

*Full Length Research Paper*

# Identification of recombinants of hepatitis B virus genotypes C from Hong Kong China

Qianqian Liu and Zeng Tu\*

Department of Microbiology, College of Basic Medical Sciences, Chongqing Medical University, Chongqing 400016, China.

Received 10 December, 2020; Accepted 8 February, 2021

**Hepatitis B virus (HBV) infections severely threaten the health of the human being. Frequent recombination of HBV within or between genotypes was reported to favor the viral evolution and adaption and the recombination may cause more severe clinical symptoms. In the study, we collected genotype C HBV from five Asian countries and detected their possible recombination events by bioinformatic analysis. There are two main subtypes, C1 and C2, within the C genotypes among the collected data in the study. Subtype C1 is most prevalent in Cambodia, Bangladesh, and China, while C2 is prevalent in Japan, Indonesia, and a small part of China. Three recombination events were detected and verified from C1 genotype HBV from Hong Kong China as demonstrated by recombinant (KJ410515), ranging from 2381 to 1861 nt. Recombinant events were detected and verified by recombination analysis in the study. It is important to filter possible recombinants when using the online-genbank data to do phylogenetic analysis.**

**Key words:** Hepatitis B virus, subgenotype C1, genotype C, subgenotype C2, recombinant analysis, HBV Hong Kong, China.

## INTRODUCTION

Hepatitis B virus (HBV) infection severely threatens the health of the human being. One-third of the world's population is infected with HBV, including 350 million of them suffering from chronic HBV infection (Lee, 1997). Thirty percent of the infected person even died from HBV-related liver disease. HBV belongs to the genus *Orthohepadnavirus*; the HBV genome is an incomplete double-stranded circular DNA molecule with 3.2 kp that encodes four overlapping open reading frames (orfs), including preC/C, preS1/preS2/S, and X genes (Westover

and Hughes, 2007; Zhang et al., 2010). Due to the lack of proofreading activity of viral polymerase during the reverse transcription step of genome replication, HBV genetic variability is high and leads to differences in nucleotide sequence (Arauz-Ruiz et al., 2002; Norder et al., 1994; Okamoto et al., 1988; Stuyver et al., 2000).

Therefore, eight genotypes (A-H) have been established based on the 7.5% inter-ethnic differences in the entire nucleotide sequence (Arauz-Ruiz et al., 2002; Norder et al., 1994; Okamoto et al., 1988; Stuyver et al., 2000).

\*Corresponding author. E-mail: [tuz1980@126.com](mailto:tuz1980@126.com).

Later, two other genotypes (I and J) were initially proposed for the HBV strains found in Vietnam and Laos (Tran et al., 2008; Olinger et al., 2008).

Gene recombination is an important mechanism of virus evolution and challenges the designation of vaccines and antiviral treatment strategies (Yang et al., 2006). The characteristics of life-long persistent infection and the prevalence of different genotypes in the same area provide a higher probability of co-infection of different genotypes in a host, thus having a high risk of viral recombination. In Northwestern China, recombinants between HBV genotypes were frequently reported in the past decade (Wang et al., 2005; Zhou et al., 2011). Subgroup Ba of genotype B, which is recombinant with genotype C, is found primarily in Southeast Asian countries and seems to be more pernicious than subgroup Bj (Huy et al., 2004). In the past, the strain of HBV genotype F3/A1 was found in the Afro-Colombian population (Alvarado-Mora et al., 2012). Evidence was supplied to prove the recombination of genotype C/D in Western China (Zhou et al., 2012) and even detected new subtype D9, which is recombined by genotypes C and D (Ghosh et al., 2013). However, there is no much research to discuss the recombination between different subtypes in the same genotype in Asia, so the goal is to research the recombination between HBV C genotype in five countries of Asia (China, Cambodia, Indonesia, Japan, and Bangladesh).

In the study, we collected 131 complete genomes of HBV genotype C in five Asian countries and constructed their phylogenetic tree. Then recombination events were detected and confirmed. Here, we reported recombination events detected in our dataset.

