Proc. Natl. Acad. Sci. USA
Vol. 77, No. 4, pp. 2129–2133, April 1980
Genetics

DNA sequence of baboon highly repeated DNA: Evidence for evolution by nonrandom unequal crossovers

(satellite DNA/nucleosome/heterochromatin/primates)

LAWRENCE DONEHOWER*†, CLEMENT FURLONG‡, DAVID GILLESPIE*§, AND DAVID KURNIT‡¶

*Section on Molecular Hybridization, Laboratory of Tumor Cell Biology, National Cancer Institute, Bethesda, Maryland 20205; and ‡Center for Inherited Diseases, RG-20, University of Washington, Seattle, Washington 98195

Communicated by Arno G. Motulsky, December 26, 1979

ABSTRACT A highly repeated DNA was isolated from the West African baboon (*Papio papio*) as a 343-base-pair fragment after digestion of total baboon DNA with the restriction endonuclease *Bam*HI. The DNA sequence of this fragment was obtained by chemical cleavage methods and is compared with the DNA sequence of related highly repeated primate DNAs from African green monkey (*Cercopithecus aethiops*) and man. The 343-base-pair baboon repeat consists of two related but nonidentical wings of 172 and 171 base pairs, respectively. The baboon 172-base-pair wing shares more homology with the African green monkey 172-base-pair repeat than with the baboon 171-base-pair wing. Comparison with the previously published monkey and human DNA sequences indicates that: (*i*) All the DNA sequences apparently arose from a common ancestral sequence. (*ii*) Evolution of the primate DNA sequences can be explained by a model involving unequal crossovers at specific points within the repeated DNA, possibly mediated by the sequence $\frac{5'\text{-AAGG-}3'}{3'\text{-TTCC-}5'}$ or its invert $\frac{5'\text{-GGAA-}3'}{3'\text{-CCTT-}5'}$. (*iii*) There are alternating domains of conserved and divergent DNA sequences within each >170-base-pair wing sequence. Taken together, the DNA sequences of these primates suggest a model whereby highly repeated DNAs are established and evolve as a consequence of unequal nonrandom exchanges of DNA duplexes. These exchanges may be mediated by short repeated nucleotide sequences and involve exchanges within and between the >170-base-pair wings.

A significant, but markedly variable, component of the eukaryotic genome is comprised of highly repeated DNA sequences that are localized in constitutive heterochromatin, that are not transcribed, and for which quantitative variation within a species (or among closely related species) is tolerated without major phenotypic effects (1, 2). Nucleic acid hybridization and DNA sequence studies demonstrate that, unlike the remainder of the genome, these highly repeated DNAs tend to be comprised of short simple sequences repeated in tandem with a high degree of quantitative variation both within and among eukaryotic species. Molecular and cytogenetic evidence suggest a model whereby unequal double-strand exchanges associated with mitotic DNA replication result in the creation and amplification of these DNA sequences (1, 3–5).

In this report, the DNA sequence is obtained for a baboon highly repeated DNA. Comparison of this DNA sequence with previously published related DNA sequences from African green monkey (6) and man (7) is consistent with the above unequal exchange model and suggests the additional constraint that the exchanges are nonrandom, possibly mediated by a short specific nucleotide sequence. This model of unequal mitotic exchanges mediated by short specific nucleotide sequences may be relevant to recent findings concerning the evolution and plasticity of the eukaryotic genome.

MATERIALS AND METHODS

Isolation of 343-Base-Pair Repeat from Baboon DNA. DNA from *Papio papio* (West African baboon) was isolated from frozen primate tissues by a modification (8) of Marmur's method (9). Total *Papio* DNA was incubated with an excess of the restriction endonuclease *Bam*HI (Bethesda Research Laboratories, Rockville, MD) (2–5 units per μg of DNA for 4 hr at 37°C) and the 343-base-pair repeat [denoted BABr (*Bam*HI)-1] was cut out and eluted from an 8% polyacrylamide gel as described (8) or by electroelution in 45 mM Tris base/41 mM boric acid/1 mM EDTA/0.1% sodium dodecyl sulfate at 100 mV for 8 hr.

