between *copia*- and *gypsy*-like types, the barley retrotransposons *Sukkula* (a LARD element), *Sabrina* (*gypsy*-like), and *BARE*-1 (*copia*-like) matched the greatest number of ESTs in the database, approached only by *Cinful* of maize.

Transcription of retrotransposons in barley has been directly demonstrated for *BARE*-1 (Suoniemi et al., 1996a), the *env*-containing *Bagy*-2 (Vicient et al., 2001a), and for TRIM elements (Kalendar et al., in prep.). The *BARE*-1 LTR, which is 1,829 bp long, contains two promoters, both of which are active in vivo. For the downstream TATA box, both positive and negative regulatory regions for transient expression in barley protoplasts were uncovered (Suoniemi et al., 1996a). The upstream segment giving the greatest positive contribution to expression resembles a GARE element (Huang et al., 1990), which responds transcriptionally to gibberellic acid by interaction with nuclear factors (Huttly et al., 1992; Sutliff et al., 1993). The role of *BARE*-1 promoter motifs within intact organs has been confirmed for barley (Chang et al., in prep.). There is also circumstantial evidence that *BARE*-1 is stress-responsive, because the copy number of this element is correlated with environmental factors associated with drought and temperature stress in natural populations of *Hordeum spontaneum*, wild barley (Kalendar et al., 2000).

Reverse transcription follows transcription, and depends upon a (–)-strand initiation site (PBS), which is generally complementary to a tRNA in LTR retrotransposons, and a (+) strand initiation site (PPT), which is purine-rich. These are highly conserved both for *BARE*-1 (Suoniemi et al., 1997) and LARD elements (Kalendar et al., 2004) in barley. Evidence for production and packaging of *BARE*-1 cDNAs has been found (Jääskeläinen et al., 1999). Translation of the LTR retrotransposons generates one single polyprotein, or two separate *gag* and *pol* products. These are processed to mature forms by retrotransposon-encoded proteinase (Irwin and Voytas, 2001) The *BARE*-1 elements contain a single open reading frame, as does yeast Ty5, and antibodies to both GAG and integrase reveal that processing to the predicted mature sizes takes place in a variety of tissues (Jääskeläinen et al., 1999). Furthermore, *BARE*-1 GAG antibodies also detect a full-length translation product and its proteolytic processing products from barley and several other cereals including rice (Vicient et al., 2001b). Translation of a third encoded domain, for the envelope protein, has not yet been directly shown in barley for the *Bagy*-2 element, although the requisite RNA splicing takes place (Vicient et al., 2001a). The GAG and integrase products of *BARE*-1 appear able to form virus-like particles, which can be visualized from sucrose gradients, associated with reverse transcriptase activity (Jääskeläinen et al., 1999).

Integration and loss of retrotransposons shape the barley genome

As described in the Introduction, cellular genes occupy roughly 5% of the barley genome. Most of the rest of the genome appears to be comprised of retrotransposons. The *BARE*-1 element is present in *Hordeum* genomes as $1–2 \times 10^4$ copies of full-length elements of 8.9 kb and $60–187 \times 10^4$ solo LTRs of 1.8 kb, depending on the accession analyzed (Vicient et al., 1999). Thus, 5–10% of the genome is comprised of this single retrotransposon family. The prevalence of the *Bagy-2* retrotransposon appears comparable to that of *BARE*-1, the genome containing about 10^4 copies of the 10-kb element (Vicient et al., 2001a). Full-length LARD elements and their solo LTRs together comprise about 1.3% of the genome (Kalendar et al., 2004). Hence, these three families together compose 8–13% of the genome, and there are many other families (Table 1; Shirasu et al., 2000) for which the copy number has not yet been estimated. Many of the highly abundant retrotransposons are also dispersed throughout all seven haploid chromosomes, except in the centromeres, telomeres, and nucleolar organizing regions (Suoniemi et al., 1996b; Vicient et al., 1999). However, the *gypsy*-like element *CEREBA* is highly, if not exclusively, localized to cereal centromeric regions (Presting et al., 1998; Hudakova et al., 2001). Centromeric retrotransposons are found elsewhere in the plant kingdom, and may represent a specialized, ancient family of *gypsy*-like elements.

