Acta Biochimica Polonica

Vol. 46 No. 4/1999

873–878

QUARTERLY

Fragments of LINE-1 retrotransposons flanked by inverted telomeric repeats are present in the bovine genome. Homology with human LINE-1 elements[°]

Grażyna Płucienniczak and Andrzej Płucienniczak[✉]

Institute of Biotechnology and Antibiotics, Starościńska 5, 02-516 Warszawa, Poland

Received: 12 July, 1999

Key words: retrotransposons, LINE-1, L1MC4, telomeric repeats

In the bovine genome we found two intrachromosomal DNA fragments flanked by inverted telomeric repeats (GenBank Accession Nos. AF136741 and AF136742). The internal parts of the fragments are homologous exclusively to the human sequences and to the consensus sequence of the L1MC4 subfamily of LINE-1 retrotransposons which are widespread among mammalian genomes. We found that distribution of homologous human sequences within our fragments is not random, reflecting a complicated pattern of insertion mechanisms of and maintenance of retrotransposons in mammalian genomes.

One of the possible explanations of the origin of LINE-1 truncated elements flanked by inverted telomeric repeats in the bovine genome is that extrachromosomal DNA fragments may be modified by telomerase and subsequently, transferred into chromosomal DNA.

Eucaryotic chromosomal DNA is modified on its ends by telomeric repeats [1]. Particularly, mammalian chromosomes carry on their ends thousands of 5′-CCCTAA-3′ repeats which, it is believed, may participate in processes determining the biological clock of the cells [2–4]. One of the arising question is whether extrachromosomal elements derived from LINE-1 retrotransposons [5–8] or their parts may be modified by telomerase. There are several indications that retrotransposons have the ability to be included within telomeric sequences, for example into yeast telomeres [9]. We suppose that this phenomenon may proceed modification of extrachromosomal elements by telomerase, which may facilitate subsequent insertion into chromosomal DNA.

There are several examples of GenBank library entries of DNA fragments flanked by mammalian telomeric repeats one in human (Alvarez, 1992, unpublished, GenBank Acces-

[°]Supported by the State Committee for Scientific Research (KBN, Poland) grant No. 5 PO6D 003 11

[✉]Correspondence to Andrzej Płucienniczak: e-mail: apl@iba.waw.pl

sion No. X65597) and five in Chinese hamster (Alvarez, 1992, unpublished, GenBank Accession Nos. X65592 .. X65596) which are localised intrachromosomally. In this work we give two examples of such DNA fragments present in the bovine genome.

MATERIALS AND METHODS

DNA was isolated from peripheral blood using Epicentre Technologies kit (catalog No. MG71100). The DNA preparations were additionally treated with phenol and chloroform/isoamyl alcohol mixture (24:1, v/v), ethanol precipitated and dissolved in 10 mM Tris/ HCl, 1 mM EDTA, pH 8.0. PCR amplification was performed with primers: TELA (5'-CTAACCCTAACCCTAACCCTAAA-3'), TELC (5'-CTAACCCTAACCCTAACCCTAAC-3'), TELG (5'-CTAACCCTAACCCTAACCCTAAG-3') and TELT (5'-CTAACCCTAACCCTAACCCTAAT-3'), using different combinations of the oligonucleotides. The PCR was performed in 50 μl reaction volume in buffer containing 50 mM KCl, 2.0 mM $MgCl_2$, 0.02 mM of each dNTP, 10 mM Tris/HCl, pH 8.9, 50 pM of primer(s), two units of Taq DNA polymerase (Ingen, Sieradz, Poland) and 100 ng of total bovine DNA as a template. Following the initial denaturation step for 2 min at 94°C, Taq DNA polymerase was added at 72°C (hot start), and PCR was performed for 12, 15, 18, 21, 24, 27, 30, 33 and 36 cycles with denaturation at 94°C for 30 s, annealing at 58°C, and elongation at 72°C for 1 min. Subsequently, the reactions were incubated for an additional 2 min at 72°C. The amplified DNA fragments were isolated by polyacrylamide gel electrophoresis as described previously [10], phosphorylated with phage T4 polynucleotide kinase, treated with the Klenow fragment of *Escherichia coli* DNA polymerase I in the presence of 1 mM dNTP to remove 3'-overhangs introduced by Taq DNA polymerase [11], treated with phenol and chloroform/isoamyl alcohol mixture (24:1, v/v), ethanol precipitated and ligated to *Hinc*II digested pBS SK+ (Stratagene) plasmid. Ligation products were transformed into the NM522 *E. coli* strain. Cells of the MV1193 *E. coli* strain were transformed with recombinant plasmids and infected with helper bacteriophage M13K07 to obtain single-stranded DNA for nucleotide sequence determination which was performed using the standard dideoxy method [12] with Taq DNA polymerase and [α-^{32}P]dCTP internal labelling [12].

