CMLS, Cell. Mol. Life Sci. 54 (1998) 383–393
1420-682X/98/050383-11 $ 1.50 + 0.20/0

CMLS Cellular and Molecular Life Sciences

Review

Retrotransposable elements in the *Dictyostelium discoideum* genome

T. Winckler

Institut für Pharmazeutische Biologie, Universität Frankfurt (Biozentrum), Marie-Curie-Str. 9, D-60439 Frankfurt am Main (Germany), Fax +49 69 798 29662, e-mail: winckler@em.uni-frankfurt.de

Received 21 November 1997; received after revision 6 January 1998; accepted 6 January 1998

Abstract. Repetitive DNA is a major component of any living cell. In eukaryotes retrotransposable elements make up several percent of the genome size, and consequently, retroelements are often identified in experiments aimed at establishing physical maps and whole genome sequences. In this review, recent progress in the characterization of retrotransposable elements in the genome of the eukaryotic microorganism *Dictyostelium discoideum* is summarized with a focus on retroelements which integrate near transfer RNA genes with intriguing position specificity.

Key words. *Dictyostelium*; repetitive DNA; retrotransposon; tRNA gene; evolution; review.

Retrotransposable elements

Transposable elements (TEs) are DNA segments which are able to move and replicate within genomes. TEs were often considered 'selfish DNA' or 'molecular parasites', replicating and evolving within a host cell genome without playing positive roles in the evolution of their carriers [1, 2]. However, there is now overwhelming evidence that TE insertions have tremendous biological consequences on host genome structure and function, and it is recognized that most spontaneous mutations having phenotypic effects may be associated with TE insertions (see refs 3 and 4 and references therein).

Retrotransposable elements are mobile genetic entities which replicate through the transcription of genomic retroelement copies into RNA intermediates. These 'genomic' RNAs are reverse-transcribed into double-stranded DNA by means of element-encoded RNA-dependent DNA polymerases (reverse transcriptases, RTs) and integrated at new locations in the host genome. Apart from the general interest in understanding their retrotransposition mechanisms, retrotransposons have gained interest as useful tools in mapping eukaryotic genomes [5] and as vectors for gene therapy [6].

Retrotransposable elements are divided into several groups depending on their structural characteristics and retrotransposition mechanisms. The group of retroviruses and retrovirus-like retrotransposons has long terminal repeats (LTRs) located at both ends of the elements (fig. 1). Retroviral genomes encode several functional proteins required for retrotransposition [7, 8], including reverse transcriptase/ribonuclease H (RT/RNase H), protease and integrase (IN) proteins. Characteristically, retroviral proteins are encoded in several open reading frames (ORFs), which are translated into polyprotein precursers due to ribosomal frame shifting or stop-codon suppression. Polyproteins are cleaved into functional proteins by the retroviral protease. LTR retrotransposons are similar to retrovirsuses in their genomic organization (fig. 1) and retrotransposition mechanism. Whereas retroviruses are able to leave one cell to infect

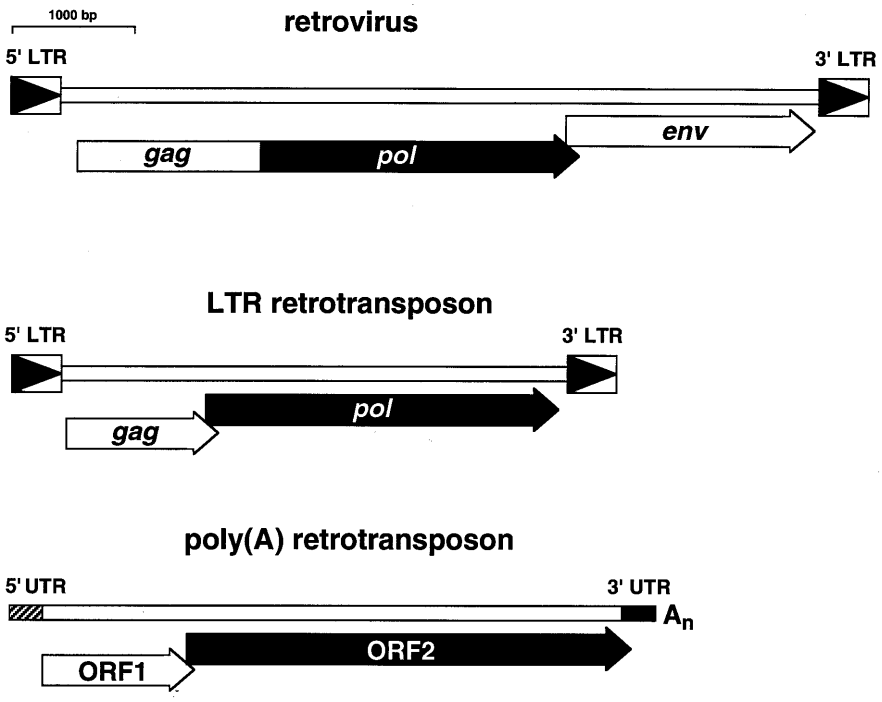

Figure 1. General structures of the three major classes of retrotransposons. Refer to the text for details.

another, LTR retrotransposons form cytosolic viruslike particles (VLPs) containing RNA [9], but these VLPs are not infectious, and retrotransposon DNAs integrate into the genome of the same cell. As a consequence, nonviral retrotransposable elements add considerable amounts of DNA (5–7 kb) to the host cell genome with each replication cycle and may comprise up to several percent of a eukaryotic genome [10, 11].

The retrotransposition mechanism of LTR retroelements is well understood (for detailed descriptions of retroviral reverse transcription and integration, see refs 7, 12, 13). Unlike all other animal viruses, retroviruses depend on the integration of mobilized protoviruses into the host cell genome for replication. Reverse transcription of plus-strand RNA is primed by host tRNAs which are also packaged into the viral particles. LTR sequences must be present at both ends of RNA transcripts for proper reverse transcription into double-stranded DNA retrotranscripts. Integration is achieved by specific processing and joining to target DNA strands of LTR ends by the retroelement-encoded IN. Integration of most LTR retrotransposons is random, although preferential integration of retroviruses into actively transcribed genes has been described [14].

Non-LTR retrotransposons are a large and diverse group of elements which encode at least one ORF with homology to retroviral RTs but lack LTRs [15]. The poly(A) retroelements are a subclass of non-LTR elements that share structural homologies, e.g. they encode one or two ORFs flanked by nonredundant untranslated regions (UTRs) and terminate with poly(A) or A-rich tails (fig. 1). The RTs of poly(A) retroelements are only distantly related to retroviral RTs and apparently lack RNase H domains [16]. In contrast to yeast mitochondrial group II introns, which are cotranscribed with adjacent genes and subsequently catalyse their own retrotransposition by means of an encoded RT protein [17, 18], poly(A) retroelements are actively transcribed from promoters located in the 5′ UTR of the elements [19–22].