## MATERIALS AND METHODS

### Sequence download

All of HBV complete genomes available before 6 November 2020 were downloaded from NCBI (<https://www.ncbi.nlm.nih.gov/>) Nucleotide Database. "HBV genotype C complete genome" and "HBV C" and "China" were used as keywords research terms, and the search results were filtered by sequence length from 2800 to 3500 nt and genotype C. 131 complete genome sequence were retrieved from the research records. After those sequences were aligned and analyzed, we had only left 12 full-sequences suspected as the recombinant genome. All complete genome can be obtained by searching for the accession number on GenBank.

### Sequence alignment and recombinants identification

The original datasets were aligned by MUSCLE listed in the Molecular Evolutionary Genetics Analysis (MEGA) version X software, and MEGA conducted phylogenetic analyses (Radjef et al., 2004). To find potential recombination signals in those isolates, recombination analysis was performed using the RDP4 program (Wang et al., 2015) Using seven algorithms, including RDP, GENECONV, MaxChi, Chimaera, 3eq, SiSican, and bootscan, 12 recombination events were detected in all 131 isolates (Table 1). In

all 12 potential recombination isolates, three recombination sequences had remarkable high certainty with a P-value of at least three algorithms  $<1 \times 10^{-6}$ , and two potential recombination strains have a high degree of certainty on account of recombinant score  $> 0.6$  (Table 1) (Martin et al., 2010). Besides, one recombination sequences have a fair probability since the recombinant score is between 0.4 and 0.6 (Wang et al., 2005). The similarity plot and bootscanning analyses were performed by Simplot software version 3.5.1. At last, only 3 sequences were proved actually recombined. The neighbor-joining of MEGAX established a Phylogenetic tree to manifest the relationship between the related fragment of recombinant isolate and their major or minor parents (Pérez et al., 2014).

## RESULTS

A total of 187 original full-length sequences of five countries (China, Cambodia, Indonesia, Japan, and Bangladesh) in Asia were retrieved, but 131 valid sequences were finally included after removing the no-subtype classification. After collating the data, we found that C1 and C2 subtypes are more common than other subtypes. Among the C genotypes, subtype C1 is prevalent in Cambodia, and most of China and Japan and a small part of China are prevalent subtype C2 in the five countries.

MEGA X ran preliminary multiple sequence alignments and the RDP4 program performed our recombinant analysis of full-sequences and recombination analysis. Twelve potential recombinations were detected in all 131 isolates. In all 12 potential recombination sequences, three recombination sequences had a prominent high degree of certainty with a P-value of at least three algorithms  $<1 \times 10^{-6}$  and two-three potential recombination sequences with a high degree of certainty on account of recombinant score  $> 0.6$  (Table 1). Also, because the recombination score is between 0.4 and 0.6, a recombination sequence has a considerable probability.

Potential recombinant isolates KJ410515 were analyzed using the Simplot 3.5.1 program with a 200-nt window moving in 20-nt steps. In the similarity plot analysis of recombinations, the horizontal axis represents the nucleotide position of the midpoint of the window from the 5' end of the query sequence (nt 2381 in HBV-KJ410515) (Figure 1). The vertical axis represents the similarity between the analyzed sequence and the reference sequence. KJ410515 displayed the highest similarity with KJ410505 isolate more than 90% from beginning to about 2381 nucleotide position. However, at 2381-2861 nucleotide positions, the sequence similarities were comparatively higher with the FJ562308 strain isolated. The analysis for KJ410515 demonstrated that recombination sites were probably located at 2381 to 2861 nucleotide positions, which was most closely related to FJ562308. Bootscanning analysis confirmed the recombination of strain KJ410515. The recombinant gene segment of Simplot is in keeping with the consequence of RDP4. Similarly, the recombinant regions of KJ410508 and KJ410512 were detected and

**Table 1.** Summary of possible recombination events in HBV genotype C isolates identified by RDP4. The isolates beginning with KJ are all from Hong Kong, China.

