Clonal Expansion of Normal-Appearing Human Hepatocytes during Chronic Hepatitis B Virus Infection[†]

William S. Mason,^{1*} Chen Liu,² Carol E. Aldrich,¹ Samuel Litwin,¹ and Matthew M. Yeh³

Fox Chase Cancer Center, 333 Cottman Avenue, Philadelphia, Pennsylvania 19111¹; Department of Pathology, Immunology, and Laboratory Medicine, University of Florida, Gainesville, Florida 32610²; and Department of Pathology, University of Washington School of Medicine, Seattle, Washington 98195³

Received 19 April 2010/Accepted 24 May 2010

Chronic hepatitis B virus (HBV) infections are associated with persistent immune killing of infected hepatocytes. Hepatocytes constitute a largely self-renewing population. Thus, immune killing may exert selective pressure on the population, leading it to evolve in order to survive. A gradual course of hepatocyte evolution toward an HBV-resistant state is suggested by the substantial decline in the fraction of infected hepatocytes that occurs during the course of chronic infections. Consistent with hepatocyte evolution, clones of >1,000 hepatocytes develop postinfection in the noncirrhotic livers of chimpanzees chronically infected with HBV and of woodchucks infected with woodchuck hepatitis virus (W. S. Mason, A. R. Jilbert, and J. Summers, *Proc. Natl. Acad. Sci. U. S. A.* 102:1139–1144, 2005; W. S. Mason et al., *J. Virol.* 83:8396–8408, 2009). The present study was carried out to determine (i) if extensive clonal expansion of hepatocytes also occurred in human HBV carriers, particularly in the noncirrhotic liver, and (ii) if clonal expansion included normal-appearing hepatocytes, not just hepatocytes that appear premalignant. Host DNA extracted from fragments of noncancerous liver, collected during surgical resection of hepatocellular carcinoma (HCC), was analyzed by inverse PCR for randomly integrated HBV DNA as a marker of expanding hepatocyte lineages. This analysis detected extensive clonal expansion of hepatocytes, as previously found in chronically infected chimpanzees and woodchucks. Tissue sections were stained with hematoxylin and eosin (H&E), and DNA was extracted from the adjacent section for inverse PCR to detect integrated HBV DNA. This analysis revealed that clonal expansion can occur among normal-appearing human hepatocytes.

Transient hepatitis B virus (HBV) infections, which generally last <6 months, do not cause cirrhosis and cause only minor increases in the risk of hepatocellular carcinoma (HCC) (3, 46). Chronic infections, typically lifelong, can cause cirrhosis and HCC (3). Of the ~350 million HBV carriers now alive, ca. 60 million will die prematurely of cirrhosis and/or HCC. Cirrhosis, which usually develops late in infection, is a significant risk factor for HCC. Early reports stated that most HCCs occur on a background of cirrhosis. However, later studies suggested that as many as 50% of HCCs may occur in noncirrhotic liver (4), that is, in patients in whom the progression of liver disease still appears rather mild. Thus, liver damage that appears severe by histologic examination is not a prerequisite for HCC.

Interestingly, during chronic HBV infections there is, in the face of persistent viremia, a decline over time in the fraction of infected hepatocytes, from 100% to as little as a few percent (5, 12–14, 16, 17, 22, 23, 27, 34, 37, 38). Along with HCC, this is perhaps the most surprising and unexplained outcome of chronic infection. The timing of this decline has not been systematically studied, but it is presumably gradual, occurring over years or decades, and dependent on persistent, albeit

low-level, killing of infected hepatocytes by antiviral cytotoxic T lymphocytes (CTLs) (20). It is believed that the liver is largely a closed, self-renewing population. Such a population might be expected to evolve under any strong or persistent selective pressure. In HBV-infected patients, the earliest and most persistent selective pressure is immune killing of infected hepatocytes, which should initially constitute the entire hepatocyte population. Persistent killing of HBV-infected hepatocytes could lead to clonal expansion of mutant or epigenetically altered hepatocytes that had lost the ability to support infection and that were not, therefore, targeted by antiviral CTLs.

Such a selective pressure may explain why foci of altered hepatocytes (FAH) and HCC are typically virus negative (1, 6, 11, 26, 29, 31, 35, 40, 41, 44). Normal or preneoplastic hepatocytes (e.g., in FAH) that have evaded the host immune response should undergo clonal expansion, because their death rate is lower than that of surrounding hepatocytes, even if they do not have a higher growth rate. Indeed, clonal expansion of hepatocytes has been detected, in the absence of cirrhosis, in woodchucks chronically infected with woodchuck hepatitis virus (WHV) (19) and in chimpanzees chronically infected with HBV (21). The presence of discrete foci of normal-appearing but virus-negative hepatocytes in chronically infected woodchuck livers (39) suggested, but did not prove, that normal-appearing hepatocytes that had lost the ability to support virus replication might clonally expand.

The purpose of the present study was, therefore, to determine if normal-appearing hepatocytes undergo clonal expansion. To address this issue, we focused on noncirrhotic livers,

* Corresponding author. Mailing address: Fox Chase Cancer Center, 333 Cottman Ave., Philadelphia, PA 19111. Phone: (215) 728-2462. Fax: (215) 728-2412. E-mail: ws_mason@fccc.edu.

[†] Supplemental material for this article may be found at <http://jvi.asm.org/>.

[‡] Published ahead of print on 2 June 2010.

because hepatocyte appearance and organization in many cirrhotic nodules are often considered to indicate premalignancy (7, 24, 25, 44), and this, together with the cellular environment in the cirrhotic liver, may explain why as many as 50% of cirrhotic nodules have been found to be made up of clonally expanded hepatocytes (2, 18, 24, 25, 28, 44). In older HBV patients, cirrhosis, the result of cumulative scarring due to ongoing tissue injury, presumably produces an evolutionary pressure on the hepatocyte population due to restricted blood flow and altered hepatic architecture.

