Biochemical and Biophysical Research Communications 553 (2021) 160–164

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Profile of different Hepatitis B virus integration frequency in hepatocellular carcinoma patients

Weiyang Li [a, b, 1], Yanwei Qi [a, 1], Hanshi Xu [c], Wei Wei [a], Xiaofang Cui [a, d, *]

[a] Jining Medical University, Jining, Shandong, 272067, China
[b] Collaborative Innovation Center for Birth Defect Research and Transformation of Shandong Province, Jining Medical University, Jining, Shandong, 272067, China
[c] Department of Nasopharyngeal Carcinoma, Sun Yat-sen University Cancer Center, Guangzhou, Guangdong, PR China
[d] Shandong Key Laboratory of Behavioral Medicine, School of Mental Health, Jining Medical University, Jining, Shandong, 272067, China

ARTICLE INFO

Article history:
Received 9 February 2021
Received in revised form
19 February 2021
Accepted 11 March 2021
Available online 24 March 2021

Keywords:
Hepatocellular carcinoma
HBV
HBV DNA integration
Integration frequency
Integration breakpoints

ABSTRACT

Hepatitis B virus (HBV) DNA integration is closely related to the occurrence of liver cancer. However, current studies mostly focus on the detection of the viral integration sites, ignoring the relationship between the frequency of viral integration and liver cancer. Thus, this study uses previous data to distinguish the breakpoints according to the integration frequency and analyzes the characteristics of different groups. This analysis revealed that three sets of breakpoints were characterized by its own integrated sample frequency, breakpoint distribution, and affected gene pathways. This result indicated an evolution in the virus integration sites in the process of tumor formation and development. Therefore, our research clarified the characteristics and differences in the sites of viral integration in tumors and adjacent tissues, and clarified the key signaling pathways affected by viral integration. Hence, these findings might be of great significance in the understanding of the role of viral integration frequency in hepatocellular carcinoma.

1. Introduction

It is very well-known that hepatitis B virus (HBV) has a very close relationship with liver cancer, especially HBV integration in the host DNA. In 2011, liver cancer was ranked second among the causes of malignant tumor mortality in the Chinese male population and fifth among the causes of malignant tumor mortality in the female one [1]. HBV patients have a higher risk of liver cancer than people not carrying this virus, and cancer risk is higher if HBV patients also have chronic liver disease and cirrhosis [2]. HBV-associated liver cancer is often accompanied with the occurrence of HBV integration, since integration is found in the 85%–90% of cancer samples from patients with HBV-related liver cancer [3,4], while HBV integration is detected only in the 30.7% of the adjacent normal tissue from the above patients [5]. HBV integration was first discovered in 1980 [6,7].

The understanding of the degree of HBV integration made great progress thanks to the development of second-generation sequencing technology. Indeed, the characteristics and functions of HBV integration sites on human genome were gradually uncovered by researchers and in 2012, researchers identified the close relationship between HBV integration and genomic instability using whole-genome sequencing technology [8]. Sung et al. used the same technology to investigate the characteristics of HBV integration in 88 pairs of liver cancer samples, and the results revealed the presence of hot-spot integration genes, in which, the ones with high-frequency were TERT, MLL4, and CCNE1. Thus, the expression of these hot-spot genes was affected by these HBV integration events [5]. Subsequently, Zhao et al. conducted a study on the regularity of HBV integration and clinical relevance in 426 matched liver cancer samples and normal ones, discovering that HBV integration prefer to occur in the hot spots region such as CPG and telomere region, than in other sites, and this result revealed why patients with HBV integration on the TERT gene usually have a poor prognosis [9]. However, despite more and more studies revealed the close relationship between HBV integration and tumorigenesis, the mechanism of HBV integration triggering tumorigenesis

* Corresponding author. Jining Medical University, Jining, Shandong, 272067, China. Tel.: 86-0537-3616220; fax: 86-0537-3616220.
E-mail addresses: 163.lwy@163.com (W. Li), qiyanweii@icloud.com (Y. Qi), xuhanshi@foxmail.com (H. Xu), aningvi@163.com (W. Wei), scucxf@163.com (X. Cui).
[1] These authors contribute equally to this work.

https://doi.org/10.1016/j.bbrc.2021.03.056

remains unclear. According to previous studies, at least three ways lead to liver cancer [1]: HBV integration into the genome causes chromosomal instability, structural variation, multiple site deletions, and increasing mutation rate [2]; HBV integration causes the changes in the expression of genes associated to cell proliferation and differentiation, and telomere reverse transcriptase [3]; HBV integration leads to the expression of key viral proteins which regulate liver cancer cell proliferation and differentiation [10–13].

