JOURNAL OF VIROLOGY, July 1981, p. 185-195
0022-538X/81/070185-11$02.00/0

Vol. 39, No. 1

Integrated Polyoma Genomes in Inducible Permissive Transformed Cells

PIERRE CHARTRAND,* NADINE GUSEW-CHARTRAND, AND PIERRE BOURGAUX

Département de Microbiologie, Faculté de Médecine, Centre Hospitalier Universitaire, Université de Sherbrooke, Sherbrooke, Quebec, Canada J1H 5N4

Received 5 January 1981/Accepted 7 April 1981

Using the approach described by Botchan, Topp, and Sambrook (Cell 9:269-287, 1976), we analyzed the organization of the integrated viral sequences in five clonal isolates from the same permissive, inducible cell line (Cyp line) transformed by the *ts*P155 mutant of polyoma virus. In all five clones, viral sequences were found that could be assigned to a common integration site, as they were joined to the cellular DNA in the same fashion in every instance. However, the sequences comprised between these points differed markedly from clone to clone, as if cell propagation had been accompanied by amplification or recombination or both within the viral insertion. When the clones were compared, no correlation could be found between the abundance, or the organization, of the integrated viral sequences and the amount, or the nature, of the free viral DNA molecules produced during induction. Altogether, our findings suggest that specific events, occurring during either the excision or the subsequent replication of the integrated viral sequences, are responsible for the predominant production of nondefective viral DNA molecules by permissive transformed cells, such as Cyp cells.

In polyoma virus (Py)-transformed rat cells, full or partial head-to-tail tandems of the viral genome are covalently linked to the high-molecular-weight cellular DNA (1, 3, 12). During cell propagation, these integrated viral sequences can undergo, at high frequency, two types of alterations, both of which appear to require a functional viral *a* gene product: (i) the loss of the full copy tandems (2), or (ii) the selective amplification, as head-to-tail tandems of internal subgenomic sequences containing the origin of replication (8) or both.

Another feature of Py-transformed rat cells is the occasional production of free viral DNA that results from a periodic and spontaneous induction of the resident viral genome in a very small proportion (<1%) of the cells (14, 18). The free viral DNA molecules thus produced, the structure of which faithfully reflects the integrated viral genomic or subgenomic tandems, are probably the result of homologous recombination within the tandems (10). Their production is dependent on the presence of a functional viral *a* gene product, i.e., the large T polypeptide, that could either amplify the viral sequences before excision (4, 6) or, alternatively, act as a recombination-promoting agent (10).

Recently, one of us isolated a mouse cell line (Cyp line) transformed at the restrictive temperature of 39°C by *ts*P155, an early temperature-sensitive mutant of Py (7). When compared with nonpermissive transformed cells, Cyp cells offer the advantage of producing virus or viral DNA in large amounts, after either superinfection (9) or transfer to the nonrestrictive temperature of 33°C. The Cyp line, which was derived from a single focus of transformed cells, was recloned at the 11th passage to give rise to a number of clones, among which were C10, C11, C12, and C13. Cultures from these clones produce no virus and little or no free viral DNA at 39°C, but again, can all be induced to synthesize free, nondefective viral DNA by transfer to 33°C. The amount of viral DNA thus synthesized varies from clone to clone, but is reproducible for a given clone: it is low for C10 and C11 and high for C12 and C13 (16). This variation may not reflect one in the frequency of induction, since for both C10 and C13, for instance, a majority of the cells are able to produce infectious virus after temperature shiftdown (7). The highly inducible C12 clone has been recloned in agar, yielding numerous subclones with a phenotype similar to that of their parent, except in the case of subclone C12/a1. When transferred to 33°C, cells from this subclone produce predominantly cyclic DNA molecules (RmI), comprising the entire viral genome linked to about 1×10^6 daltons of mouse DNA. Genome-size P155 DNA is also produced by C12/a1 cells at 33°C, in a 1:20 ratio compared with RmI (9, 16). These two forms are not segregated by further

recloning of C12/a1 cells (B. Sylla, unpublished data). Superinfection of Cyp cells at 39°C with wild-type Py of several origins results in the replication of the superinfecting, but not of the resident, Py genome (9). Since both *ts*P155 and the virus rescued from Cyp cells can be complemented by a late mutant or Py wild type, this observation suggests that the resident viral genome carried by Cyp cells propagated at 39°C is not as readily available for autonomous replication as the superinfecting genome, possibly because it is integrated.

In the present report, we show that cells from clones C10, C11, C12, C14, and C12/a1, all derived from the same transformed focus, contain integrated viral sequences, some of which appear to be linked to the cellular DNA in the same fashion in all the clones. Yet, the internal arrangement of these viral sequences is unique to each clone. Furthermore, the structure of the integrated sequences does not provide straightforward explanations either for the different degrees of inducibility of the clones or for the nature of the free viral DNA molecules that are produced after induction. In contrast to what has been observed for Py-transformed rat cells, the free viral DNA molecules produced in cells from the various Cyp clones do not faithfully reflect the integrated species. These observations imply the existence, in these inducible permissive cells, of a mechanism, acting either during or after excision, which favors the production of nondefective free viral genomes.

MATERIALS AND METHODS

Cells. The origin and cultivation of the Cyp cell line and of its clones C10, C11, C12, C13, and C12/a1 have been described previously (7, 16).

Preparation and endonuclease digestion of cellular DNA. High-molecular-weight cellular DNA was prepared by the procedure of Gross-Bellard et al. (11). Only those preparations of cellular DNA that contained little or no free viral DNA were used in this study. The analyses presented here for a given clone were all done with the same preparation of DNA, although similar results were obtained with other preparations. For each assay, 10 μg of cell DNA was digested at 37°C for 3 h, with a fivefold excess of restriction enzyme and the buffer recommended by the supplier (New England Biolabs or BRL). To monitor whether digestion had proceeded to completion, 10^{-2} μg of lambda DNA was added to each reaction mixture; after annealing with the Py DNA probe, the blots were reannealed with a lambda DNA probe.

Electrophoresis and blotting. Restricted DNA was electrophoresed for about 16 h through horizontal agarose slab gels (20 by 15 cm) in E buffer (36 mM Tris [pH 7.8]-36 mM NaH_2PO_4-1 mM EDTA). Depending upon the size of the fragments to be resolved, the agarose concentration was 0.5, 0.7, or 1.2%. Molecular weights were determined from those of the lambda DNA fragments present in each lane (see preceding paragraph) and of the *Hin*dIII fragments of lambda DNA run in a control lane.

