Deep sequencing of liver explant transcriptomes reveals extensive expression from integrated hepatitis B virus DNA

Johan Ringlander 1 | Catarina Skoglund 2,3 | Kasthuri Prakash 1 | Maria E. Andersson 1 | Simon B. Larsson 1 | Ka-Wei Tang 1 | Gustaf E. Rydell 1 | Sanna Abrahamsson 4 | Maria Castedal 2,3 | Heléne Norder 1 | Kristoffer Hellstrand 1 | Magnus Lindh 1

Abstract
Hepatitis B virus (HBV) is a major cause of hepatocellular carcinoma (HCC). Integration of HBV DNA into the human genome may contribute to oncogenesis and to the production of the hepatitis B surface antigen (HBsAg). Whether integrations contribute to HBsAg levels in the blood is poorly known. Here, we characterize the HBV RNA profile of HBV integrations in liver tissue in patients with chronic HBV infection, with or without concurrent hepatitis D infection, by transcriptome deep sequencing. Transcriptomes were determined in liver tissue by deep sequencing providing 200 million reads per sample. Integration points were identified using a bioinformatic pipeline. Explanted liver tissue from five patients with end-stage liver disease caused by HBV or HBV/HDV was studied along with publicly available transcriptomes from 21 patients. Almost all HBV RNA profiles were devoid of reads in the core and the 3′ redundancy (nt 1830-1927) regions, and contained a large number of chimeric viral/human reads. Hence, HBV transcripts from integrated HBV DNA rather than from covalently closed circular HBV DNA (cccDNA) predominated in late-stage HBV infection, in particular in cases with hepatitis D virus co-infection. The findings support the suggestion that integrated HBV DNA can be a significant source of HBsAg in humans.

KEYWORDS
HBsAg, HCC, hepatitis B virus, integration, transcriptome

1 | BACKGROUND

Chronic infection with hepatitis B virus (HBV) is a major cause of hepatocellular carcinoma (HCC) due to inflammation, formation of oncogenic viral proteins and integration of HBV DNA into the human genome. The potential role of HBV DNA integrations for the development of HCC has been addressed in many studies. Integrations are randomly distributed in all human chromosomes, but some locations are reportedly more frequent in cancer tissue. HBsAg expression from integrations has been proposed to be of importance for maintaining high levels of the hepatitis B surface antigen (HBsAg) in the blood. This possibility is supported by the absence of

Abbreviations: HBsAg, hepatitis B surface antigen; HBV, Hepatitis B virus; HCC, hepatocellular carcinoma; HDV, hepatitis delta virus; ICGC, International Cancer Genome Consortium.
HBsAg decline during prolonged antiviral therapy, by remaining HBsAg levels when HBV DNA is suppressed to low levels during the natural course of infection, and by high levels of hepatitis delta virus (HDV) with HBsAg in its envelope, despite low HBV replication.

If integrated HBV DNA were expressed, the transcripts would differ from HBV RNA derived from covalently closed circular DNA (cccDNA), which is the template for HBV replication. Transcripts from cccDNA are typically ≈3300 nt long (excluding the poly-A tail) and contain a 3′ RNA ‘redundancy’ with the 3′ end at nucleotide (nt) 1927. By contrast, integration-derived RNA lacks the 5′ part that encodes the core antigen due to absence of an upstream promoter that initiates this transcription, and ends at or slightly upstream of nt 1830.

For the present study, we utilized these predicted differences to assess the degree of expression from HBV integrations. For this purpose, we used transcriptome deep sequencing, but with greater depth and longer read length than in standard RNA assays. The RNA purification comprised a ribosomal RNA depletion step and cDNA synthesis was performed with hexamer primers, a strategy aiming at increasing the likelihood of detecting reads that contain junctions between viral and human RNA as compared with using poly-A enrichment. Also, a sensitive bioinformatics pipeline was developed and applied on both transcriptomes obtained from analysis of explanted liver tissues and from publicly available RNA data from patients with HBV-related HCC (from International Cancer Gene Consortium, ICGC). The results imply that integrations are frequently expressed in HBV-infected human liver.

