Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 6652–6656, November 1984
Biochemistry

A viable simian virus 40 variant that carries a newly generated sequence reiteration in place of the normal duplicated enhancer element

(72-base-pair repeat/enhancer consensus sequence/transcriptional control)

CANDACE SWIMMER* AND THOMAS SHENK†

Department of Microbiology, Health Sciences Center, State University of New York, Stony Brook, NY 11794

Communicated by M. J. Osborn, July 11, 1984

ABSTRACT **A segment comprising the transcriptional enhancer elements was deleted from a recombinant plasmid carrying the simian virus 40 genome. The mutated viral chromosome was excised from the plasmid and propagated through several cycles of growth in monkey kidney cells. A variant was obtained that carried reiterations of sequences that span both sides of the deleted enhancer region. The mutant virus, dup1495, displays a lag in its growth kinetics as compared to its parent, but it ultimately generates wild-type yields. The mutant virus expresses early mRNAs at near-normal levels, and the reiterated sequence functioned in *cis* to enhance transformation of mouse cells by the herpesvirus thymidine kinase gene. Thus dup1495 reiterated segments encode enhancer activity even though their primary sequence is radically different from that of the normal simian virus 40 enhancer.**

The RNA polymerase II transcriptional control region consists of a complex variety of functional domains. One of these domains, which has been identified in some but not all control regions, is termed an enhancer or activator element. These sequences potentiate transcription by an as yet unknown mechanism. One of their most striking characteristics is that they function irrespective of orientation and at the 5′ or 3′ end of a transcription unit. Enhancer elements have been identified in a number of viral and cellular transcription units (reviewed in refs. 1 and 2). The prototype enhancer is the 72-base-pair (bp) tandem repeat of simian virus 40 (SV40) DNA. This element is located >100 bp upstream from the cap site for early viral genes, and it is required in *cis* for efficient transcription of the gene (3–7).

No extensive sequence homology is evident among the various enhancers that have been identified. However, several short consensus sequences are present in multiple enhancers. One such "core" element is the sequence G-G-T-G-T-G-G-$\frac{A\,A\,A}{T\,T\,T}$-G (8). A point mutation in this sequence interferes with SV40 early gene expression, suggesting that it constitutes an essential component of the SV40 enhancer (9). This core sequence overlaps one of a pair of reiterated sequence elements separated by a short spacer, which have been shown by mutational analysis to be essential for enhancer function in bovine papilloma virus (10, 11). The reiteration is also evident in each of the SV40 72-bp repeats: G-C-T-G-T-G-G-A - (15 bp) - G-G-T-G-T-G-G-A (10). Some, but not all, enhancer elements also contain multiple potential Z-DNA-forming regions. Three such segments have been identified in the SV40 enhancer domain using Z-DNA-specific antibodies, and it has been postulated that the left-handed conformation may play a critical role in the function of at least certain enhancers (12).

Here we report the isolation and characterization of a viable SV40 variant, dup1495, which lacks the majority of the SV40 72-bp repeat sequence and carries in its place a reiteration of sequences that normally lie to either side of the enhancer region. The reiterated sequence displays enhancer activity both as a component of the SV40 chromosome and in heterologous assays.

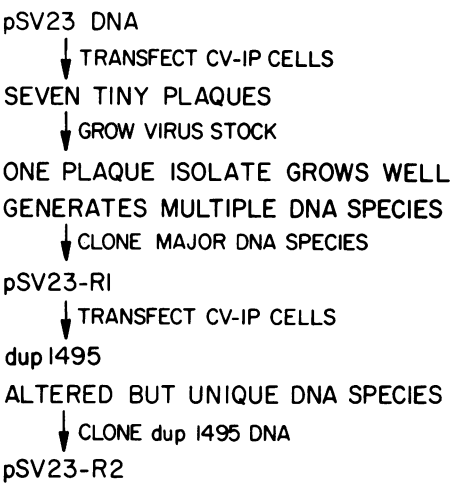

FIG. 1. Flow diagram depicting the series of manipulations that gave rise to dup1495.

