Vol. 177, No. 1, 1991 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
May 31, 1991 Pages 120-124

EVOLUTION OF ALU REPEATS SURROUNDING THE HUMAN FERREDOXIN GENE

Bon-chu Chung[+,*] and Chi-Yao Chang[+,#]

[+]Institute of Molecular Biology, Academia Sinica
Nankang, Taipei, Taiwan, 11529, Republic of China

[#]Institute of Life Sciences, Tsing-Hwa University
Hsin-Chu, Taiwan, Republic of China

Received March 27, 1991

Ferredoxin is an iron-sulfur protein that serves as an electron carrier for the mitochondrial oxidation/reduction system. During the characterization of the human ferredoxin gene, we have identified three Alu sequences surrounding it. When these Alu sequences were compared with others, all three of them are more related to the consensus Alu than the 7SL gene, the progenitor of the Alu family. It suggests that they are members of the modern Alu family. Their sequences differ from the Alu consensus sequence by about 5%, indicating that they were inserted into the chromosome about 35 million years ago.

Ferredoxin is an iron-sulfur containing protein that serves as an electron donor in the oxidation/reduction reactions catalyzed by the mitochondrial cytochromes P450 (1). The mitochondrial electron-transport system is present in the adrenal, liver, and kidney for the synthesis and metabolism of steroids, vitamin D, and bile acids (2). In this system, electrons are donated by NADPH via $NADP^+$-ferredoxin-dependent oxidoreductase and ferredoxin to the terminal cytochrome P450, which in turn transfers electrons to substrates to form products (1).

Ferredoxin has been found most abundantly in the adrenal, hence it is termed adrenodoxin (3). In the adrenal, synthesis of adrenodoxin is stimulated by adrenocorticotropin using cAMP as an intracellular messenger (4, 5). In other steroidogenic tissues such as placenta and ovary, ferredoxin synthesis is stimulated by gonadotropins via cAMP (6-8). Therefore cAMP plays an important role in the stimulation of ferredoxin transcription (9). In order to understand that structural and regulational aspects of the ferredoxin gene, we have cloned and characterized

*To whom all correspondence should be addressed.

0006-291X/91 $1.50

the human ferredoxin genes (10). In addition to the active gene, the human ferredoxin gene family contains two pseudogenes (11). We have studied the structure of the flanking regions of the ferredoxin gene and found three Alu repeats surrounding it.

Human Alu repeats are a family of DNA sequences roughly 300 bp in length that has been repeated 500,000 times in the human genome (12). Although some Alu sequences have been identified in cDNA's and hence are transcribed by either RNA polymerase II or III (13, 14), most of the Alu sequences are probably present as nonfunctional pseudogenes (12).

All Alu sequences are derived from a common progenitor 7SL gene (15), a series of mutations have accumulated during the 30 million years that Alu sequences have evolved. The 500,000 Alu sequences can be classified into many subfamilies that were derived from distinct sets of founder sequences (16-20). Jurka and Smith have classified Alu sequences into the 7SL like Alu-J and the other modern Alu-S families (21). Our present study on the Alu repeats surrounding the human ferredoxin genes suggests that these Alu's belong to the modern Alu family.

Materials and Methods

The cloning, hybridization, and subcloning experiments used standard procedures as reported previously (11). Dideoxy sequencing procedures were performed from M13 templates using a Sequenase kit purchased from USB. DNA sequences were analyzed using the program purchased from GCG (Madison, WI).

Results and Discussion

1. Characterization and Organization of Alu Repeats in the Human Ferredoxin Gene

We have reported the cloning and organization of members of the human ferredoxin gene (11). This gene family contains two pseudogenes h2 and h3 in additional to the active genes (11). During characterization of the sequences flanking the ferredoxin genes, we have identified three Alu sequences. One is present at about 1000 bp upstream from the active gene h5, termed h5-Alu. The others are each located about 100 bp upstream and downstream from the h3 pseudogene, termed h3-Alu's. All three Alu sequences contain poly(A) tails at their 3'-end and are flanked by direct repeats, therefore they arose through RNA-mediated transposition events.