DNA fragments were transferred from the gels onto nitrocellulose filter paper (BA85; Schleicher & Schuell Co.) as described by Southern (15).

Nick translation and annealing. Radioactive probes were prepared by the procedure of Maniatis et al. (13). The template DNA was the complete Py genome. It had been cloned in the *Bam*HI site of plasmid pBR322 (M. Bastin, unpublished data) and recovered as a linear molecule after cleaving the recombinant plasmid with *Bam*HI. Each reaction mixture contained 1 μg of Py DNA, 100 μCi each of α-[^{32}P]dATP and α-[^{32}P]dCTP (Amersham Corp.; ≃3,000 Ci/mmol), and 0.25 mM of both dGTP and dTTP (Sigma Chemical Co.) in 20 mM Tris-hydrochloride buffer (pH 7.5)-10 mM $MgCl_2$-10 mM 2-mercaptoethanol-50 μg of bovine serum albumin per ml (Sigma Chemical Co.). The reaction was initiated by adding 5 U of *Escherichia coli* DNA polymerase I (New England Biolabs) at 18°C and stopped after 60 min by adding 20 μl of 0.2 M EDTA. The labeled DNA was separated from radioactive triphosphate nucleosides by chromatography on a Sephadex G-75 column equilibrated with TE buffer. The specific activity of the probe was 1×10^8 cpm/μg or higher.

The annealing procedure was that of van der Ploeg and Flavell (17) with minor modifications. Before annealing, blots were incubated at 68°C for 2 h in prehybridization buffer (10× Denhardt in 3× SSC) and then for a further 2 h in hybridization buffer (prehybridization buffer plus 0.1% sodium dodecyl sulfate, 10% dextran sulfate, and 50 μg of salmon sperm DNA per ml). Thereafter, each blot was inserted in a plastic bag with 3 ml of hybridization buffer containing 5×10^6 cpm of the heat-denatured probe. The bags were sealed, and annealing was allowed to proceed for 60 h at 68°C in a shaking water bath. The filters were then removed from the bags and washed extensively with prehybridization buffer, 1× SSC, 0.3× SSC, and finally, 0.1× SSC. After being drained, they were autoradiographed at −80°C with Kodak XR-1 film and Cronex Lightning-Plus intensifying screens.

RESULTS

Number of distinct integrated viral structures in each clone (no-cut RE). To assess the number of distinct insertions of viral DNA, we digested Cyp cell DNA with three restriction endonucleases (RE) which do not cut Py DNA and are thus referred to as "no-cut" RE (*Bgl*II, *Hpa*I, and *Xho*I). *Bgl*II was the most useful, yielding small enough DNA fragments to be well separated by electrophoresis through agarose gels (Fig. 1, lanes 1 to 5). In the DNA blots of four of the clones analyzed, that is to say, all except C12/a1, two or more major species of high molecular weight annealed with the radioactive viral probe. The existence of more than one DNA band containing Py sequences in a given clone could be due to the presence of

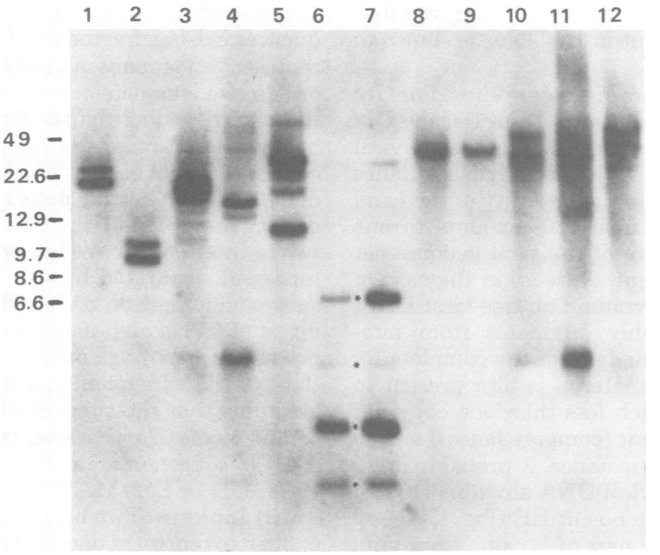

FIG. 1. *Analysis of the integrated Py DNA sequences of clones C10 (lanes 1 and 8), C11 (lanes 2 and 9), C12 (lanes 3 and 10), C12/a1 (lanes 4 and 11), and C13 (lanes 5 and 12) after cleavage with BglII (lanes 1 to 5) or HpaI (lanes 8 to 12). Ten micrograms of cleaved cellular DNA was loaded in each lane and electrophoresed as described in the text. The DNA was then blotted onto a nitrocellulose filter and annealed with a probe of high specific activity, comprising the whole of Py DNA. Lanes 6 and 7 contain undigested RmI DNA and undigested genome-size Py DNA, in the amount of 10^{-5} μg (lane 6) or 5×10^{-5} μg (lane 7) each. Starting with the lowest one, the four bands (*) in each of these two lanes are closed circular Py, closed circular RmI, open circular Py, and open circular RmI. The size markers are given in kilobase pairs (see the text).*

integrated viral genomes at two or more distinct sites in the cellular DNA. Alternatively, it could be the result of either different cells having different amounts of viral material at the same site or the whole of the unique viral insertion having been amplified, together with some of the adjacent cellular sequences. Another observation was that these major bands migrated differently in each clone. Considering that all clones had a common ancestor, it seemed likely that in each clone, at least one of the viral inserts was at a common site. Since we did not detect a common viral band with the no-cut RE, this would imply that the amount of viral DNA at that site differed from clone to clone. As will be seen later, all clones actually have viral sequences at a common integration site that account for the majority of the major Py hybridizing bands seen in the *Bgl*II digest (Fig. 1, lanes 1 to 5). At that common site, the amount of viral DNA is increased by about 5.3 kilobase pairs (kbp) in C12/a1, about 5.4 and 8.1 kbp in C12, and about 9.3 and 13.3 kbp in C10, over that present in C11 (see below). These calculated differences in size agree very well with the differences in mobility seen in the *Bgl*II digest. C13 has a complex integration pattern that renders the analysis difficult.

