Cloning and Sequencing of Viral Integration Site in Human Fibroblasts Immortalized by Simian Virus 40

Osamu Yano^{1,2}, Hideyasu Hirano¹, Yuji Karasaki¹, Ken Higashi¹, Hayato Nakamura¹, Shinobu Akiya² and Sadao Gotoh¹

¹Department of Biochemistry, University of Occupational and Environmental Health, Kitakyushu City, 807, and ²Department of Ophthalmology, University of Occupational and Environmental Health, Kitakyushu City, 807, Japan

Key words: SV40/human fibroblast/integration/immortalization

ABSTRACT. We have analyzed cellular DNA sequences at the viral genome integration site in a human fibroblast cell line VA13 immortalized by simian virus 40 (SV40). The computer analysis of the junctional cellular DNA sequences did not show any homology to the DNA sequences previously reported. This suggests that immortalization by SV40 was not induced by the destruction of any known oncogene or antioncogene at the integration site. We did not find the precise substantial sequence homology at the junctional site between the cellular DNA and SV40 DNA, indicating that the recombination mechanism involved does not require precise sequence homology and therefore, SV40 genome was probably not integrated by homologous recombination. Short direct and inverted repeats of 5 to 29 nucleotides were found in the junctional cellular and SV40 DNA.

Cellular DNA abutting SV40 DNA was found by the Northern blot analysis to be expressed in diploid human fibroblasts and SV40-transformed cells. The nature of this RNA is now under study.

Unlike rodent cells which easily form continuously growing cultures either spontaneously or after viral infection (45), diploid human fibroblasts have a limited lifespan *in vitro* (20, 21) and are resistant to transformation to an immortal form not only by spontaneous change but also physical or chemical means (27, 37). When infected with simian virus 40 (SV40), human fibroblasts can easily be transformed and can grow longer than their normal counterpart but eventually deteriorate and perish (13, 14, 24).

In all SV40-transformed cells, at least a part of the viral DNA is persistently integrated in host cellular DNA (46) and the functional large T antigen is expressed until the very end of their lifespan (14). In very rare cases, however, an immortalized cell arises from these SV40-transformed cells (13, 14, 24). Taken together, the above suggests that large T antigen is required to transform human fibroblasts but that it is not enough to immortalize them.

Some authors have suggested that the integration of the SV40 DNA in SV40-transformed and immortalized human fibroblasts occurs at random locations. However, these authors investigated integration sites only by

analyzing the assignment of SV40 DNA to specific chromosomes (7, 8, 31) or by analyzing Southern blotting patterns (5, 18, 25). Their arguments can be criticized on several points. First, if short stretches of specific cellular DNA sequences would border viral insertions, they could not be detected by Southern blot analysis. Moreover, the different alterations of cellular DNA after the integration of a virus genome at the specific sites should produce restriction fragments of various sizes in different cell lines. Furthermore, when differently rearranged SV40 genomes are integrated at many sites including the same hot spot, the Southern blotting pattern would not be same. Although in the case of rat cells, it has been reported that SV40 DNA is not integrated at the specific site by sequencing the junctional nucleotides sequences of only 20 bases or so (15, 44), no one has reported the DNA structure of SV40 DNA integration site at the nucleotide level in immortalized human fibroblasts.

In our present study, we cloned the DNA fragments containing the integrated viral DNA from SV40-transformed human fibroblast VA13 clone 1 cells into prokaryotic vectors and determined the DNA sequences encompassing the junctions between SV40 and cellular DNA. Also, we performed the Northern blot analysis (1, 32) to investigate the expression of the junctional cellular DNA sequence and found they were expressed in

Abbreviations: SV40, simian virus 40; MCF-PBS, magnesium, calcium free phosphate-buffered saline; TE, 10 mM Tris-HCl and 1 mM EDTA, pH 8.0

both normal diploid and SV40-immortalized human fibroblasts.

MATERIALS AND METHODS

Cells. WI-38 human diploid cells were derived from the normal embryonic (3 month gestation) lung tissue of a Caucasian female, and the cell line VA13 is a SV40-transformed and immortalized derivative of WI-38 and was a generous gift from Dr. H. Koprowski (26). A clone of VA13 (VA13 clone 1) was prepared by plating about 10,000 cells in 60 mm Petri dishes which contained many small fragments of microscope cover slips, then transferring a fragment that bore only one cell into a small well. This cloning procedure was repeated twice. YH-1 cells were obtained and reported elsewhere (47).

Cell DNA. High molecular weight cellular DNA was prepared by the method of Blin and Stafford (4). In short, confluent cells were harvested by trypsinization, washed with magnesium, calcium free phosphate-buffered saline (MCF-PBS) and lysed by incubation in Proteinase K (300 μ g/ml), 0.5 M EDTA, pH 8.0 and 1% sodium dodecyl sulfate (SDS) at 55°C for 3 hr. The lysate was extracted with phenol saturated with 10 mM Tris-HCl and 1 mM EDTA, pH 8.0 (TE) containing 8 hydroxyquinoline (5%). The extract was dialyzed against 50 mM Tris-HCl, 10 mM EDTA, pH 8.0 and 10 mM NaCl exhaustively, incubated with pancreatic ribonuclease A (100 μ g/ml, DNase heat inactivated) at 37°C for 3 hr and then phenol extracted again. The aqueous phase was dialyzed against the TE exhaustively.