With the aim to establish approximate number of DNA fragments flanked with telomeric repeats in the bovine genome the 248 bp fragment of bovine growth hormone gene was amplified with primers BGHP (5'-TCCTCAAGCAGACCTATGAC-3') and BGHM (5'-TATTAGGAAAGGACAGTGGG-3') for 12, 15, 18, 21, 24, 27, 30, 33 and 36 cycles under the same conditions as described above.

Database searches were performed using the National Institutes of Health ENTREZ and BLAST software [13]. Presence of repetitive sequences in the cloned DNA fragments was established by CENSOR [14] or Repeat Masker (Smit A.F.A. & Green P., unpublished, http://ftp.genome.washington.edu/cgi-bin/RepeatMasker) software.

RESULTS AND DISCUSSION

After PCR amplification of the total bovine DNA with primers TELA and TELC two DNA fragments (B1 and B2) were obtained (Fig. 1). The presence of TELC primer consisting of unchanged telomeric repeats alone in the PCR reaction mixture was not sufficient for revealing these DNA fragments, neither were other combinations of primer pairs or single primers in the PCR reaction mixture. The experiment with the parallel PCR of the part of the bGH gene and B1 an B2 showed that the fragments amplified with telomeric primers are unique or that their number of repeats is low in the bovine genome (Fig. 1A and B). Two

DNA fragments were amplified at the time which is clearly seen after digestion of the amplified fragments with *Eco*RV restriction nuclease (Fig. 1C). The results of the PCR were independent of the sex of the DNA donor.

The amplified fragments were cloned and their nucleotide sequence was determined. It appeared that both DNA fragments are flanked by telomeric repeats characteristic for the mammalian chromosomes (Fig. 2).

The nucleotide sequence analysis shows that the B1 and B2 fragments differ in their length mainly because of the 37 bp long direct repeat present in the first of them (Fig. 2). The 37 bp long repeated sequence is contiguous to the 5′ telomeric flanking region. The nucleotide sequences forming direct repeat and these localised between positions 123–213 for B1 and 85–175 for B2 show no homology to the GenBank data library sequences. Comparison of the rest of the nucleotide sequences of the B1 and B2 with data deposited in GenBank revealed homology to sequences which were found to belong to the LINE-1 elements, particularly to the L1MC4 subfamily of these elements [15]. Significant alignments were obtained exclusively to the human sequences which are common for L1MC1 and L1MC4 LINE-1 elements. Further search performed with CENSOR [14] and RepeatMasker programs showed that the best homology is observed to the consensus sequence of L1MC4 [15], subfamily of the LINE-1 retrotransposons (Fig. 2). The homology region of both B1 and B2 is localised on the 3′ UTR region of L1MC4.

