Gene, 86 (1990) 269–274
Elsevier

GENE 03376

Two subfamilies of murine retrotransposon ETn sequences

(Recombinant DNA; repetitive DNA; immunoglobulin switch regions; nucleotide sequences)

Briton E. Shell*, John T. Collins, Laura A. Elenich, Paul F. Szurek and Wesley A. Dunnick**

Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, MI 48109-0620 (U.S.A.)

Received by A.M. Skalka: 31 July 1989
Revised: 6 October 1989
Accepted: 10 October 1989

SUMMARY

Early transposon (ETn) elements are 5.7-kb retrotransposons found in the murine genome. We have sequenced large portions of two ETn elements that have apparently transposed within the DNA of a murine myeloma cell line, P3.26Bu4. One of the transposed ETn elements has 5′ and 3′ long terminal repeats (LTRs) that are exact duplicates of each other and has a 6-bp target site duplication. These results suggest that this element, which inserted into an immunoglobulin γ1 switch region, moved by a retrotransposition process. Our nucleotide sequences confirm that individual ETn elements are very similar to one another and lack open reading frames. However, the ETn sequences reported here and those previously described differ significantly near their 5′ LTRs, including 200 bp of weak similarity and 240 bp of complete disparity. Southern hybridization analysis suggests that both subfamilies of ETn sequences are represented many times in the mouse genome. The possibility that the disparate sequences have a role in transposition by ETn elements is discussed.

INTRODUCTION

Retrotransposons are genetic elements usually characterized by the presence of long terminal repeats (LTRs) of 300–500 bp at their ends, by moderately high copy number in genomic DNA, and by evidence of movement within their respective 'host' genomes (Finnegan, 1985). At least one retrotransposon, yeast Ty, is known to transpose through an RNA intermediate that is packaged into a retrovirus-like particle (Boeke et al., 1985; Garfinkel et al., 1985). Early transposon (ETn) is one of a number of murine elements thought to be retrotransposons. Cloned ETn elements are 5.6 kb in length and are flanked by direct repeats of 330 bp (Brulet et al., 1983; Sonigo et al., 1987). Within the BALB/c genome the ETn family has about 200 members, at least some of which are transcribed in early embryogenesis (Brulet et al., 1983; 1985). We have shown that ETn transcripts are also present in tumors of immunoglobin-producing cells, and that ETn elements are capable of being inserted at new loci within the genome of one such cell (Shell et al., 1987).

Sonigo et al. (1987) have used both Southern hybridization and nt sequence data to suggest that ETn is a very homogeneous family of elements. Notably, two ETn elements cloned at random from BALB/c mouse genomic DNA by these investigators are nearly identical in sequence and contain no significant ORFs. We have reported the

Correspondence to: Dr. W.A. Dunnick, Dept. of Microbiology and Immunology, 1301 E. Catherine, Ann Arbor, MI 48109-0620 (U.S.A.) Tel. (313)763-3195.

* Present address: Department of Biology, Millsaps College, Jackson, MS 39202 (U.S.A.) Tel. (601)354-5201.

** Present address: Department of Experimental Carcinogenesis, University of Texas, Science Park at Smithville, P.O. Box 389, Smithville, TX 78957 (U.S.A.) Tel. (512)237-2403.

Abbreviations: bp, base pair(s); ETn, early transposon; kb, kilobase(s) or 1000 bp; LTR, long terminal repeat; nt, nucleotide(s); ORF, open reading frame; SSC, standard saline citrate (0.15 M NaCl/0.015 M Na_3·citrate pH 7.6).

0378-1119/90/$03.50 © 1990 Elsevier Science Publishers B.V. (Biomedical Division)

interruption by two ETn elements of immunoglobulin heavy chain switch regions in the BALB/c myeloma P3.26Bu4 (Shell et al., 1987). To learn whether these apparently mobile elements differ structurally from the randomly selected ETn elements described by Sonigo and coworkers, we determined the major portion of their nt sequences. We report here a significant difference between the two pairs of sequences, and demonstrate that the BALB/c genome includes numerous representatives of both kinds of ETn elements.

