The mode of hepatitis B virus DNA integration in chromosomes of human hepatocellular carcinoma

T. Nagaya,[1] T. Nakamura, T. Tokino, T. Tsurimoto, M. Imai,[2] T. Mayumi,[2] K. Kamino,[3] K. Yamamura,[3] and K. Matsubara

Institute for Molecular and Cellular Biology, Osaka University, Suita, Osaka, 565 Japan; [2]Immunology Division, Jichi Medical School, Tochigi-Ken, 329-04 Japan; [3]The Fourth Department of Internal Medicine, Osaka University Medical School, Osaka, 530 Japan

Nineteen DNA samples that carry integrated hepatitis B virus (HBV) DNA were isolated from seven independent human hepatomas by molecular cloning, and their structures were determined. The results, combined with reported data, were analyzed so that one can obtain insights into the mechanisms of integration of this virus DNA and possible rearrangements that occur subsequently. The distribution of DNA junctions along the virus genome suggests that there are recombination-proficient regions. Thus, about half of the integrants were the Coh type, viz., one of their virus–cell DNA junctions fell within the so-called cohesive end region that lies between two 11-bp direct repeats (DR1 and DR2) in the virus genome where transcription and replication of the genome are initiated. All the integrated virus genomes were defective at least in one site around the cohesive end region, particularly within the X gene. The recombination-proficient regions are used not only for formation of virus–cell but also of virus–virus junctions. Neither virus nor cell DNA show unique sequences at the junctions, and targets for integration lie on many different chromosomes.

Hepatitis B virus (HBV) is a causative agent of human hepatitis that affects some 200 million people in the world. The viral genome, which consists of 3.2 kb DNA, often integrates into the chromosome of liver cells in the process of infection, resulting in continuous production of virus antigens. A significant number of chronic carriers later develop hepatocellular carcinoma (HCC), in which integrated HBV genomes are often detected. Studies on the process of integration and the subsequent behavior of the integrated genome would shed light on understanding the life cycle of this virus and the possible link between integration and cancer formation. However, such studies have been precluded because HBV replicates only in human and chimpanzee livers, not allowing use of model animals or cell cultures for detailed studies. The only informative way to approach these problems at present is to analyze the structure of integrated HBV DNA in HCCs and compare it with the known structure and function of the virus DNA.

Recent studies have shown that the HBV genome carries four coding frames, all of which are located on the same DNA strand. The genome also carries a unique region called the cohesive end region, which joins the linear, incomplete virus DNAs into a circle. This region is flanked by two 11-bp direct repeats called DR1 and DR2 (see Fig. 1). The whole genome is transcribed from DR1 into pregenome RNA that is used as mRNA as well as the template for minus-strand DNA synthesis. Other RNA syntheses start at sites in front of PreS and S genes, and all of them, inclusive of the pregenome RNA, terminate just right of DR1 in Figure 1, viz., outside of the cohesive end region. The minus-strand and plus-strand DNA syntheses start at DR1 and DR2, respectively, and proceed in opposite directions (Seeger et al. 1986). Thus, the cohesive end region carries most of the sites related to DNA and RNA syntheses. Some of the genes and sites are displayed schematically along the linearized virus genome in Figure 1.

Several groups have analyzed the integrated HBV DNA and proposed different integration mechanisms (Dejean et al. 1983, 1984; Koike et al. 1983; Koch et al. 1984; Shaul et al. 1984; Mizusawa et al. 1985; Yaginuma et al. 1985: Ziemer et al. 1985; Choo et al. 1986; Miller et al. 1985). For example, Dejean et al. (1984) saw in two cases that DR1 or DR2 in virus DNA is located at the virus–cell DNA junctions and proposed that DRs will play a role in integration. Yaginuma et al. (1985) observed that an integrated HBV DNA is flanked by re-

[1]Present address: Institute for Bioscience, Nippon Zeon Co., Kawasaki, 210 Japan.
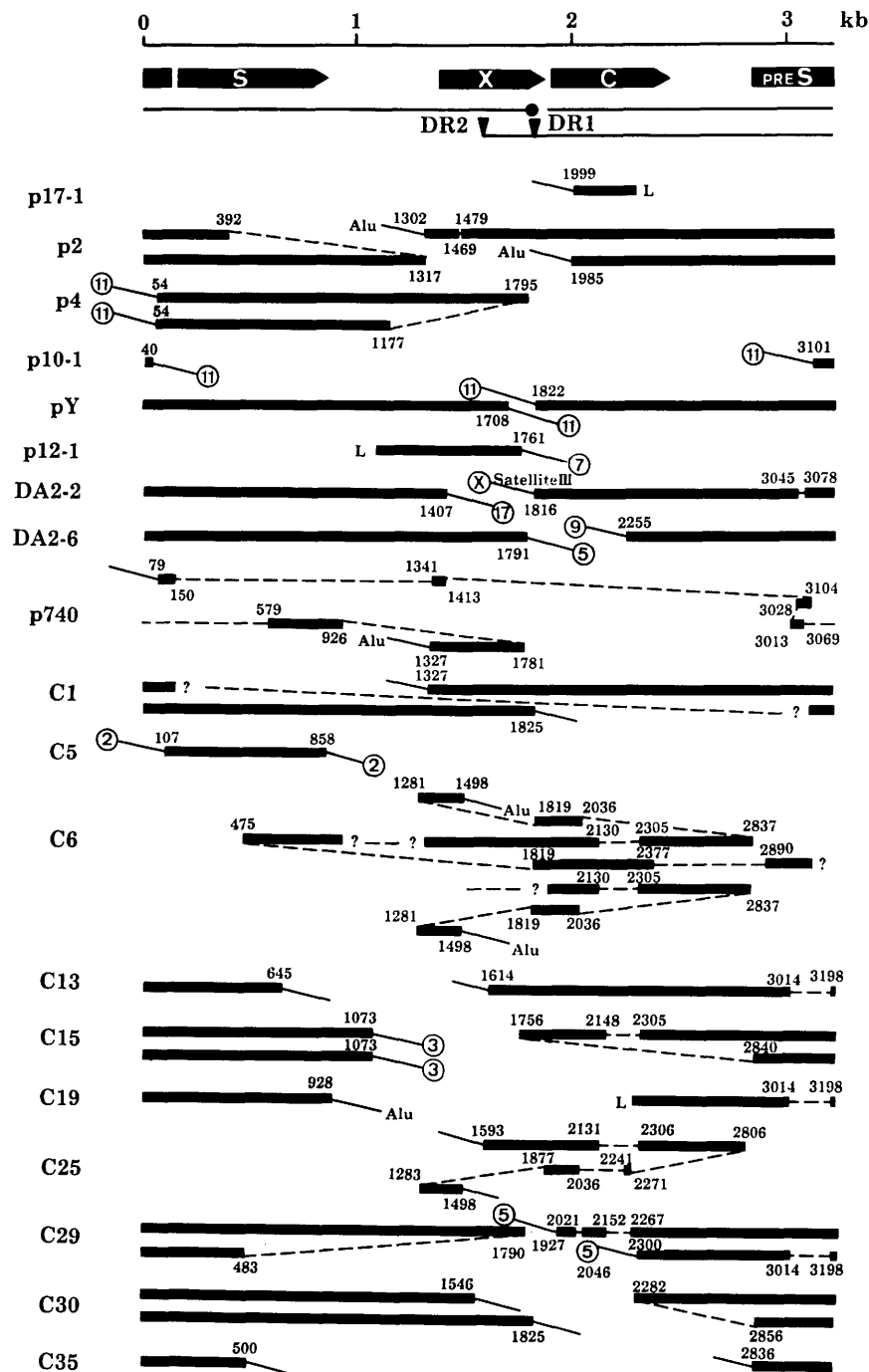
Figure 1. (See facing page for legend.)

