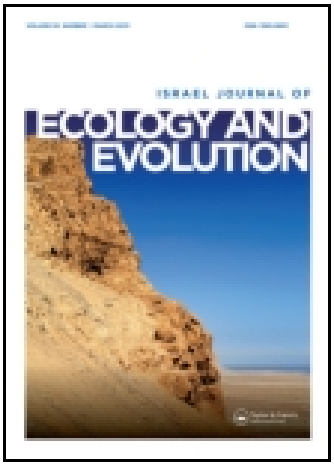

Israel Journal of Ecology & Evolution

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tiee20

Retrotransposons: Metaparasites and Agents of Genome Evolution

François Sabot [a], Ruslan Kalendar [a], Marko Jääskeläinen [a], Chang Wei [a], Jaakko Tanskanen [a b] & Alan H. Schulman [a b]

[a] MTT/BI Plant Genomics Laboratory, Insitute of Biotechnology, Viikki Biocenter, University of Helsinki, P.O. Box 56, Viikinkaari 1, FIN-00014 Helsinki, Finland

[b] Plant Genomics, Biotechnology and Food Research, MTT Agrifood Research Finland, Myllytie 10, FIN-31600 Jokioinen, Finland

Published online: 14 Mar 2013.

To cite this article: François Sabot , Ruslan Kalendar , Marko Jääskeläinen , Chang Wei , Jaakko Tanskanen & Alan H. Schulman (2006) Retrotransposons: Metaparasites and Agents of Genome Evolution, Israel Journal of Ecology & Evolution, 52:3-4, 319-330, DOI: 10.1560/IJEE_52_3-4_319

To link to this article: http://dx.doi.org/10.1560/IJEE_52_3-4_319

ISRAEL JOURNAL OF ECOLOGY & EVOLUTION, Vol. 52, 2006, pp. 319–330

RETROTRANSPOSONS: METAPARASITES AND AGENTS OF GENOME EVOLUTION

François Sabot,[a] Ruslan Kalendar,[a] Marko Jääskeläinen,[a] Chang Wei,[a]
Jaakko Tanskanen,[a,b] and Alan H. Schulman[a,b,*]
[a]*MTT/BI Plant Genomics Laboratory, Insitute of Biotechnology, Viikki Biocenter, University of Helsinki, P.O. Box 56, Viikinkaari 1, FIN-00014 Helsinki, Finland*
[b]*Plant Genomics, Biotechnology and Food Research, MTT Agrifood Research Finland, Myllytie 10, FIN-31600 Jokioinen, Finland*

ABSTRACT

Transposable elements comprise the bulk of higher plant genomes. The majority of these elements are the Class I LTR retrotransposons, which transpose via an RNA intermediate in a "Copy-and-Paste" mechanism. Because retrotransposons use cellular resources and their own enzymes to replicate independently of the genome as a whole, and have thereby become in many cases more predominant than the cellular genes, they have been considered "selfish DNA" and nuclear parasites. They are thought to share many features of the internal life cycle of retroviruses such as HIV (lentiviruses). However, whereas at least some of the retroviruses arriving in an organism during an infection must be functional in order for the infection to proceed, some LTR retrotransposon families appear to completely lack active members even though they remain mobile. Furthermore, the process of retrotransposition is inherently error-prone and mutagenic, giving rise to "pseudospecies," or clusters of imperfect copies. The non-autonomous retrotransposons are able to cis- and trans-parasitize host retrotransposons to gain mobility, much as do defective interfering particles of RNA viruses. Hence, a complex dynamic is set up, whereby the impact of retrotransposons on genomes can be under selection on the organismal level; the impact of non-autonomous retrotransposons on autonomous ones can likewise be under selection if there is selection on the autonomous elements themselves. We are exploring the retrotransposon life cycle and the causes and possible consequences of non-autonomy at each stage regarding genome evolution.

Keywords: LTR retrotransposon, genome evolution, parasitism, non-autonomy, life cycle, insertional mutagenesis

INTRODUCTION

Transposable elements, including both Class II or DNA transposons and Class I or retrotransposons, comprise the bulk of higher plant genomes, ranging from 15% of the

*Author to whom correspondence should be addressed. E-mail: alan.schulman@helsinki.fi
Accepted April 2007.

nuclear DNA in *Arabidopsis thaliana* to more than 90% in some *Liliaceae* (Vitte and Panaud, 2005). The DNA transposons are mobile DNA segments that move around the genome by a "cut and paste" mechanism requiring a transposon-encoded enzyme called transposase. The majority of transposable elements, however, are the Class I LTR retrotransposons (Fig. 1), which transpose via an RNA intermediate in a "Copy-and-Paste" mechanism (Schulman and Kalendar, 2005). Hence, most of the DNA of many genomes is derived directly from replicated RNA, a remnant of the RNA world. Because retrotransposons use cellular resources and their own enzymes to replicate independently of the genome as a whole, and have thereby become in many cases more predominant than the cellular genes, they have been considered "selfish DNA" and nuclear parasites. Their replication shares many features of the internal life cycle of retroviruses (lentiviruses) such HIV (Wilhelm and Wilhelm, 2001)

