OPEN ACCESS

Original research

Spatial transcriptomics reveals a low extent of transcriptionally active hepatitis B virus integration in patients with HBsAg loss

Xiaoqi Yu,[1] Qiming Gong,[2] Demin Yu ,[1] Yongyan Chen,[1] Ying Jing,[3] Fabien Zoulim ,[4,5,6] Xinxin Zhang [1,7]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/gutjnl-2023-330577).

For numbered affiliations see end of article.

Correspondence to
Professor Xinxin Zhang, Department of Infectious Diseases, Research Laboratory of Clinical Virology, Shanghai Jiao Tong University Medical School Affiliated Ruijin Hospital, Shanghai, China; zhangx@shsmu.edu.cn, Professor Fabien Zoulim, INSERM U1052, Cancer Research Center of Lyon (CRCL), Lyon, France; fabien.zoulim@inserm.fr and Dr Ying Jing, Center for Intelligent Medicine Research, Greater Bay Area Institute of Precision Medicine (Guangzhou), School of Life Sciences, Fudan University, Guangzhou, China; jingying@ipm-gba.org.cn

Received 4 July 2023
Accepted 17 October 2023
Published Online First
15 November 2023

To cite: Yu X, Gong Q, Yu D, *et al*. *Gut* 2024;**73**:797–809.

ABSTRACT

Objective Hepatitis B virus (HBV) can integrate into the chromosomes of infected hepatocytes, contributing to the production of hepatitis B surface antigen (HBsAg) and to hepatocarcinogenesis. In this study, we aimed to explore whether transcriptionally active HBV integration events spread throughout the liver tissue in different phases of chronic HBV infection, especially in patients with HBsAg loss.

Design We constructed high-resolution spatial transcriptomes of liver biopsies containing 13 059 tissue spots from 18 patients with chronic HBV infection to analyse the occurrence and relative distribution of transcriptionally active viral integration events. Immunohistochemistry was performed to evaluate the expression of HBsAg and HBV core antigen. Intrahepatic covalently closed circular DNA (cccDNA) levels were quantified by real-time qPCR.

Results Spatial transcriptome sequencing identified the presence of 13 154 virus-host chimeric reads in 7.86% (1026 of 13 059) of liver tissue spots in all patients, including three patients with HBsAg loss. These HBV integration sites were randomly distributed on chromosomes and can localise in host genes involved in hepatocarcinogenesis, such as *ALB, CLU* and *APOB*. Patients who were receiving or had received antiviral treatment had a significantly lower percentage of viral integration-containing spots and significantly fewer chimeric reads than treatment-naïve patients. Intrahepatic cccDNA levels correlated well with viral integration events.

Conclusion Transcriptionally active HBV integration occurred in chronically HBV-infected patients at different phases, including in patients with HBsAg loss. Antiviral treatment was associated with a decreased number and extent of transcriptionally active viral integrations, implying that early treatment intervention may further reduce the number of viral integration events.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Hepatitis B virus (HBV) integration, occurring early post-HBV infection and even in patients with a low viral load, contributes to hepatitis B surface antigen (HBsAg) expression and plays a key role in hepatocarcinogenesis. However, the in situ distribution and extent of HBV integration in patients with chronic HBV infection, even those with HBsAg loss are poorly understood.

WHAT THIS STUDY ADDS

⇒ Using spatial transcriptome sequencing, we demonstrated widespread transcriptionally active HBV integration in liver biopsies of chronically HBV-infected patients in different phases.

⇒ Antiviral treatment was associated with a reduced number and a lower extent of transcriptionally active viral integrations, and patients with HBsAg loss had fewer transcriptionally active HBV integration events.

⇒ Different patterns of intrahepatic expression of HBsAg, HBV core antigen and transcriptionally active HBV integration were observed, and covalently closed circular DNA levels correlated well with transcriptionally active viral integration events.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ This work implies that early treatment intervention may reduce the number of transcriptionally active viral integration events.

⇒ HBV integration into cancer-related genes in patients undergoing antiviral treatment suggests that hepatocellular carcinoma screening should be carefully maintained in these patients and the underlying mechanism needs to be further investigated.

INTRODUCTION

Chronic hepatitis B virus (HBV) infection is a global public health burden,[1] causing serious liver diseases including cirrhosis and hepatocellular carcinoma (HCC).[2] Loss of serum hepatitis B surface antigen (HBsAg), also considered as 'functional cure', is rarely achieved in chronic hepatitis B (CHB) patients despite long-term antiviral treatment.[3] Antiviral therapy can significantly suppress viral replication, yet the effects of long-term antiviral therapy on HBV integrations remain poorly understood.

During the HBV replication cycle, template-switching failure results in the formation of double-stranded linear HBV DNA (dslDNA),[4] which can integrate into the human genome. HBV integration is one of the major causative factors in virally

induced hepatocarcinogenesis, although the exact mechanism remains unclear.[5] In addition, integrated viral DNA may contribute to the production of HBsAg. However, the expression of other viral proteins, that is, polymerase, HBV e antigen (HBeAg) and HBV core antigen (HBcAg) from integrated HBV DNA does not occur in general, because their promoters are separated from the coding region after integration.[6] Previous studies have shown that HBV integration can be detected in the early phases of chronic HBV infection[7] and have suggested that HBV integration occurs at random sites in the host genome,[8] however, evidence regarding the number of virus-infected cells harbouring viral integrations is lacking.

The recently developed spatial transcriptomics (ST) technology, which provides high-quality genome-wide transcriptome data with intact positional information,[9] has been used to analyse spatial heterogeneity in human liver cancer[10] and other diseases. To get a comprehensive understanding of the intrahepatic features, we here used ST technology to investigate intrahepatic cell heterogeneity and the spatial distribution of transcriptionally active HBV integration events in chronically HBV-infected patients at different disease phases.

PATIENTS AND METHODS

Study population

According to the current EASL guidelines,[11] 18 patients with chronic HBV infection were recruited in this study from Ruijin Hospital from 22 September 2020 to 25 June 2021. All patients underwent a liver biopsy for histology assessment of necroinflammation activity and fibrosis stage, and tissue surplus to diagnostics were snap-frozen in optimum cutting temperature compound and stored at −80°C until subsequent use for spatial transcriptome sequencing. Matched serum samples were also collected.

Biochemical, virological and serological assessments

Blood biochemical parameters, including alanine aminotransferase (ALT) levels, were measured using an automated chemistry analyser (Beckman Coulter). Serum HBV DNA levels were quantified via real-time PCR (Shanghai KEHUA Bio-engineering) with a lower limit of detection of 50 IU/mL. Quantitative serum HBsAg levels and the presence of HBeAg and anti-HBe were measured using the Abbott Architect immunoassay system (Abbott Laboratories). HBV genotypes were determined by direct sequencing of HBV DNA extracted from serum samples (patients EP1-4and EN1-5) or liver tissues (patients SL1-3), or by a DNA-independent fluorescent lateral flow immunoassay (Beijing Wantai Biological Pharmacy Enterprise Co., Ltd.) with HBV genotype-specific monoclonal antibodies targeting epitopes in PreS2 (patients OT1-6).

Spatial transcriptome sequencing

Spatial transcriptome sequencing was performed using the Visium Technology Platform from 10X Genomics company. The raw sequencing data generated have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE243367). (dataset)[12] Liver biopsy tissues from two patients were fixed on one Visium Spatial Gene Expression Slide. For further detailed methods, refer to online supplemental information.

ST data processing

The Visium Spatial RNA-seq output and bright-field images were analysed using Space Ranger to detect tissue, align reads, generate feature-spot matrices, perform gene expression analysis, and place spots in spatial context on the slide image. For each spot, we quantified the number of genes and unique molecular identifiers (UMIs) and kept high-quality spots with 200–7500 genes, 1000–60 000 reads detected and no more than 50% mitochondrial gene counts. We removed specific mitochondrial genes, immunoglobulin genes, high abundance long intergenic noncoding RNA, and genes linked with poorly supported transcriptional models as previously described.[13]

Dimensionality reduction and clustering

After basic quality controls, R package Seurat[14] was used to integrate expression data from different sections of each patient. Principal components analysis (PCA) was performed to project spots into a low-dimensional space, defined by the first 30 PCs. These PCs were used to perform either the Barnes-Hut t-distributed stochastic neighbour embedding (tSNE) or uniform manifold approximation and projection (UMAP) for visualisation.

Cell type scoring by a signature-based strategy

We used a signature-based strategy to score cell type enrichments in each spot. We curated a set of gene signatures of 18 common stromal and immune cell types in the liver based on the xCell signatures and prior reference, encompassed normal hepatocytes, T cells, myeloid cells, fibroblasts, endothelial cells, CD8+central memory T cell (Tcm), CD8+ effective memory T Cell (Tem), CD4+T cells, CD4+naïve T cells, CD4+memory T cells, conventional dendritic cell (DC), active DC, monocytes, macrophages M1, macrophages M2, neutrophils, lymphatic (ly) endothelial cells, microvascular (mv) endothelial cells.[10] The corresponding cell type scores were determined as the average log-transformed normalisation expression values of the genes in each signature.

