Supplementary Table 1 Primers and probes used in the analysis

Primers for methylation analysis in HBx

HBx GM8F1	TGGTTGTTNGGNTGTGTTGTTAAT	309bp
HBx GM8R1	[Biotin]ACATTANTANNAATCCAAAAATCCTCTTNT	
HBx GM8S1	NGGNTGTGTTGTTAATTG	
Sequence to Analyze1	GATTTTTYGYGGGAYGTTTTTTGTTTAYGTTTYGTYGGYGTTGAA TTTYGYGGAYGATTTTTTTYGGGGTYGTTTGGGGTTTTAT	
Sequence to Analyze2	GATTTTGYGYGGGAYGTTTTTTGTTTAYGTTTYGTYGGYGTTGAA TTTYGYGGAYGATTTTTTTYGGGGTYGTTTGGGGTTTTAT	

type-C AB033550) ″GATTTTTYGYGGGAYG″ and AB205119 (Genotype-B AB205119) ″GATTTTGYGYGGGAYG″

Probes for detection of HBV using FISH technique in PLC/PRF/5 cell line

FISH probe 1	GATTCCTAGGACCCCTGCTC ATGAGGCATAGCAGCAGGAT	429
FISH probe 2	TTCCTCTTCATCCTGCTGCT AAGCCCTACGAACCACTGAA	312
FISH probe 3	ATGTTTCCCTCTTGTTGCTG ATAGAGGTAAAAAGGGACTCAAGATG	264
FISH probe 4	TATTGGGGGCCAAGTCTGTA CAAGGCAGGATAGCCACATT	302
FISH probe 5	GCTTTCACTTTCTCGCCAAC CTAGCAGCCATGGAAAGGAG	298
FISH probe 6	CGTGGAACCTTTGTGTCTCC GAAGAAGGGGACGGTAGAGG	262
FISH probe 7	TTGCCCAAGGTCTTACACAAG ATACGGGTCAATGTCCATGC	281
FISH probe 8	AGCACCATGCAACTTTTTCC GCCTGAGTGCTGTATGGTGA	536
FISH probe 9	TGGGTGGGAAGTAATTTGGA GAGGCGAGGGAGTTCTTCTT	321
FISH probe 10	CCGCTTACAGACCACCAAAT CAGAGGGCCCACATATTGTT	325
FISH probe 11	CCCTCCTTTCCTCACATTCA ACCAACCTCCCATGCTGTAG	326
FISH probe 12	CTACAGCATGGGAGGTTGGT CTGCTGGCAGTGTTGTCAAT	278

Primers for detection of integration sites using Alu PCR technique (Originated from Murakami et al)

UP5 Alu-sense	CAGUGCCAAGUGUUUGCUGACGCCAAAGUGCUGGGAUUA
T3-515 Alu-antisense	AUUAACCCUCACUAAAGCCUCGAUAGAUYRYRCCAYUGCAC
UP6 Alu-sense (tag)	CAAGTGTTTGCTGACGCCAAAG
midT3 Alu-antisense (tag)	ATTAACCCTCACTAAAGCCTCG
pUTP HB1(HBVx)	ACAUGAACCUUUACCCCGUUGC
MM37 HB2(HBVx)	TGCCAAGTGTTTGCTGACGC
MM60 HB3(HBVx)	CTGCCGATCCATACTGCGGAAC

