Volume 12 Number 5 1984 Nucleic Acids Research

'Solo' large terminal repeats (LTR) of an endogenous retrovirus-like gene family (VL30) in the mouse genome

Galit Rotman, Ahuva Itin and Eli Keshet

Department of Virology, Hebrew University, Hadassah Medical School, Jerusalem 91010, Israel

Received 12 January 1984; Accepted 6 February 1984

ABSTRACT

VL30 genetic elements constitute a murine multicopy gene family that is retrovirus-like, despite the lack of sequence homology with any known retrovirus. Over one hundred copies of VL30 units are dispersed throughout the mouse genome. We report here that the mouse genome also contains 'solo' VL30 long terminal repeats (LTRs). These are structures which contain the LTR detached from the rest of the VL30 sequences. The isolation of solo LTRs from a mouse embryonic gene library with the aid of sub-genomic VL30 probes is described. Direct DNA sequencing established that the solo LTR unit is grossly similar to a standard VL30 LTR and that the LTR is flanked by a 4-base pair duplication. The analogy to the occurance of solitary LTR units of transposable elements is discussed.

INTRODUCTION

The similarity between retrovirus proviruses and transposable elements has become apparent in the recent years (1,2). A number of properties are shared by these two groups of genetic elements. These include a basic DNA structure in which the element contains two long terminally repeated sequences (LTR), each bound by small inverted repeats. In all known cases the dinucleotides (5')TG....CA(3') define the ends of the element, and a short sequence of flanking host DNA is duplicated (3,4). Furthermore, a striking homology (>60%) was found between the LTRs of an avian retrovirus (ALV) and a copia-related element (5).

Recently, additional features have been found in eukaryotic transposons, which until now were considered unique to retroviruses. For example, the "δ" long terminal repeat of the Ty1 transposon of yeast has been shown to have a U_5-R-U_3 structure (6) (where U_3 and U_5 designate sequences unique to the 3' and 5' ends of the RNA, respectively, and R designates a short terminal repeat of the RNA). Furthermore, the presence of a sequence complementary to the 3' end of a tRNA adjacent to the 5' "δ" repeat of Ty1 suggests the use of tRNA as primer for reverse transcription (7). Even more striking

is the encapsidation of the RNA of the transposable element copia in virus-like particles that contain reverse transcriptase activity (8).

Reciprocally,there is an increasing body of evidence that demonstrates the capability of retroviruses to behave like transposons. Exogenously introduced proviruses can act as mutagens by inserting into a gene and causing its inactivation; they are also capable of excising and restoring the wild type phenotype (9). More important still, endogenous retrovirus-like elements can exert transposon-like effects. For example, insertion of an intracisternal A-particle (IAP) element onto mouse immonuglobulin light chain DNA rendered the gene functionally defective (10), whereas integration of an IAP DNA within the c-mos gene led to increased expression of the gene (11).

It is important, therefore, to define the overall repertoire of cellular retrovirus-like elements and to assess their potential to act like transposons. We are studying a murine retrovirus-like gene family known as VL30 (virus-like elements that encode 30S RNA) (12). These VL30 elements are not genetically related (i.e. have no nucleic acid homology) to any known retrovirus, but yet have retrovirus-like properties such as the capacity of the 30S RNA to be encapsidated in C-type virions (12,13), and the subsequent transmissibility of these sequences to other cell genomes via pseudovirus infection (14). We have recently shown that VL30 DNA structure strongly resembles that of proviruses and transposable elements. Specifically, VL30 elements contain long terminal repeats (15), bound by small inverted repeats (16). A small duplication of host DNA and the dinucleotide (5')TG....CA(3') at the ends of the VL30 element are also present (16). Furthermore, at the inner boundary of the 5' LTR, a tRNA binding site is found, which probably serves as the primer for the reverse transcription of the 30S RNA (16).

Here we extend the list of tranposon-like features of VL30 elements by the identification of 'solo' VL30-LTRs in the mouse genome (i.e. single LTR units that are detached from the rest of the element). This 'solo' VL30-LTR is reminiscent of the 'solo'-δ found in yeast (17). This structure is presumably the result of homologous recombination between the direct repeats and subsequent loss of the sequences enclosed between them.