2 | METHODS

2.1 | Patients and liver tissue samples

Liver explants from five patients who underwent liver transplantation for HBV-related chronic liver disease were investigated, including two with HDV co-infection. Patients were selected to represent end-stage disease of chronic HBV infection, including cases with HDV co-infection, and included all liver transplanted HBV patients at our centre. All had liver cirrhosis and two HCC. Liver tissue was obtained directly after surgery and without delay split into multiple pieces that were stored in 1.5 mL tubes at −70°C until analysed. For the analyses in the present study, we used slices from the frozen tissue pieces that were approximately 5 µm thick with an area of 1 cm².

2.2 | Quantification of HBV DNA, HBsAg and HDV RNA in serum and liver tissue

HBV DNA and HBsAg levels in serum were quantified by Cobas TaqMan or Cobas 6800 (Roche Diagnostics) and by the Architect assay (Abbott), respectively. HDV RNA levels in serum were quantified by real-time PCR using primers HDV_F, GGATGCCCAGGTCGGAC and HDV_R, CCTCTTCCGGTTCCGAC, an MGB (minor groove binding) probe with a FAM fluorophore, ATCTCCACCTCCYCG, and a serial dilution of a plasmid carrying the target region as quantification standard.

2.3 | RNA extraction and library preparation

The liver tissue was homogenized as previously described. The RNA was extracted using the RNase Kit from Qiagen, and was, after confirmation of sufficient RNA integrity by TapeStation (Agilent Technologies Inc) analysis, processed by Eurofins/GATC for RNA-seq. Before library preparation of RNA, rRNA depletion was performed in order to enrich mRNA and other non-rRNA species. For one library from patient 1, mRNA was also enriched for poly-A transcripts for comparison with rRNA depletion.

Prior to strand specific paired-ends library preparation (TruSeq Stranded Total RNA Library Prep Kit, Illumina Inc), the extracted RNA was fragmented using sonication into approximately 350 nt long fragments and converted to cDNA using random primers. No further size selection was made prior to sequencing. All RNA extracted, and all products from the library preparation were sequenced in all samples. Samples from patients 1, 2 and 4 yielded four, two and two libraries, respectively, and the sequence reads from these libraries were later combined in bioinformatics analysis. Samples from patients 3 and 5 resulted in one library each.

2.4 | Transcriptome sequencing and data analysis

Paired-ends sequencing of cDNA libraries on Illumina HiSeq (Illumina Inc) was performed by Eurofins/GATC on all libraries from all samples. Forward and reverse read length was 150 nt. The reads were trimmed and paired prior to bioinformatics analyses. We aimed to sequence 200 million reads from each sample (100 million read pairs). The number of total reads for all included samples ranged from 131 575 678 to 413 030 404.

Transcriptome data were analysed using CLC Genomics Workbench (Qiagen) to (i) determine HBV read coverage and human gene expression, (ii) identify viral/human junction points and (iii) generate graphic profiles for the HBV RNA distributions. A customized bioinformatics pipeline for detection of all HBV reads and HBV/human fusion reads was developed and applied. After trimming and quality analysis of the reads, reads that mapped to an HBV reference genome were identified using Burrows-Wheeler aligner. Reads only partly mapping to HBV were detected using a softclip script and were aligned to the human reference genome (hg19) using BLAT. In addition, paired-end reads with one read mapping to HBV in its entirety and a paired mate not mapping to HBV were extracted and also mapped to hg19 using BLAT. HBV reads with the same junction points and pair mates less than 400 base pairs (bp) from each other in the human genome were considered to represent the same unique integration. RNA splicing of HBV reads was detected using a STAR based script, and by manual inspection of read mappings.
2.5 | Assessment of the proportion of putative integration-derived RNA

This estimation was based on the assumptions that all HBV RNA contains the X region and that integration-derived RNA does not contain the core region. It also assumes that precore RNA is much rarer than core RNA and that preS1 and X RNA are much rarer than preS2 RNA. Thus, the proportion of RNA that was integration-derived was calculated as follows: (average coverage of reads in X – average coverage of reads in core)/average coverage of reads in X.

2.6 | Analysis of publicly available data in the ICGC database

To expand the assessment of HBV transcriptome profiles, we analysed sequences retrieved from the LIRI-JP collection in the International Cancer Genome Consortium (ICGC) database [https://dcc.icgc.org/projects/LIRI-JP], which contains RNA data from Illumina sequencing liver tissue samples from patients with HCC. We retrieved RNA data from 21 patients that according to available metadata had chronic HBV infection (with tumour and non-tumour samples taken at time for resection or explantation) of which 37 samples were HBV RNA positive. The RNA sequences were analysed with the same bioinformatics pipeline as those from the liver explant patients.