Digestion with restriction endonucleases and Southern blotting. Restriction endonucleases were purchased from either TAKARA SHUZO CO., LTD. (KYOTO, JAPAN) or TOYOBO CO., LTD. (OSAKA, JAPAN), and DNA samples were restricted under the conditions specified by the enzyme suppliers. Samples of digested DNA were electrophoresed on a 0.5% agarose gel and transferred to a nylon filter (PALL CORP., NEW YORK, USA) by the method of Southern (43). 32 P-labeled SV40 DNA was used as a probe to detect the integrated SV40 DNA sequence in the host genome. SV40 DNA was purchased from Bethesda Research Laboratories, Inc., (MARYLAND, USA).

Recovery of DNA fragment from an agarose gel. Total cellular DNA was digested by either EcoR I or BamH I and then electrophoresed. Subsequently referring to the results of Southern blots, the gel slice which has the target DNA fragment containing the integrated SV40 DNA sequence was cut out and DNA fragment was recovered by the method of Yang *et al.* (48).

DNA cloning. The bacteriophage λ vectors used were Charon 4A (50) and Charon 27 (40). Charon 4A DNA and Charon 27 DNA was digested with EcoR I and BamH I, respectively. The Charon 4A EcoR I arms were ligated to EcoR I digests of the cellular DNA and the Charon 27 BamH I arms to BamH I digests of the cellular DNA. The resulting concatemeric DNA was packaged into phage particles using

GIGAPACK GOLD packaging extract (STRATAGENE, CALIFORNIA, USA), and the phages were introduced to an *E. coli* strain LE392. Recombinant phages containing SV40 DNA sequence were identified by the plaque hybridization method with 32 P-labeled SV40 DNA as a probe (2). EcoR I and BamH I DNA fragments respectively containing junctional cellular and SV40 sequences were recovered from the corresponding recombinant phage DNA. Each junctional DNA fragment was digested by Sph I, Pst I, Acc I, Ava II and EcoR I and the resulting subfragments were then subcloned into the plasmid vector pUC18. The hybrid plasmids were introduced into an *E. coli* strain JM109 by the calcium precipitation method (36).

DNA sequencing. The nucleotide sequences were mainly determined by the chain termination method of Sanger *et al.* (41) but the Acc I-Sph I fragment (Fig. 2) of the cellular DNA abutting the left side of SV40 DNA which was obtained from fragment ② (Fig. 2) was determined by the method of Maxam and Gilbert (37).

Northern blotting. Total cellular RNA was extracted from normal human fibroblast YH-1 and SV40-immortalized human fibroblast VA13 clone 1 by the guanidinium thiocyanate method (3, 6). Samples of 20 μ g of total cellular RNA were electrophoresed in a 1% agarose-formaldehyde gel and were transferred to a nylon filter (Amersham International, ENGLAND) (1, 33). 32 P-labeled DNA probes used were cellular DNA fragments abutting SV40 DNA (0.25 kb Sph I-Acc I fragment or 0.27 kb Acc I fragment or 0.9 kb Sph I-BamH I fragment in Fig. 2) or β -actin gene fragment (1.8 kb Pst I fragment) which was a generous gift from Dr. B.M. Paterson (29). Hybridization was conducted under $4 \times$ SSC, 50% formamide at 42°C for 17 hr. The filters were washed at 68°C twice with $2 \times$ SSC, 0.1% SDS for 30 min, twice with $1 \times$ SSC, 0.1% SDS for 30 min and twice with $0.1 \times$ SSC, 0.1% SDS for 30 min.

Computer analysis of DNA sequence. The DNA sequence data were compared with Gen Bank data (Gen Bank release 63, c/o Intelligenetics Inc., CALIFORNIA, USA) and analyzed by our computer program (Hirano, H. 1985 DNA sequence analysis system GENE grade III, System Soft Kitakyushu Inc., FUKUOKA, JAPAN) (30) for sequence homology.

RESULTS

Molecular cloning of the junctional DNA sequence at the viral integration site in VA13 clone 1 cells. SV40 genome is integrated at many sites in most of SV40-transformed and immortalized human fibroblast cell lines (18, 25) and it is difficult in such cell lines to investigate whether the SV40 DNA integration site exhibits characteristic sequences in establishing immortalization. Therefore as a starting material for our study we chose SV40 transformed and immortalized human fibroblasts VA13 clone 1 cells, in which SV40 genome is integrated at a single site.