Sequencing of long contiguous segments of barley and other cereal chromosomes has produced a view, consistent with early data from analytical centrifugation and reannealing experiments, that these genomes are compartmented into relatively compact islands of genes, "gene space," surrounded by large "seas" of repetitive DNA, largely comprised of retrotransposons (Barakat et al., 1997; Sandhu and Gill, 2002). In a 60-kb segment of the barley genome from chromosome 4HL containing the *Mlo* locus, the three genes are clustered in a block of 18 kb (Panstruga et al., 1998). A 66-kb region on chromosome 2 HL is similar, containing three genes on an island of 18 kb flanked by an expanse that has undergone at least 15 retrotransposon integrations (Shirasu et al., 2000). A contiguous

stretch of 261 kb at the barley *Mla* locus segment contains 32 predicted genes separated into three gene islands by two complexes of nested retrotransposons (Wei et al., 2002). The *BARE*-1 elements alone comprise 17.5% of the total region. An even longer contig of 417.5 kb, assembled from four BACs, confirmed the gene island model of the barley genome (Rostoks et al., 2002). Reverse transcription renders the LTRs of a given element identical at the time of integration. However, over time, their sequences subsequently diverge at the neutral rate. This forms a clock that can be used to estimate the time that has elapsed since the integration of a particular retrotransposon copy at a given position in the genome. This principle enabled Wei and co-authors (2002) to estimate that the *Mla* locus of barley was invaded by five *BARE*-1 elements over the last two million years.

While these few case studies of particular regions in the barley genome cannot establish an overall picture of the rate of retrotransposon insertion into the genome as whole, the currently large fraction occupied by the elements implies that the genome has been generally expanding for a long time. Are the barley and other genomes, therefore, expanding without end (Bennetzen and Kellogg, 1997a, b)? Unequal crossing-over, gene conversion, and sequential deletions may not be sufficient to efficiently remove retrotransposons and halt the expansion. Although small deletions, in some animals and in plants with small genomes, may be sufficient to maintain the size of these genomes in equilibrium, in plants with large genomes mechanisms acting directly on transposable elements may be more important (Bennetzen, 2002; Petrov, 2002). Data from barley indicates that LTR – LTR recombination may play a major role in counteracting retrotransposon-based genome expansion. This form of intrachromosomal recombination in a single retrotransposon results in loss of one LTR and the internal domain, leaving behind a recombinant solo LTR. In the small genome of yeast, 85% of the retroelements are present as solo LTRs (Kim et al., 1998). Genomes of *Hordeum* species have 7–42 fold more *BARE*-1 LTRs than full-length copies of this retrotransposon (Vicient et al., 1999). Furthermore, the nests of retrotransposons in barley appear to be comprised almost entirely of solo LTRs (Shirasu et al., 2000). In the genome of rice, the >98% similarity of most pairs of LTRs in intact elements (McCarthy et al., 2002) supports the hypothesis that turnover may serve to eliminate elements and keep genomes small. Furthermore, an analysis of intergenic sequences for orthologous glutenin loci in *Triticum* revealed large-scale retrotransposon insertions and deletions as well as other kinds of rearrangements (Anderson et al., 2002; Wicker et al., 2003)

Whatever the relative contributions of the various mechanisms promoting loss of DNA in the genome, the end result in barley and other cereals are the gene islands discussed above. The current snapshot of the genome, available in long contiguous sequences, reveals variation in both gene density and retrotransposon number. Overall gene densities of one per ~ 20 kb have been observed in several contigs, within a very wide range of one per 12–103 kb (Panstruga et al., 1998; Feuillet and Keller, 1999; Shirasu et al., 2000; Rostocks et al., 2002). These yield local densities inside the gene islands themselves ranging from one in 4.6–8.1 kb (Shirasu et al., 2000; Wei et al., 2002). The highest densities approach that of one per 4.5 kb found in *Arabidopsis thaliana*. The intra-island gene densities in the large genomes mirror, therefore, the more global organization in retrotransposon-poor genomes.

