Volume 10 Number 5 1982 Nucleic Acids Research

Structural analysis of repetitive DNA sequences in the bovine corticotropin-β-lipotropin precursor gene region

Yumiko Watanabe*, Toshihiko Tsukada*+, Mitsue Notake*, Shigetada Nakanishi† and Shosaku Numa*

*Department of Medical Chemistry, +Second Division, Department of Internal Medicine, and †Institute for Immunology, Kyoto University Faculty of Medicine, Kyoto 606, Japan

Received 6 January 1982; Accepted 1 February 1982

ABSTRACT

Repetitive DNA sequences in the bovine corticotropin-β-lipotropin precursor gene region have been mapped and subjected to nucleotide sequence analysis. Two of the four repetitive DNA segments found are located in the 5'-flanking region, and one each within the intervening sequences. Each repetitive DNA segment contains one to three highly homologous unit sequences with an approximate length of 120 base pairs. All the unit sequences are flanked on the 3' side by tandem repeats. There are about 10^5 copies of the repetitive DNA in the bovine genome. Comparison of the bovine repetitive sequences with those of other mammalian species reveals the presence of a homologous segment of approximately 40 base pairs. This segment and the region preceding it in the bovine repetitive DNA exhibit sequence homology with the region encompassing the origin of DNA replication in papovaviruses.

INTRODUCTION

The genomes of higher organisms contain reiterated DNA sequences that are interspersed with unique DNA sequences (1,2). Using blot hybridization analysis, we have previously shown that four such repetitive sequences are present in the bovine genomic DNA segment of approximately 17 kilobase pairs (kb) that contains the gene encoding the common precursor of the pituitary hormones corticotropin (ACTH) and β-lipotropin (β-LPH) (3). The bovine ACTH-β-LPH precursor gene is approximately 7.3 kb long and consists of three exons divided by two large intervening sequences (introns) (3,4); one of the introns with an approximate length of 4 kb (intron A) is located within the segment encoding the 5'-untranslated region of the mRNA, and the other with an approximate length of 2.2 kb (intron B) interrupts the protein-coding sequence near the signal peptide region. Two of the repetitive DNA sequences found in this gene region are positioned in the 5'-flanking segment, and one each within the introns. All these sequences have been shown by cross-hybridization analysis to be homologous with one another. In the present investigation, we have carried out detailed mapping and nucleotide sequence analysis of the repetitive DNA in the bovine ACTH-β-LPH precursor gene region

0305-1048/82/1005-1459$2.00/0

and have estimated its reiteration frequency in the bovine genome. The repetitive DNA sequences determined are compared with one another and with other mammalian repetitive sequences.

MATERIALS AND METHODS

DNA Preparations. Recombinant DNA clones that carry bovine genomic DNA fragments containing the ACTH-β-LPH precursor gene as well as plasmids carrying the respective subcloned fragments in pBR322 were described previously (3,4). Calf thymus DNA was prepared as described by Polsky et al. (5).

Restriction Mapping and DNA Sequencing. Restriction endonucleases were obtained from Takara Shuzo Co. (Kyoto, Japan), Bethesda Research Laboratories (Rockville, USA) and New England Biolabs (Beverly, USA). Reactions were carried out under the conditions recommended by the vendors. Restriction mapping was accomplished by combining data obtained by digestion and subsequent gel electrophoresis of both 5'-end-labelled and unlabelled DNA fragments and by blot hybridization analysis with total calf thymus DNA (EcoRI-digested) labelled by nick translation with [α-^{32}P]dCTP (the Radiochemical Centre, Amersham, England) as a probe; the procedures used were as described previously (3,4). 5'-End labelling of restriction fragments and DNA sequencing were carried out by the procedure of Maxam and Gilbert (6).

Cot Analysis. Hybridization reactants were incubated at 60°C in the presence of 20 mM Tris-HCl buffer (pH 7.4), 0.6 M NaCl and 0.2 mM EDTA. The reaction mixtures were covered with liquid paraffin to prevent evaporation. Aliquots were taken at time intervals to assay the hybrid formed by measuring its resistance to S1 nuclease. For this assay, samples were incubated at 37°C for 30 min in a reaction mixture containing 50 mM sodium acetate buffer (pH 4.5), 0.25 M NaCl, 1 mM $ZnSO_4$, 17 μg/ml Escherichia coli DNA (sheared and single-stranded) and 50 units/ml S1 nuclease (P-L Biochemicals, Milwaukee, USA), and the radioactivity in trichloroacetic acid-insoluble material was determined. Equivalent Cot values (those that would obtain at 0.18 M NaCl) have been plotted (7).