Clonal expansion was detected by assaying for integrated HBV DNA by inverse PCR (19, 21). Because integration occurs at random sites in host DNA, each integration event provides a unique genetic marker for the cell in which it occurred, and for any daughter cells. Thus, the clonal expansion of these tagged hepatocytes can be measured by determining how many times a given virus-cell DNA junction is repeated in a liver fragment. Analysis of fragments of nontumorous liver from noncirrhotic HCC patients revealed that at least 1% of hepatocytes are present as clones of >1,000 cells. Examination of 5- μ m-thick sections of paraffin-embedded livers from the same patients revealed that clonally expanded hepatocytes were present in liver sections lacking preneoplastic lesions or changes. Therefore, normal-appearing hepatocytes must have undergone clonal expansion. Although clonal expansion was detected by analysis of integrated HBV DNA, the expansion did not appear to be due to the site of integration of the viral DNA into host DNA.

These results are consistent with the hypothesis that immune selection and the later emergence of liver cirrhosis, with altered lobular organization and restricted blood flow, may constitute the two major selective pressures on the hepatocyte population that culminate in hepatocellular carcinoma. More-direct proof of the role, if any, of immune selection in hepatocyte evolution and HCC will require, first of all, an assay with a greater ability to detect clonally expanded hepatocytes. The present approach is limited by a number of factors, including a need for integration near a particular restriction endonuclease cleavage site in host DNA and for conservation of particular viral sequences so that the integrated DNA can be amplified using the PCR primers chosen. These issues may explain why the fraction of clonally expanded hepatocytes reported here is much less than that suggested by histologic data showing that more than 50% of hepatocytes appear negative for virus replication in long-term carriers. Further dissection of this issue will also require localization and determination of the virologic status of hepatocyte clones present in tissue sections.

MATERIALS AND METHODS

Liver tissue. Tumor tissue and nontumorous samples of liver were collected from 5 male HBV carriers, numbered P1 to P5, during surgical resection of HCC and were stored at -80°C . Fragments (~500 mg) of each were provided for the current study. Separate samples were formalin fixed for histopathology as shown in Fig. 2. This study was reviewed and approved by the Fox Chase Cancer Center institutional review board (IRB).

Extraction of DNA from liver tissue. Total DNA was extracted, essentially as described previously, from ~1- to 3-mg fragments cut from the frozen liver blocks (21). Following incubation for 2 h at 55°C in a sodium dodecyl sulfate (SDS)-proteinase K solution, the mixture was extracted with phenol-chloroform (1:1). After the addition of 20 μg of dextran carrier, nucleic acids were precipitated by addition of 2 volumes of ethanol and overnight storage at -20°C . To assay for integrated viral DNA in paraffin-embedded tissue sections, the remain-

FIG. 1. Inverse PCR assay for integrated HBV DNA. An example of an integration junction between host DNA and the right end of DSL DNA, at nt 1831, is shown at the top. To detect integration junctions near this location in viral DNA, total DNA from infected livers was cleaved with NcoI, which cuts HBV DNA at nt 1374 and host DNA at locations that differ according to the integration site. The cut DNA was circularized by incubation with T4 DNA ligase and was cut with the BsiHKAI restriction endonuclease to produce linear molecules in which virus-cell junctions are flanked by viral sequences. These fragments were serially diluted into 96-well PCR trays and were amplified by nested PCR using the virus-specific primers described in Materials and Methods. PCR products were cut from the gels and sequenced to locate the integration site.

der of the frozen liver blocks were subjected to ethanol fixation, followed by dehydration with xylene and paraffin embedding (30). Five-micrometer-thick sections were cut and deposited on glass slides for further processing. To extract DNA, the sections were first deparaffinized in xylene and then soaked in 100% ethanol and briefly air dried. The tissue was then scraped into an SDS-proteinase K solution and was incubated for 2 h at 55°C . The nucleic acids were extracted and collected, by ethanol precipitation, as described above.

qPCR. Host cell DNA was quantified by quantitative real-time PCR (qPCR) of a region including host epsilon globin DNA (GenBank accession no. M81361) using a purified PCR-amplified fragment of epsilon globin DNA as a control. The DNA fragment used as a qPCR control was amplified with primers 5'-GAGCTAGAACTGGGTGAGAT-3' and 5'-ACCAGCACCTGAACCTCACCT-3' to yield a 976-nucleotide (nt) DNA. qPCR was carried out with the second primer and 5'-GGTGAATATTACCCTAGCAAGTTG-3' to yield a PCR product of 607 nt.

Viral DNA replication intermediates in liver DNA preparations were also quantified by qPCR, using primers 5'-TGTGCACTTCGCTTCACCTCT-3' (nt 1580 to 1600) and 5'-AAAGTTGCATGGTGTGGTG-3' (nt 1825 to 1806). These primers span the region of the HBV genome from slightly upstream of DR2 to slightly upstream of DR1 (i.e., the 5' end of HBV minus-strand DNA) and were designed based on viral DNA sequencing, as described below. HBV DNA (36) (GenBank accession no. X02763) cloned into the EcoRI site of pSP65 was cleaved with EcoRI and serially diluted for use as a qPCR control.

qPCR was carried out using the Bio-Rad SYBR green Supermix with each primer at a concentration of 0.25 μM . The reaction conditions were 2 min at 95°C , followed by 40 cycles of 15 s at 95°C , 15 s at 55°C , and 30 s at 72°C .

Sequencing of the region of the HBV genome that includes the right end of HBV DSL DNA. Inverse PCR (Fig. 1) was carried out to detect integration junctions between the right end of HBV double-stranded linear DNA (DSL DNA) (32) and host DNA (9, 43). In order to design primers for the inverse PCRs, we sequenced the right end of DSL DNA. Sequences between nt 1193 and 1860, with numbering according to Galibert et al. (8) (numbered according to the plus strand of GenBank accession no. V01460), were PCR amplified from the liver DNA of each patient by nested and heminested PCR, and the products were

subjected to direct DNA sequencing. The sequences of the inverse PCR primers for each patient were as follows. For P1, P2, P3, and P5, the outer primers were 5'-AAAGGACGTCCCGCGCAG-3' (nt 1424 to 1407) and 5'-TTCGCTTCACCTCTGCACG-3' (nt 1587 to 1605), and the inner primers were 5'-CACAGCCTAGCAGCCATGG-3' (nt 1392 to 1374) and 5'-CGCATGGAGACCACCGTGA-3' (nt 1607 to 1625). For P4, the outer primers were the same as those for the other patients, and the inner primers were 5'-CACACCCTAGCAGCCATGG-3' (nt 1392 to 1374) and 5'-CGCATGGAGACCACCGTGA-3' (nt 1607 to 1625).