Identification of transcriptionally active HBV integrations

Virus-host chimeric reads were identified using the ChimericSeq package as previously described[15] with the human reference genome GRCh38/hg38 and a collection of HBV complete genome from the NCBI Nucleotide collection provided in the ChimericSeq package. The output from ChimericSeq was further processed manually. Reads that mapped exactly to the corresponding reference were selected for further analysis. Mapped reads aligned to multiple host loci were discarded. To avoid any possibility of false alignment and to ensure the reliability of HBV integration identification, chimeric sequences with >20 bp polyadenylation (polyA) at the end and sequences with >8 bp polyA at the end and >60% nucleobase A were filtered out. The remaining sequences were verified by Basic Local Alignment Search Tool (BLAST). Duplicate reads were removed to extract the unique chimeric reads. Because multiple reads often overlap with a single integration, chimeric reads with the same HBV coordinate, human coordinate, and the strand of reads was defined as a single viral integration, otherwise, they were considered as distinct viral integrations. The list of human tumour suppressor genes was obtained from the Tumour Suppressor Gene Database,[16] oncogenes from the ONGene database[17] and cancer mutated genes from the Catalogue of Somatic Mutations in Cancer (COSMIC database).[18]

DNA extraction from liver biopsies

DNA was extracted from snap-frozen human liver needle biopsies as previously described.[19] Briefly, liver samples were homogenised using a high-speed low-temperature tissue grinding

machine (Servicebio) in homogenisation buffer (Tris HCl pH 8 50 mM, EDTA 1 mM, NaCl 150 mM), then transferred to a microtube for DNA purification (MasterPure Complete DNA and RNA Purification Kit (Lucigen)) as recommended by the manufacturer. The PK digestion step was omitted for total DNA extraction. The quantity of extracted DNA was assessed using a NanoDrop Spectrophotometer. Before covalently closed circular DNA (cccDNA) amplification, DNA was treated with 30U of Plasmid-safe DNase (Lucigen) for 2 hours at 37°C.

Quantification of total HBV DNA, cccDNA in liver samples

Quantification of intrahepatic total HBV DNA and cccDNA was performed as previously described,[19] using the SLAN 96S real-time PCR system (Shanghai Hongshi Medical Technology) with primers and probes specific for total HBV DNA or cccDNA as listed in online supplemental table S1. Serial dilutions of a plasmid containing the 1.3-fold HBV genome served as the quantification standard. To normalise the number of viral copies per cell content, the number of cellular genomes was determined using housekeeping beta-actin. Negative control was included in each real-time qPCR reaction to exclude sample contamination.

Immunohistochemistry staining

HBsAg and HBcAg immunohistochemistry (IHC) were performed on adjacent slices of the same embedded tissue for STs. Tissue slices (5 µm) were incubated with 0.3% H_2O_2 at room temperature for 30 mins and then incubated with primary HBsAg and HBcAg antibody (Origene) at 37°C for 60 min. After washing, the targets were incubated with a peroxidase-conjugated secondary antibody for 30 min and detected using the DAB detection kit. The slices were counterstained with haematoxylin II, followed by a bluing reagent. The sections were then dehydrated, cleared and mounted permanently.

Statistical analysis

Continuous variables that were not normally distributed were presented as medians (IQR). Categorical variables were described as counts (%). Values were compared using the Mann-Whitney U test when appropriate. Linear regression analysis was performed to determine the factors associated with transcriptionally active HBV integration events. R^2 was calculated as a measure of the goodness of fit of the linear model. Pearson correlation coefficients (r) were calculated for linear correlations. Graphs were plotted using R or GraphPad Prism software. Statistical analyses were performed using SPSS V.24.0 software. A two-sided $p<0.05$ was considered statistically significant.

Patient and public involvement

Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

RESULTS

Characteristics of the study population

As described in Methods section, we collected liver biopsy tissue from 18 patients with chronic HBV infection. Characteristics of the study population are summarised in online supplemental table S2. Among them, 10 were male and 8 were female, with a median age of 39.50 (37.25–45.00) years. There were nine patients with HBV genotype B and nine with genotype C. The population comprised untreated patients (n=9), on-treatment (OT) patients without HBsAg loss (n=6) and patients with HBsAg loss (n=3) at the time of liver biopsy, including two patients were still receiving antiviral treatment, and one patient had already stopped long-term antiviral treatment.

Among the nine untreated patients, four were HBeAg-positive (EP) patients named EP1-4, and five were HBeAg-negative (EN) patients named EN1-5. According to the current EASL guidelines,[11] patient EP1 fall into an indeterminate grey area, and the others were in the HBeAg-positive CHB (patients EP2, EP3 and EP4), HBeAg-negative chronic HBV infection (patient EN1) and HBeAg-negative CHB (patients EN2-5) disease phases, respectively. Among these nine patients, six were treatment naïve patients, and patients EP1, EN4 and EN5 had a history of antiviral treatment. In the OT group, six patients with CHB were receiving antiviral treatment for more than 1 year at the time of liver biopsy, and they were named OT1-6. Patients OT1, OT3 and OT6 were HBeAg-positive, whereas patients OT2, OT4 and OT5 were HBeAg-negative. The remaining three patients who had already achieved HBsAg loss (SL) after long-term antiviral treatment were named SL1-3. At the time of liver biopsy, patients SL1 and SL2 were still receiving NUC treatment, while patient SL3 had already stopped antiviral treatment. At the time of enrolment in this study, patient OT5 and OT6 had already undergone one liver biopsy before antiviral treatment, whereas others had not.

Spatial intrahepatic cell heterogeneity

We used ST sequencing to analyse the intrahepatic cell heterogeneity. Each section contains liver biopsy tissues from two patients (online supplemental figure S1). The median sequencing depth of a single spot was approximately 3882 UMIs and 1346 genes in this study (online supplemental table S3). After data processing, 13 059 tissue spots from 18 patients were obtained, and we combined these spots across different sections for dimensionality reduction and clustering. The distribution of the clusters was presented in UMAP/tSNE projection space (figure 1A) and tissue physical space (figure 1B), and we found that the clusters were spatially intertwined.

The spot diameter on the Visium Spatial Gene Expression Slide in this study was 55 µm, which means each spot may contain multiple cells. As the current ST technology is unable to provide sufficient accuracy at a single-cell scale, a signature-based strategy was used to score cell type enrichments in each spot, and scores across clusters were compared. The normal hepatocyte scores were higher in cluster 3 and lower in cluster 4, with little difference in the remaining clusters. Monocytes, as well as tissue-resident macrophages, described as the immunological sentinels of the liver, were significantly enriched in cluster 5 (figure 2A). Fibroblasts, liver endothelial cells and T cell subtypes scored higher in cluster 4, and neutrophils scored higher in cluster 6. Notably, the composition of cell clusters varied among patients, with a notably higher abundance of cluster 5 in patients EP2 and OT6 (figure 2B), reflecting different stromal and immune cell infiltration degrees.

HBV integration derived virus-host chimeric mRNA expression

To gain high resolution on the spatial distribution of transcriptionally active HBV integration, we used the ChimericSeq package to capture virus-host chimeric reads in 10 ×Visium ST data. Overall, 13 154 virus-host chimeric reads were detected in the liver biopsy tissues of all patients (table 1), with a median (IQR) chimeric reads number of 115 (72–679.5) per patient. Of note, chimeric reads were also detected in three patients who achieved HBsAg loss, although in lower numbers. Although the difference was not statistically significant, the median number