The retrotransposition mechanism of poly(A) retroelements is less understood than that of retroviruses. The lack of LTRs suggests that the transposition mechanism of poly(A) retroelements differs from that of retrovirus-like elements. It has long been controversial how the reverse transcription of poly(A) elements is primed, and whether these elements encode an IN protein. Several recent observations helped to solve this puzzle. First, it was shown that the poly(A) retroelement R2Bm from *Bombyx mori* encodes a protein domain that functions as endonuclease and produces nicks in the integration target DNA which can directly serve as substrates for reverse transcription of R2Bm RNAs [23]. Then it was noted that the poly(A) retroelement L1Tc from *Try-*

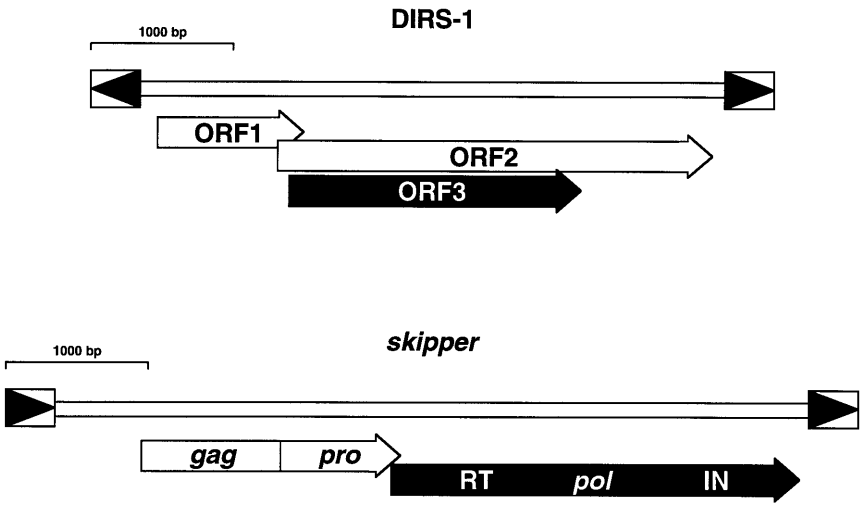

Figure 2. Structures of the *D. discoideum* LTR retroelements DIRS-1 and *skipper*.

panosoma cruzi encodes a protein domain with significant homology to apyrinic/apyrimidinic (AP) endonucleases [24]. Finally, Boeke and coworkers showed that an AP-like endonuclease domain in the human L1Hs element exhibited endonuclease functions in vitro and in vivo [25] [the AP-like endonuclease of poly(A) retroelements is now named 'EN domain']. In the current model, poly(A) retroelements replicate by a coupled target DNA-primed reverse transcription/integration process catalysed by a multifunctional EN/RT protein encoded by the retroelements. The mechanism of nick-primed integration is also used by group II introns [26, 27].

Dictyostelium discoideum

The cellular slime mold *Dictyostelium discoideum* is an amoeboid microorganism that represents an ancient branch of eukaryotic descent [28]. The growing number of known genes and gene-flanking DNA sequences have provided insights into the genome organization of *D. discoideum* cells. The *D. discoideum* genome consists of an estimated $3-4 \times 10^7$ base pairs with a very low G/C content (77% A/T average) [29]. *D. discoideum* cells feed on soil bacteria and grow vegetatively by mitosis. Depletion of nutrients induces single cells of a population to collect into multicellular organisms which undergo several morphological changes that end up in stalks which lift a ball of spores from the substratum into the air. *D. discoideum* cells differentiate into two major cell types: stalk cells and spores. In contrast to the development of higher organisms, growth and development are strictly separated in *D. discoideum*. This provides a powerful model to study developmentally regulated gene expression. The introduction of cell transformation techniques and random mutagenesis by restriction enzyme-mediated integration of plasmids (REMI) for *Dictyostelium* research [30] has allowed the detailed analysis of the *D. discoideum* developmental cycle at the molecular level. Significant progress is now being made in the identification of signal transduction pathways regulating inter- cellular signalling and cell aggregation [31–33] as well as molecular networks that regulate cell differentiation and morphogenesis in multicellular stages [34–37].

LTR retrotransposons in *D. discoideum*

DIRS-1 (Tdd-1) elements

The first retrotransposable element described in *D. discoideum* cells was independently isolated in two laboratories as an 'interspersed repeat sequence' that hybridized to several developmentally regulated mRNAs. This element was named DIRS-1 (*Dictyostelium* intermediate repeat sequence 1) [38, 39] or Tdd-1 (transposable element of *Dictyostelium discoideum*) [40]. In *D. discoideum* cells there are about 40 full-length DIRS-1 elements and more than 200 related sequences resulting from truncated DIRS-1 elements [39]. Although DIRS-1 appears to be an LTR retrotransposon, it has several unique properties. A consensus DIRS-1 element encodes three ORFs which are flanked by inverted LTRs [41] (fig. 2). The inverted LTRs of DIRS-1 elements are nonidentical, since the 3′ LTR has a 28-bp extension and several point mutations as compared with the canonical 5′ LTR sequence [42]. ORF1 and ORF2 overlap, as observed in other retroelements. In contrast with other LTR retroele-

ments, the DIRS-1 RT domain, which is most closely related to RTs of the Ty3/*gypsy* retrotransposon family [16], is encoded in a separate ORF (ORF3) that is completely contained within ORF2 (fig. 2).

The LTRs of DIRS-1 contain heat-shock promoter consensus motifs [42, 43], and it has been shown that several DIRS-1 derived transcripts are elevated in heat-shocked *D. discoideum* cells. A predominant 4.5-kb polyadenylated RNA and a heterologous population of shorter DIRS-1 RNAs are induced after starvation of *D. discoideum* cells and accumulate throughout development [39]. It is puzzling that the putative transcription orientation of the DIRS-1 LTRs is directed to the flanking regions of DIRS-1 elements [42]. However, there is evidence that the LTR promoter sequences may support bidirectional transcription [42], and the 4.5-kb RNA probably represents an LTR-to-LTR transcript. In heat-shocked cells a minus-strand RNA, 1.4 kb in length, was detected that was likely transcribed from the 3′ LTR promoter [40, 44]. However, this 'antisense' transcript does not have protein-encoding capacity, and its physiological relevance is not clear.