peating sequences originating from the host target sequence and claimed that the integration may be carried out by a mechanism similar to the one acting in retrovirus DNA integration. Koch et al. (1984), on the other hand, detected a patch homology between the integrating virus DNA and cellular DNA sequences and proposed that such a homology may play some important role. The apparent discrepancy could reflect various ways of HBV DNA integration into cellular DNA. Alternatively, the virus DNA could integrate into cellular DNA in a unique fashion, but the diversity could be brought about by subsequent rearrangements. Since results from a small number of HCC samples could simply reflect a special reaction that does not necessarily represent the general phenomena, one must handle a reasonably large number of samples to obtain insights into the mechanism acting in the integration process or the subsequent reaction.

A fair amount of data on sequences of virus–host junctions and the structure of the integrated virus genome has been published recently by two groups, both of which examined HBV DNA in cell line PLC/PRF/5
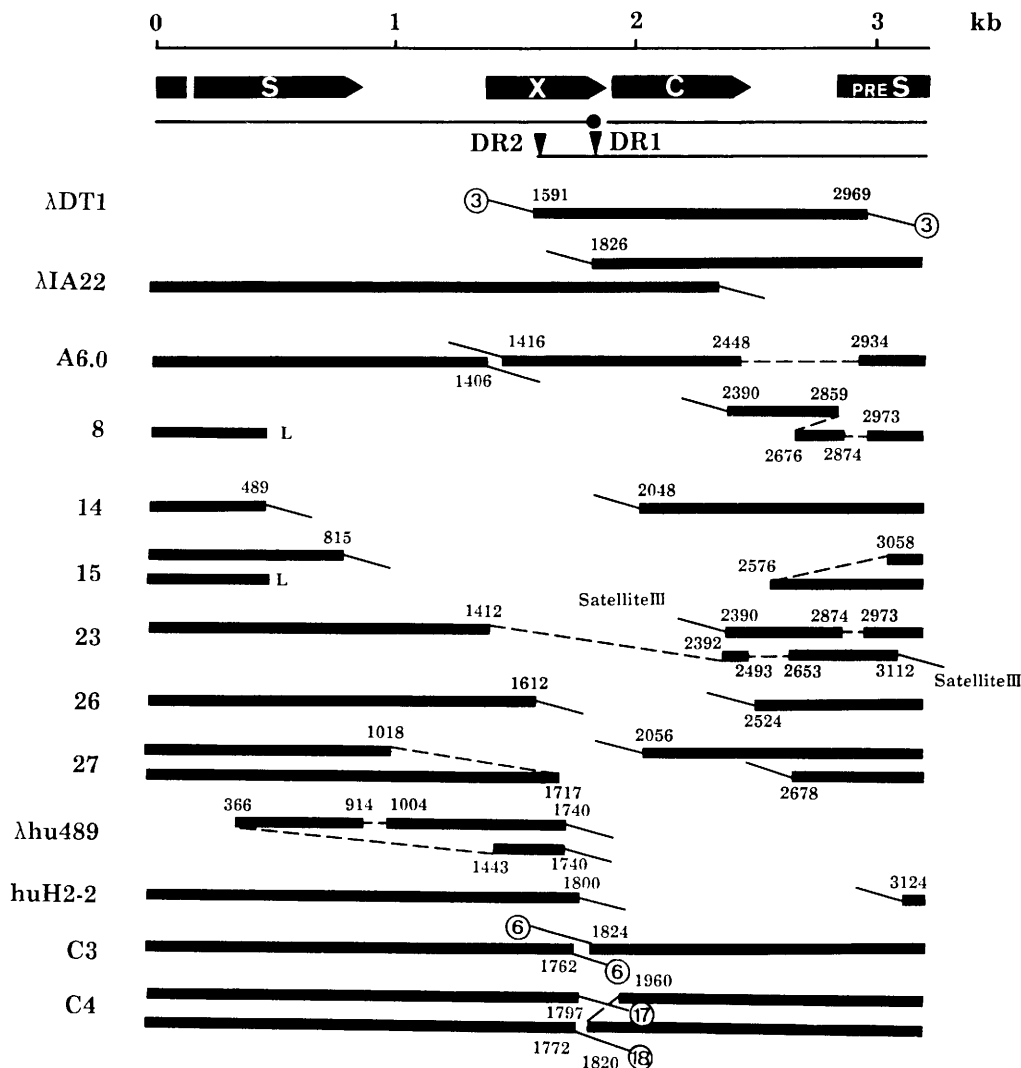
Figure 1. Structure of the integrated HBV sequences. At the top of the figure, the whole HBV genome is represented. The long thin line shows the minus strand with the DNA-linked protein (●), the short thin line represents the plus strand. The length of the HBV genome is indicated in kilobases, numbering from the hypothetical EcoRI site of sybtype adr (Ono et al. 1983; Fujiyama et al. 1983). Open reading frames and their direction of transcription are represented by filled arrows. (S) HBsAg (preS), pre-S region; (C) HBcAg; (X) X gene. The P gene is omitted for clarity. DR1, DR2 are the 11-bp direct repeats whose starting points are 1590 and 1824, respectively. The region in between DR1 and DR2 is the cohesive end region. HBV genome regions that are covered by the integrated virus DNA are shown by solid bars. The virus DNA that has one continuous portion(s) of a circular HBV genome is represented by a single solid bar aligned in the same level. Deletions in the genome are represented by dotted lines. The complex-type virus DNA that consists of two or more HBV genomes whose regions are overlapping, at least in part, is represented by multiple solid bars aligned in different levels, each representing different units of the virus genome, and connected by tilted dotted lines. Solid thin lines represent flanking human DNA. Encircled numerals show the number of the human chromosome in which the flanking DNA was assigned (see text). L indicates that the cloning vector is directly joined to the virus DNA. Alu and Satellite III indicate that the cell DNA at the junction is homologous to either Alu repeating sequence or Satellite III sequence, respectively.