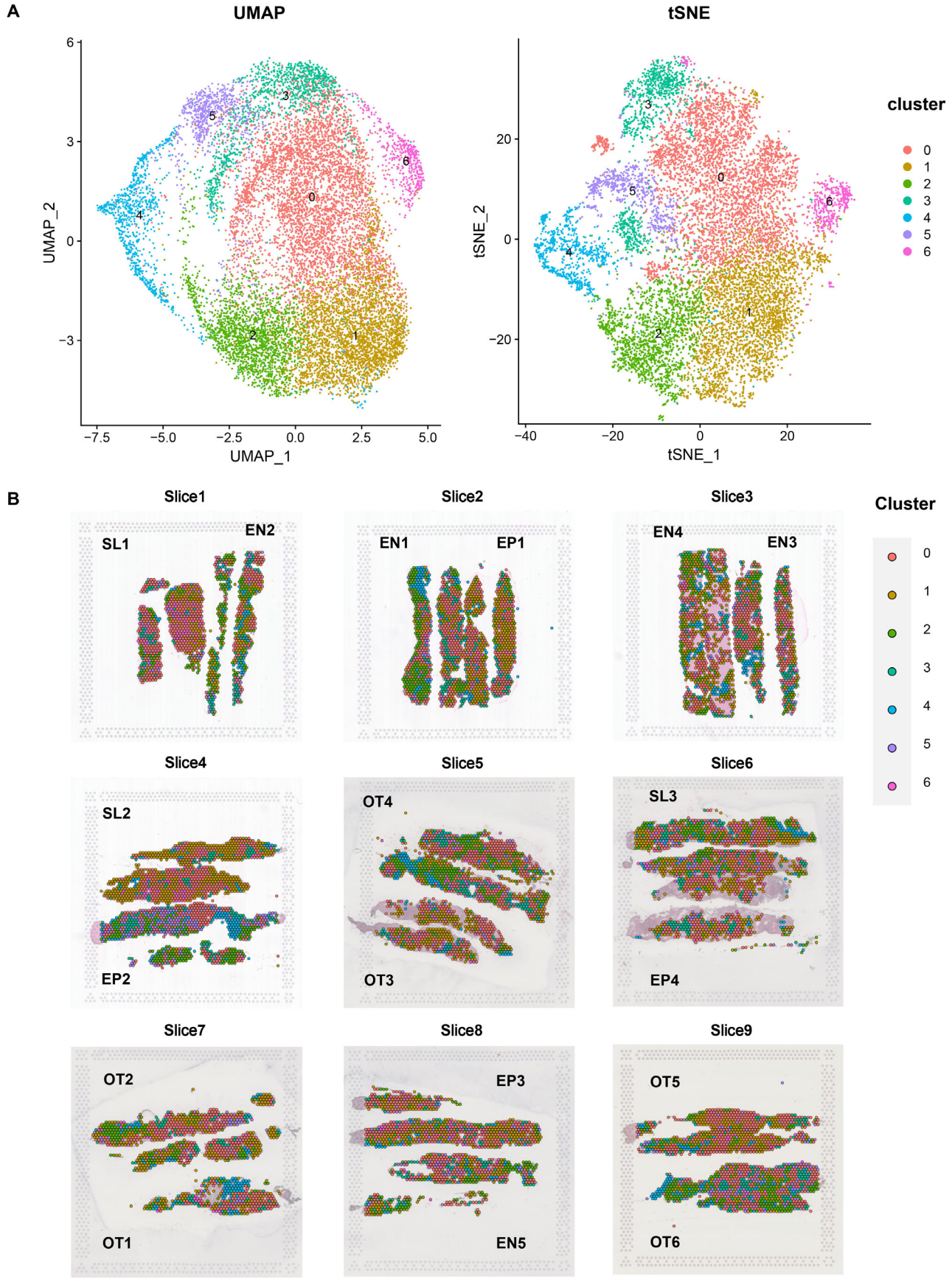

Figure 1 Spatial heterogeneities in patients with chronic HBV infection. (A) UMAP (uniform manifold approximation and projection) and tSNE (t-distributed stochastic neighbour embedding) plots of spots from all sections are coloured according to their cluster identities. (B) Spatial cluster distribution in each section. EN, HBeAg-negative; EP, HBeAg-positive; HBsAg, hepatitis B surface antigen; HBV, hepatitis B virus; OT, on-treatment; SL, HBsAg loss.

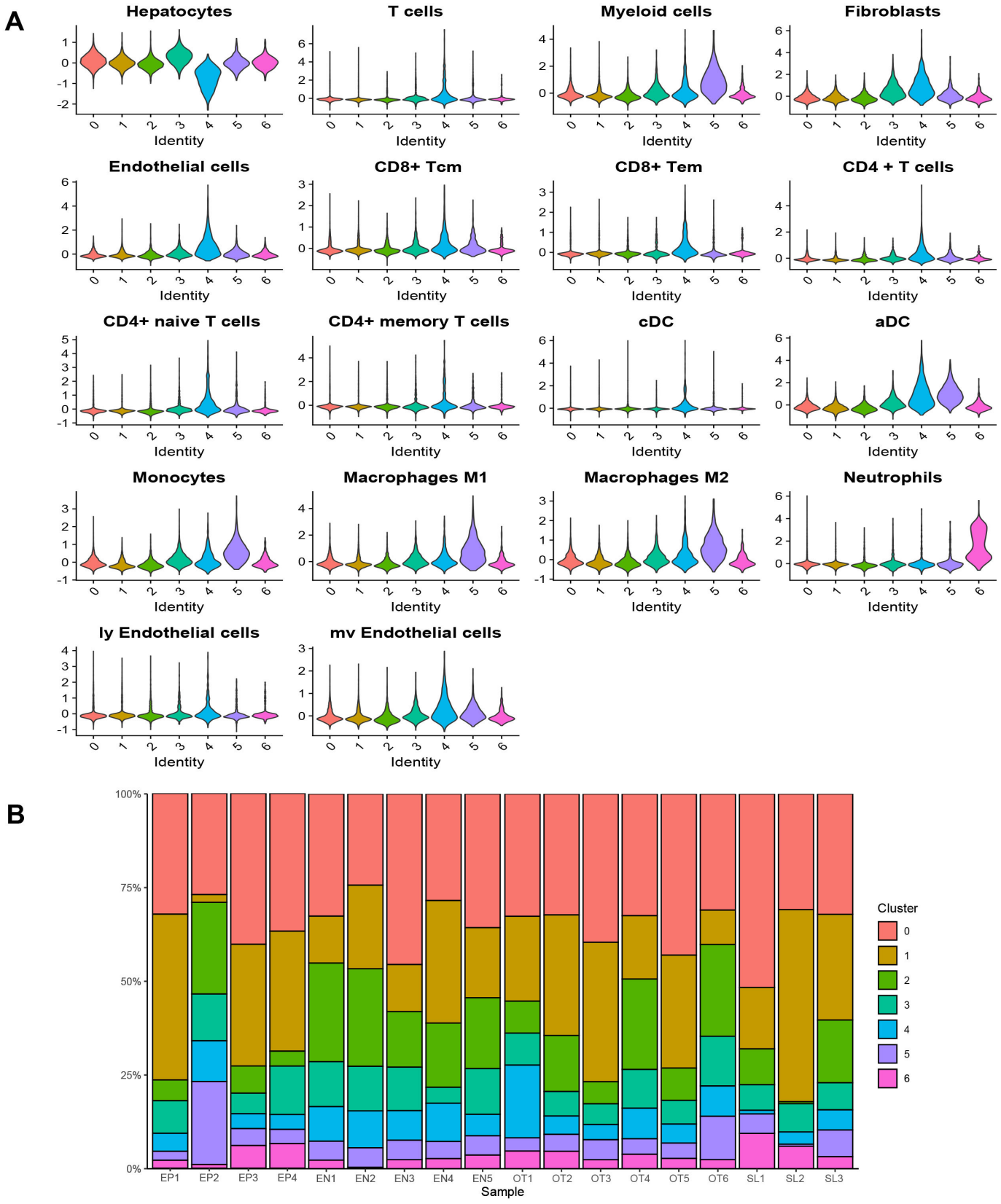

Figure 2 Cluster comparison and the cell type scoring. (A) Violin plots of the 18 stromal and immune cell type scores in each cluster. (B) Comparison of the cluster fraction for each patient. aDC, active dendritic cell; cDC, conventional dendritic cell; EN, HBeAg-negative; EP, HBeAg-positive; HBsAg, hepatitis B surface antigen; ly, lymphatic; mv, microvascular; OT, on-treatment; SL, HBsAg loss; Tcm, central memory T cell; Tem, effective memory T cell.

Table 1 Quantification of intrahepatic HBV integrations, total HBV DNA and cccDNA levels

Patient-ID	Percentage of integration-containing spots	Total chimeric reads	Unique chimeric reads	Distinct integration	Total HBV DNA (copies/cell)	cccDNA (copies/cell)
Untreated						
EP1	3.598	129	91	13	5.138	0.009
EP2	31.174	3979	1574	348	NA	NA
EP3	9.452	665	396	92	NA	NA
EP4	3.579	87	71	31	NA	NA
EN1	4.551	287	193	29	4.180	0.137
EN2	35.502	2110	995	82	19.071	0.688
EN3	4.539	723	268	27	10.636	0.061
EN4	0.112	81	25	3	NA	NA
EN5	0.954	45	24	4	1.075	0.137
On treatment						
OT1	2.647	101	65	9	4.586	0.253
OT2	0.126	1	1	1	NA	NA
OT3	2.762	378	135	9	19.120	0.087
OT4	1.784	460	209	13	6.905	0.059
OT5	4.247	90	64	25	NA	NA
OT6	34.620	3903	1350	105	6.218	0.284
HBsAg loss						
SL1	1.623	18	14	11	3.300	UD
SL2	1.745	81	59	18	1.211	0.032
SL3	0.103	16	10	1	0.977	0.029

cccDNA, covalently closed circular DNA; EN, HBeAg-negative; EP, HBeAg-positive; HBV, hepatitis B virus; NA, not available; OT, on-treatment; SL, HBsAg loss; UD, undetectable.

of chimeric reads was higher in untreated patients than in OT patients (287 vs 90), and was higher in HBeAg-positive patients than in HBeAg-negative patients (332.5 vs 88.5). Significantly fewer virus-host chimeric reads were observed in patients who were receiving or had received antiviral treatment compared with treatment-naïve patients (p=0.023), with median chimeric reads number of 85.5 and 694, respectively. The highest number of chimeric reads was 3979, detected in an HBeAg-positive treatment-naïve patient EP2. A high level of HBV integration was also observed in an HBeAg-positive, NUC-treated patient OT6, with 3903 chimeric reads detected. Notably, patient OT6 had relatively high levels of circulating biomarkers of HBV replication activity before receiving antiviral treatment, including HBV DNA, HBsAg and HBeAg levels.