Event number	Recombinant sequence(s)	Breakpoint position Begin/End	Parental sequence(s) Major/Minor	Recombinant score	P-Value for the six detection methods in RDP4						
					RDP	GENECONV	BootScan	MaxChi	SiScann	3eq	Chimaera
1	KJ410512	1869/3220	KJ410510/MG571355	0.737	4.62E-19	1.33E-17	4.04E-10	6.80E-22	1.21E -15	8.59E -49	4.32E -19
2	KJ410508	427/2370	KJ410510/KJ410505	0.690	NS	1.06E-02	NS	5.86E-08	8.01E-06	2.74E-11	4.39E-09
3	KJ410515	2381/2861	KJ410505/FJ562308	0.728	9.58E-04	2.80E-05	1.25E-03	NS	4.82E-05	2.78E-02	NS

located at 427-2370 and 1869-3220 nt, respectively (Table 1).

To further analyze the recombination of KJ410515, a phylogenetic incongruence analysis was performed using major parent and minor parent by the neighbor-joining of MEGAX for recombinant sequences and performed bootstrapping with 1000 repeats. Genome sequences were divided into two alignments, and an independent phylogenetic tree was constructed for each data set (Figure 2). The phylogenetic trees constructed with the corresponding fragments of KJ410505 and FJ562308 confirmed the chimeric pattern found in the genome of the KJ410515 strain. Most of the recombinant strain fragments were derived from a KJ410505 strain, and the middle part of the fragments was from an FJ562308 strain (Figure 2). KJ410515 Phylogenetic analysis of these sequences confirmed their relatedness, as they all possess a high bootstrap value (BV). As shown in Figure 2A, KJ410515 and its major parent KJ410505 clustered together at 0-2381 and 2861-3220, and in the nucleotide fragments of 2381-2861, KJ410515 and FJ562308 were clustered together (Figure 2b). It shows a close relationship of KJ410515 with KJ410505 and FJ562308, which indicated the reasonability of recombination. From Table 2, basic information of KJ410515 and its parents can be obtained. The collection date and countries increase the reasonability of the recombination of KJ410515.