EXPERIMENTAL AND DISCUSSION

(a) Nucleotide sequences of ETn elements

We had previously isolated a molecular clone, γM85, that represented the 5′ end of an insertion of ETn sequences into the immunoglobulin γ1 switch region of the cell line P3.26Bu4 (Shell et al., 1987). To determine if this insertion had the characteristics of a retrotransposition event, we isolated a second molecular clone, γM22, that represents the 3′ end of this disruption of the γ1 switch region (Fig. 1). To eliminate the possibility of artifactual rearrangements during the propagation of γM22, we compared the size of the insert in this molecular clone to the size of the corresponding fragment in genomic DNA. The 2.2-kb insert in γM22 and the corresponding *Eco*RI fragment in P3.26Bu4 DNA comigrate (Fig. 1B, lanes 1 and 2), suggesting that γM22 is representative of the genomic DNA. We determined the nt sequence of all of the ETn portion of γM22 and of adjacent γ1 switch sequences. Sequence comparison of γM85, γM22, and a germline ETn element (MG1) sequenced by Sonigo et al. (1987) demonstrates that together γM85 and γM22 represent insertion of an entire ETn element into the γ1 switch region (data not shown). γM85 includes flanking γ1 switch sequences, the 5′ LTR from the ETn element, and about 4500 additional bp internal to the ETn element. γM22 includes an additional 300 bp of internal ETn sequences, the 3′ LTR and flanking γ1 switch sequences. The 3′ LTR present in γM22 and the 5′ LTR present in γM85 are identical over their 330 bp lengths (Fig. 2). Six bp of the target γ1 switch sequence (Mowatt and Dunnick, 1968) are duplicated (underlined in Fig. 2) at the site of ETn insertion, and no γ1 switch sequences are missing.

We had previously isolated a molecular clone, γM72, that represented the 5′ end of the insertion of ETn sequences near a γ2a gene in P3.26Bu4 (Shell et al., 1987). Because the loss of an *Eco*RI site during the derivation of this molecular clone complicated its analysis, we have recloned from P3.26Bu4 genomic DNA the 7.2-kb *Eco*RI fragment representing this insertion. The new clone, γM73, accurately represents the genomic fragment in P3.26Bu4 DNA

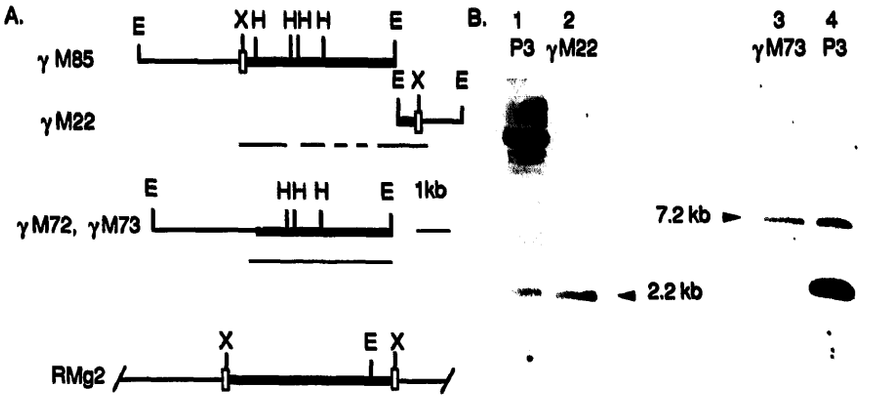

Fig. 1. ETn molecular clones and genomic sequences. A: Structure of molecular clones. ETn sequences are symbolized by thick lines, with LTRs shown as open boxes; other DNA sequences are noted by thinner lines. E (*Eco*RI), X (*Xba*I), and H (*Hin*dIII). *Hin*dIII sites were not mapped in RMg2. The extent of nt sequences determined is noted by thin lines beneath each clone. Panels B: Comparison of genomic DNA and molecular clones by Southern hybridization. P3.26Bu4 (abbreviated as P3), γM22, and γM73 DNA were digested with *Eco*RI. Samples (20 μg of genomic DNA and 200 pg of cloned DNAs) were fractionated on 0.8% agarose gels and blotted onto nitrocellulose, and the resulting filters were hybridized with the γ1 switch probe, pγ1/EH10.0 (lanes 1–2) or the γ2a switch probe, pSγ2a-1 (lanes 3–4). The large (11-kb) fragment in lane 1 is another version of the expressed γ1 switch region (Shell et al., 1987); the smaller fragment in lane 4 is the germline Sγ2a region. Lanes 1–2 and 3–4 are from independent experiments.