The LTR retrotransposons are called so because they are bounded by Long Terminal Repeats (Fig. 1). The LTRs contain the promoter for transcription of the element, the first step in replication, as well as other RNA processing signals (Suoniemi et al., 1997). The RNA is copied into double-stranded cDNA during the process of reverse transcription,

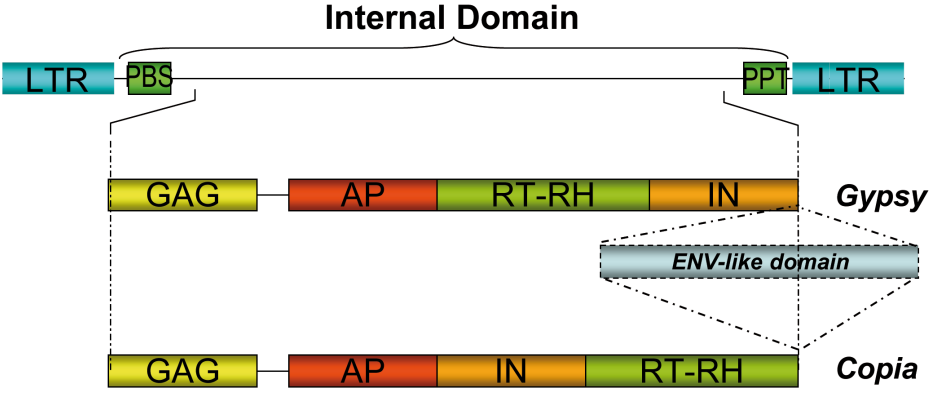

Fig. 1. Organization of the major groups of autonomous LTR retrotransposons. The LTR retrotransposons are universally bounded by long terminal repeats (LTRs). Internal to these are the (-)-strand and (+)-strand priming sites for reverse transcription (respectively, PBS and PPT). The internal domain contains the open reading frames (ORFs) for the retrotransposon-encoded proteins. Generally, the GAG and possible ENV proteins are present as distinct ORFs, whereas the other are fused into a single ORF. The order and sequence similarities of the protein-coding domains define the major clades of retrotransposons (Xiong and Eickbush, 1990). The domains are GAG, the capsid protein; AP, aspartic proteinase; RT-RNaseH, a bifunctional ORF encoding both the reverse transcriptase (RT) and RNaseH activities; and integrase (IN). See the text for a description of their roles in the retrotransposon life cycle. The major classes, *Gypsy* and *Copia*, differ in the position of the IN relative to RT-RNaseH. In addition, major groups of *Gypsy* (Vicient et al., 2001a) and *Copia* elements (Laten et al., 2005) in the plants contain a third ORF encoding a polypeptide reminiscent of a retroviral envelope protein (position of the *env* in the *env*-containing *Gypsy* and in the SIRE1 family of *Copia* elements shown by dotted lines).

which also makes the LTRs identical in the cDNA. The ends of the LTRs, as reverse-transcribed, double-stranded DNA, carry the recognition sites for integrase and are critical to the element's capacity to be inserted into the genome by this enzyme.

The LTRs enclose an internal domain that, in fully functional elements, encodes the proteins necessary for carrying out retrotransposition (Schulman and Kalendar, 2005). These proteins (Fig. 1) are present as two main Open Reading Frames (ORFs). These are: the *gag*, specifying the structural protein forming the nucleocapsid; and the *pol*, encoding the enzymatic functions. The POL is a polyprotein and is auto-processed by its AP (aspartic proteinase) domain. It also contains RT (reverse transcriptase) and RNAseH, a bifunctional polypeptide carrying out reverse transcription and IN (integrase), which inserts the new LTR retrotransposon copy into the genome (Suoniemi et al., 1998). In some elements, such as *BARE1* (Manninen and Schulman, 1993), the two ORFs are fused into one.

The LTR retrotransposons are divided into the *Gypsy* and *Copia* groups, according to whether the organization of their *pol* ORF (Fig. 1) mirrors the respective type elements of *Drosophila melanogaster*. Some *Gypsy* (Vicient et al., 2001a) and *Copia* (Laten et al., 2005) elements possess a third potential ORF or domain very similar to the *env* (envelope domain) ORF of the retroviruses (Fig. 1). The ENV product in retroviruses is involved in infectivity by mediating membrane–membrane fusion, but its function in retrotransposons is unknown. As sequence data accumulate, however, it has become clear that several structurally conserved groups of elements do not contain these canonical protein-coding domains, yet appear to propagate within the genome nevertheless. We are interested in understanding how these non-autonomous retrotransposons carry out their life cycle. Here, we discuss the question of non-autonomy from the standpoints of the retrotransposon life cycle and the impact of retrotransposons on the genome. We focus on plants because their distinctive reproductive cycle leads to special considerations regarding the evolutionary dynamics of retrotransposons.