Fig. 1. Southern blot analysis of total cellular DNA from SV40-immortalized human fibroblast (VA13 clone 1). Approximately 20 μ g of total cellular DNA from VA13 clone 1 was applied per well. 32 P-labeled SV40 DNA was used as a probe. Lane 1: 1 ng of supercoiled and relaxed circular SV40 DNA as a positive marker, Lane 2: EcoR I digests of total cellular DNA, Lane 3: BamH I digests of total cellular DNA.

Southern blotting pattern (43) of total cellular DNA from VA13 clone 1 cells which was restricted with EcoR I or BamH I is shown in Fig. 1. EcoR I and BamH I digestion produced two bands each which are hybridizable to 32 P-labeled SV40 DNA probe (EcoR I: 15 kb and 2.5 kb, BamH I: 18 kb and 2.5 kb). The 15 kb EcoR I (fragment I) and the 2.5 kb BamH I (fragment II) DNA fragments containing the SV40 DNA were

cloned into phage vectors charon 4A and charon 27, respectively (Fig. 2). The cloned EcoR I or BamH I fragment obtained as aforementioned comigrated on an agarose gel with the band detected by Southern blotting when total cellular DNA was digested with the same enzyme (data not shown).

Restriction mapping and sequencing strategy of the DNA fragments cloned from VA13 clone 1 cells. Fig. 2 shows the restriction mapping of the DNA fragments cloned from VA13 clone 1 cells. The integrated SV40 DNA is about 13.5 kb long, which is larger than SV40 unit length (5243 base pairs) and the numbering of the nucleotides of SV40 genome is not continuous all the way from end to end. Taken together, the above findings suggest that the SV40 genome is rearranged complicatedly in this cell line. Each of EcoR I-Sph I fragment (fragment ①), Sph I fragment (fragment ②), Sph I-EcoR I fragment (fragment ③) and EcoR I-BamH I fragment (fragment ④) shown in fig. 2 was subcloned into a plasmid vector pUC18 and sequenced. The central part of the integrated SV40 genome was restricted, mapped and sequenced partially. The length of each restriction fragment was collated with its partial sequences. Based on the congruency between them, we inferred that the central part of the integrated SV40 genome (13.5 Kb) probably does not contain cellular sequences, at least not long ones.

Nucleotide sequences of cellular DNA at the viral integration site. Figure 3A and 3B show the cellular DNA sequences abutting the integrated SV40 DNA in VA13

Fig. 2. Restriction mapping and sequencing strategy of the DNA fragments cloned from VA13 clone 1 cells. DNA fragment I of 15 kb and fragment II of 2.5 kb which contain both cellular and SV40 DNA were cloned from VA13 clone 1 cells. The open bar and solid lines represent integrated SV40 DNA and flanking cellular DNA sequence, respectively. The numbers above the open bar represent the nucleotide numbers in the SV40 DNA sequences after Tooze (41). Distances are approximately to scale except for the central part of the integrated SV40 genome. Upper: The fragment I represents the EcoR I fragment of about 15 kb and the fragment II represents the BamH I fragment of about 2.5 kb, Lower: The fragments ①-④ were subcloned into the plasmid vector pUC18 and then sequenced.

10	20	30	40	50
AGATCCAGCC	ACATGTGAAC	CCGTCTGTAC	TAAAAACAA	ACATTAGCCA
60	70	80	90	100
TGCATGTGCG	GCACCTGAGT	CCCAGCTACT	CGAGTGAGCA	GAGAATCACT
110	120	130	140	150
TGAAGCCAGA	GCAGAGGTGT	GTGAGTGAAG	TCACAGCCAC	TGCACTCCAG
160	170	180	190	200
CCTGGGTGAC	AGAGCGAAAC	TCCATTTGGA	CAAAAAAGAA	AGAAAGTACT
210	220	230	240	250
CCATAAATAT	TTGCAGGTGG	AAAGAATGAT	ATTCACATAT	GATTTAATCA
260	270	280	290	300
AAGACTGACA	GTTTATCTAT	TTGATAAAAA	GACTGTGAGA	AGCAGAGATT
310	320	330	340	350
GATGGAGACA	GGAGAGAAAG	TGGAGGAGAA	GAGAAGATGG	AGACAGCACA
360	370	380	390	400
CACACATGCA	CAGACTTGAA	CATGCAGACA	GACATTGCAC	ATGCACATGC
410	420	430	440	450
ATGCACATAC	ACAGAGGCAT	ATATACACAT	GTACTAACAT	ACGCATAGTC
460	470	480	490	500
ACAGATACAC	TGTAGGTGCA	AACACCTATA	TATGTGTTCA	CATTCATACA
510	520	530	540	550
CGGGAGTTAT	TATATTTGTA	CTCACACATG	TAGAGGGTAG	AATTGTGGTT
560	570	580	590	600
ATTAGATGCT	GGGAAGAGTA	GAGGAAGAGA	AGATAGGGAG	AGGCTGGTTG
610	620	630	640	650
GTGGATACAA	AGTTACACCT	GAATGAGAGA	TAATTTTGAC	TGTTCTGTAT
660	670	680	690	700
ACTGTAGGAT	GAATATGGTT	AACAATAACT	TAGTGTATAT	TTTCAAAAAG
710	720	730	740	750
CCAGAAGAGA	GGATCTTGAA	TGTTACACACC	ACAAATAAAT	GATAAATGTT
760	770	780	790	800
CGAGGTGATG	GATATGCTAC	TCTGATTTGA	TCATTACACA	TTGTATAAAT
810	820	830	840	850
GTATTGAAAA	ATCATCTGTA	TCCCATACAT	ATGTATAATT	ATTATGCATC
860	870	880	890	900
ACTAAAATTA	AAGACCATAT	ATGCACACAT	AGACATGCAC	ATGACAACAT
910	920	930	940	950
ACTCATGCAT	ATGTGTATAC	ATGCAGTTAC	ACATATACAC	AAACACACAT
960	970	980	990	1000
<u>ATATGCACAC</u>	<u>ATAGACATGC</u>	<u>ACATGAAGAC</u>	ACTATGCACA	TGTCCACACA
1010	1020	1030	1040	1050
TGCACACAGC	CACACACATA	CATATATACA	CAAATACCCA	TAAATATGCA
1060	1070	1080	1090	1100
CATACACAGA	TACTGAACAC	TGTGAACAAA	GGTTTGTA	CCAGGAAATA