There were no free viral DNA molecules that we could detect in C10 or C11, and almost none in C13. C12 had less than one free copy of the viral genome per cell genome equivalent, and C12/a1 had about five copies or more. Thus, except in the case of C12/a1, free viral genomes were not present in sufficient numbers to prevent us from using, in our analysis of the integrated sequences, RE that cut Py DNA (see below). For C12, C12/a1, and C13, a number of low-intensity high-molecular-weight bands were also noted. Presumably, these bands represented additional integrated viral sequences rather than oligomers of free viral DNA, since they were not found at the same position in digests from other no-cut RE, like *Hpa*I (lanes 10 to 12). The fragments containing viral sequences generated by both *Hpa*I (Fig. 1, lanes 8 to 12) and *Xho*I (data not shown) were of very high molecular weight and could not be properly resolved by electrophoresis. Although the use of these two enzymes added little to the information already obtained with *Bgl*II, it nevertheless assisted us in selecting, for further analysis, DNA preparations that were virtually devoid of free viral DNA molecules.

Tandems and linkers (one-cut RE). An autoradiogram obtained after digestion of the

cellular DNA with *Eco*RI and *Bam*HI, RE that introduce only one cut in Py DNA, is shown in Fig. 2.

(i) Tandems of complete genomes. For C10 (lanes 1 and 8), C12/a1 (lanes 4 and 11), and C13 (lanes 5 and 12), we notice the presence of viral material comigrating with linear Py DNA (lanes 6 and 7). Judging from the intensity of the band in all six lanes, this material amounts to one copy (lane 7) or more of the viral genome per cell genome equivalent. At least in the case of C10 and C13, which contain no free viral DNA, this material probably originates from integrated head-to-tail tandems of the complete Py genome. The same material is also present in C12 cells, but in much less than one copy per cell genome equivalent (compare lanes 3 and 10 with lane 7). In that instance, it probably originates from the free viral DNA already detected in the digest from the no-cut RE (Fig. 1, lanes 3 and 10). If, however, part of it came from viral genomes integrated in tandems, these would be present only in a minor fraction of the cell population. The C11 digest contains no material migrating like linear Py DNA (lanes 2 and 9), indicating that C11 has no integrated complete copy tandems.

(ii) Tandems of incomplete genomes. In both of the C10 and C13 digests (*Eco*RI and *Bam*HI), we detect viral sequences migrating faster than linear Py DNA. In C10, these sequences exist as fragments of about 4.0 kbp and represent an amount of viral DNA equivalent to one to two complete genomes per cell. For C13, the corresponding figures for the most prominent of these bands are 4.7 kbp and three to four copies. Material of the same kind is also present in C12, but only readily detectable in the *Bam*HI digest (lane 10); here the figures are 2.7 kbp and two to three copies. We believe that these bands represent integrated head-to-tail tandems of a subgenomic portion of Py DNA. In the *Eco*RI digest of C12, a high-molecular-weight fragment producing a strong, intense band is observed (lane 3). This fragment can be accounted for by assuming that the subgenomic tandems in C12 comprise the *Bam*HI site, but not the *Eco*RI site. No subgenomic tandems are observable in either C11 or C12/a1.

(iii) Linkers. Two bands, neither one intense enough to represent one viral genome equivalent per cell, are found at positions that vary with the RE used, but not with the clone analyzed. These bands, designated in the figure by open circles, are likely to originate from the virus-host linkers (see below). Because they are common to all five clones, these would have a common integration site, where the same viral sequences would be joined to the same cellular sequences. This is not surprising since all five clones derive

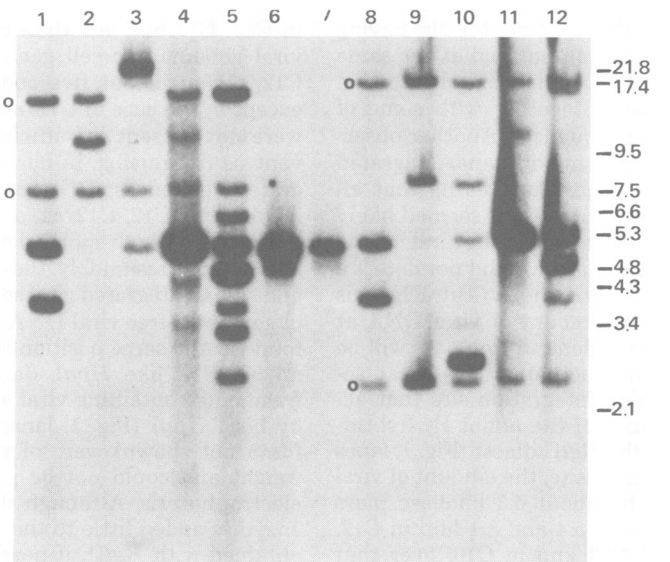

FIG. 2. *Analysis of the integrated Py DNA sequences of clones C10 (lanes 1 and 8), C11 (lanes 2 and 9), C12 (lanes 3 and 10), C12/a1 (lanes 4 and 11), and C13 (lanes 5 and 12) after cleavage with EcoRI (lanes 1 to 5) or BamHI (lanes 8 to 12). Lanes 6 and 7 contain, respectively, 5×10^{-5} μg and 10^{-5} μg of EcoRI-digested Py DNA. The size markers are in kilobase pairs. The circles designate the position of the two bands believed to be the common virus-host linkers (see the text).*

from a single focus of transformation (see above). Hereafter in the text, we will refer to this site and to these linkers as the common site and linkers. In the *Eco*RI digest of C12, the upper common linker is, however, missing (Fig. 2, lane 3). Presumably, this is due to the fact that this linker is attached to the subgenomic tandems found higher up in the gel (see above). If this interpretation is correct, it would imply that, in C12, the subgenomic tandems are located at the common integration site.

Other fragments containing viral sequences. In the case of C11, one band of about 11 kbp is visible in the *Eco*RI digest (lane 2), whereas two bands are present in the *Bam*HI digest (lane 9)—one of about 7.8 kbp and the other virtually comigrating with the common linker of 2.4 kbp. Taken together, the results shown in Fig. 2 and 3 (see below), as well as others not shown, suggest that these bands originate from viral sequences inserted at a cellular site that is unique to C11. These sequences, however, represent less than a complete Py genome. As to the additional bands seen in digests from clones C12 and C12/a1, they are of relatively low intensity and, thus, possibly originate from a minority in the cell population. No additional bands were seen in either of the digests from C10, nor in any other digest from the same clone (Fig. 3 and data not shown). As C10 seems to have only two linkers, those common to all clones, we, therefore, conclude that all of the integrated viral DNA in C10 is present at the common site. This conclusion implies that both the genomic and subgenomic tandems detected in C10 (see above) are integrated at that site. However, we have to recall that the results obtained with one no-cut RE, *Bgl*II, suggested the existence of two integrated structures, rather than one. As will be shown below, this apparent discrepancy seems to reflect the existence, in clone C10, of two kinds of viral insertions which would differ only by having one more, or one less, copy of the amplified subgenomic sequences of 4.0 kbp. These two kinds of insertions could be present in different cells or, alternatively, could both exist in all cells. This would imply that there are at least two copies of each of the common linkers per genome. As already discussed above, the integration pattern in C13 is too complex to determine the origin of the extra bands seen in Fig. 2.