The derivations of the molecular clones γM85 and γM72 clones have been described (Shell et al., 1987). To isolate γM22, 10–20 kb fragments from a partial *Eco*RI digest of P3.26Bu4 DNA were ligated into λCh35 (Loenen and Blattner, 1983). γM22 was selected as a clone that hybridized to the γ1 switch probe pγ1/EH10.0 (Mowatt and Dunnick, 1986). To isolate γM73, 10–17 kb *Bam*HI fragments of P3.26Bu4 DNA were digested with *Eco*RI and cloned into *Eco*RI-digested λCh35. γM73 was selected by hybridization to a γ2a switch region probe, pSγ2a-1 (a gift from K.B. Marcu); the insert in γM73 is a 7.2-kb *Eco*RI fragment that is nearly identical to the insert in γM72. RMg2 is a molecular clone selected from the *Sau*3A partial digestion library constructed from BALB/c DNA by Lui et al. (1980). The probe used in this selection was a 2.8 kb *Hin*dIII-*Eco*RI fragment from the 3′ end of γM85.

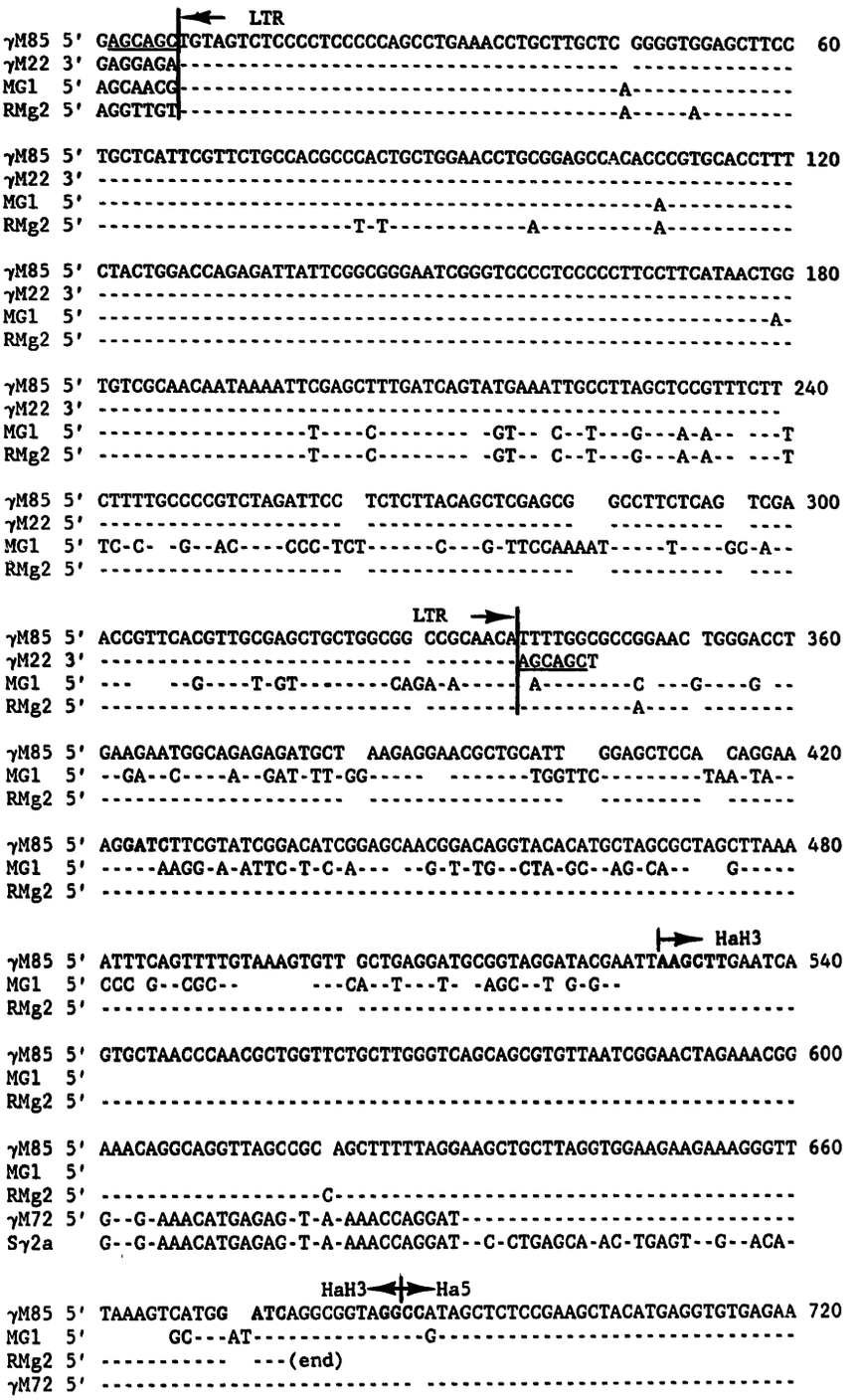

Fig. 2. Partial nt sequences of various ETn elements. Fragments derived by cleavage of γM85, γM72, γM22, γM73 or RMg2 were subcloned into M13mp18 or M13mp19 and sequenced by the dideoxy chain termination method (Sanger et al., 1980). Sequences have been deposited with the EMBL database; γM85 5′ LTR, ETn sequences in γM72, ETn sequences in γM22 and RMg2 5′ LTR (accession number X15598). The sequences of γM22, γM72, RMg2 (this study) and MG1 (Sonigo et al., 1987) are compared to the sequence of the LTR and ETn flanking sequences in γM85. Dashes indicate identity to the γM85 sequence, whereas substitutions are indicated as the appropriate nt abbreviation. Each of the sequences is contiguous; gaps are introduced in the sequences to maximize similarity. Notation (end) specifies the extent of determined sequences of RMg2. The nt 601–660 are compared to the part of the γ2a gene where ETn insertion occurred (determined by sequencing a portion of pSγ2a-1). The sequence of γM73 was determined and is identical to that of γM72 at nt 601–720. A few restriction sites are noted by boldface type. The location of two probes (HaH3 and Ha5) used in Fig. 3 are noted. The 6-bp target sequence duplication is underlined.

