Volume 13 Number 5 1985 Nucleic Acids Research

Molecular structures of mitochondrial-DNA-like sequences in human nuclear DNA

Hisayuki Nomiyama, Makoto Fukuda, Shoji Wakasugi, Teruhisa Tsuzuki and Kazunori Shimada

Department of Biochemistry, Kumamoto University Medical School, Kumamoto 860, Japan

Received 19 December 1984; Accepted 8 February 1985

ABSTRACT

Two lambda phage clones carrying mitochondrial-DNA-like (mtDNA-like) sequences isolated from a human gene library were named Lm E-1 and Lm C-2, and their DNA structures were characterized. Lm E-1 contains about 0.4 kb DNA homologous to the 5' portion of the mitochondrial 16S ribosomal RNA (rRNA) gene and Lm C-2, a 1.6 kb DNA homologous to the 3' portion of the 12S rRNA gene and to almost all of the 16S rRNA gene. Comparisons of their nucleotide sequences with those of the corresponding regions of the human mtDNA revealed no detectable DNA rearrangement and their homologies to the human mtDNA are 84% and 80%, respectively. There are neither terminal repeats in the nuclear mtDNA-like sequences nor duplications of the nuclear DNAs flanking the mtDNA-like sequences. Evolutionary relationship between these two human nuclear mtDNA-like sequences and the human and bovine mtDNAs is discussed.

INTRODUCTION

Many eukaryotic nuclear DNAs have recently been revealed to contain mitochondrial-DNA-like (mtDNA-like) sequences. These sequences have been found in the nuclear DNAs of yeast[1], locust[2], Podospora[3], sea urchin[4], maize[5], rat[6] and human[7]. Moreover, mtDNA of ascomycete Neurospora was found to contain a DNA sequence homologous to the nuclear gene coding for one of the mitochondrial ATPase subunits[8]. All these findings strongly suggest that in eukaryotic cells, movement of DNAs between nuclei and organella occurred during evolution.

One of the mtDNA-like sequences present in the nuclear DNA of yeast[1] was found to contain a rearranged configuration of a part of the mtDNA sequence. This structure is assumed to have originated from a petite mtDNA, which is amplified and often constitutes rearranged portions of the mtDNA. In the ascomy-

cete Podospora, excision and amplification of a particular mtDNA sequence, termed senDNA, are associated with mycelial senescence[9,10]. Wright and Cummings speculate that these amplified senDNAs are transferred from mitochondria to nuclei and integrate into the nuclear DNA during senescence[4]. However, neither structures of the mtDNA-like sequences present in mammalian nuclear DNAs nor the mechanisms of their integration into nuclear DNAs have been reported.

We have now elucidated nucleotide sequences of two nuclear mtDNA-like sequences isolated from a human gene library. These sequences are homologous to the mitochondrial ribosomal RNA (rRNA) gene regions, and contain no detectable mtDNA rearrangement, although they have many base-substitutions. We detected no characteristic structures at the 5' and 3' termini of these mtDNA-like regions and our results suggest that the mtDNA fragments integrate into the human nuclear DNA by an illegitimate recombination. Evolutionary relationship between these two human nuclear mtDNA-like sequences and the human and bovine mtDNAs is also discussed.

MATERIALS AND METHODS

Isolation of phage clones, Lm C-2 and Lm E-1, was described in a previous paper[7]. The clones were isolated from a human gene library using a pHIG-2 DNA insert as a probe. Structure of the pHIG-2 DNA[11] was described and it contains an insert of cDNA prepared on the mitochondrial 16S rRNA. Lm E-1 DNA was digested with EcoRI, and an 8.0 kb DNA fragment that hybridizes with pHIG-2, was subcloned into the EcoRI site of pBR322. Southern blotting, prehybridization, hybridization and washing conditions were as described[12]. Nucleotide sequences were determined by the method of Maxam and Gilbert[13], and were analyzed by the GENAS system[14] at the Kyushu University Computer Center. Enzymes and radionucleotides used were as described[12].