The progress in the HBV integration research revealed that HBV integration events are not randomly occurring, but there is a specific localization and hot spot targeted regions. However, previous studies usually focus on the regular characteristics of HBV integration and ignore the proportion of cell population in which the same HBV integration event occurred. In theory, the proportion of the same integration event in the cell population can effectively reflect cell proliferation and tumor development. Therefore, this aspect is of great significance in the understanding of HBV integration. In this study, the HBV integration events were distinguished according to the proportion of the same integration event in different cell population to study the characteristics of HBV integration. The normalized support-reads number of each breakpoint (NORM value) was evaluated because it can be used as an index for the cell ratio of HBV integration. This value was positively correlated with the ratio of cells involved in HBV integration, and then the characteristics of the breakpoints with different NORM value (NORM value > 5; NORM <5) were analyzed. Our study found that the breakpoints of normal samples had lower NORM values than tumor samples. The breakpoints with higher NORM and lower NORM value in the tumor samples had different sample frequency and significant differences were found in signaling pathway distribution. In addition, the comparison of the breakpoints between tumor samples and normal samples revealed a significant difference in the hot spots genes, gene frequency, gene elements and pathways. These findings further indicated that the HBV integration event had different patterns in different stages of tumor development.

2. Materials and methods

The breakpoints of the HBV integration sites were taken from our previous study [9]. Breakpoints sourced from tissue samples of 426 HCC patients (Table S2, Table S3, Table S4). These breakpoints of tumor samples were divided into two groups according to the NORM value. The tumor high-NORM group is the one with breakpoints with a NORM value over than 5. Tumor low-NORM group is the one with breakpoints with a NORM value lower than 5. The normal group is the one with breakpoints with a NORM value lower than 5 in the adjacent normal tissues. The functional region of each breakpoint was annotated using ANNOVAR in hg19 coordinates [14]. Pathway enrichment analysis of the breakpoints was performed using the Cluster profiler software, a bioinformatic tool for intelligent, which performs statistical analysis [15,16]. The significance threshold for altered biological processes/pathways was set at a corrected hypergeometric *p*-value of 0.05. The region data of genomic elements were downloaded from the UCSC database [17]. The chi-squared test was used to detect the difference in breakpoint distribution between tumor and normal tissues. Thus, the difference in the distribution of gene frequency among the three groups with NORM values mentioned above was detected using the Kruskal-Wallis Test. The Mann-Whitney *U* test was also used to compare the difference in gene frequency between two groups.

2.1. Detection of HBV integration

The capture probes were designed according to the DNA sequences of 8 genotypes of HBV and synthesize by MyGenostics. The DNA from each sample was broken into fragments of 150–200 bp in length by Covaris M220 (Covaris Inc., Woburn, MA). These fragments were purified, end blunted, 'A' tailed and adaptor ligated to obtain DNA library. The capture of HBV DNA in the library and paired-end 90-bp read length sequencing was performed on the HiSeq 2000 sequencer according to the manufacturer's instructions (Illumina) [9]. HIVID pipeline was used for the detections of HBV integration breakpoints [18]. To minimize the impact of sequencing data and define the integration frequency, we used normalized number of support-reads (NORM value) as the index. Norm value was equal to the supported reads number of each HBV breakpoint per million clean read pairs. The breakpoints with Norm value ≥ 2 were retained.

2.2. Integration frequency

The integration frequency was calculated according to a previous report [5]. The integration frequency is the value of the number of support-reads versus the average depth of the upstream 1 kb distance. The consistent breakpoints between the whole genome sequencing and HBV capture data were sourced from our previous study [18]. Then, these consistent breakpoints were used to calculate the correlation.