MATERIALS AND METHODS

Plasmids, Viruses, and Cells. The recombinant plasmid pSV-WT2 contains the entire SV40 chromosome inserted at its unique *Bam*HI cleavage site into the unique *Bam*HI cleavage site of pBR322. pSV23 is a derivative of pSV-WT2, which lacks the SV40 segment between *Sph* I and *Kpn* I cleavage sites (sequence position 129–298, see Fig. 2). It carries an 8-bp *Xho* I linker segment at the site of the deletion. pSV23-R1 and R2 contain variant SV40 chromosomes (inserted into pBR322 via unique *Bam*HI sites) derived by passaging the SV40 genome excised from pSV23 through CV-1P cells (see Fig. 1). SV40 sequence rearrangements in these plasmids (see Fig. 2) were elucidated by nucleotide sequence analysis (13) using a 15-nucleotide primer DNA (sequence positions 365–379) synthesized by a modification of the phosphite coupling method (14). pHSV-106 (15) contains the herpes simplex type 1 thymidine kinase gene on a 1.3-kilobase *Bam*HI fragment.

The wild-type SV40 (wt830) is a plaque-purified derivative of the SVS strain (16). The duplication variant dup1495 was generated by serial propagation of the SV40 sequences from

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.

Abbreviations: SV40, simian virus 40; bp, base pair(s).

*Present address: Department of Cellular and Developmental Biology, Harvard University, Cambridge, MA 02138.
†To whom reprint requests should be addressed at: Department of Molecular Biology, Princeton University, Princeton, NJ 08544.

pSV23 through CV-1P cells. CV-1P (17) and mouse LTK$^-$ (18) cells were propagated in medium supplemented with 10% calf serum.

Transfection of CV-1P cells with infectious SV40 DNA used DEAE dextran (17), and transfection of murine LTK$^-$ cells with plasmid DNAs was by calcium phosphate precipitation (18).

RNA Preparation and Analysis. To prepare RNAs, CV-1P cells were infected at a multiplicity of 10 plaque-forming units per cell and harvested at various times after infection. Cytoplasmic, poly(A)$^+$ RNA was prepared (19) and subjected to RNA blot analysis (20) using probe DNAs specific for either the early or late SV40 transcription unit.

RESULTS

Isolation of dup1495. This virus was isolated during the course of an experiment designed to isolate cellular sequences with enhancer activity. No cellular DNA segments were identified, but a viable SV40 variant, dup1495, grew out of a control experiment in which cells had been infected with an "enhancerless" SV40 chromosome.

The series of manipulations leading to isolation of dup1495 are diagrammed in Fig. 1. A segment encoding enhancer function was deleted from a recombinant plasmid carrying the entire SV40 chromosome. The altered plasmid is termed pSV23. It lacks the SV40 segment between sequence positions 129 and 298 (Fig. 2, shaded region). SV40 sequences were released from pSV23 by cleavage with a restriction endonuclease and used to transfect CV-1P cells. After ≈3 weeks, several tiny plaques appeared, and one generated a virus stock. The SV40 DNA produced upon infection of CV-1P cells with this virus stock was heterogeneous, and the major DNA species was cloned into pBR322 and termed pSV23-R1. The SV40 sequences from pSV23-R1 were used,