Propagation of E. coli cells carrying recombinant plasmids was carried out in accordance with the guidelines for recombinant DNA research issued by the Ministry of Education, Science and Culture, Japan.

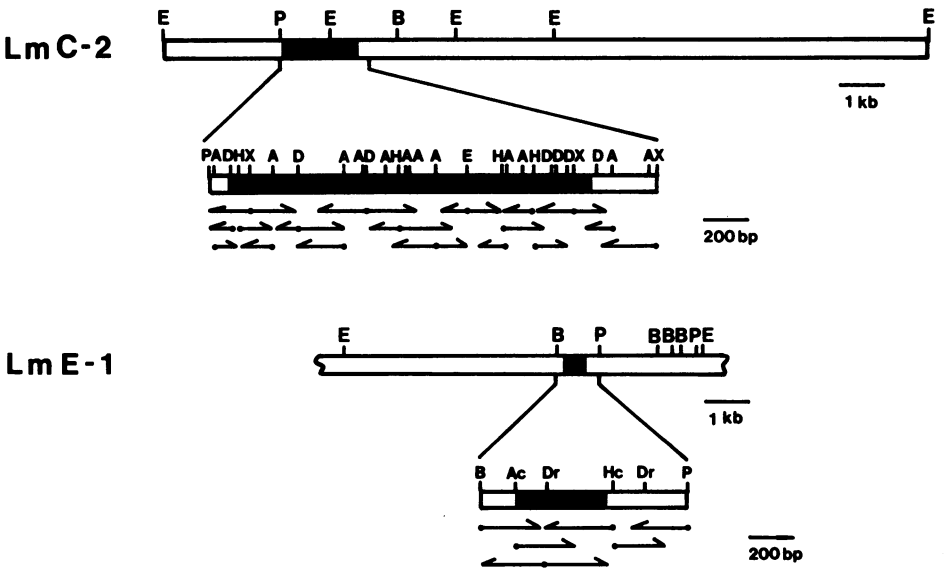

Figure 1. Restriction maps of the two genomic clones. The mtDNA-like sequences and the flanking nuclear sequences are indicated by solid and open boxes, respectively. Arrows beneath the enlarged maps represent the sequencing strategy. Each arrow indicates the length and direction of the sequences. The end-labeled sites are marked with solid circles. Restriction sites: EcoRI (E), PstI (P), BglII (B), AluI (A), DdeI (D), HaeIII (H), XbaI (X), AccI (Ac), DraI (Dr), HincII (Hc).

RESULTS AND DISCUSSION

Nucleotide sequence analysis of the two nuclear mtDNA-like regions

We isolated six phage clones from a human gene library, using a cDNA prepared on the mitochondrial 16S rRNA as a probe[7]. To elucidate the molecular structures of the nuclear mtDNA-like sequences, we analyzed nucleotide sequences of mtDNA-like regions carried by two phage clones, Lm C-2 and Lm E-1. The mtDNA-like regions were determined by Southern blot analysis, and their restriction maps and the regions homologous to the human mitochondrial rRNA genes are schematically summarized in Figure 1. Nucleotide sequences of the mtDNA-like regions carried by these two clones were determined by the strategy shown in this figure, and were compared with that of the human mtDNA[15] (Fig. 2). The mtDNA-like region of Lm E-1 was designated E-1