We analysed the spatial distribution of spots containing transcriptionally active viral integration. The results revealed transcriptionally active HBV integration events in 7.86% (1026 of 13 059) of the liver tissue spots, and all these spots are highlighted in red (figure 3A). The median percentage of spots with transcriptionally active viral integration was 3.17% (1.46–5.78) per patient, and the highest percentage was 35.50% in an HBeAg-negative treatment naïve patient EN2. Patient OT6, mentioned above with a higher number of chimeric reads, also had a higher percentage of viral integration-containing spots than other antiviral-treated patients. Likewise, the percentage of viral integration-containing spots was significantly lower in patients who were receiving or had received antiviral treatment (1.76%) compared with treatment-naïve patients (7.00%) (p=0.005). Patients who were receiving antiviral treatment had a relatively lower percentage of spots with transcriptionally active viral integration than untreated patients (1.78% vs 4.54%). Spots with transcriptionally active viral integration could barely be detected in patients with HBsAg loss (figure 3B). We also calculated the percentage of viral integration-containing spots in each cluster. As expected, cluster 3, relatively enriched in hepatocytes, had the highest proportion of viral integration-containing spots (14.37%), while cluster 4, relatively deficient in hepatocytes, had a relatively lower percentage of viral integration-containing spots (figure 3C).

Moreover, those viral integration-containing spots showed diverse chimeric reads densities, with the maximum number of total chimeric reads being 441 in a spot from patient EP2 (figure 3D). We then filtered out duplicate reads to extract unique chimeric reads, and further quantified the number of distinct transcriptionally active viral integrations. As shown in online supplemental figures S2 and S3, patient EP2 also has the maximum number of unique chimeric reads (120 per spot) and distinct integration events (10 per spot). The number of total chimeric reads (r^2=0.2425, p=0.0446), unique chimeric reads (r^2=0.2422, p=0.0448), and distinct viral integration events (r^2=0.2864, p=0.0268) were significantly but not strongly correlated with serum HBsAg levels (figure 4). The number of distinct viral integration events also correlated with serum HBV DNA levels (r^2=0.2617, p=0.0300). However, there was no significant correlation between the number of chimeric reads and patients' age, known duration of infection, inflammation grade or fibrosis stage.

Differential localisation of transcriptionally active HBV integration events

A total of 5498 unique chimeric reads were detected, with 95.43% (5427 of 5498) representing positive-strand HBV DNA. Additionally, 92.96% (5111 of 5498) of these unique chimeric reads were derived from RNAs initiated in the integrated viral genome, extending into the host genome. Furthermore, the viral breakpoints for 90.81% (4993 of 5498) of these unique chimeric reads were mapped to the HBV sequence between nucleotide positions 1590 and 1840, corresponding to the viral genomic region spanning direct repeat 2 (DR2) and direct repeat 1 (DR1),

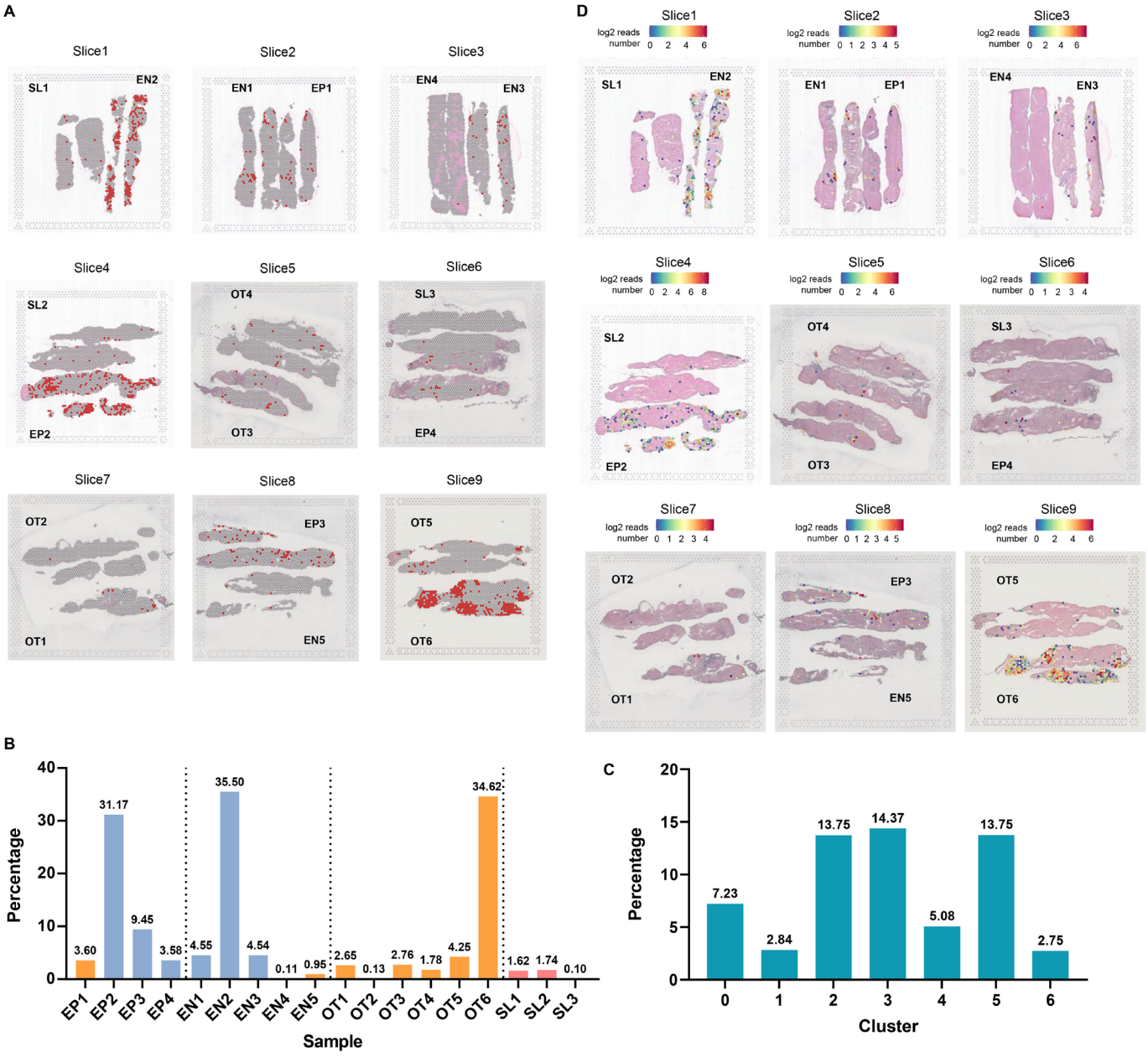

Figure 3 Spatial distribution of transcriptionally active HBV integration. (A) Spatial plot showing the distribution of transcriptionally active HBV integration in each section. Spots containing viral integration are highlighted in red. (B) The percentage of spots containing viral integration across patients who were receiving or had received antiviral treatment (orange), treatment-naïve patients (blue) and patients with SL (red). (C) The percentage of spots containing viral integration across all clusters. (D) The density of total chimeric reads detected in each spot was displayed. EN, HBeAg-negative; EP, HBeAg-positive; HBsAg, hepatitis B surface antigen; HBV, hepatitis B virus; OT, on-treatment; SL, HBsAg loss.

the region contains the enhancer II, known to upregulate viral gene expression. Few transcriptionally active HBV integration events were also observed in the open reading frame preC/Core (ORF C), ORF P, and ORF P region overlapping with ORF S (figure 5A).

A total of 783 distinct, transcriptionally active viral integrations were detected. Since Visium Spatial RNA-seq captures polyadenylated mRNA for cDNA library preparation, viral integration into untranscribed portions of the genome was not detected and only transcriptionally active HBV integration with the potential integrated HBV sequences were identified. Transcriptionally active HBV integration events were detected in cellular chromosomes without preferential hotspots (figure 5B), and were localised in 349 different human genes. Host alignments of integrated sequences fell on CDS, exon, stop codon, 3'UTR, and transcript as annotated by the loaded human GTF file. The most frequent genes associated with transcriptionally active HBV integration included *ALB* (a commonly reported HCC-related gene) and *CLU* (a tumour suppressor gene). In addition, *APOB* (a frequently reported HCC-related gene) was also found to be associated with transcriptionally active viral integration. Of the 349 host genes, 6 were oncogenes, 13 were tumour suppressor genes and 9 were known cancer-related gene (COSMIC database).[18] Notably, some of the highly detected host genes associated with transcriptionally active viral integration in patient OT6 were cancer-related genes, such as *ALCAM*, *MBL2* and *SH2D5*.