High local genic densities, through nested retrotransposon integrations, may be selectively advantageous due to the capacity of LTRs from two different retrotransposons on a chromosome to recombine (Shirasu et al., 2000; Vicient and Schulman, submitted), removing all DNA in between. If the intervening DNA contains genes, these would be lost through such a recombination, leading to loss of fitness. Retrotransposons with an affinity for integrating into genes in barley, if they exist, are expected to be rare and under the tight regulation seen for *Tos17* of rice (Hirochika et al., 1996; Hirochika, 1997).

Retrotransposons as molecular markers in barley

The retrotransposons of barley have proven to be an excellent foundation on which to build molecular marker systems for applications such as genetic fingerprinting (Leigh et al., 2003), variety identification, genetic diversity studies (Kalendar et al., 2000; Baumel et al., 2002), marker-assisted selection, and mapping (Waugh et al., 1997; Manninen et al., 2000). Several PCR-based methods that produce multiple products have been developed. They all employ a primer anchored in retrotransposon, generally in the LTR. They differ in the annealing site of the second primer. The S-SAP (sequence-specific amplified polymorphism) method, first applied in barley, uses a primer matching a restriction site adapter (Waugh et al., 1997). In IRAP (inter-retrotransposon amplified polymorphism), the second primer matches another retrotransposon, either from the same or a different family (Kalendar et al., 1999). The REMAP (retrotransposon-microsatellite amplified polymorphism) approach uses a retrotransposon primer and one, bearing a selective anchor at the 3′ end, which matches a microsatellite.

The IRAP method succeeds because, rather than being dispersed at a predicted average distance of 50 kb, retroelements are clustered in the barley genome. Likewise, microsatellites and retrotransposons tend to be clustered in the genome, increasing the utility of REMAP (Ramsay et al., 1999). The methods are robust regarding the generation of solo LTRs. Although LTR-LTR recombination may delete the internal region of an element, the remaining solo LTR leaves many such losses transparent for marker systems based on LTRs. This is because recombination does not remove the annealing site, generally near the outer edge of the LTR, for primers used in polymorphism detection. Similarly, stacked insertions of one retrotransposon into another, and subsequent deletions within such nests, will not affect the products generated from the LTRs of the original element because these events will be "behind", or 5′, to the position of the primer. Retrotransposon marker systems such as RBIP (Flavell et al., 1998) and TAM (Flavell et al., 2003), which score the presence or absence of individual insertions based on specific amplification between the flanking sequences, have not yet been adapted to barley, but are currently under development.

Prospects

The genus *Hordeum* has proven to be a good experimental system for investigating the life cycle of plant retrotransposons and their impact on the genome, and for applying them as molecular markers. To date, only a few elements from barley have been investigated in any detail. Many aspects of the life cycle remain to be clarified, and the regulatory mechanisms linking retrotransposon expression to cellular processes have not yet been revealed. A low-copy mutagenic element such as *Tos17* has not yet been found in barley. However, many investigative tools are either now available or will be soon. These include the large international genomics effort currently under way, a diversity of mutants, good genetic maps, synteny with other cereal genomes, genetically diverse wild barley, and widespread wild species. These, combined with the need for practical tools in barley breeding, ensure that more advances will be forthcoming.

References

Abbo S, Dunford RP, Foote TN, Reader SM, Flavell RB, Moore G: Organization of retro-element and stem-loop repeat families in the genomes and nuclei of cereals. Chromosome Res 3:5–15 (1995).