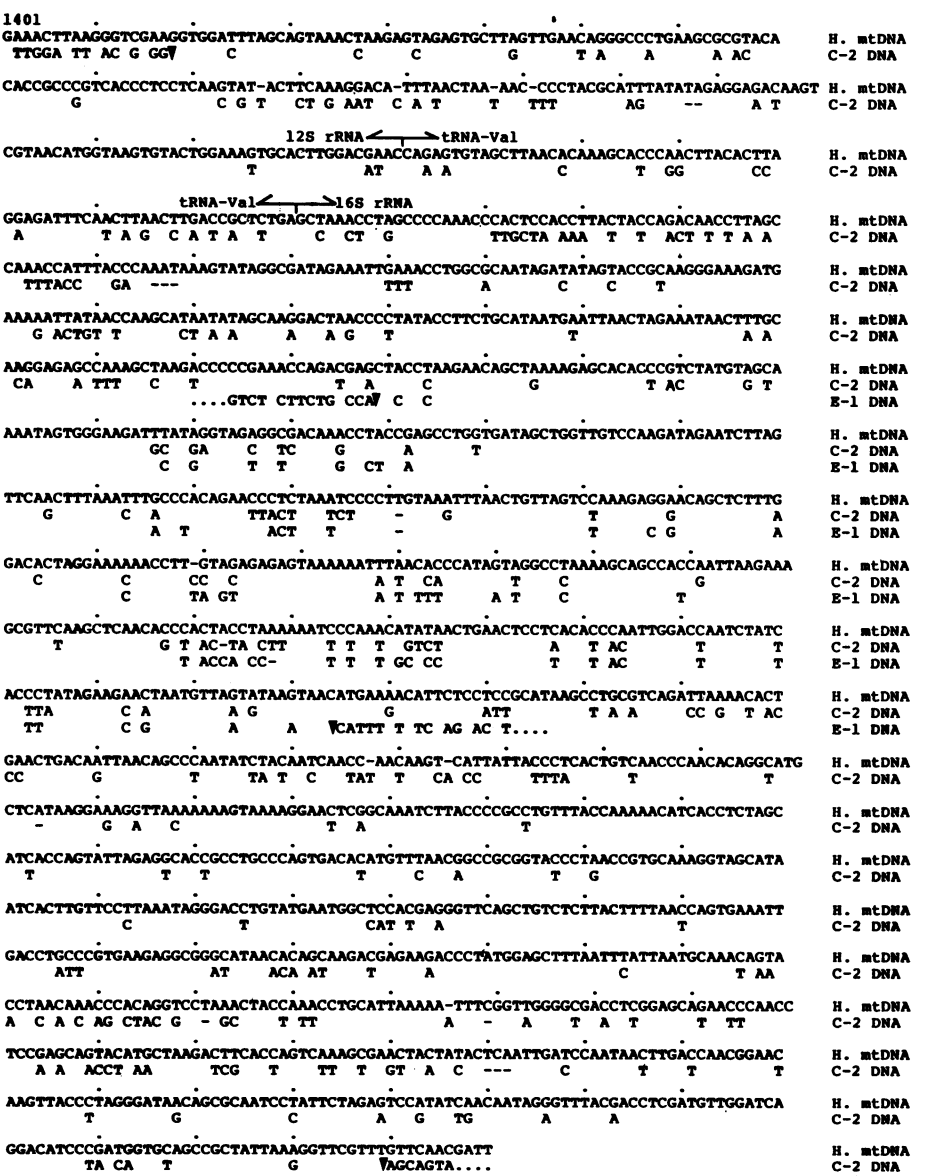

Figure 2. Nucleotide sequences of the two nuclear mtDNA-like regions and the comparison of their sequences with that of the human mtDNA[15] (H. mtDNA). Only those bases that differ from human mtDNA sequence are shown in the lines of the E-1 and C-2 DNA sequences. Arrowheads indicate the 5' and 3' ends of the mtDNA-like sequences. Additions of one or a few nucleotides to align sequences for maximum homology are indicated by (-) marks.

