monitoring and evaluation should be launched.

The specific environmental studies recommended include (i) the effects of altered precipitation on ecosystems; (ii) basic studies on plant and microorganism adaptation to seeding agents; (iii) the potential for combination of seeding agent silver with other metals, pesticides, power plant emission products, and other pollution sources; (iv) tracer studies of nucleants in seeded storm cells to locate their deposition in the environment; and (v) long-term monitoring of silver levels and dynamics in the soil-plant-aquatic environment before and after cloud-seeding activities.

Several of our findings indicate that scientific research and policy research efforts should be continued as well as monitoring and reevaluation of effects. A continuing assessment of the nation's hail suppression capability should occur in the years ahead.

References and Notes

1. D. Atlas, *Science* **195**, 139 (1977).
2. S. Changnon, *Bull. Am. Meteorol. Soc.* **58**, 20 (1977).
3. B. Farhar, Ed., *Hail Suppression: Society and Environment* (monograph 24, Institute of Behavioral Science, University of Colorado, Boulder, 1977).
4. _____ and J. Mewes, *Social Acceptance of Weather Modification: The Emergent South Dakota Controversy* (monograph 23, Institute of Behavioral Science, University of Colorado, Boulder, 1976), pp. 10–15.
5. J. Mewes and B. Farhar, in (*3*), pp. 35–45; M. Smith, B. Farhar, J. Mewes, R. Davis, S. Cox, in *ibid*, pp. 124–155; S. Cox and S. Hernandez, in *ibid.*, pp. 156–170; S. Hernandez and T. Toll, in *ibid.*, pp. 171–188; G. Fletcher, in *ibid.*, pp. 189–195; R. Davis, in *Weather and Climate Modification*, W. Hess, Ed. (Wiley, New York, 1974), pp. 767–786.
6. J. Mewes, in (*3*), pp. 80–95 and 103–116.
7. G. H. Stever, Phi Beta Kappa Lecture, Washington, D.C., 18 May 1973; S. Changnon, *Bull. Am. Meteorol. Soc.* **56**, 27 (1975); *ibid.* **57**, 234 (1976).
8. The other major contributors to this assessment study, with their specialty areas in parentheses, were R. Davis (law), E. Fosse (crop insurance), D. Friedman (property insurance), J. Haas (sociology), T. Henderson (weather modification practices), J. Ivens (science writer), M. Jones (technology assessment), D. Klein (ecology), D. Mann (political science), G. Morgan (meteorology), S. Sonka (economics), C. Taylor (economics), and J. Van Blokland (economics).
9. S. Changnon, *J. Appl. Meteorol.* **16**, 837 (1977).
10. L. Boone, *Estimating Crop Losses due to Hail* (Agricultural Economics Report 267, U.S. Department of Agriculture, Washington, D.C. 1974), pp. 4–40; S. Changnon, *J. Appl. Meteorol.* **11**, 1128 (1972); E. Fosse, *Impacts of Effective Hail Suppression on Crop Insurance Industry and Probable Policy Action Response* (Report of Crop-Hail Insurance Actuarial Association, Chicago, 1976), pp. 4–41; D. Friedman, *Hail Suppression Impact Upon Property Insurance* (Report of the Travelers Insurance Company, Hartford, Conn., 1976), pp. 1–69.
11. W. Brinkmann, *Severe Local Storm Hazard in the United States: A Research Assessment* (Institute of Behavioral Science, University of Colorado, Boulder, 1975), pp. 1–154.
12. V. Schaefer, *J. Weather Mod.* **8**, 5 (1976).
13. B. Farhar, thesis, University of Colorado, Boulder (1975).
14. Committee on Atmospheric Science, *Weather and Climate Modification, Problems and Progress* (National Academy of Sciences, Washington, D.C., 1973) pp. 100–106; *Bull. Am. Meteorol. Soc.* **54**, 694 (1973); L. Grand and J. Reid, *Workshop for an Assessment of the Present and Potential Role of Weather Modification in Agricultural Production* (Atmospheric Sciences Paper 236, Colorado State University, Fort Collins, 1975), pp. 38–236.
15. S. Changnon and G. Morgan, *J. Weather Mod.* **8**, 164 (1976).
16. B. Farhar and J. Clark, in preparation. Scientific opinion could have changed since these 1975 cross-sectional surveys [see also (*15*)], as a result of maturation and developments in the field. However, it is unlikely that it would have changed sufficiently to define hail suppression as a currently feasible and certain technology.
17. J. Haas, *Bull. Am. Meteorol. Soc.* **54**, (No. 7), 647 (1973); B. Farhar, *Mass Emergencies* **1**, 313 (1976).
18. D. Pfost, in (*3*), pp. 117–123.
19. E. Rogers and F. Shoemaker, *Communication of Innovations* (Free Press, New York, 1971); B. Farhar, "Public reaction to weather modification," paper presented National Science Foundation Advisory Panel Meeting, Boulder, Colo., 14 July 1975.
20. S. Changnon *et al.*, *Hail Suppression Impacts and Issues* (Report of Illinois State Water Survey, Urbana, 1977), pp. 1–442.
21. S. Krane, *Social Implications of the National Hail Research Experiment: A Longitudinal Study* (Human Ecology Research Services, Boulder, Colo., 1976); B. Farhar and J. Mewes, *Weather Modification and Public Opinion: South Dakota, 1973, Second Interim Report* (Institute of Behavioral Science, University of Colorado, Boulder, 1974).
22. B. Farhar, "The locus of control of weather modification: scientists vs. citizens," paper presented at the Midwest Sociological Society Meeting, Minneapolis, 15 April 1977.
23. D. Mann and B. Farhar, in (*3*), pp. 196–218.
24. B. Farhar, R. Rinkle, G. Johnson, "Cloud seeding in California: risk, control and citizen response," paper presented at the Pacific Sociological Society Meetings, Sacramento, 21 April 1977.
25. R. Davis, *Hydrology and Water Resources in Arizona and Southwest* (Proceedings 1975 Meeting of AWWA and Hydrology Section of Arizona Academy of Sciences, Tucson, 1976), vol. 6, pp. 48–64.
26. R. Davis, paper presented at National Convention of American Society of Civil Engineers, Denver, 10 December 1975.
27. D. Klein, *J. Appl. Meteorol.* **14**, 673 (1975); C. Cooper, *Modifying the Weather, A Social Assessment* (University of Victoria, Victoria, British Columbia, 1973), pp. 99–134.
28. S. Changnon, paper presented at the AAAS Annual Meeting, Denver, 22 February 1977.
29. C. Taylor and E. Swanson, *The Economic Impact of Selected Nitrogen Restrictions* (AERR-133, Department of Agricultural Economics, University of Illinois, Urbana, 1975); C. Taylor, P. Van Blokland, E. Swanson, K. Frohberg, *Two National-Equilibrium Models of Crop Production*. (AERR-147, Department of Agricultural Economics, University of Illinois, Urbana, 1977).

30. We thank R. Davis, J. Haas, J. Ivens, M. Jones, D. Klein, D. Mann, G. Morgan, S. Sonka, C. Taylor, and J. Van Blokland. Many others gave support and critical reviews, including W. Ackermann and 34 attendees at a user's workshop; and the members of the advisory panel including S. Borland, E. Bollay, J. Fivor, W. Fowler, W. Thomas, and C. Wolf. Many students contributed to the background research, and two theses based on various elements of the study were awarded at the University of Illinois. This work was supported by grant ERP 75-09980 from the National Science Foundation and by the State of Illinois.

Highly Reiterated Sequences of SIMIANSIMIANSIMIANSIMIANSIMIAN

Hagai Rosenberg, Maxine Singer, Martin Rosenberg

The DNA of eukaryotes characteristically contains multiple copies of certain nucleotide sequences that occur anywhere from a few to millions of times per genome. The more highly repetitive sequences typically occur in long tandem arrangements and can comprise a substantial portion of the genome (*1–3*). Certain of these sequences (termed satellites) can be isolated directly from total genomic DNA by virtue of their unique buoyant density (reflecting a difference in base composition from the bulk DNA). In other instances, the highly repeated sequences cannot be distinguished by density but can be obtained in relatively pure form by virtue of their rapid reannealing characteristics after denaturation of sheared total DNA. Some of the highly repetitive sequences can also be isolated by digestion of total DNA with restriction endonucleases which cleave at specific sites within the repeated sequence. In most cases studied, the repeating unit has been defined as a relatively short oligonucleotide segment (less than 20 residues) (*2, 3*); however, more recently, the existence of much longer repeat units has been indicated (*4–6*). The highly repetitive sequences often appear localized within the centromeric region of metaphase

Dr. H. Rosenberg was a Visiting Fellow at the National Cancer Institute, Bethesda, Maryland 20014, and is presently at the Israel Institute for Biological Research, Prime Minister's Office, Ness Ziona, P.O.B. 19 Israel. Drs. Singer and M. Rosenberg are members of the Laboratory of Biochemistry and Laboratory of Molecular Biology, respectively, of the National Cancer Society.