2.3. Gene frequency

Since viral integration is considered a strong *cis*-activator of flanking genes and *cis*-acting enhancers can influence their target genes over long distances [19,20] (up to 1 Mb for upstream enhancers and 850 kb for downstream enhancers), breakpoints located at less than 500 kb from the annotated genes were included to calculate the affected gene frequency in HBV-integrated samples [21].

3. Results

3.1. The distribution among sample gene frequency of HBV integration

Our results revealed a positive correlation between the NORM value and the frequency of HBV integration ($r = 0.68$, p-value <0.01, Fig. 1a). Then, the sample gene frequency of HBV integration was calculated. No significant difference was found in the gene frequency between breakpoints from tumor samples with high NORM value and low NORM value. However, breakpoints from the two types of NORM value in the tumor samples were significantly different in gene frequency with the breakpoints from normal samples (Mann-Whitney *U* Test, p-value < 0.001, Fig. 1b).

3.2. HBV integration hotspots in tumor and normal samples

The sample gene frequency of HBV integration through breakpoints in the normal, tumor high-NORM group and tumor low-NORM group was also investigated (Table S1). The hotspots of HBV integration in the tumor high-NORM group were TERT (68), KMT2B (26), DUX4 [14], PLEKHG4B [13], DUX4L7 [11], EMBP1 [11], LOC100288778 [11], FAM138D [11] and PARD6G [10] (Fig. 2a). In contrast, the hotspots of HBV integration in the tumor low-NORM group were TERT (56), MIR6859-1, PLEKHG4B, MIR6859-2 (24), LOC100288778 [21], DDX11L1 [20], FAM138D [20], WASH7P [20], KMT2B [15], DUX4 [14], EMBP1 [14], PGBD2 [13], ANKRD30BL [12], ANKRD26P1 [11], ROCK1P1 [11] and DUX4L7 [11] (Fig. 2b). Finally, the hotspots of HBV integration were FN1 [13], ANKRD30BL [6], and GPR39 [5] in the normal group (Fig. 2c).