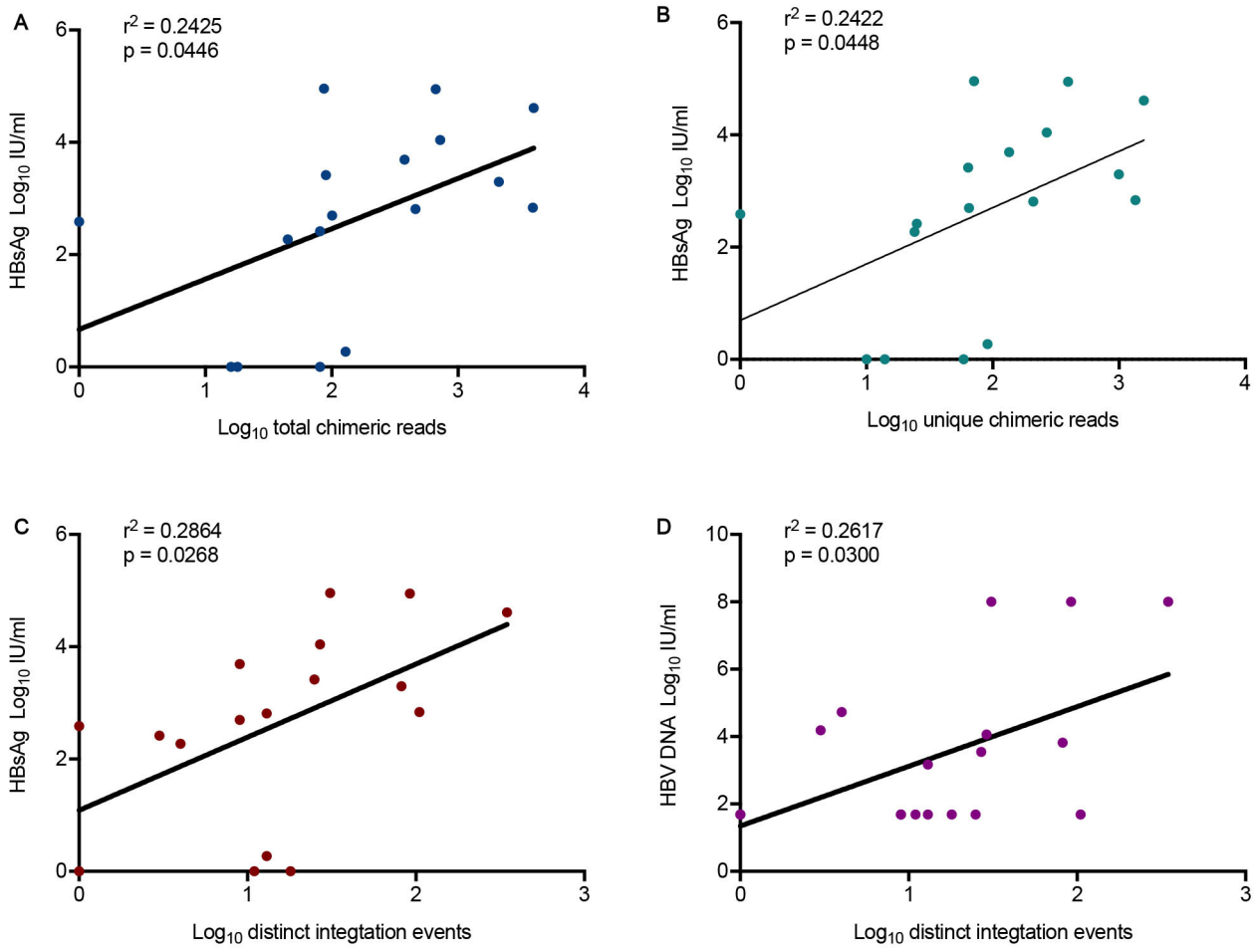

Figure 4 Correlation of transcriptionally active HBV integration events to serum HBsAg and HBV DNA level. Correlation of (A) total chimeric reads with serum HBsAg levels, (B) unique chimeric reads with serum HBsAg levels, (C) distinct viral integration events with serum HBsAg levels and (D) distinct viral integration levels with serum HBV DNA levels. Correlations were calculated using linear regression. The corresponding R^2 and p values are shown. HBsAg, hepatitis B surface antigen; HBV, hepatitis B virus.

Heterogeneously expressed intrahepatic HBsAg and HBcAg

To verify whether transcriptionally active viral integration events were consistent with HBsAg positivity in liver sections, the relative distribution of HBsAg was analysed where adequate liver sections were available for IHC. HBcAg IHC, reflecting cccDNA directed HBV expression more accurately, was also performed. Specificity of the HBsAg IHC signal was confirmed by the absence of any IHC signal in the liver of an HBsAg loss patient as a negative control. Generally, patients with higher serum HBsAg levels tended to have more HBsAg-expressing hepatocytes. Heterogeneous expression of HBsAg and HBcAg, alongside transcriptionally active viral integration were detected. Notably, patient EP3 exhibited widespread HBsAg and HBcAg positivity, whereas this patient had 9.45% viral integration-containing spots (figure 6), suggesting that transcriptionally active integration events did not occur in all infected cells. Patient OT3 with extensive HBsAg expression and highly expressed HBcAg had only 2.76% viral integration-containing spots (figure 6), suggesting that HBsAg might be mainly derived from intrahepatic cccDNA. Besides, patient EN2 with 35.5% viral integration-containing spots had a liver zone where HBsAg expression was strongly positive and transcriptionally active viral integration was detected, whereas HBcAg was negative (figure 6), suggesting that infected hepatocytes might harbour transcriptionally silenced cccDNA. Patient SL1 and SL3 who had achieved HBsAg loss were both immunohistochemically negative for HBsAg (figure 6 and online supplemental figure S4), but the presence of integration-containing spots was observed, although with a relatively lower percentage, indicating that these viral integrations might be 'silent' for HBsAg expression.

Intrahepatic HBV cccDNA contribute to HBsAg expression

We quantified the intrahepatic total HBV DNA and cccDNA in 12 of the 18 patients with available second piece of biopsy tissue. Except for patient SL1, HBV cccDNA was detected in the remaining 11 patients with low or moderate serum HBV DNA levels (table 1), ranging from 0.009 to 0.688 copies per cell. Notably, the lowest cccDNA level was detected in patient EP1 (0.009 copies/cell), who had a history of pegylated-interferon treatment. The highest cccDNA levels (0.688 copies/cell) was detected in patient EN2 with a high frequency of HBsAg-expressing hepatocytes and a high level of transcriptionally active HBV integration. As shown in figure 7, the intrahepatic cccDNA level correlated well with the percentage of integration-containing spots (r^2=0.6841, p=0.0009), total chimeric reads (r^2=0.3752, p=0.0342), unique chimeric reads (r^2=0.5123, p=0.0089) and distinct integration events (r^2=0.5106, p=0.0090).

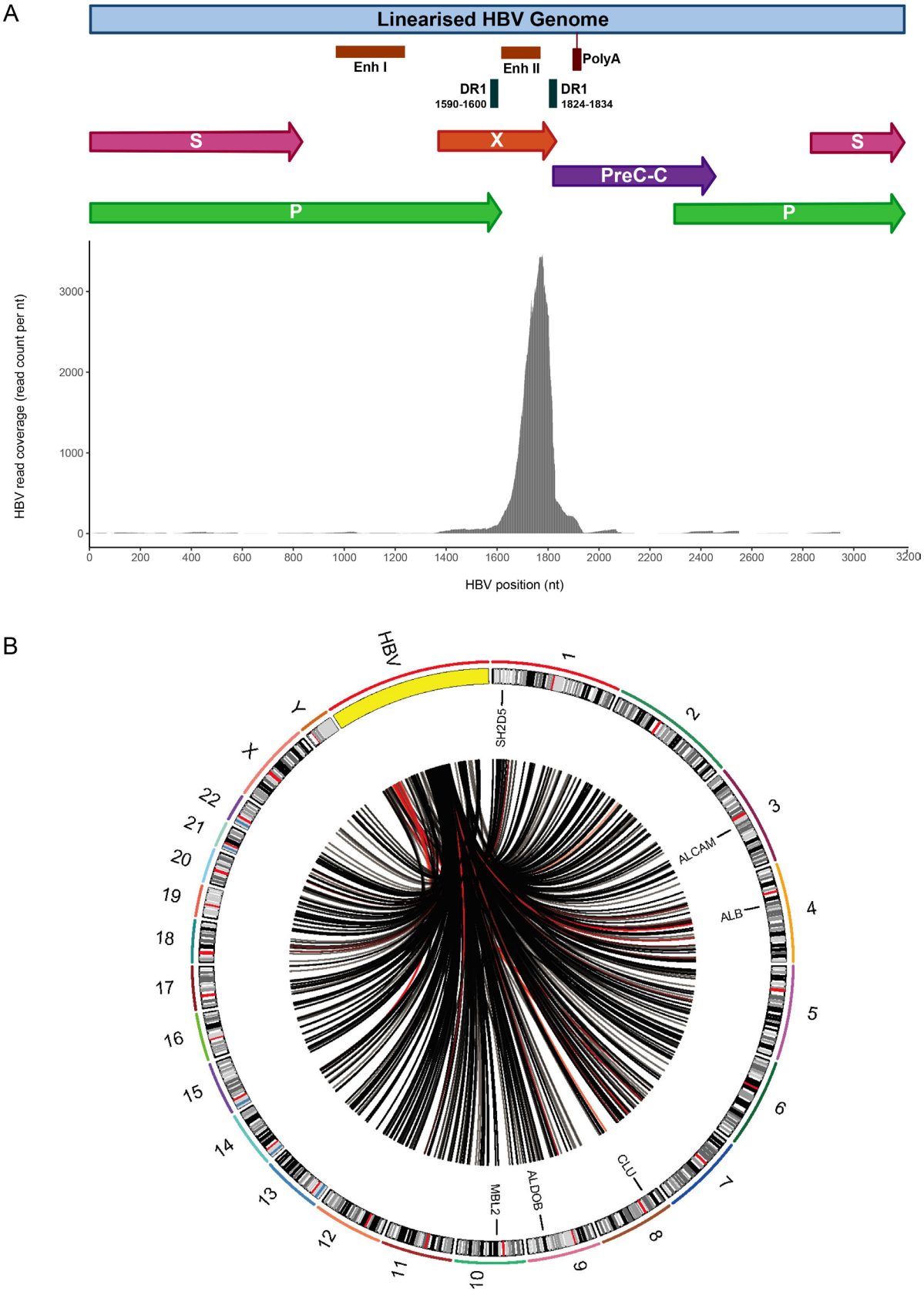

Figure 5 Visualisation of the transcriptionally active HBV integration sites on human chromosomes and the viral genome. (A) Histogram shows the frequency of transcriptionally active viral integration at each nucleotide position in the HBV genome. The locations of EnhI, EnhII, DR1, DR2 and the genes encoding HBV polymerase (green), core (violet), S (pink) and X (red) proteins are shown. Enh: enhancer; DR: direct repeat. (B) Circos plot showing host genome affected by viral integration. Each line connects the locations of breakpoints in the HBV genome to the corresponding loci in the human genome. The breakpoints in patients with HBsAg seroconversion are shown in red. HBsAg, hepatitis B surface antigen; HBV, hepatitis B virus.

