Re-analysis of hepatitis B virus integration sites reveals potential new loci associated with oncogenesis in hepatocellular carcinoma

（肝細胞がんにおけるB型肝炎ウイルス組み込みの再分析：新たな発癌関連部位の特定）

千葉大学大学院医学薬学府
先端医学薬学専攻
（主任：加藤直也 教授）
兒島 隆太

ABSTRACT

BACKGROUND

Hepatitis B virus (HBV) is a major cause of hepatocellular carcinoma (HCC). HBV DNA can get integrated into the hepatocyte genome to promote carcinogenesis. However, the precise mechanism by which the integrated HBV genome promotes HCC has not been elucidated.

AIM

To analyze the features of HBV integration in HCC using a new reference database and integration detection method.

METHODS

Published data, consisting of 426 liver tumor samples and 426 paired adjacent non-tumor samples, were re-analyzed to identify the integration sites. Genome Reference Consortium Human Build 38 (GRCh38) and Telomere-to-Telomere Consortium CHM13 (T2T-CHM13 (v2.0)) were used as the human reference genomes. In contrast, human genome 19 (hg19) was used in the original study. In addition, GRIDSS VIRUSBreakend was used to detect HBV integration sites, whereas high-throughput viral integration detection (HIVID) was applied in the original study (HIVID-hg19).

RESULTS

A total of 5361 integration sites were detected using T2T-CHM13. In the tumor samples, integration hotspots in the cancer driver genes, such as *TERT* and *KMT2B*, were consistent with those in the original study. GRIDSS VIRUSBreakend detected integrations in more samples than by HIVID-hg19. Enrichment of integration was observed at chromosome 11q13.3, including the *CCND1* promoter, in tumor samples. Recurrent integration sites were observed in mitochondrial genes.

CONCLUSION

GRIDSS VIRUSBreakend using T2T-CHM13 is accurate and sensitive in detecting HBV integration. Re-analysis provides new insights into the regions of HBV integration and their potential roles in HCC development.

Key Words: Carcinoma, Hepatocellular; Hepatitis B virus; Virus Integration

Core tip: To understand the role of HBV in HCC development, we re-analyzed HBV integration sites using publicly available data. We found that chromosome 11q13.3 is a frequently observed HBV integration site in the tumor samples. This region contains important cancer driver genes, such as *CCND1* and *FGF19*, which are amplified in HCC.

This finding supports a mechanism of carcinogenesis promoted by HBV-induced genomic instability in the liver and provides insights into treating a subset of liver cancers.

INTRODUCTION

The hepatitis B virus (HBV) is a major cause of hepatocellular carcinoma (HCC). When HBV infects liver cells, HBV DNA can be integrated into the human genome. Integration events typically occur during the early stages of an infection[1,2], and are known to promote carcinogenesis via several mechanisms: (1) increasing the expression levels of neighboring genes; (2) induction of genomic instability and somatic copy number alterations of genes; (3) deletion of tumor suppressor genes through structural mutations[3]; and (4) inducing expression of HBV X protein (HBx) or HBx fusion proteins that contribute to carcinogenesis.

To investigate the effect of HBV on hepatocarcinogenesis, several studies have been conducted using next-generation sequencing technology to identify integration sites of HBV DNA. Examples of such technologies include whole genome sequencing[4] and HBV capture sequencing[5]. These studies revealed frequent integration into the promoter regions of *TERT* and *KMT2B* in tumor tissues and *FN1* in normal tissues. In an examination of an HBV-infected human-hepatocyte chimeric mouse model, mitochondrial DNA (mtDNA) was thought to be a frequent site of integration[1]. A European study reported a lower frequency of *KMT2B* insertion and a higher frequency of integration into *ADH* genes in normal tissues[6] Most previous studies have used Genome Reference Consortium Human Build 37 (GRCh37) or human genome 19 (hg19) as the reference genomes. In GRCh37/hg19 and Genome Reference Consortium Human Build 38 (GRCh38)[7], tandem repeats, microsatellites, and minisatellites found in telomeres and centromeres remained unresolved. The complete human genome sequence, Telomere-to-Telomere Consortium CHM13 (T2T-CHM13 (v2.0))[8], was released in 2022.

Various methods have been used to detect integration breakpoints. High-throughput viral integration detection (HIVID), a detection method based on a pair-read assembly strategy[9], was applied in the analysis of 426 HCC cases[5]. GRIDSS is a multithreaded structural variant caller from a combination of assembly, split read, and read pair support[10]. VIRUSBreakend utilizes a virus-centric variant calling and assembly approach to identify viral integrations with high sensitivity and low false discovery rate, allowing the identification of integrations in repetitive host regions[11].

Here, we report new features observed by re-analyzing the published data using GRIDSS VIRUSBreakend based on GRCh38 and T2T-CHM13.

MATERIALS AND METHODS

Sequence data were obtained from the Sequence Read Archive (SRA) with accession number SRA335342[5]. The dataset consisted of 426 tumor samples and 426 paired adjacent non-tumor samples.

All reads in the dataset were aligned to the GRCh38 and T2T-CHM13 reference genomes using bwa-mem2[12,13]. VIRUSBreakend was used to detect integration sites (Supplementary Figure 1), and the analysis was performed using Nextflow[14] on Amazon Web Service. HBV integration sites were detected using GRIDSS VIRUSBreakend[11]. Integration sites were compared with the count of fragments providing breakend for the variant allele (BVF) in the variant call format (VCF) file. Statistical analysis and visualization were performed using R software, and statistical significance was set at $P < 0.05$.

