Volume 11 Number 19 1983 Nucleic Acids Research

Rat mitochondrial DNA polymorphism: sequence analysis of a hypervariable site for insertions/deletions

Gregory G.Brown* and Lee J.DesRosiers

Department of Biology, McGill University, Montreal, Quebec H3A 1B1, Canada

Received 18 July 1983; Revised and Accepted 6 September 1983

ABSTRACT

Three closely related variants of rat (Rattus norvegicus) mtDNA have been shown to differ in the number of T residues found in a run of Ts (light strand) which spans the junction between the $tRNA^{Cys}$ and $tRNA^{Tyr}$ genes. The number of Ts in the repeat varies from 6 to 8 in these DNAs. Another, less closely related, R. norvegicus variant has a run of 5 Ts at this site and in the related species, Rattus rattus, a run of 4 Ts is found. In R. norvegicus mtDNA runs of 5 As and 5 Gs are found just to the 3' side of the variable T repeat, and it is suggested that the three runs of repeated nucleotides may stabilize heteroduplexes which result from strand slippage and which give rise to the insertions and/or deletions. Among 17 mtDNA clones derived from an individual with the 8T repeat, one clone was found which possessed a 9T repeat. This variant may represent an additional DNA type originally present within the individual.

INTRODUCTION

An extraordinary amount of sequence variation is found in the mitochondrial genomes of higher animals (1-5). Comparisons of mtDNA sequences from closely related species (6,7) shows that most of this rapid evolutionary change results from a high rate of occurrence of silent, transition-type base substitutions. Although sequence diversity due to the deletion or insertion of bases is common in nuclear genes, such changes have not been frequently observed in animal mtDNA. That they occur is evident from comparisons of mtDNA sequences of distantly related mammals (8,9), but only a few examples of changes of this sort have been observed among closely related mtDNAs. These include a novel restriction fragment polymorphism in rat mtDNA (4,5), a deletion in the D-loop region of gorilla mtDNA (10), insertions and deletions between the D loops of mice and rats (11) and insertions or deletions of one or two bases in the D-loop region of different human mtDNA variants (12,13). Recent restriction enzyme studies on human mtDNA polymorphism (14), however, indicate that changes of this sort may be more common than previously thought.

The first example of intraspecific variation in animal mtDNA apparently

due to deletions or insertions was the unusual fragment polymorphism described by Brown and Simpson (4) and Brown et al. (5). Three closely related Rattus norvegicus mtDNA variants, designated NC_2, NC_3 and NC_5, were found which seemed to differ by deletions or insertions of 4 base pairs or less. These changes were the only discernable differences among these DNAs, and they all occurred within a HinfI restriction fragment approximately 150 bp in length. The confinement of these differences to such a small region on the molecule indicated that the variation occurs at a single locus. The extremely small amount of overall sequence divergence between these mtDNAs (0.1%) further suggests that the deletions and insertions are generated rapidly relative to base substitution changes. The variable fragment therefore appears to contain a preferred site for deletion and insertion events. Sequence analysis of the region was undertaken in order to better understand the nature of these changes and the means by which they are produced.

We report here the sequences of this variable region for six different cloned rat mtDNAs. The changes are shown to arise from the deletion or insertion of one or more bases in the region separating the $tRNA^{Cys}$ and $tRNA^{Tyr}$ genes. We have found the identical type of variation among the cloned DNAs obtained from a single animal. This suggests the possibility that the mtDNA sample from which the clones were derived contained more than one type of mtDNA.

MATERIALS AND METHODS

Mitochondrial DNAs were isolated from the liver and kidneys of individual wild Rattus norvegicus and Rattus rattus animals which had been trapped at different sites in North America (4), or from the livers of Sprague Dawley laboratory strain R. norvegicus animals. The procedures used have been described (4). All of the mtDNA samples used for this study had been prepared by two cycles of CsCl gradient centrifugation in the presence of ethidium bromide.