Inverse PCR to detect integrated HBV DNA. For inverse PCR (Fig. 1), an aliquot of DNA isolated from ~1- to 3-mg liver fragments was digested with 10 U of NcoI-high fidelity (HF) in 40 μ l of NEB4 (New England Biolabs, Inc.). The NcoI was then heat inactivated for 20 min at 72°C, and the volume was adjusted to 450 μ l by addition of T4 DNA ligase buffer, followed by addition of 500 U of T4 DNA ligase (New England Biolabs, Inc.). The mixture was incubated for 2 h at ~25°C, after which the ligase was heat inactivated for 20 min at 72°C, followed by the addition of 10 μ l of 10% sodium dodecyl sulfate, 10 μ l of 5 M NaCl, and 2 μ l of 20-mg/ml dextran carrier. Two volumes of 100% ethanol were added, and the nucleic acids were allowed to precipitate overnight at -20°C. The nucleic acids were collected by centrifugation; the pellet was washed with 1 ml of 100% ethanol with vortexing; and the pellet was again collected by centrifugation. The nucleic acids were dried under a vacuum and were dissolved in 40 μ l of NEB2 supplemented with BSA. The DNA was digested first with 5 U of BsiHKA1 for 1 h at 65°C and then with 5 U of SphI for 1 h at 37°C (New England Biolabs, Inc.).

The DNA extracted from frozen liver was then serially diluted and subjected to nested PCR in 96-well trays as previously described (21), using the primers described above. DNA extracted from 5- μ m-thick tissue sections was not serially diluted; instead, an equivalent amount was distributed to each PCR mixture. The first PCR was carried out using AmpliTaq Gold reagents (Applied Biosystems, Foster City, CA) and enzyme, while the nested reaction was carried out using the Promega FlexI DNA polymerase system (Promega Corp., Madison, WI).

The products of nested PCRs were subjected to electrophoresis in 1.3% agarose gels, visualized with ethidium bromide, and cut from the gel. To identify virus-cell junctions, the products were sequenced (33) with the appropriate forward primer that maps from nt 1607 to 1625. Sequence alignments with HBV DNA were carried out using the GCG program FASTA to locate virus-cell junctions. To screen for repeated virus-cell junctions, cell sequences adjacent to integrated HBV DNA were screened using Sequencher (Gene Codes Corporation). The location of flanking DNA sequences on the human genome was determined using BLAST to search the "reference genomic sequences" (refseq_genomic) in the database, and the coding potential at this location in the human genome was evaluated using the Affymetrix Integrated Genome Browser (Feb_2009 build).

Histology. The patient samples were scored for fibrosis according to the system of Ishak et al. (modified Knodell score) (15) after formalin fixation, paraffin embedding, sectioning, and Masson's trichrome staining. Staining with hematoxylin and eosin (H&E) was used to assess general histologic features.

Frozen liver tissue for histological analysis was processed, paraffin embedded, sectioned (thickness, 5 μ m), and H&E stained for histological analysis (30).

Estimation of hepatocyte clone size. The following calculations were performed to obtain an unbiased estimate of clone size, assuming that 5- μ m-thick tissue sections had been cut through a spherical clone of hepatocytes. A spherical clone of hepatocytes of radius R_s was assumed to reside in a liver with a homogeneous hepatocyte concentration of ρ cells per unit volume. Thus, the clone was presumed to have a number of hepatocytes (N) equal to $4\pi\rho R_s^3/3$. In order to estimate N , we need to estimate R_s . In collecting thin sections of liver, we were in effect obtaining thin slices of this spherical clone. This slice will have the form of a short circular cylinder of height w and radius R_c . The average value of R_c is $\pi R_s/4$, where half the tissue sections have a radius greater and half have a radius smaller than R_c . Thus, we estimated R_s as $4R_c/\pi$ and proceeded to estimate R_c .

A nucleus contained in the spherical clone will be observed if a subsegment is captured in which the chromosome containing the virus-cell junction resides. In the current experiment, the section width was 5 μ m, compared to a hepatocyte nuclear diameter of ~6 μ m. The thin-section capture volume is $\pi R_c^2(2r + w)$ where r is the nuclear radius. That is, a portion of any nucleus within a distance of r or less from the section of thickness w will be included in a section. If $p(C)$ is the chance that a random nucleus is captured (i.e., that the integrant is included in the section), then the expected number of captured nuclei, $E(N_c)$, is calculated as $\rho\pi R_c^2(2r + w)p(C)$. To find $p(C)$, we consider a random nucleus centered at the origin. The slice is represented by two vertical lines whose coordinates are $x - w/2$ and $x + w/2$, where x is the coordinate of the slice center. If $-r - w/2 < x < r + w/2$, then a fraction of the nucleus will be contained in the

slice. This fraction is $p(C)$. The volume contained between two parallel vertical planes both cutting a sphere of radius r , centered at the origin, $V(\alpha, \beta)$, is calculated as $\pi(\beta - \alpha)[r^2 - (\beta^2 + \alpha\beta + \alpha^2)/3]$, where α is the x coordinate of the left plane and β is that of the right. The slice center, x , is characterized as being in one of three regions: either $-r - w/2 < x < -r + w/2$, $-r + w/2 < x < r - w/2$, or $r - w/2 < x < r + w/2$. The first region has the slice approaching from the left until its left surface just touches the nucleus. The second has the slice completely inside the nucleus, moving from touching on the left to touching on the right. The final region has the slice moving out of the nucleus until its left surface just touches. The chance of capturing the nucleus is the fraction of the nuclear volume within the parallel planes divided by the total nuclear volume.

$$p(C) = \frac{1}{(4/3)\pi r^3} \int_{-r-w/2}^{r+w/2} V[\alpha(x), \beta(x)] \frac{dx}{2r + w} = \frac{1}{(4/3)\pi r^3(2r + w)} (I_1 + I_2 + I_3)$$

where

$$I_1 = \int_{-r-w/2}^{-r+w/2} V(-r, x + w/2) dx$$

$$I_2 = \int_{-r+w/2}^{r-w/2} V(x - w/2, x + w/2) dx$$

$$I_3 = \int_{r-w/2}^{r+w/2} V(x - w/2, r) dx$$

$E(N_c)$ is set to the observed nuclear count, and the equation is solved for R_c . To account for the expected off-equatorial position of the slice, R_c is multiplied by $4/\pi$ to estimate R_s , the clone radius. Finally, the clone size, N , is estimated by the equation $N = 4\pi\rho R_s^3/3$. Because cuts are taken at random sites through the sphere of hepatocytes, half the estimates of N may exceed the true value by as much as 2.06-fold, which would be achieved if a cut were taken through the waist of the sphere. That is, clone size could be overestimated by as much as 2.06-fold.