Fig. 3. A

clone 1 cells. The cellular flanking DNA sequences did not show an extensive homology over a wide distance to the DNA sequences of any genes previously reported, except some short stretches of 80 to 120 nucleotides which are homologous to short sequences in several genes, such as human growth hormone gene (percent match 81.7%) (11), human Alu family interspersed repeat (percent match 79.1%) (23, 9), human haptoglobin gene (percent match 80.8%) (35) and human apolipo-

protein B gene (percent match 78%) (33).

Nucleotide sequence at virus-cell DNA junctions.

Fig. 4 shows the junctional nucleotide sequences of the two restriction fragments (fragment ② and fragment ④ in Fig. 2) which involve transitions between viral and cellular DNA sequences. The transition is sharp and therefore it allows a precise and unambiguous assignment of the junctional nucleotides to the viral or cellular DNA on both sides of the viral insertion. SV40 DNA runs up

10	20	30	40	50
AAAAGCTTGA	GAAATGGCAT	TAAAAAGTGT	AGACAGCATG	CACACAGAGA
60	70	80	90	100
TACATATGCA	CACACACAGC	CACATGTATA	CACACTTGTG	CAAATGTATG
110	120	130	140	150
CACACACAGA	CACACATATG	CCCATACATA	GGCATATACA	CGTACAGATT
160	170	180	190	200
CAGATATAAA	CACATACAAA	CACTCATGTA	CATATGCAAA	CACATGTACT
210	220	230	240	250
CACACAGTCT	GCTGGCATAG	TGATGGCTAA	ACCAATGCAG	GGTAGGGCAA
260	270	280	290	300
AATGTCTCTG	GTACAAGGTC	AGTAGGATAA	AATCTGCAGG	ATCAAATGGA
310	320	330	340	350
ATGCTTAACA	ACACTCCTCT	GTGGGGTATC	AGTGTCTCCA	TTTTACAACC
360	370	380	390	400
AATGCTTAAT	CAAATGCACA	TTGAGTTGAA	CTGAATATTG	GAAAATGAAT
410	420	430	440	450
TGCTACCTTT	GAATGAATCT	ACTGAAAAAT	TATACTTGTC	CTGGGTCAGA
460	470	480	490	500
CCGTGTGAAA	AGCACTGGGT	ATTGGAGATG	AAAAGTGTTT	GGCCCTTTTT
510	520	530	540	550
CTTTGAAAGC	TTAGTCCAGT	GGGGAAGATG	GATGTGGAAG	CTGAGCAATT
560	570	580	590	600
AAAATGCAGC	TTGATAAAAG	CTGTGCTGAA	AAACCCTCAC	TGAGCTGAGC
610	620	630	640	650
AGTGGGGGCT	GGGAGGGAGC	AGGTCTTAAA	CTGAATGATC	AGGTTTAGTG
660	670	680	690	700
GGAGCTGTTT	AAATTCTGAG	CCATCAGGCA	TTTTTGCTGG	ATTCCTTTTCG
710	720	730	740	750
GTGCGCTTGT	TAATCAATTA	TAAAATAATA	ATAATGAACA	CTTCAGGAGA
760	770	780	790	800
GAGGATGGAG	GTTGAAAGAC	AGTGGAGGGC	TGGACTTGAG	GTTGGAAGAT
810	820	830	840	850
TTGCAGAGAC	CTGGATTGTG	CATTGACCT	TTACAAATTG	AAGCGAGAGG
860	870	880	890	900
CTTTGCATCT	CTGAACATCA	GGATTCCACT	CTATAAAAAG	GAGAATTGAA
910	920	930	940	950
AGCATTCAGT	TTATAGGTTT	GGTGTGAGGG	CTCAACGAGA	ATGTGGATCC