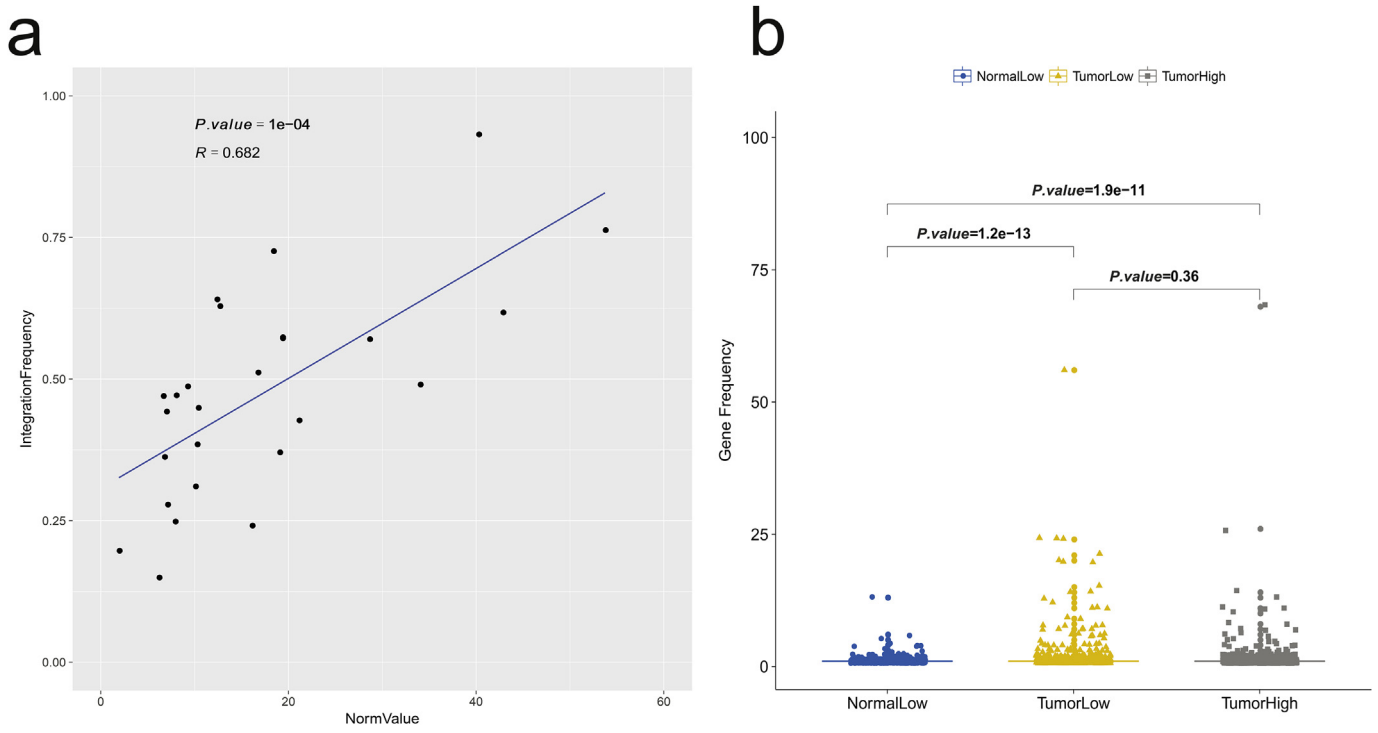

Fig. 1. Gene sample frequency of HBV integration. a) Correlation between NORM value and cell integration frequency. b) Difference in gene frequency in the three groups. Normal (blue dots), tumor low (yellow dots) and tumor high (grey dots) group. Each point represented the sample frequency of each gene. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

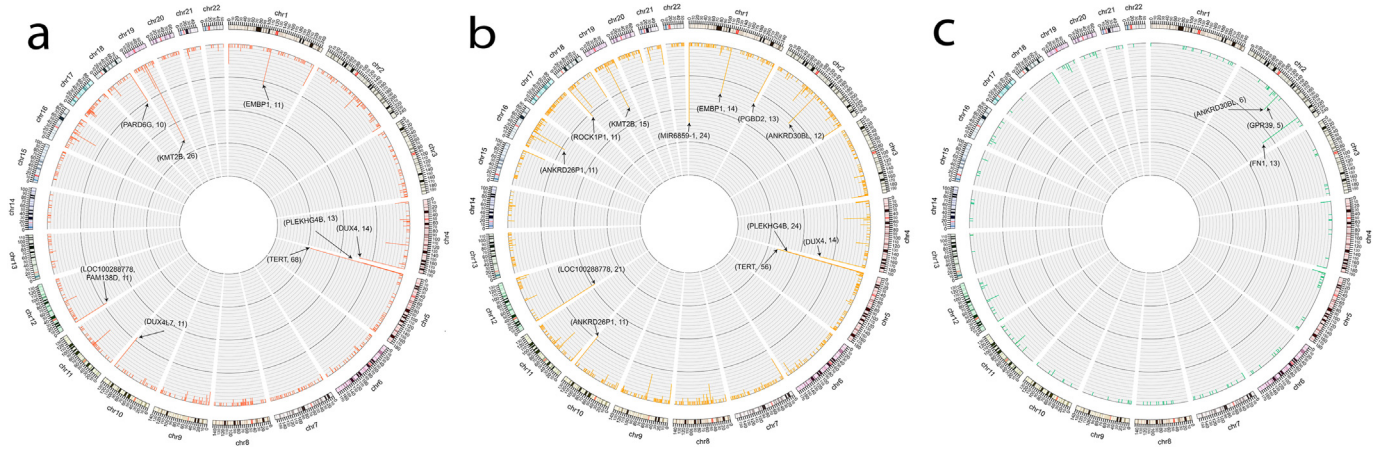

Fig. 2. Gene sample frequency of HBV integration in the three groups. a) Gene sample frequency in the tumor breakpoints with high-NORM value, b) gene sample frequency in the tumor breakpoints with low-NORM value, c) gene sample frequency in the normal samples. The figures show the difference in gene frequency in the three group samples. The height represented the sample frequency of the gene with HBV integration.