DNA Sequence Determination. BABr (*Bam*HI)-1 was labeled at the 3′ ends by adding 100 μCi (1 Ci = 3.7×10^{10} becquerels) of deoxyguanosine 5′-[α-^{32}P]triphosphate (2000–3000 Ci/mmol; Amersham) to 15 μl containing 2.5 μg of purified BABr (*Bam*HI)-1 and 1.4 units of Klenow *Escherichia coli* polymerase (large fragment) (Boehringer Mannheim) in *Hin* buffer (10 mm Tris·HCl, pH 7.4/50 mM NaCl/6.6 mM $MgCl_2$/0.1 mM dithioerythritol). This resulted in the fill-in of one unique labeled nucleotide to each of the 3′ ends of the *Bam*HI cut.

After labeling, BABr (*Bam*HI)-1 was split into two fragments with the restriction enzyme *Hph* I (New England Biolabs) (10). The cleavage products were separated on an 8% gel and electroeluted, yielding fragments for DNA sequence determination that were each labeled at a single 3′ terminus and were 200 and 142 nucleotides long, respectively (Fig. 1). The DNA sequence of these fragments was obtained by using the Maxam–Gilbert procedure for chemical cleavage (11) [the G, A > C, C, C + T, and alternate G (methylene blue modification) (12) reactions were used]. Electrophoresis was carried out on 20%, 12%, and 8% polyacrylamide (11) gels that were 0.4 mm thick on a gel apparatus that accommodated gels up to 91 cm long (Riverside Scientific Enterprises, Seattle, WA). The DNA sequences adjacent to the *Bam*HI and *Hph* I restriction sites were determined as follows. BABr (*Bam*HI)-1 was digested with *Hin*dIII (Boehringer Mannheim) and labeled by using Klenow polymerase to fill in the *Hin*dIII 3′ termini selectively in the pres-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviation: BABr, (*Bam*HI)-1, 343-base-pair repeat.

† Current address: Laboratory of Tumor Virus Genetics, Building 41, Room D243, National Cancer Institute, Bethesda, MD 20205.
§ Current address: H. L. Orlowitz Cancer Institute, Hahnemann Medical College, Philadelphia, PA 19102.
¶ Current address: Children's Hospital Medical Center, Genetics Division, 300 Longwood Avenue, Boston, MA 02115.