Fig. 3. B

Fig. 3. Cellular DNA sequences at the SV40 integration site. A: Cellular DNA sequence adjacent to the left end of the integrated SV40 DNA in Fig. 2 and boxed sequence is of the junctional SV40 DNA. Numbering is from left to right toward SV40 sequence and No. 1 was chosen arbitrarily in the cellular DNA. B: Cellular DNA sequence adjacent to the right end of the integrated SV40 DNA in Fig. 2 and boxed sequence is of the junctional SV40 DNA. Numbering is from left to right away from SV40 DNA and No. 1 was chosen arbitrarily in the SV40 DNA.

to nucleotides 1023 and 1001 at either end of the SV40 genome, and then is followed by the sequence of cellular origin. We did not find the precise homology at the junctional sites between cellular and viral DNAs, indicating that the recombination mechanism involved did

not require precise sequence homology in VA13 clone 1 cells. However, we did find a "Patchy" homology (15, 16, 17, 42) between cellular and viral DNA (Fig. 4).

Integration junction points are located near short direct repeats of the cellular DNA. As in SV40-trans-

Fig. 4. Nucleotide sequences at the cell-virus DNA junctions in the VA13 clone 1 cells. (A) and (B) are nucleotide sequences at the left and right junctions, respectively. Upper sequence represents the junctional cellular and viral DNA sequence of fragment ② and fragment ④ in Fig 2 as determined by us (VA13). Lower sequence is the corresponding SV40 sequence (SV). Homologous sequence is indicated with double dots. Nucleotide sequences overlined and underlined represent the "Patchy" homology between VA13 cellular and SV40 DNA. The arrowhead represents the junctional point between cellular and viral DNA.

formed rat cells (19), we found short direct and inverted repeat motifs in SV40 and cellular DNA near the junctional points. Some of the repeat motifs occurred on both sides of the integrated SV40 DNA in either direct or inverted form. Table I shows only the representatives of these motifs composed of 7 nucleotides, which occurred more than 6 times in cellular DNA abutting the integrated SV40 DNA. These motifs are scattered in the flanking cellular DNA, but the distribution of these motifs is not even. The cellular flanking sequences of both sides were divided into 5 parts, each containing 200 bp. Then, 65% (51 among 78 motifs) of these

Table I. THE REPRESENTATIVE MOTIFS COMPOSED OF 7 NUCLEOTIDES.

Motif	Repetition		Total number of Repetition
	^{a)} Direct	Inverted	
5'-ATGCACA-3'	14	1	15
5'-ACACATA-3'	7	2	9
5'-ACATGCA-3'	8	1	9
5'-ACACACA-3'	9	0	9
5'-CACACAT-3'	8	0	8
5'-CACATGT-3'	6	1	7
5'-GCACACA-3'	7	0	7
5'-CATATAT-3'	4	3	7
5'-TGCACAT-3'	7	0	7

These motifs occurred more than 6 times in cellular DNA abutting the integrated SV40 DNA in direct or inverted or both orientations.

a) "Direct" means from left to right in Fig. 2.

motifs were found in the closest parts to the left and right virus-cell junctional points. By contrast, only 5% (4 among 78 motifs) of these motifs were distributed in the farthest parts at both sides. Namely, the closer these motifs were to the junctional points, the denser their distribution.

The motifs composed of more than 7 nucleotides also occurred in the flanking cellular DNA and one such motif composed of 29 nucleotides occurred twice in the cellular DNA abutting the left side of the integrated SV40 DNA sequences (underlined in Fig. 3A). Some direct and inverted motifs occurred not only in the cellular DNA but also in SV40 DNA. Many motifs composed of less than 6 nucleotides are also found scattered in the cellular and SV40 DNA either in a direct or inverted form (data not shown).

Expression of cellular DNA sequences at the viral integration site. We were interested in seeing whether or not the integrated SV40 DNA affected the expression of host genes neighboring the SV40 integration site. We extracted the total cellular RNA from VA13 clone 1 and normal human fibroblast YH-1 cells (47) and performed the Northern blot analysis using as a probe a radio-labeled cellular DNA fragment neighboring the SV40 integration site in VA13, which consists of either Sph I-Acc I or Acc I subfragment of fragment ② or Sph I-BamH I subfragment of fragment ④ in Fig. 2. In both VA13 clone 1 and YH-1 cells, the comparable amount of RNA of the same size (about 400 nucleotides) was de-

tected, only when Sph I-Acc I subfragment was used as a probe.