Cloning. A few nanograms of each mtDNA sample were digested with BamHI and combined with pBR322 DNA which had been digested with BamHI and treated with alkaline phosphatase. The mixture was precipitated by the addition of 2.5 volumes of ethanol, the DNA was collected by centrifugation, treated with DNA ligase and used to transform E. coli strain K802 (recA-). Ampicillin resistant, tetracycline sensitive clones were selected and analyzed for the presence of recombinant plasmids (15). Cloning procedures for R. norvegicus type NA mtDNA and R. rattus type RE mtDNA have been described (6).

DNA Analysis. Restriction digests, 3' end-labeling and gel electrophoresis were performed as described (4). For Southern blot analysis, fragments were transfered to a "Gene Screen" membrane (New England Nuclear) and hybridized with a [^{32}P]-labeled nick translated probe according to the manufacturer's instructions. DNA sequence analysis was by the Maxam-Gilbert method (16) with minor modifications. In order to sequence the cloned HinfI fragments used in this study, it was usually necessary to treat the isolated, end-labeled fragments with HaeIII prior to strand separation in order to remove contaminating pBR322 sequences.

RESULTS

Cloning of the variant rat mtDNAs

In order to obtain sufficient quantities of DNA for sequencing, the BamHI fragments of types NC_2, NC_3 and NC_5 rat mtDNAs were cloned in pBR322. BamHI cleaves all known R. norvegicus mtDNAs at two sites, producing fragments of approximately 11 and 5 kb (4, 5, 17-20). The location of these sites are indicated in Figure 1. Clones containing plasmids with inserts of these fragments were obtained for the three mtDNA variants, and in each case the mtDNA sample used for the cloning was derived from a single animal. Restriction of the large BamHI fragment containing plasmids with HinfI revealed length variation in an 150 bp fragment which paralleled the variation seen in the corresponding mtDNAs (Figs. 2A&2B). Clones possessing these large fragments were termed NC201 for type NC_2, NC303 for type NC_3 and NC506 for type NC_5.

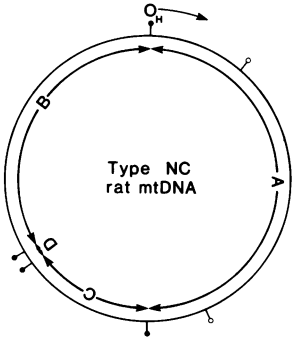

Figure 1. (A) BamHI (open circles) and HindII (closed circles) restriction map of type NC R. norvegicus mtDNA. O_H, origin of heavy strand replication (D-loop). The arrow denotes the direction of heavy strand replication.

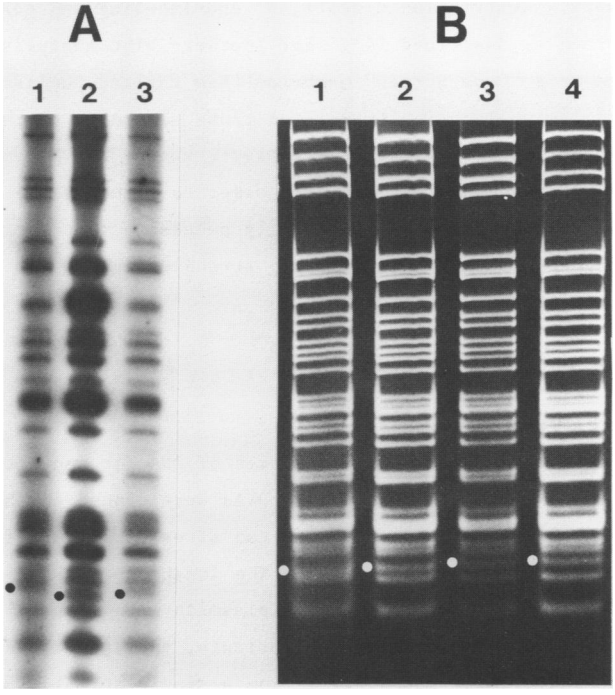

Figure 2. (A): HinfI fragments of type (1) NC_3, (2) NC_2 and (3) NC_5 mtDNAs. (B): HinfI fragments of clones (1) NC201, (2) NC506, (3) NC303, (4) NC308. (A) is an autoradiograph of fragments labeled at their 3' ends by the use of DNA polymerase I large fragment and [^{32}P] dATP. (B) is an ethidium bromide stained gel. Variable fragments are indicated by dots.