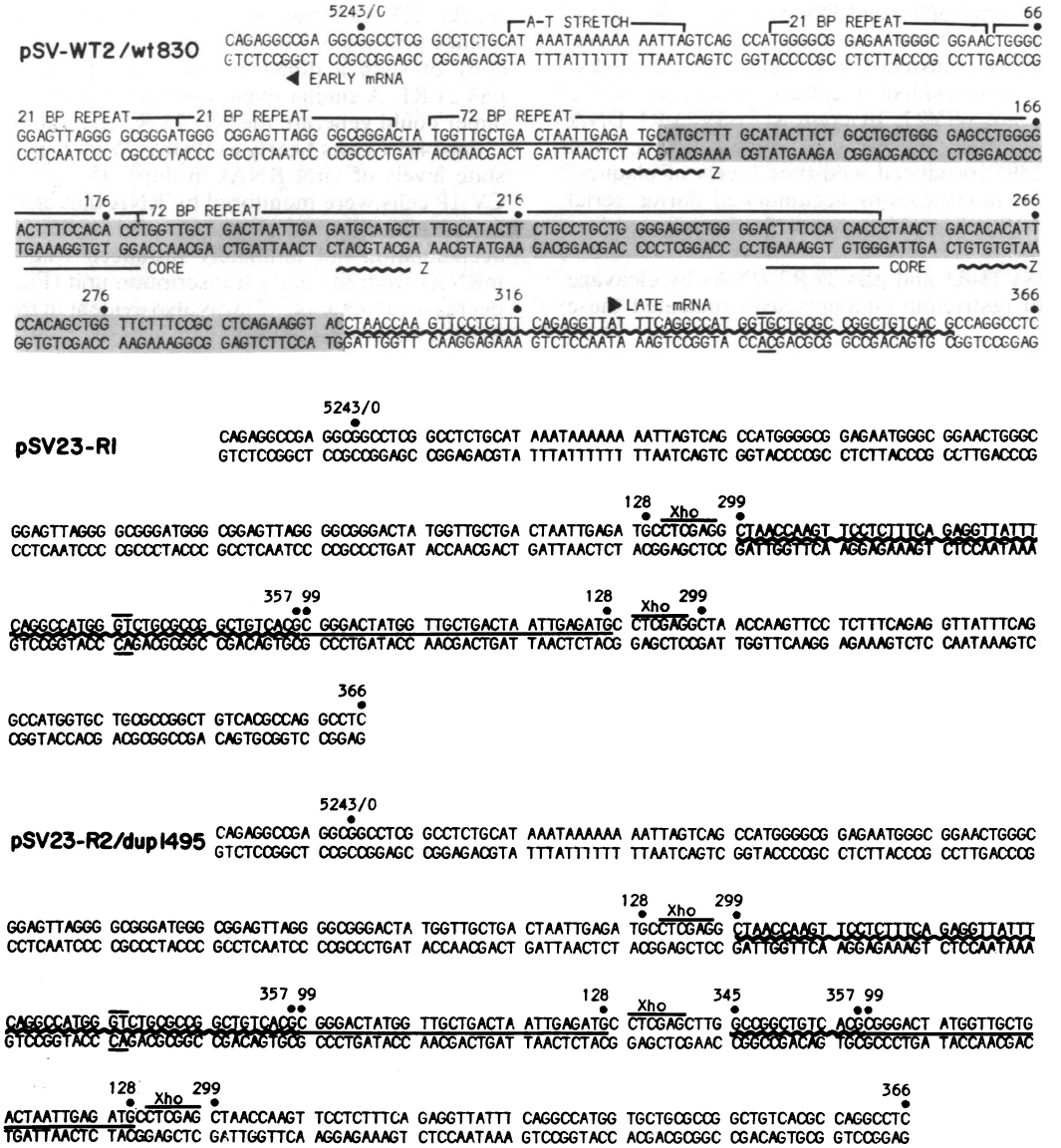

FIG. 2. Sequence of the SV40 genome that encodes the early and late transcriptional control regions. The sequence is numbered according to Tooze (21). pSV-WT2/wt830 depicts the wild-type sequence. Major features of the primary sequence are indicated above the double strands (A-T stretch, 21-bp repeats, 72-bp repeats), and sequences that have been proposed as critical elements of the enhancer region are underscored (core and Z). Arrowheads indicate major cap sites for early and late mRNAs. The region deleted to generate the enhancerless variant (pSV23) is shaded. Straight and wavy lines between DNA strands designate sequence to the early and late sides, respectively, of the pSV23 deletion, which are reiterated in the variants that have recovered enhancer activity. pSV23-R1 and pSV23-R2/dup1495 depict the sequences of naturally arising viable variants derived from the defective enhancerless mutant. pSV23-R1 has regained modest enhancer activity and pSV23-R2 exhibits a substantial increase in activity. The *Xho* I cleavage site designates an 8-bp linker DNA inserted at the site of the pSV23 deletion. Straight and wavy lines between the base pairs designate reiterated sequences.

Table 1. Plaque-forming efficiencies of viral DNAs

DNA	Plaque-forming units per μg of DNA	Plaque diameter, mm
pSV-WT2	2.9×10^5	3
pSV23	$<1.0 \times 10^1$	<0.5
pSV23-R1	7.5×10^3	0.5–1
pSV23-R2	3.6×10^5	2

CV-1P cells were transfected with *Bam*HI-cleaved plasmid DNAs. The cleavage released the viral chromosome intact from plasmid sequences. Plaques were counted on day 14. pSV23 gave rise to tiny plaques by day 30 ($\approx 10^3$ plaque-forming units per μg).

in turn, to transfect CV-1P cells, and they gave rise to an altered but stable SV40 variant, dup1495. Virus-specific DNA within dup1495-infected cells appears to be homogeneous. This DNA species was also cloned into pBR322 and termed pSV23-R2.

The growth potential of variant DNAs was monitored by plaque assay (Table 1). Viral sequences were excised from plasmids and used to transfect CV-1P cells. After 2 weeks, very few plaques were evident in cultures that received the enhancerless variant pSV23. In contrast, pSV23-R1 DNA displayed a substantially increased plaquing efficiency and pSV23-R2 (dup1495) produced wild-type levels of plaques. We conclude that modifications accumulated during serial passaging in CV-1P cells, which overcame the enhancerless defect.