Anderson OD, Larka L, Christoffers MJ, McCue KF, Gustafson JP: Comparison of orthologous and paralogous DNA flanking the wheat high molecular weight glutenin genes: sequence conservation and divergence, transposon distribution, and matrix-attachment regions. Genome 45:367–380 (2002).

Arabidopsis Genome Initiative: Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815 (2000).

Arnaud P, Goubely C, Pélissier T, Deragon J-M: SINE retroposons can be used in vivo as nucleation centers for de novo methylation. Mol Cell Biol 20: 3434–3441 (2000).

Baltimore D: RNA-dependent DNA polymerase in virions of RNA tumor viruses. Nature 226:1209–1211 (1970).

Barakat A, Carels N, Bernardi G: The distribution of genes in the genomes of Gramineae. Proc Natl Acad Sci USA 94:6857–6861 (1997).

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

Belyayev A, Raskina O, Nevo E: Chromosomal distribution of reverse transcriptase-containing retroelements in two Triticeae species. Chromosome Res 9:129–136 (2001).

Bennetzen JL: The contributions of retroelements to plant genome organization, function and evolution. Trends Microbiol 4:347–353 (1996).

Bennetzen JL: Mechanisms and rates of genome expansion and contraction in flowering plants. Genetica 115:29–36 (2002).

Bennetzen JL, Kellogg EA: Do plants have a one-way ticket to genome obesity? Plant Cell 9:1509–1514 (1997a).

Bennetzen JL, Kellogg EA: Reply. Plant Cell 9:1901–1902 (1997b).

Boeke JD: LINEs and Alus – the polyA connection. Nat Genet 16:6–7 (1997).

Boyko E, Kalendar R, Korzun V, Gill B, Schulman AH: Combined mapping of *Aegilops tauschii* by retrotransposon, microsatellite, and gene markers. Plant Mol Biol 48:767–790 (2002).

Burr B, Burr FA: Ds controlling elements of maize at the shrunken locus are large and dissimilar insertions. Cell 29:977–986 (1982).

Chavanne F, Zhang DX, Liaud MF, Cerff R: Structure and evolution of *Cyclops*: a novel giant retrotransposon of the Ty3/Gypsy family highly amplified in pea and other legume species. Plant Mol Biol 37:363–375 (1998).

Cheng C, Tsuchimoto S, Ohtsubo H, Ohtsubo E: Evolutionary relationships among rice species with AA genome based on SINE insertion analysis. Genes Genet Syst 77:323–334 (2002).

Deragon J-M, Capy P: Impact of transposable elements on the human genome. Ann Med 32:264–273 (2000).

Echenique V, Stamova B, Wolters P, Lazo G, Carollo L, Dubcovsky J: Frequencies of *Ty1-copia* and *Ty3-gypsy* retroelements within the Triticeae EST databases. Theor Appl Genet 104:840–844 (2002).

Fedoroff N, Wessler S, Shure M: Isolation of the transposable maize controlling elements Ac and Ds. Cell 35:235–242 (1983).

Feuillet C, Keller B: High gene density is conserved at syntenic loci of small and large grass genomes. Proc Natl Acad Sci USA 96:8265–8270 (1999).

Flavell RB, Bennett MD, Smith JB, Smith DB: Genome size and the proportion of repeated nucleotide sequence DNA in plants. Biochem Genet 12:257–269 (1974).

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

Flavell AJ, Bolshakov VN, Booth A, Jing R, Russell J, Ellis TH, Isaac P: A microarray-based high throughput molecular marker genotyping method: the tagged microarray marker (TAM) approach. Nucleic Acids Res 31:e115 (2003).

García-Martínez J, Martínez-Izquierdo JA: Study on the evolution of the *Grande* retrotransposon in the *Zea* genus. Mol Biol Evol 20:831–841 (2003).