2.7 | Ethics

The study was approved by the Regional Ethical Review Board in Gothenburg (registration number 835-17), and the patients gave informed oral and written consent to participate. This study is conforming according to the ethical guidelines of the 1975 Declaration of Helsinki.
3 | RESULTS

The two sets of HBV RNA sequence data—from explant tissue and from the ICGC database—were processed applying the same bioinformatics strategy that extracted HBV reads to obtain RNA profiles and to identify viral-human fusions.

3.1 | HBV RNA profiles

The HBV transcriptome profiles are shown in Figure 1 (explant liver tissue) and Figure 2 (ICGC database sequences). Patient 1 had relatively high HBV DNA levels in serum (5.55 log IU/mL) and a transcriptome profile with moderate coverage of the core region, and ten times greater coverage in the S and X regions, suggesting that more than 90% of the S and X RNA derived from integrations (Table 1).

Patient 2 had a low HBV DNA level in serum (2.09 log IU/mL) as a result of tenofovir treatment for four months. The total number of HBV reads was low (2108 reads) in non-tumour tissue, with coverage in core that was similar to that in S and X regions, indicating that RNA originated from cccDNA. By contrast, tumour tissue showed a high HBV coverage (316 033 reads) and a profile with a 1000-fold greater depth in S and X compared with the core region, indicating absence of HBV replication in the tumour, but abundant transcription of S and X RNA, most likely from integrated HBV DNA.

Patients 3-5 had received treatment with tenofovir or entecavir for at least two years and lacked detectable HBV DNA in serum at the time of transplantation. Since antivirals do not inhibit transcription, the HBV RNA profile might still contain expression of HBV genes from cccDNA or integrations. In patient 3, with end-stage HBV-induced cirrhosis, only 238 HBV RNA reads were detected, all in the S and X regions, indicating very low HBV transcription, presumably mainly from integrations. Patient 4, with HDV-induced cirrhosis, showed no detectable HBV DNA in serum but relatively high HBsAg levels (3.27 log IU/mL). A total of 4637 HBV RNA reads were found in liver tissue of which >99% were in the S and X regions. Likewise, patient 5 (HDV-induced cirrhosis and HCC) had no HBV DNA in serum and a transcriptome (in non-tumour tissue) with 3027 HBV reads, which were confined to S and X regions.

In theory, all RNA species from cccDNA should end at a common polyadenylation site at nt 1927. By contrast, RNA from

![FIGURE 2](image-url) A-E. A shows a merged profile based on all HBV RNA reads in the 21 patients of the ICGC data set mapped to the HBV reference genome. B-E shows four RNA profiles from two ICGC cases (tumour and non-tumour). The profiles in blue show the HBV RNA reads coverage (max coverage left of the graphs). The bars above each profile show HBV genomic positions of each integration point, and the height of the bar represents the number of HBV/human fusion reads. The graphs have different Y-axis scales.
integrated HBV DNA (preS1, preS2 or X mRNA) should extend no further than nt 1830, but could be shorter if the genome has been truncated during integration. Shorter transcripts can also be generated from both cccDNA and integrations if an upstream polyadenylation signal was used.\(^{25,26}\) As shown in Figure 1, transcripts that extended beyond nt 1830, indicating that cccDNA was the source, were found in patient 1 (the only patient with high HBV DNA in serum) and in non-tumour tissue from patient 2, that is only in samples that also contained significant amounts of core RNA.

The profiles of the HBV transcriptome in the ICGC data set were similar to those observed in explant tissue from patients. Figure 2A shows a compilation of all RNA reads from 21 patients with HBV-related HCC (tumour and non-tumour samples), and 2B-E exemplifies three of these cases. Individual data for all cases are presented in the Figure S1. In almost all cases, there were very low fractions of core region reads and absence of reads in the segment between nt 1830 and 1930, that is a lack of reads that represent RNA derived from cccDNA. The median number of HBV reads in these samples was 18 794 (range 11 568 244; IQR 53 109),