RESULTS

The four DNA segments containing repetitive sequences in the ACTH-β-LPH precursor gene region, designated as a, b, c and d in the direction of transcription of the gene, were mapped by blot hybridization analysis with the total DNA probe as shown in Fig. 1. These DNA segments and their

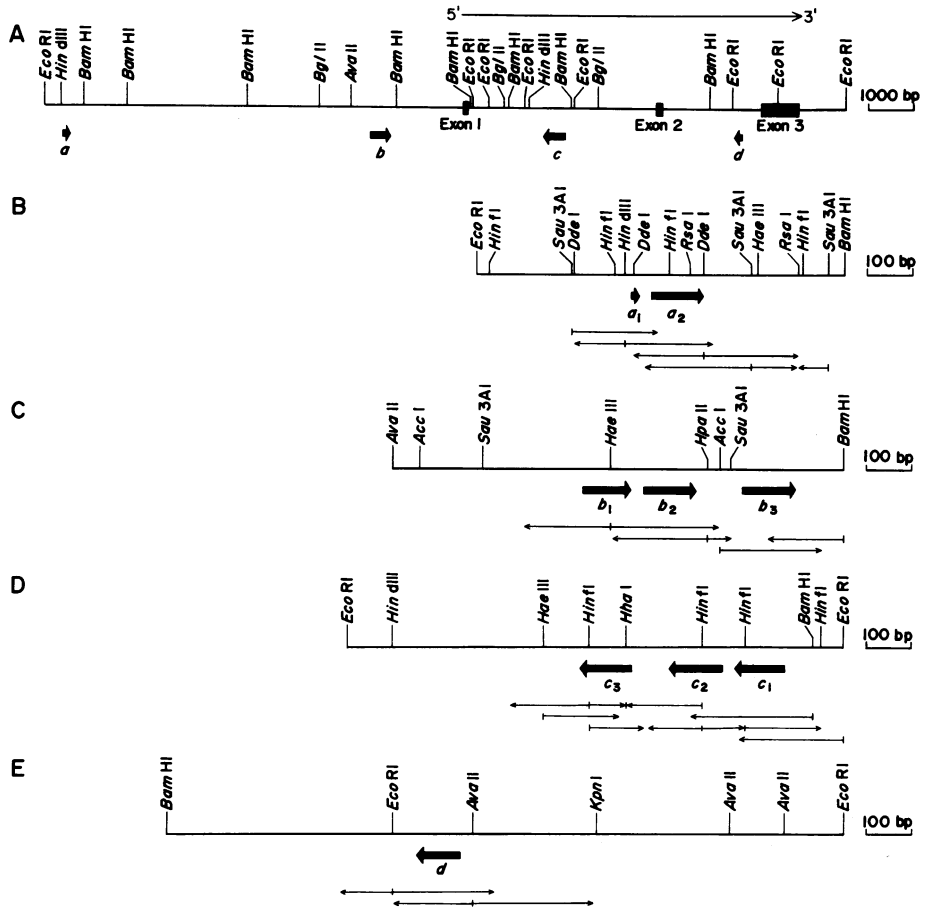
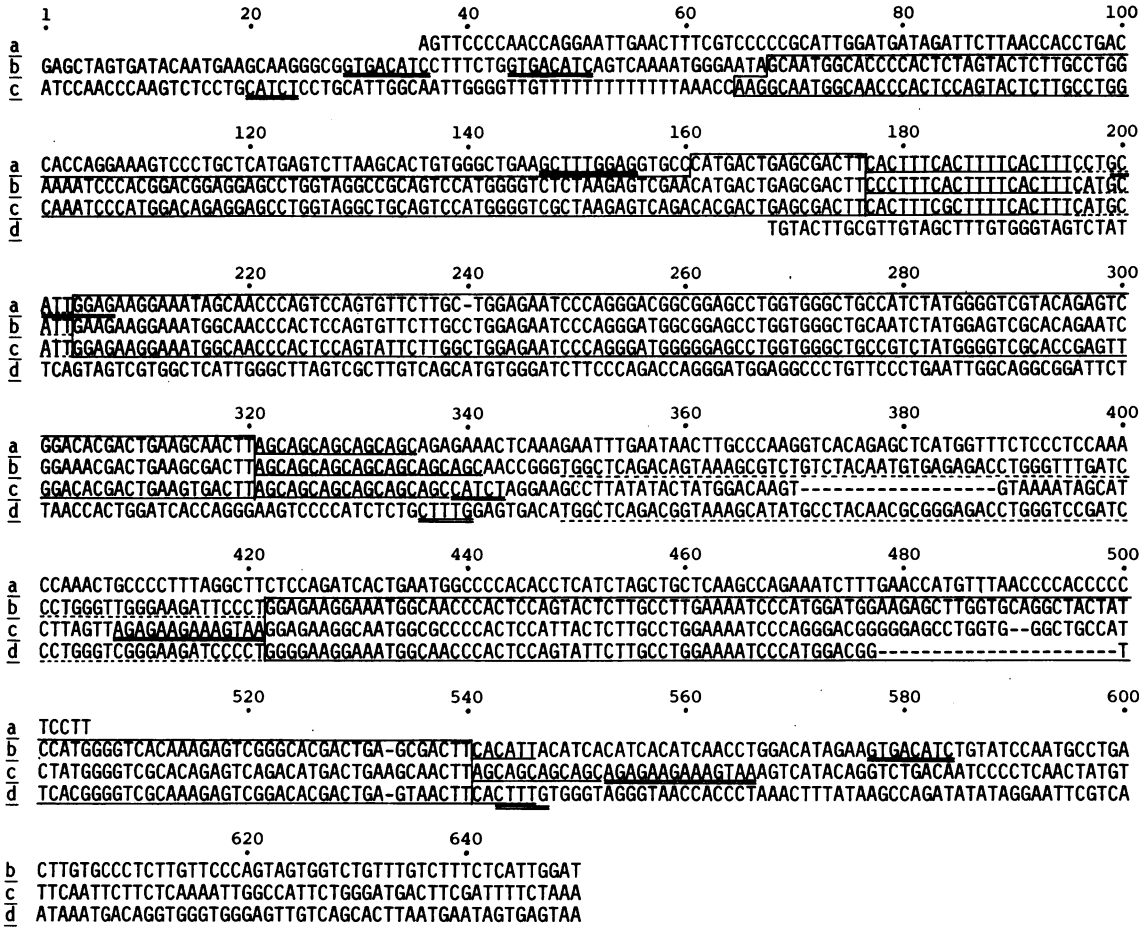