MATERIALS AND METHODS

Recombinant DNA clones

Individual copies of VL30 elements were cloned from a gene library of the inbred mouse strain Balb/C in bacteriophage λ Charon 4A (constructed by ligating partially digested EcoRI fragments to the Charon 4A arms) as previously

described (18). The mouse genomic inserts containing VL30 sequences were subsequently transferred to the bacterial plasmid pBR322 and restriction enzyme cleavage maps were obtained. The locations of the LTRs were determined by heteroduplex analysis (15). Specific VL30 probes were obtained by subcloning fragments from defined regions of a mapped VL30 clone.

Restriction enzyme analysis and blot-hybridization

Restriction endonucleases were purchased from New England Bioloabs. Generally, the digestion conditions recommended by the supplier were used. Mapping of restriction enzyme recognition sites was carried out by the use of different combinations of restriction enzymes and size analysis of resultant fragments. DNA fragments were electrophoresed through 0.7%-1.2% horizontal agarose slab gels and visualized by ethidium bromide staining. DNA fragments were blotted from agarose gels onto nitrocellulose filter paper essentially as described by Southern (19). The filters were subsequently baked for 2 hrs at 80°C. Filters were prehybridized for 3 hrs at 65°C in 6XSSC containing 0.02% polyvinylpyrrolidone (PVP), 0.02% Ficoll, and 50 µg/ml of heat-denatured sonicated salmon sperm DNA. Hybridization was for about 20 hrs at 65°C in the same solution, with the addition of 0.1% SDS and 0.5-1.0x10^6 cpm/ml of the indicated hybridization probe. After hybridization filters were washed for 1 hr at 50°C in 0.1XSSC containing 0.1% SDS. Autoradiography was carried out at -70°C with Agfa Curix RP2 X-ray films using enhancing screens.

Hybridization probes

Labelling of cloned DNA by nick-translation was carried out by the procedure of Roop et.al. (20), to specific activities of 1x10^8 to 5x10^8 cpm/µg of DNA. P^{32}-labeled complementary DNA (cDNA) was prepared by the method of Taylor et.al. (21). Calf thymus oligodeoxynucleotides were used as primers in order to synthesize cDNA in which all the RNA sequences are uniformally represented.

DNA sequencing

DNA fragments were end-labeled by digesting the DNA with a combination of restriction enzymes that allowed the incorporation of a certain (P^{32})-dNTP exclusively at one end (by the use of the large fragment of DNA PolI). DNA sequencing was carried out essentially according to the method of Maxam and Gilbert.

RESULTS

Isolation of VL30 solo-LTR units

The general strategy for isolating solo-LTR units was to search for

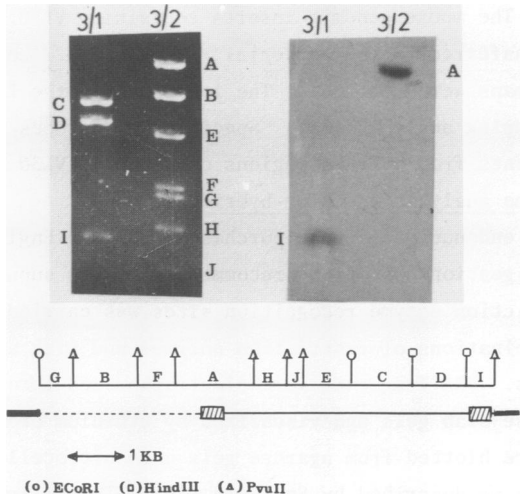

Figure 1. Identification of VL30 sequences present in clone VLS-1. A schematic representation of a clone that contains a complete VL30 unit, and at least some flanking DNA sequences at both sides, is shown (clone 3 in ref. 15, from which most flanking sequences on the right side have been removed). Indicated in the figure are the locations of the VL30 unique sequences (solid line), large terminal repeats (hatched boxes), flanking cellular sequences (broken line) and the junctions with the vector (heavy line). VL30 DNA sequences residing to the right and to the left of the single internal EcoRI site were subcloned in plasmid pBR322 (subclones 3/1 and 3/2 respectively). In order to distribute VL30 sequences at electrophoretically separable fragments, clone 3/1 was multiply digested with EcoRI, HindIII and PvuII and clone 3/2 with EcoRI and PvuII (the relevant restriction sites are indicated in the figure). Fragments were electrophoresed through 1.2% agarose gel. A photograph of the ethidium bromide stained gel is shown (left). (For illustrative purpose, the larger vector DNA fragments were cut out). DNA fragments were blotted, hybridized with P^{32}-labeled VLS-1 DNA and autoradiographed (right). The only fragments that hybridized: fragment A (1.4Kb) and fragment I (0.5Kb) - are the LTR-containing fragments.

genomic clones that hybridize exclusively with VL30-LTR sequences and fail to hybridize with any internal-VL30 sequences.