0036-8075/78/0428-0394$02.00/0

chromosomes but may also occupy noncentromeric positions (*7–9*). Little is known about the function of these highly repeated sequences, although theories abound. They apparently do not serve as templates for RNA transcription (*10*).

Interesting questions about the structure of the highly repeated sequences remain. Are the various copies of a repeated sequence identical or do they differ? What is the nature and extent of the sequence variations which may occur? Are these sequence divergences random or specific nonrandom alterations? How are variant repeat sequences arranged relative to one another? What is the DNA structure in the genome adjoining these repeat units? Are the number of copies of a repetitive sequence and the arrangement of the variant copies identical in all tissues of a given organism and in various members of a given species? Do common structural features occur between the repetitive sequences of organisms from different species? How are the repetitive sequences organized within the nucleoprotein structure of chromatin?

These questions and others may now be examined more precisely than was allowed by previous techniques. Technical advances offer new approaches for the study of the highly repetitive eukaryotic DNA. The use of DNA restriction endonucleases coupled with methods for rapidly determining the exact nucleotide sequence of DNA segments permit a detailed study of the structure of the repeating units. In addition, the capability now exists for obtaining cloned segments of repetitive DNA and then comparing these with the uncloned repetitive DNA component derived directly from the eukaryote genome. The discovery that, upon infection of mammalian cells, the genomes of certain viruses can recombine with the highly repetitive component of the host DNA and yield efficiently replicating recombinants (*11*) affords one procedure for obtaining useful quantities of a cloned repetitive DNA segment. The studies reported here represent our initial application of these approaches to the highly repetitive DNA of the African green monkey (*Cercopithecus aethiops*). Our results indicate that the highly repetitive DNA component of a eukaryote may be more complex than was previously supposed.

Summary. A 172–base pair segment of DNA that is repeated several million times in the genome of the African green monkey has been characterized. Sequence analysis revealed that the many repeats of this complex unit are not all identical but represent a set of closely related segments: Sequence divergence occurs at various positions in the segment in a nonrandom manner. The uncloned segment obtained from monkey DNA is compared with a cloned segment of the same DNA which was recombined into the genome of simian virus 40 during permissive infection.

Digestion of the total African green monkey DNA with restriction endonuclease Hind III and fractionation of the products by electrophoresis on polyacrylamide slab gels yield a series of discrete fragments derived from highly reiterated DNA (*12–14*). The larger fragments appear to be size multimers of the smallest fragment [designated as AGMr(Hind III)-1], which we show here to be 172 residues in chain length. Similar analyses of the purified α component [defined as a rapidly reannealing fraction of total monkey DNA with a density of 1.699 g/cm^3 (*15, 16*)] indicate that it is composed largely of AGMr(Hind III)-1 and its multimers (*12, 14–16*). In this article, we describe a complete sequence specifying the most abundant nucleotide residue at every position of the uncloned AGMr(Hind III)-1 segment. The data permit certain conclusions and invite additional speculations concerning the organization and extent of sequence divergence of AGMr(Hind III)-1 within the monkey genome. Most important, we conclude that the repeat length of highly repetitive DNA can be quite long, in the present instance 172 base pairs (bp) with few, if any, internal repetitions and that the repeated units consist of a set of closely related but variant sequences. While certain of the possible variations are observed, others are not, suggesting that the set is not simply the result of random divergence from some initial sequence during the course of evolution. Thus, primary sequence information reveals a level of complexity in highly repetitive DNA that could not be readily assessed by hybridization experiments or restriction enzyme analysis alone.

In earlier work (*13*) we reported the characterization and sequence determination of a 184-bp DNA segment obtained from the genome of a defective, substituted variant of simian virus (SV40) that had been serially passaged in monkey cells (*17*). Hybridization data indicated that most of this cloned segment was homologous to highly repetitive monkey DNA (reiteration frequency of 1.6×10^6) and specifically to AGMr-

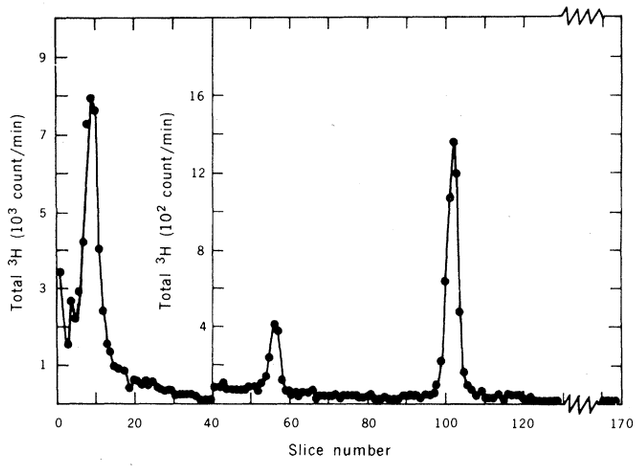

Fig. 1. Products of digestion of BSC-1 DNA with Hind III. Total DNA was isolated (*6*) from uninfected cells of the BSC-1 line (*28*) derived from African green monkey kidney. ^{3}H-Labeled BSC-1 DNA was obtained by adding [^{3}H]thymidine to the culture medium (*33*). ^{3}H-Labeled BSC-1 DNA (10 μg; 10,200 count/min per microgram) was incubated in 50 μl of solution containing 6 m*M* tris-HCl, *p*H 7.5, 6 m*M* $MgCl_2$, 6 m*M* 2-mercaptoethanol, and 50 units of Hind III (New England Biolabs) for 18 hours at 37°C. The mixture was adjusted to contain 0.1 percent sodium dodecyl sulfate and 0.2 mg of transfer RNA per milliliter (final volume, 0.1 ml); it was then extracted two times with phenol, and DNA was precipitated from the aqueous layer with 0.3*M* sodium acetate and 2.5 volumes of ethanol. The precipitated DNA was dissolved in 50 μl of water, and the mixture was subjected to electrophoresis through a cylindrical (6 mm by 18 cm) polyacrylamide gel (5 percent acrylamide; 20:1 acrylamide : bisacrylamide) for 8 hours at 120 volts. The direction of migration was left to right. The electrophoresis buffer consisted of 0.05*M* tris, 0.05*M* boric acid, and 1 m*M* EDTA (*p*H 8.3). At the end of the run, the gel was sliced into 1-mm slices, the slices were air-dried and incubated overnight at 60°C with 0.5 ml of 30 percent H_2O_2, and the radiactivity was counted in a Triton-toluene scintillation fluid. Of the initial 1×10^5 count/min, a total of 0.75×10^5 was recovered from the gel.

(Hind III)-1 and its multimers (*13*), little or no wild-type SV40 sequence was detected. However, the data did not permit firm conclusions concerning the precision or extent of the homology between the fragment isolated from the defective virus and that isolated from monkey DNA. We are now able to compare the sequence of AGMr(Hind III)-1 with that previously obtained for the cloned sequence isolated from the substituted SV40 variant. The DNA insert found in the defective virion appears to be a single member of the set of sequences that comprise AGMr(Hind III)-1. Our results also define the junctions between the AGMr(Hind III)-1 sequence and other sequences within the defective virus and, thus, may provide useful information concerning the recombinational events that lead to the formation of these substituted virions. Unless indicated otherwise, all procedures and materials were as previously described (*13*).

Characterization of DNA Fragment AGMr(Hind III)-1

Exhaustive treatment of total DNA from the BSC-1 line of monkey kidney cells with the restriction endonuclease Hind III yields a set of discrete fragments containing highly repetitive DNA sequences as well as a heterogenous mass of fragments of relatively high molecular weight (Fig. 1) (*12–14*). The two smallest fragments, called here AGMr(Hind III)-1 and -2, are 172 bp and approximately 340 bp long, respectively, and are well resolved by electrophoresis on polyacrylamide gels (slice numbers 102 and 56, respectively, in Fig. 1). Calculation of the percentage of the total ^{3}H-labeled DNA on the gel that is recovered as AGMr(Hind III)-1 or -2 permits an estimation of the relative amount of each fragment in the total genome. The average values obtained from several experiments like that shown in Fig. 1 indicate that AGMr(Hind III)-1 and -2 account for approximately 7 and 1 percent of the total DNA, respectively. Similar experiments with unlabeled DNA isolated from the liver of the African green monkey gave a value of 10 percent for AGMr(Hind III)-1.