RESULTS

Comparison of HBV integration sites

In total, 5361 and 5198 integration breakpoints were detected with T2T-CHM13 and GRCh38, respectively. The breakpoints were similar between the references using GRCh38 and T2T-CHM13 (Figures 1A and B). Consistent with previous studies, integration breakpoints were enriched in the *TERT* promoter region in tumor samples. In contrast, integration into *FN1* was frequently observed in non-tumor samples.

Compared with the original study, our analysis detected integrations in more samples (357 vs 328 in tumors; 288 vs 160 in non-tumors) (Table 1). In addition, we detected integration in the *TERT* region in 105 tumor samples, whereas the original study observed integration in 95 tumor samples (Table 2). In contrast, the number of breakpoints detected in tumors was lower than that in the original study (Table 1). In our study, only breakpoints validated by VIRUSBreakend were counted (Supplementary Figure 2). Integration of *DDX11L* was frequently detected in the original study, but no integration breakpoints were detected in our study (Table 2). The *DDX11L* gene family is frequently detected as a target for integration using a capture sequencing approach, but it is possible that fragments were mapped incorrectly owing to repetitive sequences[15,16]. In the non-tumor samples, our study detected more integrations, both in the number of samples and breakpoints. For example, we detected 97 integration breakpoints in the *FN1* gene from 56 non-tumor samples. The earlier analysis detected only 19 breakpoints from 17 non-tumor samples (Table 2). Few oncogenic regions were affected in the non-tumor samples.

Breakpoints were most frequent around direct repeat 1 of the HBV genome (Figure 1C).

Table 1 Comparison of HBV integration breakpoints among reference genomes.

	GRCh38	T2T-CHM13	Original
Tumor			
Number of breakpoints	2439	2487	3486
Number of samples	357	355	328
Non-tumor			
Number of breakpoints	2759	2874	739
Number of samples	288	288	160

Table 2 Comparison of frequent integration breakpoints in the samples.

Gene	GRCh38		Original	
	Breakpoints (n)	Samples (n)	Breakpoints (n)	Samples (n)
Tumor				
TERT	150	105	160	95
KMT2B	56	33	55	30
DDX11L1	0	0	36	23
CCNA2	12	7	14	8
CCNE1	13	9	14	7
Non-tumor				
FN1	97	56	19	17
TERT	12	10	8	3
IQGAP2	7	5	1	1
KMT2B	7	4	5	3

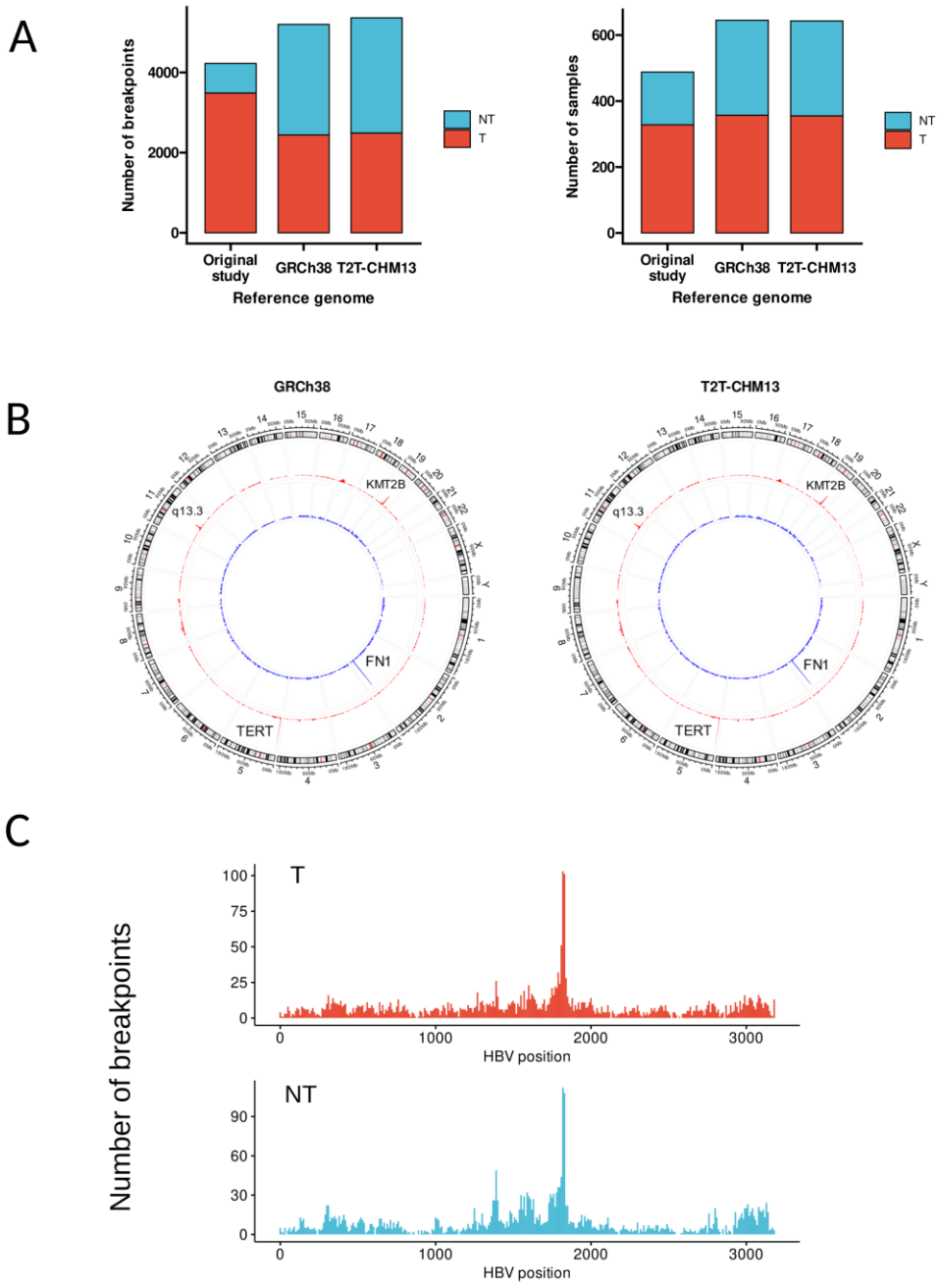

Figure 1 Hepatitis B virus integration breakpoints across the reference genomes. A: Integration breakpoints in the human reference genomes in tumor and non-tumor samples;

B: Circos plot of integration breakpoints. Red represents tumor samples, and blue represents non-tumor samples; C: Hepatitis B virus genome integration breakpoints. T: Tumor; N: Non-tumor.