Samples of a Balb/C mouse genomic library, comprising close to a genome equivalent, were plated and plaque imprints on two duplicate nitrocellulose filters were prepared. One nitrocellulose filter was hybridized with a 'VL30-internal' probe (which contains sequences located at the center of the VL30 element), while the duplicate filter was hybridized with a 'VL30-LTR' probe (which contains only VL30-LTR DNA sequences). Numerous overlapping positive signals were detected by both probes. These clones are probably complete VL30 genomes. The LTR probe, however, detected additional signals

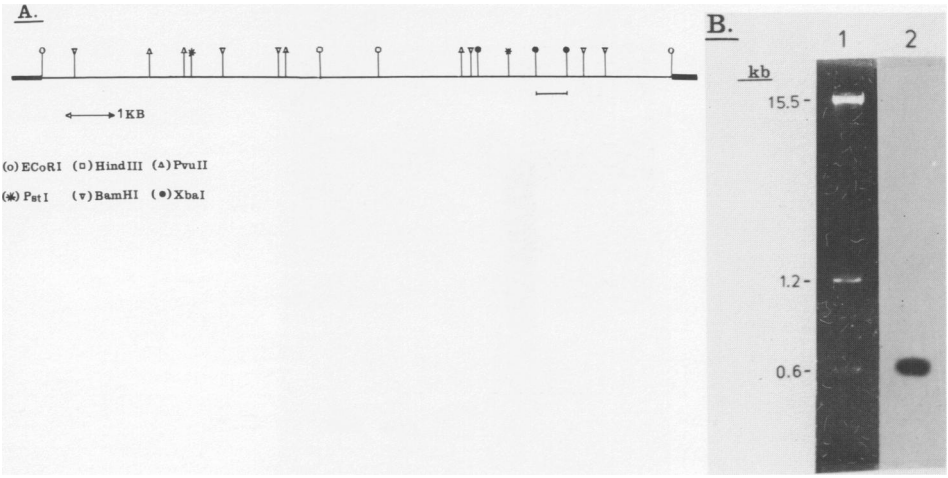

Figure 2. Localization of VL30 LTR sequences within clone VLS-1 DNA.
A. To facilitate restriction enzyme mapping of clone VLS-1, the single EcoRI insert fragment was transferred from the λ vector to the EcoRI site of plasmid pBR322. Recognition sites of the indicated restriction enzymes were mapped by standard methods.
B. Electrophoretically resolved DNA fragments produced in mapping experiments were blotted and hybridized with an 'LTR-specific' probe. The probe used was a deleted VL30 clone generated during propagation in bacteria that contains no VL30 sequences other than LTR sequences (ref.15). The example shown is an XbaI digest of clone VLS-1 in pBR322. Three XbaI fragments were obtained (lane 1). The large 15.5Kb fragment is composed of pBR322 DNA and sequences from both sides of the VLS-1 insert. Most of the hybridization was confined to the 0.6Kb XbaI fragment. Under longer exposure additional weak hybridization to the 1.2Kb XbaI fragment was detected.

that the internal probe failed to detect, indicating that in these LTR-containing clones, at least some VL30 sequences are missing. Several such clones were plaque-purified and further analyzed. The characterization of one representative clone (designated clone VLS-1) is detailed below.

VL30 sequences, other than the LTR, are not present in VLS-1 DNA

In order to study which VL30 sequences are present in clone VLS-1, a typical full size VL30 DNA clone was digested with restriction enzymes; the resultant fragments were blotted and hybridized with P^{32}-labeled DNA of clone VLS-1. As can be seen in Figure 1, VLS-1 DNA hybridized only with the two LTR-containing fragments, indicating the absence, to any significant extent, of VL30 sequences other than the LTR in this particular genomic DNA fragment.