AUTONOMY AND PARASITISM WITHIN THE GENOME

The proteins of the retrotransposon are necessary for the classical life cycle of both retrotransposons and retroviruses (Fig. 2). Once the cellular RNA polymerase transcribes the retroelement RNA and the cellular translation machinery produces the encoded proteins, the next step is the packaging of the transcript into a particle comprised of GAG. The RNA is converted to double-stranded cDNA by the reverse transcriptase, the particle and cDNA are targeted to the nucleus in some manner, and the cDNA is then transferred to the nucleus. The integrase completes the process by making a staggered cut in the genomic DNA and ligating the retrotransposon into that site, which finally undergoes gap repair by cellular mechanisms.

The precise correspondence of the retrotransposon and retrovirus life cycles has not been established for each step of replication. However, the shared major translation products, replication signals, and structural features indicate that they are likely very similar. One major difference is the existence of large, non-autonomous groups

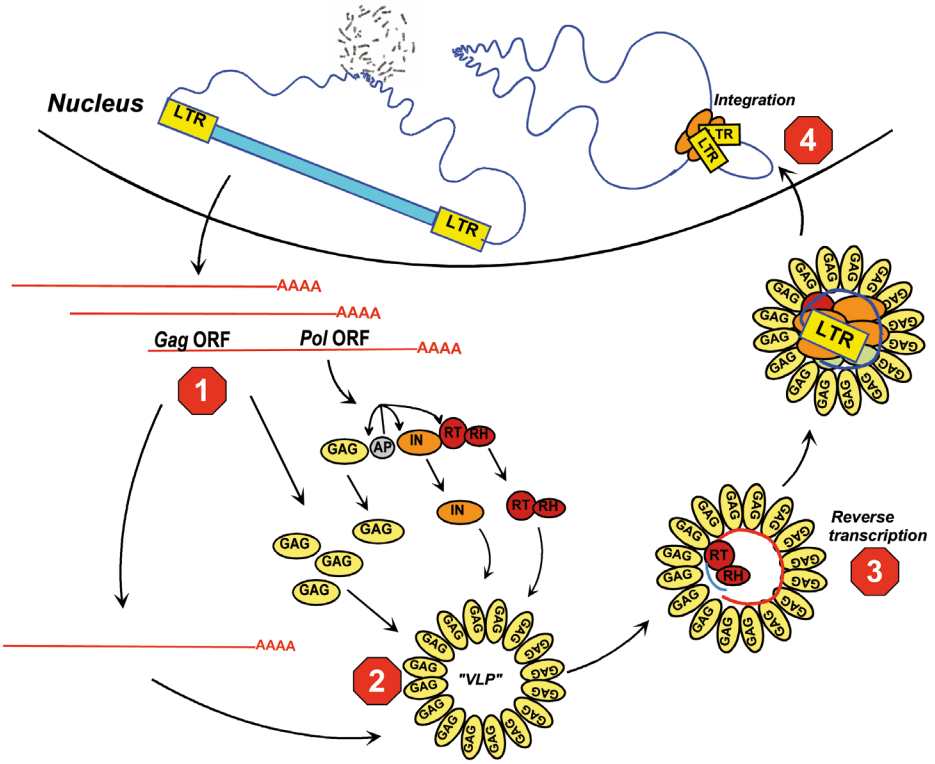

Fig. 2. Retrotransposon life cycle and its possible blocks. The proposed life cycle is presented diagrammatically. A retrotransposon integrated in genomic DNA (wavy line) is first transcribed and the transcript exported across the nuclear envelope (curved line). Transcripts (lines with AAA tails) are translated generally as two open reading frames (ORFs) producing GAG and POL products (described in the text and in Fig. 1). Non-autonomous elements partially or completely lack the ORFs (block 1; blocks shown as octagons). Packaging of the transcript requires recognition by homologous or heterologous GAG, a second potential block to replication (octagon 2). Reverse transcription requires recognition by a reverse transcriptase (block 3) following co-packaging and priming. Following localization to the nucleus, a homologous or heterologous integrase must recognize the LTR ends (block 4) to insert the cDNA copy into the genome.

of retrotransposons (Sabot and Schulman, 2006). Non-autonomy for retroelements can be defined as the lack of coding domains for the main proteins required for replication. Retrotransposons have been present as integral genomic components in eukaryotes since before the split between the evolutionary lines leading to animals, fungi, and plants. Hence, there has been ample time for the appearance of non-autonomous retrotransposon groups able to replicate as parasites on autonomous retrotransposons. We will describe below several LTR retrotransposon groups that appear to completely lack translationally competent members. Nevertheless, their polymorphic insertion pattern indicates that they are able to move within genomes.