### TABLE 1

Characteristics and results of RNA deep sequencing of liver explant tissue from five patients

|                      | Patient 1 | Patient 2 | Patient 3 | Patient 4 | Patient 5 |
|----------------------|-----------|-----------|-----------|-----------|-----------|
| Age at transplantation | 50.5      | 54.8      | 34        | 53.3      | 47.2      |
| Sex                   | Male      | Male      | Male      | Male      | Male      |
| Geographic origin     | Balkan    | Middle East| East Africa| Middle East| Middle East|
| Antiviral treatment   | Lamivudine resistance\(^a\) | Tenofovir 4 mo | Tenofovir >2 y | Tenofovir 2 y | Entecavir |
| HBsAg serum (log IU/mL) | 3.79     | 3.61      | 1.00      | 3.27      | 2.86      |
| HBV DNA serum (log IU/mL) | 5.55     | 2.09      | Neg       | Neg       | Neg       |
| HDV RNA serum (log copies/mL) | 4.99   | 3.94      |           |           |           |
| Alpha-fetoprotein (μg/L) |          | 30        |           | 1         |           |
| RNA-seq. liver tissue | T         | NT        | T         | NT        |           |
| Total million reads in library | 249      | 366       | 132       | 205       | 413       |
| HBV reads (total)     | 38 544    | 316 033   | 2108      | 238       | 4637      |
| Max coverage HBV      | 5343      | 54 553    | 194       | 34        | 595       |
| Normalized RNA-seq liver tissue |          |           |           |           |           |
| Normalization quote\(^b\) | 0.80     | 0.55      | 1.52      | 0.97      | 0.48      |
| Max coverage HBV normalized | 4295   | 29 772    | 295       | 33        | 288       |
| Average coverage, core region normalized | 182      | 3         | 142       | 0         | 1         |
| Average coverage, S region normalized | 1254     | 6078      | 126       | 0         | 94        |
| Average coverage, X region normalized | 2854     | 11 431    | 175       | 0         | 190       |
| Max coverage, core region normalized | 462      | 9         | 226       | 0         | 3         |
| Max coverage, S region normalized | 1902     | 19 320    | 179       | 33        | 275       |
| Max coverage, X region normalized | 4295     | 19 092    | 295       | 19        | 152       |
| Putative fraction integration-derived HBV RNA (non-core region reads) | 94% | 100% | 19% | 100% | 99% N/A\(^c\) |
| Total number of fusion reads | 968      | 8227      | 4         | 16        | 235       |
| Number of unique fusion reads | 37       | 34        | 1         | 1         | 15        |
| Fusion reads fraction (out of all HBV reads) | 3% | 3% | 0% | 7% | 5% 0% |
| HDV reads (total)     | 1 227 776 | 557       | 9686      |           |           |
| Max coverage HDV      | 174 663   | 85        | 1447      |           |           |

Abbreviations: NT, non-tumour tissue; T, tumour.

\(^a\)M204V mutation in the reverse transcriptase region of HBV.

\(^b\)Normalized to standard library size of 200 million reads.

\(^c\)Not applicable because of low or no coverage.
and the median proportion of putative integration-derived RNA (based on coverage in X and core regions as described in methods) was 95% (range 24%-100%; IQR 13%). A different expression profile was found in the non-tumour sample from patient RK126 (Figure 2D), where HBV expression was restricted to the region representing the preS2 and S regions until position 454 where a large number of reads carried HBV-human fusion points in GTF2I gene transcripts.

3.2 Fusion reads

A more direct way to demonstrate expression of integrated HBV DNA is to identify fusion reads consisting of both viral and host RNA. As shown in Figure 1, fusion reads were observed in all patients, but not all samples. More than 99% of fusion reads were composed of a 5′ HBV part and a 3′ human part. The number of fusion reads differed markedly between the samples (details in Table S1), but the proportion of HBV reads that were fusion-derived was similar (range 0%-7%). The number of unique HBV/human fusion reads (with ≥2 reads coverage) ranged between 0 and 37. The tumour tissue in the sample with HBV-induced HCC (patient 2) contained a large number of fusion reads with the same junction point.

Overall, fusion reads were very frequent in the ICGC data set (range 0%-18% of all HBV reads), and almost all had an HBV 5′ part ending in the region 1750-1830 followed by a human sequence. All ICGC sample integrations are presented in Table S2. In addition to the integrations previously reported in the ICGC data set, our analysis detected many new unique HBV integrations and the total number of fusion reads was also higher.