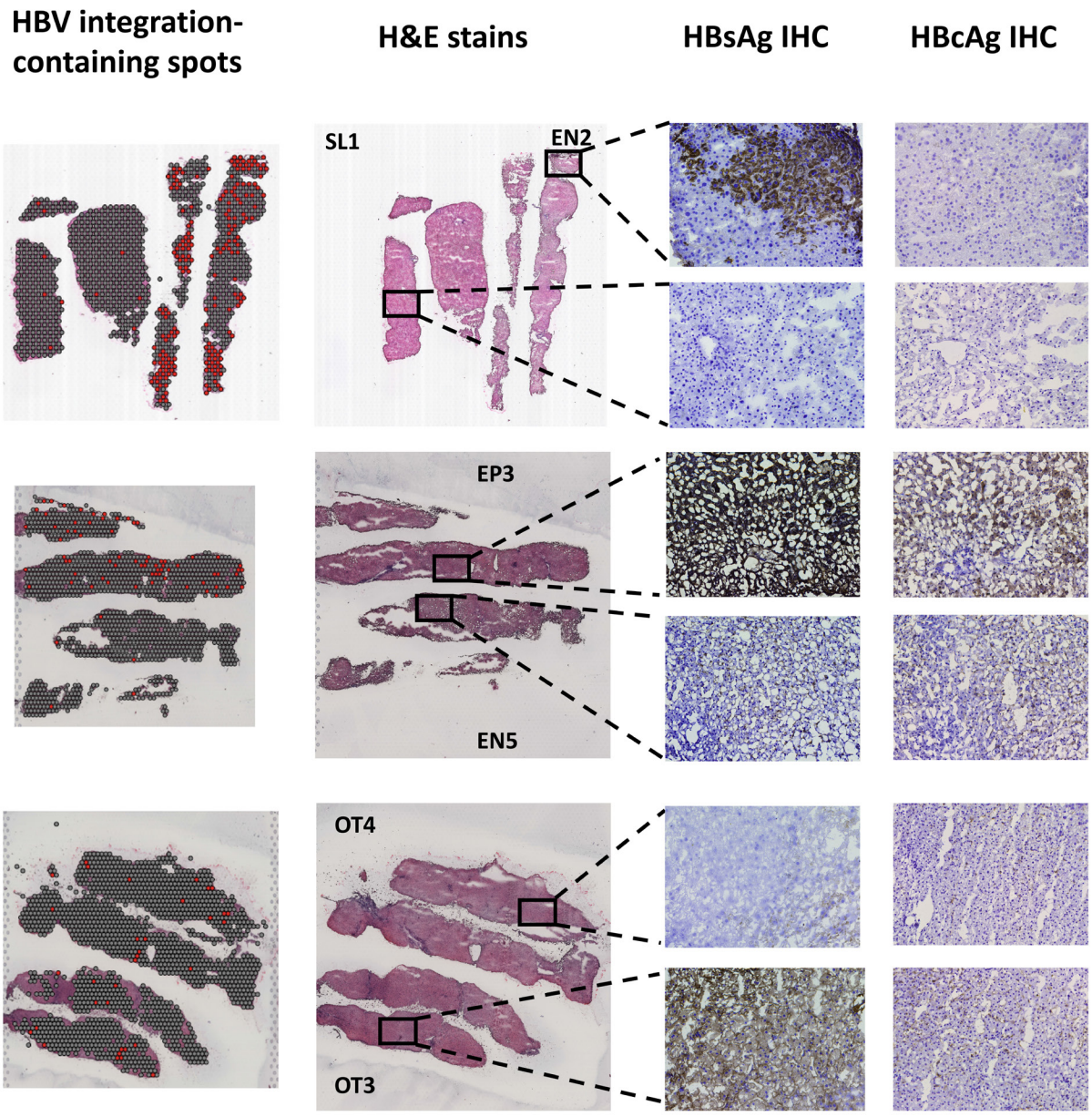

Figure 6 Intrahepatic HBsAg and HBcAg were heterogeneously expressed. HBsAg and HBcAg IHC were performed on adjacent slices from the same OCT-embedded samples of patients with adequate liver sections. The IHC results are displayed with matched H&E staining and spatial plots showing the distribution of transcriptionally active HBV integration. Spots containing viral integration are highlighted in red. Brown, HBsAg or HBcAg; blue, nuclei. EN, HBeAg-negative; EP, HBeAg-positive; HBsAg, hepatitis B surface antigen; HBV, hepatitis B virus; IHC, immunohistochemistry; OCT, optimum cutting temperature; OT, on-treatment; SL, HBsAg loss.

Intrahepatic total HBV DNA, including both HBV DNA replicative intermediates and integrated HBV DNA, exceeded one copy per cell in patients with intrahepatic HBsAg expression, regardless of low viraemia, while patients with HBsAg loss displayed relatively lower total HBV DNA levels. Patient OT3 with the ubiquitous presence of HBsAg had the highest intrahepatic total HBV DNA levels (19.120 copies/cell) and a moderate cccDNA level of 0.087 copies per cell. Intrahepatic total HBV DNA levels were not significantly correlated with the number of chimeric reads and distinct integration events, but were correlated with serum HBsAg levels (r^2=0.4901, p=0.0165).

DISCUSSION

In this study based on liver biopsies from 18 patients with chronic HBV infection, we explored whether actively transcribed HBV integration events could spread throughout the liver tissue. Using the emerging ST technology which could provide both cell spatial location and transcript information, we demonstrated the

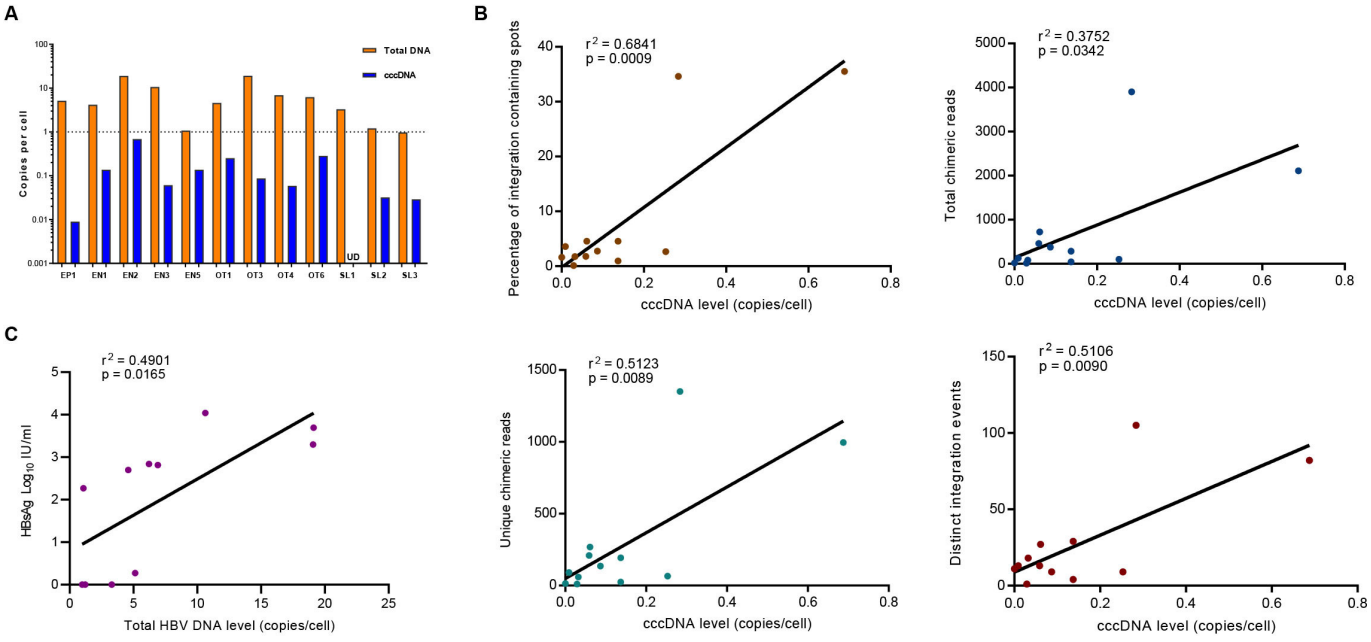

Figure 7 Quantification of intrahepatic total HBV DNA and cccDNA levels and their correlation to transcriptionally active HBV integration events. (A) Quantification of intrahepatic total HBV DNA (orange) and cccDNA (blue) levels in 12 patients. (B) Correlation of intrahepatic cccDNA levels with the percentage of integration-containing spots, total chimeric reads, unique chimeric reads, and distinct viral integration events. (C) Correlation of intrahepatic total HBV DNA level with serum HBsAg levels. Correlations were calculated using linear regression. The corresponding R^2 and p value are shown. cccDNA, covalently closed circular DNA; EN, HBeAg-negative; EP, HBeAg-positive; HBsAg, hepatitis B surface antigen; HBV, hepatitis B virus; OT, on-treatment; SL, HBsAg loss; UD, undetectable.

presence of transcriptionally active viral integration at different phases of chronic HBV infection, even in patients who achieved functional cure. At the current Visium ST resolution, each spot may contain multiple cells, precluding the assignment of a specific cell type to each spot. Therefore, we took advantage of a signature-based strategy to compare cell type enrichments across different clusters. The results showed that stromal and immune cell scores varied among clusters, and the composition of cell clusters exhibited a high degree of diversity, reflecting diverse stromal and immune cell infiltration degrees among patients. In addition to cluster 3, which was enriched in hepatocytes, cluster 5, which was enriched in monocytes and macrophages, also had a relatively higher proportion of viral integration-containing spots. Despite using this approach, definitive confirmation of cell type presence or absence in a cluster remains challenging, and based on these findings, the possibility of viral integration into cell types other than hepatocytes cannot be ruled out. Further studies are needed to clarify this point using the emerging single-cell and spatial multiomics technologies.