Chromosome 11q13.3 is a frequent site for HBV integration

When the chromosome region was explored, we found that the integration breakpoint at 11q13.3 was enriched with T2T-CHM13 and GRCh38 (Figures 2A-C). Breakpoints at 11q13.3 were more frequent in the tumor samples than in the non-tumor samples (16 (3.8%) of tumor samples compared to 1 or (0.02%) of non-tumor samples, Figure 2B). 11q13.3 is characterized by the evolutionarily well-conserved genes *CCND1*, *FGF19*, *FGF4*, and *FGF3*[17], where copy number amplification frequently occurs in tumors (Figure 2D)[18,19]. Some breakpoints were within the genic and promoter regions of the genes, including *CCND1* and *FGF4*. Integration appeared to be distributed more in the non-genic regions (Figure 2B). When fragments from the integration site were counted using BVF, the values were higher in tumor samples than in non-tumor samples (Figure 2E). High BVF value formed a peak in the 11q13.3 in addition to the peak in the *TERT*, *KMT2B*, and *CCNE1* genes in the tumor samples (Figure 2E and Supplementary Figure 3).

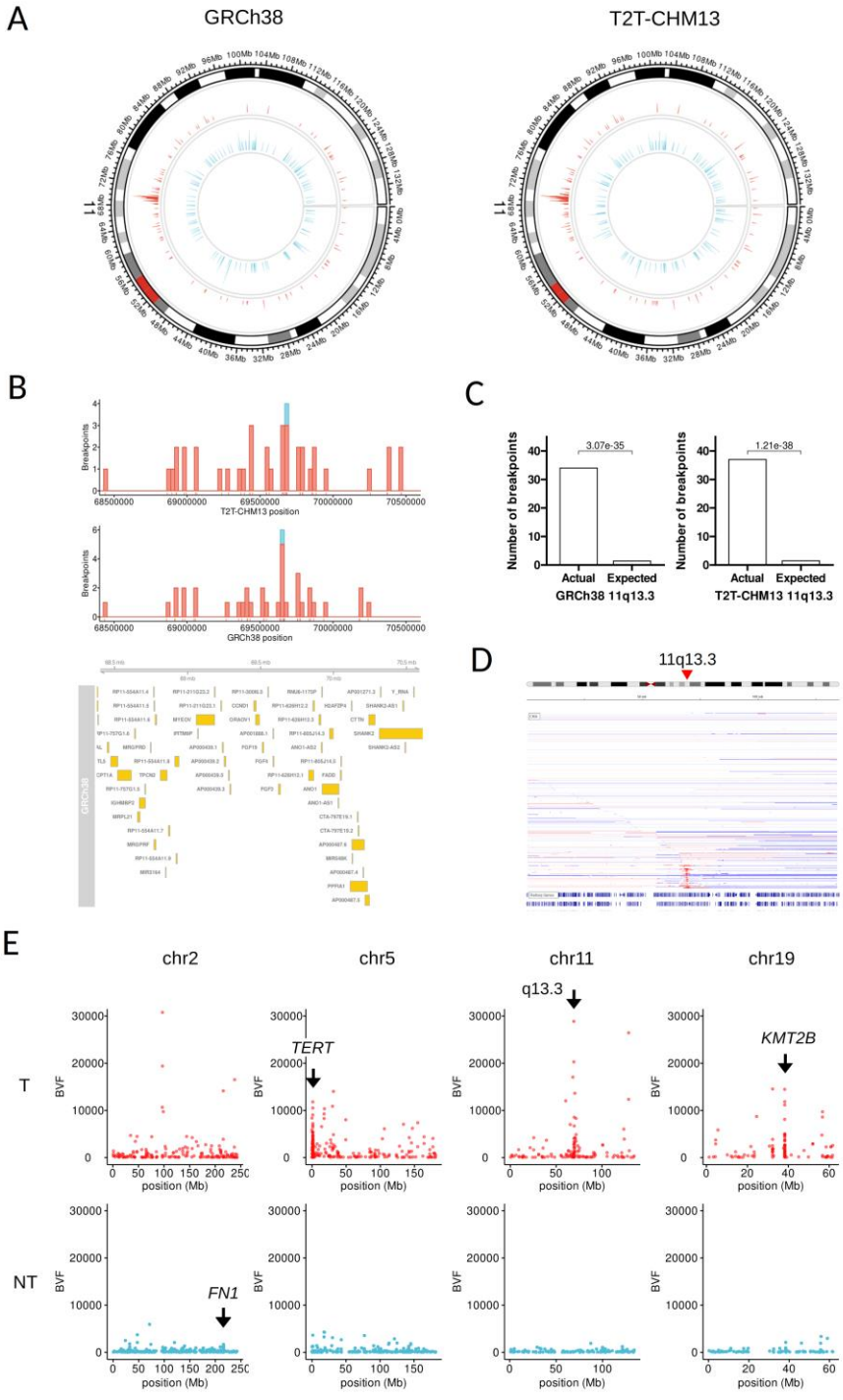

Figure 2 Integration breakpoints at chromosome 11. A: Circos plot of breakpoints at chromosome 11 in the human reference genomes; B: Integration breakpoints around 11q13.3

in relation to coding genes retrieved from Ensembl. Red represents tumor samples, and blue represents non-tumor samples; C: Comparison of integration breakpoints around 11q13.3 in the tumor samples. Actual represents actual number of integration breakpoints. Expected represents expected number of integration breakpoints assuming random distribution; D: Copy number of liver cancer samples from cBioPortal[18,19]. Red represents amplification, and blue represents deletion. E: Distribution of the number of fragments that provide breakend for the variant allele (BVF). T: Tumor; NT: Non-tumor.