3.3. Distribution of breakpoints in different gene elements

The HBV breakpoints in the tumor high-NORM, tumor low-NORM and normal samples showed similar distribution in TFBS, Fragile, CPG sites (Table S2, Table S3, Table S4). However, The HBV integration sites in the tumor high-NORM samples were more prone to be located in Alu and CPG compared to the tumor low-NORM and normal samples ($p < 0.001$, Chi-squared test, Fig. 3a). The HBV breakpoints in the normal samples were significantly prone to be located in the LINE region than in the tumor samples ($p < 0.001$, Chi-squared test, Fig. 3a). No significant difference was observed among the three groups in the TFBS region. The distribution characteristics of the HBV breakpoints in normal, tumor low-NORM and tumor high-NORM group revealed that HBV breakpoints were more prone to be located in the INTERGENIC region in the tumor low-NORM and tumor high-NORM than in the normal group ($p < 0.001$, Chi-squared test, Fig. 3b), so as to promoter regions. However, HBV breakpoints were more prone to be located in the INTRON in the normal samples than the tumor samples ($p < 0.001$, Chi-squared test, Fig. 3b).

3.4. Pathway analysis

The results revealed that the enriched pathway was the one of

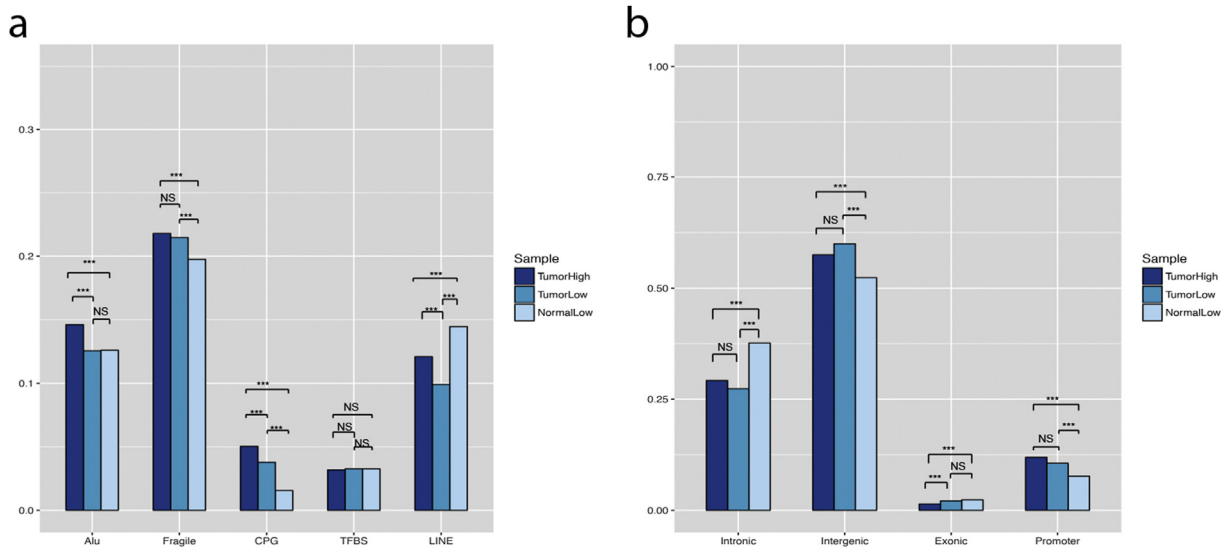

Fig. 3. Distribution ratio of breakpoints. a) Distribution of breakpoints in genome element, b) Distribution of breakpoints in genetic elements. The X axis represents different regions; the Y axis represents the ratio of the integration breakpoints. The normal samples (light blue) and the tumor samples (high-NORM group: dark blue; low-NORM group: sky blue) percentages of breakpoints are shown. The *p* values were calculated by Chi-squared test and corrected by Fisher exact test. ***$p < 0.001$. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Axon guidance in the tumor high-NORM group (Fig. 4, TableS5). The enriched pathways in the tumor low-NORM group were Human papillomavirus infection, Retrograde endocannabinoid signaling, and PI3K-Akt signaling pathway (Fig. 4,TableS6). However, the main enriched pathways of breakpoints from DNA samples were the Calcium signaling pathway, Inflammatory mediator regulation of TRP channels, p53 signaling pathway, and Glioma and Axon guidance (Fig. 4, TableS7).