DISCUSSION

In SV40-immortalized rat cells, it has been reported that SV40 DNA is not integrated at the specific site by sequencing the junctional nucleotides (15, 44). However, no one has thus far reported the cellular DNA sequences at the SV40 genome integration site in human cells. Therefore, we wanted to characterize the DNA sequences at the SV40 integration site in human immortalized fibroblasts. Several authors have argued that there is no site specificity for integration of SV40 genome into cellular DNA in SV40-immortalized human cells on the grounds of the different chromosomal assignment of integrated SV40 genome (7, 8, 31) or the different Southern blot patterns among different cell lines (5, 18, 25). However, we consider their arguments and their evidence as not sufficiently strong to elucidate sequential characteristics at the viral integration site (see INTRODUCTION). Therefore, in our present study, we determined the DNA sequences encompassing the junction between viral and cellular DNA in SV40-immortalized human fibroblasts in which the viral genome is integrated at only a single site.

At the junctional site the transition between viral and cellular DNA sequence is sharp (Fig. 4). We did not find the precise homology between SV40 and cellular DNA (Fig. 4), indicating that the integration of SV40 DNA into the host cellular DNA is not brought about by a homologous recombination. However, we found a "Patchy" homology (15, 16, 17, 42) exists between SV40 and cellular DNA (Fig. 4), which has been reported to define a site of interaction between DNAs undergoing recombination in eukaryotic cells (15, 16, 17, 42).

Also, we found the short direct and inverted repeats in cellular DNA near the junctional points (Table I), the significance of which is thus far unknown. However, the fact that the closer these motifs are to the cell-virus junctions, the denser their distribution, suggests their positive contribution to the recombinational integration event of the SV40 genome. It has been reported that direct repeat sequences and palindromic DNA sequences may have a single strand-forming potential *in vivo* (10, 12, 22, 34, 49) and that recombinations preferentially occur in the single stranded part of the gene (28).

The cellular flanking DNA sequences were subjected to a computer program for sequence homology analysis, but we could not find any extensive homology to the DNA sequence of genes previously reported, such as oncogenes or antioncogenes, except for some short homologous stretches of 80 to 120 nucleotides found in several genes. Therefore, it seems unlikely that the inte-

Fig. 5. Northern blot analysis of total cellular RNA. Samples of (20 μ g) from VA13 clone 1 (lane 1) and YH-1 (lane 2) were analyzed. A: 32 P-labeled cellular DNA fragments abutting SV40 DNA (0.25 kb Sph I-Acc I fragment in Fig. 2) were used as a probe, B: 32 P-labeled β -actin gene fragment (1.8 kb Pst I fragment) was used as a probe (29).

grated SV40 DNA itself directly affected the function of an oncogene or an antioncogene at the integration site in VA13 clone 1 cells.

It was proposed that immortalized cells can be classified into four complementation groups and so there are at least four genes involved in mechanisms regulating mortality or immortality (39). Integration of SV40 DNA either in or near one of these genes may affect the function and the expression of these genes, which may contribute to immortalizing human cells. We performed the Northern blot analysis to investigate the expression of the junctional cellular DNA in SV40-immortalized human fibroblast (VA13 clone 1) and in normal human fibroblast (YH-1). In both VA13 clone 1 and YH-1 cells, the comparable amount of RNA of the same size (about 400 nucleotides) was detected (Fig. 5). It is unknown whether the RNA transcribed from the junctional cellular DNA sequence in VA13 clone 1 and YH-1 cells have the same functional activity in both cells. The study of the nature of these RNAs is now under way.

Acknowledgments. We thank Prof. D. Schlessinger of Washington University School of Medicine for his valuable discussion. We are deeply indebted to Dr. Y. Sakaki and Dr. M. Hattori of Kyushu University for their useful technical instructions and for provision of pUC18; and to Drs. T. Matsui, H. Saiga and T. Nomoto, Department of Molecular Biology, University of Occupational and Environmental Health, for their useful discussion. Charon 27 phage vector is a generous gift from Dr. Frederik R. Blattner of University of Wisconsin-Madison. This research was supported in part by grants from the Fukuoka Cancer Society and the Ministry of Education, Science and Culture of Japan.

The nucleotide sequence data reported in this paper will appear in the DDBJ, EMBL and Gen Bank Nucleotide Sequence Databases under the accession number D00846 and D00847.