RESULTS

Detection of clonally expanded hepatocytes in samples from human livers chronically infected with HBV. Hepadnaviruses typically have a relaxed circular, partially double stranded, DNA genome (rcDNA). However, about 10 to 20% of virus particles in the bloodstream have a double-stranded, linear DNA genome (DSL DNA), which is formed as a result of aberrant plus-strand priming during rcDNA formation (32). A second form of DSL DNA, with a terminal redundancy corresponding to the cohesive overlap of rcDNA, is also thought to occur. However, this larger form is apparently created via strand displacement synthesis through the cohesive overlap. This is thought to occur when rcDNA enters the hepatocyte nucleus to form covalently closed circular viral DNA (cccDNA), the template for viral RNA synthesis (42).

Integration of viral DNA via illegitimate recombination occurs at random sites in host DNA during the course of an infection. The predominant substrate for integration is the DSL DNA (9, 10, 43), formed as a result of aberrant plus-strand priming. Since each integration of viral DNA occurs at a random site in host DNA, repeated occurrence of a particular integration in a liver fragment can be attributed to clonal expansion of the hepatocyte in which that integration occurred. Assaying for highly repeated integration sites can therefore be used to estimate the clonal expansion of hepatocytes within a liver tissue fragment.

Integrated viral DNA can be detected and quantified by inverse PCR (Fig. 1). The inverse PCR procedure used in this study was designed to detect junctions between the right side of

TABLE 1. Hepatocytes analyzed for integrated DNA by inverse PCR^a

Noncirrhotic patient no.	Ishak fibrosis score ^b	Total no. of HBV DNA copies/hepatocyte	Total no. of hepatocytes extracted for inverse PCR (10 ⁶) ^c
1	3	1.7	1.45
2	1	2.4	0.6
3	1	2.9	3.4
4	2	2.6	2.4
5	3	1.4	1.1
Total			8.95

^a For each patient, total cellular DNA was extracted from four fragments (weight, ~1 to 3 mg) of liver. Total HBV DNA and host cellular DNA were quantified by qPCR, as described in Materials and Methods.

^b Fibrosis scores based on the Ishak staging system range from 0 (no cirrhosis) to 6 (definite cirrhosis) (15). Stage 1 is characterized by fibrous expansion of some portal areas, with or without short fibrous septa; stage 2, by fibrous expansion of most portal areas, with or without short fibrous septa; stage 3, by fibrous expansion of most portal areas with occasional portal-to-portal bridging. Stage 4 includes marked portal-to-portal and portal-to-central bridging as well. Stages 5 and 6 are defined as incomplete cirrhosis and definite or probable cirrhosis, respectively.

^c The sum of the total number of hepatocytes in the four samples from each liver is based on the assumption that hepatocytes constitute 70% of the cellular population of the liver.

DSL DNA (32) and cellular DNA. The right end of DSL DNA includes the viral X gene and terminates at the 5' end of the minus strand (see Fig. 4).

In the present study, HBV DNA integration sites were quantified in nontumorous liver samples acquired from 5 HBV patients, P1 to P5, during surgical resection of HCC (Table 1; Fig. 2). As illustrated in Fig. 2 and summarized in Table 1, fibrosis ranged from mild (P2 and P3) to severe (P1 and P5); none of the patients were cirrhotic. DNA was extracted from 4 liver fragments (weight, ~1 to 3 mg) cut from each patient sample. Host DNA and viral DNA were quantified by qPCR (Table 1). As reported in other studies of late-stage HBV infections (e.g., reference 37), the amounts of viral DNA in the liver were small, with an average of only a few copies per hepatocyte. Inverse PCR was next carried out to detect integrants (19, 21). The assay, involving serial endpoint dilution prior to nested PCR, to amplify integrated DNAs (19, 21), was capable in theory of detecting repeats of integrations when at least ~50 to 100 were present in a liver fragment; in practice, the presence of high-copy-number clones in a DNA sample may mask the detection of low-copy-number clones. As shown in Fig. 3, hepatocyte clones, ranging in size from ~100 to ~60,000 hepatocytes, were found.

All but two of the mapped virus-cell junctions occurred upstream of the right end of DSL DNA (Fig. 4); none occurred immediately downstream, as might have been expected if rcDNA or cccDNA were a common precursor to integrated viral DNA. Thus, for HBV DNA, as for duck hepatitis B virus (DHBV) and WHV DNA (9, 10, 42, 43), the data are consistent with the idea that a linear viral DNA, and not rcDNA, serves as the major substrate for integration into host DNA.

Sites of integration of viral DNA did not explain the clonal proliferation of hepatocytes. Integration sites were mapped on the human genome as described in Materials and Methods. The results are summarized in Table 2 and are presented in

FIG. 2. H&E-stained liver tissues from HCC patients. Formalin-fixed, paraffin-embedded liver tissues (nontumorous) from patients P1, P2, P4, and P5 were stained with H&E and Masson's trichrome stain (MT) in order to assess the degree of deposition of collagen fiber, which is stained blue. Separate blocks of tissue were frozen for inverse PCR, as shown in Fig. 5. Magnifications, $\times 200$ (H&E) and $\times 40$ (Masson's trichrome stain). P3 liver tissue (not shown) was similar in appearance to P2 liver tissue.

more detail in Table S5 in the supplemental material. Among 12 clones that had at least 1,000 cells, the integration sites in 11 were mapped to 9 different chromosomes. Of these 11 integration sites, 7 were located in intergenic regions (Table 3) and 4 in the introns of known or predicted genes. Among the remaining 32 integration sites, not present in large clones, 29 were mapped, 9 to introns of known or predicted genes and the

FIG. 3. Large hepatocyte clones in liver samples from patients 1 to 5. Four frozen liver fragments from each patient were extracted and analyzed. A summary of the large clones found in these samples is shown. Integration sites corresponding to clones of 1,000 or more hepatocytes are summarized in Table 3 and, in greater detail, in Table S5 in the supplemental material.