The DIRS-1 element lacks a tRNA binding site to prime reverse transcription, as is typically found in other LTR retroelements. Hence, the DIRS-1 reverse transcription mechanism cannot be described in terms of the characteristic retroviral mechanism. An alternative mechanism to explain the maintenance of the unique DIRS-1 structure upon retrotransposition was suggested by Lodish and coworkers [41]. The reverse transcription of the 4.5-kb RNA transcript may be initiated at the A/T-rich 3′ end of the 4.5-kb transcript by an unknown primer. Integration of DIRS-1 elements does not create target-site duplications (TSDs) and is independent of certain DNA target sequences. However, DIRS-1 elements show strong preference for integration into other DIRS-1 elements [45].

Skipper

The LTR retrotransposon *skipper* was isolated from a complementary DNA (cDNA) library of growing *D. discoideum* cells [45a]. The *skipper* element encodes protein domains known from other LTR retrotransposons including *gag* (nucleocapsid), *pro* (protease), and *pol* (RT/RNase H and integrase) (fig. 2). *Skipper* belongs to the Ty3/*gypsy* subfamily of LTR retrotransposons as deduced from the alignment of *skipper* RT with other related enzymes. The organization of ORFs within the *skipper* element is somewhat unusual in that *gag–pro* fusions may be generated by stop-codon suppression. The *pro* ORF is separated from the *pol* ORF by a putative +1 translational frame shift also observed in yeast Ty elements. *Skipper* appears to be strongly transcribed into full-length RNAs. However, most *skipper* transcripts must be assumed incompetent for transposition, since only 15–20 *skipper* copies are usually present in the *D. discoideum* genome. Like DIRS-1, *skipper* apparently lacks a tRNA primer-binding site characteristic for other retrovirus-like elements, and the exact priming mechanism for reverse transcription of *skipper* RNAs is unknown. Integration of *skipper* elements appears to be random. In one case a *skipper* element was found inserted into a DIRS-1 element, but it is unknown whether there is a preference of integration into DIRS-1 due to DNA sequence homologies among the elements.

Transfer RNA gene-associated retrotransposons in *D. discoideum*

Eukaryotic organisms contain more than 100 different tRNA gene species. tRNAs are remarkably similar in their secondary and tertiary structures [46], and the structural concept of tRNA gene transcription is conserved in all tRNA genes [47]. tRNA gene transcription by RNA polymerase III is regulated by TATA-less, tRNA gene–internal promoter elements (termed A box and B box). The B box is the primary target for sequence-specific binding of transcription factor IIIC (TFIIIC). TFIIIC bound to a tRNA gene recruits an additional factor, TFIIIB, into a ternary complex that is recognized by RNA polymerase III [48].

In *D. discoideum*, about 80 different tRNA genes have been isolated (reviewed in ref. 49). Characteristically, 90% of the analysed tRNA genes have one or two extra B-box consensus motifs located 30–40 base pairs downstream of the coding regions of the tRNA genes (exB boxes). Although binding of *D. discoideum* TFIIIC to exB motifs has been demonstrated in vitro [50], and likely also occurs in vivo, the physiological significance of exB boxes for transcription of *D. discoideum* tRNA genes is unclear. Permanent binding of regulatory factors to their binding sites near promoter regions may be of importance in reorganizing the nucleosome structure of the chromatin, allowing binding of additional factors required for high-efficiency transcription [51]. Thus, binding of TFIIIC to exB boxes may be important in regulating tRNA gene transcription in *D. discoideum* cells by organizing the chromatin structure near the tRNA genes.

DRE elements

DRE elements (*Dictyostelium* repetitive elements) were identified during the analysis of tRNA genes in *D. discoideum* when it was noticed that conserved DNA sequences occurred in the 5′ flanking regions of many tRNA genes [52, 53]. DRE elements are inserted in a

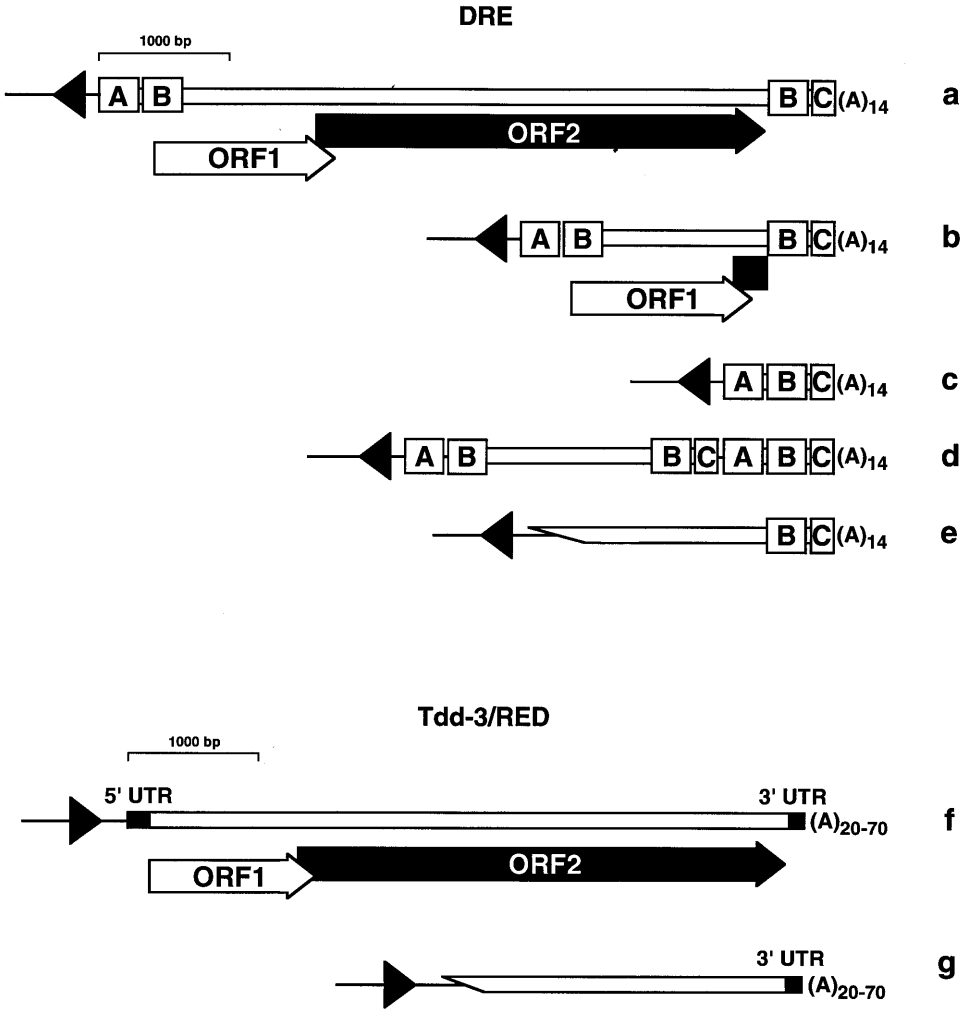

Figure 3. Structures of the *D. discoideum* poly(A) retroelements DRE, Tdd-3 and RED. Several structural variants of the elements are shown: (*a*) the consensus full-length DREa element; (*b*) the DREb element, which has a large deletion in ORF2; (*c*) recombinations among B-modules, which often result in the complete deletion of protein-encoding regions; (*d*) DRE tandems, resulting from integration of a DRE element upstream of a tRNA gene in a preexisting DRE::tDNA situation; (*e*) 5′ trunctions of DRE; (*f*) consensus full-length structures of Tdd-3 and RED elements; (*g*) 5′ deletions of Tdd-3 and RED elements.