Figure 1. Restriction mapping of repetitive sequences in the bovine ACTH-β-LPH precursor gene region and strategy for DNA sequencing. (A) 17 kb DNA segment containing the ACTH-β-LPH precursor gene; (B), (C), (D), (E) portions of the 17 kb DNA segment containing the repetitive sequences a, b, c and d, respectively. The direction of transcription of the ACTH-β-LPH precursor gene is from left to right. Scales are given on the right side of each restriction map. For reference, the locations of the exons are shown by closed boxes in (A), and those of the repetitive sequences by thick arrows beneath the restriction maps; the direction of the arrows has been assigned by comparison with the human Alu family sequences in such a way that it would indicate the direction of in vitro transcription of these sequences by RNA polymerase III (26). Only relevant restriction sites are displayed for AvaII in (A), for DdeI and RsaI in (B) and for AvaII and KpnI in (E). The thin arrows under the restriction maps in (B), (C), (D) and (E) indicate the direction and extent of sequence determinations.


1 20 40 60 80 100
a AGTTCCCCAACCAGGAATTGAACTTTCGTCCCCCGCATTGGATGATAGATTCTTAACCACCTGAC
b GAGCTAGTGATACAATGAAGCAAGGGCGGTGACATCCTTTCTGGTGACATCAGTCAAAATGGGAATAGCAATGGCACCCCACTCTAGTACTCTTGCCTGG
c ATCCAACCCAAGTCTCCTGCATCTCCTGCATTGGCAATTGGGGTTGTTTTTTTTTTTTTAAACCAAGGCAATGGCAACCCACTCCAGTACTCTTGCCTGG
120 140 160 180 200
a CACCAGGAAAGTCCCTGCTCATGAGTCTTAAGCACTGTGGGCTGAAGCTTTGGAGGTGCCCATGACTGAGCGACTTCACTTTCACTTTTCACTTTCCTGC
b AAAATCCCACGGACGGAGGAGCCTGGTAGGCCGCAGTCCATGGGGTCTCTAAGAGTCGAACATGACTGAGCGACTTCCCTTTCACTTTTCACTTTCATGC
c CAAATCCCATGGACAGAGGAGCCTGGTAGGCTGCAGTCCATGGGGTCGCTAAGAGTCAGACACGACTGAGCGACTTCACTTTCGCTTTTCACTTTCATGC
d TGTACTTGCGTTGTAGCTTTGTGGGTAGTCTAT
220 240 260 280 300
a ATTGGAGAAGGAAATAGCAACCCAGTCCAGTGTTCTTGC-TGGAGAATCCCAGGGACGGCGGAGCCTGGTGGGCTGCCATCTATGGGGTCGTACAGAGTC
b ATTGAAGAAGGAAATGGCAACCCACTCCAGTGTTCTTGCCTGGAGAATCCCAGGGATGGCGGAGCCTGGTGGGCTGCAATCTATGGAGTCGCACAGAATC
c ATTGGAGAAGGAAATGGCAACCCACTCCAGTATTCTTGGCTGGAGAATCCCAGGGATGGGGGAGCCTGGTGGGCTGCCGTCTATGGGGTCGCACCGAGTT
d TCAGTAGTCGTGGCTCATTGGGCTTAGTCGCTTGTCAGCATGTGGGATCTTCCCAGACCAGGGATGGAGGCCCTGTTCCCTGAATTGGCAGGCGGATTCT
320 340 360 380 400
a GGACACGACTGAAGCAACTTAGCAGCAGCAGCAGCAGAGAAACTCAAAGAATTTGAATAACTTGCCCAAGGTCACAGAGCTCATGGTTTCTCCCTCCAAA
b GGAAACGACTGAAGCGACTTAGCAGCAGCAGCAGCAGCAGCAACCGGGTGGCTCAGACAGTAAAGCGTCTGTCTACAATGTGAGAGACCTGGGTTTGATC
c GGACACGACTGAAGTGACTTAGCAGCAGCAGCAGCAGCCATCTAGGAAGCCTTATATACTATGGACAAGT----------GTAAAATAGCAT
d TAACCACTGGATCACCAGGGAAGTCCCCATCTCTGCTTTGGAGTGACATGGCTCAGACGGTAAAGCATATGCCTACAACGCGGGAGACCTGGGTCCGATC
420 440 460 480 500
a CCAAACTGCCCCTTTAGGCTTCTCCAGATCACTGAATGGCCCCACACCTCATCTAGCTGCTCAAGCCAGAAATCTTTGAACCATGTTTAACCCCACCCCC
b CCTGGGTTGGGAAGATTCCCTGGAGAAGGAAATGGCAACCCACTCCAGTACTCTTGCCTTGAAAATCCCATGGATGGAAGAGCTTGGTGCAGGCTACTAT
c CTTAGTTAGAGAAGAAAGTAAGGAGAAGGCAATGGCGCCCCACTCCATTACTCTTGCCTGGAAAATCCCAGGGACGGGGGAGCCTGGTG--GGCTGCCAT
d CCTGGGTCGGGAAGATCCCCTGGGGAAGGAAATGGCAACCCACTCCAGTATTCTTGCCTGGAAAATCCCATGGACGG----------T
520 540 560 580 600
a TCCTT
b CCATGGGGTCACAAAGAGTCGGGCACGACTGA-GCGACTTCACATTACATCACATCACATCAACCTGGACATAGAAGTGACATCTGTATCCAATGCCTGA
c CTATGGGGTCGCACAGAGTCAGACATGACTGAAGCAACTTAGCAGCAGCAGCAGAGAAGAAAGTAAAGTCATACAGGTCTGACAATCCCCTCAACTATGT
d TCACGGGGTCGCAAAGAGTCGGACACGACTGA-GTAACTTCACTTTGTGGGTAGGGTAACCACCCTAAACTTTATAAGCCAGATATATAGGAATTCGTCA
620 640
b CTTGTGCCCTCTTGTTCCCAGTAGTGGTCTGTTTGTCTTTCTCATTGGAT
c TTCAATTCTTCTCAAAATTGGCCATTCTGGGATGACTTCGATTTTCTAAA
d ATAAATGACAGGTGGGTGGGAGTTGTCAGCACTTAATGAATAGTGAGTAA

