Influence of long terminal repeat retrotransposons in the genomes of fission yeasts

Mikel Zaratiegui*1

*Department of Molecular Biology and Biochemistry, Rutgers, the State University of New Jersey, Nelson A133, 604 Allison Road, Piscataway NJ 08854, U.S.A.

Abstract

LTR (long terminal repeat) RTs (retrotransposons) are almost ubiquitous in eukaryotic genomes. Their abundance and selfish properties make them a major influence in the regulation and evolution of their host genome. Recently, several striking properties of the LTR RTs of fission yeast have been uncovered, affecting important cellular processes such as gene regulation, nuclear architecture and genome integrity.

The present review summarizes the current information and puts it in the context of the wider search for understanding the influence of transposable elements on the host genome.

Introduction

TEs (transposable elements) are almost universally present in both prokaryotic and eukaryotic genomes. Class I TEs, or RTs (retrotransposons), move by way of an RNA intermediate that can be retrotranscribed into new insertable DNA copies. This stands in contrast with Class II TEs, DNA transposons, which must excise the existing insertions to move them. LTR (long terminal repeat) RTs are highly related to retroviruses, with whom they share a retrotranscription mechanism. They are characterized by two direct repeats, the LTRs that flank the coding sequences and contain the promoter, polyadenylation signal and other sequences necessary for replication and insertion. Lacking functional env genes, LTR RTs are locked in repeating cycles of transposition within the host genome. This amplification mechanism has led to their accumulation in many eukaryotic genomes [1], sometimes to dramatic proportions; 90% of the maize genome consists of LTR RTs or their mutated remnants.

From the moment of their discovery, it was recognized that LTRs are extremely recombinogenic [2]. The only mechanism counteracting LTR RT sequence accumulation is the mitotic HR (homologous recombination) between the two LTRs, which evicts the coding sequences leaving a solo LTR behind as a telltale trace of a transposition event [3]. Studies of full-length copies and solo LTRs of different TE families [4] as well as direct measurement of the rates of transposition [5] and inter-LTR recombination [6] have led to the conclusion that LTR RTs live teetering on the equilibrium between these two counteracting processes. Inter-LTR recombination can potentially take place between any LTRs in the genome. This frequently occurs between LTRs placed in close proximity, leading to deletions of the flanked DNA, and is a common source of chromosomal rearrangements [7]. The progressive accumulation of LTR RTs and solo LTRs in the host genomes makes their recombination potential a major driver of eukaryotic genome evolution.

TEs can influence their host genomes beyond the structural changes that they induce. Transposition can lead to disruption of coding and regulatory sequences, and all organisms have evolved defence mechanisms to control TE activity. Studies in plants and fungi and the germline of several metazoans have shown that the RNAi pathway generates sRNA (small RNA) derived from the primary transcriptional products of many TEs in almost all organisms where RNAi is active [8]. These sRNAs can guide post-transcriptional degradation of their targets’ mRNA as well as heterochromatin-remodelling activities that silence TE transcription. In this manner, RNAi provides a form of sequence-specific detection of...
parasitic elements. However, not all TEs are silenced by RNAi, and not all organisms with silenced transposons have active RNAi (a notable example is Saccharomyces cerevisiae), suggesting that the host organisms have evolved additional mechanisms to recognize and silence transposons. Heterochromatin silencing of TEs can sometimes influence the regulation of nearby genes [9,10], so TEs have a clear influence in the evolution of regulatory networks.

In summary, the selfish properties of LTR RTs and the reaction of the host genome to their parasitic behaviour, multiplied by their widespread presence, make them a major influence in the evolution of eukaryotic genomes. The influence of LTR RTs on genome regulation and stability has been studied extensively in yeast. The present review summarizes the recent discoveries on the regulation of LTR RTs, and their influence on transcriptional silencing, genome integrity and nuclear architecture in the fission yeasts of the genus Schizosaccharomyces (Figure 1).