In a reciprocal experiment, clone VLS-1 DNA was digested and blot-hybridized with either an LTR-probe (Fig. 2) or a VL30 cDNA probe, where all VL30

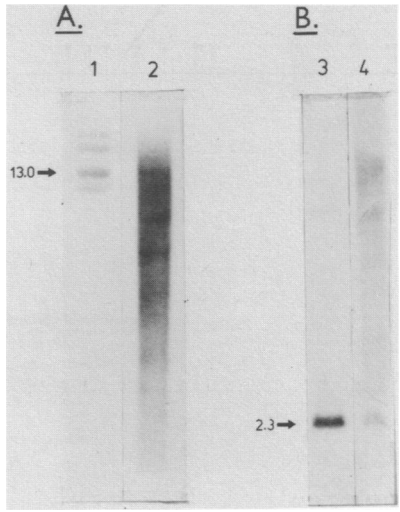

Figure 3. Detection of VLS-1 DNA containing fragments in the mouse genome.
A. 5 ng of clone VLS-1 DNA (in λ Charon 4A) (lane 1) and 15 μg DNA from Balb/C mouse embryo (lane 2) were digested with EcoRI, electrophoresed in two parallel lanes of a 0.7% agarose gel, blotted and hybridized with P^{32}-labeled DNA of clone VLS-1 (in the same λ vector). The hybridization bands in lane 1 are the Charon 4A 'arms' (18Kb and 10Kb) and the 13Kb mouse insert (indicated by the arrow). The upper band is the 'arms' adjoined through their sticky-ends.
B. VLS-1 DNA and Balb/C DNA were digested with BamHI (lanes 3 and 4, respectively) and analysed as in A, except that the hybridization probe was a 1.2Kb XbaI fragment adjacent to the LTR. (This fragment is contained within a 2.3Kb BamHi fragment- see Fig. 2).

sequences were uniformally represented (data not shown). In both cases, the hybridization was confined mostly to a 0.6Kb XbaI fragment; additional weak hybridization was observed with the adjacent 1.2Kb XbaI fragment only under longer exposures. The identical hybridization pattern observed with both probes suggested that no VL30 sequences, other than the LTR, are present in clone VLS-1. The LTR-cross reactive XbaI fragments define the approximate location of the VL30 LTR. The exact boundaries of this LTR were established by direct DNA sequencing (see below).

VLS-1 DNA is an authentic genomic fragment

The propagation in phage of elements containing direct repeats may result in homologous recombination involving the repeated DNA. Such an event with cloned VL30 DNA would lead to the loss of all VL30 sequences, leaving one copy of the LTR. Therefore it was essential to demonstrate that this 'solo' LTR is present in the mouse genome in the exact molecular context as in clone

```
       10         20         30         40         50         60         70         80
TGAAGAATAAAGATAAAAAATTACTGAACTCTTCCTCACCCCAGAGCCCGACCCCTCCCATCTAGAGGTTGTTCCCAGAA

       90        100        110        120        130        140        150        160
CACTCCTGAACTCTTCACCCCAGAATGCATTCCTGAACTCCTCACCCTAGAGTTCGAACCCTCCCAACTAAAGACTGTTC

      170        180        190        200        210        220        230        240
TAAGAACATTTTTGAGATAAGGGCCTCCTGAACAACCTCAGAATTGAACCGGGTACATTGTCAAATAATAGGCATCACTC

      250        260        270        280        290        300        310        320
CTTAGTTATGTAGAATTCCCTTGGCATGACCCCTTGTCCCTTGGCAGAACCCCTTAGTTATGCAAACTTGTACTCCCTTG

      330        340        350        360        370        380        390        400
CCCCGCTCTCCCCCCTTGAGTTTTTCCTATATAAGCCGTGAAAAATTTGGCTCAAGGTCGATTCACCTCTACACCATACT

      410        420        430        440        450        460        470        480
AGGTGTATGAGTTTCGACCCCAGAGCTCTGGTCTATGTGCTTTCTTGCTGTTACTTTATTAAATCTTGCCCACTACAATT

      490        500        510        520        530        540        550        560
TGAGTTCGGTCTCAGTGTCTTCTTGGGTCCGCGGCTGTCTCGAGGCTTGAGTGAGGGTCTCCCTTCGGGTCTTTCA
```

```
                    LTR
.....GGGGGAAAGGAT TGAA.....TTCA GGATCAGGTATGGGGG.....
```

Figure 4. Sequence analysis of VLS-1 LTR and its adjacent sequences. Sequencing was carried out as described under 'Materials and Methods'. Boundaries of the LTR were defined by comparison with the DNA sequences of other VL30-LTR units determined in our laboratory. Bottom: LTR was boxed in order to illustrate the 4 bp inverted repeats at the LTR termini and the 4 bp duplication of flanking DNA (underlined sequence).