GROUPS OF NON-AUTONOMOUS RETROTRANSPOSONS

The process of retrotransposition is inherently error-prone and mutagenic. The RNA polymerase II and reverse transcriptase each have error rates greater than that of cellular DNA polymerase. Therefore replication via transcription and reverse transcription leads to an overall mutation rate in newly inserted daughter copies of retroelements that is 10^3–10^4-fold higher than that of the mother copies that are copied by DNA polymerase II as part of the chromosome (Preston, 1996; Boutabout et al., 2001). Furthermore, chromosomal copies likely display neutral or nearly-neutral rates of decay over time due to replication errors, accumulation of indels, and large-scale insertions and recombinations. Hence, a variable fraction of a given LTR retrotransposon family or subfamily will be fully intact and functional. This gives rise to the concepts of "pseudospecies," or clusters of imperfect copies, for retroviruses and retrotransposons. For example, the *BARE1* family of retrotransposons, which is competent for transcription (Suoniemi et al., 1996; Vicient et al., 1999a), translation (Jääskeläinen et al., 1999; Vicient et al., 2001b), and virus-like particle formation (Jääskeläinen et al., 1999), and which displays high levels of insertional polymorphism (Waugh et al., 1997; Kalendar et al., 2000; Leigh et al., 2003), nevertheless contains many copies with open reading frames interrupted by stop codons (Suoniemi et al., 1998). Other active elements display the same variation (Casacuberta et al., 1997).

Individual copies may, therefore, be to varying degrees transcriptionally or translationally competent (translation leading to a functional protein) or active. Active elements may complement the life cycle blocks (Fig. 2) of inactive or incompetent members of the same family in *cis* and of retrotransposon families or groups in *trans* (Sabot and Schulman, 2006). If a translationally incompetent group of elements sequesters or competes for the protein products of an active group, it is fair to regard this phenomenon as parasitism. In a conceptually similar way, negative interfering viruses are parasitic or block replication of otherwise virulent viruses (Hu et al., 1997; Flichman et al., 1999). Individual elements can easily acquire stop codons either during replication or following integration. However, one can envisage a translationally incompetent element that is particularly successful at propagation emerging as a new subfamily and ultimately family of non-autonomous elements. This would occur if, following integration, it were able to transcribe replication-competent RNAs that successfully compete for the protein products of fully autonomous elements.

Non-autonomous retrotransposon groups, therefore, differ from the swarm of individual variants or fossils of autonomous elements. Non-autonomous elements will lack all or part of their open reading frame, and a hallmark of a successfully propagating non-autonomous group is its emergence as a population with a fairly uniform structure. An example of a non-autonomous element family that nevertheless retains most of its autonomous character is *BARE2* (Vicient and Schulman, 2005; Sabot and Schulman, 2006; Tanskanen et al., 2006), diagrammed in Fig. 3. Members of the *BARE2* family are highly similar to *BARE1* and *Wis2*, but have conserved deletion of the first start codon, which leads to loss of the capacity to translate the capsid protein GAG. A more extreme deletion derivative is the *Morgane* family (Sabot et al., 2006), illustrated in

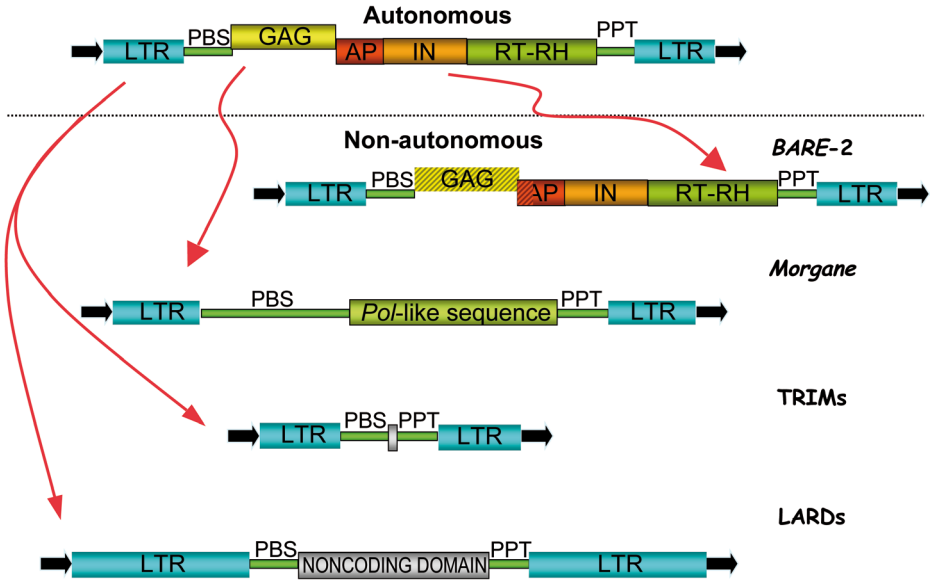

Fig. 3. Autonomous and non-autonomous retrotransposons. A representative autonomous retrotransposon is shown (*Copia*) above the major groups of non-autonomous elements so far discovered. A conserved deletion of the start codon in the *BARE2 gag* oblates all of the GAG and part of the AP as well (denoted with hatching). *Morgane* bears a DNA deletion that removed all of the *gag* and most of the *pol*, which is also highly degenerate. TRIMs and LARDs represent the non-autonomous groups where the ORFs have disappeared completely. In LARDs, the ORF has been replaced by a non-coding domain conserved in RNA secondary structure. TRIMs have lost the internal domain virtually completely, and retain only the signals for transcription and reverse transcription. Retrotransposon nomenclature is described in Fig. 1 and in the text. Arrows flanking the LTRs indicate the TSDs generated at integration. Elements are not drawn to the same relative scale.