Mitochondrial DNA has sites where HBV DNA is frequently integrated

There is some debate regarding whether mtDNA is a frequent site of HBV integration. A study using a mouse model by Furuta et al. found that mtDNA was frequently integrated early in infection[1]. More recently, a preprint suggested that mtDNA is indeed a site for integration[20,21]. Although the original paper on which this study was based did not mention integration into mitochondria, we detected many integration breakpoints into mtDNA and identified repeat integration sites (Table 3 and Figure 3). Integration breakpoints in mtDNA were observed in both tumor and non-tumor samples. Recurrent integration events were observed in *ND4*. Of these, eight events were from non-tumor samples, and two were from tumor samples. Microhomologous sequences were observed in some regions. For example, the GCCNTTCTCATC sequence, where N represents any nucleotide or gap, was observed at the junction of the *ND4* gene (Chromosome M:11079) and the HBV genome (HBV:1559). In contrast, the GCTTCACC sequence was observed at the junction of the *ND4* gene (Chromosome M:11104) and the HBV genome (HBV:1590). It is also possible that these integration breakpoints exist in nuclear-mitochondrial segments.

Table 3 HBV integration breakpoints in mitochondrial DNA.

Sample	Tumor/Non-tumor	Chromosome	Position	HBV	Quality score
SRR3104746	NT	chrM	5367	1247	6933.33
SRR3105143	NT	chrM	5682	1513	11817.58
SRR3105012	T	chrM	8220	112	2080.04
SRR3104491	NT	chrM	8524	2482	33322.16
SRR3105095	NT	chrM	8694	471	4209.06
SRR3105101	NT	chrM	11079	1559	2937.33
SRR3105143	NT	chrM	11079	1559	14876.83
SRR3105001	NT	chrM	11104	1590	45538.21
SRR3105149	NT	chrM	11104	1590	16156.63
SRR3105251	NT	chrM	11104	1590	2276.31
SRR3105293	NT	chrM	11104	1590	24562.69
SRR3104643	T	chrM	11104	1590	13683.16
SRR3105172	T	chrM	11126	1621	2610.44
SRR3105251	NT	chrM	11130	1625	6793.43
SRR3104939	NT	chrM	11139	1729	30830.54
SRR3105149	NT	chrM	12453	323	2345.03
SRR3104636	NT	chrM	12735	1381	26840.64
SRR3105083	NT	chrM	13273	1755	4230.85
SRR3104696	NT	chrM	13433	363	2963.36
SRR3104643	T	chrM	13964	2809	2222.86
SRR3104823	NT	chrM	14052	1017	1788.18
SRR3105049	NT	chrM	14892	1768	21102.87
SRR3104982	T	chrM	15679	1768	32371.78
SRR3105185	NT	chrM	16319	1788	2013.53