(Fig. 1B, lanes 3 and 4), and nt sequences at the junction of the γ2a gene and the ETn element in both γM72 and γM73 are identical. These sequences reveal that the entire 5′ LTR and about 300 additional nt of ETn sequence are missing from the inserted ETn element (Fig. 2, nt 601–660). It is impossible to determine whether these sequences (and perhaps γ2a sequences as well) were lost before or after the insertion of ETn near the γ2a gene.

More than 95% of the ETn sequences present in γM72 and 70% of the ETn sequences present in γM85 were determined. Both of these clones end near *Eco*RI sites within their ETn portions, and therefore each is missing 700 bp of 3′ ETn sequences. The sequences of these two clones confirm that they represent movement of distinct members of the ETn family; the two sequences mismatch at 4.5% of their nt. The major structural difference between our cloned ETn elements and those described previously is near their 5′ ends. The 5′ LTRs of γM85 and a previously described ETn element (MG1; Brulet et al., 1983; Sonigo et al., 1987) are nearly identical in sequence in their 5′ portions. Similarity is less apparent in the 3′ part of the 5′ LTRs, and becomes insignificant 100 bp 3′ of the end of 5′ LTRs (Fig. 2, nt residue 425). The next 90 bp of MG1 sequence show complete disparity with the next 240 bp of γM85 sequence (Fig. 2, nt residues 426–666). An ETn element cloned from germline DNA (RMg2) is nearly identical in sequence throughout its 5′ portion to the γM85 element, and therefore differs from the MG1 element in this region of disparity. Because the ETn sequences in γM72 begin 630 nt downstream of those in γM85, only a small portion of the region of disparity is available for comparison. However, in the short portion of this region that it retains, γM72 is identical in sequence to γM85 and RMg2 (Fig. 2, nt 632–666). Beyond the region of disparity (which ends at nt 674 in Fig. 2), and extending 4.2 kb to the 3′ ends of γM85 and γM72, our cloned ETn elements are very similar to MG1 (γM85 and MG1 are identical at 94.5% of compared nt; γM72 and MG1 at 92.5%; data not shown). By determining all of the ETn sequences present in γM22, we established that this similarity extends to the 3′ end of the ETn element interrupting Sγ1 (not shown).