Three regions of our fragments (positions 276–327 and 447–493 on B1 and also sequences containing telomeric repeats) produce significant alignments with the human sequences very frequently (Fig. 2). Distribution of the homologous sequences within our B1 and B2 DNA fragments is not random. The

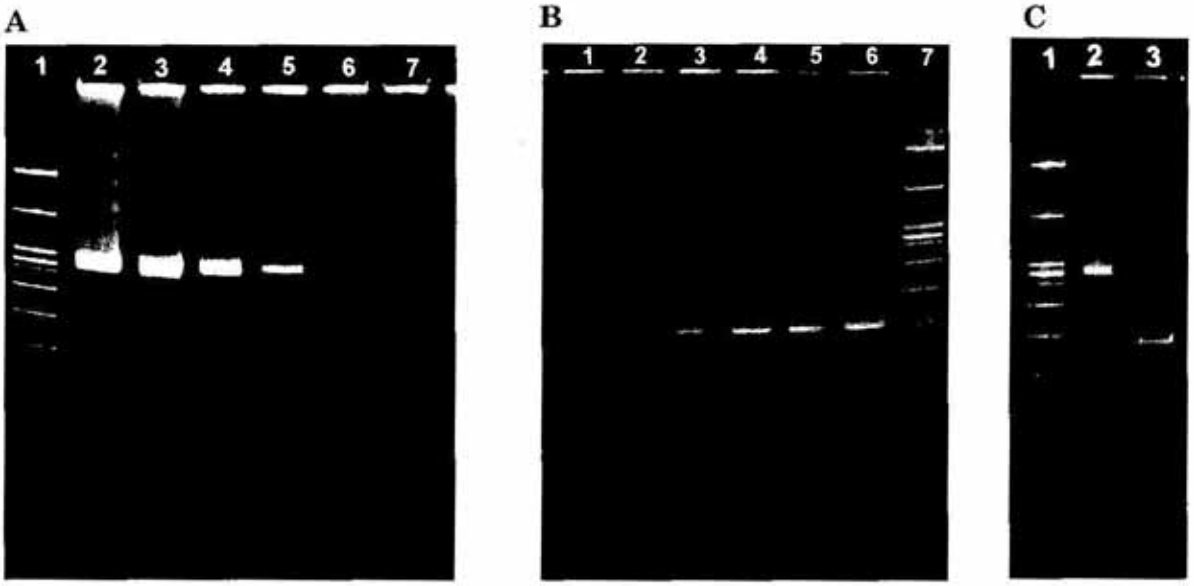

Figure 1. Electrophoretic pattern of the amplification products carried out with telomeric primers (A and C), and the bovine growth hormone gene fragment, primers BGHP and BGHM (B).

A, lines 2–7, number of cycles carried out: 36, 33, 30, 27, 24 and 21, respectively. B, lines 1–6, number of cycles carried out: 21, 24, 27, 30, 33 and 36, respectively. Lanes 1 in A and C represent the mixture of DNA fragments obtained after separate digestion of the pUC19 plasmid with *Taq*I and *Msp*I restriction nucleases (fragment length in bp: 1444, 736, 501, 489, 476, 404, 331, 242, 190, 147 and 111+110). The same mixture is present in B, line 7. C, line 2, DNA fragments amplified with telomeric primers; line 3, the same fragments digested with *Eco*RV restriction nuclease.