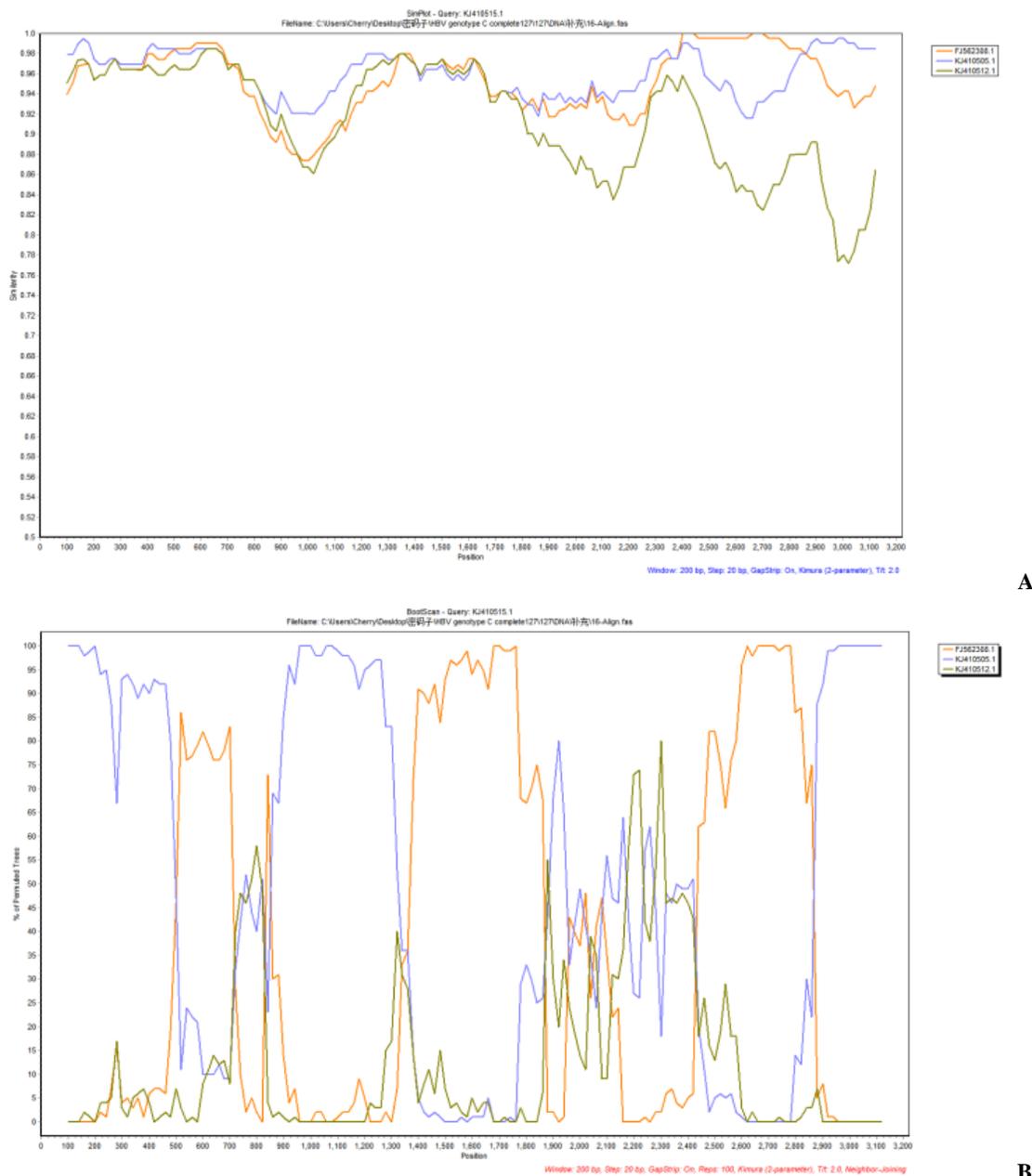
## DISCUSSION

HBV infection is endemic in many Asian countries, especially in China. HBV is endowed with a variety of genetic diversity. The long-term prevalence of HBV infection favors the frequent occurrence of genetic mutation and recombination within or between genotypes or subgenotypes. Rich HBV data available online encourages us to investigate the possible recombination event, originally ignored by the submitter. In the study, we collected 131 full-length HBV genotype C from five Asian countries and found three underdescribed recombinations (KJ410515, KJ410508, and KJ410512) in isolates from Hong Kong, China using RDP4 recombination detection software. The recombination regions are spanning or within the Core or P gene region, a common region for reported variation and recombination (Kay and Zoulim, 2007; Araujo, 2015). Different locations on a phylogenetic tree of non-recombinant and recombinant regions confirmed recombination events, as shown by the case of KJ410515. Those recombinations may have functional roles for the recombination region of one of the recombinant strains KJ410512 which is highly similar to the Ba subtype of type B/C recombination of the preC region and this type of recombinant strain is widely popular until now, and the Ba subtype HBV can cause more serious clinical diseases (Sugauchi et al., 2002). The significance of HBV recombination deserves

investigation.

The recombination of KJ410515 is located in the P gene region, and its function is to encode polymerase. The HBV Pol protein is the focus of basic research and translational research. The Viral polymerase is the target of all current HBV drug therapies (except interferon), and it is the only area that is usually sequenced during treatment escape (Rhee et al., 2010; Buhlig et al., 2020). Genetic variation resulting from recombination can allow immune escape and treatment resistance. An intergenotypic B/I recombinant and B/C recombinants, the new subtype C17 (Feng et al., 2020), this type of recombination between different genotypes are prevalent, but there are few studies on recombination between different subtypes.