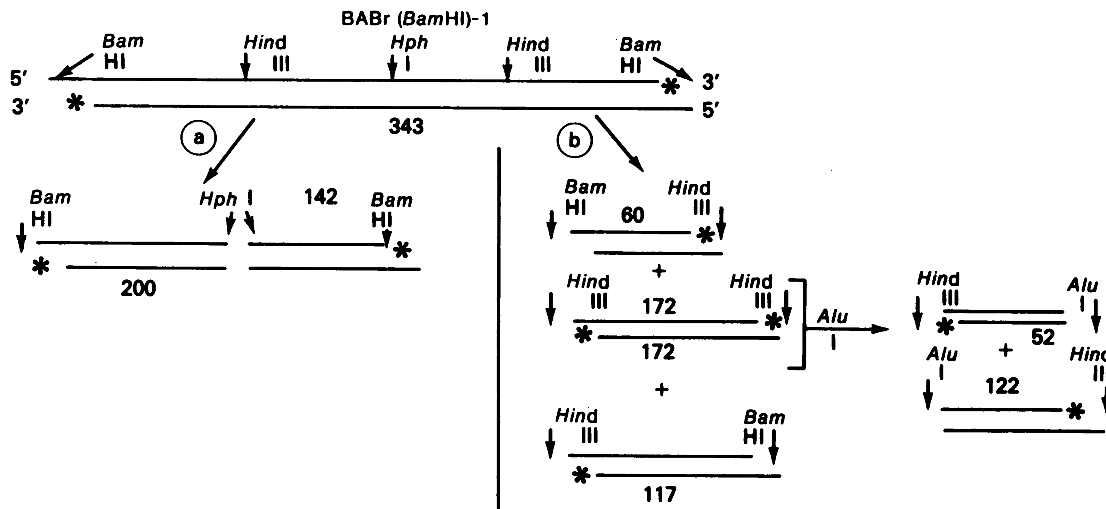

FIG. 1. DNA sequence determination strategy. (*a*) The 343-base pair BABr (*BAM*HI)-1 fragment was end-labeled at the 3′ ends and then cleaved with *Hph* I. Single-stranded 3′-end-labeled fragments of length 200 and 142 base pairs were thereby made available for sequence determination. Asterisks indicate the sites of 3′-end labeling. (*b*) BABr (*Bam*HI)-1 was cleaved with *Hin*dIII and then end-labeled with a protocol that filled in the 3′ ends of the *Hin*dIII sites but left the *Bam*HI ends unlabeled. This generated 60- and 117-base-pair fragments labeled at a single 3′ end which were used directly for sequence determination towards the unlabeled *Bam*HI sites. In addition, a 172-base-pair fragment labeled at both 3′ ends was generated; this fragment was cleaved with *Alu* I and the sequence of the resulting 122-base-pair fragment which contained the *Hph* I recognition and cleavage sites (10) was determined. Taken together, the protocols in *a* and *b* allowed for sequence determination of the entire 343-base-pair BABr (*Bam*HI)-1 fragment.

ence of 50 μM dATP/50 μM dGTP/50 μCi of deoxycytidine 5′-[α-^{32}P]triphosphate (2000–3000 Ci/mmol; Amersham) (Fig. 1). The *Bam*HI termini were not labeled with [^{32}P]dCTP, given the omission of dTTP from the reaction mixture. This procedure generated 117- and 60-base-pair fragments labeled at a single 3′ terminus and a 172-base-pair fragment [equivalent to the "*Papio* 2" wing (Fig. 2)] labeled at both 3′ termini. After separation on an 8% gel and electroelution, the DNA sequence adjacent to the *Bam*HI site was determined towards the (unlabeled) ends of the 117- and 60-base-pair fragments. The 172-base-pair fragment was cleaved with *Alu* I resulting in labeled fragments 52 and 122 base-pairs in length which were separated on an 8% polyacrylamide gel and electroeluted (Fig. 1). The DNA sequence adjacent to the *Hph* I cleavage site was determined by establishing the sequence of the 122-base-pair fragment. Fig. 1 depicts this strategy schematically. Confirmatory sequencing for the major portion of the sequence was obtained by determining the sequence of the complementary strand after labeling at the 5′ end of BABr (*Bam*HI)-1 with adenosine 5′-[γ-^{32}P]triphosphate by using the bacterial alkaline phosphatase–T4 polynucleotide kinase (Bethesda Research Laboratories, Rockville, MD) regimen of Maxam and Gilbert (11) and subsequent cleavage with *Hph* I or *Hin*dIII. The G, A > G, C, and C + T reactions outlined in ref. 11 were utilized to determine the sequence of the 5′-end-labeled *Hph* I fragments of BABr (*Bam*HI)-1.

FIG. 2. Nucleotide sequence and restriction endonuclease map of *P. papio* highly repeated DNA. The nucleotide sequence of the 343-base-pair highly repeated fragment of baboon DNA isolated after *Bam*HI digestion [denoted BABr (*Bam*HI)-1] is presented. The format of presentation is such that the sequence has been divided into two wings (*Papio* 1 and *Papio* 2) to facilitate comparison with each other and with the known DNA sequence for other primate >170-base-pair repeats. (*Upper*) Nucleotides 1–85 of *Papio* 1 and 2. (*Lower*) Nucleotides 86–172 of *Papio* 1 and 2 (note that *Papio* 1 has one less nucleotide, indicated by a blank at residue 90 in the *Papio* 1 sequence). The previously determined restriction endonuclease recognition sites (8), which are in agreement with the presented sequence, are denoted by arrows.