HBV DNA integration in patients with low viral loads has been discussed in a recent study, indicating that integration can extend to the entire liver and support the ubiquitous expression of HBsAg.[20] Using ST technology, our study revealed widespread transcriptionally active HBV integration in different phases of chronic HBV infection, even in patients with HBsAg loss, indicating that HBV integration is more pervasive than previously thought. Furthermore, 95.43% (5427 of 5498) of the unique virus-host chimeric reads represented HBV positive-strand, indicating that these HBV integrations are actively transcribed. In addition, transcripts from the integrated HBV tended to fuse with the human genome across the DR1 locus on the HBV genome, which is consistent with previous studies identifying DR1 as the major junction for viral integrations.[21]

Patients who were receiving or had received antiviral treatment had fewer transcriptionally active viral integration events than treatment-naïve patients, which raises issues regarding the optimal timing of therapeutic intervention in clinical practice.[22] A previous study revealed that inhibition of viral replication is associated with a reduced number of transcriptionally active distinct HBV-host DNA integration.[23] Another study also showed that NUC treatment resulted in a reduction of HBV DNA integration, but viral integration was still detectable in all patients even after 10 years of treatment.[24] We speculate that blocking viral replication by antiviral treatment likely reduces various intrahepatic HBV DNA forms, including dslDNA, thereby preventing new rounds of HBV infection and reducing de novo HBV DNA integration, meanwhile, uninfected hepatocytes may expand, diluting the number of hepatocytes with viral integrations in the whole liver, thus reducing the proportion of hepatocytes containing integrated HBV DNA during antiviral treatment. In addition, no significant correlation was observed between the number of chimeric reads and known duration of infection, suggesting that during the process of random death and regeneration, non-transformed hepatocytes carrying HBV integration may be lost over time if they do not have a selective advantage and could be replaced by non-infected cells if antiviral therapy is potent enough to reduce the pool of infectious virus in the blood circulation. However, another study performed in the woodchuck model, where almost all hepatocytes are infected harbouring both cccDNA and integrated viral DNA sequences, showed that while HBV cccDNA may decrease during treatment, integrated viral DNA showed no discernible decrease.[25] Furthermore, even in patients with HBsAg loss, HBV can be present in the liver as transcriptionally inactive cccDNA or integrated HBV DNA.[26] Taken together with the above findings, it was not unexpected to see few viral sequences integrated into

the host genome in patients under treatment and patients with HBsAg loss, as a result of integration events occurring in the earlier infection phase. This is consistent with the in vitro observation that HBV integration may occur very early after infection in cultured hepatocytes.[27]

The accumulation of HBV integration in the human genome imposes a mutation burden by disrupting important regulatory genes, driving aberrant gene expression by fusion of viral-human sequences, and inducing genomic rearrangements,[28] thus increasing the risk of oncogenic transformation.[29 30] The most frequently reported mutated genes involved in hepatocarcinogenesis,[31] such as *TRER, TP53 and CTNNB1*,[32] were not found in the HBV integration sites in our study, which seems quite reasonable as none of these patients have developed HCC so far (2–3 years post liver biopsy). However, *ALB* and *APOB*, commonly reported oncogenic genes in HCC,[33] were found in the HBV integration sites in this study. Moreover, some HBV integration sites in patient OT6 are cancer-related genes, such as *ALCAM*, which may contribute to HCC development,[34] and *SH2D5*, which is triggered by the HBV X protein (HBx), inducing HCC cell proliferation,[35] and *MBL2,* which may result in compromised innate immunity and elevated HCC risk.[36] Theoretically, antiviral therapy in non-cirrhotic and non-HCC patients could minimise the genetic damage to the hepatocyte population caused by viral integration and reduce the selective expansion of hepatocyte lineages containing HBV integration, which constitutes a risk factor of HCC.[37] Early or late antiviral intervention could have potentially the same effect in preventing new rounds of infection and subsequent integration events; however, in the case of late treatment intervention, the burden of integrated sequences would be significantly higher and associated with a higher likelihood of procarcinogenic events also linked to the duration of infection. Transcriptionally active HBV integration in cancer-related genes suggests that hepatocarcinogenesis could be already underway even in patients receiving antiviral treatment. Therefore, HCC screening should be carefully maintained in patients undergoing antiviral therapy, and therapeutics targeting infected cells with integrated viral DNA may be needed in the future. Further studies are required to explore the merits of antiviral therapy in preventing disease progression and development of HCC by reducing HBV integration.

We also performed HBsAg and HBcAg IHC to investigate the colocalisation between the expression of viral antigens and transcriptionally active HBV integration and observed a complex distribution of HBsAg, HBcAg and viral integration. Liver zones with the ubiquitous presence of intrahepatic HBsAg and highly expressed HBcAg but fewer viral integration-containing spots suggested that transcriptionally active viral integration events did not occur in all infected cells. HBcAg-expressing hepatocytes may have a high level of HBV replication and high cccDNA content,[38] therefore, we speculated that HBsAg might be mainly derived from intrahepatic cccDNA in these cases. There was no significant correlation between reductions in HBsAg and transcriptionally active HBV integration in patients receiving NUC treatment,[24] supporting that HBsAg can also be expressed from residual cccDNA during treatment. The observation that hepatocytes expressing HBsAg colocalising with transcriptionally active HBV integration and core antigens suggests that both viral integration and cccDNA may serve as transcriptional templates for HBsAg expression. There were liver zones where HBsAg expression was strongly positive and with viral integration-containing spots, whereas HBcAg was negative, implying that the infected hepatocytes might harbour transcriptionally silenced cccDNA. Additionally, patients with HBsAg loss were both immunohistochemically negative for HBsAg, but they may still harbour integration-containing spots, although at a relatively lower level, indicating that some viral integrations may be 'silent' for HBsAg expression. A previous study suggested that a selective process for hepatocyte turnover can occur and might result in the emergence of hepatocyte clones that are resistant to T-cell killing or some hepatocyte lineages that are more responsive to growth signals to divide.[7] Therefore, the HBV 'silent' hepatocytes lineages may expand and escape immune responses in the liver of these patients.

HBV cccDNA is known to persist in the livers of infected patients even after long-term therapy[39] and even after HBsAg loss and seroconversion.[40] In our study, except for one patient with HBsAg loss, cccDNA was detected in patients regardless of their low or moderate serum HBV DNA levels. Moreover, cccDNA levels were correlated with transcriptionally active viral integration events.

In summary, we used ST technology to provide evidence that transcriptionally active HBV integration can be detected throughout the liver tissue in different phases of chronic HBV infection, even in patients who achieved functional cure, and patients who were receiving or had received antiviral treatment tended to have lower transcriptionally active viral integration levels.

Author affiliations
[1]Department of Infectious Diseases, Research Laboratory of Clinical Virology, Shanghai Jiao Tong University Medical School Affiliated Ruijin Hospital, Shanghai, China
[2]Department of Infectious Diseases, Shanghai Jiao Tong University Medical School Affiliated Ruijin Hospital, Shanghai, China
[3]Center for Intelligent Medicine Research, Greater Bay Area Institute of Precision Medicine (Guangzhou), School of Life Sciences, Fudan University, Guangzhou, China
[4]INSERM U1052- Cancer Research Center of Lyon (CRCL), Lyon, France
[5]University of Lyon, UMR_S1052, CRCL, Lyon, France
[6]Department of Hepatology, Croix Rousse Hospital, Hospices Civils de Lyon, Lyon, France
[7]Clinical Research Center, Shanghai Jiao Tong University Medical School Affiliated Ruijin Hospital, Shanghai, China

Twitter Ying Jing @yingjing06

Acknowledgements The authors are grateful to Barbara Testoni for her great support on cccDNA work. We would like to thank the 18th ISVHLD Global Hepatitis Summit 2023 for the opportunity to present this study and publish the abstract.[41] We also thank all the patients involved in this study

Contributors XZ and XY conceived and designed the study. XZ and QG recruited patients and collected liver biopsy samples. XY, QG and YC were responsible for collecting and summarising the clinical data. XY and DY performed the experiments. XY and YJ analysed the data. XY wrote the manuscript supported by XZ and FZ. FZ and XZ critically revised the manuscript and provided meaningful input. XZ acted as guarantor. All authors reviewed and approved the final version.