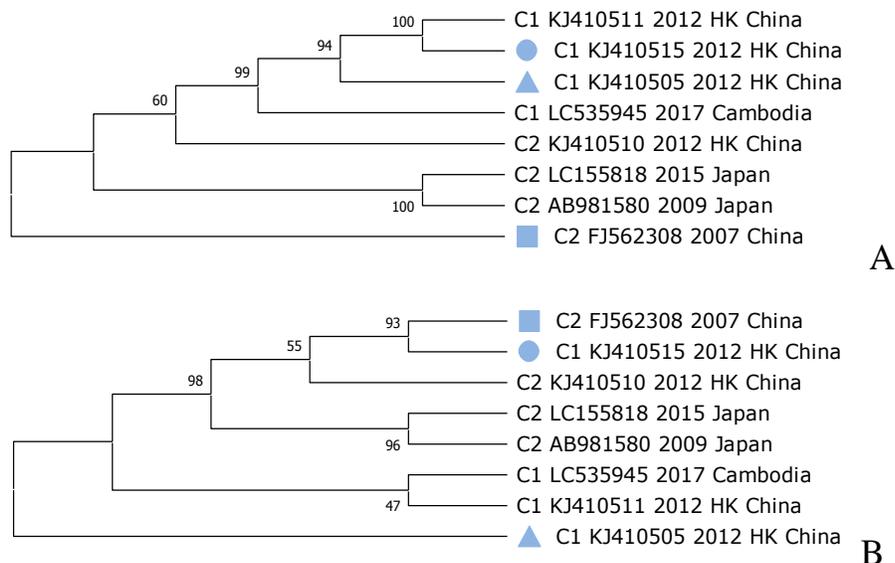
Genetic recombination is an important mechanism for virus evolution, and it can be found in some studies, and it is common for viruses to undergo genetic recombination. In the past years, C/D inter-genotype recombinants of type I (breakpoints at nt 50 and 1450) and type II (breakpoints at nt 10 and 799) have been reported from the Qinghai-Tibet Plateau (Wang et al., 2005). Moreover, in 2011, the prevalence of recombinants between C/D genotypes in chronic hepatitis B patients in Northwest China was declared to be high (Zhou et al., 2011). B/C inter-genotype recombination has also been reported in Thailand, Vietnam, Indonesia, and South China (Luo et al., 2004). The emergence of recombinants



**Figure 1.** Similarity plot and bootscanning analysis of the complete genome sequences of the KJ410515 (A, B) strains. Gene structure organization and bootscanning analysis (upper panel, A), similarity plot (lower panel, B) of complete HBV genomes, using a sliding window of 200 nt moving in 20-nt steps. The KJ410515 isolate was used as a query sequence. For each bootscanning analysis, the names of the viruses used as the reference strains were indicated in the square's upper right corner.

may change the epidemiology of the virus. A few HBV/E strains also were identified as minority genotypes further east in Mozambique and Madagascar or further north in Tunisia (Kramvis et al., 2005). However, phylogenetic analysis showed that genotype E was dominant in Sudan (51%), followed by genotype D (41.5%). Genetic recombination probably produces new subtypes. The HBV/D-E recombinant of a new HBV/D subtype that has been identified and spread in Nigeria is considered a new

subtype D8 (Abdou Chekaraou et al., 2010). In short, genetic recombination as an important way of virus evolution may change its clinical characteristics and epidemiology, and the identification of genetic recombination is only the first step to detect the emergence and prevalence of new subtypes. However, only the emergence of dangerous recombinants is found as early as possible to control the virus infection further. Therefore, the identification of recombination is of great



**Figure 2.** Phylogenetic trees of HBV genotype C isolates. A, a Phylogenetic tree based on the fragment of 0-2381 and 2861-3220 (major parent) in isolates. B Phylogenetic tree based on the fragment of 2381-2862 (minor parent) in isolates. Trees were constructed with a MEGA software package (v5.0) using the neighbor-joining method and maximum likelihood with 1000 bootstrap. The guide value is indicated at the node. Three different patterns ●, ▲ and ■ before the title, represent recombinant isolate, major parent and minor parent, respectively.

**Table 2.** Basic information about KJ410515 and its parent strain.

Event number	Isolate	Collection date	Country	Genotype
1	KJ410515	24-Jan-2012	Hong Kong (China)	C1
2	KJ410505	24-Jan-2012	Hong Kong (China)	C1
3	FJ562308	10-Cct-2007	Bei Jing (China)	C2

significance.