Analysis of pSV23-R1 and pSV23-R2 DNAs by cleavage with a variety of restriction endonucleases indicated these DNAs carried inserts in the vicinity of the enhancer region deletion present in pSV23. Accordingly, the deletion proximal region was sequenced in pSV23-R1 and pSV23-R2 (Fig. 2). Both variants contained duplications of sequences normally located to either side of the deletion in pSV23. pSV23-R1 carries a single reiteration of sequences to either side of the original deletion. pSV23-R2 (dup1495) carries a double reiteration and, therefore, contains three copies of each reiterated element.

The poor growth of the enhancerless variant could be rescued by inserting the reiteration into pSV23 at the site of the enhancer deletion. Thus, it is clear that the reiterations are responsible for the enhanced growth properties of dup1495.

The mechanism by which the reiterations arose is uncertain. Perhaps they were generated during resolution of dimer to monomer SV40 chromosomes. Tandem head-to-tail dimers of the SV40 genome are present in substantial quantities within infected cells ($\approx$1% of the SV40 DNA population; see ref. 22). An illegitimate recombination event between sequence position 357 in one copy and position 99 in the other copy of such a dimer would generate the reiteration in pSV23-R1. A similar event during resolution of a pSV23-R1 dimer could generate the pSV23-R2 reiteration.

dup1495 Grows Nearly as Well as Wild-Type SV40. Steady-state levels of viral RNAs in dup1495- and wt830-infected CV-1P cells were monitored by RNA blot analysis. Cells infected with dup1495 showed a modest lag ($\approx$12 hr) in the accumulation but ultimately produced wild-type levels of mRNAs from the early transcription unit (Fig. 3*A*). A slight decrease (a factor of <2) was also evident in the levels of late mRNAs synthesized in mutant-infected cells (Fig. 3*B*). Production of the late polypeptide VP-1 was decreased by a factor of $\approx$3 in cells infected with dup1495 as compared to the wild type (data not shown). The rate of accumulation of mutant and wild-type DNA subsequent to infection was also examined (Fig. 4). Again, little difference between viruses was observed. Finally, the growth kinetics and yield of dup1495 and wt830 were compared (Fig. 5). Virus production from dup1495-infected cultures initially lagged well be-