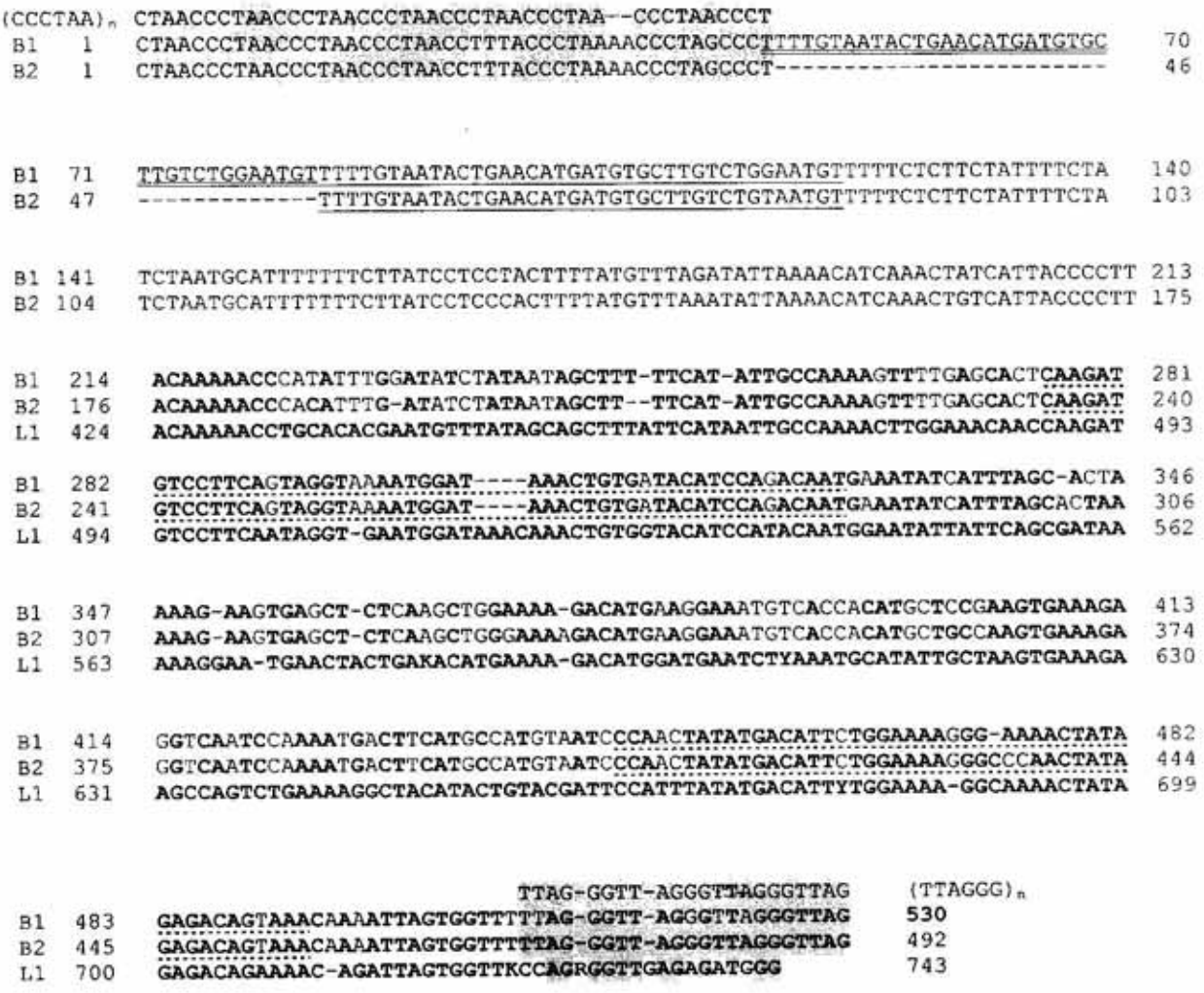

Figure 2. The nucleotide sequences of bovine LINE-1 elements (B1 and B2) flanked by inverted telomeric repeats and their comparison with the consensus sequence of L1MC4 subfamily [15] of LINE-1.

The fragment of the L1MC4 consensus sequence is shown in bold. Regions of the bovine sequences which are identical with the L1MC4 nucleotide sequence are also shown in bold. The nucleotides forming 37 bp direct repeat on the B1 DNA fragment are underlined and doubly underlined. The regions of B1 and B2 producing the most frequent significant alignments with human DNA sequences are underlined with a dotted line. The hexanucleotides forming telomeric repeats are shadowed. For comparison, canonical telomeric repeats $(5'\text{-CCCTTA-}3')_n$ and $(5'\text{-TAAGGG-}3')_n$ are also shown.

positions of the first and the last nucleotides of homology regions show narrow distributions, being much sharper for the beginnings of significant alignments (Fig. 3). These results reflect some processes which govern reverse transcription of LINE-1 elements, their insertion into genomes and further events: mutation rates at a given position, secondary insertions of different elements into LINE-1 sequences and recombination processes. At present it is difficult to evaluate which one of the mentioned factors is the most important for the observed pattern.