Viral sequences at the site of integration (two-cut RE). In Fig. 3A, we show the autoradiogram of a blot obtained after digestion of Cyp cell DNA with *Hin*dIII, which cuts Py DNA twice, thereby producing *Hin*dIII fragments A and B, of about 3.0 and 2.3 kbp, respectively. *Hin*dIII fragment A is present in all clones and in only one copy in C11 and C12. In the case of C12, it is only visible after a short exposure because of the presence of an intense radioactive band located immediately below. In contrast, the DNA of C10, C12/a1, and C13 contains multiple copies of *Hin*dIII-A, as was expected from the presence of complete copy tandems (Fig. 2).

Whereas C12/a1 and C13 contain multiple copies of *Hin*dIII-B, C10 seems to have only one, compared with more than one copy of *Hin*dIII-A. Therefore, the genomic-size tandems in this clone probably comprise less than two full copies of the Py genome. C12 has much less than one copy of *Hin*dIII-B per cell genome equivalent, possibly originating from free viral DNA molecules (see no-cut RE). Finally, C11 has no copy of *Hin*dIII-B.

Two bands (marked by open circles) are again common to all clones. They do not comigrate with the similarly cut Py DNA marker and represent less than one viral copy per cell genome equivalent. These presumably are the common linkers, as already discussed above (one-cut RE). Since C11 contains an inducible, infectious Py genome and no viral sequences migrating as *Hin*dIII-B after cleavage with the corresponding RE, we have to conclude that the *Hin*dIII-B sequences are part of these linkers, as if integration in C11 had occurred within the *Hin*dIII-B sequences of Py DNA. Since *Hin*dIII-B is also underrepresented in C10 and C12 and since all clones have common linkers, we therefore conclude that integration at the site common to all Cyp clones has occurred in the *Hin*dIII-B sequences. This conclusion was confirmed by using probes made from the *Hin*dIII fragments of Py DNA; indeed, the *Hin*dIII-B but not the *Hin*dIII-A probe was found to anneal with the fragments identified as the common linkers (data not shown).

The subgenomic tandems already observed for C10, C12, and C13 (see one-cut RE) are also obvious in Fig. 3. In the case of C10, an intense 4.0-kbp band indicates that the subgenomic unit comprises only one of the two *Hin*dIII sites. This is also the case for C12, where we see the intense 2.7-kbp band already detected in the *Bam*HI digest (Fig. 2). For C13, on the contrary, the 4.7-kbp repeat detected after *Bam*HI or *Eco*RI treatment contains both of the *Hin*dIII sites; as seen in Fig. 3, this results in the production of normal *Hin*dIII-A (3.0 kbp) and of a fragment of about 1.7 kbp. This observation suggests that the viral sequences that are not represented in the subgenomic repeat of C13 are part of the *Hin*dIII-B.

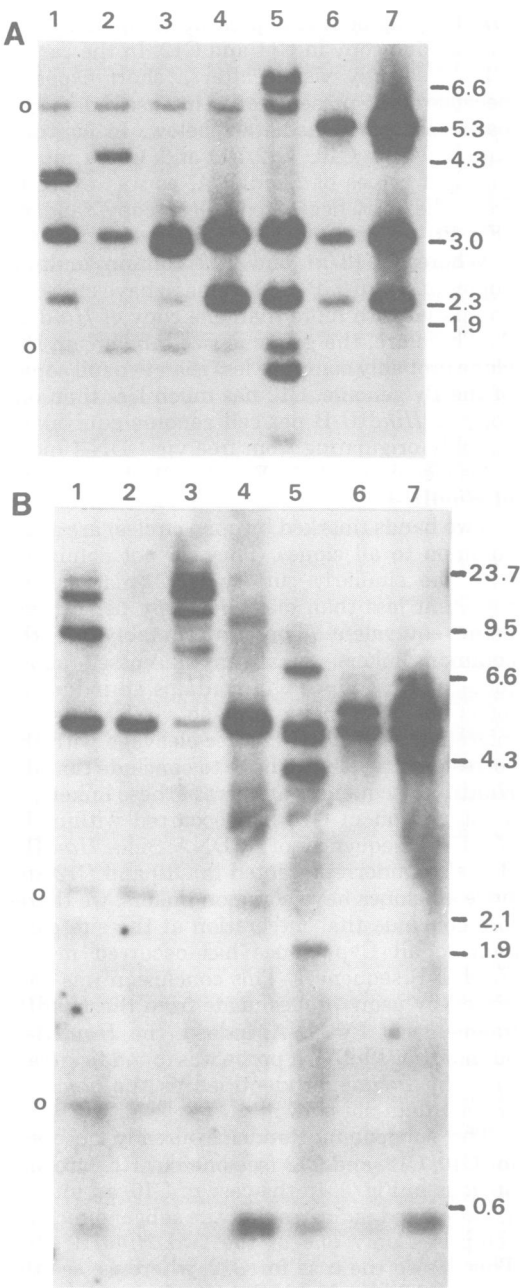

FIG. 3. *Analysis of the integrated Py DNA sequences of clones C10 (lane 1), C11 (lane 2), C12 (lane 3), C12/a1 (lane 4), and C13 (lane 5) after cleavage with HindIII (A) or HincII (B). Lanes 6 and 7 contain the size markers: (A) EcoRI-cleaved Py DNA (5.3 kbp), together with HindIII-fragment A (3.0 kbp) and HindIII fragment B (2.3 kbp) of that DNA. (B) EcoRI-cleaved Py DNA (5.3 kbp), together with HincII fragment A (4.8 kbp) and HincII fragment B (0.5 kbp) of the same DNA. In both cases, lane 6 contains about 10^{-5} μg of each EcoRI- and HindIII- or HincII-cleaved DNA, whereas lane 7 contains 5×10^{-5} μg of each. The size markers are in kilobase pairs. The circles indicate the position of the two bands believed to be the common virus-host linkers (see the text). Due to an electrophoresis artifact, the bands corresponding to one of the linkers formed a curved line in one of the gels (B).*

In the *Hin*dIII digest of C11, C12, and C13, other bands, whose significance was discussed above (see one-cut RE), are seen. Notice that there are no such additional bands in C10, as expected from the presence of viral sequences only at the common integration site in that clone.