Variation in the mtDNA clones obtained from a single animal

In the process of screening clones for variation in HinfI fragments it was noticed that a different large BamHI fragment clone of type NC_3 mtDNA, NC308, possessed a slightly larger HinfI fragment than the variant fragment of NC303 (Fig. 2B). This clone was derived from the same mtDNA sample as clone NC303, and since this mtDNA sample was derived from a single animal, then either variation existed in the population of mtDNAs within this animal or the variation occurred during the cloning process.

Southern blot experiments established that the bulk of the variable HinfI fragment lay within the smallest HindII fragment (fragment D) of rat mtDNA (Fig. 1), and it was possible to clearly distinguish between the HindII fragments of clones NC303 and NC308 by polyacrylamide gel electrophoresis. In order to estimate the likelihood that the variation between clones NC303 and

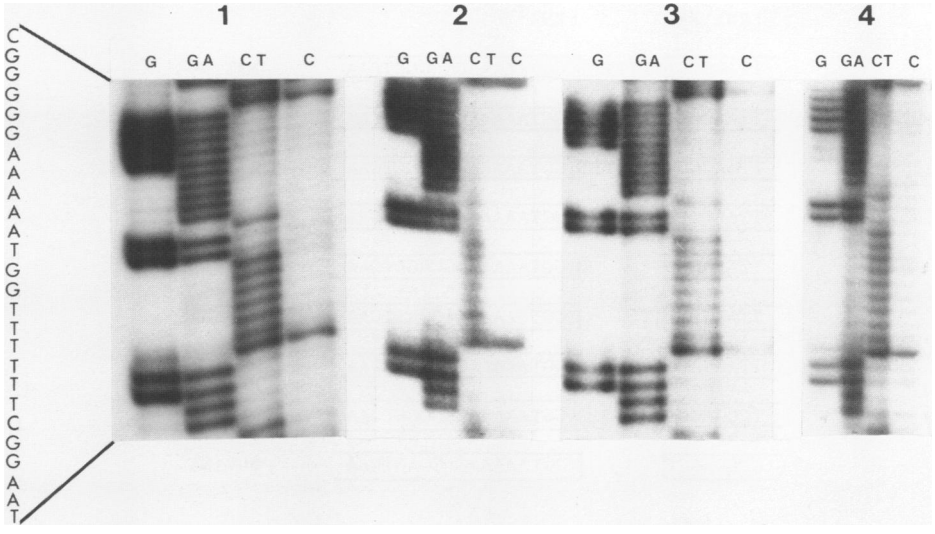

Figure 3. Maxam-Gilbert sequencing gels of the variable region of clones: 1, NC201; 2, NC506; 3, NC303; 4, NC308.

NC308 occurred during their propagation, plasmids from 37 individual colonies derived from NC303 and 42 individual colonies derived from NC308 were analyzed with HindII. No variation was observed among any of the NC303 DNAs or the NC308 DNAs. All the NC303 DNAs could be distinguished from the NC308 DNAs, however, and vice-versa. These results suggest that the length of the variable fragment is stable during propagation of the clones and, presumably, throughout the cloning process.

In order to assess the degree of heterogeneity present among different original clones obtained from the NC_3 mtDNA sample, 15 additional large fragment containing clones were analyzed with HindII. All were found to possess a fragment equal in mobility to that of NC303. This result indicated that the bulk of the mtDNA in the initial sample was that represented by clone NC303 and that if the NC308 type occurred at all it was present only at a relatively low level. No variation was observed among several different initial clones of the NC_2 and NC_5 mtDNA samples, indicating that in each case the cloned sequences were representative of the bulk of the mtDNA samples. Furthermore no differences in mobility could be observed between the variable HinfI fragments of the cloned DNAs other than NC308 and those of the corresponding original mtDNA samples.