highly position- and orientation-specific manner 50 ± 4 bp upstream of tRNA genes, producing 14 ± 2 bp target-site duplications upon integration [52]. DRE elements encode two overlapping ORFs (fig. 3) which may be translated by a -1 ribosomal frame shifting. Protein domains homologous to the AP-like endonucleases identified in human and insect poly(A) retroelements are located in ORF2. Protein sequence alignments of the DRE RT domain encoded in ORF2 with RTs of other retroelements allow one to group DRE into the class of poly(A) retrotransposons. This classification is supported by the lack of LTRs, the generation of 5′ truncated copies upon retrotransposition, and the existence of terminal oligo(A) stretches in DRE elements [54–56].

DRE elements are flanked by nonredundant terminal 'repeats' (TRs) which were named A-module and C-module (fig. 3). In addition, DRE elements contain an identical sequence motif adjacent of the A- and C-modules, termed B-module (fig. 3). DRE elements are present in several structural variants in the *D. discoideum* genome. DREa represents the full-length 5659 bp consensus sequence of the element, whereas DREb elements contain a large internal deletion of ORF2 [54]. Both types of DRE elements are flanked by TRs, including the B-modules (fig. 3). Many tRNA genes are associated with two DRE elements arranged in tandem, where the analysis of TSDs suggest that the tandems result from second-site integrations (that is, DRE elements integrating upstream of tRNA genes near the 5′

ends of the A-modules of preexisting DRE::tDNA associates). Many other structural variants of DRE were observed, many of which appear to result from recombination events among genomic copies of the various modules. As typical for poly(A) retroelements, many DRE copies show 5′ deletions. It is important to emphasize that all identified DRE copies, including the 5′-truncated ones, are flanked by 14-bp TSDs and inserted 50 bp upstream of tRNA genes.

D. discoideum strains have been divided into high-copy DRE (HCD) and low-copy DRE (LCD) strains. In HCD strains (e.g. NC4 and its axenic derivatives), about 100 copies each of the DREa and DREb elements were detected, and only a few 5′-truncated elements were found. In LCD strains very few full-length DREa elements (e.g. 3 copies in strain V12) were found, and instead large numbers of 5′-truncated elements were identified [56]. Comparison of allelic tRNA gene copies in various *D. discoideum* strains for the presence of DRE elements indicated that DRE may be transposing in laboratory and wild strains.

The A-module of DRE has promoter activity and drives the synthesis of plus-strand RNAs [57]. Plus-strand RNA transcripts start at nucleotide +1 of DRE elements and probably are generated by RNA polymerase II. The B-module has a strong silencing effect on the promoter activity of the A-module when assayed with a downstream reporter gene or when determining the steady-state level of A-module-regulated transcripts [57]. This may be due to the formation of a hairpin within the 5′ part of the B-module in the RNA transcript that may destabilize the RNA transcripts or interfere with protein translation. A similar type of downregulation of protein translation from retroelement transcripts has been observed for other retroelement species as well [58]. Alternatively, the B-module may encode a *cis*-acting DNA sequence binding to a repressor protein. No binding of nuclear proteins to either the A-module or the B-module has yet been detected in vitro in gel retardation assays (unpublished results from our laboratory).

The C-module is a highly A/T-rich DNA sequence at the end of DRE elements, which terminates with A_{14} homopolymeric stretches [55]. Plus-strand RNAs expressed from the A-module promoter terminate within the C-module about 20 or 80 bp upstream of the A_{14} homopolymeric stretch [57]. This would undoubtedly affect the integrity of the element upon retrotransposition. However, unlike other poly(A) retroelements, DRE elements produce minus-strand RNA transcripts. These transcripts start within the A_{14} homopolymeric stretch and extend to various lengths toward the 5′ end of DRE elements [57]. The C-module has promoter activity and may be responsible for regulating minus-strand transcription [57]. It was proposed that hybridization of complementary RNAs may prime reverse transcription of the element in both directions, and that RNA/DNA duplexes may be integrated into new DNA target sites [57]. Another attractive model to explain the physiological role of DRE minus-strand transcripts is that hybridization of these transcripts with plus-strand RNAs may downregulate the steady-state level of plus-strand DRE transcripts required for translation of DRE-encoded proteins and transposition. Regulation of RNA transcripts by the endogenous expression of antisense RNAs exists in *D. discoideum* in the regulation of *EB4* gene transcripts [59].

Regulation of DRE transposition through the expression of minus-strand RNAs would require regulation of C-module promoter activity. A nuclear factor (CMBF) was identified which bound to the C-module in a sequence-specific manner [60]. This factor is a good candidate to function as a regulator of DRE minus-strand

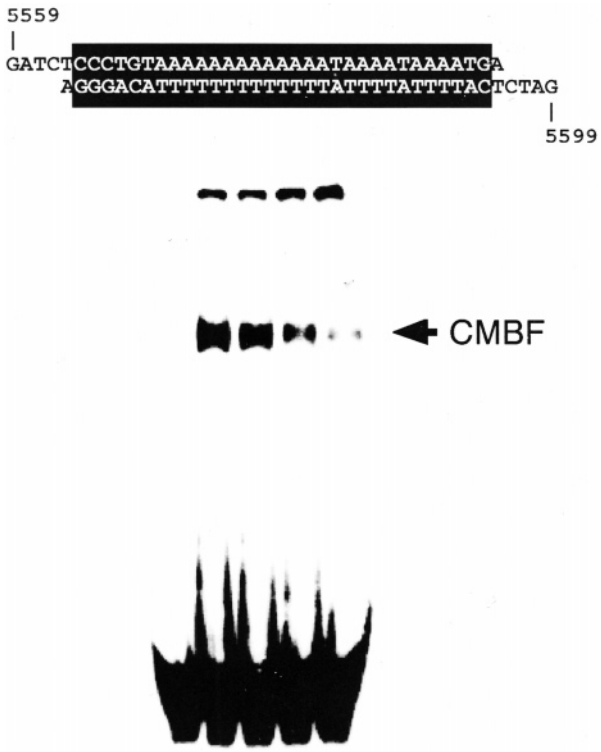

Figure 4. Interaction in vitro of the nuclear protein CMBF and an isolated binding site derived from the C-module of DREa. Two complementary oligonucleotides were annealed (upper panel) and radiolabelled. The double-stranded oligonucleotide was used in gel retardation assays using nuclear extracts of *D. discoideum* cells. The numbers in the upper panel refer to the nucleotide positions in the DREa element. Lane 1, no protein added; lanes 2–5, competition of CMBF binding (lane 2) with 0.1, 0.5 and 1 μg of plasmid containing a complete C-module as specific competitor (lanes 3–5). See ref. 60 for detailed description of the assay conditions.