In Fig. 3B, we present the results obtained after digestion of the DNA with *Hin*cII. This RE also cuts Py DNA twice, into fragments of about 4.8 kbp (*Hin*cII-A) and 0.5 kbp (*Hin*cII-B). In brief, C10, C11, C12/a1, and C13 are found to have one or more copies of *Hin*cII-A. This is less obvious in the case of C13 (lane 5) because of the presence of another, slightly faster, fragment containing viral sequences. The presence of a *Hin*cII-A in lane 2 indicates that sequences representing over 90% of the Py genome are present uninterrupted in the integrated viral DNA of C11. This result is important, as we have already shown that C11 contains no tandems of complete Py genomes and no *Hin*dIII-B. It thus allows us to narrow down the viral component of the common linkers to the *Hin*cII-B, whose sequences are a subset of those present in the *Hin*dIII-B. C12 (lane 3) has much less than one copy of *Hin*cII-A per cell genome equivalent, presumably originating from free viral DNA, as already discussed. Thus, the sequences included in the *Hin*dII-A are not present uninterrupted at the integration site, possibly as a result of the internal amplification of the subgenomic unit of 2.7 kbp already described above (Fig. 2, lane 10, and Fig. 3A, lane 3). If this were true, one would expect the amplified sequences to be found somewhere high up in the gel. Actually, two intense bands corresponding to about 13 and 10 kbp of DNA, can be seen in the autoradiogram (Fig. 3B, lane 3). Since they are separated by about 2.7 kbp, they may again originate from two different kinds of integrated viral structures, one having one more copy of the subgenomic unit than the other. This would explain why the two major viral bands in the no-cut digest were also separated by about 2 to 3 kbp (Fig. 1, lane 3).

*Hin*cII-B is detectable in C10, C12/a1, and C13, but not in C11 or C12. As discussed previously for *Hin*dIII-B, this may mean that the viral sequences that constitute the common linkers are part of *Hin*cII-B.

In the *Hinc*II digest of C10, we also detect two strong bands corresponding to fragments of about 13 and 9 kbp. We believe that they originate from the subgenomic tandems already detected in this clone. The high molecular weight of these species would imply that there are no *Hinc*II sites in those amplified sequences. Furthermore, the fact that there are two such species is consistent with the idea that the two bands seen in the no-cut *Bgl*II digest reflect the existence in clone C10 of two kinds of integrated viral sequences which differ only by the presence of an extra copy of the amplified 4.0-kbp subgenomic unit (see above). The band migrating slightly faster than the *Hinc*II-A in the C13 digest is likely to originate from the tandems of the 4.7-kbp subgenomic unit. This would indicate that one of the two *Hinc*II sites is thus amplified with these sequences.

Cellular sequences at the site of integration. As stated above, clone C12/a1 produces, after transfer to 33°C, a DNA molecule, referred to as RmI, consisting of the entire viral genome plus 1×10^6 daltons of mouse cellular DNA linked into a covalently closed circle (16). RmI represents 95% of the low-molecular-weight DNA synthesized after induction of C12/a1 cells; the remaining 5% is genome-size P155 DNA. A physical map of RmI is shown in Fig. 4. The cellular DNA in RmI comprises one *Bam*HI site, one *Hin*dIII site, and one *Bgl*II site. RmI is thus converted into a linear structure by *Bgl*II, cut into two fragments by *Bam*HI, and into three by *Hin*dIII. Some of these fragments (in Fig. 4, *Bam*HI-1 and -2 and *Hin*dIII-2 and -3) could be called "linkers," as they are composed of both viral and cellular sequences.

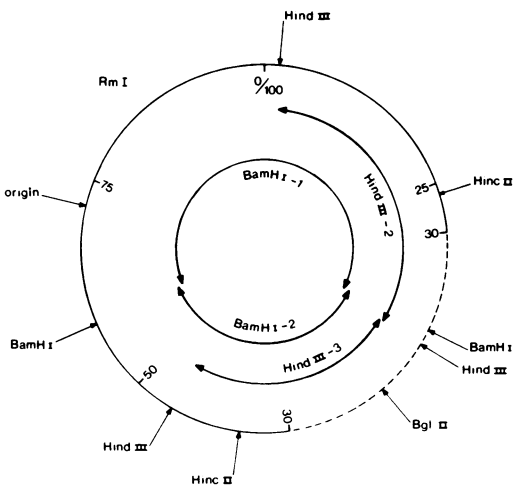

FIG. 4. *Physical map of RmI. The latter is shown as a complete Py genome with an addition of about* 1×10^6 *daltons of DNA at map position 30 (16). This DNA is of cellular origin and comprises one cleavage site for a RE that does not cut Py DNA, BglII (B. Sylla and P. Bourgaux, unpublished data). Notice that the cellular insertion also contains one cleavage site for BamHI and one for HindIII (16). Because these enzymes also cleave Py DNA, RmI is thus cut into two fragments by BamHI and into three fragments by HindIII. Of these five fragments, four contain both cellular and viral sequences. The coordinates of these four fragments (or linkers) are shown on the map.*

We have thus compared the linkers of the integrated viral genomes in the Cyp cells with the "linkers" in RmI. Before the usual analysis, cell DNAs were digested either with *Bam*HI alone or the *Bam*HI-*Bgl*II combination (Fig. 5) or, alternatively, with *Hin*dIII alone or the *Hin*dIII-*Bgl*II combination (Fig. 6). As indicated in the figures by the open circles, we could identify in each clone a linker that comigrates with a fragment of RmI. When digestion was with one RE only, this fragment was *Bam*HI-2 or *Hin*dIII-3 (Fig. 4). When either combination of two enzymes was used, the particular linker was shortened by *Bgl*II to the same extent as the corresponding RmI fragment. Thus, as judged from the respective positions of these RE sites, the common insertion and RmI share an identical arrangement of sequences on either side of one of their viral-cellular joints. This is what one would expect if RmI had been excised from the common integration site, together with some of the flanking cellular sequences.

Physical maps of the integrated viral genomes. From the results shown above and others not shown, we have constructed the physical maps of the integrated viral sequences present in the Cyp clones at the common integration site (Fig. 7).