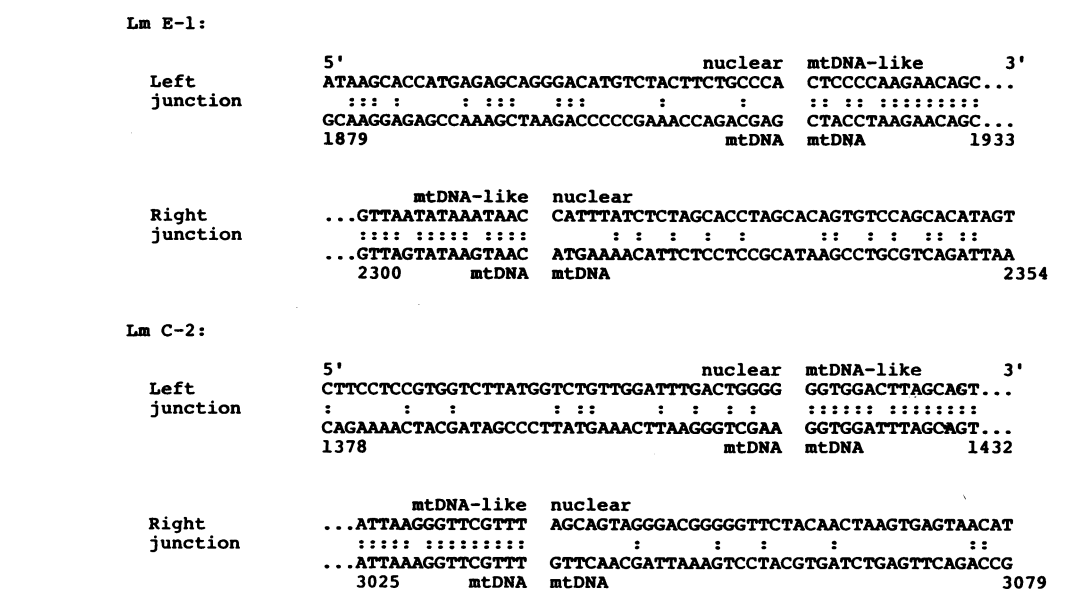

Figure 3. Nucleotide sequences of the junctions between the mtDNA-like sequences and the flanking nuclear sequences. Aligned below the mtDNA-like sequences is that of the human mtDNA[15]. Colons indicate the matched bases.

DNA and it corresponds to the 5' portion of the 16S rRNA gene, and that of Lm C-2, designated C-2 DNA, corresponds to the 3' portion of the 12S rRNA gene and to almost all of the 16S rRNA gene. Total lengths of the E-1 and C-2 DNAs are about 0.4 and 1.6 kb, respectively.

To elucidate the mechanisms of their integrations, we first examined structures of the terminal regions of these two nuclear mtDNA-like sequences, and found that their terminal regions do not contain characteristic structures such as direct or inverted repeats (Fig. 2). Moreover, there is no duplication of the nuclear DNA sequences at the junctions of the mtDNA-like and the nuclear DNA sequences (Figs. 2 & 3). We aligned sequences of the left and right junctions of the two nuclear mtDNA-like regions with that of the corresponding regions of the human mtDNA and found no significant homology between these regions (Fig. 3). There is also no sequence homology between the two nuclear target sites present in the Lm E-1 and Lm C-2 DNAs, indicating that

these mtDNA fragments did not integrate into specific sites on the nuclear DNA (Fig. 3).

In mammalian cells, recombinations are frequently observed even between DNAs that show little homology (illegitimate recombination). Mechanisms of integration of DNA tumor viruses into host nuclear DNAs have been extensively studied as model systems for the analysis of recombination mechanisms within mammalian cells[16-21]. These studies have shown that there is no specific target site for viral DNA integration and that there is no apparent DNA homology between the interacting sites of the viral and of the host nuclear DNAs so far examined. Although in some cases, a short stretch of homology is observed around junctions of viral and host nuclear DNAs[17-19,21], such patchy homology may have little statistical significance in recombination[22]. From the elucidated DNA structures of the two nuclear mtDNA-like regions, we speculate that the integration of these mtDNA fragments into nuclear DNA is mediated by mechanisms similar to that of viral DNA integration into host nuclear DNAs.

The manner in which mtDNAs present in mitochondria are transported into nuclei has not been clarified, however, there are implications in the observations that mitochondria are sometimes present within the nuclei of human and other mammalian cells[23-26]. These intranuclear mitochondria may soon be degraded within nuclei, the mtDNAs subsequently released and integrate into the nuclear DNA.