The LTR RTs of fission yeasts

Four fission yeasts of the genus Schizosaccharomyces have been completely sequenced: S. pombe, S. octosporus, S. cryophilus and S. japonicus [11,12]. Like other yeasts, they exhibit clear signs of repeated invasions by several types of LTR RT. The best studied are those found in S. pombe, two closely related RTs called Tf1 and Tf2 [13–15]. Only Tf2 is present in the sequenced type genome, along with more than 200 solo LTRs of Tf2, Tf1 and other Tf-type elements. S. octosporus contains no distinguishable RTs, whereas S. cryophilus only contains one LTR RT (Tcry1). S. japonicus shows a wide variety of LTR RTs, with no fewer than ten distinct elements (named Tj1–Tj10) with signs of recent activity. Taken together, the Schizosaccharomyces LTR RTs can be grouped into two phylogenetic clades. The first one, which includes Tj2, Tj3, Tj8, Tj7, Tj8, Tj9 and Tj10 show extensive similarity to the Ty3 LTR RTs found in S. cerevisiae. The other clade includes Tf1 and Tf2, Tcry1, Tj3, Tj4 and Tj6 [12]. The two clades exhibit one major biological difference: whereas the Ty3-like elements initiate retrotranscription from a tRNA primer, Tf-like elements create a self-primer through fold back and cleavage of the 5′ end of the transcript [16].

The biology of the Tf transposons has been studied extensively. Once a comprehensive catalogue of Tf2 elements and solo LTRs was available from the genome sequence, it was observed that Tf2 RTs exhibited a strong preference for insertion into the promoters of protein-coding genes [13]. Using a Tf1 overexpression system to newly invade the S. pombe genome followed by low- and high-throughput characterization of insertion sites [17,18], demonstrated that Tf1 also inserts preferentially into promoters, with a preference for those of stress response genes. This suggests that they share a targeting mechanism. LTR RT element tropism is usually guided by protein–protein interactions between the integrase and host sequence-specific DNA-binding factors [19]. Consistent with the preference for stress-response promoters, Tf1 integrase interacts with the stress-related transcription factor Atf1, and Atf1 deletion inactivates some insertion hotspots. However, others were unaffected, and not all Tf1 insertion hotspots exhibit Atf1 binding, indicating that additional factors are involved [20,21]. The preference for stress-related genes prompted the hypothesis that Tf element mobilization could provide the means for rapid evolution of gene regulatory networks that would confer an adaptive response to stress conditions. Indeed, insertion of Tf1 into heat-activatable promoters can bring enhancer activity that activates transcription [22]. These results lend new credence to the hypothesis originally proposed by Britten and Davidson [23], which posited that TEs were the source of regulatory sequences.

Silencing of the fission yeasts’ LTR RTs

The LTR contains the promoter of the RT. Transcription from Tf1/Tf2 solo LTRs and full-length transposons is generally considered to be low under basal conditions, but can be activated by heat, hypoxia and other stresses [24]. A detailed exploration of S. pombe genome-wide transcription data focused on RT sequences suggested that they exhibited various amounts of basal transcription, with some elements having transcription as high as that of histone genes, whereas others were silent [25]. The reasons for this disparity are unknown. One possibility is that sequence variability could affect the binding of regulatory factors. For example, the LTR of Tf2-11 lacks a sterol-regulatory element that in other Tf2 copies activates transcription under hypoxic conditions through recruitment of the transcription factor Sre1 [24]. Transcription of full-length RTs is restricted to the sense strand, whereas solo LTRs display transcription in both strands [25].
Silencing of the Tf transposons implicates several mechanisms, probably partially redundant. These include chromatin-mediated mechanisms, such as the participation of CENP-B (centromere protein B) [26] and histone chaperones [27], as well as RNA-mediated mechanisms such as RNAi and the exosome [28,29].