The existence in the three genomes of interstitial DNA sequences containing inverted telomeric repeats at their ends (Alvarez, 1992, unpublished and this work) arises a question about the mechanisms leading to their origin and present localisation. We see at least two ways to explain our results. The simplest way

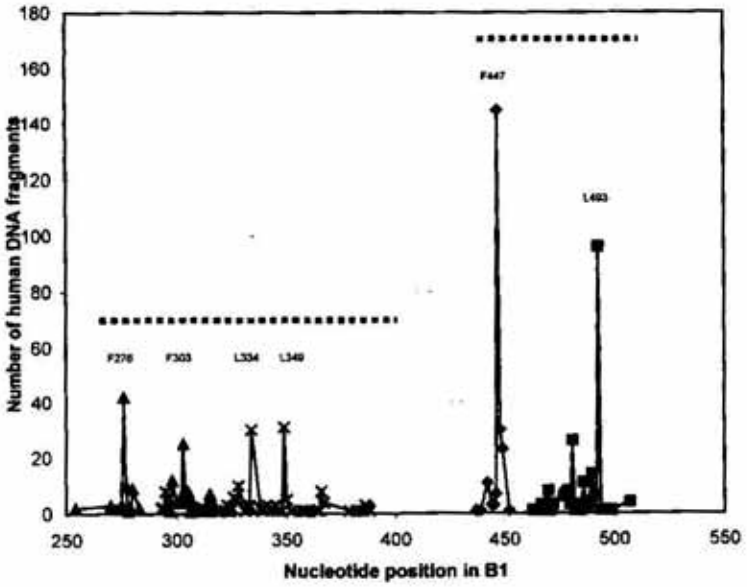

Figure 3. Distribution of the first and the last positions of B1 regions producing significant alignments with human DNA sequences.

The two regions of homology to the fragments of human DNA sequences are shown with broken lines above the plots. Positions of maximal numbers of significant alignments are shown above the peaks (F276, F303, L334, L349, F447 and L493; F and L show the first and the last positions, respectively).

is to assume that extrachromosomal DNA fragments, truncated parts of LINE-1 elements, arising during the transposition process, were modified by telomerase and after are similar to the telomeric repeats (Fig. 2). This may facilitate modification of truncated members of the family by telomerase.

If telomerase action is not considered, the

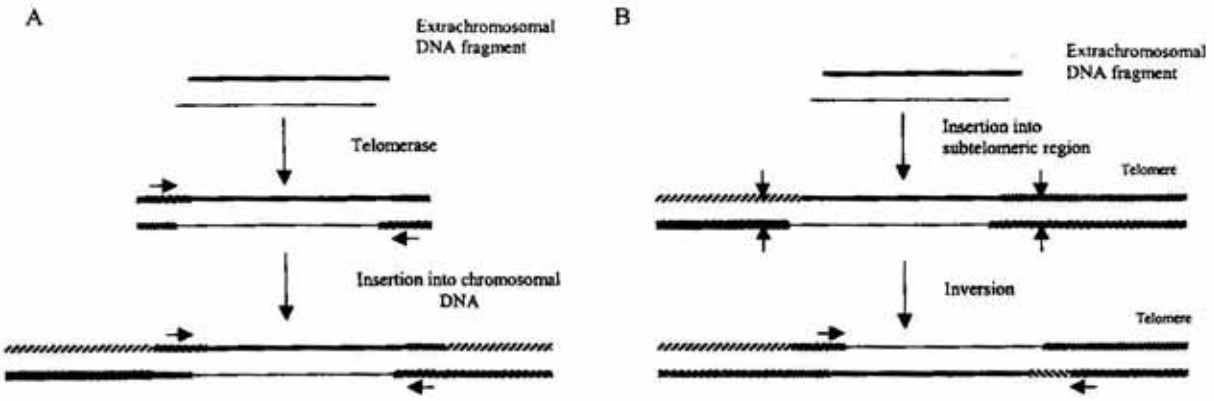

Figure 4. Two ways leading to the formation of interstitial DNA sequences containing inverted telomeric repeats at their ends.