The DNA of clone C10 contains both complete copies of Py DNA and tandems of a subgenomic unit of 4.0 kbp that includes map units 58 to 11.6 (*Bam*HI and *Kpn*I sites), but not map units 44.6 or 26.5 (*Hin*dIII and *Hinc*II sites). However, the analysis of the *Hin*dIII digest suggests that there are fewer uninterrupted copies of *Hin*dIII fragment B than of *Hin*dIII fragment A. Both the complete copies and the subgenomic tandems are at the common integration site, as C10 has only two virus-host linkers, the common ones. It is likely, however, that at that site, there are two kinds of integrated viral sequences which differ only by the presence of an additional copy of the 4.0-kbp unit, as suggested by the results obtained with *Bgl*II and *Hinc*II. Viral sequences originating from the common integration site should be linked to the same amount of cellular DNA after cleavage with a no-cut RE, regardless of the clone they come from. As this amount can be estimated to about 4.5 kbp in the *Bgl*II digest

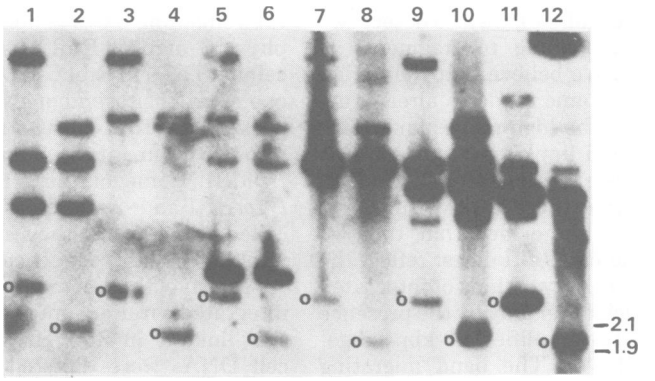

FIG. 5. *Fragments containing Py sequences in the DNA from the Cyp clones (C10, lanes 1 and 2; C11, lanes 3 and 4; C12, lanes 5 and 6; C12/a1, lanes 7 and 8; C13, lanes 9 and 10) and in RmI (lanes 11 and 12, about* 5×10^{-5} *μg in each), after digestion with BamHI (odd-numbered lanes) or double digestion with BamHI-BglII (even-numbered lanes). The size markers are indicated in kilobase pairs. The circles designate the fragments containing viral sequences that migrate similarly in all digests, including that of RmI, after one treatment or the other.*

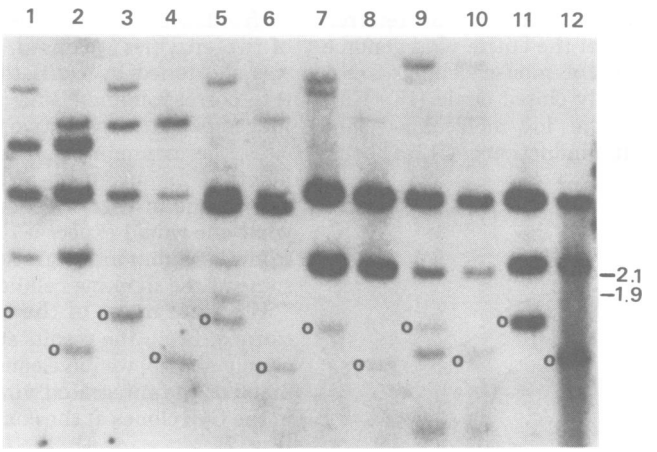

FIG. 6. *Fragments containing Py sequences in the DNA from the Cyp clones (C10, lanes 1 and 2; C11, lanes 3 and 4; C12, lanes 5 and 6; C12/a1, lanes 7 and 8; C13, lanes 9 and 10) and in RmI (lanes 11 and 12, about* 5×10^{-5} *μg in each), after digestion with HindIII (odd-numbered lanes) or a double digestion with HindIII-BglII (even-numbered lanes). The size markers are indicated in kilobase pairs. The circles designate the fragments containing viral sequences that migrate similarly in all digests, including that of RmI, after one treatment or the other.*

of C11 (see below), this would leave roughly 14 and 18 kbp of viral sequences in C10, depending on the number of 4.0 kbp units present.

In clone C12, we find head-to-tail tandems of a subgenomic unit of about 2.7 kbp. These tandems are part of the viral sequences integrated at the common site, since they migrate as though they were fused to one of the common linkers in the *Eco*RI digest. However, they are internally positioned with respect to the two viral *Hinc*II sites. They include map units 44.6 to 92.1 (*Hin*dIII and *Pvu*II sites) but not map units 36.0 or 100 (*Hinc*II and *Eco*RI sites) (Fig. 7). As in the case of C10, the analysis of *Bgl*II and *Hinc*II digests suggests that there are two kinds of integrated viral sequences at the common site which differ only in the number of copies of the 2.7-kbp subgenomic unit. However, it should be stressed that, in contrast with C10, the integrated structure in C12 does not include a complete colinear copy of Py DNA. The implication of this structural feature for the production of infectious viral DNA during induction will be dealt with below.

In the *Bgl*II digest of C11, the integrated viral sequences are present as two distinct bands with

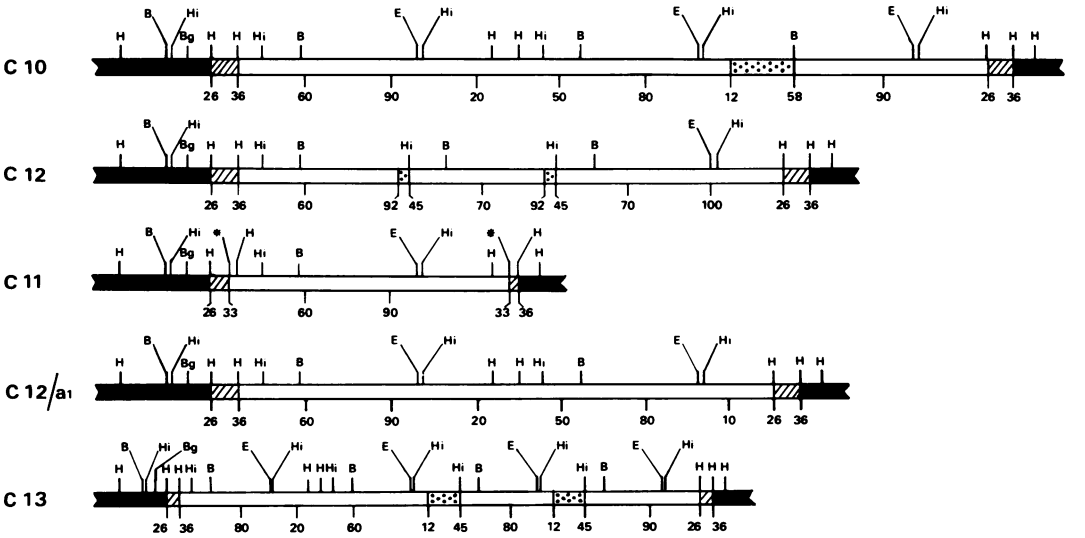

FIG. 7. *Physical maps of the integrated viral genomes present at the common integration site in the Cyp clones. The maps were constructed from the data shown in Fig. 1-3, 5, and 6 and from data obtained with other digests not shown. The Py sequences are shown in white, and the cellular sequences are shown in black. The hatched boxes on each side define the extreme limits inside of which cellular and Py DNA are covalently joined together. The dotted areas refer to the unknown endpoints of the amplified subgenomic fragments. Only the EcoRI (E), BamHI (B), BglII (Bg), HindIII (Hi), and HincII (H) sites are shown on the maps. The numbers refer to standard viral map units. The drawing scale used for C13 is half that used for the others.*