REFERENCES

- ALWINE, J.C., KEMP, D.J., and STARK, G.R. (1977). Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. *Proc. Natl. Acad. Sci. USA*, **74**: 5350–5354.
- BENTON, W.D. and DAVIS, R.W. (1977). Screening λ gt recombinant clones by hybridization to single plaque *in situ*. *Science*, **196**: 180–182.
- BERGER, S.L. and CHIRGWIN, J.M. (1989). Isolation of RNA. *Methods Enzymol.*, **180**: 3–13.
- BLIN, N. and STAFFORD, D.W. (1976). A general method for isolation of high molecular weight DNA from eukaryotes. *Nucleic Acids Res.*, **3**: 2303–2308.
- CAMPO, M.S., CAMERON, I.R., and ROGERS, M.E. (1978). Tandem integration of complete and defective SV40 genomes in mouse-human somatic cell hybrids. *Cell*, **15**: 1411–1426.
- CHIRGWIN, J.M., PRZYBYLA, A.E., MACDONALD, R.J., and RUTTER, W.J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonucleases. *Biochemistry*, **18**: 5294–5299.
- CROCE, C.M., GIRARDI, A.J., and KOPROWSKI, H. (1973). Assignment of the T-Antigen of Simian Virus 40 to human chromosome C-7. *Proc. Natl. Acad. Sci. USA*, **70**: 3617–3620.
- CROCE, C.M. (1977). Assignment of the integration site for simian virus 40 to chromosome 17 in GM54VA, a human cell line transformed by simian virus 40. *Proc. Natl. Acad. Sci. USA*, **74**: 315–318.
- DEININGER, P.L., JOLLY, D.J., RUBIN, C.M., FRIEDMANN, T., and SCHMID, C.W. (1981). Base sequence studies of 300 nucleotide renatured repeated human DNA clones. *J. Mol. Biol.*, **151**: 17–33.
- DEURING, R., WINTERHOFF, U., TAMANOI, F., STABEL, S., and DOERFLER, W. (1981). Site of linkage between adenovirus type 12 and cell DNAs in hamster tumour line CLAC3. *Nature*, **293**: 81–84.
- FIDDES, A.C., SEEBURG, P.H., DENOTO, F.M., HALLEWELL, R.A., BAXTER, J.D., and GOODMAN, H.M. (1979). Structure of genes for human growth hormone and chorionic somatomammotropin. *Proc. Natl. Acad. Sci. USA*, **76**: 4294–4298.
- GLIKIN, G.C., GARGIULO, G., RENA-DESCALZI, L., and WORCEL, A. (1983). Escherichia coli single-strand binding protein stabilizes specific denatured sites in superhelical DNA. *Nature*, **303**: 770–774.
- GIRARDI, A.J., JENSEN, F.C., and KOPROWSKI, H. (1965). SV40-induced transformation of human diploid cells: Crisis and Recovery. *J. Cell Comp. Physiol.*, **65**: 69–84.
- GOTOH, S., GELB, L., and SCHLESSINGER, D. (1979). SV40-transformed human diploid cells that remain transformed throughout their limited lifespan. *J. Gen. Virol.*, **42**: 409–414.
- GUTAI, M.W. and NATHANS, D. (1978). Evolutionary variants of Simian Virus 40: Cellular DNA sequences and sequences at recombinant joints of substituted variants. *J. Mol. Biol.*, **126**: 275–288.
- GUTAI, M.W. (1981). Recombination in SV40-infected cells: Nucleotide sequences at viral-viral recombination joints in naturally arising variants. *Virology*, **109**: 344–352.
- GUTAI, M.W. (1981). Recombination in SV40-infected cells: Viral DNA sequences at sites of circularization of transfecting linear DNA. *Virology*, **109**: 353–365.
- HARA, H. and KAJI, H. (1987). Random integration of SV40 in SV40-transformed, immortalized human fibroblasts. *Exp. Cell Res.*, **168**: 531–538.
- HASSON, J.-F., MOUGNEAU, E., CUZIN, F., and YANIV, M. (1984). Simian Virus 40 illegitimate recombination occurs near short direct repeats. *J. Mol. Biol.*, **177**: 53–68.
- HAYFLICK, L. and MOORHEAD, P.S. (1961). The serial cultivation of human diploid cell strains. *Exp. Cell Res.*, **25**: 585–621.
- HAYFLICK, L. (1965). The limited *in vitro* lifetime of human diploid cell strains. *Exp. Cell Res.*, **37**: 614–636.
- HENTSCHEL, C.C. (1982). Homocopolymer sequences in the spacer of a sea urchin histone gene repeat are sensitive to S1 nuclease. *Nature*, **295**: 714–716.
- HOUCK, C.M., RINEHART, F.P., and SCHMID, C.W. (1979). A ubiquitous family of repeated DNA sequences in human genome. *J. Mol. Biol.*, **132**: 289–306.
- HUSCHTSUHA, L.I. and HOLLIDAY, R. (1983). Limited and unlimited growth of SV40-transformed cells from human diploid MRC-5 fibroblasts. *J. Cell. Sci.*, **63**: 77–99.
- HWANG, S.P. and KUCHERLAPATI, R. (1980). Localization and organization of integrated simian virus 40 sequences in human cell line. *Virology*, **105**: 196–204.
- JENSEN, F., KOPROWSKI, H., PAGANO, J.S., PONTEN, J., and RAVDIN, R.G. (1964). Autologous and homologous implantation of human cells transformed *in vitro* by simian virus 40. *J. Nat. Cancer Inst.*, **32**: 917–937.
- KAKUNAGA, T. (1978). Neoplastic transformation of human diploid fibroblast cells by chemical carcinogens. *Proc. Natl. Acad. Sci. USA*, **75**: 1334–1338.
- KOSHY, R., KOCH, S., FREYTAG VON LORINGHOVEN, A., KAHMANN, R., MURRAY, K., and HOPFSCHNEIDER, P.H. (1983). Integration of Hepatitis B virus DNA: Evidence for integration in single-stranded gap. *Cell*, **34**: 215–233.
- KOST, T.A., THEODORAKIS, N., and HUGHES, S.H. (1983). The nucleotide sequence of the chick cytoplasmic β -actin gene. *Nucleic Acids Res.*, **11**: 8287–8301.
- KUBOMURA, S., OBARA, M., KARASAKI, Y., TANIGUCHI, H., GOTOH, S., TSUDA, T., HIGASHI, K., OHSATO, K., and HIRANO, H. (1987). Genetic analysis of the cell binding domain region of the chicken fibronectin gene. *Biochim. Biophys. Acta*, **910**: 171–181.
- KUCHERLAPATI, R., HWANG, S.P., SHIMIZU, N., MCDUGALL, J.K., and BOTCHAN, M.R. (1978). Another chromosomal assignment for a simian virus 40 integration site in human cells. *Proc. Natl. Acad. Sci. USA*, **75**: 4460–4464.
- LEHRACH, H., DIAMOND, D., WOZNEY, J.M., and BOEDTKER, H. (1977). RNA molecular weight determination by gel electrophoresis under denaturing condition, a critical reexamination. *Biochemistry*, **16**: 4743–4751.
- LUDWIG, E.H., BLACKHART, B.D., PIEROTTI, V.R., CAIATI, L., FORTIER, C., KNOTT, T., SCOTT, J., MAHLEY, R.W., LEVY-WILSON, B., and MCCARTHY, B.J. (1987). DNA sequence of the human apolipoprotein B gene. *DNA*, **6**: 363–372.
- MACE, H.A.F., PELHAM, H.R.B., and TRAVERS, A.A. (1983). Association of an S1 nuclease-sensitive structure with short direct repeats 5' of *Drosophila* heat shock genes. *Nature*, **304**: 555–557.
- MAEDA, N. (1985). Nucleotide sequence of the haptoglobin and haptoglobin-related gene pair: The haptoglobin-related gene contains a retrovirus-like element. *J. Biol. Chem.*, **260**: 6698–6708.