Fig. 3. *Morgane* retains only a relict of its coding domain, and must rely on complementation by the translation products of other retrotransposons for its function. Although it appears highly likely, based on sequence and structural similarities, that *BARE2* could be complemented by *BARE1* or *Wis2* elements, the parasitic host for *Morgane* remains undefined.

Recently, large, structurally uniform retrotransposon groups in which no member contains the *Gag*, *Pol*, or *Env* internal domains have been identified (Witte et al., 2001; Kalendar et al., 2004; Antonius-Klemola et al., 2006). These groups are non-autonomous in replication, even if individual elements may be active or inactive transcriptionally. These fully non-autonomous groups have been named LARDs (LArge Retrotransposon Derivatives) and TRIMs (Terminal Repeats In Miniature). LARDs contain long LTRs and a long, conserved internal domain that shows no protein coding capacity (Kalendar et al., 2004). In contrast, TRIMs are highly reduced, with short LTRs and internal domains that contain only the signals for reverse transcription (Witte et al., 2001). No *trans*-activating (*trans*-parasitic) element yet has been demonstrated for any of these groups. Nevertheless, members of the non-autonomous groups appear to be more poly-

morphic than the autonomous elements by methods such as IRAP and REMAP that display insertion site polymorphism of retrotransposons.

HOW DOES A NON-AUTONOMOUS ELEMENT MOVE?

The presence of large homogenous groups of non-autonomous retrotransposons displaying insertion site polymorphism argues that these elements have the capacity to both replicate and integrate. However, they lack coding capacity. We are currently working to solve the puzzle of how these elements carry out their life cycle. The possible blocks to the replication of non-autonomous retrotransposons are presented in Fig. 2. Because these elements lack coding capacity, they must parasitize the protein products of autonomous elements.

Transcription is the first step of replication, and for this all retrotransposons use cellular, rather than their own, enzymes and transcription factors. The following step is packaging of the transcript into a virus-like particle (VLP) comprised of the GAG protein, which is encoded by the retrotransposon itself. Packaging is driven, at least in retroviruses, by the PSI sequence (Packaging Signal) within the transcript, which forms an RNA secondary structure that is recognized by the GAG (Evans et al., 2004). The PSI has not been well studied for retrotransposons. However, we have found (Sabot et al., unpublished) conserved sequences and structures at the PSI position in LARD elements, arguing for a common mechanism in non-autonomous elements that allows them to parasitize GAG. The *BARE2* family of retrotransposons has a specific, conserved deletion ablating its GAG (Tanskanen et al., 2006). However, the highly conserved *BARE1* and *BARE2* RNA structures make it likely that *BARE2* is packaged into *BARE1* particles (Tanskanen et al., 2007).

Following their packaging, the RNA templates (generally two per particle) are reverse transcribed. Reverse transcription of non-autonomous elements must rely on reverse transcriptase from translationally competent elements. Reverse transcription is not template-specific, as long as the RNA template is primed. Both TRIM and LARD elements contain highly conserved reverse transcriptase priming sites. (Witte et al., 2001; Kalendar et al., 2004). This is a further indicator that both groups parasitize the proteins of autonomous elements, and that both have undergone selection for maintenance of this parasitic function, even while the open reading frames have been lost. Because of the high error rate of reverse transcriptase, active retrotransposons in general must undergo purifying selection to maintain function.

The final step of the retrotransposon life cycle is insertion back into the genome, called integration, carried out by integrase, which is a protein encoded by retrotransposons. The enzymology of integration is conserved throughout life, in all Kingdoms (Rice and Baker, 2001). Integrases recognize the termini of the LTRs, make staggered cuts in genomic DNA, and simultaneously ligate the LTR end to the genomic DNA insertion site. When the staggered cuts are repaired they generate target site duplications (TSDs), which are diagnostic for the role of an integrase or transposase in the insertion of a transposable element. Hence, for integrase recognition, non-autonomous elements must either possess LTR motifs that can be recognized by the autonomous integrases with

which they are packaged or they must possess motifs that can be recognized by multiple retrotransposon groups. The generality of LTR recognition is unexplored among the plants, although it has been investigated in the yeast *Ty1* (Moore and Garfinkel, 2000). As a third alternative, non-autonomous elements may integrate with the use of integrase, as does the T-DNA of *Agrobacterium* and the DNA of bombarded plant transformation vectors. Some plant viruses as well, such as members of the geminivirus, badnavirus, and caulimovirus groups, have been shown to occasionally integrate without catalysis by integrase (Harper et al., 2002). Atypical integration sites lacking TSDs are often seen for LARDs (unpublished results), although TRIMs and *Morgane* elements appear to integrate by the canonical integrase pathway (Witte et al., 2001; Sabot and Schulman, 2006; Sabot et al., 2006).