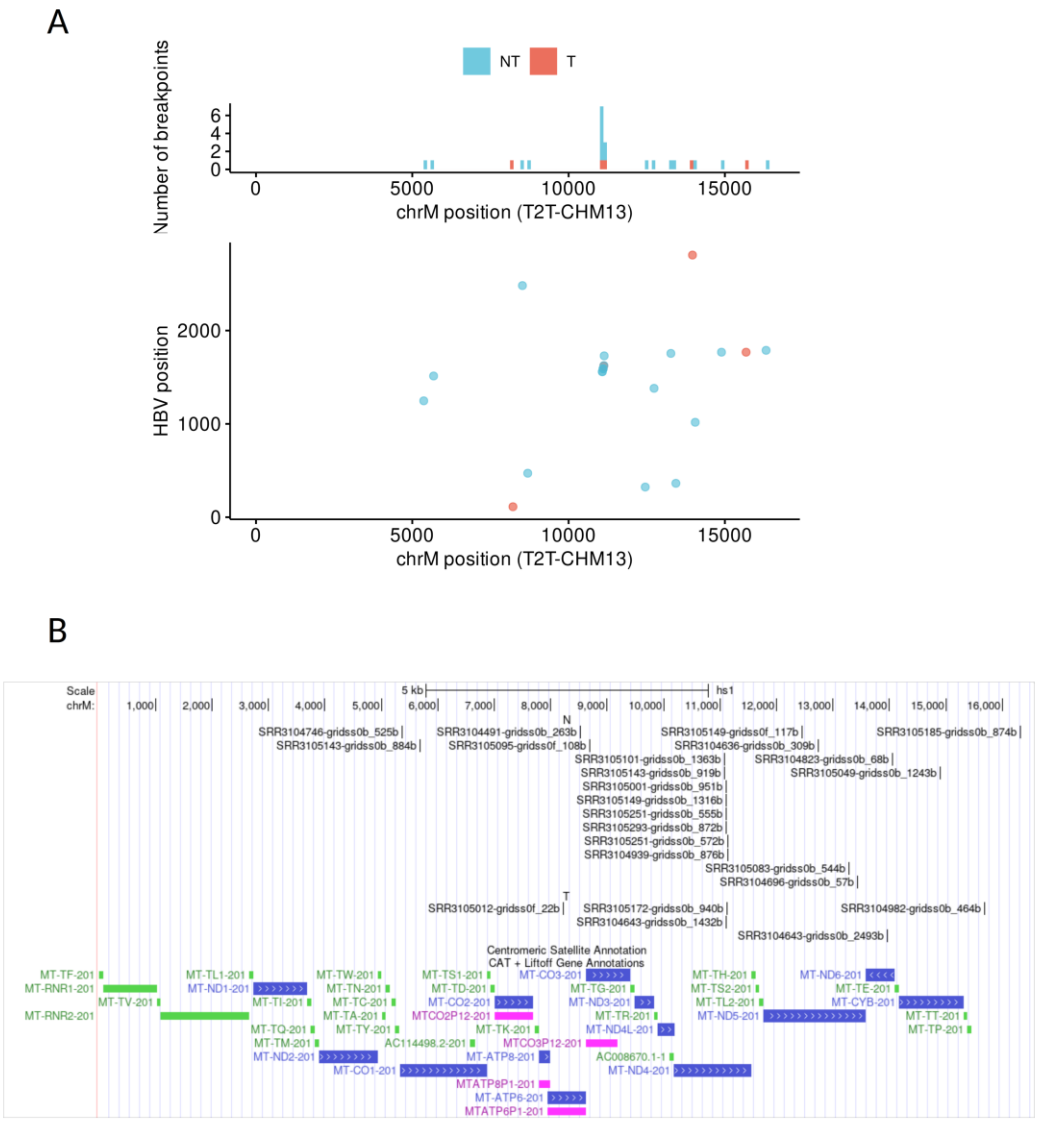

Figure 3 Integration breakpoints in the mitochondrial genome. A: The upper panel displays integration breakpoints across mitochondrial genomes according to tumor and non-tumor

samples, and the lower panel shows integration breakpoints along the human and hepatitis B virus genomes. Red represents tumor samples, and blue represents non-tumor samples; B: Integration breakpoints on the mitochondrial genome annotated using UCSC genome browser (NT: Non-tumor; T: Tumor)[22].

DISCUSSION

In this study, GRIDSS VIRUSBreakend, with an updated human reference genome, was used to detect HBV integration using public sequencing data from liver tumor and non-tumor samples. HBV integration was detected in more samples than in the original analysis (Table 1). The difference in methods could account for the discordant results. We investigated an example of HBV integration sites in the *TERT* region detected by GRIDSS VIRUSBreakend, but not in the original study (Supplementary Figures 4 and 5). In the original study, the HIVID pipeline, based on paired-end read assembly, was applied to detect integration[9]. In the sequencing data, some paired-end reads could not be assembled because of the absence of overlapping bases. These reads were also included in our analysis to detect integration sites more accurately. It should be noted that the GRIDSS VIRUSBreakend uses genotype D HBV for viral genome reference, whereas genotype C HBV is dominant in the current dataset, which may affect the sensitivity of virus detection.

We found HBV integration clusters in the 11q13.3 region (Figure 2). Unlike previously known single gene integration sites, such as *TERT* and *KMT2B*, 11q13.3 spans multiple gene regions. Although these clusters can be observed in the supplemental data of the original paper, to our knowledge, it has not been previously mentioned. Enrichment of 11q13.3 was more significant in tumors than in non-tumor tissues. *CCND1*, *FGF19*, *FGF4*, and *FGF3* are located at 11q13.3, where copy number amplification frequently occurs in tumors.

Integration into *CCND1*, located at 11q13.3, is a potential driver event[23], but its frequency is not high. Although recent studies have not detected integration at 11q13.3[1,6], several studies have detected these events only as supplementary data[4,5,24] and they have been reported since 1988[25,26]. According to a study by Bok et al., the expression levels of cancer-related genes, including *CCND1* and *FGF19*, are elevated near the viral integration site on 11q13.3 in an HCC cell line[27]. HBV integration at this locus may be linked to cancer gene activation, as *FGF19* amplification was associated with chronic HBV infection.[28,29].

HBV integration may be associated with copy number alterations[3]. Chromosomal instability often leads to copy number alterations in the short and long arms of the chromosome. However, 11q13.3 causes strong copy number amplification in a localized region in the middle of the chromosome (Figure 2D). Previous results using whole genome sequencing indicated that the integration allele frequency was high in the tumor samples, especially in the recurrent integration in tumors such as *TERT*[4]. By comparing fragment counts from the integration site using BVF, the values were found to be higher in the tumor samples than in the non-tumor samples. Some of the integration breakpoints at 11q13.3 had extremely high fragment counts (Figure 2E and Supplementary Figure 3). If BVF correlates with the integration allele frequency, it is possible that these events reflect the clonal expansion of

tumors with integration breakpoints or the amplification of integrated genes. *CCND1-FGF19* amplification occurred at later points in the evolution of HCC[30]. Further research is needed to investigate the relationships between integration, copy number alteration, and cancer gene activation at 11q13.3.

In our analysis, HBV integration in the mtDNA was observed in 2.3% (20/852) of the samples, and the *ND4* gene was a frequent target of HBV integration (Table 3). According to a previous study, HBV integration into mtDNA has occurred in only 0.1% of human clinical liver tissues[1]. Mouse model experiments have suggested that this integration primarily occurs during the early stages of HBV infection through microhomology-mediated end joining[1]. It is also possible that HBV integration occurs in nuclear copies of mtDNA sequences rather than in the mitochondria. Giosa et al. detected HBV integration in DNA isolated from mitochondria[20]. The D-loop region is the target of HBV integration. Our analysis suggests that *ND4* genes may also be targeted for integration through microhomology-mediated mechanisms.