Gilbert N, Arnaud P, Lenoir A, Warwick SI, Picard G, Deragon J-M: Plant S1 SINEs as a model to study retroposition. Genetica 100:155–160 (1997).

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, et al: A draft sequence of the rice genome *(Oryza sativa* L. ssp. *japonica)*. Science 296:92–100 (2002).

Grandbastien M-A, Spielmann A, Caboche M: *Tnt1*, a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics. Nature 337:376–380 (1989).

Hake S, Walbot V: The genome of *Zea mays*, its organization and homology to related grasses. Chromosoma 79:251–270 (1980).

Heslop-Harrison JS: Comparative genome organization in plants: from sequence and markers to chromatin and chromosomes. Plant Cell 12:617–636 (2000).

Hirochika H: Retrotransposons of rice: their regulation and use for genome analysis. Plant Mol Biol 35:231–240 (1997).

Hirochika H, Fukuchi A, Kikuchi F: Retrotransposon families in rice. Mol Gen Genet 233:209–216 (1992).

Hirochika H, Sugimoto K, Otsuki Y, Tsugawa H, Kanda M: Retrotransposons of rice involved in mutations induced by tissue culture. Proc Natl Acad Sci USA 93:7783–7788 (1996).

Huang N, Sutliff TD, Litts JC, Rodriguez RL: Classification and characterization of the rice alpha-amylase multigene family. Plant Mol Biol 14:655–668 (1990).

Hudakova S, Michalek W, Presting GG, ten Hoopen R, dos Santos K, Jasencakova Z, Schubert I: Sequence organization of barley centromeres. Nucl Acids Res 29:5029–5035 (2001).

Huttly AK, Phillips AL, Tregear JW: Localization of *cis* elements in the promoter of a wheat alpha-*Amy2* gene. Plant Mol Biol 19:903–911 (1992).

Irwin PA, Voytas DF: Expression and processing of proteins encoded by the *Saccharomyces* retrotransposon *Ty5*. J Virol 75:1790–1797 (2001).

Jääskeläinen M, Mykkänen A-H, Arna T, Vicient C, Suoniemi A, Kalendar R, Savilahti H, Schulman AH: Retrotransposon *BARE*-1: Expression of encoded proteins and formation of virus-like particles in barley cells. Plant J 20:413–422 (1999).

Jiang N, Bao Z, Zhang X, Hirochika H, Eddy SR, McCouch SR, Wessler SR: An active DNA transposon family in rice. Nature 421:163–167 (2003).

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

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

Kalendar R, Vicient CM, Peleg O, Anamthawat-Jonsson K, Bolshoy A, Schulman AH: LARD retroelements: Conserved, non-autonomous components of barley and related genomes. Genetics 166:1437–1450 (2004).

Kankaanpää J, Mannonen L, Schulman AH: The genome sizes of *Hordeum* species show considerable variation. Genome 39:730–735 (1996).

Kapitonov V, Jurka J: Molecular paleontology of transposable elements from *Arabidopsis thaliana*. Genetica 107:27–37 (1999).

Kazazian HH Jr, Goodier JL: LINE drive: Retrotransposition and genome instability. Cell 110:277–280 (2002).

Kikuchi K, Terauchi K, Wada M, Hirano HY: The plant MITE mPing is mobilized in anther culture. Nature 421:167–170 (2003).

Kim JM, Vanguri S, Boeke JD, Gabriel A, Voytas DF: Transposable elements and genome organization: A comprehensive survey of retrotransposons revealed by the complete *Saccharomyces cerevisiae* genome sequence. Genome Res 8:464–478 (1998).

Kumar A, Bennetzen J: Plant retrotransposons. Annu Rev Genet 33: 479–532 (1999).

Kurata N, Umehara Y, Tanoue H, Sasaki T: Physical mapping of the rice genome with YAC clones. Plant Mol Biol 35:101–113 (1997).