surrounding regions were subjected to nucleotide sequence analysis by the procedure of Maxam and Gilbert (6) according to the strategy indicated in Fig. 1B, C, D and E. The nucleotide sequences determined are presented in Fig. 2. Each repetitive DNA segment contains one or more highly homologous sequences with an approximate length of 120 base pairs (bp). Three such 120 bp unit sequences are present within the segments b and c (b_1, b_2, b_3 and c_1, c_2, c_3, respectively). The sequences lying between these 120 bp units are also partly homologous in the segments b and c, so that the homology extends over 390 bp. On the other hand, the segment a contains one 120 bp homologous unit (a_2) and a portion of it (a_1), and the segment d consists of one such unit with a 20 bp deletion in the middle. The homologous units a_2, b_2 and c_2 are preceded by an 8 bp homologous sequence, and the units b_3 and d by a 73 bp homologous sequence. Because small portions of repetitive sequences or extensively divergent repetitive sequences may not be detected by blot hybridization analysis with the total DNA probe (see the accompanying paper), the possibility that such aberrant repetitive sequences are additionally present in this gene region cannot be excluded.

In Fig. 3, the sequences of the 120 bp homologous units contained in the segments a, b, c and d are aligned, and a consensus sequence is deduced. Thus aligned, the unit sequences a_1, b_1, c_1, a_2, b_2, c_2, b_3, c_3 and d exhibit 78-100 % homology; percent homology was calculated according to the formula described by Pan et al. (8), except that gaps were counted as one substitution regardless of their length. All the unit sequences are flanked on the 3' side by tandem repeats of $(AGC)_x$ or $(CACT_y)_z$ (x = 4-7; y = 2-4; z = 3; in the case of b_3 and d, z = 1) (Fig. 2). The structural organizations of the four repetitive DNA segments are schematically illustrated in Fig. 4.

The number of copies of the repetitive DNA in the bovine genome was estimated by analysis of reassociation rates. As shown in Fig. 5, the renaturation reached 50 % of the endpoint at Cot = 0.025; E. coli DNA re-

Figure 2. Nucleotide sequences of the four DNA segments containing repetitive sequences. Nucleotide residues are numbered in the direction from 5' to 3' as indicated by the thick arrows in Fig. 1. Gaps (-) have been inserted by inspection to achieve maximum homology. The 120 bp homologous units (a_1, a_2, b_1, b_2, b_3, c_1, c_2, c_3 and d) are boxed. The tandem repeats flanking the 120 bp unit sequences (single lines), the homologous regions preceding the sequences a_2, b_2, and c_2 as well as those preceding the sequences b_3 and d (dashed lines) and the non-conserved short direct repeats (double lines) are underlined.

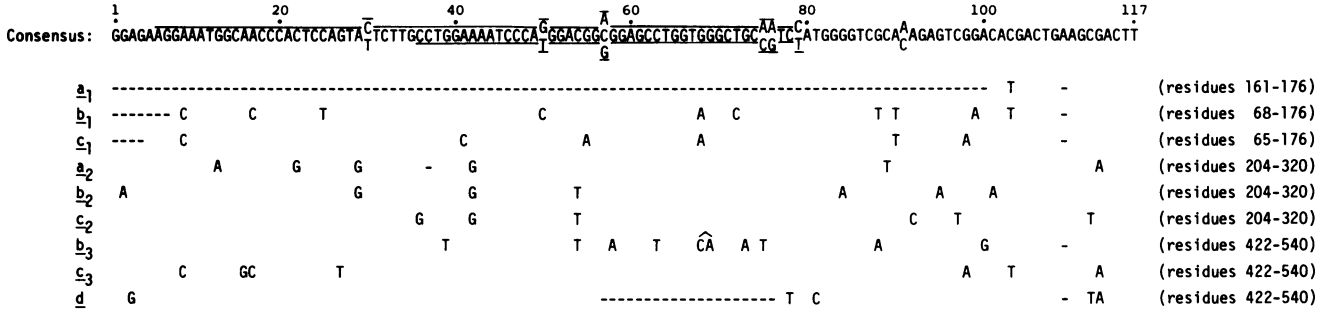


Figure 3. Sequence comparison of the 120 bp homologous units. The consensus sequence is given above, and only the divergencies from it are indicated for each unit. The locations of additions (^) and deletions (-) have been determined by inspection to achieve maximum homology. The nucleotide numbers given on the right side refer to those in Fig. 2 and differ from those for the consensus sequence. The segment exhibiting homology with the repetitive sequences of other mammalian species is underlined, and that showing homology with the regions encompassing the replication origins of human BK virus and simian virus 40 is overlined.