(Ziemer et al. 1985; Koch et al. 1984). Altogether seven independent virus DNA integrants, which seem to represent almost all of the integrated HBV genomes in this cell line were analyzed. The results showed that all the integrated HBV DNAs are in fragmented or rearranged forms. No unique target sequences for insertion were found in either cellular or HBV DNA. However, here again, these samples were obtained from only one particular HCC cell line; therefore, one cannot eliminate a possibility that these results have biases that reflect a

special event associated with the integration or the subsequent rearrangements in the founder cell. Apparently, gathering more experimental data from as many independent HCC samples as possible is needed to extract the general feature of integration.

We prepared and analyzed 19 DNA samples that carry virus–host DNA junctions originating from seven independent human hepatomas. The results showed that in about half of the samples, the joining of virus to cell DNA occurred within the region between the two DRs

in the virus genome. The results also showed, as a general feature, that the integrated virus genomes would not serve as active templates for synthesis of virus pregenome RNA that acts as an mRNA as well as a template for the first DNA strand synthesis (Summers and Mason 1982; Seeger et al. 1986). Given these observations, the possible intermediates for integration are discussed.

Results and discussion

All the integrated virus genomes are defective

From seven independent HCCs, we obtained 19 λ clones that carry integrated virus DNA (Table 1). Sixteen of them had both "left" and "right" junctions of the integrated virus DNA, allowing analyses of their structures. Each cloned DNA corresponded to a band in Southern blotting profiles prepared from the original cell using the same enzyme.

We determined the structure of integrated virus genomes by constructing fine cleavage maps of the virus DNA and by sequencing or by hybridizing against subgenomic HBV DNA probes. Nineteen such structures, including unfinished ones (C6 and C1) that are too complex to analyze, are displayed in Figure 1. Because the major aim of this work is to obtain insights into general features of integrated HBV DNA, we also incorporated into the same figure 13 other reported cases (Dejean et al. 1984; Koch et al. 1984; Mizusawa et al. 1985; Yaginuma et al. 1985; Ziemer et al. 1985; Hino et al. 1986).

A remarkable feature of the integrant virus DNAs is that none of them carry an intact, contiguous, whole genome of the virus. This means that none of the integrant virus DNA in HCC can serve as a template for synthesis of pregenome RNA, which will be used as an mRNA or later as a template for virus DNA replication (Summers and Mason 1982; Tsurimoto et al. 1987). HBV

Table 1. *HCC tissues used and λ clones that carry the integrated HBV genome*

HCC tissue[a]	Serological markers[b]	λ Clones obtained
3	sAg, cAb	p17-1 (*Hind*III)[c]
7	sAg	p2, p4, p10-1 (*Eco*RI)
8	sAg	pY (*Hind*III)
11	sAg, eAb	p12-1 (*Hind*III)
21	sAg	DA2-2, DA2-6 (*Eco*RI)
22	sAg	p740 (*Eco*RI)
PLC342	sAg	C1, C5, C6, C13, C15, C19, C25, C29, C30, C35 (*Eco*RI)

[a] All the tissues, except for 22 and PLC342, are dissected HCCs. They were obtained from: (3) Department of Microbiology of Tohoku University; (7, 8, and 11) Fourth Department of Internal Medicine, Osaka University Medical School; (21) Department of Pathology, University of Tokyo (22 and PLC342) originate from dissected HCCs, propagated in nude mice (Matsui et al. 1986).
[b] Only checked markers are listed.
[c] Enzymes used for cloning are shown in parentheses.

DNA replication in HCC, if it occurs, may do so by using a free, unintegrated virus genome as a template. One exception is λIA22 (Dejean et al. 1984), which seems to carry a contiguous genome more than one unit in size. Its structure, however, has not been critically examined. It is likely that λIA22 consists of more than two genome fragments whose junction(s) has been overlooked.

A majority of the integrated HBV genomes are deleted in or around the cohesive end region, and very few of them carry both intact DR1 and DR2. For the same reason, gene X, which spans the cohesive end region, is defective in almost all of the cases. An enhancer sequence is located at about 200 bp upstream from the X gene (Shaul et al. 1985). The deletion of the X gene region or cohesive end region upon integration may sometimes bring a cellular gene close to the enhancer; in such cases, the integrated virus genome will act as a portable enhancer for the cellular gene.

Coh integrants that carry virus DNA integrated within or around the cohesive end region

To examine more closely the sites or sequences in the virus DNA used for integration, we extracted subfragments that carry virus–cell DNA junctions from the λ clones. The results of sequencing studies with 32 junctions are shown in Figure 2, with the host sequence on the left side and the virus sequence on the right side.

As previously reported (Ziemer et al. 1985), the sequences at the switch points are nonunique. However, five clones of the 19 samples (pY, DA2-2, C1, C25, C30) have one of the switch points very close to DR1 or DR2 sequences (shown by boxes in Fig. 2). Three more clones (p12-1, DA2-6, C13) have the switch points within the virus cohesive end region, which lies in between DR1 and DR2 (Fig. 3; see below). Thus, out of 19 clones, eight have at least one of the virus DNAs ending at the DR or within the cohesive end region. For the sake of simplicity, we will call them Coh integrants.

In addition to the data above, seven clones among 13 reported samples (λDT1, λIA22, 26, λhu489, huH2-2, C3, C4) also are Coh integrants. Combining these results, Coh integrants amount to 15 out of 32 clones analyzed. These results demonstrate that it is the cohesive end region of the virus genome, rather than the DRs (Dejean et al. 1984), that predominantly appears at the switch point. The predominant appearance of Coh integrants is likely to reflect the mechanism of integration.