3.3 Expression of human genes adjacent to HBV integrations

To explore the potential impact of HBV integration on the expression of human genes, we compared the human RNA data from our samples with published mRNA data for liver tissue from healthy individuals devoid of HBV infection. Most integrations were found in introns or intergenic areas and had no significant impact on the human gene expression; 12% were found in exons (details in Table S1). TERT expression was increased approximately 30 reads (and with ≥2 reads coverage) located near or upstream of nt 1830.

The ICGC data set had HBV integrations in TERT found in one tumour sample (RK010). HBV integrations were also found in MLL4 (KMT2B), strongly associated with liver cancer [5], in five tumour samples. Human RNA expression of ICGC samples has previously been published, showing marked overexpression of TERT and moderate overexpression of MLL4.

3.4 Reads representing HBV splicing or recombination

Spliced HBV RNA forms were found in six out of seven explant samples but represented less than 1% of the total HBV RNA. The most common spliced RNA showed ends joining between nt 2067 and nt 489, and was found in >2000 reads in the tumour sample from patient 2. The second most common was nt 2985-489, found in 55 reads in patient 1. In the ICGC data set, HBV splicing was not abundant and did not seem to affect the transcriptome profiles. Out of the >1 500 000 HBV reads in the ICGC data, only 306 reads represented splicing at previously reported donor/acceptor HBV splice sites, and the most common was nt 458-489 joining (28 reads in total). RNA reads suggesting recombination or deletions of (presumably integrated) HBV DNA rather than splicing were present in several samples both in the liver explants and in the ICGC data set, and the latter had a total of 47 178 reads indicating such events, representing 3% of all HBV reads.

3.5 Hepatitis D virus RNA

Analysis of HDV RNA reads in liver tissue was performed on the transcriptome data from patients 4 and 5. The maximum HDV RNA coverage was 174 663× in the sample from patient 4 (HDV RNA in serum 4.99 log copies/mL) and 1448× in non-tumour and 85× in tumour tissue sample from patient 5 (HDV RNA in serum 3.94 log copies/mL). Reads indicative of HDV infection were not found in any of the ICGC samples.

4 DISCUSSION

Three findings in this study argue that a large proportion of the HBV RNA in liver tissue, and indirectly much of HBsAg in serum, likely originates from integrated HBV DNA. First, most samples had a transcriptome profile with a much lower number of reads in the core region than in the S regions, indicative for a linear template that lacks a promoter upstream of the core gene. Second, reads representing the 3′ ‘redundancy’ between nt 1830 and 1927 were rare, in agreement with the expected absence of this part in RNA from integrated HBV DNA. Third, we detected a large number of fusion reads, that is RNA with a 5′ viral part followed by a 3′ human part, almost all located near or upstream of nt 1830.

By comparing the number of reads mapping to the core and to the S region, we estimated that in most cases more than 90% of the HBV RNA were from integrations. An alternative calculation, based on the number of fusion reads, estimated that 10%-70% of all HBV RNA was derived from integrations. The latter estimation was obtained by the finding that fusion reads constituted 1%-7% of all the HBV reads and the assumption (based on read length and genome size) that fusion reads should represent 10% of all HBV RNA from integrations. The lower rate of integrations indicated by the frequency...
of fusion rates might to some extent be explained by polyadenylation of some of the integration-derived RNA at an upstream human poly-A signal, because then they would not be identified as fusion reads originating from integrations. Another possible explanation is that either the HBV or human part of some fusion reads was too short to be mapped and BLAT-ed to the corresponding reference genome, and thus not be counted as an HBV integration. Despite these differences, both these analyses of RNA read counts indicate that expression of integrated DNA was significant and likely encoded most of the HBV RNA.

In a previous study in chimpanzees treated with an RNAi drug, similar HBV RNA profiles with few reads mapping in the core region were observed in three HBeAg-negative animals, and in one of these animal host-viral fusions were shown to be frequent by using single-molecule real-time sequencing. Our observations corroborate these results in humans and in a larger number of individuals. A recent study analysed liver tissue (non-tumour and tumour) from five patients, mainly by targeted amplifications of different HBV transcripts followed by traditional sequencing, and found that HBV RNA from cccDNA overall was in minority or lacking. 

In contrast, our study provides well-supported RNA profiles covering the whole HBV genome, allows quantitative estimates of the degree of integration and demonstrates a strong spatial link between HBV integrations breakpoints and the HBV transcriptome profile, as seen in Figures 1 and 2. The large number of identified fusion reads might be explained by our bioinformatics pipeline being more sensitive regarding detection of fusion reads compared to previous methods. For example, our pipeline included initial mapping to the HBV genome and included a script for discordant read pairs, that is, pairs with one read mapping to the HBV reference and the read mate solely to a human gene.