Kurata N, Nomomura K-I, Harushima Y: Rice genome organization: the centromere and genome interactions. Ann Bot 90:427–435 (2002).

Laten H, Majumdar A, Gaucher EA: *SIRE-1*, a *copia/Ty1*-like retroelement from soybean, encodes a retroviral envelope-like protein. Proc Natl Acad Sci USA 95:6897–6902 (1998).

Leeton PR, Smyth DR: An abundant LINE-like element amplified in the genome of *Lilium speciosum*. Mol Gen Genet 237:97–104 (1993).

Leigh F, Kalendar R, Lea V, Lee D, Donini P, Schulman AH: Comparison of the utility of barley retrotransposon families for genetic analysis by molecular marker techniques. Mol Genet Genomics 269:464–474 (2003).

Lenoir A, Lavie L, Prieto J-L, Goubely C, Côté J-C, Pélissier T, Deragon J-M: The evolutionary origin and genomic organization of SINEs in *Arabidopsis thaliana*. Mol Biol Evol 18:2315–2322 (2001).

Malik HS, Eickbush TH: Phylogenetic analysis of ribonuclease H domains suggests a late, chimeric origin of LTR retrotransposable elements and retroviruses. Genome Res 11:1187–1197 (2001).

Malik HS, Henikoff S, Eickbush TH: Poised for contagion: evolutionary origins of the infectious abilities of invertebrate retroviruses. Genome Res 10: 1307–1318 (2000).

Manninen I, Schulman AH: *BARE*-1, a *copia*-like retroelement in barley (*Hordeum vulgare* L.). Plant Mol Biol 22:829–846 (1993).

Manninen O, Kalendar R, Robinson J, Schulman AH: Application of *BARE*-1 retrotransposon markers to map a major resistance gene for net blotch in barley. Mol Gen Genet 264:325–334 (2000).

McCarthy EM, Liu J, Lizhi G, McDonald JF: Long terminal repeat retrotransposons of *Oryza sativa*. Genome Biology 3:research00053 (2002). http://genomebiology.com/2002/3/10/Research/0053.

McClintock B: Controlling elements and the gene. Cold Spring Harbor Symp Quant Biol 21: 197–216 (1956).

McIntyre CL, Clarke C, Appels R: Amplification and dispersion of repeated DNA sequences in the *Triticeae.* Pl Syst Evol 160:39–59 (1988).

Meyers BC, Tingey SV, Morgante M: Abundance, distribution, and transcriptional activity of repetitive elements in the maize genome. Genome Res 11:1660–76 (2001).

Miyao A, Tanaka K, Murata K, Sawaki H, Takeda S, Abe K, Shinozuka Y, Onosato K, Hirochika H: Target site specificity of the *Tos17* retrotransposon shows a preference for insertion within genes and against insertion in retrotransposon-rich regions of the genome. Plant Cell 15:1771–1780 (2003).

Moore G, Cheung W, Schwarzacher T, Flavell R: BIS 1, a major component of the cereal genome and a tool for studying genomic organization. Genomics 19:469–476 (1991).

Nakazaki T, Okumoto Y, Horibata A, Yamahira S, Teraishi M, Nishida H, Inoue H, Tanisaka T: Mobilization of a transposon in the rice genome. Nature 421:170–172 (2003).

Panstruga R, Büschges R, Piffanelli P, Schulze-Lefert P: A contiguous 60 kb genomic stretch from barley reveals molecular evidence for gene islands in a monocot genome. Nucl Acids Res 26:1056–1062 (1998).

Pelissier T, Tutois S, Deragon JM, Tourmente S, Genestier S, Picard G: *Athila*, a new retroelement from *Arabidopsis thaliana*. Plant Mol Biol 29:441–452 (1995).

Peterson-Burch BD, Wright DA, Laten HM, Voytas DF: Retroviruses in plants? Trends Genet 16:151–152 (2000).