2. These Alu sequences belong to the Modern Alu family

To study the mutational rates of the h3- and h5-Alu sequences, we compared them with the progenitor 7SL gene and a consensus Alu sequence that was generated by Jurka and Smith (Fig. 1). Jurka and Smith have classified all Alu

```
7SL          GCCGGGCGCGGTGG  CGCGTGCCTGTAGTC  CCAGCTACT-CGGGA
Alu-cons     GGCCGGGCGCGGTGG  CTCACGCCTGTAATC  CCAGC-ACTTTGGGA
H3-5'           T    A A         T T
H3-3'            A   AT A -

7SL          GGCTGAGGCTGGAGG  ATCGCTTGAGTCCAG  GAGTTC----CCAGC
Alu-cons     GGCCGAGGCGGGCGG  ATCACCTGAGGTCAG  GAGTTCGAGACCAGC
H3-5'        A   A    AT          T     C
H3-3'                C  TA        --T A             A

7SL          CTGGGCAACATAGCG  AGACCCCGTCTCT
Alu-cons     CTGGCCAACATGGTG  AAACCCCGTCTCTAC  TAAAAATACAAAAA-
H3-5'                -        T                      G
H3-3'                            T                          A

7SL             GCCGGGCGCGGT  GGCGCGTGCCTGTAG  TCCCAGCTACTCGGG
Alu-cons     TTAGCCGGGCGTGGT  GGCGCGCGCCTGTAA  TCCCAGCTACTCGGG
H3-5'           A- T          T  ATAT              A G  T
H3-3'          T   T           G                       G
H5            *TG C T          A A        G     T

7SL          AGGCTGAGGCTGGAG  GATCGCTTGAGTCCA  GGAGTTCTGGGCTGT
Alu-cons     AGGCTGAGGCAGGAG  AATCGCTTGAACCCG  GGAGGCGGAGGTTGC
H3-5'                T       G   --------  ------------
H3-3'        G T             T          A            C
H5                            A        T

7SL          AGTGCGCCTGTGA--  --GCCACTGCACTCC  AGCCTGGGCAACATA
Alu-cons     AGTGAGCC-GAGATC  GCGCCACTGCACTCC  AGCCTGGGCGACAGA
H3-5'           A     A       A                        A
H3-3'                    T   T T              A
H5                       T   T          A           AGA T

7SL          GCGAGACCCCGTCTC
Alu-cons     GCGAGACTCCGTCTC
H3-5'          A      --
H3-3'         A  A    A
H5            A     C  A G
```

Fig. 1. Comparison of alu sequences surrounding the ferredoxin gene family with consensus Alu sequence and the homologous regions of human 7SL DNA. Sequences of h3-5', h3-3', and h5-Alu are shown only when they are different from the consensus. Dashes indicate base deletions. The asterisk marks the start site of h5-alu.

sequences into the ancient Alu-J and modern Alu-S subfamilies (1988). Fifteen positions that are typically different between the ancient 7SL DNA and the consensus Alu are used as diagnostic positions for comparison. About 83% of the modern Alu-S sequences would be identical to the consensus at these 15 positions and conversely, 69% of the the J-Class Alu sequence would be identical to the 7SL DNA at these positions (Table 1). The h3-5', h3-3', and h5-Alu sequences have 60%, 73%, and 71% identity to the consensus in these diagnostic positions and only 13%, 13%, and 28% identity to the 7SL DNA. Therefore we conclude that they belong to the modern Alu-S subfamily (Table 1).

Table 1. Frequency of bases identical to that of consensus at diagnostic positions in ferredoxin Alu sequences and the major Alu subfamilies

Alu	Frequency Match with Consensus	Frequency Match with 7SL DNA	Frequency Mismatch
H3-5′	0.60 (9/15)	0.13 (2/15)	0.26 (4/15)
H3-3′	0.73 (11/15)	0.13 (2/15)	0.13 (2/15)
H5	0.71 (5/7)	0.28 (2/7)	0 (0/7)
J-Class*	0.17	0.69	0.16
S-Class*	0.83	0.09	0.08

Sequences of h3- and h5-Alu were compared with the consensus and 7SL DNA at diagnostic positions that are underlined in Fig. 1. The percentage of bases that is identical to both the consensus, the 7SL DNA, or none of the two is listed.

*The percentage of Alu-S and Alu-J sequences that are identical to the consensus, the 7SL DNA, or nothing has been calculated from the numbers taken from Jurka and Smith (1988).