Figure 3 displays the virus–cell junctions represented by arrowheads pointing from the cell DNA toward the integrated virus DNA. This figure also shows the recombination-proficient regions along the virus genome, as defined by having at least four DNA junctions in Figure 1 within a stretch of less than 24 bp. These regions are marked by Roman numerals. Regions II–IV lie within the cohesive end region. Regions II and IV coincide with DR2 and DR1, respectively.

There are 18 virus–cell junctions in the cohesive end region carried by the 15 Coh integrants described above. Of these junctions, 13 are clustered around DR1, nine leftward and four rightward. Five junctions are clustered

around DR2, two leftward and three rightward. In total, half of the virus–cell junctions that lie in this region are concentrated around DR1 with leftward direction.

These observations strongly suggest, though do not prove, that intermediates of replication or transcription act as "substrates for integration"; such intermediates will provide the single-stranded DNA at or around the cohesive region that is used for strand invasion into cell DNA. If minus-strand DNA invaded, then virus DNA that lies to the left of the junction in Figure 3 should appear. The nine junctions accumulating around DR1, could fit this model. If plus-strand DNA invaded, then rightward junctions in this region should appear. There are several rightward junctions around DR2. However, invasion of minus-strand DNA looks much more active than that of plus-strand DNA.

One could argue that circular viral DNA or tandemly arranged genomes could also act as intermediates for in-

tegration; such integration must involve preferential "activation" of the cohesive end region, in analogy with integration of retrovirus genomes (Varmus and Swanstrom 1985). However, there is no available evidence at the present time that supports the presence of such an integration system. Neither is there any evidence to show the appearance of tandem structures.

One could also argue that the virus DNA might have integrated randomly along its genome, but random integration was followed by preferential loss of DNA in the cohesive end region. However, this model requires another assumption to explain why and how such preferential loss takes place. In addition, the observation by Tsurimoto et al. (1987) that three tandemly arranged, complete HBV genomes can be perpetuated in integrated form in a hepatoblastoma cell line presents a view opposed to the obligatory deletion of the viral DNA integrants.

Clone	Cell (5')	Virus (3')
p17-1	TGGTTATGTCTATTGCCTCA	CGCCTCTGCTCTGTATCGGG 1999 →
p2	ACAAAATTAAGCCGGGCGCG Alu	GATCTCCTCGACACCGCCTC 1985 →
	CCCTCCCGAGTAGCTGGGAC Alu	CAGCCGGTCTGGAGCGAAAC 1302 →
p4	AGCCACACCGTGCTCCTGGG	TTCCTGCTGGTGGCTCCAGT 54 →
p10-1	TTTTTCCTACAGAAACCTCC	CACTCTGGGATCTAACAGAG 40 ←
	GCCTTTGCCCATTGACAGTC	TGACAACATCGCCAGCTGCG 3101 →
pY	TGAGGTTTATGTAGTTCCCC	T|TTCACCTCTGC|TAATCA 1822 →
	TGGAATAAACTACAAATCAT	AAGTATGCCTXXAGGTCGGT 1708 ←
p12-1	CCCCTCTAGGGTAATGAAGG	TTAACCTAATCTCCTCCCCC 1761 ←
DA2-2	GAATAGAATAGAATAGAATA SateliteIII	GCAACTTT|TTCACCTCTGC|C 1816 →
	GAAAATAACCCATTTTCTTT	CAGGATCCAGTTGGCAGCAC 1407 ←
DA2-6	AAATCCTCATTTGTGTTCTC	TTGGTGTCTTTTGGAGTGTG 2255 →
	CAACAAGATAGGAATGATGG	TATGCCTACAGCCTCCTAGT 1791 ←
p740	TCCACTGTGGCCCTGAGTGA	AACAGTAAACCCTGTTCCGA 79 →
	AAAAAAACAAAAAAAACTCC Alu	GGGACTGACAACTCTGTTGT 1327 →
C1	CCAGGACCTTTTTCTCCAAA	GGGACTGACAACTCTGTTGT 1327 →
	AGAGCCATCGACAACAGTGC	AAAAAGTTGCATGGTGCTGG 1825 ←

Clone	Cell (5')	Virus (3')
C5	TTTAAAGTATGGGCCTGTAT	TCACCCATAACGTCAAACTT 107 →
	GAAGTGGCAATGGAGCGAGA	CCATCTTTTGGTTTTATTAG 858 ←
C6	CTCAAGTAGATGGGACTACA Alu	GAAGAAGGGGACGGTAAACC 1498 ←
C13	TTTCCTTTAAAAAAAACTAT	ACCACCGTGAACGCCCACCA 1614 →
	CTGGATTGGTTACATCACAG	TCCCATAGGAATCTTGCGAA 645 ←
C15	TCCTAAAAGTGTGATTTCTC	ATACATGCATATAAAGGCAT 1073 ←
C19	TTTTTTTTTGAGAACGGAGT Alu	GTACAATGTGTTCTTGTGGT 928 ←
C25	TCTCTCAATTTTAATTAAAA	ACCTCTGCACGTCGCATGGA 1593 →
	AAGTCAGATCGGGACTCACA	GAAGAAGGGGACGGTAGAGC 1498 ←
C29	TGGCCTCCCAAAATGCTGGG	GAGCCACCAAATGCCCCTAT 2298 →
	CGGGGCTTGGTGTCACGCGC	TGGAGCTTCTGTGGAGTTAC 1927 →
C30	XAGCCATCGACAACAGTGCA	AAAAAGCCGCATGGTGCTGG 1825 ←
	TATCTTCTACTGTCTGTCGG	AGACCGCGTAAAGAGAGGTG 1546 →
C35	GAAAGCCTGTTTAGATTGTA	TAGTTGAGGTTCCTGGAAGT 500 ←
	TGGTATCTGCTGTAATGTCT	AAGAGCTACAGCATAGGAGG 2836 →