This study has several limitations. First, the analysis was conducted using existing data and the findings were not validated using independent data. Second, the original data were based on HBV capture sequencing, and gene copy numbers were not available. Finally, the integration data were obtained from short-read sequencing and have not been validated using long-read sequencing data.

CONCLUSION

HBV integration in HCC samples has been characterized using the complete human reference. GRIDSS VIRUSBreakend using T2T-CHM13 is accurate and sensitive in detecting HBV integration. HBV frequently integrates at the 11q13.3 region, where the *CCND1* gene is located, and this region is frequently amplified in several types of cancer, including HCC. Further research is needed to examine how HBV integration interacts with driver gene expression and copy number alteration.

DATA AVAILABILITY

All the data supporting this study are stored in the SRA database with accession number SRA335342. Supplementary Tables are available for download in the Zenodo repository (https://zenodo.org/, Digital Object Identifiers: 10.5281/zenodo.8118882).

ACKNOWLEDGEMENTS

We would like to thank Feroz Meeran for English language editing.

SUPPLEMENTARY INFORMATION

- Supplementary Figure 1 Overview of HBV integration discovery pipeline

- Supplementary Figure 2 An example of HBV integration breakpoints in the *TERT* promoter region of SRR3104641
- Supplementary Figure 3 Scatterplot of BVF across chromosomes
- Supplementary Figure 4 An example of HBV integration breakpoints in the *TERT* promoter region of SRR3104862
- Supplementary Figure 5 Paired-end reads lacking overlapping bases

Generate IGV reference[31]

Supplementary Table (https://doi.org/10.5281/zenodo.8118882)

- Integration Sites (T2T-CHM13_human, T2T-CHM13_hbv, GRCh38_human, GRCh38_hbv)
- Annotation using GRCh38
- Sample metadata

REFERENCES

1 Furuta M, Tanaka H, Shiraishi Y, Uchida T, Imamura M, Fujimoto A, Fujita M, Sasaki-Oku A, Maejima K, Nakano K, Kawakami Y, Arihiro K, Aikata H, Ueno M, Hayami S, Ariizumi S-I, Yamamoto M, Gotoh K, Ohdan H, Yamaue H, Miyano S, Chayama K, Nakagawa H. Characterization of HBV integration patterns and timing in liver cancer and HBV-infected livers. *Oncotarget* 2018; **9**: 25075–25088. [DOI: 10.18632/oncotarget.25308]

2 Chauhan R, Michalak TI. Earliest hepatitis B virus-hepatocyte genome integration: Sites, mechanism, and significance in carcinogenesis. *Hepatoma Research* 2021; **7**: 20. [DOI: 10.20517/2394-5079.2020.136]

3 Álvarez EG, Demeulemeester J, Otero P, Jolly C, García-Souto D, Pequeño-Valtierra A, Zamora J, Tojo M, Temes J, Baez-Ortega A, Rodriguez-Martin B, Oitaben A, Bruzos AL, Martínez-Fernández M, Haase K, Zumalave S, Abal R, Rodríguez-Castro J, Rodriguez-Casanova A, Diaz-Lagares A, Li Y, Raine KM, Butler AP, Otero I, Ono A, Aikata H, Chayama K, Ueno M, Hayami S, Yamaue H, Maejima K, Blanco MG, Forns X, Rivas C, Ruiz-Bañobre J, Pérez-del-Pulgar S, Torres-Ruiz R, Rodriguez-Perales S, Garaigorta U, Campbell PJ, Nakagawa H, Van Loo P, Tubio JMC. Aberrant integration of Hepatitis B virus DNA promotes major restructuring of human hepatocellular carcinoma genome architecture. *Nat Commun* 2021; **12**: 6910. [DOI: 10.1038/s41467-021-26805-8]

4 Sung W-K, Zheng H, Li S, Chen R, Liu X, Li Y, Lee NP, Lee WH, Ariyaratne PN, Tennakoon C, Mulawadi FH, Wong KF, Liu AM, Poon RT, Fan ST, Chan KL, Gong Z, Hu Y, Lin Z, Wang G, Zhang Q, Barber TD, Chou W-C, Aggarwal A, Hao K, Zhou W, Zhang C, Hardwick J, Buser C, Xu J, Kan Z, Dai H, Mao M, Reinhard C, Wang J, Luk JM. Genome-wide survey of recurrent HBV integration in hepatocellular carcinoma. *Nat Genet* 2012; **44**: 765–769. [DOI: 10.1038/ng.2295]

5 Zhao L-H, Liu X, Yan H-X, Li W-Y, Zeng X, Yang Y, Zhao J, Liu S-P, Zhuang X-H, Lin C, Qin C-J, Zhao Y, Pan Z-Y, Huang G, Liu H, Zhang J, Wang R-Y, Yang Y, Wen

W, Lv G-S, Zhang H-L, Wu H, Huang S, Wang M-D, Tang L, Cao H-Z, Wang L, Lee T-L, Jiang H, Tan Y-X, Yuan S-X, Hou G-J, Tao Q-F, Xu Q-G, Zhang X-Q, Wu M-C, Xu X, Wang J, Yang H-M, Zhou W-P, Wang H-Y. Genomic and oncogenic preference of HBV integration in hepatocellular carcinoma. *Nat Commun* 2016; **7**: 12992. [DOI: 10.1038/ncomms12992]

6 Péneau C, Imbeaud S, La Bella T, Hirsch TZ, Caruso S, Calderaro J, Paradis V, Blanc J-F, Letouzé E, Nault J-C, Amaddeo G, Zucman-Rossi J. Hepatitis B virus integrations promote local and distant oncogenic driver alterations in hepatocellular carcinoma. *Gut* 2022; **71**: 616–626. [DOI: 10.1136/gutjnl-2020-323153]