expression. The most intriguing characteristic of CMBF is that it binds to two extremely A/T-rich DNA motifs within the C-module in vitro. One such 40-bp CMBF binding site was synthesized from complementary oligonucleotides. The double-stranded DNA oligonucleotide (fig. 4, upper panel) was specifically bound by purified CMBF (fig. 4) and, when cloned into a *D. discoideum actin15* promoter, generated a specific CMBF-binding site in the background of other A/T-rich DNA stretches within the *actin*15 promoter [60]. A/T-rich DNA sequences are common in *D. discoideum* RNA polymerase II promoters and often affect the transcription rates of the downstream genes [61].

Tdd-3 and RED elements

It was noted several years ago by Firtel and coworkers [62] that repetitive DNA elements were responsible for restriction-length polymorphisms and considerable genetic instability near the discoidin I gene locus of *D. discoideum* cells. The authors isolated genomic DNA fragments which contained two different transposable elements, Tdd-2 and Tdd-3. It appeared that a Tdd-3 element was inserted into a Tdd-2 element, which had previously inserted downstream of the coding region of a $tRNA^{Met}$(CUA) gene. In subsequent studies it was shown that Tdd-3 elements insert independently of Tdd-2 into genomic loci downstream of tRNA genes [52, 53, 63]. Tdd-2 and Tdd-3 are highly diverged in DNA sequence except for a 22-bp DNA motif near the 5′ end of the element that is almost identical [62, 64]. This suggested that both elements evolved from a common ancestor capable of inserting downstream of tRNA genes.

In a more recent study [65] the full-length 5.2-kb DNA sequence of Tdd-3 was obtained. Tdd-3 elements encode two overlapping ORFs flanked by nonredundant UTRs. ORF2 encodes the functional domains also found in DRE and other members of the poly(A) retroelement family, including RT and EN domains (fig. 3). Tdd-3 elements appear to be transcribed into full-length RNAs during vegetative growth of *D. discoideum* [65], a prerequisite for transposition. The promoter regulating the transcription of Tdd-3 RNAs is located within the first 360 bp at the 5′ end of the Tdd-3 element (T.W., unpublished results). About 30 copies of Tdd-3 exist in the NC4-derived axenic *D. discoideum* strains. Restriction-length polymorphisms of Tdd-3 elements in present laboratory strains show that active transposition of Tdd-3 elements occurs [62]. All Tdd-3 clones isolated so far are inserted downstream of tRNA genes. Of 14 Tdd-3 clones, 11 were inserted 80–100 bp downstream of the B boxes of the corresponding tRNA genes [65]. Three Tdd-3 elements were inserted about 130 bp downstream of a B box and 70–90 bp of an exB box, respectively. Thus it appears that B-box equivalents determine the positioning of Tdd-3 integration. Several isolated genomic DNA fragments contained Tdd-3 tandem arrays. Although only the borders of such tandems have been sequenced [63], it is plausible to assume that the 5′ Tdd-3 elements of tandems are inserted downstream of tRNA genes. Analysis of the spacing DNA sequences between Tdd-3 elements in the tandems showed no conserved sequences. Tdd-3 tandems may result from integration of a Tdd-3 element downstream of a B box of a tRNA gene locus that was already associated with another Tdd-3 element inserted downstream of an exB box of the same tRNA gene.

Computer data base analyses of the deduced protein sequence encoded in ORF2 of Tdd-3-identified RED (repetitive elements of *Dictyostelium*) elements as close relatives to Tdd-3 RED elements were first identified in genomic DNA fragments screened for Tdd-3 because RED elements frequently form RED/Tdd-3 tandems [63]. Many RED/Tdd-3 tandems contain exB motifs in the spacing sequence between the elements, suggesting that RED/Tdd-3 tandems can result from integration of Tdd-3 elements downstream of exB boxes located downstream of RED elements. A full-length RED element is very similar to Tdd-3 in terms of genomic organization and structural properties, e.g. all isolated RED clones are inserted downstream of tRNA genes (T.W., unpublished data). RED elements encode all protein domains found in Tdd-3, and their deduced protein sequences are highly homologous [65]. About 30 RED copies are identified on Southern blots probed with 3′ ends of RED, and many 5′-truncated RED copies were isolated in screenings aimed at determining the full-length RED sequence. However, all RED clones isolated so far have multiple mutations leading to frame shifts and stop codons. In reverse transcriptional polymerase chain reaction (RT-PCR) analyses no RED-specific transcripts could be detected. Thus, unlike Tdd-3 elements, it may be that RED elements are not actively transposing in present *D. discoideum* laboratory strains (unpublished results in our laboratory).

The H3R element: an unidentified tRNA gene-specific LTR retrotransposon?

The analysis of tRNA gene-flanking sequences in *D. discoideum* identified a short repeated DNA element termed H3R (Hind3 repeat), which was found inserted 10–30 bp upstream of nine isolated tRNA genes [53]. H3R elements are frequently truncated due to the insertion of DRE elements 50 bp upstream of tRNA genes [53]. The H3R element is 268 bp long and consists of an extremely A/T-rich sequence. The ends of this element, 5′-TGTAATAAAG . . . CTTTATTACA-3′ are complementary and fit into the consensus DNA sequence of