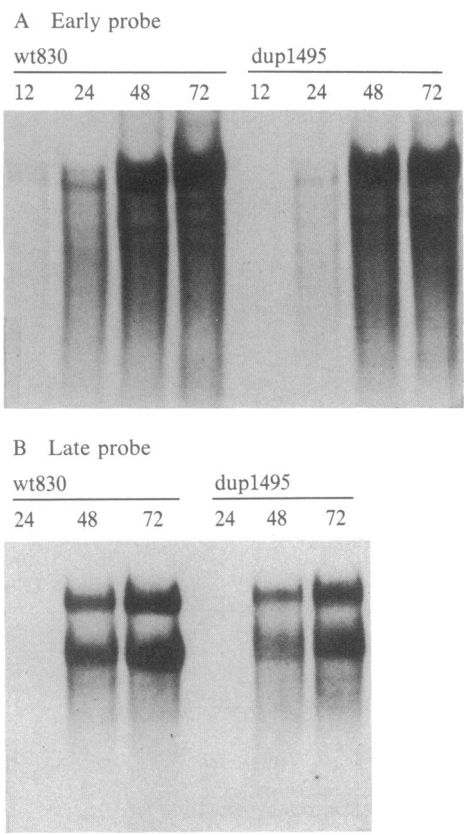

FIG. 3. RNA blot analysis of viral RNAs produced in wild-type and dup1495-infected CV-1P cells. Cytoplasmic polyadenylylated RNAs were prepared at various times (hr) after infection at a multiplicity of 10 plaque-forming units per cell and were analyzed using a probe corresponding to either the early or late viral transcription unit. An identical amount of each polyadenylylated RNA preparation was analyzed, using excess probe DNA to permit accurate quantitation.

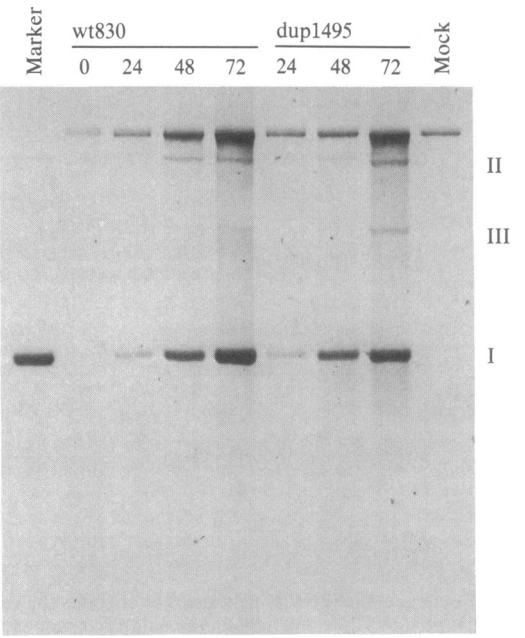

FIG. 4. Electrophoretic analysis of viral DNAs produced in wild-type and dup1495-infected CV-1P cells. Infections were at a multiplicity of 10 plaque-forming units per cell; viral DNA was prepared by Hirt extraction (23) at the indicated times (hr) and was analyzed by electrophoresis in an agarose gel. Forms I, II, and III SV40 species are labeled.