The DNA fragment AGMr(Hind III)-1 was obtained in microgram quantities from BSC-1 DNA by exhaustive digestion with Hind III (*13*) and purification by electrophoretic separation on 5 percent polyacrylamide slab gels. AGMr-(Hind III)-1 was eluted from the gels electrophoretically (*18*). Figure 2 shows

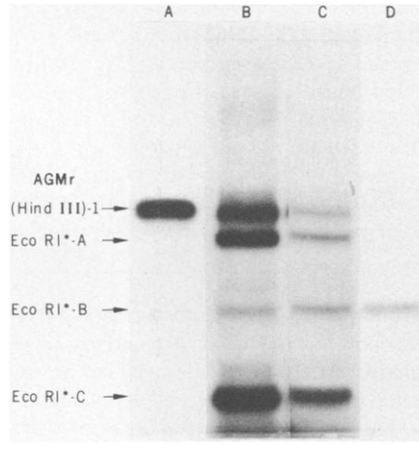

Fig. 2. Cleavage of AGMr(Hind III)-1 fragment with Eco RI*. The 5′-termini of the AGMr(Hind III)-1 fragment were labeled with ^{32}P as described (*13*). The end-labeled fragment was digested under standard Eco RI* conditions (*19*) with 30 units of Eco RI (Miles). The reaction was terminated by adjusting the reaction mixture to 0.01*M* EDTA, 0.5 percent sodium dodecyl sulfate, and 5 percent glycerol, and the products were analyzed directly by electrophoresis on a 5 percent polyacrylamide slab gel. (A) (Hind III)-1 fragment untreated; (B) (Hind III)-1 fragment (0.1 μg) digested for 8 hours; (C) (Hind III)-1 fragment (0.01 μg) digested for 12 hours; (D) Eco RI*-A fragment eluted from the gel (column B) and cleaved again as in (B).

the results obtained when AGMr(Hind III)-1 is treated with restriction endonuclease Eco RI* [that is, Eco RI under conditions of relaxed specificity (*19*)]. Cleavage occurs initially (Fig. 2B) at positions (31 to 35) in the sequence (see Fig. 3), yielding fragment Eco RI*-C (35 residues, that is, base pairs), and fragment Eco RI* -A (145 residues); a small amount of fragment Eco RI*-B (80 residues) is also formed (Fig. 2B). With higher enzyme concentrations and longer reaction times relatively more Eco RI*-B accumulates (Fig. 2C) as a result of cleavage at positions (95 to 98) [the additional expected 60-bp-long fragment is not seen since the AGMr(Hind III)-1 was terminally labeled]. Exhaustive treatment of purified Eco RI*-A with Eco RI* results in complete conversion to Eco RI*-B (Fig. 2D). Thus, the two Eco RI* sites [AATT (A, adenosine; T, thymidine) (*19*)] are not equally susceptible to cleavage, presumably because of the one different flanking base pair.

AGMr(Hind III)-1 is also cleaved by restriction endonuclease Hph (*20*) into two discrete fragments estimated to be approximately 35 and 145 residues in length by mobility on polyacrylamide electrophoresis (data not shown, but see Fig. 3).

Exhaustive treatment of AGMr-(Hind III)-1 with either Eco RI* or Hph results in complete digestion, indicating that the fragment is essentially free of unrelated DNA segments.

The Nucleotide Sequence of AGMr(Hind III)-1

Both RNA and DNA sequencing techniques were used to determine an unambiguous sequence for the 172 residues of the uncloned AGMr(Hind III)-1 fragment (Fig. 3). A ^{32}P-labeled complementary RNA (cRNA) transcript of the DNA fragment was prepared and characterized by procedures described previously (*13*). Polyacrylamide gel electrophoresis of the cRNA (*13*) afforded purification of the major product, which was an essentially full-length copy of one strand of AGMr(Hind III)-1. The oligonucleotide products generated from the major cRNA by both complete and partial digestion with T1 and pancreatic ribonucleases (Fig. 4, B and D) were analyzed by standard RNA sequencing techniques (*18, 21*) as described (*13*).

Many, but not all, of the oligonucleotides were identical to those previously obtained and characterized in our analysis of the cRNA made to a DNA fragment designated (Hind II/Hind III)-E, isolated from the substituted defective SV40 variant (Fig. 4, A and C) (*13*). The nucleotide sequence of (Hind II/Hind III)-E as well as that of an additional DNA segment contiguous to (Hind II/Hind III)-E in the defective SV40 variant [(Hind II/Hind III)-C] is presented at the bottom of Fig. 3 for comparison.

Direct DNA sequence analysis was also carried out on AGMr(Hind III)-1 with both partial snake venom phosphodiesterase analysis (*22, 23*) and the dimethyl sulfate–hydrazine (DMS-HZ) procedures of Maxam and Gilbert (*24*). For these procedures the AGMr(Hind III)-1 was first labeled with ^{32}P at both its 5′-hydroxy termini as described (*13, 22*). The ^{32}P-labeled material was then digested with either Hph or Eco RI*; each of the resulting radioactive fragments was purified by gel electrophoresis (Fig. 2) and subjected to sequence analysis from its labeled Hind III terminus. The relative positions of the Hph and Eco RI* sites allowed sequence data (Fig. 3) which confirmed the structures of these restrictions sites to be obtained.

Analysis of the products of partial snake venom exonuclease digestion by both two-dimensional "homochromatography" (*21*) (Fig. 5) and one- and two-

dimensional paper electrophoresis gave nucleotide sequences for 20 to 25 residues at each end of AGMr(Hind III)-1. Direct DNA sequencing by the DMS-HZ techniques (*24*) gave sequences for 80 to 100 residues of each strand from its 5′ Hind III end. The sequences were deduced from a comparison of several gel patterns, representative examples of which are shown in Fig. 6.

The RNA and DNA sequencing methods gave completely consistent results. In conjunction, these methods allow deduction of an unambiguous sequence with far lower probability of error than would be the case with either method alone. The sequence of AGMr(Hind III)-1 is given in Fig. 3.

It is important to emphasize that AGMr(Hind III)-1 was not cloned but was derived from total monkey DNA. Nevertheless both the RNA and DNA sequencing methods permitted the unambiguous assignment of a single predominant nucleotide residue at each position in the fragment. The significance of this observation is discussed below.

Characteristics of AGMr(Hind III)-1

As we have indicated (*13*), the sequences defined as AGMr(Hind III)-1 are presumably derived from the most abundant class of African green monkey repetitive DNA, namely, the α component (*12, 14–16, 25*). Treatment of α component with Hind III results in the almost complete conversion of the high-molecular-weight DNA to a fragment the same size as AGMr(Hind III)-1 as well as the characteristic multimers (*12, 14–16*). The α component, of which 77 percent corresponds to the Hind III monomer (*14*), accounts for about 13 (*14*) to 20 (*16*) percent of total African green monkey DNA.

We have estimated that AGMr(Hind III)-1 represents about 7 and 10 percent of the total DNA from BSC-1 cells and monkey liver, respectively. The location of the Eco RI* site at residues 32 to 35 within the sequence of AGMr(Hind III)-1 (Fig. 3) is consistent with previous reports on the location of one Eco RI* site relative to the Hind III site in α component (*14–16*). Another report (*12*) indicated an Eco RI site at the same relative position, but it seems likely that the digestion conditions in fact permitted relaxed specificity for the endonuclease. Previous reports (*12, 14–16*) demonstrated that the Hind III