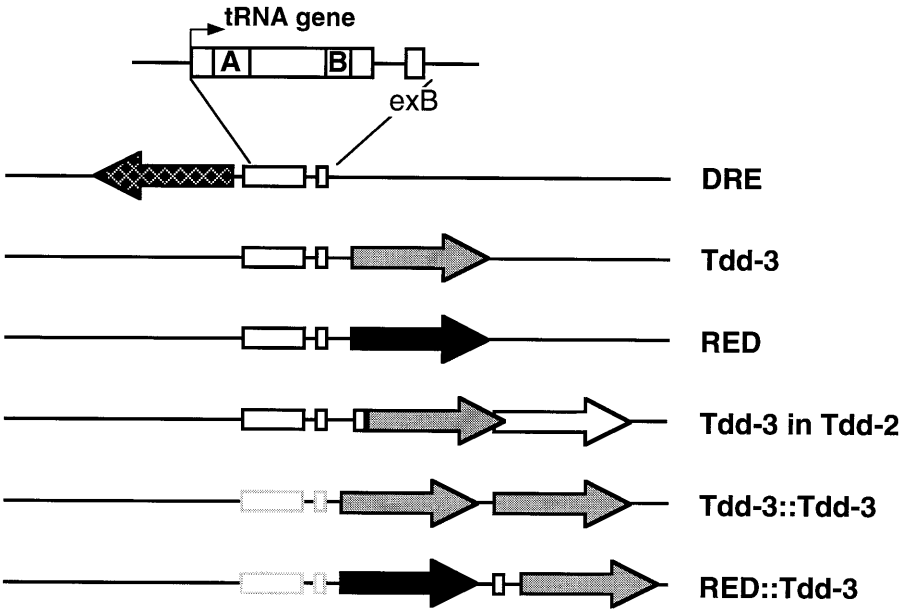

Figure 5. Summary of tDNA::retroelement associates. A tRNA gene with the gene-internal promoter elements (A-block and B-block) are shown. The position of exB boxes and the transcription orientation of the tRNA gene are indicated by the white box and small arrow, respectively. The tRNA gene-associated retroelements are shown with the arrowheads pointing in the direction of plus-strand transcription of the elements. Note that Tdd-3::Tdd-3 and RED::Tdd-3 tandems have only been sequenced at the tandem junctions between the elements. Therefore, the tRNA genes which are likely to be present 5′ of the tandems are shown in grey.

LTR ends (5′-TG . . . CA-3′) of retroviruses and LTR retrotransposons. The 5′-TG . . . CA-3′ motif is critical for processing LTRs and joining reverse-transcribed LTR retroelements to target-site DNA [13]. In the yeasts *Saccharomyces cerevisiae* and *Candida albicans* the existence of transfer RNA gene-associated solo LTRs has been reported [66, 67]. Solo LTRs most likely derive from intact LTR retrotransposons by recombination [68]. Hence, it may be considered that the H3R element is a solo LTR of a yet unknown full-length LTR retroelement that targets upstream of *D. discoideum* tRNA genes.

Possible mechanisms of tRNA gene-specific integration

As outlined in this review, several poly(A) retrotransposons of *D. discoideum* use tRNA genes as landmarks for integration independent of certain DNA sequences at the integration sites. The genomic organization of tRNA gene-associated retroelements so far known in *D. discoideum* is summarized in figure 5. An analogous targeted integration of retrotransposons near tRNA genes is found in yeast, where several LTR retrotransposons are clustered near tRNA genes and may themselves serve as targets for the integration of additional retroelement copies [69, 70]. Ty3 elements integrate 1–4 bp upstream of the transcription starts of tRNA genes. Ty1 is less specific with respect to the integration position, but most Ty1 elements are found within a range of ~400 bp of tRNA genes. The functional proteins encoded by Ty LTR retroelements are well characterized and can be used in in vitro integration assays [71]. In such assays integration of Ty3 depended on active tRNA genes, and inactivation of the B-box promoter element prevented Ty3 integration. Ty3 integration depended on the RNA polymerase III transcription factors TFIIIB and TFIIIC bound to tRNA genes [71]. The presence of RNA polymerase III inhibited Ty3 integration in vitro, suggesting that the Ty3 integration apparatus and RNA polymerase III compete for interaction sites on transcription factors bound to tRNA genes [72]. Thus, it is most likely that protein–protein interactions between the Ty3 integration complex and TFIIIB or TFIIIC determine the position specificity of Ty3 integration.

The integration specificity of DRE is comparable to that of Ty3, except that DRE integrates 50 bp upstream of tRNA genes. To date no experimental data are available that describe DRE integration, mostly

due to the less-characterized DRE-encoded proteins and the lack of an in vitro integration system. However, it is assumed that DRE integrase interacts with RNA polymerase III transcription factors, presumably TFIIIB. Analysis of several Tdd-3 and RED 5′-flanking regions allows one to speculate that B-box equivalents act in *cis* to determine the insertion position of these elements. As already suggested for DRE integration, it is assumed that protein–protein interactions of a component of the RNA polymerase III complex (most likely TFIIIC) and a Tdd/RED-specific integration factor may be responsible for guiding the retroelement DNA to the target integration sites.

Conclusions

The replication of 'selfish' retroelements within a host cell genome may cause two types of problems. First, transposition of mobilized retroelements results in the addition of significant amounts of DNA to the host cell genome. An increased genome size is a major disadvantage especially in rapidly dividing cells such as microorganisms, since it is an energetic burden and requires more time to double the genome upon cell division. This disadvantage may be compensated if we assume that the activity of retroelements enhances the genome flexibility and phenotypic fitness of their hosts. On the other hand, nonspecifically integrating retroelements may cause insertion mutations eventually leading to decreased evolutionary fitness or cell death. It is intriguing that nature apparently invented tRNA gene-specific integration of retroelements several times. The integration of retroelements upstream of tRNA genes known from yeast and *Dictyostelium* are analogous, but may not be directly linked in evolution, since Ty3 and DRE elements represent different classes of retroelements and insert at different distances to the tRNA genes (and probably use different sites of interaction with RNA polymerase III transcription factors). The integration specificity of Tdd-3 and RED are different from DRE because they insert downstream of tRNA genes. Integration of retroelements near tRNA genes may circumvent the problem of insertion mutagenesis of structural genes, since the flanking regions of tRNA genes are usually devoid of coding regions, and the large number of tRNA genes within a cell present multiple landing pads for mobilized retroelements. According to the 'phenotype paradigm' [1], 'the major and perhaps only way in which a gene can ensure its own perpetuation is by ensuring the perpetuation of the organism it inhabits' [1]. Thus, integration of retroelements near tRNA genes may be the compromise in a coevolution process allowing retroelements to maintain a viable population without being deleterious to the host cell.

Acknowledgements. I am grateful to G. Schumann for making available a preprint describing the *skipper* element. I thank T. Dingermann for his substantial support and J. B. Bell for comments on the manuscript. The work in our laboratory is supported by grants from the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie.