Figure 2. Nucleotide sequences of the human–virus DNA junctions. Human sequences are shown on the left side and virus sequences on the right side. The nucleotide numbers of the virus genome appearing at the junctions are presented under the sequence. The arrowheads show the direction of the viral genome, pointing toward increasing nucleotide numbers. The nucleotide sequences that appear in the DRs of the HBV genome are shown by boxes. The virus sequences that are located in the cohesive region are underlined. *Alu* and Satellite III represent those human sequences that are homologus to *Alu* repeating sequence (Houck et al. 1979) or Satellite III sequence (Deininger et al. 1981), respectively.
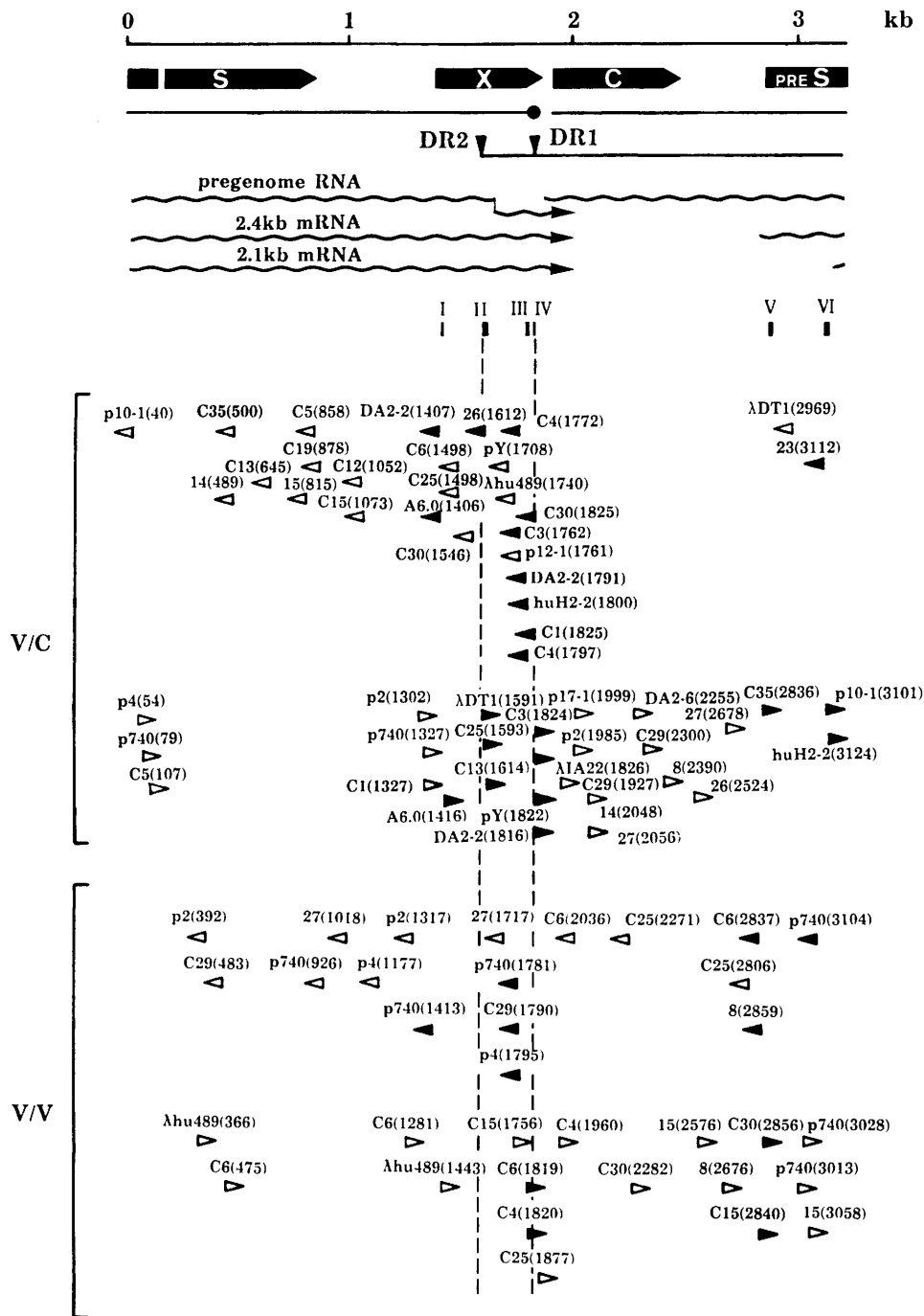
Figure 3. Distribution of virus–cell (V/C) and virus–virus (V/V) recombination sites along HBV genome. The virus–virus junctions that can be accounted for by deletions in the genome are not displayed. The minus and plus strands of the HBV genome, coding regions, and direction of transcription and nucleotide numbering are indicated as in Fig. 1. The three wavy lines represent the HBV mRNAs. Each arrowhead represents the position of the junction and the direction pointing toward the HBV DNA. For example, rightward arrowheads show that virus DNA lies to the right of the junction. Leftward arrowheads show the reverse situation. The borders of the cohesive end region that lie in between DR1 and DR2 are indicated by dotted vertical lines. Recombination-proficient regions, which are defined by having at least four DNA junctions within a stretch of less than 24 bp, are represented by filled boxes and marked by Roman numerals: Region I, 1406–1416; II, 1593–1614; III, 1780–1800; IV, 1816–1825; V, 2836–2856; VI, 3101–3124. Junctions located in such recombination-proficient regions are represented by filled arrowheads.

We were not able to find any unique features for the other end of the integrated virus genome. Therefore, at this stage we assume that it is formed by random recombination with cell DNA.

Other recombination-proficient regions along the HBV genome

Upon a closer look at Figure 3, the distribution of the virus–cell DNA junctions suggests that the HBV

genome has some recombination-proficient regions other than the cohesive end region. Examples are the regions V and VI that lie in front of the pre-S gene and S gene, respectively. Ziemer et al. (1985) have made a similar observation. These sites correspond to the initiation sites for 2.4-kb and 2.1-kb mRNAs (Rall et al. 1983; Standring et al. 1983; Cattaneo et al. 1984).

Altogether, these observations favor the idea that there is some correlation between RNA production and integration or recombination, suggesting intervention of reverse transcriptase. More data are needed to clarify these problems.

The virus–virus DNA junctions

Interestingly, the cohesive end region of the virus genome also seems to be preferred for the formation of virus–virus junctions. In Figure 3, seven junctions are located in or around the cohesive end region, and of the 13 integrants previously reported, two have the same feature. This suggests that such virus–virus DNA junctions have been formed by a mechanism similar to that acting in production of the Coh integrants. Such questions as when one virus genome integrates into another and why the virus–virus junctions appear so frequently await further elucidations.

Temporary augmentation of integration and recombination reactions

Tsurimoto et al. (1987) constructed an HCC cell line that produces Dane-like particles and demonstrated that almost all the newly produced virus DNA molecules are in encapsidated form, suggesting that such molecules will not see cell DNA. In patient liver, on the other hand, very active virus replication is observed sometimes, and a large amount of virus DNA is detected in "free" form. It may be only under such circumstances that the "substrates for integration" are available. One must doubt if such a condition will last for many cell generations.