FIG. 4. Evidence that HBV DSL DNA is the precursor to integrated HBV DNA. The junction sites for each distinct integration event are plotted. As can be seen by comparison to a map of the viral genome, numbering from the unique EcoRI site, all integration junctions but two occurred upstream of the 5' end of the minus strand, which defines the right end of DSL DNA. None were detected immediately downstream, although virus-cell junctions were common in this region when integration junctions near the left end of HBV DSL DNA were analyzed (21).

remaining 20 to intergenic regions. Thus, the distribution between integration into intergenic sites and integration into host genes was about the same for the two groups of samples.

The inverse PCR assay detected a highly repeated integration site in the HCC from only one of the patients (P2; integration in an intergenic region of chromosome 4), possibly because viral integration junctions in the remaining HCCs were not near an NcoI site in host DNA (see Fig. 1).

In summary, extensive clonal expansion does not appear to be associated with the site of integration of HBV DNA. Similar results were obtained in a previous study of clonal expansion in chimpanzees chronically infected with HBV (21).

Normal-appearing hepatocytes undergo clonal expansion during chronic HBV infection. Studies were next carried out to determine if clonal expansion occurred among normal-appearing hepatocytes or was strictly correlated with the emergence of FAH, often considered preneoplastic. To address this issue, DNA was extracted from tissue sections of ethanol-fixed, paraffin-embedded liver from the same tissue blocks, from patients P1 to P5, analyzed in Fig. 3. One 5- μ m-thick section from each block was H&E stained. DNA was extracted from an immediately adjacent tissue section, and a portion of the extracted DNA was analyzed for integrated HBV DNA by inverse PCR (Fig. 5; Table 4). For purposes of comparison, the section thickness of 5 μ m compares to a hepatocyte diameter of $\sim$ 20 μ m and a hepatocyte nuclear diameter of $\sim$ 6 μ m. Evidence that the tissue sections intersected large clones of hepatocytes was obtained with blocks from two of the four patients, P3 and P4. These clones were distinct from the clones

summarized in Fig. 3. As illustrated in Fig. 5 for patient P3 and as found for all five patients, the tissue section adjacent to the section from which DNA was extracted did not contain preneoplastic lesions, indicating that the clones represented hepatocytes that appear normal by conventional histologic analysis.

Evidence that inverse PCR using a single restriction endonuclease underestimates the prevalence of hepatocyte clones. The inverse PCR assay described in Fig. 1 would be expected to underestimate the frequency of hepatocyte clones, because detection of an integrant requires that the integration occur no more than about 1,000 nt from an NcoI cleavage site in host DNA. To test this, we asked if additional clones would be detected if another restriction endonuclease, DpnII, were used in place of NcoI (see Fig. 1). DpnII has a 4-base recognition sequence and a 4-base overhang. Three of the DNA preparations used to generate the data in Fig. 3 were reanalyzed using DpnII. Approximately as many clones of $>1,000$ hepatocytes were found as with NcoI (data not shown). The data suggest that the actual number of hepatocytes that have undergone significant clonal expansion may be much larger than would be estimated by using only one, or even two, restriction endonucleases to delimit detectable integration sites in host DNA.

TABLE 2. Widespread distribution of HBV DNA integration sites in noncirrhotic liver by chromosome^a

Chromosome no.	No. of clones of $>1,000$ hepatocytes ^b	Total no. of distinct virus-cell junctions	Chromosome length (bp)	No. of integrations per 10^8 bp
1		0	2.5×10^8	0.0
2	1	7	2.4×10^8	2.9
3	2	3	1.98×10^8	1.5
4	2	3	1.91×10^8	1.6
5	1	2	1.81×10^8	1.1
6		3	1.71×10^8	1.8
7	1	2	1.59×10^8	1.3
8		2	1.46×10^8	1.4
9		1	1.41×10^8	0.7
10	1	2	1.36×10^8	1.5
11		1	1.35×10^8	0.7
12		4	1.34×10^8	2.9
13		0	1.15×10^8	0.0
14		1	1.07×10^8	0.9
15		0	1.03×10^8	0.0
16	1	2	8.9×10^7	2.5
17		0	8.1×10^7	0.0
18		0	7.8×10^7	0.0
19	1	1	5.9×10^7	1.7
20	1	3	6.3×10^7	4.8
21		0	4.8×10^7	0.0
22		1	5.1×10^7	2.0
X		2	1.55×10^8	1.3
Y		0	5.9×10^7	0.0

^a Forty distinct integration sites detected on human chromosomes were mapped and are normalized in the rightmost column to integrations per 10^8 bp, to illustrate that the distribution by chromosome did not show any obvious preferences. These data are summarized in greater detail in Table S5 in the supplemental material.

^b Out of 12 clones with $>1,000$ hepatocytes, the integration site in only 1 could not be assigned to a specific chromosome. Among the remaining 32 integration sites (in smaller clones), 3 could not be mapped.

TABLE 3. Clonal expansion of hepatocytes did not correlate with integration into any obvious chromosomal locations^a

Patient	Clone size (no. of cells)	Chromosome	Gene
1	2,430	7	Sema domain, immunoglobulin domain (Ig), short basic domain, secreted, semaphorin 3E
	2,430	20	PLCB (phospholipase C, beta 1 [phosphoinositide specific])
	7,290	5	Intergenic
	14,300	10	Intergenic
2	2,430	2	Intergenic
3	2,160	3	ATPase, Ca ²⁺ transporting, plasma membrane 2
	7,290	2	Intergenic
	14,580	4	Kinase insert domain receptor (a type III receptor tyrosine kinase)
4	7,290	19	Intergenic
5	23,760	3	Intergenic
	58,320	16	Intergenic

^a Data for clones of >1,000 hepatocytes are shown. Clones from Fig. 3 for which integration sites could not be mapped are not summarized here. All results are summarized in greater detail in Table S5 in the supplemental material.

DISCUSSION

This study demonstrates clonal expansion of human hepatocytes during chronic HBV infections. It extends to humans previous findings made with the woodchuck and chimpanzee models of chronic hepadnavirus infections (19, 21) in showing that large clones of hepatocytes emerged after hepadnavirus infection. In addition, we show here that normal-appearing hepatocyte are able to undergo extensive clonal expansion, an issue left unresolved in the chimpanzee and woodchuck studies.