36. MANDEL, M. and HIGA, A. (1970). Calcium-dependent bacteriophage DNA infection. *J. Mol. Biol.*, **53**: 159-162.
37. MAXAM, A.M. and GILBERT, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavage. *Methods Enzymol.*, **65**: 499-560.
38. NAMBA, M., NISHITANI, K., FUKUSHIMA, F., KIMOTO, T., and YUASA, Y. (1988). Multi-step neoplastic transformation of normal human fibroblasts by Co-60 gamma rays and Ha-ras oncogene. *Mutation Res.*, **199**: 415-423.
39. PEREIRA-SMITH, O.M. and SMITH, J.R. (1988). Genetic analysis of indefinite division in human cells: Identification of four complementation groups. *Proc. Natl. Acad. Sci. USA*, **85**: 6042-6046.
40. RIMM, D.L., HORNESS, D., KUCERA, J., and BLATTNER, F.R. (1980). Construction of coliphage lambda Charon vector with BamH I cloning sites. *Gene*, **12**: 301-309.
41. SANGER, F., NICKLEN, S., and COULSON, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA*, **74**: 5463-5467.
42. SAVAGEAU, M.A., METTER, R., and BROCKMAN, W.W. (1983). Statistical significance of partial base-pairing potential: implications for recombination of SV40 DNA in eukaryotic cells. *Nucleic Acids Res.*, **11**: 6559-6570.
43. SOUTHERN, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.*, **98**: 503-517.
44. STRINGER, J.R. (1981). Integrated Simian Virus 40 DNA: Nucleotide sequences at cell-virus recombinant junction. *J. Virol.*, **38**: 671-679.
45. TODARO, G. and GREEN, H. (1963). Quantitative studies of the growth of mouse embryo cell in culture and their development into established lines. *J. Cell Biol.*, **17**: 299-313.
46. TOOZE, J. (1980). The molecular biology of tumor viruses, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
47. YANAGISAWA, K., SUENAGA, Y., NISHIO, K., and GOTOH, S. (1983). Establishment of a human diploid cell strain. *J. UOEH*, **5**: 49-54.
48. YANG, R.C.-A., LIS, J., and WU, R. (1979). Elution of DNA from agarose gels after electrophoresis. *Methods Enzymol.*, **68**: 176-182.
49. WEINTRAUB, H. (1983). A dominant role for DNA secondary structure in forming hypersensitive structures in chromatin. *Cell*, **32**: 1192-1203.
50. WILLIAMS, B.G. and BLATTNER, F.R. (1979). Construction and characterization of the hybrid bacteriophage lambda Charon vector for DNA cloning. *J. Virol.*, **29**: 555-575.

(Received for publication, October 4, 1990
and in revised form, November 26, 1990)