(b) Representation of γM85-like ETn sequences in BALB/c DNA

We used two γM85 subfragments as probes for ETn sequences in BALB/c DNA. One is a 150-bp *Hae*III fragment (Ha5) which is shared by our molecular clones and those of Sonigo et al., (1987). The second probe is a 140-bp *Hin*dIII-*Hae*III fragment (HaH3), which is found in γM85, but not in the MG1 clone of Sonigo and his colleagues. The common probe (Ha5) should detect a 2150-bp *Sst*I fragment representing ETn elements similar to MG1,

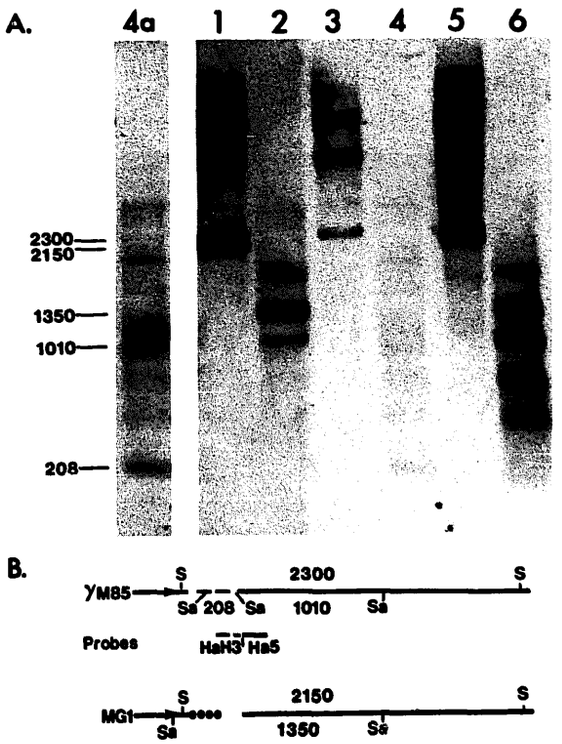

Fig. 3. Subfamilies of ETn sequences in BALB/c genomic DNA. Panel A: BALB/c liver DNA was cut with either *Sst*I (lanes 1, 3, and 5) or *Sau*3A (lanes 2, 4, 4a and 6) and processed for Southern hybridization analysis (Southern, 1975; Shell et al., 1987). Probes used (see map B below and Fig. 2) were Ha5 (lanes 1 and 2), HaH3 (lanes 3, 4, and 4a) and RMg2 (whole ETn in Figs. 1 and 2; lanes 5 and 6). The sizes of certain restriction fragments are noted. Lane 4a represents a fivefold longer exposure of lane 4. B: Restriction maps, with selected restriction sites and fragment sizes, are shown. S, *Sst*I; Sa, *Sau*3A. Similar portions of the molecular clones γM85 and MG1 are noted by solid lines. The region of sequence disparity is represented by a dashed line in the γM85 map and a dotted line in the MG1 map. The gap in the MG1 map represents the 150-bp deletion in MG1 relative to γM85 (the 'gaps' in the dashed part of the γM85 map are not meant to indicate deletions). Locations of Southern hybridization probes are noted.