As shown in Figure 2, there is no detectable mtDNA rearrangement in the E-1 and C-2 DNAs, and their homologies to human mtDNA are 84% and 80%, respectively. Although there are several insertions and deletions of one or a few bases in these mtDNA-like sequences, divergence of their sequences from that of the current human mtDNA is mainly due to base-substitutions, and which are almost randomly spread out over the E-1 and C-2 DNAs.

We estimated the approximate time of integration of the mtDNA fragments into nuclear DNA, on the following assumptions. First, these two nuclear mtDNA-like sequences are similar to pseudogenes, because they apparently diverged from human mtDNA, and because neither contain complete mitochondrial rRNA genes. Second, mutations in pseudogenes are fixed in germ lines of

mammals at the rate of 1.26% per million years[27]. Since this mutation rate is also comparable to that calculated for genes present on mtDNAs (0.5 to 2.0%)[28], we used this value to assess the evolutionary relationship between the nuclear mtDNA-like sequences and the human mtDNA, and calculated that the E-1 and C-2 DNAs were transferred from mitochondria into nuclei about 12 and 15 millions years ago, respectively.

Evolutionary relationship between the nuclear mtDNA-like sequences and the human and bovine mtDNAs

Brown et al. compared nucleotide sequences of mtDNAs among closely related species and found that transitional base-substitutions (G⇌A and T⇌C) are predominant over transversional ones (purine⇌pyrimidine), while proportions of transitions are apparently lower in the mtDNAs among less closely related species[28].

To acquire more information for estimation of the time of integration of the mtDNA-like sequences into the human nuclear DNA, we aligned mtDNA sequence of bovine[29] with that of human mtDNA and of the two nuclear mtDNA-like sequences (Fig. 4), and compared properties of base-substitutions present among these DNAs. The bovine species probably diverged from humans about 80 million years ago. The 5' end region of the 16S rRNA gene of bovine mtDNA shows 75% homology to the corresponding region of the human mtDNA. Within this region, proportions of transitional base-substitutions are 63% in the E-1 and 58% in the C-2 DNAs, respectively, while that in the bovine mtDNA is 44%. The proportions of transitional base-substitutions are in good correlation with the percentages of homologies observed among these regions (Table 1).

These results support the idea that the E-1 and C-2 DNAs are transferred from mitochondria to nuclei sometime after the human and bovine diverged. Accordingly, the proportion of transitions among base-substitutions seems to be one useful index to analyze the evolutionary distance between a mtDNA and its nuclear counterparts. However, it is apparent from these analyses that, to estimate the time of integration of mtDNA more precisely, nucleotide sequences of the human nuclear mtDNA-like regions should be compared with those of the mtDNAs in species

Figure 4. Comparison of the two nuclear mtDNA-like sequences and bovine mtDNA[29] (B. mtDNA) with human mtDNA[15] (H. mtDNA). Common base-substitutions present in the two nuclear mtDNA-like and bovine mtDNA sequences are indicated by arrowheads. Positions of insertions of one or several bases in bovine mtDNA are indicated by arrows.

more closely related to human. These studies are underway.

When we compared the bovine mtDNA and the human nuclear mtDNA-like sequences with the human mtDNA, we noticed that among

Table 1. Sequence homologies of the two nuclear mtDNA-like sequences and bovine mtDNA compared with the 5' terminal region of the 16S rRNA gene of human mtDNA and proportions of transitions in the base-substitutions.

	Homology to Human mtDNA	Proportion of Transitions
E-1 DNA	84%	63%
C-2 DNA	81%	58%
Bovine mtDNA	75%	44%

Homologies are calculated from Fig. 4.

many base-substitutions, the fifteen indicated by arrowheads in Figure 4, are all substituted by exactly the same bases. We assume that some of these common base-substitutions represent the base sequences present in the mtDNA at the time of integration into the nuclear genome. After their integrations, these bases are conserved only in the nuclear mtDNA-like sequences, while the human mtDNA evolved independently and mutations occurred in these bases.