THE SCALES OF RETROTRANSPOSON ACTION: INTER-ELEMENT, GENOMIC, ORGANISMAL, AND POPULATION

Looking beyond the mechanics of parasitism by non-autonomous elements, we must consider what selective forces and dynamics may be at work at the various levels of organization: the retrotransposon families, the genome, the organism, and the population. The observation that non-autonomous elements can be more insertionally polymorphic than are autonomous ones implies that the non-autonomous elements are in some way streamlined for replication. Whether this also means that they limit the propagation of autonomous elements, in a way reminiscent of negative interfering viruses and their parasitism on otherwise virulent viruses, remains to be seen. However, it is not difficult to imagine either scenario: selection for interference in order to limit propagation in the genome and its concomitant effects; selection or tolerance of integration of both autonomous and non-autonomous elements in order to drive genome size growth. The relative rarity of TSDs flanking LARDs suggests that LARDs may be serving to limit expansion of their host retrotransposon family. In this scenario, LARDs would counteract the propagation of their host family by out-competing them for packaging with the host GAG and enzymes, but only poorly integrating back into the genome. The integration rate would only need to be sufficient to overcome inactivation through mutation of cDNA copies and loss of genomic ones.

The *BARE2* family, as discussed above, appears to parasitize the *BARE1* autonomous elements. Our investigations show that cultivated barley and wild barley (*H. spontaneum*) both contain even more *BARE2* elements than *BARE1* elements, unlike the other species in the *Hordeum* genus that have fewer *BARE1* copies (Vicient et al., 1999b; Tanskanen et al., 2007). The generally higher copy genomic number of both *BARE* families correlates with the larger genomes of cultivated and wild barley compared to the other species (Vicient et al., 1999b), but is due mainly to an increase in *BARE2* copy number (Tanskanen et al., 2007). This suggests both a burst of *BARE1* and -2 amplification after the divergence of *H. spontaneum* from the other *Hordeum* clades and an additive rather than interfering effect from the non-autonomous family. Likewise, the MITE elements of plants, which are highly reduced and non-autonomous

DNA transposons that have in a few cases been linked to autonomous partners, are more abundant than the elements that drive their mobility (Jiang et al., 2004; Loot et al., 2006). On the cellular level, we see that the relative pool sizes of *BARE1* and -2 transcripts parallel the genomic abundance (Tanskanen et al., 2007). This suggests that *BARE2* is getting packaged and integrated equally efficiently.

In contrast to the LARDs discussed above, there is little evidence that *BARE2* is acting to limit the propagation of *BARE1*; this may be connected to the integration pattern of *BARE1* and -2. Both *BARE1* and -2 have very similar LTRs in the domain recognized by integrase and show similar integration patterns. Both integrate primarily into the repetitive fraction of the genome. Hence, *BARE* elements appear to rarely inactivate a gene through insertion. In consequence, very high copy numbers appear to be tolerated. Earlier work by Kalendar and coworkers (2000) showed that, over very small geographical distances, there is selection for gain and against loss of *BARE* retrotransposons that is correlated with a microclimate gradient. Wild barley is largely self-fertilizing, with limited gene flow; this may provide a mechanism for stratification and fixation of increased *BARE* numbers. In contrast, elements such as *Tos17* of rice (Miyao et al., 2003) and *Tnt1* of tobacco (Benlloch et al., 2006) are both mutagenic and rare. It remains to be seen if such rare, deleterious elements generally are kept in check by an interfering role of non-autonomous groups, or if cellular repression mechanisms at the transcriptional and post-transcriptional levels such as for *Tos17* (Cheng et al., 2006) do the job.

The combination of a sexual transmission mechanism and a lack of lethality allows retrotransposons not only to propagate through plant populations, but also to develop non-autonomous forms. In this way, retrotransposons differ from both plant viruses and animal retroviruses. Retrotransposons are able to move between plants via pollination as part of the genome of the sperm nucleus. A non-autonomous retrotransposon, likewise, can continue to propagate following its transfer as part of a haploid genome during fertilization. In contrast, a non-autonomous plant virus is unlikely to succeed in infecting a host.

Although the sexual transfer of retroviruses is conceptually similar to pollen flow, their presence in the seminal fluid and not in the spermatozoa (Kato et al., 2006) suggests that non-autonomous copies would be defective in infection. Hence, the interaction of autonomous and non-autonomous elements is not seen with retroviruses. Although, in principle, integrated retroviruses could give rise to non-autonomous, parasitic groups, this has been relatively rare. The lethality of some retroviruses such as HIV may hinder the evolution of integrated, but non-autonomous versions; the host dies before they can propagate. Human and other mammalian genomes do, however, contain so-called endogenous retroviruses or ERVs (Dewannieux et al., 2006), which are ancient integrated retroviruses. They may retrotranspose through parasitic complementation by either LINE or co-infecting retroviral proteins (Bannert and Kurth, 2006). Mobile ERVs are, however, the exceptions rather than the rule.

The differential success of non-autonomous retrotransposon groups in the plants may be related to the distinct developmental scheme of plants. In animals, the germ line is laid down early in development. In plants, however, the germ line differentiates

following meiosis in the floral meristem, which is itself induced following multiple, mitotic cell divisions. The clonal cell lines giving rise to the floral meristems provide an opportunity for selection to act on retrotransposons during somatic divisions; the successful ones will be passed into the germ line. Providing that the deleterious effect does not increase twice as rapidly as the copy number (because the gamete contributes half of the new genome), pollen flow will tend to lead to transfer of new retrotransposon copies within a plant population.