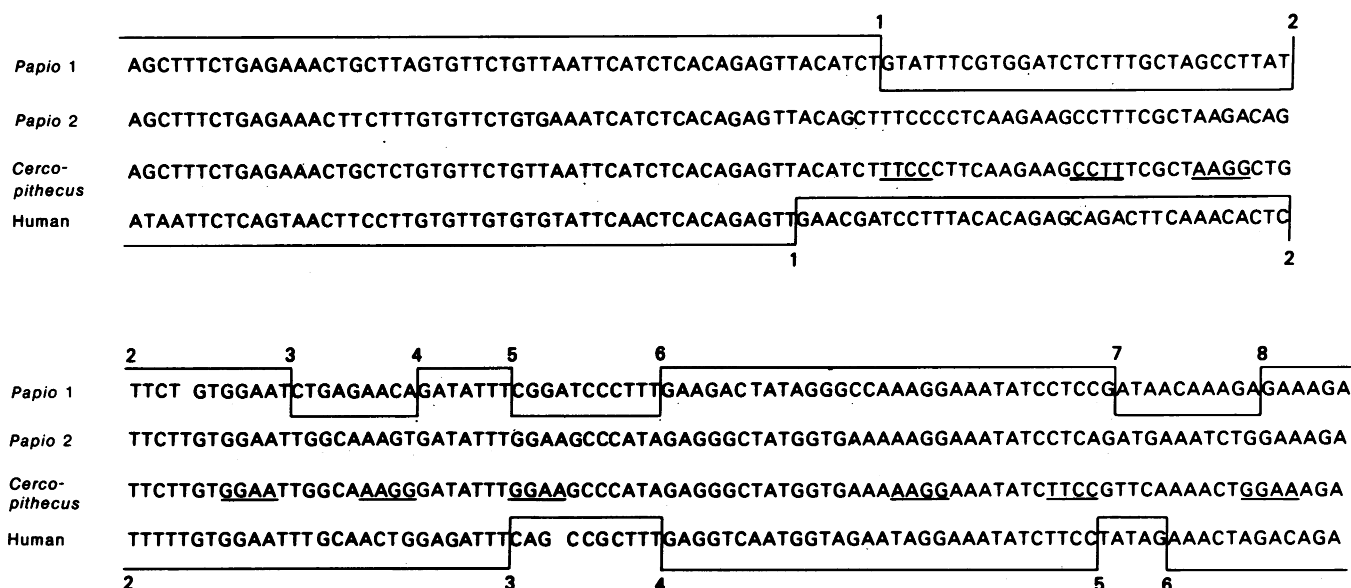

FIG. 3. Comparison of DNA sequence of highly repeated primate DNAs. The DNA sequences of *Papio* 1 and *Papio* 2 wings are presented in the same format as in Fig. 1 in conjunction with the published DNA sequences for a 172-base pair repeat from *C. aethiops* (African green monkey) (6) and a 171-base-pair wing of a 342-base-pair repeat from *Homo sapiens* (man) (7). By utilizing the *Cercopithecus* DNA sequence as a reference, regions of homology among the primate highly repeated DNA sequences are indicated by inclusion between the solid lines; conversely, regions excluded by the lines indicate sequences that lack superimposable homology with the *Cercopithecus* sequence. The boundaries between homologous and nonhomologous regions are numbered separately for the *Papio* 1 and human repeats. These boundaries may represent the sites of nonrandom unequal exchanges (see text). The occurrence of the short nucleotide sequence $\frac{5'\text{-AAGG-}3'}{3'\text{-TTCC-}5'}$ and its invert $\frac{5'\text{-GGAA-}3'}{3'\text{-CCTT-}5'}$ has been denoted by underlining.

RESULTS

The DNA sequence obtained for BABr (*Bam*HI)-1 is given in Fig. 2. The DNA sequence is in agreement with the known restriction map for this sequence (8). The 343-base-pair fragment is comprised of two wings, denoted "*Papio* 1" and "*Papio* 2" in Fig. 2. The sequence is presented in a format that begins at the *Hin*dIII restriction site of each wing to facilitate comparison with the DNA sequences from other primates (Fig. 3) [in particular, the 172-base pair repeat of AGMr (*Hin*dIII)-1 from *C. aethiops* (6) which was originally sequenced from the *Hin*dIII site]. The *Bam*HI site utilized for isolation and end-labeling of BABr (*Bam*HI)-1 occurs at positions 114–119 of the *Papio* 1 wing.

No definitive evidence for sequence heterogeneity within the 343-base-pair fragment was apparent on the DNA sequencing gels. However, a number of minor bands were seen on an 8% polyacrylamide gel used to isolate fragments of BABr (*Bam*HI)-1 after cleavage and selective labeling at the *Hin*dIII sites. Thus, although sequence heterogeneity does exist, it did not result in ambiguities in interpretation of the sequencing gels.