sizes of 11 and 9 kbp. In this instance, we believe the two bands indicate the presence of viral insertions at two different sites, since we can detect in some of the digests other linkers in addition to the common ones. The examination of the *Eco*RI digest (Fig. 2, lane 2) tells us that neither of the corresponding structures contains tandems of either the complete genome or of a subgenomic fragment that would include an *Eco*RI site. In that digest, indeed, the only viral material detectable besides that in the common linkers is present in a high-molecular-weight fragment (of about 11 kbp) that contains less than one copy of Py DNA per cell genome equivalent. Being thus partly cellular, this fragment is more likely to come from an integrated structure unique to C11 than from that at the common site, whose linkers are both clearly identifiable in the same gel. Actually, the analysis of numerous RE digests of C11 DNA suggests that this second integrated structure probably consists of a subgenomic segment that includes sequences between map units 53.8 and 92.1 (*Hpa*II and *Pvu*II sites). Since this second integrated structure has fewer viral sequences than that at the common site, it probably corresponds to the less intense of the two bands (that of about 11 kbp) observed for the *Bgl*II digest (Fig. 1, lane 2). As to the viral structure at the common site, it includes sequences colinear with at least 90% of the Py genome, that is to say, the sequences present in *Hin*cII fragment A. Actually, further analysis of C11 DNA with five more RE (*Kpn*I, *Pvu*II, *Pst*I, *Hha*I, and *Hpa*II; data not shown) finally led us to the conclusion that the viral sequences at the common site consist of little more than a single colinear Py genome, joined to the cellular DNA at about map unit 30 (Fig. 7). There is possibly a short duplication of the viral sequences at the joints, since the *Pst*I site at map unit 33.1 seems to be present twice (shown by an asterisk). This duplication would not extend beyond map unit 26.5 (*Hin*cII site) on one side, or map unit 36.0 (*Hin*cII site) on the other side, of the viral insert.

As seen in the *Bgl*II digest (Fig. 1, lane 4), clone C12/a1 has only one major integrated viral structure of about 15 kbp. This represents a supplement of about 5 kbp of viral DNA, as compared with the equivalent structure in C11. Since C12/a1 does not seem to contain subgenomic tandems and has the same linkers as C11, we believe that the integrated structure in C12/a1 is similar to that in C11, but with two complete copies of the Py genome in head-to-tail tandems, instead of one. Because of the presence of free viral DNA in C12/a1, it is difficult to be more specific about this clone.

Clone C13 seems to have more distinct integrated viral structures than any of the other

clones, at least judging from the *BglII* digest (Fig. 1, lane 5). Because of an overall similarity between the features observed for C10 and C13 (tandems of complete genomes, subgenomic tandems, and common linkers), we tentatively propose that in C13, the structure at the common site consists of head-to-tail tandems of Py DNA, with an internal amplification of a subgenomic segment that includes viral map units 44.6 to 11.6 (*Hin*dIII and *Kpn*I sites). It would be too much of a coincidence for the amplified subgenomic segment to be entirely colinear with the integrated viral genome in C11, and not part of the sequences integrated at the common site. We have no evidence to indicate that tandems of complete Py DNA should be present at that site. If they are not, the structure at the site would then be closer to the type seen in C12 than to that in C10.

All physical maps presented in Fig. 7 show at least one complete, uninterrupted early region. These maps also indicate that the amplified subgenomic units present in C10, C12, and C13 all comprise sequences from the same region of the Py genome, that which includes the origin of replication.

DISCUSSION

The data reported above indicate that some of the integrated Py DNA sequences present in clones C10, C11, C12, C12/a1, and C13, all derived from passage 11 of the same Cyp line, can be accounted for on the basis of a single integration event in the common ancestor cell. When DNA from these cells is digested with an RE that cuts in Py DNA, certain fragments can be detected that contain viral sequences but do not comigrate with any of the fragments from similarly cleaved Py DNA. Some of these fragments produce, after annealing with the viral probe, radioactive bands that are less intense than would be expected from exclusively viral fragments of a similar molecular weight. When the DNA from the various clones is digested with a given RE, two such fragments can be found in all digests that display the same electrophoretic mobility (Fig. 2 and 3). This is what one would expect if the same cellular sequences were linked to the same viral sequences in all clones. The Cyp clones therefore appear to have one integration site in common. As to the viral sequences comprised between these identical joints, they obviously differ from clone to clone (Fig. 7). We believe that the differences noted may be due to two phenomena already described by Basilico and his colleagues (1, 8), namely, the loss of full-copy tandems and the amplification of subgenomic segments. If they were to occur at random, such modifications would be likely to introduce heterogeneity in any cell population. This would explain why, in the case of a clone like C10, for instance, the analysis performed with no-cut RE indicates the presence of more than one kind of insert of viral DNA, whereas that carried out with both one-cut and two-cut RE does not detect more than two viral cellular joints. Another possible explanation would be that, in some cells at least, the whole viral insertion at the common site has been duplicated, together with the adjacent cellular sequences. In that hypothesis, every such cell would have acquired two copies of each of the common viral cellular joints. Subsequent rearrangements within the viral sequences themselves would allow these two initially identical insertions to diverge significantly. Both of these possibilities could apply to other Cyp clones.