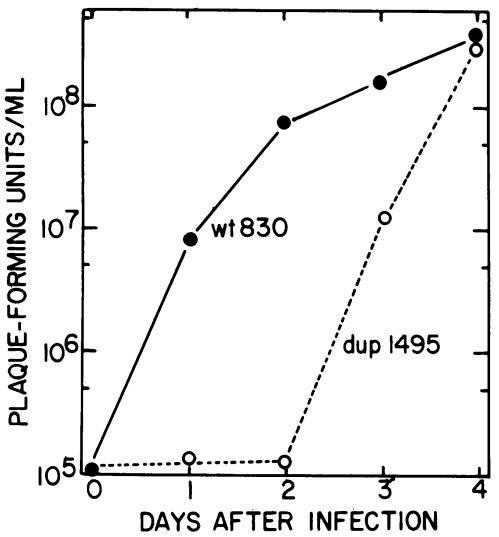

FIG. 5. Growth kinetics of wild-type and dup1495 viruses in CV-1P cells. Infections were at a multiplicity of 3 plaque-forming units per cell; cultures were harvested at the indicated times, and the virus titer was measured by plaque assay on CV-1P cells.

hind that of the wild type. The final yield of the two viruses, however, was identical. The reason for the substantial lag in virus production is not clear, given the modest delay in accumulation of dup1495-encoded products within infected cells. Perhaps a virus-encoded product we have not measured is seriously perturbed (agno protein?). Alternatively, the series of minor alterations might have an additive effect on the overall dynamics of virus production.

We conclude that dup1495-infected cells accumulate near wild-type levels of viral mRNAs, viral DNA, and progeny virions. The dup1495 sequence reiteration functions effectively to replace the normal SV40 enhancer elements.

The dup1495 Sequence Reiteration Displays Enhancer Activity. A variety of enhancer elements have been shown to increase the frequency of tranformation by plasmid DNAs carrying marker genes. The reiterated sequences in pSV23-R1 and pSV23-R2 (dup1495) were tested for this effect using the herpesvirus thymidine kinase gene. Four SV40-derived fragments were inserted into a plasmid carrying this gene in various orientations either 5′ or 3′ to the coding region. The fragments had the same end points and spanned the enhancer domain. Interestingly, insertion of a segment derived from the enhancerless variant reduced the frequency of transformation by the thymidine kinase-containing plasmid (Table 2, compare no insert to pSV23). The wild-type region substantially increased the frequency relative to the enhancerless insert, and segments carrying the reiterated sequences gave intermediate levels of enhancement (Table 2, pSV-WT2, pSV23-R1, and pSV23-R2). The duplication in pSV23-R1 displays low enhancer activity. As expected from the analysis of plaque formation in Table 1, the pSV23-R2 (dup1495)-derived segment was a more effective enhancer element than that from pSV23-R1. The pSV23-R2 (dup1495) reiteration appears about one-half as potent as the wild-type element in this assay.

Table 2. Reiterated sequences from pSV23-R2 (dup1495) enhance transformation by the herpesvirus thymidine kinase gene

Insert	Location	Colonies per μg of DNA
None		70
pSV-WT2	5′A	610
	5′B	390
	3′A	1020
	3′B	390
pSV23	5′A	16
pSV23-R1	5′A	90
pSV23-R2	5′A	390
	5′B	280
	3′A	490
	3′B	250

Parental plasmids carrying the herpesvirus thymidine kinase (*TK*) gene were obtained from M. Ostapchuk and A. Levine. They were derived from pHSV106, which contains the *TK* gene on a 1.3-kilobase *Bam*HI fragment and each carries a unique *Xho* I linker segment inserted at a *Pvu* II site either 5′ or 3′ to the *TK* coding region. SV40 segments tested for enhancer activity were *Hin*dIII fragments spanning from sequence position 5171 to 1046. SV40 segments were inserted into the unique *Xho* I cleavage sites of the pHSV106 derivatives. Inserts were either 5′ or 3′ (−200 or +1600 relative to the TK cap site) in either the same (A) or opposite (B) orientation as the *TK* gene. Murine LTK$^-$ cells were transfected with 500 ng of plasmid DNA and TK$^+$ colonies were quantified 14 days later.