DISCUSSION

Theme and Variations: An Apparent Core Sequence. In contrast with various highly repeated eukaryotic DNAs that are characterized by tandem repeats of short oligonucleotide sequences [e.g., crab (13) and *Drosophila* (14) DNAs] (for review, see ref. 2), the related primate highly repeated DNA sequences do not show a simple short nucleotide repeat pattern. However, the internal sequence is nonrandom, and shows evidence of a "theme and variation" pattern with the suggestion of various permutations on a common prototype sequence. Over 50% of the sequence is characterized by polypyrimidine:polypurine runs of four or more nucleotides; in particular, the sequences $\frac{5'\text{-CTTT-}3'}{3'\text{-GAAA-}5'}$ and its invert $\frac{5'\text{-TTTC-}3'}{3'\text{-AAAG-}5'}$ and $\frac{5'\text{-CCTT-}3'}{3'\text{-AAGG-}5'}$ and its invert $\frac{5'\text{-TTCC-}3'}{3'\text{-GGAA-}5'}$ are over-represented in these runs. Thus, the baboon sequence reported in Fig. 2 resembles mouse, guinea pig, and kangaroo rat satellite DNAs (15, 16) in having a complex internal pattern albeit with evidence of a basic core sequence. In this regard, similarities between the core sequence of primates given above and those reported for mouse (14), *Drosophila* (17), and human (18) satellite DNAs are of note.

Comparison of Primate Highly Repeated DNA Sequences. The highly repeated fraction of baboon DNA that shows a regular 343-base-pair periodicity after cleavage with *Bam*HI (8) is comprised of two wings of length 172 and 171 base pairs, respectively. These two nonidentical wings (19) are closely related to a 171-base-pair wing of a 342-base-pair repeated human DNA fragment (7) and to a 172-base-pair repeated sequence denoted AGMr (*Hin*dIII)-1 isolated from the African green monkey (6) as shown in Fig. 3. The similarities among the DNA sequences in Fig. 3 support the concept that all these highly repeated primate DNA sequences were derived from a common ancestral sequence (8, 16, 20). Among the primate repeated sequences in Fig. 3, the *Cercopithecus* sequence most closely resembles the commonest sequence (derived by taking the commonest base at each position), differing by only four bases. In comparison, *Papio* 2 differs by 7 bases, *Papio* 1 differs by 37 bases, and human differs by 43 bases. It is of note that *Papio* 2 shares more sequence homology with *Cercopithecus* than with *Papio* 1.

Comparison of the known primate repeated DNA sequences in Fig. 3 demonstrates that the sequence divergence from the common ancestral sequence has been nonrandom in that there are regions of relatively high nucleotide sequence conservation interspersed with regions of extensive sequence divergence. In Fig. 3, the boxed regions within the lined domains indicate relatively constant regions in which primate highly repeated DNA sequences share close homology with the known sequence for *Cercopithecus* (6). These are interspersed with more "variable" regions in which *Papio* 1 or human diverge sharply

from the *Cercopithecus* and *Papio* 2 sequences. The "variable" regions of *Papio* 1 and human cannot be superimposed readily *in toto* on other parts of the *Cercopithecus* sequence or on each other. However, the DNA sequence in these regions is nonrandom as evidenced by the maintained over-representation of the basic core oligonucleotide sequence (see above) in such regions.

The most plausible explanation for the evolution of the *Papio* 1 and human highly repeated DNAs from an ancestral sequence is again via unequal crossovers. The degree and discreteness of divergence and nonsuperimposability of divergent sequences in the "variable" regions disfavors models depending on replication slippage or unselected drift to explain the occurrence of this divergence. Assuming an independent mutational history of the four sequences in Fig. 3, the high degree of sequence conservation between *Papio* 2 and *Cercopithecus* in the "variable" regions makes it untenable that the variable regions represent areas with an intrinsically high mutation rate. It follows that the pattern of divergence is consistent with a model in which the "variable" regions represent exchanges of stretches of DNA in an ancestral sequence with similar but nonidentical DNA sequences. The existence of such nonidentical but closely related DNA sequences adjacent to the original >170-base pair repeat would be an expected feature of the unequal crossover model (3, 4). This proximity coupled with the availability of short-range nucleotide homology would facilitate exchange of related but nonidentical DNA fragments. Thus, as illustrated in Fig. 3, the origin of *Papio* 1 from an ancestral sequence more closely representing the *Cercopithecus* sequence can be explained by eight crossover events within the >170-base pair repeat; in the case of the human sequence, six such internal crossover sites are indicated in Fig. 3. Comparison of the DNA sequences in the variable regions of *Papio* 1 and human highly repeated DNAs shows that the crossovers have occurred nonrandomly; of the 51 bases in the human "variable" regions and 59 bases in the *Papio* 1 "variable" regions, 44 are apposed at the same point along the sequence. There are two possible interpretations of this nonrandomness: the postulated crossovers resulted from the same crossover events in an ancestral sequence common to human and *Papio* or independent crossover events that occurred were restricted in their potential location. Although the two smaller human "variable" regions (between human crossovers 3–4 and 5–6) could be related in part to the *Papio* crossovers 5–6 and 7–8, respectively, this is not the case for the largest crossovers (1–2 in each sequence), which were apparently independent events in human and *Papio*, respectively. Computer analysis by Monte Carlo simulations and Markov chain analysis of the sequences in Fig. 3 indicate that divergence from a *Cercopithecus*-like common ancestral sequence may have occurred about 5 × 10^7 years ago (unpublished results).