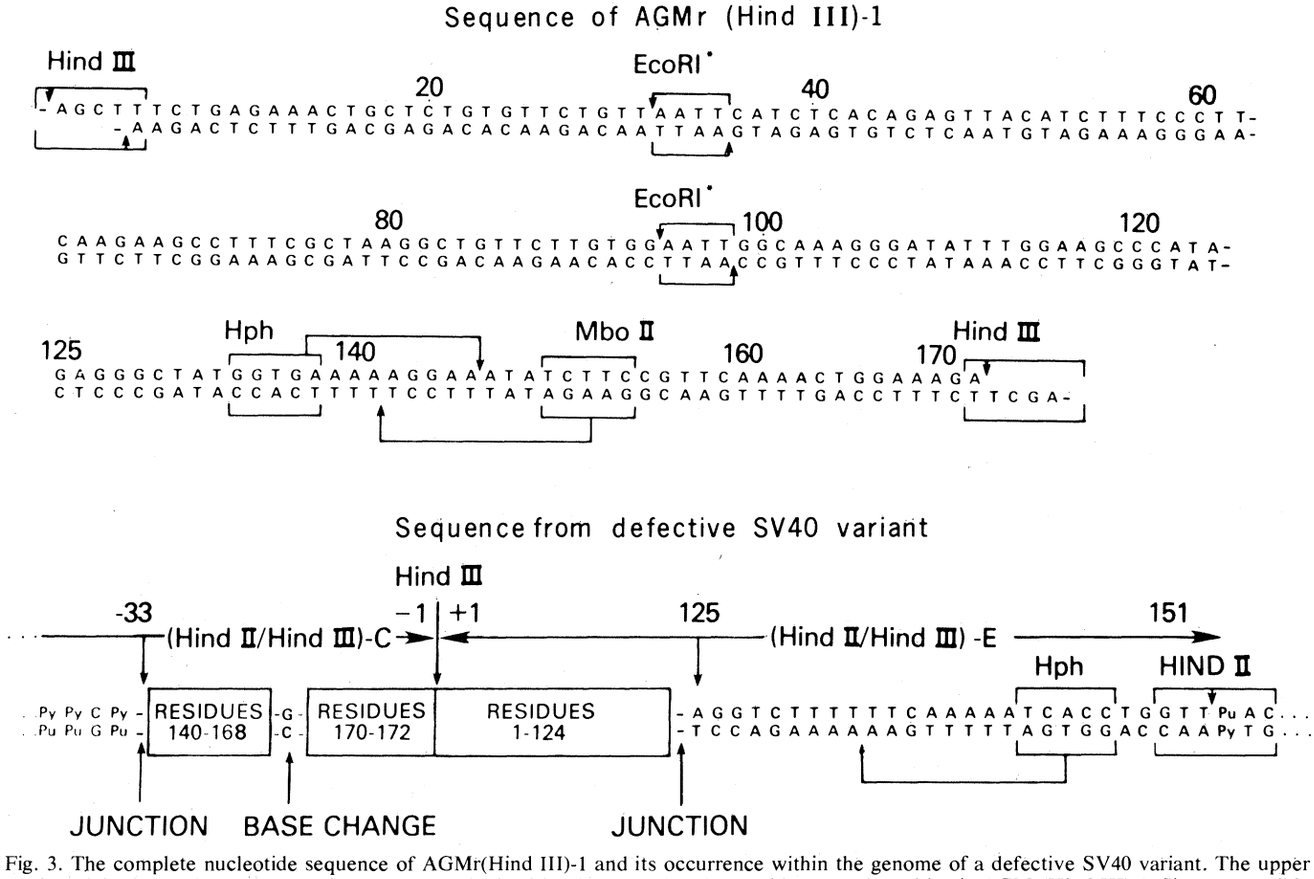

Fig. 3. The complete nucleotide sequence of AGMr(Hind III)-1 and its occurrence within the genome of a defective SV40 variant. The upper section of the figure shows the complete sequence obtained for the most abundant residue at each position in AGMr(Hind III)-1. Sites susceptible to cleavage by known restriction endonucleases are shown. Inverted repeat sequences (dyad symmetries $\geq$ 5 bp) occur at positions 4 to 8 and 10 to 14; 68 to 73 and 78 to 84; 117 to 121 and 127 to 131; 120 to 125 and 130 to 135; 107 to 113 and 145 to 151. Those that contain unsymmetrical core sequences give rise to potential stem and loop structures in the single-stranded cRNA or DNA. True palindromes (>7 bp) occur at positions 140 to 147 and 158 to 165. The significance of these various symmetries is unknown. The lower section of the figure shows, schematically, the relation between the cloned sequence of monkey DNA isolated from the defective SV40 variant and the sequence determined for uncloned AGMr(Hind III)-1. The sequences shown were derived from two restriction endonuclease fragments, termed (Hind II/Hind III)-C and -E that are contiguous in the defective SV40 (*17*) and separated by a Hind III site. The (Hind II/Hind III)-C sequences are given negative numbers, starting from the Hind III site and going leftward. The (Hind II/Hind III)-E sequences are given positive numbers, also starting from the Hind III site but going rightward. In the diagram, residues identical to those in AGMr(Hind III)-1 are given in blocks showing the corresponding numbers in the AGMr(Hind III)-1 sequence shown above. Residues that differ from those in AGMr(Hind III)-1 are shown. The term "junction" describes those positions where AGMr(Hind III)-1 sequences end and sequences of unknown origin begin. The determination of the sequences in the defective SV40 are described in (*13*).