Funding This study was funded by a grant from National Natural Science Foundation of China (81974301), a grant from National Key Projects of Infectious Diseases (2018ZX10302204-001-003), and a grant from Shanghai Municipal Science and Technology Major Project (ZD2021CY001) awarded to XZ. FZ received fundings from the HBV cure task force ANRS-EID (Agence Nationale pour la Recherche sur le SIDA et les hépatites virales-Maladies infectieuses émergentes), the French National Research Agency «Investissements d'Avenir program» (CirB-RNA project – ANR-17-RHUS-0003), the European Union's Horizon 2020 research and innovation program under grant agreement n°847939 (IP-cure-B project).

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Consent obtained directly from patient(s).

Ethics approval This study involves human participants and was approved by the Ethics Committee of Shanghai Ruijin Hospital (RJHKY2018-234). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. All data relevant to the study are included in the article or uploaded as online supplemental information. Raw sequencing data have been submitted to Gene Expression Omnibus (GEO), and the GEO accession number is GSE243367.

ORCID iDs
Demin Yu http://orcid.org/0000-0001-9441-8616
Fabien Zoulim http://orcid.org/0000-0002-2245-0083
Xinxin Zhang http://orcid.org/0000-0002-0598-6425

REFERENCES

1 Polaris Observatory C. Global prevalence, treatment, and prevention of hepatitis B virus infection in 2016: a Modelling study. *Lancet Gastroenterol Hepatol* 2018;3:383–403.
2 Tong S, Revill P. Overview of hepatitis B viral replication and genetic variability. *J Hepatol* 2016;64:S4–16.
3 Levrero M, Testoni B, Zoulim F. HBV cure: Why, how, when? *Curr Opin Virol* 2016;18:135–43.
4 Tu T, Zhang H, Urban S. Hepatitis B virus DNA integration: in vitro models for investigating viral pathogenesis and persistence. *Viruses* 2021;13:180.
5 Li X, Zhang J, Yang Z, *et al*. The function of targeted host genes determines the Oncogenicity of HBV integration in hepatocellular carcinoma. *J Hepatol* 2014;60:975–84.
6 Tu T, Budzinska MA, Shackel NA, *et al*. HBV DNA integration: molecular mechanisms and clinical implications. *Viruses* 2017;9:75.
7 Mason WS, Gill US, Litwin S, *et al*. HBV DNA integration and Clonal hepatocyte expansion in chronic hepatitis B patients considered immune tolerant. *Gastroenterology* 2016;151:986–98.
8 Mason WS, Low H-C, Xu C, *et al*. Detection of Clonally expanded hepatocytes in chimpanzees with chronic hepatitis B virus infection. *J Virol* 2009;83:8396–408.
9 Vickovic S, Eraslan G, Salmén F, *et al*. High-definition spatial Transcriptomics for in situ tissue profiling. *Nat Methods* 2019;16:987–90.
10 Wu R, Guo W, Qiu X, *et al*. Comprehensive analysis of spatial architecture in primary liver cancer. *Sci Adv* 2021;7:eabg3750.
11 European Association for the Study of the Liver. Electronic address EEE, European Association for the study of the L EASL, clinical practice guidelines on the management of hepatitis B virus infection. 2017;67:370–98.
12 Yu X, Gong Q, Yu D, *et al*. Spatial Transcriptomics reveals a low extent of Transcriptionally active hepatitis B virus integration in patients with HBsAg loss. 2023. Available: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE243367
13 Li H, van der Leun AM, Yofe I, *et al*. Dysfunctional Cd8 T cells form a proliferative, dynamically regulated compartment within human Melanoma. *Cell* 2019;176:775–89.
14 Butler A, Hoffman P, Smibert P, *et al*. Integrating single-cell Transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 2018;36:411–20.
15 Shieh F-S, Jongeneel P, Steffen JD, *et al*. Chimericseq: an open-source, user-friendly interface for analyzing NGS data to identify and characterize viral-host Chimeric sequences. *PLoS One* 2017;12:e0182843.
16 Zhao M, Kim P, Mitra R, *et al*. Tsgene 2.0: an updated literature-based Knowledgebase for tumor Suppressor genes. *Nucleic Acids Res* 2016;44:D1023–31.
17 Liu Y, Sun J, Zhao M. Ongene: A literature-based database for human Oncogenes. *J Genet Genomics* 2017;44:119–21.
18 Tate JG, Bamford S, Jubb HC, *et al*. COSMIC: the catalogue of somatic mutations in cancer. *Nucleic Acids Res* 2019;47:D941–7.
19 Allweiss L, Testoni B, Yu M, *et al*. Quantification of the hepatitis B virus cccDNA: evidence-based guidelines for monitoring the key obstacle of HBV cure. *Gut* 2023;72:972–83.
20 Meier M-A, Calabrese D, Suslov A, *et al*. Ubiquitous expression of HBsAg from integrated HBV DNA in patients with low viral load. *J Hepatol* 2021;75:840–7.
21 Podlaha O, Wu G, Downie B, *et al*. Genomic modeling of hepatitis B virus integration frequency in the human genome. *PLoS One* 2019;14:e0220376.
22 Zoulim F, Mason WS. Reasons to consider earlier treatment of chronic HBV infections. *Gut* 2012;61:333–6.
23 Hsu Y-C, Suri V, Nguyen MH, *et al*. Inhibition of viral replication reduces Transcriptionally active distinct hepatitis B virus Integrations with implications on host gene dysregulation. *Gastroenterology* 2022;162:1160–70.
24 Chow N, Wong D, Lai C-L, *et al*. Effect of antiviral treatment on hepatitis B virus integration and hepatocyte Clonal expansion. *Clin Infect Dis* 2023;76:e801–9.
25 Summers J, Mason WS. Residual integrated viral DNA after Hepadnavirus clearance by nucleoside analog therapy. *Proc Natl Acad Sci U S A* 2004;101:638–40.
26 Wong GLH, Gane E, Lok ASF. How to achieve functional cure of HBV: stopping Nucs, adding interferon or new drug development? *J Hepatol* 2022;76:1249–62.
27 Tu T, Budzinska MA, Vondran FWR, *et al*. Hepatitis B virus DNA integration occurs early in the viral life cycle in an in vitro infection model via sodium Taurocholate Cotransporting polypeptide-dependent uptake of enveloped virus particles. *J Virol* 2018;92:e02007-17.
28 Arzumanyan A, Reis HMGPV, Feitelson MA. Pathogenic mechanisms in HBV- and HCV-associated hepatocellular carcinoma. *Nat Rev Cancer* 2013;13:123–35.
29 Tremblay M-P, Armero VES, Allaire A, *et al*. Global profiling of alternative RNA splicing events provides insights into molecular differences between various types of hepatocellular carcinoma. *BMC Genomics* 2016;17:683.
30 Álvarez EG, Demeulemeester J, Otero P, *et al*. Aberrant integration of hepatitis B virus DNA promotes major restructuring of human hepatocellular carcinoma genome architecture. *Nat Commun* 2021;12:6910.
31 Villanueva A. Hepatocellular carcinoma. reply. *N Engl J Med* 2019;381:10.1056/NEJMc1906565#sa3.
32 Zucman-Rossi J, Villanueva A, Nault J-C, *et al*. Genetic landscape and biomarkers of hepatocellular carcinoma. *Gastroenterology* 2015;149:1226–39.
33 Llovet JM, Pinyol R, Kelley RK, *et al*. Molecular pathogenesis and systemic therapies for hepatocellular carcinoma. *Nat Cancer* 2022;3:386–401.
34 Lu X-Y, Chen D, Gu X-Y, *et al*. Predicting value of ALCAM as a target gene of microRNA-483-5p in patients with early recurrence in hepatocellular carcinoma. *Front Pharmacol* 2017;8:973.
35 Zheng Y, Ming P, Zhu C, *et al*. Hepatitis B virus X protein-induced Sh2 domain-containing 5 (Sh2D5) expression promotes hepatoma cell growth via an Sh2D5-transketolase interaction. *J Biol Chem* 2019;294:4815–27.
36 Su C, Lin Y, Cai L, *et al*. Association between mannose-binding lectin variants, Haplotypes and risk of hepatocellular carcinoma: A case-control study. *Sci Rep* 2016;6:32147.
37 Tu T, Budzinska MA, Shackel NA, *et al*. Conceptual models for the initiation of hepatitis B virus-associated hepatocellular carcinoma. *Liver Int* 2015;35:1786–800.
38 Zhang X, Lu W, Zheng Y, *et al*. In situ analysis of Intrahepatic virological events in chronic hepatitis B virus infection. *J Clin Invest* 2016;126:1079–92.
39 Werle-Lapostolle B, Bowden S, Locarnini S, *et al*. Persistence of cccDNA during the natural history of chronic hepatitis B and decline during Adefovir Dipivoxil therapy. *Gastroenterology* 2004;126:1750–8.
40 Yuen M-F, Wong DK-H, Sablon E, *et al*. Hbsag Seroclearance in chronic hepatitis B in the Chinese: virological, histological, and clinical aspects. *Hepatology* 2004;39:1694–701.