The basis of the clonal expansion, especially of that of the normal-appearing hepatocytes, remains unclear. One explanation would involve cellular transformation leading to unregulated growth, either through expression of viral genes from integrated DNA or via host cell DNA mutations. However, this does not explain clones with no obvious morphological transformation (Fig. 5 and Table 4). A second explanation could involve random death and regeneration within the entire hepatocyte population. This scenario would produce hepatocyte clones, but it seems unlikely to explain very large clones of 10,000 or more hepatocytes. This is because the generation of clones of even a few thousand hepatocytes by this route would require several hundred turnovers of the entire hepatocyte population (21), a possibility for which there is currently no experimental support.

Another explanation for large clones of hepatocytes involves the resident stem/progenitor cells. If mature hepatocytes become senescent due to persistent pressure to divide in order to sustain liver size, then replacement would eventually depend on the maturation of hepatic stem cells to produce young hepatocytes, which would become responsible, during long-term infections, for the bulk of hepatocyte replacement. That is, the bulk of the hepatocyte population loses the capacity for self-renewal, which now depends entirely on only a small sub-

FIG. 5. Inverse PCR for integrated HBV DNA in fixed liver tissue from an HBV carrier. Fixation and embedding of frozen liver tissue for sectioning was carried out after the removal of tissue fragments for the experiments described in the legend to Fig. 3 and in Tables 1 and 2; thus, these integration events were distinct from those summarized there and in Table S5 in the supplemental material. (A) Representative area of an ethanol-fixed, H&E-stained 5-μm-thick liver tissue section from patient 3. (B) Inverse PCR of DNA extracted from an immediately adjacent 5-μm-thick section containing ~20,000 hepatocytes. Inverse PCR for integrated HBV was carried out on one-eighth of the total extracted DNA. Following the inversion reaction (Fig. 1), the DNA was distributed to a 96-well tray for nested PCR. Results of gel electrophoresis of the PCR products are shown. The bands were extracted and sequenced. The results are summarized in Table 4 and are presented in detail in Table S6 in the supplemental material.

TABLE 4. Hepatocyte clones are detected by inverse PCR of DNA extracted from 5- μ m-thick sections of paraffin-embedded liver^a

Patient	Slide(s)	No. of repeats detected	Estimated focus size(s) (clone size)	Chromosome	Location
P1		4	32	4	Intergenic
		3	24	4	Intergenic
P2		2	16	13	Intergenic
P3	1, 3	26, 33	208 (19,300), 264 (27,600)	2	Intergenic
	1, 3	2, 1	24		
	1	4	32	17	Intron of rabaptin, RAB GTPase binding effector protein 1
	1, 3	3, 5	24, 40	11	Intron of brain-derived neurotrophic factor (BDNF) opposite strand (non-protein coding)
	1, 3	2, 4	16, 32	7	Intron of dipeptidyl-peptidase 6 transcript variant 3
	3	2	16		
	3	3	24	12	Intergenic
P4		13	104 (6,800)	1	Intron of vacuolar protein sorting 13 homolog D (<i>Saccharomyces cerevisiae</i>)
		13	104 (6,800)	3	Exon of RAB6B, member of the RAS oncogene family
		3	24	11	Intron of NEL-like 1 (chicken)
		2	16	15	Intron of suppression of tumorigenicity 18 (breast carcinoma) (zinc finger protein)
		2	16	9	Intergenic
		2	16	14	Intergenic
P5	3	6	48	16	Intergenic

^a Of the total DNA extracted from single paraffin sections, 12.5%, containing ~17,000 to 28,000 hepatocytes, based on hepatocyte nuclear counts, was analyzed by inverse PCR for integrated viral DNA. For P3, two slides were analyzed, flanking either side of the H&E-stained slide shown in Fig. 5. Four of seven clones were present on both slides. The approximate focus size was estimated by multiplying the observed repeats by 8. Clone size was estimated as described in Materials and Methods, by assuming that hepatocyte clones were spherical. The separate calculations for P3 are for the same clone. The approach used for estimating the clone size could lead to an overestimate of as much as 2.06-fold, if the 5- μ m-thick sections cut through the equator of a sphere. Data on these integrations are summarized in greater detail in Table S6 in the supplemental material.

set of hepatocytes. To our knowledge, there are no data to support or refute this possibility.

Finally, as discussed earlier, we favor a model in which immune evasion is the basis for clonal expansion. That is, clonal expansion involves hepatocytes that have lost the ability to support virus replication and are no longer targets of the immune response to viral antigens. Since we have shown here that clones can be detected in tissue sections (Fig. 5; Table 4), this should be testable.

A practical issue also needs to be resolved. Inverse PCR is useful for detecting the integration of viral DNA near particular restriction endonuclease cleavage sites in host DNA, but, as shown above, this approach underestimates the amount of integration and of clonal expansion, because not all integrations will be detected using only a single enzyme. In addition, the use of a serial endpoint dilution approach to quantifying clonal expansion will be biased toward detection of the largest clones in a DNA sample, with smaller clones being missed or underestimated due to competition between different PCR products in the same reaction. Thus, the estimate of 1% for the number of hepatocytes present in clones of >1,000 cells may be a gross underestimate of the true number. Other approaches are clearly needed to resolve this problem.

Finally, the current study was carried out using nontumorous liver samples acquired during surgical resection of HCC. An important but so far unresolved question is whether extensive clonal expansion is characteristic of human HBV infections in general, or only of those ~25% of infections with the highest risk of progressing to cirrhosis and HCC during a patient's lifetime, presumably those with the highest level of cumulative liver damage. This cannot be decided from the woodchuck

model, because ~100% of chronically infected woodchucks develop HCC. The finding of clonal expansion in the chimpanzee model, often considered not to develop HCC, might argue against this possibility, but in fact the risk of cirrhosis and HCC in chimpanzees is unknown, because very few chimpanzees have ever been subjected to long-term monitoring. Two questions may therefore be worth assessing in human samples. (i) Is clonal expansion early in life (i.e., in the immune tolerance phase of infection [45]) a risk factor for cirrhosis and/or HCC in later life? (ii) Is the appearance of HBV-resistant hepatocytes in young patients, as a major fraction of the hepatocyte population, a surrogate marker of clonal expansion and/or an independent risk factor for the development of cirrhosis and HCC in later life? This should be a fairly straightforward issue to address if frozen tissue is available for histologic and biochemical analyses, as illustrated in Fig. 5.