Southern blotting analyses done with many HCCs in many laboratories have shown that most of the tumor cells are monoclonal and that such cells carry multiple integrants. Cases of single integrants have been reported, but they are not common. Although multiple, the integrated HBV DNAs in HCCs are stable, as shown by monoclonality of their Southern blotting profiles.

These observations strongly suggest that integrations took place for a relatively limited period of time under limited conditions (augmented period for integration) in the infected cells; such cells must have survived while allowing active recombination and/or accumulation of active "substrates," and their activity for integration terminated by changes in the rate of virus growth or cellular physiological conditions. It may be that such cells were rendered refractory to virus replication or became unable to keep producing free virus DNA "substrates" or to supply recombination systems.

Cell DNA used as a target for virus integration

Yaginuma et al. (1985) found a case where direct repeats,

apparently derived from the target cell DNA sequence, flank the integrated HBV DNA. Given this observation, they proposed that HBV DNA may integrate via a mechanism analogous to that acting in the integration of retrovirus DNA or in the formation of a pseudogene. Among the 27 cases that carry both "left" and "right" junctions, theirs is the sole case that has the flanking direct repeats. Therefore, it seems that their proposed model is a rather remote one for the HBV DNA integration in general. Koch et al. (1984) saw a small-patch homology between the virus and cell DNA sequences outside the switch point. Such short homologous sequences can be found with almost any combination of two DNA sequences; therefore, their significance is not clear.

No unique cellular site(s) for HBV DNA integration has been found so far (see Fig. 2), except that *Alu* repeating sequences (Houck et al. 1979) were used as the targets for integration in four cases (p2, p740, C6, C19). In one case (DA2-2), the target was human satellite III–related DNA (Deininger et al. 1981), which was also found at the junction in clone 23 (A123) (Shaul et al. 1986). Except for the weak preference shown for repeating sequences, the HBV DNA seems to select no unique cell DNA sequence as the target for integration.

The present data do not, however, eliminate a possibility that a unique cell sequence(s) is selected in the integration process, but it does not lie right at the site of virus–cell DNA switch point.

Chromosome assignment of the integrated site

DNA fragments that carry a nonrepetitious, unique sequence were prepared as hybridization probes from the flanking cellular DNAs in each integrant (see Table 2).

Table 2. *Assignment of the flanking cellular DNA probe to a group of chromosomes*

Probe	Restriction site and size of the probe[a]		Assigned group of chromosomes
p4	EX	0.3 kb	9, 10, 11, 12
p10-1L	EHc	0.4 kb	9, 10, 11, 12
p10-1R	XbHc	0.3 kb	9, 10, 11, 12
pYL	HG	0.3 kb	9, 10, 11, 12
pYR	BG	0.6 kb	9, 10, 11, 12
p12-1	HHc	0.3 kb	6, 7, 8, X
DA2-2L	*Hae*III	0.7 kb	6, 7, 8, X
DA2-2R	*Hae*III	0.55 kb	16, 17, 18
DA2-6L	EHc	0.2 kb	9, 10, 11, 12
DA2-6R	*Hae*III	0.4 kb	3, 4, 5
C5L	EX	1.3 kb	1, 2
C5R	*Hpa*II	0.2 kb	1, 2
C15	EX	0.8 kb	3, 4, 5
C29L	*Bgl*II	1.2 kb	3, 4, 5
C29R	*Hae*III	0.3 kb	3, 4, 5

Probes representing either side of the integrating HBV DNA are discriminated by the symbol L or R. Thus, for example, two probes from "left" and "right" flanking cellular DNAs of p10-1 are designated p10-1L and p10-1R.

[a] Restriction enzyme abbreviations are shown as follows: (B) *Bam*HI; (E) *Eco*RI; (G) *Bgl*II; (H) *Hind*III; (Hc) *Hinc*II; (X) *Xho*I; (Xb) *Xba*I. For example, the probe prepared from p4 is 0.3 kb in size; one end is cut with *Eco*RI and the other end with *Xho*I.

Condensed chromosomes in a karyotypically normal human lymphoblastoid cell line were prepared and sorted into eight fractions, which carry chromosomes 1, 2; 3, 4, 5; 6, 7, 8, and X; 9, 10, 11, 12; 13, 14, 15; 16, 17, 18; 19, 20; 21, 22, using fluorescence-activated cell sorter. DNA in each fraction was subjected to Southern analysis using the probe described above. This allowed us to assign the probe DNAs to a group of chromosomes, as shown in Table 2. To assign each probe to a unique chromosome, a small set of human–mouse hybrid cell panels was employed for discordant assay. Since we have the results of Table 2, intensive discordant assays were unnecessary. The results are shown in Table 3 and are collectively displayed in Figure 1 (assigned chromosome numbers are represented by encircled numerals). The flanking cellular sequences from both sides were assigned to the same chromosome in the majority of the samples, but in DA2-2 and DA2-6, they were assigned to different chromosomes, showing that there must have been rearrangements of chromosomes.

Chromosomal assignments of HBV DNA integration sites have been done by three groups. Seven integrants have thus been studied and were assigned to chromosomes 3 (Dejean et al. 1986), 6 (Hino et al. 1986), 11p13-p14 (Rogler et al. 1985), 11q22, 15q22-q23, 18q12 (Bowcock et al. 1985), 17, and 18 (Hino et al. 1986). Altogether, these results indicate that targets for HBV integration lie on at least 11 different chromosomes and suggest that HBV DNA does not show a strong preference for a particular chromosome for integration. However, three of our clones and two reported cases are assigned to chromosome 11, a figure that is higher than that expected for random distribution. Further work, including the assignment of locations in individual chromosomes, is needed in this respect.

Conclusion

Our studies have shown the importance of handling a reasonably large number of samples to infer the process of integration when analyses of the integrants are the only approach to study the problem. There have been several conflicting models based on analyses of only a limited number of samples, but now most of them have been eliminated. To penetrate more deeply into the problem, however, a culture cell system that allows virus DNA integration is awaited.

Materials and methods

DNA

Table 1 lists the HCC tissues or HCC-originating cells propagated in nude mice used in this work. DNA was extracted from the HCC tissues or cells using the procedure developed by Fasano (Fasano et al. 1984). Each DNA, subjected to Southern hybridization (Southern 1975), showed multiple integration profiles (data not shown).