CENP-B regulation of LTR silencing and subnuclear architecture

CENP-B is a family of DNA-binding factors that have been domesticated from the Pogo/Tigger Class II transposase. The structure of these proteins is inherited from the ancestral transposase, so they bind to DNA through their N-terminal domain and dimerize through their C-terminal domain [30]. Members of this family are present in several species, from fission yeast to humans, but appear to have arisen from independent exaptation events [31]. Human CENP-B localizes to pericentromeric α-satellite repeats, where it binds to a sequence, the 17 bp CENP-B box, that is present in centromeric satellite repeats from a wide variety of species. The function of the centromeric CENP-Bs is unclear, but it is necessary for either centromere assembly or de novo deposition of heterochromatic marks over α-satellite arrays, depending on whether there is an active centromere elsewhere in the chromosome [32].

S. pombe has three CENP-B homologues: Abp1/Cbp1, Cbh1 and Cbh2 [33,34]. These have also been implicated in pericentromeric heterochromatin assembly [35], but a study on the genome-wide localization of Abp1 and Cbh1 revealed that the elements with the strongest CENP-B enrichment were the LTRs, where they bind to specific sequences [26]. Deletion of Abp1 and Cbh1 results in a dramatic activation of Tf2 transcription, indicating that CENP-Bs silence the LTRs. The mechanism for CENP-B-mediated LTR silencing involves the recruitment by Abp1 of the histone deacetylases Clr3 and Clr6, as well as collaboration with the Set1 methyltransferase to control sense and antisense transcription [26,36].

Additionally, the dimerizing activity of CENP-B brings together solo LTRs and full-length transposon sequences into a subnuclear region called the Tf body [26]. This clustering requires Abp1-mediated recruitment of the Ku complex that in turn recruits condensin [37]. The Tf1 body is usually placed at the nuclear periphery, a subnuclear region usually associated with transcriptional repression. However, silencing and clustering processes are separate, because mutation of Ku and deletion of the Abp1 dimerization domain abrogate Tf bodies, but do not activate transcription. The Tf bodies are lost during S-phase and upon DNA damage, due to H3K56 (histone H3 Lys56) acetylation interfering with Ku recruitment. As a consequence of their clustering into Tf bodies, LTRs play a major role in the organization of chromatin in subnuclear domains, which affects important processes such as transcriptional regulation of nearby genes and DNA repair.

RNAi and the exosome

Tf2 RTs are not silenced by RNAi. Deep sequencing approaches showed little sRNA originating from LTR or Tf2 coding sequences [38–40], and genome-wide surveys of heterochromatin revealed that H3K9 (histone H3 Lys9) methylation and the HP1 (heterochromatin protein 1) homologue Swi6, typical of RNAi-dependent heterochromatin regions, were markedly absent [38]. Consistently, mutation of RNAi factors and the H3K9 methyltransferase Clr4 have very mild effects on Tf2 and LTR transcription [41]. However, the RNAi machinery does localize at LTRs, as revealed by a sensitive DamID (DNA adenine methyltransferase identification) study [28]. Intriguingly, this localization does not need the participation of the sRNA-binding effector Ago1 (Argonaute 1), arguing for a sRNA-independent function. RNAi factors localize to and transcriptionally repress euchromatic targets in association with the nuclear pore complex, regulating the transcriptional heat stress response in an sRNA-independent manner [42]. Considering the localization of Tf bodies at the nuclear periphery, the mild repression of LTR transcription by RNAi may represent an example of this process.

The Tj transposons of S. japonicus do elicit a robust RNAi response, complete with generation of sRNA and methylated H3K9 deposition [12]. In fact, it would appear that S. japonicus has co-opted clusters of Tj transposons to act as pericentromeric and mating type locus heterochromatin. In contrast with S. pombe, S. japonicus has no obvious CENP-B homologues, indicating that the two main branches of Schizosaccharomyces yeasts diverge in the mechanism they use to silence LTR RTs [12]. However, S. pombe is not entirely incapable of generating a RNAi response to the Tf2 RTs: in the absence of the exosome, the RNAi machinery produces sRNA mapping to the coding sequence of Tf2 RTs, which gains methylated H3K9 [29]. The exosome and RNAi therefore compete for Tf2 transcripts. When active, RNAi against Tf2 depends on the RNA surveillance factor Red1 and the poly(A) polymerase Pla1, indicating that Tf mRNA becomes a substrate for RNAi as part of an RNA surveillance pathway. These results provide fresh mechanistic insight into the RNAi control of RTs.