CONCLUSIONS

Advances in recent years have shown that retrotransposons are far from the "junk" they were initially considered (Morgante, 2005). They are a dynamic component of the genome, affecting both gene expression and overall genome structure. Autonomous retrotransposons comprise the most abundant gene families of the genome. Although retrotransposons are not junk, their ancient fossils are found throughout the genome. Until very recently, the non-autonomous retrotransposons were considered the junk of the junk. We are now establishing that they are, rather, the parasites of the parasites. It will likely prove that the dynamics of their interaction with the autonomous elements affect genes, genomes, and plants on all scales. Hence, ecological approaches apply also within the genome. The concept of genome ecology is being developed by others for DNA transposons and other biological systems, particularly insects (Le Rouzic et al., 2006). Here, we have confined ourselves mostly to the retrotransposons of plants, especially in the Triticeae cereals.

ACKNOWLEDGMENTS

Anne-Mari Narvanto and Ursula Lönnqvist are thanked for their always excellent technical assistance. F. Sabot was supported by a University of Helsinki Post-doctoral Fellowship. Work described here was carried out under a grant from the Academy of Finland, Projects 106949 and 207485.

REFERENCES

Antonius-Klemola, K., Kalendar, R., Schulman, A.H. 2006. TRIM retrotransposons occur in apple and are polymorphic between varieties but not sports. Theor. Appl. Genet. 112: 999–1008.

Bannert, N., Kurth, R. 2006. The evolutionary dynamics of human endogenous retroviral families. Annu. Rev. Genomics Hum. Genet. 7: 149–173.

Benlloch, R., d'Erfurth, I., Ferrandiz, C., Cosson, V., Beltran, J.P., Canas, L.A., Kondorosi, A., Madueno, F., Ratet, P. 2006. Isolation of mtpim proves Tnt1 a useful reverse genetics tool in *Medicago truncatula* and uncovers new aspects of AP1-like functions in legumes. Plant Physiol. 142: 972–983.

Boutabout, M., Wilhelm, M., Wilhelm, F.X. 2001. DNA synthesis fidelity by the reverse transcriptase of the yeast retrotransposon Ty1. Nucl. Acids Res. 29: 2217–2222.

Casacuberta, J.M., Verhettes, S., Audeon, C., Grandbastien, M.-A. 1997. Quasispecies in retrotransposons: a role for sequence variability in Tnt1 evolution. Genetica 100: 109–117.

Cheng, C., Daigen, M., Hirochika, H. 2006. Epigenetic regulation of the rice retrotransposon *Tos17*. Mol. Genet. Genomics 276: 378–390.

Dewannieux, M., Harper, F., Richaud, A., Letzelter, C., Ribet, D., Pierron, G., Heidmann, T. 2006. Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements. Genome Res. 16: 1548–1556.

Evans, M.J., Bacharach, E., Goff, S.P. 2004. RNA sequences in the Moloney murine leukemia virus genome bound by the Gag precursor protein in the yeast three-hybrid system. J. Virol. 78: 7677–7684.

Flichman, D., Cello, J., Castano, G., Campos, R., Sookoian, S. 1999. *In vivo* down regulation of HIV replication after hepatitis C superinfection. Medicina (B. Aires) 59: 364–366.

Harper, G., Hull, R., Lockhart, B., Olszewski, N. 2002. Viral sequences integrated into plant genomes. Annu. Rev. Phytopathol. 40: 119–136.

Hu, W.S., Bowman, E.H., Delviks, K.A., Pathak, V.K. 1997. Homologous recombination occurs in a distinct retroviral subpopulation and exhibits high negative interference. J. Virol. 71: 6028–6036.

Jääskeläinen, M., Mykkänen, A.-H., Arna, T., Vicient, C., Suoniemi, A., Kalendar, R., Savilahti, H., Schulman, A.H. 1999. Retrotransposon *BARE1*: Expression of encoded proteins and formation of virus-like particles in barley cells. Plant J. 20: 413–422.

Jiang, N., Feschotte, C., Zhang, X., Wessler, S.R. 2004. Using rice to understand the origin and amplification of miniature inverted repeat transposable elements (MITEs). Curr. Opin. Plant Biol. 7: 115–119.

Kalendar, R., Tanskanen, J., Immonen, S., Nevo, E., Schulman, A.H. 2000. Genome evolution of wild barley (*Hordeum spontaneum*) by *BARE1* retrotransposon dynamics in response to sharp microclimatic divergence. Proc. Natl. Acad. Sci. USA 97: 6603–6607.

Kalendar, R., Vicient, C.M., Peleg, O., Anamthawat-Jonsson, K., Bolshoy, A., Schulman, A.H. 2004. LARD retroelements: conserved, non-autonomous components of barley and related genomes. Genetics 166: 1437–1450.