Cloning

The HCC DNA was digested with *Eco*RI or *Hind*III, which does not cleave within typical genomes of HBV subtype adr (Fu-

Table 3. *Discordant chromosome assignments*

Hybrid cell	Chromosome 1	2	Probe C5L	C5R
TA1A	+	−	−	−
II5	−	+	+	+
Conclusions	Chromosome		2	2

Hybrid cell	Chromosome 3	4	5	Probe C15	C29L,R	DA2-6R
HM31	−	+	−	−	−	−
TA3A	+	−	−	+	−	−
III'27910	−	−	+	−	+	+
III'2	−	+	+	−	+	+
Conclusions	Chromosome			3	5	5

Hybrid cell	Chromosome 6	7	8	X	Probe p12-1	DA2-2L
HM22	+	−	−	−	−	−
HM23	+	−	−	−	−	−
TA1A	−	+	−	−	+	−
Conclusions	Chromosome				7	8 or X[a]

Hybrid cell	Chromosome 9	10	11	12	Probe p4	p10-1L,R	pYL,R	DA2-6L
HM06	−	+	+	−	+	+	+	−
TA1A	−	+	−	+	−	−	−	−
p23-7-10-13	−	−	+	−	+	+	+	−
Conclusions	Chromosome		11		11		11	9

Hybrid cell	Chromosome 16	17	18	Probe DA2-2R
HM23	−	−	+	−
HM06	−	+	+	+
HM09	−	+	−	+
TA1A	+	−	+	−
Conclusion	Chromosome			17

Mouse–human hybrid cells were obtained from Dr. Y. Kaneda (Osaka University), except for TA1A and TA3A, which were provided by Dr. N. Shimizu (Keio University). Only the relevant chromosome karyotypes are shown. DNA from each hybrid cell was subjected to Southern blot analyses using the probes as shown. (+) and (−) indicate the presence and absence, respectively, of positive bands in the blots.
[a] Later assigned to X by using a cell line that carries reciprocally translocated chromosomes (T. Tokino et al., in prep).

jiyama et al. 1983; Ono et al. 1983). The digested DNA fragments were cloned into the *Eco*RI site of λ Charon 28 (Blattner et al. 1978) or the *Hind*III site of λ Styx 28, a derivative of Charon 28 whose *Bam*HI site was converted into a *Hind*III site using a synthetic DNA linker (this work). Clones carrying the integrated HBV genome were selected by using full length HBV genome DNA, that has been separated from pBR322 DNA by preparative agarose gel electrophoresis (Fujiyama et al. 1983) as a probe.

Determination of the structure of integrated HBV genomes and the virus–host junctions

The clones carrying human DNA fragments with integrated HBV genomes were purified, and the inserts were recloned into pBR322. Fragments were subjected to restriction enzyme map-

ping and Southern hybridization analysis using HBV DNA probes, and those that carry the virus–cell DNA junctions were subjected to sequencing by the dideoxy method (Sanger et al. 1977).

Chromosome assignment of the integrated site

Chromosomal assignments were done as according to the procedures developed by Fukushige et al. (1986). A human lymphoblastoid cell line, GM0131, carrying a normal karyotype was obtained from the Human Genetic Mutant Cell Repository (Camden, NJ). Condensed chromosome suspensions from these cells were prepared according to the polyamine methods of Sillar and Young (1981) and sorted into eight fractions by using a fluorescence-activated cell sorter CHROSS-1 (Japan Spectroscopic Co., Ltd.), and the DNA from each sorted chromosome (5×10^5 chromosomes) was digested by a restriction endonuclease, followed by electrophoresis and Southern blottings. The blots were hybridized to unique cellular DNA probes as prepared from the cellular DNAs flanking the integrated HBV DNAs. To prepare the probes, the cellular DNA in each integrant clone was extracted, fragmented by several restriction enzymes, and then selected for those that do not give smears when used as probes in Southern hybridization of total human genome DNA. Next, each probe was assigned to a unique chromosome by discordant assay using a small set of human–mouse hybrid cell panels (Results and discussion). Southern hybridizations were done under stringent conditions with ^{32}P-labeled DNA probes using a Multiprime DNA labeling system (Amersham).

Acknowledgments

References

Blattner, F.R., A.E. Blechel, K. Denniston-Thompson, H.E. Faber, J.E. Richard, J.L. Slightom, P.W. Tucker, and O. Smithies. 1978. Cloning human fetal γ globin and mouse α-type globin DNA: Preparation and screening of shotgun collections. *Science* **202:** 1279–1284.

Bowcock, A.M., M.R. Pinto, E. Bey, J.M. Kuyl, G.M. Dusheiko, and R. Bernstein. 1985. The PLC/PRF/5 human hepatoma cell line. II. Chromosomal assignment of hepatitis B virus integration sites. *Cancer Genet. Cytogenet.* **18:** 19–26.

Cattaneo, R., H. Will, and H. Schaller. 1984. Hepatitis B virus transcription in the infected liver. *EMBO J.* **3:** 2191–2196.

Choo, K.B., M.S. Liu, P.C. Chang, S.M. Wu, M.W. Su, C.C. Pan, and S.H. Han. 1986. Analysis of six distinct integrated hepatitis B virus sequences cloned from the cellular DNA of a human hepatocellular carcinoma. *Virology* **154:** 405–408.

Deininger, P.L., D.J. Jolly, C.M. Rubin, T. Friedman, and C.W. Schmidt. 1981. Base sequence studies of 300 nucleotide renatured repeated human DNA clones. *J. Mol. Biol.* **151:** 17–33.

Dejean, A., C. Brechot, P. Tiollais, and S. Wain-Hobson. 1983.

Characterization of integrated hepatitis B viral DNA cloned from a human hepatoma and the hepatoma-derived cell line PLC/PRF/5. *Proc. Natl. Acad. Sci.* **80:** 2505–2509.

Dejean, A., P. Sonigo, S. Wain-Hobson, and P. Tiollais. 1984. Specific hepatitis B virus integration in hepatocellular carcinoma DNA through a viral 11-base-pair direct repeat. *Proc. Natl. Acad. Sci.* **81:** 5350–5354.

Dejean, A., L. Bougueleret, K.M. Grzeschik, and P. Tiollais. 1986. Hepatitis B virus DNA integration in a sequence homologous to v-*erbA* and steroid receptor genes in a hepatocellular carcinoma. *Nature* **322:** 70–72.