ACKNOWLEDGMENTS

We are grateful to Allison Jilbert (University of Adelaide), Jesse Summers (University of New Mexico), and John Taylor, Christoph Seeger, and Thomas Tu (University of Adelaide) for helpful discussions during the course of this work and to Glenn Rall, Jesse Summers, and John Taylor for critical reading of the manuscript.

We acknowledge assistance from the research support facilities of the Fox Chase Cancer Center. W.S.M. was supported by grants from the National Institutes of Health (5R01AI018641; CA06927) and by an appropriation from the Commonwealth of Pennsylvania.

REFERENCES

1. Abe, K., T. Kurata, T. Shikata, and B. C. Tennant. 1988. Enzyme-altered liver cell foci in woodchucks infected with woodchuck hepatitis virus. *Jpn. J. Cancer Res.* 79:466–472.
2. Aoki, N., and W. S. Robinson. 1989. State of hepatitis B viral genomes in cirrhotic and hepatocellular carcinoma nodules. *Mol. Biol. Med.* 6:395–408.

3. Beasley, R. P. 1982. Hepatitis B virus as the etiologic agent in hepatocellular carcinoma: epidemiologic considerations. *Hepatology* 2:21S–26S.
4. Bosch, F. X., J. Ribes, R. Cléries, and M. Diaz. 2005. Epidemiology of hepatocellular carcinoma. *Clin. Liver Dis.* 9:191–211.
5. Chu, C. M., C. T. Yeh, R. N. Chien, I. S. Sheen, and Y. F. Liaw. 1997. The degrees of hepatocyte nuclear but not cytoplasmic expression of hepatitis B core antigen reflect the level of viral replication in chronic hepatitis B virus infection. *J. Clin. Microbiol.* 35:102–105.
6. Fausto, N. 2004. Liver regeneration and repair: hepatocytes, progenitor cells, and stem cells. *Hepatology* 39:1477–1486.
7. Furuya, K., M. Nakamura, T. Yamamoto, K. Togel, and H. Otsuka. 1988. Macroregenerative nodule of the liver. A clinicopathologic study of 345 autopsy cases of chronic liver disease. *Cancer* 61:99–105.
8. Galibert, F., E. Mandart, P. Fitoussi, P. Tiollais, and P. Charnay. 1979. Nucleotide sequence of the hepatitis B virus genome (subtype ayw) cloned in *E. coli*. *Nature* 281:646–650.
9. Gong, S. S., A. D. Jensen, C. J. Chang, and C. E. Rogler. 1999. Double-stranded linear duck hepatitis B virus (DHBV) stably integrates at a higher frequency than wild-type DHBV in LMH chicken hepatoma cells. *J. Virol.* 73:1492–1502.
10. Gong, S. S., A. D. Jensen, H. Wang, and C. E. Rogler. 1995. Duck hepatitis B virus integrations in LMH chicken hepatoma cells: identification and characterization of new episomally derived integrations. *J. Virol.* 69:8102–8108.
11. Govindarajan, S., A. Conrad, B. Lim, B. Valinluck, A. M. Kim, and P. Schmid. 1990. Study of preneoplastic changes in liver cells by immunohistochemical and molecular hybridization techniques. *Arch. Pathol. Lab. Med.* 114:1042–1045.
12. Gowans, E. J., C. J. Burrell, A. R. Jilbert, and B. P. Marmion. 1985. Cytoplasmic (but not nuclear) hepatitis B virus (HBV) core antigen reflects HBV DNA synthesis at the level of the infected hepatocyte. *Intervirology* 24:220–225.
13. Gowans, E. J., C. J. Burrell, A. R. Jilbert, and B. P. Marmion. 1983. Patterns of single- and double-stranded hepatitis B virus DNA and viral antigen accumulation in infected liver cells. *J. Gen. Virol.* 64:1229–1239.
14. Hirohashi, S., Y. Shimosato, Y. Ino, and K. Kishi. 1982. Distribution of hepatitis B surface and core antigens in human liver cell carcinoma and surrounding nontumorous liver. *JNCI* 69:565–568.
15. Ishak, K., A. Baptista, L. Bianchi, F. Callea, J. De Groote, F. Gudat, H. Denk, V. Desmet, G. Korb, R. N. MacSween, et al. 1995. Histological grading and staging of chronic hepatitis. *J. Hepatol.* 22:696–699.
16. Kojima, M., K. Udo, Y. Takahashi, H. Yoshizawa, F. Tsuda, Y. Itoh, Y. Miyakawa, and M. Mayumi. 1977. Correlation between titer of antibody to hepatitis B core antigen and presence of viral antigens in the liver. *Gastroenterology* 73:664–667.
17. Lamothe, F., J. Laurecin-Piche, J. Cote, R. Guevin, A. Viallet, and G. Richer. 1976. Detection of surface and core antigens of hepatitis B virus in the liver of 164 human subjects. A study by immunoperoxidase and orcein staining. *Gastroenterology* 71:102–108.
18. Mashal, R. D., S. C. Lester, and J. Sklar. 1993. Clonal analysis by study of X chromosome inactivation in formalin-fixed paraffin-embedded tissue. *Cancer Res.* 53:4676–4679.
19. Mason, W. S., A. R. Jilbert, and J. Summers. 2005. Clonal expansion of hepatocytes during chronic woodchuck hepatitis virus infection. *Proc. Natl. Acad. Sci. U. S. A.* 102:1139–1144.
20. Mason, W. S., S. Litwin, and A. R. Jilbert. 2008. Immune selection during chronic hepadnavirus infection. *Hepatology* 47:2–16.
21. Mason, W. S., H.-C. Low, C. Xu, C. E. Aldrich, C. A. Scougall, A. Grosse, A. Clouston, D. Chavez, S. Litwin, S. Peri, A. R. Jilbert, and R. E. Lanford. 2009. Detection of clonally expanded hepatocytes in chimpanzees with chronic hepatitis B virus infection. *J. Virol.* 83:8396–8408.
22. Nayak, N. C., A. Dhar, R. Sachdeva, A. Mittal, H. N. Seth, D. Sudarsanam, B. Reddy, U. L. Waghlikar, and C. R. M. Reddy. 1977. Association of human hepatocellular carcinoma and cirrhosis with hepatitis B virus surface and core antigens in the liver. *Int. J. Cancer* 20:643–654.