1 Doolittle R. and Sapienza C. (1980) Selfish genes, the phenotypic paradigm and genome evolution. Nature **284:** 601–603
2 Orgel L. and Crick F. (1980) Selfish DNA: the ultimate parasite. Nature **284:** 604–607
3 McDonald J. (1993) Evolution and consequences of transposable elements. Curr. Opin. Genet. Dev. **3:** 855–864
4 Kidwell M. and Lisch D. (1997) Transposable elements as sources of variation in animals and plants. Proc. Natl. Acad. Sci. USA **94:** 7704–7711
5 Devine S., Chissoe S., Eby Y., Wilson R. and Boeke J. (1997) A transposon-based strategy for sequencing repetitive DNA in eukaryotic genomes. Genome Res. **7:** 551–563
6 Hodgson C. (1996). Retro-vectors for Human Gene Therapy, R. G. Landes Company, Austin, Texas
7 Varmus H. and Brown P. (1989). Retroviruses. In: Mobile DNA, pp. 53–108, Howe M. M. and Berg D. E. (eds), American Society for Microbiology, Washington, DC
8 Katz R. and Skalka A. (1994) The retroviral enzymes. Ann. Rev. Biochem. **63:** 133–173
9 Sandmeyer S. (1992) Yeast retrotransposons. Curr. Opin. Genet. Dev. **2:** 705–711
10 Smit A. F. A. (1996) The origin of interspersed repeats in the human genome. Curr. Opin. Genet. Dev. **6:** 743–748
11 Boeke J. D. (1997) LINEs and Alus – the polyA connection. Nat. Genet. **16:** 6–7
12 Boeke J. and Corces V. (1989) Transcription and reverse transcription of retrotransposons. Ann. Rev. Microbiol. **43:** 403–434
13 Boeke J. and Chapman K. (1991) Retrotransposition mechanisms. Curr. Opin. Cell Biol. **3:** 502–507
14 Pryciak P. and Varmus H. (1992) Nucleosomes, DNA-binding proteins and DNA sequence modulate retroviral integration target site selection. Cell **69:** 769–780
15 Eickbush T. (1992) Transposing without ends: the non-LTR retrotransposable elements. New Biol. **4:** 430–440
16 Xiong Y. and Eickbush T. (1990) Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J. **9:** 3353–3362
17 Michel F. and Ferat J. (1995) Structure and activities of group II introns. Ann. Rev. Biochem. **64:** 435–461
18 Moran J., Zimmerly S., Eskes R., Kennell J., Lambowitz A., Butow R. et al. (1995) Mobile group II introns of yeast mitochondrial DNA are novel site-specific retroelements. Mol. Cell. Biol. **15:** 2828–2838
19 Mizrokhi L., Georgieva S. and Ilyin Y. (1988) *jockey*, a mobile *Drosophila* element similar to mammalian LINEs, is transcribed from the internal promoter by RNA polymerase II. Cell **54:** 685–691
20 Swergold G. (1990) Identification, charaterization, and cell specificity of a human LINE-1 promoter. Mol. Cel. Biol. **10:** 6718–6729
21 Minchiotti G. and Di Nocera P. (1991) Convergent transcription initiates from oppositely oriented promoters within the 5′ end regions of *Drosophila melanogaster* F elements. Mol. Cell. Biol. **11:** 5171–5180
22 McLean C., Bucheton A. and Finnegan D. (1993) The 5′ untranslated region of the I factor, a long interspersed nuclear element-like retrotransposon of *Drosophila melanogaster*, contains an internal promoter and sequences that regulate expression. Mol. Cell. Biol. **13:** 1042–1050
23 Luan D., Korman M., Jakubczak J. and Eickbush T. (1993) Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. Cell **72:** 595–605