4. Discussion

In this study, a positive correlation was found between the NORM value and the cell HBV integration ratio. Therefore, the NORM value could be used as an index to evaluate the cell ratio of HBV integration. According to this study, the integration sites distinguished according to the NORM value were used to investigate the characteristics of the sample gene frequency of HBV integration, the gene distribution and the involved pathways.

The comparison of the gene frequencies of the breakpoints among different groups revealed that the integration sites in groups with different NORM values had different sample gene frequency. The adjacent normal tissues generally had lower sample frequency and cell integration ratio (NORM value), which is consistent with the early stage of tumor with a lower cell ratio of monoclonal propagation. The breakpoints of different NORM value in cancer tissues had different hots spot genes and sample frequency, and it might indicate that new HBV integration events might be continuously generated through cell propagation. The ratio of breakpoint enrichment increased with the increase of the NORM value in the CPG and PROMOTER region, potentially suggesting that breakpoints played a certain evolutionary role in tumor development. Our study indicated that the occurrence of viral integration events was closely related to tumorigenesis, although the function of HBV integration remains unclear. Moreover, the enrichment of breakpoints in some pathways became more and more evident as the NORM value increased. The study revealed an evident different tendency for HBV integration in different groups. Hence, our study not only provided new ideas in the research of HBV integration but also provided a theoretical basis for the relationship between liver cancer and HBV integration.

Funding

This study was supported by Research Fund for Natural Science Foundation of Shandong Province (Grant ID:ZR2019MC024), Lin He's Academician Workstation of New Medicine and Clinical Translation, Genomics Innovation Team of Systemic Lupus Erythematosus of Jining Medical University, and the Teacher Support Funding of Jining Medical University (Grant ID:JYFC2018KJ034).

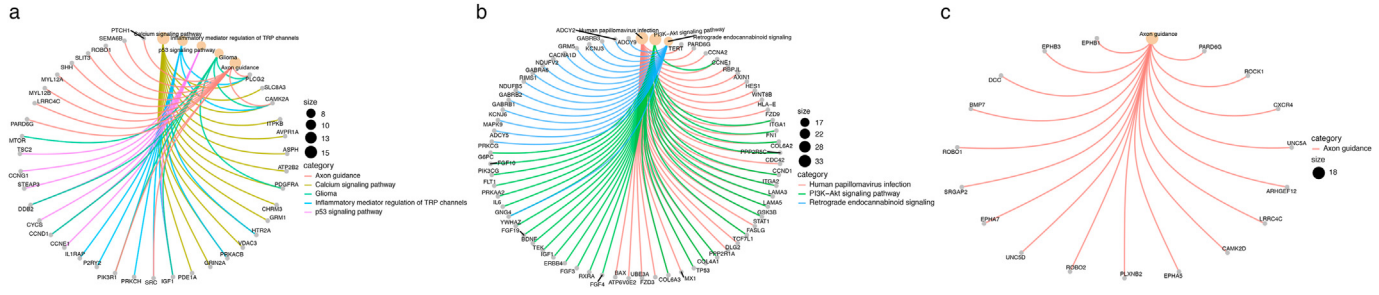

Fig. 4. The key pathways in the three groups. a) Normal group b) Tumor Low-NORM group c) Tumor High-NORM group.

Ethical approval and informed consent

All procedures performed in this study involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The study was approved by the Ethics Review Committee of the Jining Medical University and informed consent was obtained from all subjects.

Author contributions

Xiaofang Cui and Weiyang Li conceived and designed the research. Y.W.Q, H.S.X and W.W performed the experiments and interpreted the data. X.F.C., W.Y.L wrote the main manuscript. All authors reviewed the manuscript.

Disclosure of competing interest

The authors declare no competing interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbrc.2021.03.056.