VLS-1. For this purpose, high molecular weight DNA from Balb/C mouse embryo was digested with EcoRI, electrophoresed alongside the EcoRI digest of clone VLS-1, blotted and hybridized with VLS-1 DNA that had been labeled by nick-translation. A hybridization band of genomic DNA, comigrating with the EcoRI insert of clone VLS-1, was indeed detected (Fig. 3A). The detection of additional bands is probably a reflection of the reiteration frequency in the mouse genome of both the VL30 DNA and the mouse sequences that flank this solo-LTR.

In order to minimize this hybridization background, a similar experiment was performed, but this time sequences flanking the solo-LTR in VLS-1 were used as the hybridization probe. This probe detected a 2.3Kb fragment in a BamHI digest of mouse DNA, which is identical in size to the BamHI fragment of VLS-1 DNA that comprise the probe sequences (Fig. 3B). These results establish that VLS-1 DNA is an authentic genomic fragment.

Nucleotide sequence of the solo-LTR unit and its adjacent DNA

We determined the nucleotide sequence of the single LTR contained in

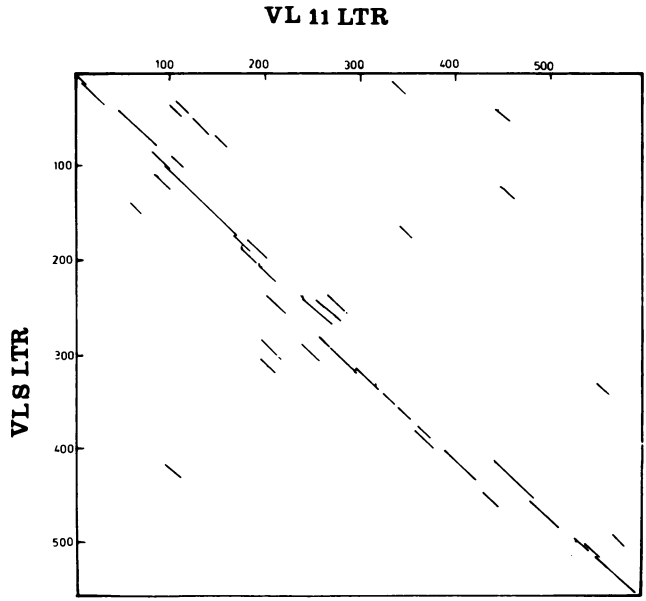

Figure 5. Dot matrix homology comparison between VLS-1 LTR and a standard VL30 LTR. Sequence of VLS-1 LTR was compared with the sequence of the LTR possessed by a standard complete VL30 unit (designated VL11-LTR), using the system described by Larson (23) with some modifications. A stringency of 10 out of 12 matches was employed.

VLS-1 DNA. Comparison with VL30 LTR sequences, that we had obtained before, allowed to define the exact boundaries of the LTR.

The sequence data (Fig. 4) yielded two important conclusions: First, the LTR is flanked by a 4 base-pair duplication of mouse DNA. This result confirmed the true solo-nature of the LTR in clone VLS-1, since all proviruses and transposable elements (wether a complete transposon or a single repeat unit), are flanked by a small duplication of the adjacent DNA (3,4). Second, dot-matrix comparison between the sequence of the LTR in clone VLS-1 and the sequence of a typical VL30-LTR unit indicated that these LTRs are grossly similar (Fig. 5). This finding eliminated the possibility that clone VLS-1 represents a genetic element distinct from VL30, whose LTR shares partial homology with the VL30-LTR.

DISCUSSION

In this study we have demonstrated the existance of 'solo' VL30-LTRs as natural constituents of the mouse genome.