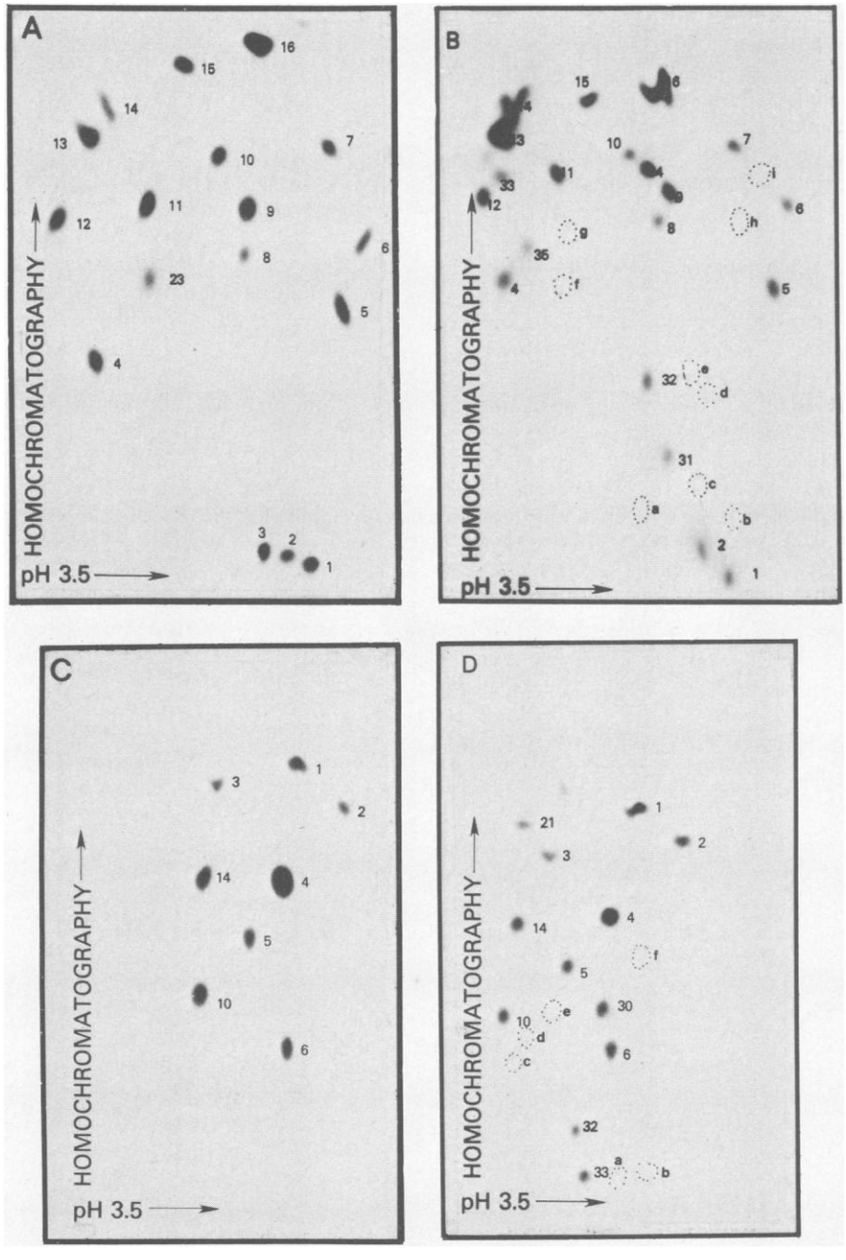

Fig. 4. Autoradiographs of two-dimensional chromatographic oligonucleotide patterns ("fingerprints") of T1 and pancreatic ribonuclease digestion products of cRNA prepared with AGMr(Hind III)-1 [(B) T1 ribonuclease; (D) pancreatic ribonuclease] and with fragment (Hind II/Hind III)-E of the defective SV40 variant [(A) T1 ribonuclease; (C) pancreatic ribonuclease)]. All procedures used for the preparation, purification, and "fingerprinting" of the cRNA's have been described (*9, 13*). The cRNA products (labeled with [α-^{32}P]GTP) were found to be predominantly asymmetric, nearly full-length transcripts of the DNA fragments and were readily purified by electrophoresis on 5 percent polyacrylamide gels in 8*M* urea (*13*). The oligonucleotide patterns shown were obtained from the purified predominant cRNA species (*13*). The numbering of the oligonucleotide products corresponds to the numbering previously used in reporting the sequences in the defective SV40 (*13*). Thus, the same numbers are used in (A) and (B) and in (C) and (D) to indicate identical oligonucleotides. Numbers 30 and above in (B) and (D) are those oligonucleotides found only in AGMr(Hind III)-1. In (B), oligonucleotides 1 and 2 are incorrectly numbered; they should be numbered 2 and 3, respectively. The stippled circles in (B) and (D) (oligonucleotides marked with letters) depict certain minor oligonucleotide products (see text) which are reproducibly and uniquely characteristic of the cRNA made with the AGMr(Hind III)-1 fragment: these minor oligonucleotides do not occur in the cRNA made with the (Hind II/Hind III)-E fragment of the defective SV40 and, in particular, do not occur among the minor oligonucleotides that were previously observed (after long radiographic exposure) and characterized [see, for example, figure 3 in reference (*13*)] and are derived from transcription of the opposite strand of (Hind II/Hind III)-E fragment. T1 ribonuclease oligonucleotides *a, b, c, f,* and *h* (B) were characterized further: after elution they were digested with pancreatic ribonuclease, and the resulting products were analyzed by chromatography. Similarly pancreatic ribonuclease oligonucleotides *a, e,* and *f* (D) were analyzed by digestion with T1 ribonuclease.

monomer derived from the α component is arranged in multiple tandem repeats within the α component, and that the configuration of the AGMr(Hind III)-1 sequence within the defective SV40 variant that we studied is consistent with a tandem arrangement in the monkey genome (see below).

AGMr(Hind III)-1 contains a strikingly asymmetric distribution of purine and pyrimidine residues: examination of either strand indicates that these stretches vary from between 6 and 12 consecutive purines or pyrimidines. There are also several regions of dyad symmetry and several true palindromes in AGMr-(Hind III)-1 (see Fig. 3).

The nucleotide sequence of AGMr-(Hind III)-1 is markedly different from the sequences reported previously for highly reiterated satellite and spacer regions of eukaryote genomes (*2*) since it contains no extensive internal repeats of relatively short oligonucleotide segments. The existence of satellite and highly repetitive DNA's lacking extensive internal repetition has previously been inferred from hybridization data and restriction endonuclease analysis of DNA from a variety of species (*4–6*). Thus, two different classes of highly repetitive sequences exist; one class, the simple-sequence repetitive DNA, is characterized by extensive repeats of short oligonucleotide segments and the other class, the complex-sequence repetitive DNA, is characterized by long repeat length with little internal repetition. Both classes may exist in a single species (*4*). Whether or not the two structural classes reflect different functions is not known. Indeed, as pointed out by Swift (*26*), little is known about the function of the highly repeated sequences, the many current theories notwithstanding.

The function of the highly repetitive α component in monkey DNA is also unknown. There is evidence suggesting that sequences homologous to highly repetitive African green monkey DNA are rare in primates except in members of the same genus (*27*). The sequences of α component [and thus of AGMr(Hind III)-1] are located in centromeres of many of the chromosomes of the species (*8, 9*) and also in some chromosome arms (*9*). There is some evidence suggesting that variants of the sequence of AGMr(Hind III)-1 exist (this article; *14, 15*) and that particular variants of the sequence may be clustered together (*14*). But the nature and significance of the variant sequences is not understood. As described below, our data suggest that the observable variations may be nonrandom.

Sequence Divergence Within AGMr(Hind III)-1

The DNA fragment AGMr(Hind III)-1 was not cloned but, rather, was isolated directly from total cellular DNA after Hind III digestion and thus must be presumed to contain a family of related sequences. Sequence variation might occur at any or all positions within the molecule, and the overall extent of variation will depend on the number of base changes as well as any relative amplification of particular diverged sequences that may have occurred over time. In this context, time includes both evolutionary time and time since the original establishment of the BSC-1 line in 1961 (*28*). The data obtained from both the RNA and DNA sequencing techniques indicate that a single unique nucleotide residue predominates at each position within the fragment population (Fig. 3). Thus, we can eliminate the possibility that the residue at any position or positions varies to such a degree as to make the nucleotide assignment at that position ambiguous. Thus, if AGMr(Hind III)-1 comprises a population of *n* different sequences, most of them have the residues shown in Fig. 3 in most positions. Nevertheless, it is important to recognize that none of the *n* members of the population need have the sequence reported in Fig. 3; indeed that sequence may not occur at all.

To what extent can the sequencing methods used reveal relatively low levels of divergence from the predominant residue at each position? Infrequent variation in sequence within the fragment population would be observed as some increased level of background "noise"—that is, minor bands in the DNA sequencing gels or minor spots in the two-dimensional RNA oligonucleotide patterns ("fingerprints")—in the autoradiographic analysis of the data. The extent to which such "noise" would be evident depends on the relative frequency of a particular divergent sequence and on the sensitivity of the sequencing methods employed.

We previously sequenced, by identical techniques, DNA segments containing sequences homologous to AGMr-(Hind III)-1 but derived from a defective substituted SV40 variant (*13*). These DNA fragments represent cloned African green monkey DNA segments presumably derived from a single site within the monkey genome. Direct comparison of the noise levels observed in the data from the SV40 variant fragments with those observed with the uncloned monkey fragment give some indication as to

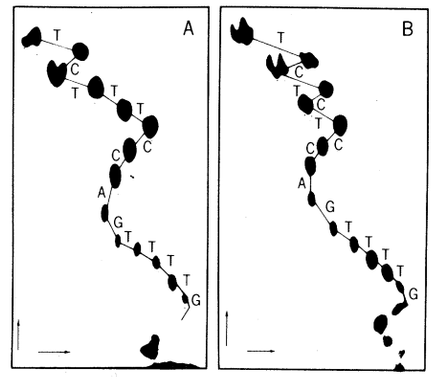

Fig. 5. Autoradiograph of the two-dimensional fractionation of the products resulting from partial digestion with snake venom phosphodiesterase of (A) the AGMr(Hind III)-1 Eco RI*-A fragment (see Fig. 2) and (B) the (Hind II/Hind III)-C fragment of the SV40 defective variant. The fragments were labeled with ^{32}P at their 5′-hydroxyl Hind III termini. All procedures were as described (*13*). In both (A) and (B) the smallest oligonucleotide observed (upper left) is pApGpCpT [characterized as described (*13*)], and the indicated residues denote the single 3′-terminal nucleotide sequentially removed by 3′-exonucleolytic action from the increasingly larger products. The only difference between (A) and (B) is at position 4, which is eight residues from the labeled 5′-terminus.

the extent and possible specificity of sequence divergence within AGMr-(Hind III)-1. Inspection of the DMS-HZ data indicates that sequence alterations occurring at or below the 10 to 15 percent level at each residue would be difficult to detect. This is due to the relatively high background noise inherent in the procedure (*24*) even when applied to a homogeneous fragment. At this level of detection, no sequence divergence was observed in the analysis of AGMr-(Hind III)-1.

The methods that employ specific enzymatic cleavage and two-dimensional fractionation of oligonucleotide products allow a more sensitive and detailed analysis (that is, T1 and pancreatic ribonuclease and resolution of the products of partial digestion with snake venom phosphodiesterase). Since the RNA polymerase used to prepare the cRNA favored selective initiation at one end of both the cloned (*13*) and uncloned DNA fragments (near position 10 of AGMr(Hind III)-1), it can be assumed that the cRNA is representative of all or almost all of the molecules in the population. Thus each T1 or pancreatic ribonuclease oligonucleotide product resolved in a two-dimensional "fingerprint" of the digested gel-purified cRNA provides data on the relative occurrence of that group of contiguous nucleotides in the fragment population. Sequence divergence within the DNA template would result in a shift of the relative positions of the corresponding altered oligonucleotides in the "fingerprint." Ability to detect these infrequent oligonucleotides as minor background spots ("noise") depends on the nature of a given sequence alteration and its frequency of occurrence within the population as well as on the resolution in the "fingerprint." In regions where resolution is good, such oligonucleotide products can be detected at approximately the 1 percent level. Close examination of two-dimensional oligonucleotide patterns obtained from several cRNA and partial snake venom phosphodiesterase digestions of AGMr(Hind III)-1 consistently indicated a variety of reproducible minor "background" oligonucleotides. These oligonucleotides were not detected in the identical analyses of the cloned monkey fragment derived from the SV40 variant (*13*) (Fig. 4) nor could they be identified as those oligonucleotides that are known to result from opposite strand transcription of the fragment. The location of "opposite strand" oligonucleotides is in fact known from the sequencing data on the fragment from the defective virion (*13*). It is unlikely that the minor oligonucleotides arise from transcripts of contaminating unrelated sequences, in view of the apparent purity of AGMr-(Hind III)-1.