For each particular clone, it is interesting to compare the organization of the integrated viral sequences, on one hand, with the responsiveness to temperature shiftdown, or the nature of the DNA synthesized after such a shift, on the other hand. As to the first comparison, our analysis indicates that it is difficult to establish a correlation between the structure of the integrated viral genome and the amount of viral DNA or virus produced after a transfer to 33°C (16). C10, the least inducible clone, carries complete viral genomes integrated in tandems, whereas C12, one of the most inducible clones, has no such tandems (Fig. 7). Even more surprisingly, C12 does not even carry a complete colinear copy of the viral genome, since the continuity of the inserted genome is interrupted by the repetition of part of its sequences. This implies that two distinct recombination events are needed to generate the numerous complete copies of the Py genome that are produced in C12 cells after induction. As to the second comparison, it seems obvious that the overall organization of the integrated viral sequences is not directly reflected by that of the free viral DNA molecules synthesized after transfer to 33°C. This conclusion is based on two main observations. One is that in most Cyp clones, the free viral DNA produced at 33°C appears to consist essentially of monomeric, nondefective Py DNA molecules (9, 16), even when the integrated viral sequences are predominantly subgenomic tandem repeats. This situation differs from that described for Py-transformed rat cells, in which free viral DNA molecules faithfully mimic the integrated viral tandems, whether genomic or subgenomic (10). Another striking observation was made by comparing the linkers with some of the fragments that can be generated from RmI. Whereas RmI is produced in detectable amounts in C12/a1

only (16), it appears that one of the linkers common to all clones is identical to one of the linkers in RmI. Since, for reasons stated above, the free viral DNA molecules generated by clones C10, C11, C12, and C12/a1 probably originate from the common integration site, one may thus wonder why one clone would accumulate mostly RmI, whereas the others accumulate monomeric viral DNA, upon transfer to 33°C. Together, the two observations we just summarized indicate that the accumulation of free viral DNA molecules occurring at 33°C in Cyp cells involves a selective mechanism through which, except in C12/a1 cells, the production of genomic molecules is favored. Whether this selective mechanism intervenes before, during, or after excision is unclear at the moment.

In the accompanying paper (9), it is demonstrated that superinfection of Cyp cells with wild-type Py results in the autonomous replication of the superinfecting, but not of the resident, viral genome. Such a failure of the resident genome to replicate autonomously after superinfection was observed for the five clones characterized in the present study, which differ widely with respect to the organization of the integrated viral sequences. It is thus difficult to envisage that the organization of integrated viral sequences in cells, all of which are inducible by temperature shiftdown, could explain why superinfection is not followed by excision.

ACKNOWLEDGMENTS

We acknowledge the expert technical assistance of Louise Bouvrette and Elizabeth Herring. We are grateful to Danielle Bourgaux-Ramoisy and Louis Delbecchi for fruitful discussions, and also to Marcel Bastin for making available cloned polyoma DNA.

This work was supported by grants from the Medical Research Council of Canada and from the National Cancer Institute of Canada. P.C. is the recipient of a scholarship from the Medical Research Council of Canada.

LITERATURE CITED

1. **Basilico, C., S. Gattoni, D. Zouzias, and G. Della-Valle.** 1979. Loss of integrated viral DNA sequences in polyoma-transformed cells is associated with an active viral A function. Cell **17**:645–659.
2. **Basilico, C., D. Zouzias, G. Della-Valle, S. Gattoni, V. Colantuoni, R. Fenton, and L. Dailey.** 1980. Integration and excision of polyoma virus genomes. Cold Spring Harbor Symp. Quant. Biol. **44**:611–620.
3. **Birg, F., R. Dulbecco, M. Fried, and R. Kamen.** 1979. State and organization of polyoma virus DNA sequences in transformed rat cell lines. J. Virol. **29**:633–648.
4. **Botchan, M., J. Stringer, T. Mitchison, and J. Sambrook.** 1980. Integration and excision of SV40 DNA from the chromosome of a transformed cell. Cell **20**: 143–152.
5. **Botchan, M., W. Topp, and J. Sambrook.** 1976. The arrangement of SV40 sequences in the DNA of transformed cells. Cell **9**:269–287.
6. **Botchan, M., W. Topp, and J. Sambrook.** 1978. Studies on simian virus 40 excision from cellular chromosomes. Cold Spring Harbor Symp. Quant. Biol. **43**:709–719.
7. **Bourgaux, P., L. Delbecchi, K. K. Y. Yu, E. Herring, and D. Bourgaux-Ramoisy.** 1978. A mouse embryo cell line carrying an inducible, temperature sensitive, polyoma virus genome. Virology **88**:348–360.
8. **Colantuoni, V., L. Dailey, and C. Basilico.** 1980. Amplification of integrated viral DNA sequences in polyoma virus-transformed cells. Proc. Natl. Acad. Sci. U.S.A. **77**:3850–3854.
9. **Delbecchi, L., D. Gendron, and P. Bourgaux.** 1981. Inducible permissive cells transformed by a temperature-sensitive polyoma virus: superinfection does not allow excision of the resident viral genome. J. Virol. **39**: 196–206.
10. **Gattoni, S., V. Colantuoni, and C. Basilico.** 1980. Relationship between integrated and non-integrated viral DNA in rat cells transformed by polyoma virus. J. Virol. **34**:615–626.
11. **Gross-Bellard, M., D. Oudet, and P. Chambon.** 1973. Isolation of high molecular weight DNA from mammalian cells. Eur. J. Biochem. **36**:32–38.
12. **Lania, L., D. Gandini-Attardi, M. Griffiths, B. Cooke, D. De Cicco, and M. Fried.** 1980. The polyoma virus 100K large T-antigen is not required for the maintenance of transformation. Virology **101**:217–232.
13. **Maniatis, T., A. Jeffrey, and D. G. Kleid.** 1975. Nucleotide sequence of the rightward operator of phage. Proc. Natl. Acad. Sci. U.S.A. **72**:1184–1188.
14. **Prasad, I., D. Zouzias, and C. Basilico.** 1976. State of the viral DNA in rat cells transformed by polyoma virus. I. Virus rescue and the presence of non-integrated viral DNA molecules. J. Virol. **18**:436–444.
15. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
16. **Sylla, B. S., D. Bourgaux-Ramoisy, and P. Bourgaux.** 1980. Induction of viral DNA synthesis in clonal derivatives of a permissive cell line transformed by a temperature-sensitive polyoma virus. Virology **100**:357–369.
17. **van der Ploeg, L. H. T., and R. A. Flavell.** 1980. DNA methylation in the human $\gamma\delta\beta$-globin locus in erythroid and nonerythroid tissues. Cell **19**:947–958.
18. **Zouzias, D., I. Prasad, and C. Basilico.** 1977. State of the viral DNA in rat cells transformed by polyoma virus. II. Identification of the cells containing non-integrated viral DNA and the effect of viral mutations. J. Virol. **24**:142–150.