The liver explant patients represent different types of end-stage liver disease: three had HBV-induced liver cirrhosis, and two of these also had HCC. Patient 1 had a relatively high level of HBV DNA in serum, and significant amounts of RNA reads in core and the 3′ redundancy were found in liver tissue (indicating an origin from cccDNA). Yet, more than 90% of the total intrahepatic HBV RNA in this case were estimated to be derived from integrations. In patient 2, only 19% of the HBV RNA in non-tumour tissue, but 100% in tumour tissue, were estimated to be integration-derived. Patient 3 had been treated with tenofovir for several years and had a very low HBV DNA level in serum. Although the antiviral treatment has no direct effect on HBV RNA levels, there were very few HBV RNA reads and no reads at all in the core region, indicating that the antiviral treatment might still have resulted in a marked reduction of cccDNA as well as integrated HBV DNA. A reduction in cccDNA is in line with a previous study reporting that cccDNA had declined by >99% after only 2 years of adefovir treatment. Patients 4 and 5 were also on antiviral treatment and had no detectable HBV DNA in serum, but had relatively high levels of HDV RNA and HbsAg. In these cases, the transcriptome profiles almost completely lacked core region reads, suggesting that essentially all of the HBV RNA was integration-derived. The findings in patients 4 and 5, with high HDV RNA levels in serum and many HDV RNA reads in liver tissue in the absence of HBV core reads, suggest that HDV might replicate in hepatocytes that lack cccDNA and HBV replication, solely relying on HbsAg from HBV DNA integrations. This has to our knowledge not been observed in any previous clinical study, but is supported by in vitro data and experimental infection of humanized mice with HDV.

The potential oncogenic effect of HBV integrations was not in focus of this study. Notably, however, the tumour tissue from patient 2 contained 34 unique HBV integrations, of which a few predominated, probably as the result of mono/oligoclonal expansion. Also, the HBV integration with the second highest coverage was found in TERT, an oncogene previously associated with HBV-induced HCC. The expression of TERT was 1000-fold compared with healthy liver tissue. Interestingly, the tumour in patient 5 showed essentially no HBV reads, suggesting that in this case oncogenesis occurred independently of HBV integration, probably mainly as a result of HDV-induced inflammation.

Previous studies have described presence of spliced HBV RNA in liver tissue and serum, and some have reported associations between spliced forms and clinical stage or interferon treatment. Our pipeline searched for reads containing a splicing point or read pairs indicative of splicing using known splicing donor acceptor sequences, including previously described HBV splicing sites. Overall, we observed low number of reads representing spliced HBV RNA forms in the explants, <1% in the core region reads, suggesting that essentially all of the HBV RNA was derived from integration in the human CCNA2 gene, which has been suggested to have an oncogenic role. The spliced transcript in the CCNA2 gene was not detected in any of our samples, but non-spliced RNA from an integration in CCNA2 was detected in one tumour sample in the ICGC data set. Several samples contained reads with rearranged HBV RNA sequences that were not at previously reported splicing sites, but likely represent recombined integrations as previously observed.

The analysis of data from tumour and/or non-tumour samples from the 21 patients in the ICGC database showed transcriptome profiles that were very similar to those obtained in the liver explants, with few reads in the core region or beyond the typical fusion site near nt 1830. In most of the ICGC cases, the HBV/human fusions were located to positions between nt 1750 and 1830 in the HBV genome. However, in one non-tumour case the predominant fusion point was at position 454, and this was strikingly reflected in the HBV transcriptome profile, as shown in Figure 2.

In summary, these results support that integrated HBV DNA could be an important source of HbsAg in patients with late-stage...
chronic HBV infection, and possibly even more so in patients with HDV infection.\textsuperscript{12,13} Since HBsAg negativity is considered a requisite for cure, future HBV treatment may have to target HBsAg production from cccDNA as well as from HBV integrations.\textsuperscript{30} Further studies on the abundance, histological distribution and clinical importance of HBV integrations are warranted in all phases of infection.

ACKNOWLEDGEMENTS

The authors would like to thank the clinical contributors and the data producers of the International Cancer Genome Consortium who have provided data for the LIRI-JP data set\textsuperscript{29}. ICGC data were used in accordance with ICGC guidelines (https://icgc.org/icgc/goals-structure-policies-guidelines).