Most of the minor "background" oligonucleotides could not be analyzed further, and thus we can only suggest that their unique occurrence in the "fingerprints" prepared from the cRNA to the uncloned AGMr(Hind III)-1 fragment and their relative positions within these "fingerprints" is consistent with their being minor variants of the major oligonucleotides. A few of the minor oligonucleotides, however, were well-enough resolved and present in sufficient quantity (at 2 to 5 percent level) to be further characterized (Fig. 4, legend). These analyses were consistent with their having single nucleotide alterations from the corresponding major sequence (as is discussed below). In addition, many possible sequence variants that would have been well resolved were not detected, suggesting relatively low levels (less than 1 percent) of divergence at these residue positions. The presence of only specific sets of minor oligonucleotides and their variability in relative intensities suggests that these products arise from one or more of the *n* sequences comprising AGMr(Hind III)-1 and that the sequence changes within this highly reiterated DNA are selective rather than random. However, the techniques employed were not suitable for

monitoring or quantifying the number or extent of sequence changes precisely.

One of the predominant minor oligonucleotides (Fig. 4B, *a*) had a sequence consistent with a change from T·A to G·C (G, guanosine; C, cytidine) at position 31 in the sequence at about the level of 4 to 5 percent. Although changes of T·A to A·T or C·G at this position could also have been observed, no such changes were detected in the "fingerprints." This alteration is particularly interesting in that it reflects the conversion of one Eco RI* site (position 32 to 36) to an Eco RI site (31 to 37) and predicts that a small percentage of AGMr(Hind III)-1 should be cleaved into two discrete fragments 32 and 135 nucleotides long by Eco RI. Using an amount of ^{32}P-end-labeled AGMr(Hind III)-1 such that cleavage of as little as 1 percent of the molecules could be readily detected (Fig. 7), we found that a small percentage (2 to 5 percent) of the fragment was indeed cleaved into fragments of the appropriate size. No other cleavage products were detected, although a similar single nucleotide change from G·C to C·G at position 99 in AGMr(Hind III)-1 would also have produced an Eco RI site. This alteration was not detected either by inspection of "fingerprints" or by treatment of ^{32}P-end-labeled AGMr(Hind III)-1 with Eco RI. These data suggest some nonrandom variation in the AGMr(Hind III)-1 sequence.

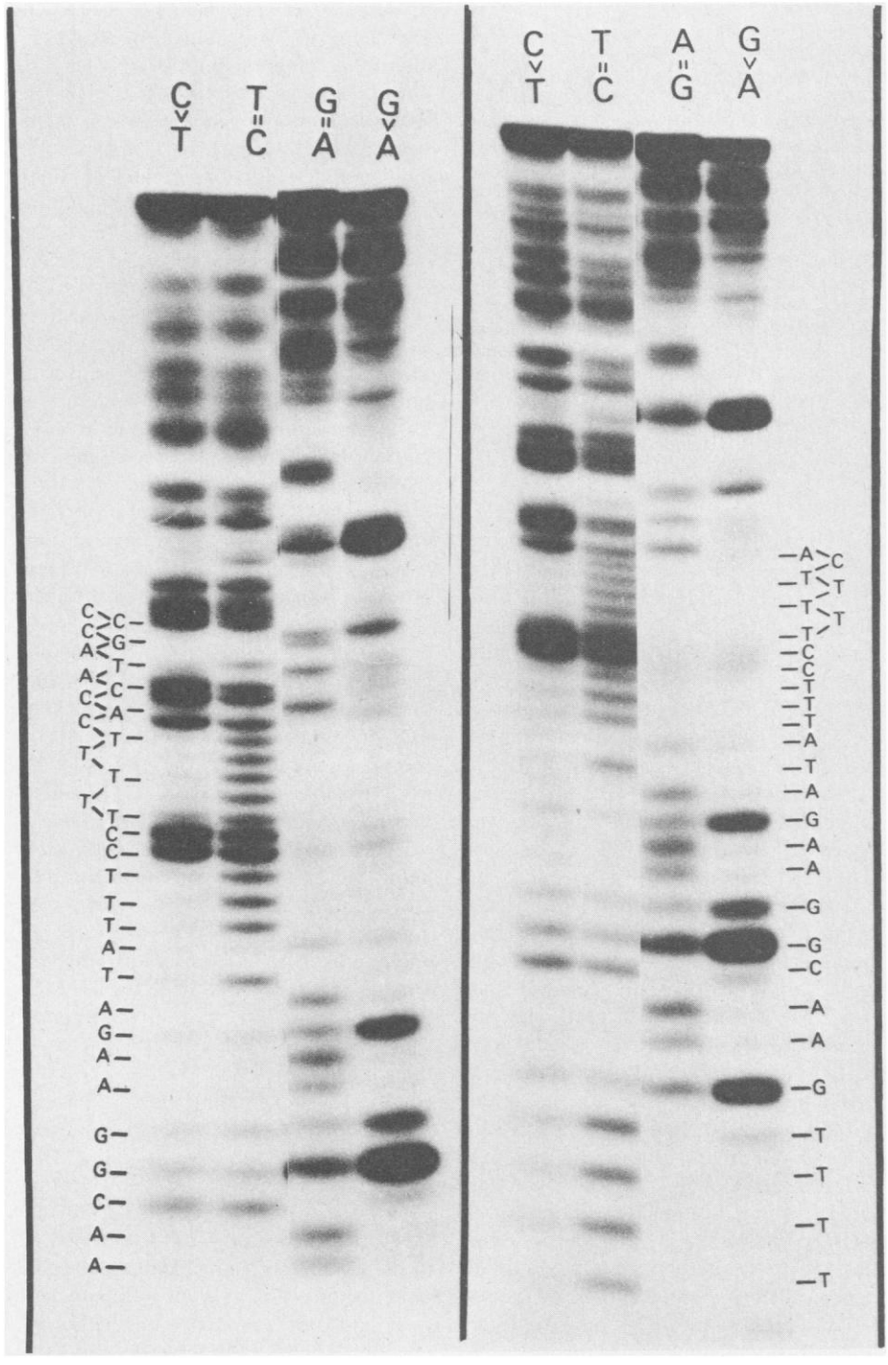

Fig. 6. Representative autoradiograph of a DNA sequencing gel displaying the results of the chemical sequencing methods (*24*) as applied to the AGMr(Hind III)-1 Eco RI*-A fragment (see Fig. 2), labeled at its Hind III end. The two sides of the figure represent identical samples subjected to different times of electrophoresis. On the right, the first discernible residue (at the bottom) is the T at position 163 (Fig. 3), 14 residues from the ^{32}P-labeled 5′ end. On the left, the first discernible residue (at the bottom) is the A at position 158 (Fig. 3), 19 residues from the labeled 5′ end.

Similar analyses made with other restriction enzymes lend support to the suggestion of nonrandom divergence. For example, there are seven sites within AGMr(Hind III)-1 where a single base change would yield a site for restriction endonuclease Taq I (TCGA) (*29*). However, cleavage of AGMr(Hind III)-1 labeled at both 5′ termini with ^{32}P with Taq I results in the conversion of about 2 to 4 percent of the fragment to only two discrete bands approximately 64 and 110 bp in length (Fig. 7C). The size calculation from gel mobility is not precise enough to distinguish among the potential cleavage sites at positions 62 to 65, 74 to 77, 113 to 116, or 123 to 126. However, only one of the potential Taq I sites has presumably diverged at a level greater than 1 percent to an authentic Taq I site. Examination of the minor cRNA oligonucleotides indicates that one T1 ribonuclease product and one pancreatic ribonuclease product, both present at the level of about 2 percent (Fig. 4B, *b* and Fig. 4D, *e*, respectively) appear consistent with a base change of A·T to G·C at position 64 of the sequence. This would, in turn, be consistent with the Taq I cleavage site observed by gel analysis; however, because of insufficient material, this alteration could not be confirmed. The cRNA analysis also indicated that the region of AGMr(Hind III)-1 between positions 102 and 115 has little detectable divergence (that is, < 1 percent), and this probably rules out Taq I cleavage at residues 113 to 116. Clearly, the data again suggest different extents and specificities of sequence divergence among the potential Taq I cleavage sites.