DISCUSSION

The deletion in pSV23 removes sequences suspected to be key elements of the SV40 enhancer domain. pSV23 lacks both copies of the enhancer core G-G-T-G-T-G-G-$\frac{A\,A\,A}{T\,T\,T}$-G described by Weiher *et al.* (9), the related reiteration described by Weiher and Botchan (11), as well as all three stretches of alternating purine/pyrimidine sequence shown to display Z conformation by Nordheim and Rich (12). The SV40 sequences in pSV23 are not viable (Table 1), and the deleted enhancer domain does not function in a heterologous enhancer assay (Table 2). dup1495 (pSV23-R2) is a viable variant, which arose subsequent to introduction of the enhancerless SV40 sequences from pSV23 into cultured monkey cells. It has acquired a reiteration of sequences that span both sides of the deletion in pSV23 (Fig. 2). Its reiteration possesses enhancer activity judged both by the restoration of viral gene expression and near wild-type growth properties (Figs. 3–5) and by function of the reiteration in a heterologous enhancement assay (Table 2).

Why does the dup1495 reiteration restore enhancer activity? The answer is not obvious but several possibilities deserve mention. Perhaps the junctions created by the reiterations encode enhancer activity. The junction at position 357/99 creates the sequence C-A-C-G-C-G, the longest run of alternating purine/pyrimidine residues in the reiterated segments. Conceivably, this sequence generates a Z conformation, replacing the 8-bp segments deleted in the variant. However, it is difficult to predict Z-DNA conformation from DNA sequence, and this stretch is shorter than the sequences previously shown to display Z character on the SV40 chromosome (12). It is also possible that sequences with weak enhancer activity lie to either or both sides of the pSV23 deletion. Reiteration of these sequences might generate a more effective enhancer domain, as observed in dup1495. The reiterations do contain homologies to the SV40 enhancer core described by Weiher *et al.* (9) at sequence positions 305–299 (T-G-G-T-T-A-G) and 110–105 (T-G-G-T-A-T). Such homologies are common on the SV40 chromosome, occurring once every 300 bp, and their contribution to dup1495 enhancer activity must remain suspect. However, these homologies are located at 25-bp intervals within the dup1495 repeats, reminiscent of the reiterated homologies described by Weiher and Botchan (11). The dup1495 interval is ≈1 turn of the DNA helix longer than in the case of naturally occurring reiterations (15–16 bp). Perhaps, then, these core homologies have been activated by their relative spatial arrangement. Beyond these homologies, a computer-aided search revealed no striking primary sequence similarities between the dup1495 reiterations and the wild-type element. Homologies might, of course, exist, which we fail to detect

because they are dispersed as small segments between random sequence stretches.

The dup1495 reiteration functions more effectively as an enhancer than does its progenitor duplication in pSV23-R1 (Fig. 5; Table 2). One might expect a clue to the basis of dup1495 enhancer function to lie in the additional iteration (54 bp), which this DNA carries as compared to the duplication in pSV23-R1. The segment reduplicates a core homology (positions 110–105) and the junction, which creates a run of alternating purine/pyrimidine residues. Beyond that, its sequence appears unremarkable.

Conceivably, the reiterations themselves contribute to the dup1495 enhancement effect independently of their primary sequence. So far, the effect of reiterations *per se* on enhancer function has not been resolved. Modest effects on enhancer efficiency of one versus two copies of the 72-bp repeat have been noted in some studies (refs. 3 and 7; W. Herr, personal communication), but not in others (24).

As this manuscript was in preparation, Weber *et al.* (25) reported the isolation of two variants similar to dup1495, which arose subsequent to infection with enhancerless SV40 DNAs. These variants have also acquired duplications, but, in contrast to dup1495, the duplicated sequences were derived exclusively from one side of the deletion (the side facing the late mRNAs: positions 295–379). This duplication has the sequence 299–357 in common with the dup1495 reiteration, and it includes the enhancer core homology at positions 305–299. The variants display no common features at their duplication junctions.