Fasano, O., D. Birnbaum, L. Edlund, J. Fogh, and M. Wigler. 1984. New human transforming genes detected by a tumorigenicity assay. *Mol. Cell. Biol.* **4:** 1695–1705.

Fujiyama, A., A. Miyanohara, C. Nozaki, T. Yoneyama, N. Ohtomo and K. Matsubara. 1983. Cloning and structural analysis of hepatitis B virus DNAs, subtype adr. *Nucleic Acids Res.* **11:** 4601–4610.

Fukushige, S., T. Murotsu, and K. Matsubara. 1986. Chromosomal assignment of human genes for gastrin, thyrotropin (TSH)-β subunit and c-erbB-2 by chromosome sorting combined with velocity sedimentation and Southern hybridization. *Biochem. Biophys. Res. Commun.* **134:** 477–483.

Hino, O., T.B. Shows, and C.E. Rogler. 1986. Hepatitis B virus integration site in hepatocellular carcinoma at chromosome 17;18 translocation. *Proc. Natl. Acad. Sci.* **83:** 8338–8342.

Houck, C.M., F.P. Rinehart, and C.W. Schmid. 1979. An ubiquitous family of repeated DNA sequences in the human genome. *J. Mol. Biol.* **132:** 289–306.

Koch, S., A. Freytag von Loringhoven, R. Kahmann, P.H. Hofschneider, and R. Koshy. 1984. The genetic organization of integrated hepatitis B virus DNA in the human hepatoma cell line PLC/PRF/5. *Nucleic Acids Res.* **12:** 6871–6886.

Koike, K., M. Kobayashi, H. Mizusawa, E. Yoshida, K. Yaginuma, and M. Taira. 1983. Rearrangement of the surface antigen gene of hepatitis B virus integrated in the human hepatoma cell lines. *Nucleic Acids Res.* **11:** 5391–5402.

Matsui, T., M. Takano, K. Miyamoto, Y. Itoh, H. Yoshizawa, M. Koike, T. Mochizuki, E. Tanaka, H. Okamoto, M. Imai, S. Mishiro, Y. Miyakawa, and M. Mayumi. 1986. Nude mice bearing human primary hepatocellular carcinoma that produces hepatitis B surface, core and e antigens, as well as deoxyribonucleic acid polymerase. *Gastroenterology* **90:** 135–142.

Miller, R.H., S.C. Lee, Y.F. Liaw, and W.S. Robinson. 1985. Hepatitis B virus DNA in infected human liver and in hepatocellular carcinoma. *J. Infect. Dis.* **151:** 1081–1092.

Mizusawa, H., M. Taira, K. Yaginuma, M. Kobayashi, E. Yoshida, and K. Koike. 1985. Inversely repeating integrated hepatitis B virus DNA and cellular flanking sequences in the human hepatoma-derived cell line huSP. *Proc. Natl. Acad. Sci.* **82:** 208–212.

Ono, Y., H. Onda, R. Sasada, K. Igarashi, Y. Sugino, and K. Nishioka. 1983. The complete nucleotide sequences of the cloned hepatitis B virus DNA: Subtype *adr* and *adw*. *Nucleic Acids Res.* **11:** 1747–1757.

Rall, L.B., D.N. Standring, O. Laub, and W.J. Rutter. 1983. Transcription of hepatitis B virus by RNA polymerase II. *Mol. Cell. Biol.* **3:** 1766–1773.

Rogler, C.E., M. Sherman, C.Y. Su, D.A. Shafritz, J. Summers, T.B. Shows, A. Henderson, and M. Kew. 1985. Deletion in chromosome 11p associated with a hepatitis B integration site in hepatocellular carcinoma. *Science* **230:** 319–322.

Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chainterminating inhibitors. *Proc. Natl. Acad. Sci.* **74:** 5463–5467.

Seeger, C., D. Ganem, and H.E. Varmus. 1986. Biochemical and

genetic evidence for the hepatitis B virus replication strategy. *Science* **232:** 477–484.

Shaul, Y., M. Ziemer, P.D. Garcia, R. Crawford, H. Hsu, P. Valenzuela, and W.J. Rutter. 1984. Cloning and analysis of integrated hepatitis virus sequences from a human hepatoma cell line. *J. Virol.* **51:** 776–787.

Shaul, Y., W.J. Rutter, and O. Laub. 1985. A human hepatitis B viral enhancer element. *EMBO J.* **4:** 427–430.

Shaul, Y., P.D. Garcia, S. Schonberg, and W.J. Rutter. 1986. Integration of hepatitis B virus DNA in chromosome-specific satellite sequences. *J. Virol.* **59:** 731–734.

Siller, R., and B.D. Young. 1981. A new method for the preparation of metaphase chromosomes for flow analysis. *J. Histochem. Cytochem.* **29:** 74–78.

Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98:** 503–517.

Standring, D.N., L.B. Rall, O. Laub, and W.J. Rutter. 1983. Hepatitis B virus encodes an RNA polymerase III transcript. *Mol. Cell. Biol.* **3:** 1774–1782.

Summers, J. and W.S. Mason. 1982. Replication of the genome of a hepatitis B-like virus by reverse transcription of an RNA intermediate. *Cell* **29:** 403–415.

Tsurimoto, T., A. Fujiyama, and K. Matsubara. 1987. Stable expression and replication of hepatitis B virus genome in an integrated state in a human hepatoma cell line transfected with the cloned viral DNA. *Proc. Natl. Acad. Sci.* **84:** 444–448.

Varmus, H. and R. Swanstrom. 1985. Replication of retroviruses. In *RNA tumor viruses* (ed. R. Weiss et al.), pp. 75–134. Cold Spring Harbor Laboratory, New York.

Yaginuma, K., M. Kobayashi, E. Yoshida, and K. Koike. 1985. Hepatitis B virus integration in hepatocellular carcinoma DNA: Duplication of cellular flanking sequences at the integration site. *Proc. Natl. Acad. Sci.* **82:** 4458–4462.

Ziemer, M., P. Garcia, Y. Shaul, and W.J. Rutter. 1985. Sequence of hepatitis B virus DNA incorporated into the genome of a human hepatoma cell line. *J. Virol.* **53:** 885–892.

The mode of hepatitis B virus DNA integration in chromosomes of human hepatocellular carcinoma.

T Nagaya, T Nakamura, T Tokino, et al.

References	This article cites 33 articles, 16 of which can be accessed free at: **http://genesdev.cshlp.org/content/1/8/773.full.html#ref-list-1**
License	
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**