LTR influence in genome integrity

In S. pombe, CENP-B homologues are the main effectors of LTR silencing. However, the most striking phenotype of CENP-B mutants is slow growth and cell death, and the triple Δabp1/Δcbh1/Δcbh2 knockout is completely inviable [43]. Transposition activation seldom results in dramatic loss of viability, indicating that CENP-B has additional essential functions. One such function was shown to be the regulation of DNA replication at the Tf LTRs. The LTRs contain a directional RFB (replication fork barrier) located in their 5′ end that depends on the DNA-binding factor Sap1, which also performs the same function at rDNA (ribosomal DNA) to regulate the directionality of replication [44]. In the absence
of CENP-B factors, the paused fork becomes destabilized and leads to numerous double-strand breaks in all chromosomes, followed by engagement of HR. Consistently, a mutation of Sap1 that abrogates fork barrier activity rescues the viability of CENP-B triple mutants, and mutation of HR factor Rhp51 is synthetic lethal with Δabp1/Δcbh1 double mutants. These results suggest that the presence of the LTRs in the S. pombe genome constitutes a source of genome-wide replicative stress. In agreement with this conclusion, a survey of the modified histone γ-H2A, implicated in DNA damage signalling, found strong enrichment in transposon sequences [45].

The RFB activity of LTR, and the ensuing HR reaction, could explain the contribution of transposon sequences to structural variation in the eukaryotic genome. As is often the case, our knowledge of S. pombe genomic variation lags behind that of S. cerevisiae. One possible example of LTR-caused rearrangement is the mechanism of sod2 gene amplification. Upon selection in a high concentration of LiCl, survivors show amplification of the sod2 gene, which codes for an Li+ antiporter, in acentric linear episomes. A frequent episomal form consists of a palindromic duplication of ~110 kb centred on two LTRs in a head-to-head orientation [46]. The rearrangement could arise as a consequence of the arrest of two converging replication forks in both LTRs that would leave an unreplicated spacer between them. This scenario has been shown to result in inverted duplications by HR with template exchange [47].

Are RFBs a conserved feature of LTR RTs, or just an idiosyncrasy of the Tf family? The Ty LTR RTs of S. cerevisiae contribute very intensely to genomic structural variation [7], and Ty LTRs also accumulate γ-H2A as a result of replicative stress, indicating that they too form an impediment to the progression of replication [48]. The cause of Ty LTR replication stress has not been investigated, so the participation of host DNA-binding factors or cis-acting elements such as hairpin-forming regions is not known. If RFBs constitute a conserved feature of LTRs, they must play a role in the evolutionary success of RTs, perhaps by regulating inter-LTR recombination [44]. Regardless of the mechanism of action or their implication in the life cycle of the RTs, it is likely that the long-known high recombinogenic potential of LTRs is due to the obstacle they pose to DNA replication.

Concluding remarks
The LTRs must pack many regulatory functions in a short sequence. As a consequence of this, and of the peppering of LTRs on eukaryotic genomes due to repeated cycles of transposition and inter-LTR recombination, they probably constitute a major source of genomic regulatory function. Investigating RT biology will no doubt contribute to our understanding of the conserved principles behind the evolution of eukaryotic genomes, and their interaction with their constant companions, the transposons. Owing to the recent progress in our mechanistic understanding of the critical roles that the LTRs of S. pombe Tf RTs play in gene regulation, nuclear architecture and genome integrity, they are shaping up to be a very interesting model for this research.

Funding
M.Z. is supported by the Searle Scholars Program from the Kinship Foundation.

References