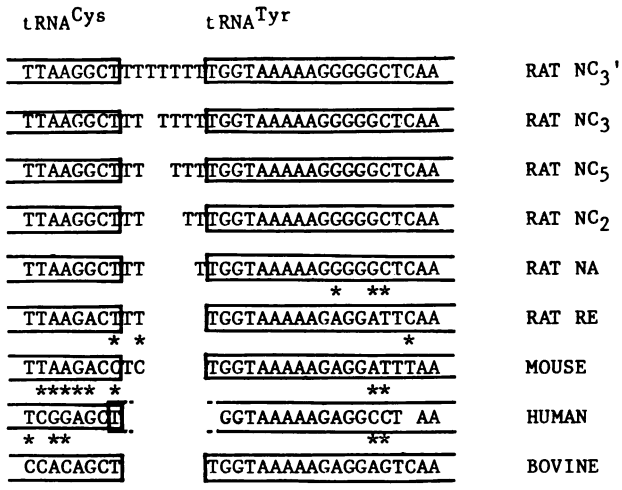

Figure 4. Alignment of the variable site on the R. norvegicus type NC DNAs with the same region on other mtDNAs. Sequences of human, mouse and bovine mtDNAs are from references 21, 8 and 9, respectively. Mismatched bases are indicated by asterisks(*). The minor variant of the rat NC_3 type represented by clone NC308 is designated type NC_3'. Rat RE refers to R. rattus type E mtDNA.

DNA sequence of the variable region

The varying HinfI fragments of clones NC201, NC303, NC308 and NC506 were isolated and their sequence was determined. The sequences were found to differ only in the number of thymidine residues (light strand) occurring in a run of Ts found in the central region of the fragment (Fig. 3). Clone NC201 has the shortest run, with 6 Ts at this location, while NC506 has 7 Ts, NC303 8 Ts and NC308 9 Ts. The HinfI fragments of the variant mtDNAs therefore differ in length by increments of 1 base. This is inconsistent with the 3-4 bp increment differences estimated by polyacrylamide gel electrophoresis (4,5). Presumably, the insertion/deletion of an additional A-T base pair at this site affects the mobility of the fragment in the gel by changing properties other than its length alone. Just on the 3' side of the variable T repeat, runs of 5A's followed immediately by 5G's are found in all four type NC mtDNAs.

Sequence of the variable region in related mtDNAs

In order to gain a more general picture of the variation which occurs in this region, the sequence of the HindII D fragment, which contains most of the variable HinfI fragment of the NC DNAs, was determined for R. norvegicus type A mtDNA and also for Rattus rattus type E mtDNA, in which the fragment also

occurs. The HindII D fragment was used because it was unclear which HinfI fragment possessed the variable region in these DNAs. Since one of the restriction sites which bounds this fragment is present in mouse mtDNA (18), it was easily possible to align the rat sequences with the mouse mtDNA sequence. The aligned regions of the six rat mtDNAs, mouse mtDNA (8), bovine mtDNA (9), and human mtDNA (21) are shown in Fig. 4. The T run of variable length occurs between the $tRNA^{Cys}$ and $tRNA^{Tyr}$ genes. These genes are separated by from 4 to 7 Ts in the rat type NC mtDNAs, by three Ts in the laboratory rat (type NA) mtDNA, and by two Ts in R. rattus mtDNA. Two nucleotides also separate the genes in mouse mtDNA, the genes overlap in human mtDNA, and are "butt-jointed" in bovine mtDNA. The runs of repeated As and Gs are found only in the R. norvegicus mtDNAs near the 5' end of the $tRNA^{Tyr}$ gene. One of the changes noted here, that between rat types NA and NC_2 mtDNAs, has also been found to occur in the mtDNA of Yoshida rat sarcoma cells (22).

DISCUSSION

In addition to this report, two other instances of intraspecific variation in animal mtDNA due to deletions and insertions have been described at the nucleotide sequence level. Greenberg and colleagues (12, 13) have shown that deletions or insertions of one or two bases occur in the D-loop region of human mtDNA, and Taira et al. (22) have found small deletions and insertions within and surrounding the tRNA genes or rat tumor mitochondria. The present work describes a hotspot for these changes in rat mtDNA. All three studies implicate nucleotide repeats as a preferred site for such events.