23. Omata, M., A. Afroudakis, C.-T. Liew, M. Ashcavai, and R. L. Peters. 1978. Comparison of serum hepatitis B surface antigen (HBsAg) and serum anti-core with tissue HBsAg and hepatitis B core antigen (HBcAg). *Gastroenterology* 75:1003–1009.
24. Paradis, V., I. Laurendeau, M. Vidaud, and P. Bedossa. 1998. Clonal analysis of macronodules in cirrhosis. *Hepatology* 28:953–958.
25. Piao, Z., Y. N. Park, H. Kim, and C. Park. 1997. Clonality of large regenerative nodules in liver cirrhosis. *Liver* 17:251–256.
26. Radaeva, S., Y. Li, H. J. Hacker, V. Burger, A. Kopp-Schneider, and P. Bannasch. 2000. Hepadnaviral hepatocarcinogenesis: in situ visualization of viral antigens, cytoplasmic compartmentation, enzymic patterns, and cellular proliferation in preneoplastic hepatocellular lineages in woodchucks. *J. Hepatol.* 33:580–600.
27. Ray, M. B., V. J. Desmet, A. F. Bradburne, J. Desmyter, J. Fevery, and J. De Groote. 1976. Differential distribution of hepatitis B surface antigen and hepatitis B core antigen in the liver of hepatitis B patients. *Gastroenterology* 71:462–467.
28. Robinson, W. S., L. Klote, and N. Aoki. 1990. Hepadnaviruses in cirrhotic liver and hepatocellular carcinoma. *J. Med. Virol.* 31:18–32.
29. Rogler, C. E. 1991. Cellular and molecular mechanisms of hepatocarcinogenesis associated with hepadnavirus infection. *Curr. Top. Microbiol. Immunol.* 168:103–140.
30. Sainte-Marie, G. 1962. A paraffin embedding technique for studies employing immunofluorescence. *J. Histochem. Cytochem.* 10:250–256.
31. Sell, S., J. M. Hunt, B. J. Knoll, and H. A. Dunsford. 1987. Cellular events during hepatocarcinogenesis and the question of premalignancy. *Adv. Cancer Res.* 48:37–111.
32. Staprans, S., D. D. Loeb, and D. Ganem. 1991. Mutations affecting hepadnavirus plus-strand DNA synthesis dissociate primer cleavage from translocation and reveal the origin of linear viral DNA. *J. Virol.* 65:1255–1262.
33. Summers, J., A. R. Jilbert, W. Yang, C. E. Aldrich, J. Saputelli, S. Litwin, E. Toll, and W. S. Mason. 2003. Hepatocyte turnover during resolution of a transient hepadnaviral infection. *Proc. Natl. Acad. Sci. U. S. A.* 100:11652–11659.
34. Suzuki, K., T. Uchida, R. Horiuchi, and T. Shikata. 1985. Localization of hepatitis B surface and core antigens in human hepatocellular carcinoma by immunoperoxidase methods. Replication of complete virions of carcinoma cells. *Cancer* 56:321–327.
35. Toshkov, I., H. J. Hacker, M. Roggendorf, and P. Bannasch. 1990. Phenotypic patterns of preneoplastic and neoplastic hepatic lesions in woodchucks infected with woodchuck hepatitis virus. *J. Cancer Res. Clin. Oncol.* 116:581–590.
36. Valenzuela, P., P. Gray, M. Quiroga, J. Zaldivar, H. M. Goodman, and W. J. Rutter. 1979. Nucleotide sequence of the gene coding for the major protein of hepatitis B virus surface antigen. *Nature* 280:815–819.
37. Volz, T., M. Lutgehetmann, P. Wachtler, A. Jacob, A. Quaas, J. M. Murray, M. Dandri, and J. Petersen. 2007. Impaired intrahepatic hepatitis B virus productivity contributes to low viremia in most HBeAg-negative patients. *Gastroenterology* 133:843–852.
38. Wursthorn, K., M. Lutgehetmann, M. Dandri, T. Volz, P. Buggisch, B. Zollner, T. Longerich, P. Schirmacher, F. Metzler, M. Zankel, C. Fischer, G. Currie, C. Brosgart, and J. Petersen. 2006. Peginterferon alpha-2b plus adefovir induce strong cccDNA decline and HBsAg reduction in patients with chronic hepatitis B. *Hepatology* 44:675–684.
39. Xu, C., T. Yamamoto, T. Zhou, C. E. Aldrich, K. Frank, J. M. Cullen, A. R. Jilbert, and W. S. Mason. 2007. The liver of woodchucks chronically infected with the woodchuck hepatitis virus contains foci of virus core antigen-negative hepatocytes with both altered and normal morphology. *Virology* 359:283–294.
40. Yang, D., E. Alt, and C. E. Rogler. 1993. Coordinate expression of N-myc 2 and insulin-like growth factor II in precancerous altered hepatic foci in woodchuck hepatitis virus carriers. *Cancer Res.* 53:2020–2027.
41. Yang, D., R. Faris, D. Hixson, S. Affigne, and C. E. Rogler. 1996. Insulin-like growth factor II blocks apoptosis of N-myc-2-expressing woodchuck liver epithelial cells. *J. Virol.* 70:6260–6268.
42. Yang, W., W. S. Mason, and J. Summers. 1996. Covalently closed circular viral DNA formed from two types of linear DNA in woodchuck hepatitis virus-infected liver. *J. Virol.* 70:4567–4575.
43. Yang, W., and J. Summers. 1999. Integration of hepadnavirus DNA in infected liver: evidence for a linear precursor. *J. Virol.* 73:9710–9717.
44. Yeh, S. H., P. J. Chen, W. Y. Shau, Y. W. Chen, P. H. Lee, J. T. Chen, and D. S. Chen. 2001. Chromosomal allelic imbalance evolving from liver cirrhosis to hepatocellular carcinoma. *Gastroenterology* 121:699–709.
45. Yim, H. J., and A. S. Lok. 2006. Natural history of chronic hepatitis B virus infection: what we knew in 1981 and what we know in 2005. *Hepatology* 43:S173–S181.
46. Yu, M. C., J. M. Yuan, R. K. Ross, and S. Govindarajan. 1997. Presence of antibodies to the hepatitis B surface antigen is associated with an excess risk for hepatocellular carcinoma among non-Asians in Los Angeles County, California. *Hepatology* 25:226–228.