In the present dataset, 12 potential recombinant sequences were detected within 131 isolates, indicating a high possibility of recombination with the submitted sequences. Thus, it is of significance to screen recombination when using the NCBI data to inquire about the evolution of the HBV. One limitation of our research is that the reliability of recombination could be improved by more detection methods, and the origin and evolution of specific recombination need to be investigated.

## CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

## REFERENCES

Abdou Chekaraou M, Brichler S, Mansour W, Le Gal F, Garba A, Dény P, Gordien E (2010). A novel hepatitis B virus (HBV) subgenotype D (D8) strain, resulting from recombination between genotypes D and

E, is circulating in Niger along with HBV/E strains. *Journal of General Virology* 91(6):1609-1620.

Alvarado-Mora MV, Romano CM, Gomes-Gouvêa MS, Gutierrez MF, Carrilho FJ, Pinho JR (2012). Phylogenetic analysis of complete genome sequences of hepatitis B virus from an Afro-Colombian community: presence of HBV F3/A1 recombinant strain. *Virology Journal* 9:244.

Araujo NM (2015). Hepatitis B virus intergenotypic recombinants worldwide: An overview. *Infection, Genetics and Evolution* 36:500-510.

Arauz-Ruiz P, Norder H, Robertson BH, Magnius LO (2002). Genotype H: a new Amerindian genotype of hepatitis B virus revealed in Central America. *Journal of General Virology* 83(8):2059-2073.

Buhlig TS, Bowersox AF, Braun DL, Owsley DN, James KD, Aranda AJ, Kendrick CD, Skalka NA, Clark DN (2020). Molecular, evolutionary, and structural analysis of the terminal protein domain of hepatitis B virus polymerase, a potential drug target. *Viruses* 12(5):570.

Feng Y, Ran J, Feng YM, Miao J, Zhao Y, Jia Y, Li Z, Yue W, Xia X (2020). Genetic diversity of hepatitis B virus in Yunnan, China: identification of novel subgenotype C17, an intergenotypic B/I recombinant, and B/C recombinants. *Journal of General Virology* 101(9):972-981.

Ghosh S, Banerjee P, Deny P, Mondal RK, Nandi M, Roychoudhury A, Das K, Banerjee S, Santra A, Zoulim F, Chowdhury A, Datta S (2013). New HBV subgenotype D9, a novel D/C recombinant, identified in patients with chronic HBeAg-negative infection in Eastern