References

[1] W. Chen, Cancer statistics: updated cancer burden in China, Chin. J. Canc. Res. 27 (1) (2015) 1, https://doi.org/10.3978/j.issn.1000-9604.2015.02.07.
[2] T. Ishikawa, Clinical features of hepatitis B virus-related hepatocellular carcinoma, World J. Gastroenterol. 16 (20) (2010) 2463–2467, https://doi.org/10.3748/wjg.v16.i20.2463.
[3] C. Brechot, D. Gozuacik, Y. Murakami, P. Paterlini-Brechot, Molecular bases for the development of hepatitis B virus (HBV)-related hepatocellular carcinoma (HCC), Semin. Canc. Biol. 10 (3) (2000) 211–231, https://doi.org/10.1006/scbi.2000.0321.
[4] M. Minami, Y. Daimon, K. Mori, H. Takashima, T. Nakajima, Y. Itoh, T. Okanoue, Hepatitis B virus-related insertional mutagenesis in chronic hepatitis B patients as an early drastic genetic change leading to hepatocarcinogenesis, Oncogene 24 (27) (2005) 4340–4348, https://doi.org/10.1038/sj.onc.1208628.
[5] W.K. Sung, H. Zheng, S. Li, R. Chen, X. Liu, Y. Li, N.P. Lee, W.H. Lee, P.N. Ariyaratne, C. Tennakoon, F.H. Mulawadi, K.F. Wong, A.M. Liu, R.T. Poon, S.T. Fan, K.L. Chan, Z. Gong, Y. Hu, Z. Lin, G. Wang, Q. Zhang, T.D. Barber, W.C. Chou, A. Aggarwal, K. Hao, W. Zhou, C. Zhang, J. Hardwick, C. Buser, J. Xu, Z. Kan, H. Dai, M. Mao, C. Reinhard, J. Wang, J.M. Luk, Genome-wide survey of recurrent HBV integration in hepatocellular carcinoma, Nat. Genet. 44 (7) (2012) 765–769, https://doi.org/10.1038/ng.2295.
[6] P.R. Chakraborty, N. Ruiz-Opazo, D. Shouval, D.A. Shafritz, Identification of integrated hepatitis B virus DNA and expression of viral RNA in an HBsAg-producing human hepatocellular carcinoma cell line, Nature 286 (5772) (1980) 531–533, https://doi.org/10.1038/286531a0.
[7] C. Brechot, C. Pourcel, A. Louise, B. Rain, P. Tiollais, Presence of integrated hepatitis B virus DNA sequences in cellular DNA of human hepatocellular carcinoma, Nature 286 (5772) (1980) 533–535, https://doi.org/10.1038/286533a0.
[8] Z. Jiang, S. Jhunjhunwala, J. Liu, P.M. Haverty, M.I. Kennemer, Y. Guan, W. Lee, P. Carnevali, J. Stinson, S. Johnson, J. Diao, S. Yeung, A. Jubb, W. Ye, T.D. Wu, S.B. Kapadia, F.J. de Sauvage, R.C. Gentleman, H.M. Stern, S. Seshagiri, K.P. Pant, Z. Modrusan, D.G. Ballinger, Z. Zhang, The effects of hepatitis B virus integration into the genomes of hepatocellular carcinoma patients, Genome Res. 22 (4) (2012) 593–601, https://doi.org/10.1101/gr.133926.111.
[9] L.H. Zhao, X. Liu, H.X. Yan, W.Y. Li, X. Zeng, Y. Yang, J. Zhao, S.P. Liu, X.H. Zhuang, C. Lin, C.J. Qin, Y. Zhao, Z.Y. Pan, G. Huang, H. Liu, J. Zhang, R.Y. Wang, Y. Yang, W. Wen, G.S. Lv, H.L. Zhang, H. Wu, S. Huang, M.D. Wang, L. Tang, H.Z. Cao, L. Wang, T.L. Lee, H. Jiang, Y.X. Tan, S.X. Yuan, G.J. Hou, Q.F. Tao, Q.G. Xu, X.Q. Zhang, M.C. Wu, X. Xu, J. Wang, H.M. Yang, W.P. Zhou, H.Y. Wang, Genomic and oncogenic preference of HBV integration in hepatocellular carcinoma, Nat. Commun. 7 (2016) 12992, https://doi.org/10.1038/ncomms12992.