It is of note that the nucleotide sequence 5′-AAGG-3′/3′-TTCC-5′ or its invert 5′-TTCC-3′/3′-AAGG-5′ (underlined in the *Cercopithecus* sequence in Fig. 3) is in close proximity to every hypothetical crossover point except number 6 for *Papio* 1 and number 4 for human where 5′-AGGG-3′ is present (which could have been AAGG in an ancestral sequence). Further, the tetranucleotide AAGG is at or near the recombinational joints of several rearrangements involving the *Cercopithecus* >170-base-pair repeat, which have been detected in defective variants of simian virus 40 (see figure 3 of ref. 21 and figure 5 of ref. 22). This short nucleotide sequence could function as a recognition site for mammalian site-specifc nuclease (23) relevant to mammalian DNA recombination. Alternatively, the existence of this short repeat could promote recombination simply by providing a short stretch of nucleotide sequence homology as outlined in the model of Smith (3).

Constraints on Exchanges. The first major constraint on exchanges, detailed above, is nonrandomness. The second major constraint is that the lengths of DNA fragments interchanged are identical (or nearly so) as indicated by the high degree of conservation of the >170-base-pair repeating units in primates (5). The third apparent constraint (which will require verification by determining the sequence of other highly repeated DNAs) is that there may be highly conserved regions within the >170-base-pair sequence. The final constraint is that quantitative amplification of the primate DNA sequences occurs in units of integral multiples of >170 base pairs. The mechanisms responsible for these constraints are open to speculation. An attractive hypothesis, put forward by Maio and coworkers (5), is that major constraints are in fact imposed by nucleosome structure. This hypothesis was prompted by their finding that the repeat length of African green monkey highly repeated DNA of 172 base pairs represents the length of the nucleosome for this DNA (5, 24). Structural limitations based on nucleosome size may explain the homogeneity of primate DNA repeat length and the finding that the exchanges hypothesized in Fig. 3 were "unequal" in that nonhomologous segments of sister chromatids were involved but were "equal" in the lengths of DNA that were interchanged. An additional hypothesis would be that highly conserved regions may be important for binding of the particular nonhistone proteins (24) that are presumably responsible for the constitutively heterochromatic character of at least one of these highly repeated sequences (25). Thus, although speculative at present, it may be possible to elucidate important structure–sequence relationships from the study of highly repeated DNAs.

Exchanges and Sequence Organization. The findings in this paper support and extend the hypothesis that highly repeated DNAs arose from a series of unequal exchanges, possibly mediated by short repeated oligonucleotides during mitotic DNA replication (1, 3–5). The potential ubiquity of these crossovers is underscored by the multiplicity of levels of sequence organization at which they have been postulated to occur. In the case of the primate DNAs reported in this paper, the initial crossovers presumably involved the core oligonucleotide sequences 5′-CTTT-3′, 5′-CCTT-3′, and their inverts, which are over-represented in these DNAs (Fig. 3). These crossovers are hypothesized to have been critical in the formation of the >170-base-pair basic repeats, which characterize various primate highly repeated DNAs (8) and which correspond with the size of a nucleosome in the case of *C. aethiops* highly repeated DNA (5). Maio *et al.* (5) have noted that unequal exchanges of nucleosome (or multiples of nucleosomes)-sized DNA fragments associated with DNA replication might explain the pattern of long-range periodicities seen in a number of highly repeated DNAs. The finding that baboon highly repeated DNA evolved from the amplification of a >340-base-pair repeat comprised by two related but nonidentical >170-base-pair wings suggests that unequal exchanges occurred within >170-base-pair basic repeats as well as between them (Fig. 3). These postulated internal crossovers exhibit several features consonant with previously postulated exchanges: nonrandomness, conservation of the basic >170-base-pair repeat length, and possible association with specific oligonucleotide sequences at crossover points. Finally, as noted in a previous review (1), cytologically detectable variations of highly repeated DNAs may also be explained by a model involving unequal mitotic crossovers.