- India. *Journal of Viral Hepatitis* 20(3):209-218.
- Huy TT, Ushijima H, Quang VX, Win KM, Luengrojanakul P, Kikuchi K, Sata T, Abe K (2004). Genotype C of hepatitis B virus can be classified into at least two subgroups. *Journal of General Virology* 85(2):283-292.
- Kay A, Zoulim F (2007). Hepatitis B virus genetic variability and evolution. *Virus Research* 127(2):164-176.
- Kramvis A, Restorp K, Norder H, Botha JF, Magnus LO, Kew MC (2005). Full genome analysis of hepatitis B virus genotype E strains from South-Western Africa and Madagascar reveals low genetic variability. *Journal of Medical Virology* 77(1):47-52.
- Lee WM (1997). Hepatitis B virus infection. *The New England Journal of Medicine* 11; 337(24):1733-1745.
- Luo K, Liu Z, He H, Peng J, Liang W, Dai W, Hou J (2004). The putative recombination of hepatitis B virus genotype B with pre-C/C region of genotype C. *Virus Genes* 29(1):31-41.
- Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefevre P (2010). RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics* 26(19):2462-2463.
- Norder H, Couroucé AM, Magnus LO (1994). Complete genomes, phylogenetic relatedness, and structural proteins of six strains of the hepatitis B virus, four of which represent two new genotypes. *Virology* 198(2):489-503.
- Okamoto H, Tsuda F, Sakugawa H, Sastrosoewignjo RI, Imai M, Miyakawa Y, Mayumi M (1988). Typing hepatitis B virus by homology nucleotide sequence: comparison of surface antigen subtypes. *Journal of General Virology* 69(10):2575-2583.
- Olinger CM, Jutavijittum P, Hübschen JM, Yousukh A, Samounry B, Thammavong T, Toriyama K, Muller CP (2008). Possible new hepatitis B virus genotype, southeast Asia. *Emerging Infectious Diseases* 14(11):1777-1780.
- Pérez R, Calleros L, Marandino A, Sarute N, Iraola G, Grecco S, Blanc H, Vignuzzi M, Isakov O, Shomron N, Carrau L, Hernández M, Francia L, Sosa K, Tomás G, Panzera Y (2014). Phylogenetic and genome-wide deep-sequencing analyses of canine parvovirus reveal co-infection with field variants and emergence of a recent recombinant strain. *PLoS One* 9(11):e111779.
- Radjef N, Gordien E, Ivaniushina V, Gault E, Anaïs P, Drugan T, Trinchet JC, Roulot D, Tamby M, Milinkovitch MC, Dény P (2004). Molecular phylogenetic analyses indicate a wide and ancient radiation of African hepatitis delta virus, suggesting a deltavirus genus of at least seven major clades. *Journal of Virology* 78(5):2537-2544.
- Rhee SY, Margeridon-Thermet S, Nguyen MH, Liu TF, Kagan RM, Beggel B, Verheyen J, Kaiser R, Shafer RW (2010). Hepatitis B virus reverse transcriptase sequence variant database for sequence analysis and mutation discovery. *Antiviral Research* 88:269-275.
- Stuyver L, De Gendt S, Van Geyt C, Zoulim F, Fried M, Schinazi RF, Rossau R (2000). A new genotype of hepatitis B virus: complete genome phylogenetic relatedness. *Journal of General Virology* 81(1):67-74.
- Sugauchi F, Orito E, Ichida T, Kato H, Sakugawa H, Kakumu S, Ishida T, Chutaputti A, Lai CL, Ueda R, Miyakawa Y, Mizokami M (2002). Hepatitis B virus of genotype B with or without recombination with genotype C over the precore region plus the core gene. *Journal of General Virology* 76(12):5985-5992.
- Tran TT, Trinh TN, Abe K (2008). New complex recombinant genotype of hepatitis B virus identified in Vietnam. *Journal of Virology* 82(11):5657-5663.
- Wang D, Yu C, Wang G, Shi K, Li F, Yuan X (2015). Phylogenetic and recombination analysis of Tobacco bushy top virus in China. *Virology Journal* 12:111.
- Wang Z, Liu Z, Zeng G, Wen S, Qi Y, Ma S, Naoumov NV, Hou J (2005). A new intertype recombinant between genotypes C and D of hepatitis B virus identified in China. *Journal of General Virology* 86(4):985-990.
- Westover KM, Hughes AL (2007). Evolution of cytotoxic T-lymphocyte epitopes in hepatitis B virus. *Infection, Genetics and Evolution* 7(2):254-262.
- Yang J, Xing K, Deng R, Wang J, Wang X (2006). Identification of Hepatitis B virus putative intergenotype recombinants by using fragment typing. *Journal of General Virology* 87(8):2203-2215.
- Zhang D, Chen J, Deng L, Mao Q, Zheng J, Wu J, Zeng C, Li Y (2010). Evolutionary selection associated with the multi-function of overlapping genes in the hepatitis B virus. *Infection, Genetics and Evolution* 10(1):84-88.
- Zhou B, Wang Z, Yang J, Sun J, Li H, Tanaka Y, Mizokami M, Hou J (2012). Novel evidence of HBV recombination in family cluster infections in western China. *PLoS ONE* 7(6):e38241.
- Zhou B, Xiao L, Wang Z, Chang ET, Chen J, Hou J (2011). Geographical and ethnic distribution of the HBV C/D recombinant on the Qinghai-Tibet Plateau. *PLoS ONE* 6(4):e18708.