The restriction endonuclease Hae III cleaves from 1 to 3 percent of AGMr-(Hind III)-1 (labeled at both 5′-termini with ^{32}P) to yield three fragments approximately 40, 65, and 100 bp in length (Fig. 7). We assume that another

fragment, about 130 to 140 bp in length, was also produced but is obscured by the undigested material. There are seven single base-pair changes that might occur in AGMr(Hind III)-1 to yield an Hae III site (GGCC) (*30*). Only a limited number of the possible sites are detectable. The observed fragments might have arisen from base changes within the potential Hae III sites at positions (68 to 71), (99 to 102), (127 to 130), or (128 to 131). Cleavage at potential sites (81 to 84), (117 to 120), and (118 to 121) was not observed.

It should be pointed out that the use of restriction enzyme cleavage to monitor and (more importantly) to quantitate the existence of specific recognition sites occurring at low frequencies may be somewhat misleading. The ability of certain restriction enzymes such as Eco RI (*19*) and Bsu (*31*) to exhibit altered or relaxed specificity is well documented, and we have already demonstrated that Eco RI* cleaves AGMr(Hind III)-1 more readily at position 32 than at position 95. In fact, Grüss and Sauer (*12*) reported that all of AGMr(Hind III)-1 is cleaved by Eco RI to yield a product about 20 bp shorter than the original fragment. Similarly Fittler (*14*) has reported that 12 to 18 percent of AGMr(Hind III)-1 is cleaved by Eco RI and, further, that 9 to 14 percent is cleaved by Bsu (an isoschizomer of Hae III) to yield fragments about 45 and 131 bp long. Our results are in contrast to these reports and suggest that the digestion conditions used by these workers (*12, 14*) allowed cleavages by the two enzymes at imperfect recognition sites. The restriction endonuclease Bsu, like Eco RI, is known to exhibit reduced specificity (*31*), whereas Hae III has been reported to demonstrate stringent specificity for the sequence GGCC (*31*).

AGMr(Hind III)-2 as well as the higher multimers of AGMr(Hind III)-1 represent variations of the AGMr(Hind III)-1 sequence in which one or more of the six base pairs in a Hind III site is altered. From the relative amounts of AGMr-(Hind III)-1 and -2, about 7 percent of the Hind III sites are divergent; additional data on the sequence of AGMr-(Hind III)-2 will be required before any statements on the nature of the divergence can be made. Studies on the cleavage of the altered Hind III site by restriction endonuclease Alu I (*14*) were interpreted as consistent with random distribution of base pair variation; however, they could also be explained by nonrandom divergence.

Thus, all the available data are consistent with nonrandom divergence among the *n* sequences comprising AGMr-(Hind III)-1. Regarding the origin of the divergent sequences, we cannot differentiate between selective, nonrandom base changes and selective amplification of certain initially randomly altered sequences.

We point out, however, that a nonrandom population, even if the number of the various sequences comprising the AGMr(Hind III)-1 population is large and their relative abundances greatly different, is consistent with some selective pressure on the evolution of the sequences.

AGMr(Hind III)-1 versus Monkey Sequences from Defective SV40 Variant

Figure 3 shows the sequence of AGMr(Hind III)-1 as well as that of a 184-bp-long segment isolated from a substituted defective variant of SV40 (*13, 17*). Residues 1 to 124 in AGMr(Hind III)-1 are identical to the 124 residues starting at the Hind III site in the fragment from the defective virus. Residues 125 to 139 of AGMr(Hind III)-1 are not found in the defective variant. However, residues 140 through 172 of AGMr-(Hind III)-1 are, with the exception of one base pair, identical to residues −1 to −33 on the opposite side of the Hind III site in the substituted defective variant. Beyond residue position −33 from the Hind III site, the sequence of the defective differs from AGMr(Hind III)-1. Thus, it appears that the defective virion contains a single contiguous segment of the repeated AGMr(Hind III)-1 sequence (residues 140 to 172/1 to 124) spanning the Hind III cleavage site. Furthermore, comparison of the two sequences allows us to define the junction points within the defective variant between the inserted AGMr(Hind III)-1 sequence and adjacent sequences, as well as a single nucleotide base change within the reiterated sequence.

The A·T base pair which occurs at position 169 of AGMr(Hind III)-1 is found as a G·C pair (position −3) in the sequence of the defective virus (see Fig. 5). While no evidence for the existence of a minor component of AGMr(Hind III)-1 containing a G·C pair at this residue was detected by the sequencing techniques, such a component may be present as a rare species. It is, of course, possible that the base change resulted from a mutation that occurred after insertion of the monkey sequence into the SV40 variant.

The AGMr(Hind III)-1 segment occurs in tandem multiple repeats in the monkey genome (see above). The sequences shown in Fig. 3 indicate that the AGMr(Hind III)-1 segment was incorporated into the defective SV40 variant as a segment starting at residue 140 of AGMr(Hind III)-1 and ending at residue 124 and omitting residues 125 to 139. Thus residues 140 and 124 represent junctions between the AGMr(Hind III)-1 sequence and other sequences within the defective SV40 genome. The junction at residue 140 (corresponding to residue −34 leftward from the Hind III site in the defective fragment) may represent a junction between highly reiterated and infrequently reiterated sequences within the monkey genome itself, since no wild-type SV40 sequences are detectable for at least an additional 200 bp (*17*). Alternatively, multiple recombinations may have occurred during the evolution of the defective and the junction would then not represent any preexisting sequence. The other junction, at residue 124, may reflect a similar situation. As we have indicated previously, the residues beyond 124 from the Hind III site in the defective virus do not appear to be wild-type SV40 sequences (*13*).

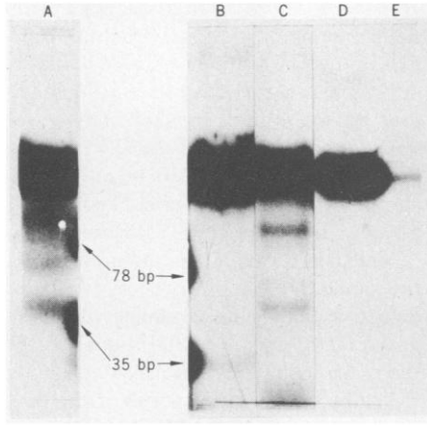

Fig. 7. Autoradiograph of a 5 percent polyacrylamide gel fractionation of the products resulting from a limit digestion of AGMr-(Hind III)-1 with restriction endonucleases Hae III, Eco RI, and Taq I. AGMr(Hind III)-1 labeled at both 5′ Hind III termini with ^{32}P (10^5 count/min; about 12 ng) was incubated (final volume 50, μl) for 8 hours under the following conditions: (A) Hae III, 5 units (*34*), in tris-HCl, *p*H 7.4 (6.6 m*M*), $MgCl_2$ (6.6 m*M*), β-mercaptoethanol (6.6 m*M*) at 37°C; (B) Eco RI, 10 units (Miles) in tris-HCl, *p*H 7.5 (100 m*M*), NaCl (50 m*M*), $MgCl_2$ (10 m*M*) at 37°C; (C) Taq I, 5 units (BRL), in Hepes, *p*H 8.4 (10 m*M*), $MgCl_2$ (6 m*M*), β-mercaptoethanol (6 m*M*), ammonium sulfate (25 m*M*) at 55°C; (D) untreated AGMr(Hind III)-1; (E) same as (D) but only 1 percent as much (about 10^3 count/min). Reactions were terminated (as in Fig. 2) and the mixtures were applied directly to the gel. The indicated size markers (only a portion of the track containing the marker is shown) are the products of digestion of end-labeled AGMr(Hind III)-1 with Eco RI*.