and a 2300 bp *Sst*I fragment representing ETn elements similar to γM85 (Fig. 3B). The same probe should detect a 1350-bp *Sau*3A fragment for MG1-like elements and a 1010-bp *Sau*3A fragment for γM85-like elements. The existence of these fragments in BALB/c genomic DNA is confirmed in the Southern hybridization experiment shown in Fig. 3A (lanes 1 and 2). As expected, these fragments are among those detected by a probe representing an entire ETn element (RMg2, lanes 5 and 6). Of the two *Sst*I fragments described above, only the larger (2300 bp), which represents γM85-like elements, should be detectable with the γM85-specific probe HaH3. The same probe should detect a 208-bp *Sau*3A fragment which the common probe should not detect. The existence of these fragments is also confirmed in the Southern blot shown in Fig. 3A (lanes 3, 4 and 4a). These results suggest that many of the ETn ele-

ments in the BALB/c genome share the sequences that distinguish γM85 from MG1. Densitometric scanning of *Sau*3A bands detected by the common probe (Ha5; Fig. 3A, lane 2) indicates that perhaps 40 of the 200 BALB/c ETn elements are members of a γM85-like subfamily (data not shown). Given the faint hybridization of Ha5 and HaH3 to other bands in *Sst*I and *Sau*3A digests, there could be additional subfamilies of ETn elements in BALB/c which are distinct from both the γM85-like and the MG1-like subfamilies.

(c) Transposition of ETn elements in P3.26Bu4 DNA

In an earlier paper we reported that two ETn elements have been inserted into the immunoglobulin heavy-chain-encoding locus in P3.26Bu4 cells (Shell et al., 1987). This report includes further evidence that one of the insertions resulted from retrotransposition. The ETn element interrupting the Sγ1 region in P3.26Bu4 (represented by γM85 and γM22) is flanked by a duplication of a 6-bp target site, as are other ETn elements described by Brulet and his colleagues (Kaghad et al., 1985). In addition, the LTRs of the γM85/γM22 element are identical in sequence (Fig. 2). Both of these features are necessary consequences of the mechanism of cDNA synthesis and insertion proposed for retroviral replication (Varmus, 1982) and, thus, for retrotransposition (Boeke et al., 1985; Garfinkel et al., 1985).

Previously sequenced ETn elements, selected at random from BALB/c genomic libraries, contain no significant ORFs (Sonigo et al., 1987). The nt sequences of mobile ETn elements reported here, with the exception of a limited region of disparity (Fig. 2), are very similar to known ETn sequences, and also contain no significant ORFs. In fact, given the structural homogeneity among ETn elements observed in both Southern blots and molecular clones, it is possible that no member of the BALB/c ETn family is able to encode proteins (Sonigo et al., 1987; this paper). Without sequencing every one of the 200 elements in the murine genome, it is impossible to formally rule out that all ETn elements are unable to encode proteins. Therefore, ETn retrotransposons might be able to catalyze their own transposition. On the other hand, transposable ETn elements, if they indeed exist, might depend on other retrotransposon families to provide *trans*-acting transposition functions. A number of other families of murine retrotransposon-like elements have been described, any of which might have members capable of acting as helpers for ETn transposition. One particularly good candidate is the IAP (for intercisternal A-type particles) family, members of which are both transcribed and transposed in plasmacytomas (Kuff and Lueders, 1989).

Many families of retrotransposons, including IAP (Shen-Ong and Cole, 1982) and Ty (Roeder and Fink, 1983), are comprised of structurally distinct subfamilies. There are at least two subfamilies of BALB/c ETn elements, whose members differ in sequences adjacent to their 5′ LTRs (Fig. 2). Both the γM85-like and MG1-like subfamilies are represented by many individual members in BALB/c DNA (Fig. 3). It is noteworthy that to date only the γM85-like ETn subfamily contains members that have been shown to insert at new loci (γM85 and γM72/γM73); perhaps members of this subfamily transpose with greater frequency than other ETn elements. If so, differences in transposition frequency may reflect a functional importance for the structural differences that define ETn subfamilies. The region of sequence disparity defining γM85-like and MG1-like ETn elements is in a location (just downstream from the 5′ LTR) similar to that of sequences known to be recognized in *cis* for retroviral RNA packaging (Watanabe and Temin, 1982; Mann et al., 1983; Sorge et al., 1984). If a packaged transcript is a necessary intermediate in retrotransposition (Garfinkel et al., 1983), one mechanism accounting for differences in transposition frequency among members of various ETn subfamilies might be differential efficiency of packaging of ETn genomic RNAs.