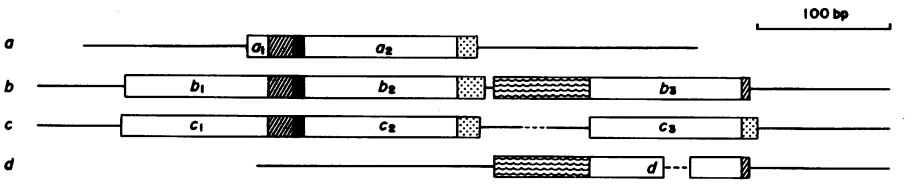


Figure 4. Schematic representation of the four repetitive DNA segments. The direction from left to right corresponds to those of the arrows shown in Fig. 1. Homologous regions are indicated by bars, non-homologous regions by solid lines, and deletions by dashed lines. The 120 bp homologous units are represented by open bars, the homologous regions preceding the sequences a_2, b_2 and c_2 by closed bars, those preceding the sequences b_3 and d by waved bars, the $(\overline{AGC})_x$ tandem repeats by stippled bars, and the $(C\overline{ACT}_y)_z$ tandem repeats by hatched bars. A scale of 100 bp is given.

annealed at a $Cot_{1/2}$ of 5.6 under the identical conditions (data not shown). Because the ratio of the $Cot_{1/2}$ value for calf thymus unique sequence DNA to that for E. coli DNA has been reported to be 500 (1), it is estimated that there are about 10^5 copies of the repetitive DNA in the bovine genome.

DISCUSSION

In the present investigation, the interspersed repetitive sequences in the bovine ACTH-β-LPH precursor gene region have been analyzed. The bovine repetitive DNA segments each contain one to three highly conserved units with an approximate length of 120 bp. The other mammalian repetitive sequences thus far documented include the human Alu family (8-10; see also the

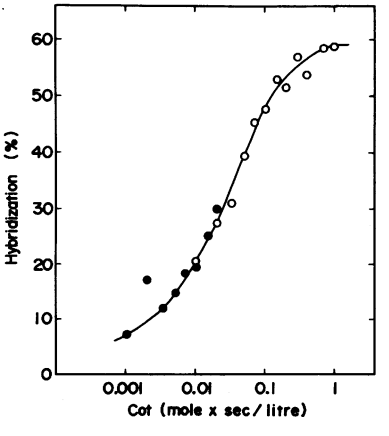


Figure 5. Analysis of the reiteration frequency of the repetitive DNA. Total calf thymus DNA, sonicated to yield 600-800 bp fragments, was incubated at a concentration of 0.174 μg/ml (●) or 1.74 μg/ml (○) with 0.3 ng/ml or 1.0 ng/ml of the nick-translated 666 bp HaeIII-EcoRI fragment containing the repetitive DNA segment c (2.5 x 10^8 cpm/μg) as a tracer at 60°C for varying times after heat denaturation. Duplex formation was assayed by S1 nuclease digestion and has been plotted against the logarithm of Cot [product of DNA concentration (in moles of nucleotides per litre) and time (in seconds)].

accompanying paper), mouse B1 sequences (11), rat repetitive sequences found in an intron of the growth hormone gene (12,13) and Chinese hamster ovary cell Alu-equivalent family (14,15). Most of these repetitive sequences exhibit two common features; first, a dA-rich sequence present at their 3' end (for each unit of the dimeric or trimeric structure of the human and rat repetitive sequences) and secondly, direct repeats of a short non-conserved sequence flanking their both sides. In the case of the bovine repetitive DNA sequences, no dA-rich sequence is found at their 3' end, which, however, is flanked by tandem repeats of $(AGC)_x$ or $(CACT_y)_z$ (x = 4-7, y = 2-4, z = 3 or 1). The dA-rich sequences observed in the repetitive sequences of the other mammalian species may be regarded in general as tandem repeats of A_xN_y where a variable number of A residues are followed by one or a few residues of C, T or G.

The repetitive unit sequence $\underline{c}_3$ is flanked by direct repeats of 14 nucleotides, AGAGAAGAAAGTAA (residues 408-421 and 553-566 in Fig. 2). No other typical direct repeats are found to flank the bovine repetitive segments extending over one 120 bp unit sequence or its multiples. But some atypical direct repeats are noted; for example, $GC\begin{matrix}T\\A\end{matrix}TTGGAG$ (residues 147-155 and 199-207) outside the sequence $\underline{a}_1$; GTGACATC (residues 29-36, 44-51 and 577-584) upstream of $\underline{b}_1$ and downstream of $\underline{b}_3$; CATCT (residues 20-24 and 339-343) upstream of $\underline{c}_1$ and downstream of $\underline{c}_2$; CTTTG (residues 336-340 and 543-547) outside $\underline{d}$. That the repetitive sequences are bounded by non-conserved short direct repeats implies that their presence may be the result of transposition events (16-19).