Finally, it is important to emphasize that the reiterations in dup1495 probably arose in response to a variety of functional pressures. The original deletion not only excised enhancer elements, it also removed a region coding for a variety of late mRNA 5′ ends (26, 27) and impinged closely on the major late mRNA 5′ end, and it removes sequences suspected to play a direct role in the late transcriptional control region (28). Thus, the specific nature of the reiterations likely reflects a complex array of functional requirements and not simply reacquisition of enhancer activity.

This work was supported by a grant from the National Cancer Institute (CA28919).

1. Khoury, G. & Gruss, P. (1983) *Cell* **33,** 313–314.
2. Gruss, P. (1984) *DNA* **3,** 1–5.
3. Banerji, J., Rusconi, S. & Schaffner, W. (1981) *Cell* **27,** 299–308.
4. Benoist, C. & Chambon, P. (1981) *Nature (London)* **290,** 304–310.
5. Ghosh, P. K. & Lebowitz, P. (1981) *J. Virol.* **40,** 224–240.
6. Gruss, P., Dhar, R. & Khoury, G. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 943–947.
7. Fromm, M. & Berg, P. (1982) *J. Mol. Appl. Genet.* **1,** 457–481.
8. Laimins, L., Khoury, G., Gorman, C., Howard, B. & Gruss, P. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 6453–6457.
9. Weiher, H., Konig, M. & Gruss, P. (1982) *Science* **219,** 626–631.
10. Lusky, M., Berg, L., Weiher, H. & Botchan, M. (1983) *Mol. Cell, Biol.* **3,** 1108–1122.
11. Weiher, H. & Botchan, M. R. (1984) *Nucleic Acids Res.* **12,** 2901–2916.
12. Nordheim, A. & Rich, A. (1983) *Nature (London)* **303,** 674–679.
13. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
14. Matteucci, M. D. & Caruthers, M. H. (1981) *J. Am. Chem. Soc.* **103,** 3185–3190.
15. McKnight, S. L. & Gavis, E. R. (1980) *Nucleic Acids Res.* **8,** 5931–5948.
16. Takemoto, K. K., Kirchstein, R. L. & Habel, K. (1966) *J. Bacteriol.* **92,** 990–994.
17. Mertz, J. E. & Berg, P. (1974) *Virology* **62,** 112–124.
18. Wigler, M., Silverstein, S., Lee, L.-S., Pellicer, A., Cheng, Y.-C. & Axel, R. (1977) *Cell* **11,** 223–232.
19. McGrogan, M., Spector, D. J., Goldenberg, C., Halbert, D. N. & Raskas, H. J. (1979) *Nucleic Acids Res.* **6,** 593–608.
20. Hearing, P. & Shenk, T. (1983) *Cell* **33,** 695–703.
21. Tooze, J., ed. (1981) *DNA Tumor Viruses,* Molecular Biology of Tumor Viruses (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), 2nd Ed.
22. Goff, S. P. & Berg, P. (1977) *J. Virol.* **24,** 295–302.
23. Hirt, B. (1967) *J. Mol. Biol.* **26,** 365–369.
24. Gorman, C. M., Moffat, L. F. & Howard, B. H. (1982) *Mol. Cell. Biol.* **2,** 1044–1051.
25. Weber, F., deVilliers, J. & Schaffner, W. (1984) *Cell* **36,** 983–992.
26. Ghosh, P. K., Reddy, V. B., Swinscoe, J., Choudary, P. V., Lebowitz, P. & Weissman, S. M. (1978) *J. Biol. Chem.* **253,** 3643–3647.
27. Contreras, R., Gheysen, D., Knowland, J., Van de Voorde, A. & Fiers, W. (1982) *Nature (London)* **300,** 500–505.
28. Hartzell, S. W., Byrne, B. J. & Subramanian, K. N. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 23–27.