ACKNOWLEDGEMENTS

We thank Rich Jove and Michael Imperiale for their careful reading of the manuscript. This work was supported by NCI grant CA39068. B.E.S. was a predoctoral fellow of the National Science Foundation during part of these studies.

REFERENCES

Boeke, J.D., Garfinkel, D.J., Styles, C.A. and Fink, G.R.: Ty elements transpose through an RNA intermediate. Cell 40 (1985) 491–500.

Brulet, P., Kaghad, M., Xu, Y.-S., Croissant, O. and Jacob, F.: Early differential tissue expression of transposonlike repetitive DNA sequences of the mouse. Proc. Natl. Acad. Sci. USA 80 (1983) 5641–5645.

Brulet, P., Condamine, H. and Jacob, F.: Spacial distribution of transcripts of the long repeated ETn sequence during early mouse embryogenesis. Proc. Natl. Acad. Sci. USA 82 (1985) 2054–2058.

Finnegan, D.J.: Transposable elements in eukaryotes. Int. Rev. Cytol. 93 (1985) 281–326.

Garfinkel, D.J., Boeke, J.D. and Fink, G.R.: Ty element transposition: reverse transcriptase and virus-like particles. Cell 42 (1985) 507–517.

Kaghad, M., Maillet, L. and Brulet, P.: Retroviral characteristics of the long terminal repeat of murine ETn sequences. EMBO Journal 4 (1985) 2911–2915.

Kuff, E.L. and Lueders, K.K.: The intercisternal A-particle gene family: structure and functional aspects. In Klein, G. and Weinhouse, S. (Eds.), Advances in Cancer Research. Academic Press, New York, 1989, pp. 183–276.

Loenen, W.A.M. and Blattner, F.R.: Lambda Charon vectors (Ch32, 33, 34, 35) adapted for DNA cloning in recombination-deficient hosts. Gene 26 (1983) 171–179.

Lui, C.-P., Tucker, P.W., Mushinski, F. and Blattner, F.R.: Mapping of heavy chain genes for mouse immunoglobulins M and D. Science 209 (1980) 1348–1353.

Mann, R., Mulligan, R.C. and Baltimore, D.: Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. Cell 33 (1983) 153–159.

Mowatt, M.R. and Dunnick, W.A.: DNA sequence of the murine gamma-1 switch regions reveals novel structural elements. J. Immunol. 136 (1986) 2674–2683.

Roeder, G.S. and Fink, G.R.: Transposable elements in yeast. In J.A. Shapiro (Ed.), Mobile Genetic Elements. Academic Press, New York, 1983, pp. 299–328.

Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A.: Cloning in single stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. 143 (1980) 161–178.

Shell, B., Szurek, P. and Dunnick, W.: Interruption of two immunoglobulin heavy chain switch regions in a murine plasmacytoma P3.26Bu4 by insertion of retroviruslike element ETn. Mol. Cell. Biol. 7 (1987) 1364–1370.

Shen-Ong, G.L.C. and Cole, M.D.: Differing populations of intercisternal A particle genes in myeloma tumors and mouse subspecies. J. Virol. 42 (1982) 411–421.

Sonigo, P., Wain-Hobson, S., Bougeleret, L., Toillas, P., Jacob, F. and Brulet, P.: Nucleotide sequence and evolution of ETn elements. Proc. Natl. Acad. Sci. USA 84 (1987) 3768–3771.

Sorge, J., Wright, D., Erdmann, V.D. and Cutting, A.E.: Amphotropic retrovirus vector system for human gene transfer. Mol. Cell. Biol. 4 (1984) 1730–1737.

Southern, E.M.: Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98 (1975) 503–517.

Varmus, H.E.: Form and function of retroviral proviruses. Science 216 (1982) 812–820.

Watanabe, S. and Temin, H.: Encapsidation sequences for spleen necrosis virus, an avian retrovirus, are between the 5′ long terminal repeat and the start of the *gag* gene. Proc. Natl. Acad. Sci. USA 79 (1982) 5986–5990.