Parallel and vertical arrows show polarity of telomeric repeats and sites of the DNA strand scissions, respectively.

that inserted into chromosomal DNA. Such an assumption explains in a simple way the presence of inverted repeats of telomeric origin at the flanks of the amplified fragments (Fig. 4A). This variant of explanation is supported to some extent by the fact that DNA fragments coming from the LINE-1 family of retrotransposons comprise sequences which explanation is more complicated. In this case, because of the presence of the inverted repeats at the flanks, it is necessary to assume that insertions of extrachromosomal fragments into the regions of chromosomes containing telomeric repeats were followed by double strand scissions in the neighborhood of the inserts and inversions of the tras-

poson's fragments together with arrays of telomeric hexanucleotides (Fig. 4B). Different variants of the second explanation are possible.

Our results do not allow to choose between the two possibilities.

REFERENCES

1. Blackburn, E.H. (1991) Structure and function of telomeres. *Nature* **350**, 569–573.
2. Counter, C.M., Avilion, A.A., LeFeuvre, C.E., Stewart, N.G., Greider, C.W., Harley, C.B. & Bacchetti, S. (1992) Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. *EMBO J.* **11**, 1921–1929.
3. Bryan, T.M., Englezou, A., Gupta, J., Bacchetii, S. & Reddel, L.L. (1995) Telomere elongation in immortal human cells without detectable telomerase activity. *EMBO J.* **14**, 4240–4248.
4. Lundblad, V. & Wright, W.E. (1996) Telomeres and telomerase: A simple picture becomes complex. *Cell* **87**, 369–375.
5. Skowronski, J., Fanning, T.G. & Singer, M.F. (1988) Unit length line-1 transcripts in human teratocarcinoma cells. *Mol. Cell. Biol.* **8**, 1385–1397.
6. Dombroski, B.A., Mathias, S.L., Nanthakumar, E., Scott, A.F. & Kazazian, H.H., Jr. (1991) Isolation of an active human transposable element. (1991) *Science* **254**, 1805–1808.
7. Dombroski, B.A., Scott, A.F. & Kazazian, H.H., Jr. (1993) Two additional potential retrotransposons isolated from the human L1 subfamily that contains an active retrotransposable element. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6513–6517.
8. Sassaman, D.M., Dombroski, B.A., Moran, J.V., Kimberland, M.L., Naas, T.P., DeBerardinis, R.J., Gabriel, A., Swergold, G.D. & Kazazian, H.H., Jr. (1997) Many human L1 elements are capable of retrotransposition. *Nature Genet.* **16**, 37–43.
9. Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M., Louis, E.J., Mewes, H.W., Murakami, Y., Philippsen, P., Tettelin, H. & Oliver, S.G. (1996) Life with 6000 genes. *Science* **274**, 546–567.
10. Dybczyński, I. & Płucienniczak, A. (1988) A protocol for DNA fragment extraction from polyacrylamide gels. *Biotechniques* **6**, 924–926.
11. Clark, J.M. (1988) Novel non-templated nucleotide addition reactions catalysed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res.* **16**, 9677–9686.
12. Sanger, F., Nicklen, S. & Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467.
13. Altshul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Miller, W. & Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST, a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402.
14. Jurka, J., Klonowski, P., Dagman, V. & Pelton, P. (1996) CENSOR – a program for identification and elimination of repetitive elements from DNA sequences. *Computers and Chemistry* **20**, 119–122.
15. Smit, A.F.A., Toth, G., Riggs, A.D. & Jurka, J. (1995) Ancestral, mammalian-wide subfamilies of LINE-1 repetitive sequences. *J. Mol. Biol.* **246**, 401–417.