The modern Alu sequence evolves at a neutral drift rate with the exception of the 25 CpG dinucleotides which are mutational hot spots (Britten et al., 1988; Coulondre et al., 1978). Excluding CpG mutations, both h3-Alu sequences accumulated 15 nucleotide differences from the consensus, while h5-Alu has 9 nucleotide differences in its half sequence. This translates into a mutation rate of 5.3 and 5.7%. Taking the drift rate as 0.15% for the non-CpG Alu repeat (Britten, 1986; Britten et al., 1988), they were calculated to have been inserted into the chromosomes about 35 Myr ago. Britten *et al.* classified modern Alu sequences into four classes according to their order of appearance (1988). H5-Alu and h3-Alu would fall into a class between III and IV following such classification.

Acknowledgments

This work was supported by National Science Council, #NSC-78-0203-B001-11 and Academia Sinica, Republic of China.

References

1. Nakamura, Y., and Otsuka, H. (1966) Biochim. Biophys. Acta, 122, 34-42.
2. Waterman, M. R., John, M. E., and Simpson, E. R. (1986) In Ortiz de Montellano, P. R. (ed), Cytochrome P-450. Plenum Press, New York and London, pp.345-386.
3. Padmanabhan, R., and Kimura,T. (1970) J. Biol. Chem., 245, 2469-2475.
4. Kramer, R. E., Anderson, C. M., Peterson, J. A., Simpson, E. R., and Waterman, M. R. (1982) J. Biol. Chem., 257, 14921-14925.
5. Simpson, E. R., and Waterman, M. R. (1988) Ann. Rev. Physiol., 50, 427-440.

6. Picado-Leonard, J., Voutilainen, R., Kao, L.–c., Chung, B.–c., Strauss, J. F. III, and Miller, W. L. (1988) J. Biol. Chem. 263, 3240-3244.

7. Golos, T. G., Miller, W. L. and Strauss, J . F. III (1987) J. Clin. Invest. 80, 897-899.

8. Anderson, C. M., and Mendelson, C. R. (1985) Arch. Biochem. Biophys., 238, 378-387.

9. John, M. E., John, M. C., Boggaram, V., Simpson, E. R., and Waterman, M. R. (1986) Proc. Natl. Acad. Sci. USA, 83, 4715-4719.

10. Chang, C.–Y., Wu, D.–A., Lai, C.–C., Miller, W. L. and Chung, B.–c. (1988) DNA, 7, 609-615.

11. Chang, C.–Y., Wu, D.–A., Mohandas, T. K., and Chung, B.–c. (1990) DNA Cell Biol., 9, 205-212.

12. Schmid, C. W., and Jelinek, W. R. (1982) Science, 216, 1065-1070.

13. Chuang, L.–M., Lin, B. J., Lee, S.–C., Tai, T.–Y., and Chen, D.–S. (1988) Biochem. Biophys. Res. Commun. 156, 1287-1292.

14. Hess, J., Perez-Stable, C., Wu, G. J., Weir, B., Tinoco, I. Jr., and Shen, C.–K. J. (1985) J. Mol. Biol. 184, 7-21.

15. Ullu, E., and Tschudi, C. (1984) Nature, 312, 171-172.

16. Slagel, V., Flemington, E., Traina-Dorge, V., Bradshaw, H. and Deininger, P. (1987) Mol. Biol. Evol., 4, 19-29.

17. Willard, C., Nguyen, H. T., and Schmid, C. W. (1987) J. Mol. Evol., 26, 180-186.

18. Britten, R. J., Baron, W. F., Stout, D. B., and Davidson, E. H. (1988) Proc. Natl. Acad. Sci. USA, 85, 4770-4774.

19. Labuda, D., and Striker, G. (1989) Nucl. Acids Res., 17, 2477-2491.

20. Matera, A. G., Hellmann, U., Hintz, M. F., and Schmid, C. W. (1990) Nucl. Acids Res., 18, 6019-6023.

21. Jurka, J., and Smith, T. (1988) Proc. Natl. Acad. Sci. USA, 85, 4775-4778.

22. Coulondre, C., Miller, J. H., Farabaugh, P. J., and Gilbert, W. (1978) Nature, 274, 775-780.

23. Britten, R. J. (1986) Science, 231, 1393-1398.