It may be of interest to point out that residue 124 of AGMr(Hind III)-1 is also the center of a dyad symmetry in the DNA running from residue 117 through 131 (see Fig. 3). In addition there are other interesting structural features surrounding this region, including other dyad symmetries, true palindromes, and symmetrically arranged restriction sites. Whether these interesting symmetries are significant in relation to the events leading to the formation of the defective remains to be determined. It is also interesting that small regions of homology occur between AGMr(Hind III)-1 and sequences (*32*) near the origin of replication of wild-type SV40 DNA. Residues 64 to 99 in AGMr(Hind III)-1 show about 70 percent homology with sequences between residues 317 and 362 of the wild-type sequence (*32*).

Thus, our experiments demonstrate the recombination of a relatively large unaltered segment of the monkey genome with SV40 DNA. The monkey fragment inserted within the defective may in fact be longer than 184 bp; the analysis of adjacent sequences remains to be determined. It is still unclear whether the incorporation of particular monkey sequences into the SV40 genome results in any selective advantage for the particular defective virion. Also the molecular mechanisms by which these intact blocks of DNA may be recombined in eukaryote cells remain obscure.

Conclusion

Digestion of African green monkey DNA with restriction endonuclease Hind III results in conversion of about 7 percent of the total DNA to fragments 172 bp long. These fragments are derived from the most highly reiterated portion of the DNA that was previously characterized as α component DNA and they comprise a set of closely related sequences. A unique nucleotide sequence representing the most abundant residue at each of the 172 positions but not necessarily the structure of any particular member of the set, has been determined. The sequence is not internally repetitive, thus distinguishing this repetitive DNA from the characterized satellites of certain other species and supporting the existence of two classes of highly repetitive sequences, simple and complex. Although a single nucleotide residue predominates at each position within the sequence, the data clearly indicate some level of divergence among the population of fragments. Thus, the isolated fragments represent a mixture of molecules of identical chain length which differ by the presence of a nucleotide other than the most abundant one at one or more positions. No particular divergent member represents more than a few percent of the total.

The divergence from the determined sequence within the members of the set is not random. That is, divergence at certain nucleotide residues occurs more frequently than divergence at other residues. The available data indicate that at least 90 percent of the molecules in the isolated set contain the most abundant nucleotide at each position, but that some positions probably contain the most abundant nucleotide in more than 99 percent of the molecules. While direct quantitation and hybridization kinetics indicate that the set as a whole occurs several million times within the monkey genome, the reiteration frequency of particular members of the set is not known.

A defective variant of SV40 that is substituted with both highly and infrequently reiterated monkey DNA sequences contains 157 of the 172 residues found in the fragments isolated by cleavage of monkey DNA with Hind III. In contrast to the uncloned fragments obtained directly from the monkey DNA, the cloned DNA preparation from the defective shows no evidence of divergence. Comparison of the sequences of the cloned and uncloned DNA segments indicates that the sequences are identical except at one position. Thus, the particular DNA segment which was recombined into the defective viral DNA may be derived from one of the more infrequent members of the set comprising the monkey fragments.

References and Notes

1. For reviews, see E. H. Davidson and R. J. Britten, *Q. Rev. Biol.* **48**, 565 (1973); R. C. Angerer, E. H. Davidson, R. J. Britten, *Cell* **6**, 29 (1975).
2. P. A. Biro, A. Carr-Brown, E. M. Southern, P. M. B. Walker, *J. Mol. Biol.* **94**, 71 (1975); S. A. Endow, M. L. Polan, J. G. Gall, *ibid.* **96**, 665 (1975); W. Salser *et al., Fed. Proc. Fed. Am. Soc. Exp. Biol.* **35**, 23 (1976); G. G. Brownlee, E. M. Cartwright, D. D. Brown, *J. Mol. Biol.* **89**, 703 (1974); D. Brutlag, R. Appels, E. S. Dennis, W. J. Peacock, *ibid.* **112**, 31 (1977).
3. W. J. Peacock, D. Brutlag, E. Goldring, R. Appels, C. W. Hinton, D. L. Lindsley, *Cold Spring Harbor Symp. Quant. Biol.* **38**, 405 (1973).
4. M. Carlson and D. Brutlag, *Cell* **11**, 371 (1977).
5. M. Botchan, *Nature (London)* **251**, 288 (1974).
6. G. Roizes, *Nucleic Acid Res.* **3**, 2677 (1976).
7. M. L. Pardue and J. G. Gall, *Science* **168**, 1356 (1970).
8. D. M. Kurnit and J. J. Maio, *Chromosoma* **42**, 23 (1973).
9. S. Segal, M. Garner, M. F. Singer, M. Rosenberg, *Cell* **9**, 247 (1976).
10. J. R. Hartman, O. Laub, Y. Aloni, E. Winocour, personal communication; E. L. Kuff, F. J. Ferdinand, E. Khoury, *J. Virol.* **25**, 25 (1978). W. G. Flamm, P. M. B. Walker, M. McCallum, *J. Mol. Biol.* **40**, 423 (1969).
11. S. Lavi and E. Winocour, *J. Virol.* **9**, 309 (1972).
12. P. Grüss and G. Sauer, *FEBS Lett.* **60**, 85 (1975).
13. M. Rosenberg, S. Segal, E. L. Kuff, M. F. Singer, *Cell* **11**, 845 (1977).
14. F. Fittler, *Eur. J. Biochem.* **74**, 343 (1977).
15. F. L. Brown, P. R. Musich, J. J. Maio, *J. Mol. Biol.*, in press.
16. J. J. Maio, *ibid.* **56**, 579 (1971).
17. G. R. K. Rao and M. F. Singer, *J. Biol. Chem.* **252**, 5124 (1977).
18. M. Rosenberg, S. Weissman, B. de Crombrugghe, *ibid.* **250**, 4755 (1975).
19. B. Polisky, P. Greene, D. E. Garfin, B. J. McCarthy, H. M. Goodman, H. W. Boyer, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 3310 (1975).
20. D. Kleid, Z. Humayun, A. Jeffrey, M. Ptashne, *ibid.* **73**, 293 (1976).
21. G. G. Brownlee, in *Laboratory Techniques in Biochemistry: Determination of Sequences in RNA*, T. S. Work and E. Work, Eds. (American Elsevier, New York, 1972); B. G. Barrell, in *Procedures in Nucleic Acid Research*, G. Cantoni and D. Davies, Eds. (Harper & Row, New York, 1971), vol. 2, pp. 751–779; G. G. Brownlee and F. Sanger, *Eur. J. Biochem.* **11**, 395 (1969); F. Sanger, G. G. Brownlee, B. G. Barrell, *J. Mol. Biol.* **13**, 373 (1965); R. A. Kramer and M. Rosenberg, *Nucleic Acid Res.* **3**, 2411 (1976).
22. T. Maniatis, A. Jeffrey, D. G. Kleid, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 1184 (1975).
23. F. Sanger, J. E. Donelson, A. R. Coulson, H. Kossel, D. Fischer. *ibid.* **70**, 1209 (1973).
24. A. M. Maxam and W. Gilbert, *ibid.* **74**, 560 (1977).
25. D. M. Kurnit, B. R. Shafit, J. J. Maio, *J. Mol. Biol.* **81**, 273 (1973).
26. H. Swift, *Cold Spring Harbor Symp. Quant. Biol.* **38**, 963 (1973).
27. D. Gillespie, *Science* **196**, 889 (1977). *Note added in proof*: Recent data [D. Dinger and D. Gillespie, unpublished data] indicate the presence of homologous sequences in DNA from baboon (*Theropithecus gelada*).
28. H. E. Hopps, B. C. Berhneim, A. Nisalak, J. H. Tijio, J. E. Smadel, *J. Immunol.* **91**, 416 (1963).
29. S. Sato, C. A. Hutchison III, J. I. Harris, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 542 (1977).
30. R. J. Roberts, J. B. Breitmeger, N. F. Tabachnik, *J. Mol. Biol.* **91**, 121 (1975).
31. K. Heininger, W. Horz, H. G. Zachau, *Gene* **1**, 291 (1977).
32. R. Dhar, K. N. Subramanian, J. Pan, S. M. Weissman, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 827 (1977).
33. Y. Aloni, E. Winocour, L. Sachs, J. Torten, *J. Mol. Biol.* **44**, 333 (1969).
34. Gift of Dr. R. diLauro.

35. We are grateful to M. Nasoff and C. Brady for technical assistance, to Dr. L. Rosenthal for preparation of BSC-1 DNA, to R. Steinberg for help with illustrations, and to Dr. D. Singer for many helpful discussions. Dr. M. Ptashne and his colleagues kindly supplied us with endo R·Hph. We also thank Dr. J. Maio for making results available to us prior to publication. Address correspondence and reprint requests to M.S. or M.R.