ACKNOWLEDGEMENTS

We thank Mariko Ohara of Kyushu University for reading the manuscript. This work was supported by grants from the Ministry of Education, Science and Culture of Japan.

REFERENCES

1. Farrelly, F. and Butow, R. A. (1983) Nature **301**, 296-301.
2. Gellissen, G., Bradfield, J. Y., White, B. N. and Wyatt, G. R. (1983) Nature **301**, 631-634.
3. Wright, R. M. and Cummings, D. J. (1983) Nature **302**, 86-88.
4. Jacobs, H. T., Posakony, J. W., Grula, J. W., Roberts, J. W., Xin, J-H., Britten, R. J. and Davidson, E. H. (1983) J. Mol. Biol. **165**, 609-632.
5. Kemble, R. J., Mans, R. J., Gabay-Laughnan, S. and Laughnan, J. R. (1983) Nature **304**, 744-747.
6. Hadler, H. I., Dimitrijevic, B. and Mahalingam, R. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 6495-6499.
7. Tsuzuki, T., Nomiyama, H., Setoyama, C., Maeda, S. and Shimada, K. (1983) Gene **25**, 223-229.
8. van den Boogaart, P., Samallo, J. and Agsteribbe, E. (1982) Nature **298**, 187-189.
9. Wright, R. M., Horrum, M. A. and Cummings, D. J. (1982) Cell **29**, 505-515.
10. Cummings, D. J. and Wright, R. M. (1983) Nucleic Acids Res. **11**, 2111-2119.
11. Tsuzuki, T., Nomiyama, H., Setoyama, C., Maeda, S., Shimada, K. and Pestka, S. (1983) Biochem. Biophys. Res. Commun. **114**, 670-676.
12. Nomiyama, H., Tsuzuki, T., Wakasugi, S., Fukuda, M. and Shimada, K. (1984) Nucleic Acids Res. **12**, 5225-5234.
13. Maxam, A. M. and Gilbert, W. (1980) Meth. Enzym. **65**, 499-560.
14. Kuhara, S., Matsuo, F., Futamura, S., Fujita, A., Shinohara, T., Takagi, T. and Sakaki, Y. (1984) Nucleic Acids Res. **12**, 89-99.
15. Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R. and Young, I. G. (1981) Nature **290**, 457-465.
16. Stringer, J. R. (1981) J. Virol. **38**, 671-679.

17. Gutai, M. W. (1981) Virology **109**, 344-352.
18. Stringer, J. R. (1982) Nature **296**, 363-366.
19. Bullock, P., Forrester, W. and Botchan, M. (1984) J. Mol. Biol. **174**, 55-84.
20. Mounts, P. and Kelly, T. J. Jr (1984) J. Mol. Biol. **177**, 431-460.
21. Schulz, M. and Doerfler, W. (1984) Nucleic Acids Res. **12**, 4959-4976.
22. Savageau, M. A., Metter, R. and Brockman, W. W. (1983) Nucleic Acids Res. **11**, 6559-6570.
23. Brandes, D., Schofield, B. H., Anton, E. (1965) Science **149**, 1373-1374.
24. Klug, H. (1966) Naturwissenschaften **53**, 339.
25. Bloom, G. D. (1967) J. Cell Biol. **35**, 266-268.
26. Oliva, H., Valle, A., Diaz Flores, L. and Rivas, M. C. (1973) Virchows Arch. Abt. B Zellpath. **12**, 189-194.
27. Miyata, T. and Yasunaga, T. (1981) Proc. Natl. Acad. Sci. USA **78**, 450-453.
28. Brown, W. M., Prager, E. M., Wang, A. and Wilson, A. C. (1982) J. Mol. Evol. **18**, 225-239.
29. Anderson, S., de Bruijn, M. H. L., Coulson, A. R., Eperon, I. C., Sanger, F. and Young I. G. (1982) J. Mol. Biol. **156**, 683-717.