Petrov DA: Mutational equilibrium model of genome size evolution. Theor Popul Biol 61:531–544 (2002).

Presting GG, Malysheva L, Fuchs J, Schubert I: A *Ty3/gypsy* retrotransposon-like sequence localizes to the centromeric regions of cereal chromosomes. Plant J 16:721–728 (1998).

Ramsay L, Macaulay M, Carle L, Morgante M, degli Ivanissevich S, Maestri E, Powell W, Waugh R: Intimate association of microsatellite repeats with retrotransposons and other dispersed repetitive elements in barley. Plant J 17:415–425 (1999).

Rostoks N, Park YJ, Ramakrishna W, Ma J, Druka A, Shiloff BA, SanMiguel PJ, Jiang Z, Brueggeman R, Sandhu D, Gill K, Bennetzen JL, Kleinhofs A: Genomic sequencing reveals gene content, genomic organization, and recombination relationships in barley. Funct Integr Genomics 2:51–59 (2002).

Rous P: A sarcoma of the fowl transmissible by an agent separable from the tumor cells. J Exp Med 13:397–411 (1911).

Sandhu D, Gill KS: Gene-containing regions of wheat and other grass genomes. Plant Physiol 128:803–811 (2002).

Sasaki T, Matsumoto T, Yamamoto K, Sakata K, Baba T, et al: The genome sequence and structure of rice chromosome 1. Nature 420:312–316 (2002).

Schmuths H, Meister A, Horres R, Bachmann K: Genome size variation among accessions of *Arabidopsis thaliana.* Ann Bot 93:317–321 (2004).

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

Shepherd NS, Schwarz-Sommer Z, Blumberg vel Spalve J, Gupta M, Weinand U, Saedler H: Similarity of the *Cin1* repetitive family of *Zea mays* to eukaryotic transposable elements. Nature 307: 185–187 (1984).

Shinnick TM, Lerner RA, Sutcliffe JG: Nucleotide sequence of Moloney murine leukaemia virus. Nature 293:543–548 (1981).

Shirasu K, Schulman AH, Lahaye T, Schulze-Lefert P: A contiguous 66 kb barley DNA sequence provides evidence for reversible genome expansion. Genome Res 10:908–915 (2000).

Smilde WD, Haluskova J, Sasaki T, Graner A: New evidence for the synteny of rice chromosome 1 and barley chromosome 3H from rice expressed sequence tags. Genome 44:361–367 (2001).

Springer PS, Edwards KJ, Bennetzen JL: DNA class organization on maize Adh1 yeast artificial chromosomes. Proc Natl Acad Sci USA 91:863–867 (1994).

Suoniemi A, Narvanto A, Schulman AH: The *BARE*-1 retrotransposon is transcribed in barley from an LTR promoter active in transient assays. Plant Mol Biol 31:295–306 (1996a).

Suoniemi A, Anamthawat-Jónsson K, Arna T, Schulman AH: Retrotransposon *BARE*-1 is a major, dispersed component of the barley (*Hordeum vulgare* L.) genome. Plant Mol Biol 30:1321–1329 (1996b).

Suoniemi A, Schmidt D, Schulman AH: *BARE*-1 insertion site preferences and evolutionary conservation of RNA and cDNA processing sites. Genetica 100:219–230 (1997).

Suoniemi A, Tanskanen J, Pentikäinen O, Johnson MS, Schulman AH: The core domain of retrotransposon integrase in *Hordeum*: Predicted structure and evolution. Mol Biol Evol 15:1135–1144 (1998a).

Suoniemi A, Tanskanen J, Schulman AH: Gypsy-like retrotransposons are widespread in the plant kingdom. Plant J 13:699–705 (1998b).

Sutliff TD, Lanahan MB, Ho T-HD: Gibberellin treatment stimulates nuclear factor binding to the gibberellin response complex in a barley α-amylase promoter. Plant Cell 5:1681–1692 (1993).