Comparison of the bovine repetitive sequences with the human Alu family and related sequences (the mouse B1 sequences and Chinese hamster ovary cell type 1 Alu-equivalent sequences) as well as with the rat repetitive sequences and related sequences (Chinese hamster ovary cell type 2 Alu-equivalent sequences) reveals the presence of a homologous segment of approximately 40 bp (Fig. 6). The 3' half of this segment includes the 14 bp sequence that is known to be common between the Alu family and related sequences and the replication origins of papovaviruses (20) (see also Fig. 7). It is to be noted that this 14 bp sequence is entirely deleted in the bovine repetitive sequence $\underline{d}$. Furthermore, the segment of the bovine repetitive sequences preceding the 14 bp sequence is extensively homologous with the corresponding regions of the human BK virus (21) and simian virus 40 genomes (22) (Fig. 7). Some homology is also noted between the corresponding regions of the bovine repetitive sequences and the polyoma virus genome (23). The homology observed

Figure 6. Comparison of the bovine repetitive sequences with those of other mammalian species. The segment of the bovine repetitive sequences that exhibits homology with other mammalian repetitive sequences (the underlined segment of the consensus sequence in Fig. 3) is displayed in comparison with the human (8-10; see also the accompanying paper) and rat counterparts (12,13). Gaps (-) have been inserted by inspection to achieve maximum homology. Matches are marked by vertical lines.

between the bovine repetitive sequences and the viral sequences at or near the replication origins supports the view that interspersed repetitive sequences are involved in DNA replication.

Evidence has been obtained to indicate that some of the interspersed repetitive DNA sequences are transcribed into low molecular weight RNAs in vitro (15,24-26) as well as in vivo (8,15,26). Furthermore, the middle one-third of the 120 bp conserved unit of the bovine repetitive sequences (positions 39-84 of the consensus sequence in Fig. 3) exhibits homology with the sequence identified as a possible intragenic RNA polymerase III recognition site (27). These findings suggest that the bovine repetitive DNA sequences are also transcribed. The resulting transcripts may interact with the repetitive DNA segments in the ACTH-β-LPH precursor gene region. Alternatively, these transcripts may interact with the ACTH-β-LPH precursor hnRNA, assuming that the bovine repetitive sequences are transcribed in the

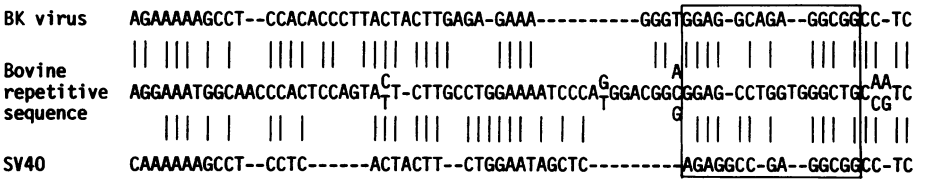


Figure 7. Comparison of the bovine repetitive sequences with the regions encompassing the replication origins of human BK virus and simian virus 40 (SV40). The overlined segment of the bovine consensus sequence shown in Fig. 3 is compared with the viral sequences, the sources of which are given in the text. Gaps (-) have been inserted by inspection to achieve maximum homology. Matches are marked by vertical lines. The 14 bp common sequence is boxed.

direction assigned by comparison with the human Alu family sequences (see Fig. 1 and Fig. 6). We have previously shown that bovine pituitary nuclear RNA contains the primary transcript of the entire ACTH-β-LPH precursor gene (3). Thus, it is possible that interspersed repetitive sequences may be involved in the transcription of the ACTH-β-LPH precursor gene or in the processing of its transcript. Another concept is that such sequences may be examples of selfish or parasite DNA which makes no specific contribution to gene expression (28,29).

ACKNOWLEDGEMENTS

We thank Dr. Takashi Miyata and Dr. Tatsuo Ooi for helpful discussions. This investigation was supported in part by research grants from the Ministry of Education, Science and Culture of Japan, the Institute of Physical and Chemical Research, the Mitsubishi Foundation and the Japanese Foundation of Metabolism and Diseases.