7 Schneider VA, Graves-Lindsay T, Howe K, Bouk N, Chen H-C, Kitts PA, Murphy TD, Pruitt KD, Thibaud-Nissen F, Albracht D, Fulton RS, Kremitzki M, Magrini V, Markovic C, McGrath S, Steinberg KM, Auger K, Chow W, Collins J, Harden G, Hubbard T, Pelan S, Simpson JT, Threadgold G, Torrance J, Wood JM, Clarke L, Koren S, Boitano M, Peluso P, Li H, Chin C-S, Phillippy AM, Durbin R, Wilson RK, Flicek P, Eichler EE, Church DM. Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly. *Genome Res* 2017; **27**: 849–864. [DOI: 10.1101/gr.213611.116]

8 Nurk S, Koren S, Rhie A, Rautiainen M, Bzikadze AV, Mikheenko A, Vollger MR, Altemose N, Uralsky L, Gershman A, Aganezov S, Hoyt SJ, Diekhans M, Logsdon GA, Alonge M, Antonarakis SE, Borchers M, Bouffard GG, Brooks SY, Caldas GV, Chen N-C, Cheng H, Chin C-S, Chow W, de Lima LG, Dishuck PC, Durbin R, Dvorkina T, Fiddes IT, Formenti G, Fulton RS, Fungtammasan A, Garrison E, Grady PGS, Graves-Lindsay TA, Hall IM, Hansen NF, Hartley GA, Haukness M, Howe K, Hunkapiller MW, Jain C, Jain M, Jarvis ED, Kerpedjiev P, Kirsche M, Kolmogorov M, Korlach J, Kremitzki M, Li H, Maduro VV, Marschall T, McCartney AM, McDaniel J, Miller DE, Mullikin JC, Myers EW, Olson ND, Paten B, Peluso P, Pevzner PA, Porubsky D, Potapova T, Rogaev EI, Rosenfeld JA, Salzberg SL, Schneider VA, Sedlazeck FJ, Shafin K, Shew CJ, Shumate A, Sims Y, Smit AFA, Soto DC, Sović I, Storer JM, Streets A, Sullivan BA, Thibaud-Nissen F, Torrance J, Wagner J, Walenz BP, Wenger A, Wood JMD, Xiao C, Yan SM, Young AC, Zarate S, Surti U, McCoy RC, Dennis MY, Alexandrov IA, Gerton JL, O'Neill RJ, Timp W, Zook JM, Schatz MC, Eichler EE, Miga KH, et al. The complete sequence of a human genome. *Science* 2022; **376**: 44–53. [DOI: 10.1126/science.abj6987]

9 Li W, Zeng X, Lee NP, Liu X, Chen S, Guo B, Yi S, Zhuang X, Chen F, Wang G, Poon RT, Fan ST, Mao M, Li Y, Li S, Wang J, JianWang, Xu X, Jiang H, Zhang X. HIVID: An efficient method to detect HBV integration using low coverage sequencing. *Genomics* 2013; **102**: 338–344. [DOI: 10.1016/j.ygeno.2013.07.002]

10 Cameron DL, Baber J, Shale C, Valle-Inclan JE, Besselink N, van Hoeck A, Janssen R, Cuppen E, Priestley P, Papenfuss AT. GRIDSS2: Comprehensive characterisation of somatic structural variation using single breakend variants and structural variant phasing. *Genome Biol* 2021; **22**: 202. [DOI: 10.1186/s13059-021-02423-x]

11 Cameron DL, Jacobs N, Roepman P, Priestley P, Cuppen E, Papenfuss AT. VIRUSBreakend: Viral Integration Recognition Using Single Breakends. *Bioinformatics* 2021; **37**: 3115–3119. [DOI: 10.1093/bioinformatics/btab343]

12 Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. (e-pub ahead of print 2013; doi:10.48550/ARXIV.1303.3997).

13 Vasimuddin Md, Misra S, Li H, Aluru S. Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems. In: *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*. Rio de Janeiro, Brazil: IEEE, 2019: 314–324.

14 Di Tommaso P, Chatzou M, Floden EW, Barja PP, Palumbo E, Notredame C. Nextflow enables reproducible computational workflows. *Nat Biotechnol* 2017; **35**: 316–319. [DOI: 10.1038/nbt.3820]

15 Tatsuno K, Midorikawa Y, Takayama T, Yamamoto S, Nagae G, Moriyama M, Nakagawa H, Koike K, Moriya K, Aburatani H. Impact of AAV2 and Hepatitis B Virus Integration Into Genome on Development of Hepatocellular Carcinoma in Patients with Prior Hepatitis B Virus Infection. *Clin Cancer Res* 2019; **25**: 6217–6227. [DOI: 10.1158/1078-0432.CCR-18-4041]

16 Midorikawa Y, Tatsuno K, Moriyama M. Genome-wide analysis of hepatitis B virus integration in hepatocellular carcinoma: Insights next generation sequencing. *Hepatobiliary Surgery and Nutrition* 2021; **10**: 54852–54552. [DOI: 10.21037/hbsn-21-228]

17 Katoh M, Katoh M. Evolutionary conservation of CCND1-ORAOV1-FGF19-FGF4 locus from zebrafish to human. *Int J Mol Med* (e-pub ahead of print 1 July 2003; doi:10.3892/ijmm.12.1.45).