CONFLICT OF INTEREST

None of the authors has any conflict of interest.

ORCID

Johan Ringlander https://orcid.org/0000-0002-8483-6504
Helène Norder https://orcid.org/0000-0002-7528-3872
Magnus Lindh https://orcid.org/0000-0002-8729-7462

REFERENCES

1. GBD 2013 Mortality and Causes of Death Collaborators. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet. 2015;385:117–171.
2. Sherlock S, Fox RA, Niazi SP, Scheuer PJ. Chronic liver disease and primary liver-cell cancer with hepatitis-associated (Australia) antigen in serum. Lancet. 1970;1:1243–1247.
3. Hai H. Role of hepatitis B virus DNA integration in human hepatocarcinogenesis. W.V.G. 2014;20:6236–6239.
4. Sung W-K, Zheng H, Li S, et al. Genome-wide survey of recurrent HBV integration in hepatocellular carcinoma. Nat Genet. 2012;44:765–769.
5. Tang K-W, Alaei-Mahabadi B, Samuelsson T, Lindh M, Larsson E. The landscape of viral expression and host gene fusion and adaptation in human cancer. Nat Commun. 2013;4:2513.
6. Edman JC, Gray P, Valenzuela P, Rall LB, Rutter WJ. Integration of hepatitis B virus sequences and their expression in a human hepatoma cell. Nature. 1980;286:535–538.
7. Yoo S, Wang W, Wang Q, et al. A pilot systematic genomic comparison of recurrence risks of hepatitis B virus-associated hepatocellular carcinoma with low- and high-degree liver fibrosis. BMC Med. 2017;15:1–17.
8. Budzinska MA, Shackel NA, Urban S, Tu T. Cellular genomic sites of hepatitis B virus DNA integration. Genes (Basel). 2018;9:365.
9. Zhao L-H, Liu X, Yan H-X, et al. Genomic and oncogenic preference of HBV integration in hepatocellular carcinoma. Nat Commun. 2016;7:12992.
10. Furuta M, Tanaka H, Shiraishi Y, et al. Characterization of HBV patterns and timing in liver cancer and HBV-infected livers. Oncotarget. 2018;9:25075–25088.
11. Tu T, Budzinska M, Shackel N. Urban S. HBV DNA integration: molecular mechanisms and clinical implications. Viruses. 2017;9(4):75.
12. Wooddell CI, Yuen M-F, Chan HL-Y, et al. RNAi-based treatment of chronically infected patients and chimpanzees reveals that integrated hepatitis B virus DNA is a source of HBsAg. Sci Transl Med. 2017;9:eaan0241.
13. Cornberg M, Wong VW-S, Locarnini S, Brunetto M, Janssen HLA, Chan HL-Y. The role of quantitative hepatitis B surface antigen re-visited. J Hepatol. 2017;66:398–411.
14. Larsson SB, Malmsröm S, Hannoun C, Norkrans G, Lindh M. Mechanisms downstream of reverse transcription reduce serum levels of HBV DNA but not of HBsAg in chronic hepatitis B virus infection. Viral J. 2015;12:213.
15. Larsson SB, Elard A, Malmsröm S, et al. HBsAg quantification for identification of liver disease in chronic hepatitis B virus carriers. Liver Int. 2014;34:e238–e245.
16. Hu B, Wang R, Fu J, et al. Integration of hepatitis B virus gene impacts on hepatitis B surface antigen levels in patients with antiviral therapy. J Gastroenterol Hepatol. 2018;33:1389–1396.
17. Tripodi G, Larsson SB, Norkrans G, Lindh M. Smaller reduction of hepatitis B virus DNA in liver tissue than in serum in patients losing HBsAg. J Med Virol. 2017;89:1937–1943.
18. Pollicino T, Raffa G, Santantonio T, et al. Replicative and transcriptional activities of hepatitis B virus in patients coinfected with hepatitis B and hepatitis delta viruses. J Virol. 2011;85:432–439.
19. Freitas N, Lukash T, Gunewardena S, Chappell B, Slagle BL, Gudina SO. Relative abundance of integrant-derived viral RNAs in infected tissues harvested from chronic hepatitis B virus carriers. J Virol. 2018;92:672–739.
20. Tang K-W, Larsson E. Tumor virology in the era of high-throughput genomics. Philos Trans R Soc Lond B Biol Sci. 2017;372:20160265.
21. Larsson SB, Tripodi G, Raimondo G, et al. Integration of hepatitis B virus DNA in chronically infected patients assessed by Alu-PCR. J Med Virol. 2018;90:1568–1575.
22. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754–1760.
23. Kent WJ. BLAT—the BLAST-like alignment tool. Genome Res. 2002;12:656–664.
24. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15:21.
25. Schütz T, Kairat A, Schröder CH. DNA sequence requirements for the activation of a CATAAA polyadenylation signal within the hepatitis B virus X reading frame: rapid detection of truncated transcripts. Virology. 1996;223:401–405.
26. Su Q, Wang SF, Chang TE, et al. Circulating hepatitis B virus nucleic acids in chronic infection: representation of differently polyadenylated viral transcripts during progression to nonreplicative stages. Clin Cancer Res. 2001;7:2005–2015.
27. Fagerberg L, Hallström BM, Oksvold P, et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Mol Cell Proteomics. 2014;13:397–406.
28. Shiraishi Y, Fujimoto A, Furuta M, et al. Integrated analysis of whole genome and transcriptome sequencing reveals diverse transcriptional aberrations driven by somatic genomic changes in liver cancers. PLoS One. 2014;9(12):e114263.
29. Yang W, Summers J. Integration of hepadnavirus DNA in infected liver: evidence for a linear precursor. J Virol. 1999;73:9710–9717.
30. Lindh M, Rydell GE, Larsson SB. Impact of integrated viral DNA on the goal to clear hepatitis B surface antigen with different therapeutic strategies. Curr Opin Virol. 2018;30:24–31.
31. Wurstthorn K, Lutgehetmann M, Dandi M, et al. Peginterferon alpha-2b plus adefovir induce strong cccDNA decline and HBsAg reduction in patients with chronic hepatitis B. Hepatology. 2006;44:675–684.
32. Freitas N, Cunha C, Menne S, Gudima SO. Envelope proteins derived from naturally integrated hepatitis B virus DNA support assembly and release of infectious hepatitis delta virus particles. J Virol. 2014;88:5742–5754.
33. Giersch K, Bhadra OD, Volz T, et al. Hepatitis delta virus persists during liver regeneration and is amplified through cell division both in vitro and in vivo. Gut. 2019;68:150-157.
34. Tu T, Jilbert AR. Detection of hepatocyte clones containing integrated hepatitis B virus DNA using inverse nested PCR. Methods Mol Biol. 2017;1540:97-118.
35. Kawai-Kitahata F, Asahina Y, Tanaka S, et al. Comprehensive analyses of mutations and hepatitis B virus integration in hepatocellular carcinoma with clinicopathological features. J Gastroenterol. 2015;51:473-486.
36. Li C-L, Li C-Y, Lin Y-Y, et al. Androgen receptor enhances hepatic telomerase reverse transcriptase gene transcription after hepatitis B virus integration or point mutation in promoter region. Hepatology. 2019;69:498-512.
37. Günther S, Sommer G, Iwanska A, Will H. Heterogeneity and common features of defective hepatitis B virus genomes derived from spliced pregenomic RNA. Virology. 1997;238:363-371.
38. Bayliss J, Lim L, Thompson AJV, et al. Hepatitis B virus splicing is enhanced prior to development of hepatocellular carcinoma. J Hepatol. 2013;59(5):1022-1028.
39. Soussan P, Tuveri R, Nalpas B, et al. The expression of hepatitis B spliced protein (HBSP) encoded by a spliced hepatitis B virus RNA is associated with viral replication and liver fibrosis. J Hepatol. 2003;38(3):343-348.
40. Chen J, Wu M, Wang F, et al. Hepatitis B virus spliced variants are associated with an impaired response to interferon therapy. Sci Rep. 2015;5(1):1-13.
41. Chiu Y-T, Wong JKL, Choi S-W, et al. Novel pre-mRNA splicing of intronically integrated HBV generates oncogenic chimera in hepatocellular carcinoma. J Hepatol. 2016;64:1256-1264.

SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Ringlander J, Skoglund C, Prakash K, et al. Deep sequencing of liver explant transcriptomes reveals extensive expression from integrated hepatitis B virus DNA. J Viral Hepat. 2020;27:1162-1170. https://doi.org/10.1111/jvh.13356