24 Martín F., Marañón C., Olivares M., Alonso C. and López C. (1995) Characterization of a non-long terminal repeat retrotransposon cDNA (L1Tc) from *Trypanosoma cruzi*: homology of the first ORF with the Ape family of DNA repair enzymes. J. Mol. Biol. **247:** 49–59
25 Feng Q. H., Moran J. V., Kazazian H. H. and Boeke J. D. (1996) Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell **87:** 905–916
26 Zimmerly S., Guo H., Perlman P. and Lambowitz A. (1995) Group II intron mobility occurs by target DNA-primed reverse transcription. Cell **82:** 545–554
27 Zimmerly S., Guo H., Eskes R., Yang L., Perlman P. and Lambowitz A. (1995) A group II intron RNA is a catalytic component of a DNA endonuclease involved in intron mobility. Cell **83:** 529–538
28 Loomis W. F. and Smith D. W. (1995) Consensus phylogeny of *Dictyostelium*. Experientia **51:** 1110–1115
29 Marx K. A., Hess S. T. and Blake R. D. (1993) Characteristics of the large (dA)(dT) homopolymer tracts in *D. discoideum* gene flanking and intron sequences. J. Biomol. Struct. Dyn. **11:** 57–66
30 Kuspa A., Dingermann T. and Nellen W. (1995) Analysis of gene function in *Dictyostelium*. Experientia **51:** 1116–1123
31 Reymond C. D., Schaap P., Veron M. and Williams J. G. (1995) Dual role of cAMP during *Dictyostelium* development. Experientia **51:** 1166–1174
32 Van Haastert P. J. M. (1995) Transduction of the chemotactic cAMP signal across the plasma membrane of Dictyostelium cells. Experientia **51:** 1144–1154
33 Chen M. Y., Insall R. H. and Devreotes P. N. (1996) Signaling through chemoattractant receptors in *Dictyostelium*. Trends Genet. **12:** 52–57
34 Gross J. D. (1994) Developmental decisions in *Dictyostelium discoideum*. Microbiol. Rev. **58:** 330–351
35 Firtel R. A. (1995) Integration of signaling information in controlling cell fate decisions in *Dictyostelium*. Genes Devel. **9:** 1427–1444
36 Loomis W. F. (1996) Genetic networks that regulate development in *Dictyostelium* cells. Microbiol. Rev. **60:** 135
37 Schaap P., Tang Y. H. and Othmer H. G. (1996) A model for pattern formation in *Dictyostelium discoideum*. Differentiation **60:** 1–16
38 Zuker C. and Lodish H. F. (1981) Repetitive DNA sequences cotranscribed with developmentally regulated *Dictyostelium* mRNAs. Proc. Natl. Acad. Sci. USA **78:** 5386–5390
39 Zuker C., Cappello J., Chisholm R. L. and Lodish H. F. (1983) A repetitive *Dictyostelium* gene family that is induced during differentition and by heat shock. Cell **34:** 997–1005
40 Rosen E., Sivertsen A. and Firtel R. A. (1983) An unusual transposon encoding heat shock inducible and developmentally regulated transcripts in *Dictyostelium*. Cell **35:** 243–251
41 Cappello J., Handelsman K. and Lodish H. F. (1985) Sequence of *Dictyostelium* DIRS-**1:** an apparent retrotransposon with inverted terminal repeats and an internal circle junction sequence. Cell **43:** 105–115
42 Zuker C., Cappello J., Lodish H. F., George P. and Chung S. (1984) *Dictyostelium* transposable element DIRS-1 has 350-base-pair inverted terminal repeats that contain a heat shock promoter. Proc. Natl. Acad. Sci. USA **81:** 2660–2664
43 Cappello J., Zuker C. and Lodish H. F. (1984) Repetitive *Dictyostelium* heat-shock promoter functions in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **4:** 591–598
44 Cohen S. M., Cappello J., Zuker C. and Lodish H. F. (1984) Transcription of DIRS-1, an unusual *Dictyostelium* transposable element. In: Molecular Biology of Development, pp. 491–508, Davidson E. H. and Firtel R. A. (eds), A. R. Liss, New York
45 Cappello J., Cohen S. M. and Lodish H. F. (1984) *Dictyostelium* transposable element DIRS-1 preferentially inserts into DIRS-1 sequences. Mol. Cell. Biol. **4:** 2207–2213
45a Leng P., Klatte D. H., Schumann G., Boeke J. and Steck T. L. (1998) Identification of *skipper*, a novel retrotransposon in *Dictyostelium*. Nucleic Acids Res., in press
46 Kim S. (1976) Three-dimensional structure of transfer RNA. In: Progress in Nucleic Acids Research and Molecular Biology, pp. 181–216, Cohn W. (ed.), Academic Press, New York
47 Geiduschek E. and Tocchini-Valentini G. (1988) Transcription by polymerase III. Ann. Rev. Biochem. **57:** 873–914
48 Willis I. (1993) RNA polymerase III: genes, factors and transcriptional specificity. Eur. J. Biochem. **212:** 1–11
49 Marschalek R. and Dingermann T. (1991) Structure, organization and function of transfer RNA genes from the cellular slime mold *Dictyostelium discoideum*. Adv. Gene Technol. **2:** 103–143
50 Bukenberger M., Dingermann T., Meissner W., Seifart K. H. and Winckler T. (1994) Isolation of transcription factor IIIC from *Dictyostelium discoideum*. Eur. J. Biochem. **220:** 839–846
51 Pazin M., Bhargava P., Geiduschek E. and Kadonaga J. (1997) Nucleosome mobility and the maintenance of nucleosome positioning. Science **276:** 809–812
52 Marschalek R., Brechner T., Amon-Böhm E. and Dingermann T. (1989) Transfer RNA genes: landmarks for integration of mobile genetic elements in *Dictyostelium discoideum*. Science **244:** 1493–1496
53 Hofmann J., Schumann G., Borschet G., Gosseringer R., Bach M., Bertling W. M. et al. (1991) Transfer RNA genes from *Dictyostelium discoideum* are frequently associated with repetitive elements and contain consensus boxes in their 5′-flanking and 3′-flanking regions. J. Mol. Biol. **222:** 537–552
54 Marschalek R., Hofmann J., Schumann G. and Dingermann T. (1992) Two distinct subforms of the retrotransposable DRE element in NC4 strains of *Dictyostelium discoideum*. Nucleic Acids Res. **20:** 6247–6252
55 Marschalek R., Hofmann J., Schumann G., Gösseringer R. and Dingermann T. (1992) Structure of DRE, a retrotransposable element which integrates with position specificity upstream of *Dictyostelium discoideum* tRNA genes. Mol. Cell. Biol. **12:** 229–239
56 Marschalek R., Hofmann J., Schumann G., Bach M. and Dingermann T. (1993) Different organization of the transfer RNA-gene-associated repetitive element, DRE, in NC4-derived strains and in other wild-type *Dictyostelium discoideum* strains. Eur. J. Biochem. **217:** 627–631
57 Schumann G., Zündorf I., Hofmann J., Marschalek R. and Dingermann T. (1994) Internally located and oppositely oriented polymerase II promoters direct convergent transcription of a LINE-like retroelement, the *Dictyostelium* Repetitive Element, from *Dictyostelium discoideum*. Mol. Cell. Biol. **14:** 3074–3084
58 McMillan J. and Singer M. (1993) Translation of the human LINE-1 element, L1Hs. Proc. Natl. Acad. Sci. USA **90:** 11533–11537
59 Hildebrandt M. and Nellen W. (1992) Differential antisense transcription from the *Dictyostelium* EB4 gene locus – implications on antisense-mediated regulation of messenger RNA stability. Cell **69:** 197–204
60 Geier A., Horn J., Dingermann T. and Winckler T. (1996) Nuclear protein factor binds specifically to the 3′-regulatory module of the long-interspersed-nuclear-element-like *Dictyostelium* repetitive element. Eur. J. Biochem. **241:** 70–76
61 Hori R. and Firtel R. A. (1994) Identification and characterization of multiple A/T-rich *cis*-acting elements that control expression from *Dictyostelium* actin promoters: the *Dictyostelium* actin upstream activating sequence confers growth phase expression and has enhancer-like properties. Nucleic Acids Res. **22:** 5099–5111
62 Poole S. J. and Firtel R. A. (1984) Genomic instability and mobile genetic elements in regions surrounding two discoidin I genes of *Dictyostelium discoideum*. Mol. Cell. Biol. **4:** 671–680
63 Marschalek R., Borschet G. and Dingermann T. (1990) Genomic organization of the transposable element Tdd-3 from *Dictyostelium discoideum*. Nucleic Acids Res. **18:** 5751–5757
64 Firtel R. (1989). Mobile genetic elements in the cellular slime mold Dictyostelium discoideum. In: Mobile DNA, pp. 557–566, Howe M. M. and Berg D. E. (eds), American Society for Microbiology, Washington, DC

65 Winckler T., Tschepke C., de Hostos E., Jendretzke A. and Dingermann T. (1997) Tdd-3, a transfer RNA gene-associated poly(A) retrotransposon from *Dictyostelium discoideum*. Mol. Gen. Genet., in press
66 Brodeur G., Sandmeyer S. and Olson M. (1983) Consistent association between *sigma* elements and tRNA genes in yeast. Proc. Natl. Acad. Sci. USA **80:** 3292–3296
67 Perreau V., Santos M. and Tuite M. (1997) *beta*, a novel repetitive DNA element associated with tRNA genes in the pathogenic yeast *Candida albicans*. Mol. Microbiol. **25:** 229–236
68 Roeder G. and Fink G. (1980) DNA rearrangements associated with a transposable element in yeast. Cell **21:** 239–249
69 Voytas D. and Boeke J. (1993) Yeast retrotransposons and tRNAs. Trends in Genet. **9:** 421–427
70 Voytas D. (1996) Retroelements in genome organization. Science **274:** 737–738
71 Kirchner J., Conolly C. and Sandmeyer S. (1995) Requirement of RNA polymerase III transcription factors for in vitro position-specific integration of a retroviruslike element. Science **267:** 1488–1491
72 Connolly C. and Sandmeyer S. (1997) RNA polymerase III interferes with Ty3 integration. FEBS Lett. **405:** 305–311