Solo-LTRs have been encountered before in other systems which include both transposons and proviruses. A precursor-product relationship, between the complete transposing element and its solitary repeat, has been demonstrated (17,22). These findings formed the basis for the suggestion that solo LTRs are generated by excision of the complete element, via recombination between the long terminal repeats. It is most likely that the solo-LTR reported here was also formed through a similar excision of a VL30 element. It should be emphasized that we are looking at footprints of past events, and more specifically at those that occured in germ cells (otherwise they would not have been fixed in the genome). The fact that several solo VL30-LTRs are present in the Balb/C genome, together with 100-200 copies of complete VL30 units, attests for an unnegligible frequency of this event.

The question arises wether excision of endogenous retrovirus-like elements, leaving behind solo-LTRs, may be accompanied by alterations of the cell phenotype. It has been recently shown that indeed such an excision event may restore the wild type phenotype from a mutation caused by the insertion of an ecotropic endogenous provirus (22).

We have recently found, using in-vitro constructs, that VL30 LTRs have the capacity to efficiently drive the expression of adjacent genes (manuscript in prep.). Solo-LTR units may retain this capacity, and thereby effect the expression of nearby resident genes.

In bacterial systems it has been shown that insertion elements (IS), which are the repeated terminal unit of bacterial transposons, may autonomously transpose (4). Wether solo VL30-LTRs share this capacity remains to be seen.

In conclusion, the observation of solo VL30-LTRs reported here may be viewed as an example for non-programed DNA rearrangements, which may occasionally cause alterations in gene expression.

ACKNOWLEDGEMENT

This work was supported by a grant from the Leukemia Research Foundation Inc.

REFERENCES

1. Finnegan, D.J. (1981) Nature 292, 800-801.
2. Temin, H.M. (1980) Cell 21, 599-600.
3. Varmus, H.E. (1982) Science 216, 812-820.
4. Calos, M.P. and Miller J.H. (1980) Cell 20, 579-595.
5. Kugimiya, W., Ikenaga, H. and Saigo, K. (1983) Proc. Nat. Acad. Sci. (USA) 80, 3193-3197.

6. Elder, R.T., Loh, E.Y. and Davis, R.W. (1983) Proc. Nat. Acad. Sci. (USA) 80, 2432-2436.
7. Eibel, H., Gafner, J., Stotz, A. and Philippsen, P. (1980) Cold Spring Harb. Symp. Quant. Biol. 45, 609-617.
8. Shiba, T. and Saigo, K. (1983) Nature 302, 119-124.
9. Varmus, H.E., Quintrell, N. and Ortiz, S. (1981) Cell 25, 23-26.
10. Kuff, E.L., Feenstra, A., Leuders, K., Smith, L., Hawley, R., Hozumi, N. and Shulman, M. (1983) Proc. Nat. Acad. Sci. (USA) 80, 1992-1996.
11. Rechavi, G., Givol, D., and Canaani, E. (1982) Nature 300, 607-611.
12. Besmer, P., Olshevsky, U., Baltimore, D., Dolberg, D. and Fan, H. (1979) J. Virol. 29, 1168-1176.
13. Howk, R.S., Troxler, D.H., Lowy, D., Duesberg, P.H. and Scolnick, E.M. (1978) J. Virol. 25, 115-123.
14. Scolnick, E.M., Vaas, W.C., Howk, R.S. and Duesberg, P.H. (1979) J. Virol. 29, 964-972.
15. Keshet, E. and Shaul, Y. (1981) Nature 289, 83-85.
16. Itin, A. and Keshet, E. (1983) J. Virol. 47, 656-659.
17. Roeder, G.S. and Fink, G.R. (1980) Cell 21, 239-249.
18. Keshet, E., Shaul, Y., Kaminchik, J. and Aviv, H. (1980) Cell 20, 431-439.
19. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
20. Roop, D.R., Norstrom, J.L., Tsai, S.Y., Tsai, M.J., and O'Malley, B.W. (1978) Cell 15, 671-685.
21. Taylor, J.M., Illmensee, R. and Summers, J. (1976) Biochim. Biophys. Acta 442, 324-330.
22. Copeland, N.G., Hutchison, K.W., and Jenkins, N.A. (1983) Cell 33, 379-387.
23. Larson, R. (1982) Nucleic Acids Res. 10, 39-46.