The occurrence of small deletion and insertion mutations has often been found to correlate with the presence of repeated bases in the DNA sequence (23-27). As initially postulated by Streisinger et al. (23), runs of repeated nucleotides can promote such changes by allowing relatively stable heteroduplexes to form when complementary DNA strands are misaligned. In the case of the variable region on rat mtDNA, such misalignment could occur via DNA strand slippage during the replication of the run of Ts. Relatively stable duplexes resulting from such events would be generated if slippage occurred only within the run of Ts or if the misalignment extended into the nearby runs of As and Gs. Following a subsequent cycle of DNA replication, a mutated DNA with an additional T in the run would be generated. Strand slippage could also occur during repair or recombination (23), but these processes may not occur in animal mitochondria (28,29).

The prevalence of deletions and insertions at this particular site in R.

norvegicus mtDNA may be due to the coincidence of neighboring runs of repeated nucleotides at and near an intergenic region. The observed changes do not result in a change in a gene product and are therefore not likely to be seriously deleterious. Changes of this type in most other regions, e.g. a frameshift change in a protein gene, would be far more likely to lead to a defective mitochondrion, and we would expect such changes to be eliminated from the mtDNA population by natural selection. The high rate of deletion and insertion changes in a site at which no change results in a gene product is similar in some ways to the high rate of silent base substitution in protein coding genes (6,7).

The occurrence of these deletion and insertion changes at a single site in mtDNAs which are otherwise extremely similar and perhaps even identical is striking. Since repeated nucleotides such as those described here are likely to occur at other sites on the mtDNA molecule, it is probable that this type of mutation occurs frequently in mtDNA. In this regard, we have recently become aware of a similar hypervariable site for such changes occurring within the D-loop region of bovine mtDNA (W. Hauswirth, personal communication). In addition, new observations on restriction site variation in human mtDNA indicate the deletion/insertion polymorphisms can occur at at lease nine distinct sites, seven of which lie outside the D loop (14). The fact that such changes have not often been observed in the comparison of mtDNA restriction fragments is probably primarily a manifestation of the small size changes involved.

While the occurrence of heteroplasmic individuals, i.e. animals possessing more than one kind of mtDNA, is necessary for mtDNA evolution to take place, the finding of variation at restriction sites in the mtDNA population of an individual by the use of conventional techniques has not been reported. The animal from which the differing clones NC303 and NC308 were obtained may well have been heteroplasmic. This view is supported by two findings. First, no variation could be detected among the many recloned NC308 and NC303 colonies, indicating that the sequence is stable during clone propagation and hence that the variation probably did not occur during cloning. Second, the type of variation which occurs between clones NC303 and NC308 is identical to that which occurs between different individuals. This is what one would expect if an intra-individual polymorphism represented an intermediate stage in the establishment of a new variant mtDNA in the rat population. Greenberg et al. (13) have likewise described two distinguishable mtDNA clones derived from the same individual. In this case the clones differed with respect to the presence of a single HaeIII site. The stability of these sequences during

clone propagation, however, was not reported.

The finding of polymorphism at this particular site among the mtDNA clones of an individual may well be more than simply fortuitous. The high degree of variation at this site relative to restriction site variation among the different rat mtDNAs indicates the region is particularly susceptible to change, and these changes may occur with sufficient frequency that variation within individuals may be common. We have begun to screen clones from other animals possessing type NC mtDNA in order to obtain additional data bearing on this point.

ACKNOWLEDGEMENTS

We would like to thank R. Cann, W. Hauswirth, J. Slightom, A. Wilson and M. Zuker for helpful discussions; R. Cann and A. Wilson for sending us their manuscript (14) prior to its publication; S. Gibbs for helpful comments on the present manuscript; R. Prussick for cloning type NC_5 mtDNA; and S. Gauthier for assistance in preparation of the manuscript. The mtDNAs used in this study were isolated by G.G.B. in M.V. Simpson's laboratory in the Department of Biochemistry, SUNY, Stony Brook, and G.G.B. acknowledges the support and encouragement he received there. This work was supported by grants from NSERC (A7624) and FCAC (83EQ2177).