Szafranski K, Dingermann T, Glockner G, Winckler T: Template jumping by a LINE reverse transcriptase has created a SINE-like 5S rRNA retropseudogene in *Dictyostelium*. Mol Genet Genomics 271:98–102 (2004).

Tatout C, Warwick S, Lenoir A, Deragon J-M: SINE insertions as clade markers for wild crucifer species. Mol Biol Evol 16:1614–1621 (1999).

Temin HM: Nature of the provirus of Rous sarcoma. Natl Cancer Inst Monogr 17:557–570 (1964).

Temin HM, Mizutani S: RNA-directed DNA polymerase in virons of Rous sarcoma virus. Nature 226:1211–1213 (1970).

Tikhonov AP, Lavie L, Tatout C, Bennetzen JL, Avramova Z, Deragon J-M: Target sites for SINE integration in *Brassica* genomes display nuclear matrix binding activity. Chromosome Res 9:325–337 (2001).

Turner BG, Summers MF: Structural biology of HIV. J Mol Biol 285:1–32 (1999).

Vershinin AV, Druka A, Alkhimova AG, Kleinhofs A, Heslop-Harrison JS: LINE and *gypsy*-like retrotransposons in *Hordeum* species. Plant Mol Biol 49:1–14 (2002).

Vicient CM, Schulman AH: *Copia*-like retrotransposons in the rice genome: few and assorted. Genome Lett 1:35–47 (2002).

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

Vicient CM, Kalendar R, Schulman AH: Envelope-containing retrovirus-like elements are widespread, transcribed and spliced, and insertionally polymorphic in plants. Genome Res 11:2041–2049 (2001a).

Vicient CM, Jääskeläinen M, Kalendar R, Schulman AH: Active retrotransposons are a common feature of grass genomes. Plant Physiol 125:1283–1292 (2001b).

Voytas DF, Ausubel FM: A copia-like transposable element family in *Arabidopsis thaliana.* Nature 336:342–344 (1988).

Watkins WS, Rogers AR, Ostler CT, Wooding S, Bamshad MJ, Brassington AM, Carroll ML, Nguyen SV, Walker JA, Prasad BV, et al: Genetic variation among world populations: inferences from 100 Alu insertion polymorphisms. Genome Res 13:1607–1618 (2003).

Waugh R, McLean K, Flavell AJ, Pearce SR, Kumar A, Thomas WTB, Powell W: Genetic distribution of *BARE*-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). Mol Gen Genet 253:687–694 (1997).

Wei F, Wing RA, Wise RP: Genome dynamics and evolution of the *Mla* (powdery mildew) resistance locus in barley. Plant Cell 14:1903–1917 (2002).

Wicker T, Stein N, Albar L, Feuillet C, Schlagenhauf E, Keller B: Analysis of a contiguous 211 kb sequence in diploid wheat *(Triticum monococcum)* reveals multiple mechanisms of genome evolution. Plant J 26:307–316 (2001).

Wicker T, Yahiaoui N, Guyot R, Schlagenhauf E, Liu ZD, Dubcovsky J, Keller B: Rapid genome divergence at orthologous low molecular weight glutenin loci of the A and Am genomes of wheat. Plant Cell 15:1186–1197 (2003).

Witte CP, Le QH, Bureau T, Kumar A: Terminal-repeat retrotransposons in miniature (TRIM) are involved in restructuring plant genomes. Proc Natl Acad Sci USA 98:13778–13783 (2001).

Wright DA, Voytas DF: *Athila4* of *Arabidopsis* and *Calypso* of soybean define a lineage of endogenous plant retroviruses. Genome Res 12:122–131 (2002).

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

Yu J, Hu S, Wang J, Wong GK-S, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, et al: A draft sequence of the rice genome *(Oryza sativa* L. ssp. *indica)*. Science 296:79–91 (2002).