18 Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discovery* 2012; **2**: 401–404. [DOI: 10.1158/2159-8290.CD-12-0095]

19 Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C, Schultz N. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *Sci Signal* 2013; **6**. [DOI: 10.1126/scisignal.2004088]

20 Giosa D, Lombardo D, Musolino C, Chines V, Raffa G, Casuscelli di Tocco F, D'Aliberti D, Caminiti G, Saitta C, Alibrandi A, Aiese Cigliano R, Romeo O, Navarra G, Raimondo G, Pollicino T. Mitochondrial DNA is a target of HBV integration. *Commun Biol* 2023; **6**: 1–14. [DOI: 10.1038/s42003-023-05017-4]

21 Giosa D, Lombardo D, Musolino C, Chines V, Raffa G, Casuscelli di Tocco F, D'Aliberti D, Saitta C, Alibrandi A, Aiese Cigliano R, Romeo O, Navarra G, Raimondo G, Pollicino T. A new high-throughput HBV integration sequencing approach shows that mitochondrial DNA is frequently targeted by virus integration in liver cells with active HBV replication. *Digestive and Liver Disease* 2022; **54**: S8–S9. [DOI: 10.1016/j.dld.2022.01.020]

22 Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler and D. The Human Genome Browser at UCSC. *Genome Res* 2002; **12**: 996–1006. [DOI: 10.1101/gr.229102]

23 Ally A, Balasundaram M, Carlsen R, Chuah E, Clarke A, Dhalla N, Holt RA, Jones SJM, Lee D, Ma Y, Marra MA, Mayo M, Moore RA, Mungall AJ, Schein JE, Sipahimalani P, Tam A, Thiessen N, Cheung D, Wong T, Brooks D, Robertson AG, Bowlby R, Mungall K, Sadeghi S, Xi L, Covington K, Shinbrot E, Wheeler DA, Gibbs RA, Donehower LA, Wang L, Bowen J, Gastier-Foster JM, Gerken M, Helsel C, Leraas KM, Lichtenberg TM, Ramirez NC, Wise L, Zmuda E, Gabriel SB, Meyerson M, Cibulskis C, Murray BA, Shih J, Beroukhim R, Cherniack AD, Schumacher SE, Saksena G, Pedamallu CS, Chin L, Getz G, Noble M, Zhang H, Heiman D, Cho J, Gehlenborg N, Saksena G, Voet D, Lin P, Frazer S, Defreitas T, Meier S, Lawrence M, Kim J, Creighton CJ, Muzny D, Doddapaneni HV, Hu J, Wang M, Morton D, Korchina V, Han Y, Dinh H, Lewis L, Bellair M, Liu X, Santibanez J, Glenn R, Lee S, Hale W, Parker JS, Wilkerson MD, Hayes DN, Reynolds SM, Shmulevich I, Zhang W, Liu Y, Iype L, Makhlouf H, Torbenson MS, Kakar S, Yeh MM, Jain D, Kleiner DE, Jain D, Dhanasekaran R, El-Serag HB, et al. Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell* 2017; **169**: 1327–1341.e23. [DOI: 10.1016/j.cell.2017.05.046]

24 Yoo S, Wang W, Wang Q, Fiel MI, Lee E, Hiotis SP, Zhu J. 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**: 214. [DOI: 10.1186/s12916-017-0973-7]

25 Hatada I, Tokino T, Ochiya T, Matsubara K. Co-amplification of integrated hepatitis B virus DNA and transforming gene hst-1 in a hepatocellular carcinoma. *Oncogene* 1988; **3**: 537–540.

26 Tokino T, Matsubara K. Chromosomal sites for hepatitis B virus integration in human hepatocellular carcinoma. *J Virol* 1991; **65**: 6761–6764. [DOI: 10.1128/jvi.65.12.6761-6764.1991]

27 Bok J, Kim KJ, Park M-H, Cho S-H, Lee H-J, Lee E-J, Park C, Lee J-Y. Identification and extensive analysis of inverted-duplicated HBV integration in a human hepatocellular carcinoma cell line. *BMB Rep* 2012; **45**: 365–370. [DOI: 10.5483/bmbrep.2012.45.6.279]

28 Ahn S-M, Jang SJ, Shim JH, Kim D, Hong S-M, Sung CO, Baek D, Haq F, Ansari AA, Lee SY, Chun S-M, Choi S, Choi H-J, Kim J, Kim S, Hwang S, Lee Y-J, Lee J, Jung

W, Jang HY, Yang E, Sung W-K, Lee NP, Mao M, Lee C, Zucman-Rossi J, Yu E, Lee HC, Kong G. Genomic portrait of resectable hepatocellular carcinomas: Implications of RB1 and FGF19 aberrations for patient stratification. *Hepatology* 2014; **60**: 1972–1982. [DOI: 10.1002/hep.27198]

29 Kang HJ, Haq F, Sung CO, Choi J, Hong S-M, Eo S-H, Jeong HJ, Shin J, Shim JH, Lee HC, An J, Kim M-J, Kim K, Ahn S-M, Yu E. Characterization of Hepatocellular Carcinoma Patients with ***FGF19*** Amplification Assessed by Fluorescence in situ Hybridization: A Large Cohort Study. *Liver Cancer* 2019; **8**: 12–23. [DOI: 10.1159/000488541]

30 Zhou S-L, Zhou Z-J, Song C-L, Xin H-Y, Hu Z-Q, Luo C-B, Luo Y-J, Li J, Dai Z, Yang X-R, Shi Y-H, Wang Z, Huang X-W, Fan J, Zhou J. Whole-genome sequencing reveals the evolutionary trajectory of HBV-related hepatocellular carcinoma early recurrence. *Sig Transduct Target Ther* 2022; **7**: 1–17. [DOI: 10.1038/s41392-021-00838-3]

31 Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. *Nat Biotechnol* 2011; **29**: 24–26. [DOI: 10.1038/nbt.1754]

World Journal of Virology

Vol.12 No.3

２０２３年６月２５日公表済

DOI: 10.5501/wjv.v12.i3.209