*To whom correspondence should be addressed

REFERENCES

1. Dawid, I.B. (1972) Dev. Biol. 29, 139-151.
2. Brown, W.M., George, M., Jr., and Wilson, A.C. (1979) Proc. Natl. Acad. Sci. USA 76, 6319-6323.
3. Avise, C.J., Giblin-Davidson, C., Laerm, J., Patton, J.C., and Lansman, R.A. (1979) Proc. Natl. Acad. Sci. USA 76, 6694-6698.
4. Brown, G.G. and Simpson, M.V. (1981) Genetics 97, 125-143.
5. Brown, G.G., Castora, F.J., Frantz, S.C. and Simpson, M.V. (1981) Ann. N.Y. Acad. Sci. 361, 135-153.
6. Brown, G.G. and Simpson, M.V. (1982) Proc. Natl. Acad. Sci. USA 79, 3246-3250.
7. Brown, W.M., Prager, E.M., Wang, A. and Wilson, A.C. (1982) J. Mol. Evol. 18, 225-239.
8. Bibb, M.J., Van Etten, R.A., Wright, C.T., Walberg, M.W. and Clayton, D.A. (1981) Cell 26, 167-180.
9. Anderson, S., DeBruijn, M.H.L., Coulson, A.R., Eperon, I.C., Sanger, F. and Young, I.G. (1982) J. Mol. Biol. 156, 683-717.
10. Ferris, S.D., Wilson, A.C. and Brown, W.M. (1981) Proc. Natl. Acad. Sci. 78, 2432-2436.
11. Walberg, M.W. and Clayton, D.A. (1981) Nucleic Acids Res. 9, 5411-5421.
12. Aquadro, C.F. and Greenberg, B.D. (1983) Genetics 103, 287-312.
13. Greenberg, B.D., Newbold, J.E. and Sugino, A. (1983) Gene 21, 33-49.

14. Cann, R. and Wilson, A.C. (1983) Genetics 104, 699-711.
15. Birboim, H.C. and Doly, J. (1979) Nucleic Acids Res. 62, 1511-1523.
16. Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 490-560.
17. Kroon, A.M., Pepe, G., Bakker, H., Holtrop, M., Bollen, J.E., Van Bruggen, E.J.F., Cantatore, P., Terpstra, P. and Saccone, C. (1977) Biochim. Biophys. Acta 478, 128-145.
18. Parker, R.C. and Watson, R.M. (1977) Nucleic Acids Res. 4, 1291-1299.
19. Francisco, J.F., Brown, G.G. and Simpson, M.V. (1979) Plasmid 2, 426-436.
20. Castora, F.J., Arnheim, N.A. and Simpson, M.V. (1980) Proc. Natl. Acad. Sci. USA 77, 6415-6419.
21. Anderson, S., Bankier, A.T., Barrell, B.G., DeBruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P. Roe, B.A., Sanger, F., Schreier, P., Smith, A.J.H., Staden, R. and Young, I.G. (1981) Nature (London) 290, 457-465.
22. Taira, M., Yoshida, E., Kobayashi, M., Yaginuma, K., and Koike, K. (1983) Nucleic Acids Res. 11, 1635-1643.
23. Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E. and Inouye, M. (1966) Cold Spring Harbor Symp. Quant. Biol. 31, 77-84.
24. Isono, K. and Yourno, J. (1974) Proc. Natl. Acad. Sci. USA 71, 1612-1617.
25. Farabaugh, P.J., Schmeissner, U., Hofer, M. and Miller J.H. (1978) J. Mol. Biol. 126, 847-863.
26. Pribnow, D., Sigurdson, D.C., Gold, L., Singer, B.S., Brosius, J., Dull, T.J. and Noller, H.F. (1981) J. Mol. Biol. 149, 337-376.
27. Shen, S., Slightom, J.L. and Smithies, O. (1981) Cell 26, 191-203.
28. Lansman, R.A. and Clayton, D.A. (1975) J. Mol. Biol. 99, 761-776.
29. Ho, C. and Coon, H.C. (1979) in Proceedings of the ICN-UCLA Symposium on Extrachromosomal DNA. Cummings, D., Borst, P., Dawid, I.B., Weissman, S. and Fox, C.F. eds. (Academic, New York), Vol. 15, pp. 501-514.