In summary, unequal exchanges mediated by repeated oligonucleotide sequences and associated with mitotic DNA replication may underly the evolution and amplification of highly repeated DNAs at multiple levels of sequence organi-

zation. The nature and extent of these varied exchanges may be modified by constraints dependent upon nucleosomal and chromosomal structure.

After this work was completed, we learned that a nearly identical repeating sequence has been characterized in bonnet monkey, a primate closely related to baboon (26).

We thank the staff and students of the 1978 DNA sequencing course at the Given Institute of Pathobiology for assistance; M. Olson, G. McKenna, J. Maio, M. Farquhar, and S. Latt for helpful discussions; and Constance Furlong (Riverside Scientific Enterprises) for provision of gel electrophoresis equipment. The research reported in this paper was performed in partial fulfillment of L.D.'s requirements for the doctoral degree in genetics at George Washington University. Support was from National Institutes of Health National Research Service Award 1F326MO6113-01 and Grant GM 15253.

1. Kurnit, D. M. (1979) *Hum. Genet.* **47**, 169–186.
2. Skinner, D. M. (1977) *BioScience* **27**, 790–796.
3. Smith, G. P. (1976) *Science* **191**, 528–535.
4. Perelson, A. S. & Bell, G. I. (1977) *Nature (London)* **265**, 304–310.
5. Maio, J. J., Brown, F. L. & Musich, P. R. (1977) *J. Mol. Biol.* **117**, 637–655.
6. Rosenberg, H., Singer, M. & Rosenberg, M. (1978) *Science* **200**, 294–394.
7. Manuelidis, L. & Wu, J. C. (1978) *Nature (London)* **276**, 92–94.
8. Donehower, L. & Gillespie, D. (1979) *J. Mol. Biol.* **134**, 805–834.
9. Marmur, J. (1961) *J. Mol. Biol.* **3**, 208–218.
10. Kleid, D., Humayun, Z., Jeffrey, A. & Ptashne, M. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 293–297.
11. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
12. Friedmann, T. & Brown, D. M. (1978) *Nucleic Acids Res.* **5**, 615–622.
13. Skinner, D. M., Beattie, W. G., Blattner, F. B., Start, B. P. & Dahlberg, J. E. (1974) *Biochemistry* **13**, 3930–3937.
14. Gall, J. G. & Atherton D. D. (1974). *J. Mol. Biol.* **33**, 319–344.
15. Biro, P. A., Carr-Brown, A., Southern, E. & Walker, P. M. B. (1975) *J. Mol. Biol.* **94**, 71–86.
16. Fry, K. & Salser, W. (1977) *Cell* **12**, 1069–1084.
17. Brutlag, D., Carlson, M., Fry, K. & Salser, W. (1978) *Cold Spring Harbor Symp. Quant. Biol.* **42**, 1137–1146.
18. Manuelidis, L. (1978) *Chromosoma* **66**, 1–21.
19. Singer, D. & Donehower, L. (1979) *J. Mol. Biol.* **134**, 835–842.
20. Gillespie, D. (1977) *Science* **196**, 889–891.
21. Wakamiya, T., McCutchan, T., Rosenberg, M. & Singer, M. (1979) *J. Biol. Chem.* **254**, 3584–3591.
22. McCutchan, T., Singer, M. & Rosenberg, M. (1979) *J. Biol. Chem.* **254**, 3592–3597.
23. Brown, F. L., Musich, P. R. & Maio, J. J. (1978) *Nucleic Acids Res.* **5**, 1093–1107.
24. Musich, P. R., Brown, F. L. & Maio, J. J. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3297–3301.
25. Kurnit, D. M. & Maio, J. J. (1973) *Chromosoma* **42**, 23–36.
26. Rubin, C. M., Houck, C. M., Deininger, P. L. & Schmid, C. W. (1980) *J. Mol. Biol.* **136**, 151–167.