[10] J. Zucman-Rossi, P. Laurent-Puig, Genetic diversity of hepatocellular carcinomas and its potential impact on targeted therapies, Pharmacogenomics 8 (8) (2007) 997–1003, https://doi.org/10.2217/14622416.8.8.997.
[11] A.J. Gehring, Z.Z. Ho, A.T. Tan, M.O. Aung, K.H. Lee, K.C. Tan, S.G. Lim, A. Bertoletti, Profile of tumor antigen-specific CD8 T cells in patients with hepatitis B virus-related hepatocellular carcinoma, Gastroenterology 137 (2) (2009) 682–690, https://doi.org/10.1053/j.gastro.2009.04.045.
[12] T. Tu, M.A. Budzinska, N.A. Shackel, A.R. Jilbert, Conceptual models for the initiation of hepatitis B virus-associated hepatocellular carcinoma, Liver Int. 35 (7) (2015) 1786–1800, https://doi.org/10.1111/liv.12773.
[13] T. Tu, M.A. Budzinska, N.A. Shackel, S. Urban, HBV DNA integration: molecular mechanisms and clinical implications, Viruses 9 (4) (2017), https://doi.org/10.3390/v9040075.
[14] K. Wang, M. Li, H. Hakonarson, ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data, Nucleic Acids Res. 38 (16) (2010) e164, https://doi.org/10.1093/nar/gkq603.
[15] G. Yu, L.G. Wang, Y. Han, Q.Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters, OMICS 16 (5) (2012) 284–287, https://doi.org/10.1089/omi.2011.0118.
[16] W. Huang da, B.T. Sherman, R.A. Lempicki, Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists, Nucleic Acids Res. 37 (1) (2009) 1–13, https://doi.org/10.1093/nar/gkn923.
[17] M. Haeussler, A.S. Zweig, C. Tyner, M.L. Speir, K.R. Rosenbloom, B.J. Raney, C.M. Lee, B.T. Lee, A.S. Hinrichs, J.N. Gonzalez, D. Gibson, M. Diekhans, H. Clawson, J. Casper, G.P. Barber, D. Haussler, R.M. Kuhn, W.J. Kent, The UCSC Genome Browser database: 2019 update, Nucleic Acids Res. 47 (D1) (2019) D853–D858, https://doi.org/10.1093/nar/gky1095.
[18] W. Li, X. Zeng, N.P. Lee, X. Liu, S. Chen, B. Guo, S. Yi, X. Zhuang, F. Chen, G. Wang, R.T. Poon, S.T. Fan, M. Mao, Y. Li, S. Li, J. Wang, X. Xu, H. Jiang, X. Zhang, HIVID: an efficient method to detect HBV integration using low coverage sequencing, Genomics 102 (4) (2013) 338–344, https://doi.org/10.1016/j.ygeno.2013.07.002.
[19] L.A. Lettice, S.J. Heaney, L.A. Purdie, L. Li, P. de Beer, B.A. Oostra, D. Goode, G. Elgar, R.E. Hill, E. de Graaff, A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly, Hum. Mol. Genet. 12 (14) (2003) 1725–1735, https://doi.org/10.1093/hmg/ddg180.
[20] L. Li, J.A. Zhang, M. Dose, H.Y. Kueh, R. Mosadeghi, F. Gounari, E.V. Rothenberg, A far downstream enhancer for murine Bcl11b controls its T-cell specific expression, Blood 122 (6) (2013) 902–911, https://doi.org/10.1182/blood-2012-08-447839.
[21] Z. Hu, D. Zhu, W. Wang, W. Li, W. Jia, X. Zeng, W. Ding, L. Yu, X. Wang, L. Wang, H. Shen, C. Zhang, H. Liu, X. Liu, Y. Zhao, X. Fang, S. Li, W. Chen, T. Tang, A. Fu, Z. Wang, G. Chen, Q. Gao, S. Li, L. Xi, C. Wang, S. Liao, X. Ma, P. Wu, K. Li, S. Wang, J. Zhou, J. Wang, X. Xu, H. Wang, D. Ma, Genome-wide profiling of HPV integration in cervical cancer identifies clustered genomic hot spots and a potential microhomology-mediated integration mechanism, Nat. Genet. 47 (2) (2015) 158–163, https://doi.org/10.1038/ng.3178.