Kato, S., Hanabusa, H., Kaneko, S., Takakuwa, K., Suzuki, M., Kuji, N., Jinno, M., Tanaka, R., Kojima, K., Iwashita, M., Yoshimura, Y., Tanaka, K. 2006. Complete removal of HIV-1 RNA and proviral DNA from semen by the swim-up method: assisted reproduction technique using spermatozoa free from HIV-1. AIDS: 967–973.

Laten, H.M., Havecker, E.R., Farmer, L.M., Voytas, D.F. 2005. SIRE1, an endogenous retrovirus family from *Glycine max*, is highly homogeneous and evolutionarily young. Mol. Biol. Evol. 20: 1222–1230.

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

Le Rouzic, A., Dupas, S., Capy, P. 2006. Genome ecosystem and transposable elements species. Gene 390: 214–220.

Loot, C., Santiago, N., Sanz, A., Casacuberta, J.M. 2006. The proteins encoded by the *pogo*-like *Lemi1* element bind the TIRs and subterminal repeated motifs of the *Arabidopsis Emigrant* MITE: consequences for the transposition mechanism of MITEs. Nucl. Acids Res. 34: 5238–5246.

Manninen, I., Schulman, A.H. 1993. *BARE1*, a *Copia*-like retroelement in barley (*Hordeum vulgare* L.). Plant Mol. Biol. 22: 829–846.

Miyao, A., Tanaka, K., Murata, K., Sawaki, H., Takeda, S., Abe, K., Shinozuka, Y., Onosato, K., Hirochika, H. 2003. 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.

Moore, S.P., Garfinkel, D.J. 2000. Correct integration of model substrates by Ty1 integrase. J. Virol. 74: 11522–11530.

Morgante, M. 2005. Plant genome organisation and diversity: the year of the junk! Curr. Opin. Biotechnol. 17: 168–173.

Preston, B.D. 1996. Error-prone retrotransposition: rime of the ancient mutators. Proc. Natl. Acad. Sci. USA 93: 7427–7431.

Rice, P.A., Baker, T.A. 2001. Comparative architecture of transposase and integrase complexes. Nat. Struct. Biol. 302–307.

Sabot, F., Schulman, A.H. 2006. Parasitism and the retrotransposon life cycle in plants: a hitchhiker's guide to the genome. Heredity 97: 381–388

Sabot, F., Sourdille, P., Chantret, N., Bernard, M. 2006. *Morgane*, a new LTR retrotransposon group, and its subfamilies in wheats. Genetica 128: 439–447.

Schulman, A.H., Kalendar, R. 2005. A movable feast: diverse retrotransposons and their contribution to barley genome dynamics. Cytogenet. Genome Res. 110: 598–605.

Suoniemi, A., Narvanto, A., Schulman, A.H. 1996. The *BARE1* retrotransposon is transcribed in barley from an LTR promoter active in transient assays. Plant Mol. Biol. 31: 295–306.

Suoniemi, A., Schmidt, D., Schulman, A.H. 1997. *BARE1* insertion site preferences and evolutionary conservation of RNA and cDNA processing sites. Genetica 100: 219–230.

Suoniemi, A., Tanskanen, J., Pentikäinen, O., Johnson, M.S., Schulman, A.H. 1998. The core domain of retrotransposon integrase in *Hordeum*: predicted structure and evolution. Mol. Biol. Evol. 15: 1135–1144.

Tanskanen, J.A., Sabot, F., Vicient, C., Schulman, A.H. 2007. Life without GAG: The *BARE2* retrotransposon as a parasite's parasite. Gene 390: 166–174.

Vicient, C.M., Kalendar, R., Anamthawat-Jonsson, K., Schulman, A.H. 1999a. Structure, functionality, and evolution of the *BARE1* retrotransposon of barley. Genetica 107: 53–63.

Vicient, C.M., Suoniemi, A., Anamthawat-Jónsson, K., Tanskanen, J., Beharav, A., Nevo, E., Schulman, A.H. 1999b. Retrotransposon *BARE1* and its role in genome evolution in the genus *Hordeum*. Plant Cell 11: 1769–1784.

Vicient, C.M., Kalendar, R., Schulman, A.H. 2001a. Envelope-containing retrovirus-like elements are widespread, transcribed and spliced, and insertionally polymorphic in plants. Genome Res. 11: 2041–2049.

Vicient, C.M., Jääskeläinen, M., Kalendar, R., Schulman, A.H. 2001b. Active retrotransposons are a common feature of grass genomes. Plant Physiol. 125: 1283–1292.

Vicient, C.M., Schulman, A.H. 2005. Variability, recombination, and mosaic evolution of the barley *BARE1* retrotransposon. Mol. Biol. Evol. 61: 275–291.

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

Waugh, R., McLean, K., Flavell, A.J., Pearce, S.R., Kumar, A., Thomas, W.T.B., Powell, W. 1997. Genetic distribution of *BARE1*-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). Mol. Gen. Genet. 253: 687–694.

Wilhelm, M., Wilhelm, F.X. 2001. Reverse transcription of retroviruses and LTR retrotransposons. Cell. Mol. Life Sci. 58: 1246–1262.

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

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