REFERENCES

1. Britten, R.J. and Kohne, D.E. (1968) Science 161, 529-540.
2. Davidson, E.H., Graham, D.E., Neufeld, B.R., Chamberlin, M.E., Amenson, C.S., Hough, B.R. and Britten, R.J. (1973) Cold Spring Harbor Symp. Quant. Biol. 38, 295-301.
3. Nakanishi, S., Teranishi, Y., Watanabe, Y., Notake, M., Noda, M., Kakidani, H., Jingami, H. and Numa, S. (1981) Eur. J. Biochem. 115, 429-438.
4. Nakanishi, S., Teranishi, Y., Noda, M., Notake, M., Watanabe, Y., Kakidani, H., Jingami, H. and Numa, S. (1980) Nature 287, 752-755.
5. Polsky, F., Edgell, M.H., Seidman, J.G. and Leder, P. (1978) Anal. Biochem. 87, 397-410.
6. Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499-560.
7. Britten, R.J., Graham, D.E. and Neufeld, B.R. (1974) Methods Enzymol. 29, 363-418.
8. Pan, J., Elder, J.T., Duncan, C.H. and Weissman, S.M. (1981) Nucleic Acids Res. 9, 1151-1170.
9. Baralle, F.E., Shoulders, C.C., Goodbourn, S., Jeffreys, A. and Proudfoot, N.J. (1980) Nucleic Acids Res. 8, 4393-4404.
10. Bell, G.I., Pictet, R. and Rutter, W.J. (1980) Nucleic Acids Res. 8, 4091-4109.
11. Krayev, A.S., Kramerov, D.A., Skryabin, K.G., Ryskov, A.P., Bayev, A.A. and Georgiev, G.P. (1980) Nucleic Acids Res. 8, 1201-1215.
12. Page, G.S., Smith, S. and Goodman, H.M. (1981) Nucleic Acids Res. 9, 2087-2104.
13. Barta, A., Richards, R.I., Baxter, J.D. and Shine, J. (1981) Proc. Natl. Acad. Sci. USA 78, 4867-4871.
14. Haynes, S.R., Toomey, T.P., Leinwand, L. and Jelinek, W.R. (1981) Mol. Cell. Biol. 1, 573-583.
15. Haynes, S.R. and Jelinek, W.R. (1981) Proc. Natl. Acad. Sci. USA 78, 6130-6134.
16. Calos, M.P. and Miller, J.H. (1980) Cell 20, 579-595.

17. Farabaugh, P.J. and Fink, G.R. (1980) Nature 286, 352-356.
18. Dunsmuir, P., Brorein, W.J.Jr., Simon, M.A. and Rubin, G.M. (1980) Cell 21, 575-579.
19. Jagadeeswaran, P., Forget, B.G. and Weissman, S.M. (1981) Cell 26, 141-142.
20. Jelinek, W.R., Toomy, T.P., Leinwand, L., Duncan, C.H., Biro, P.A., Choudary, P.V., Weissman, S.M., Rubin, C.M., Houck, C.M., Deininger, P.L. and Schmid, C.W. (1980) Proc. Natl. Acad. Sci. USA 77, 1398-1402.
21. Yang, R.C.A. and Wu, R., (1979) Science 206, 456-462.
22. Reddy, V.B., Thimmappaya, B., Dhar, R., Subramanian, K.N., Zain, B.S., Pan, J., Ghosh, P.K., Celma, M.L. and Weissman, S.M. (1978) Science 200, 494-502.
23. Friedmann, T., LaPorte, P. and Esty, A. (1978) J. Biol. Chem. 253, 6561-6567.
24. Duncan, C., Biro, P.A., Choudary, P.V., Elder, J.T., Wang, R.R.C., Forget, B.G., DeRiel, J.K. and Weissman, S.M. (1979) Proc. Natl. Acad. Sci. USA 76, 5095-5099.
25. Elder, J.T., Pan, J., Duncan, C.H. and Weissman, S.M. (1981) Nucleic Acids Res. 9, 1171-1189.
26. Duncan, C.H., Jagadeeswaran, P., Wang, R.R.C. and Weissman, S.M. (1981) Gene 13, 185-196.
27. Doolittle, W.F. (1981) FEBS Lett. 126, 147-149.
28. Doolittle, W.F. and Sapienza, C. (1980) Nature 284, 601-603.
29. Orgel, L.E. and Crick. F.H.C. (1980) Nature 284, 604-607.