Primers for methylation analysis of integration sites in PLC/PRF/5

PLC/PRF/5 Chr3a-312-F1	AGTTTTTTTATGTAAGATTTTGGGTAGG	312
PLC/PRF/5 Chr3a-312-R1	[Biotin]CTTTTATAACAAATCAAACCACCTTTTA	
PLC/PRF/5 Chr3a-312-S1	GTAAGATTTTGGGTAGGA	
PLC/PRF/5 Chr3b-266-F2	AGAGTTTGGAAAGATAAATTTTGATT	266
PLC/PRF/5 Chr3b-266-R2	ACCCCCACTAACTAAAACTTAACTATA	
PLC/PRF/5 Chr3b-266-S2	TGGAAAGATAAATTTTGATTTTATT	
PLC/PRF/5 Chr4-342-F1	GGGAGTAATGGAATGATAGAATAGTTATG	342
PLC/PRF/5 Chr4-342-R1	ACTCTATATTTCCCTCATATTACTATACAA	
PLC/PRF/5 Chr4-342-S1	ATTAATTGTTTTATGAAGGGT	
PLC/PRF/5 Chr5-309-F1	TGGGGTTTGGTTATGGGTTATTA	309
PLC/PRF/5 Chr5-309-R1	AAATTCAAAACAATATCCACACT	
PLC/PRF/5 Chr5-309-S1	GTTTGGTTATGGGTTATTAG	
PLC/PRF/5 Chr8-332-F2	GGATTAAAGATAGGGATAGTAGAGGA	332
PLC/PRF/5 Chr8-332-R2	ACAAATATTTACTAAACCATACCCAATTAT	
PLC/PRF/5 Chr8-332-S2	ATGAGTTTAAGGAATATTAATATTG	
PLC/PRF/5 Chr10-221-F2	TGATGATTGTATTYGGGTTTATGGATTA	221
PLC/PRF/5 Chr10-221-R2	ATTCAAACAATCCAAACTAAAACTTCA	
PLC/PRF/5 Chr10-221-S2	ATTTTATTTTATTTTATTAGATGAT	
PLC/PRF/5 Chr11a-301-F2	GGGAGGGGTTGAATTTGGT	301
PLC/PRF/5 Chr11a-301-R2	AACCCACAAATTATTAACACCTAT	
PLC/PRF/5 Chr11a-301-S2	GATGTATTGAGTTGTAGTTTAGTTA	
PLC/PRF/5 Chr11b-262-F2	GGTTGGAAGGYGTGGTTTTAGT	262
PLC/PRF/5 Chr11b-262-R2	ACTAACACCCRCCATAAACACTTACAT	
PLC/PRF/5 Chr11b-262-S2	GGAAGGYGTGGTTTTAGTT	
PLC/PRF/5 Chr12a-294-F1	GAGTTGGGTAAATTGAGTTATAGAAGT	294
PLC/PRF/5 Chr12a-294-R1	CCTATATTCCCATCCCATCATCT	
PLC/PRF/5 Chr12a-294-S1	AATTGAGTTATAGAAGTTGG	
PLC/PRF/5 Chr12b-340-F1	GTTTTGATTGGTTGGTTTGAAGT	340
PLC/PRF/5 Chr12b-340-R1	ACTAACACCCCCCATAAACACTTACAT	
PLC/PRF/5 Chr12b-340-S1	ATTGGTTGGTTTGAAGTA	
PLC/PRF/5 Chr13-245-F1	GGTTTTTTYGTTTGTTTTTGTTTATTTG	245
PLC/PRF/5 Chr13-245-R1	ACCTAACAACCTCCATATATAATTTATCAT	
PLC/PRF/5 Chr13-245-S1	GTGTATATTTTGTTTTATTTTTGTA	
PLC/PRF/5 Chr16a-241-F1	GAATTTTAGAGGGTTGGGGATAGATA	241
PLC/PRF/5 Chr16a-241-R1	CTAAACAAAAAAATTAAACCTTATTATCCT	
PLC/PRF/5 Chr16a-241-S1	AGTTTTTGTTTTTAAGAATATGG	
PLC/PRF/5 Chr16b-391-F1	TGTGTGGTATTTTGTATGGTAGTT	391
PLC/PRF/5 Chr16b-391-R1	AAAAAATTCTTCTTCTAAAAAACCTAC	
PLC/PRF/5 Chr16b-391-S1	TTTGTATGGTAGTTTTAGTAAA	
PLC/PRF/5 Chr17a1-305-F1	GATGGTATTGTAGGATGTTGAGAGTAGT	305
PLC/PRF/5 Chr17a1-305-R1	CCCACCCAAATAACTAAAATCATTAATTCC	
PLC/PRF/5 Chr17a1-305-S1	TTTTAGAGTTTTATTTTAGGTGGA	
PLC/PRF/5 Chr17b-266-F1	GGATTGAGTGTGATAATAGGAGAGTA	266
PLC/PRF/5 Chr17b-266-R1	ACCAACCTCCTATCCTCCAATTTA	
PLC/PRF/5 Chr17b-266-S1	GGGGAGGGAGGTTTA	

Supplementary Table 2 Correlation analysis between the methylation pattern of the integrated HBV DNA and that of the human genome

	Human (200bp)																**HBV (200bp)**															
		repetitive sequence																**repetitive sequence**														
Cell lines	GC content (%)	SINEs (%)	LINEs (%)	LTR (%)	DNA elements (%)	Unclassified (%)	Small RNA (%)	Satellites (%)	Simple repeats (%)	Low complexity (%)	Average Methylation (%)	A nucleotide (%)	C nucleotide (%)	G nucleotide (%)	T nucleotide (%)	N nucleotide (%)	GC content (%)	SINEs (%)	LINEs (%)	LTR (%)	DNA elements (%)	Unclassified (%)	Small RNA (%)	Satellites (%)	Simple repeats (%)	Low complexity (%)	Average Methylation (%)	A nucleotide (%)	C nucleotide (%)	G nucleotide (%)	T nucleotide (%)	N nucleotide (%)
PLC/PRF/5 Chromosome 3	44.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	N/A	23.5	18.0	26.5	30.0	2.0	65.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	87.6	25.5	25.0	40.0	9.5	0.0
PLC/PRF/5 Chromosome 3	31.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.0	28.0	16.5	15.0	40.5	0.0	52.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.5	30.0	27.0	25.5	17.5	0.0
PLC/PRF/5 Chromosome 5	53.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	35.0	22.5	25.0	28.5	24.0	0.0	54.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	9.6	18.0	28.5	26.0	27.5	0.0
PLC/PRF/5 Chromosome 8	23.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	N/A	34.5	11.5	11.5	40.0	2.5	42.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.0	30.5	18.5	24.0	27.0	0.0
PLC/PRF/5 Chromosome 10	39.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	22.5	0.0	N/A	22.0	24.0	15.0	39.0	0.0	56.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.5	19.0	25.5	31.0	24.5	0.0
PLC/PRF/5 Chromosome 11	64.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.1	13.0	40.2	24.5	22.0	0.5	35.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	N/A	32.0	20.0	15.5	32.5	0.0
PLC/PRF/5 Chromosome 11	61.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	15.0	23.5	28.5	33.0	15.0	0.0	37.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	15.0	30.0	21.0	16.5	32.5	0.0
PLC/PRF/5 Chromosome 12	37.5	37.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.7	30.0	21.0	16.5	32.5	0.0	45.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.8	38.0	23.5	21.5	17.0	0.0
PLC/PRF/5 Chromosome 12	74.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	30.5	0.0	2.5	9.7	27.9	46.8	15.6	0.0	35.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	30.5	21.0	14.5	34.0	0.0
PLC/PRF/5 Chromosome 16	51.0	0.0	0.0	89.0	0.0	0.0	0.0	0.0	0.0	0.0	7.3	29.0	26.5	24.5	20.0	0.0	41.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	18.0	26.5	18.0	23.5	32.0	0.0
PLC/PRF/5 Chromosome 17	63.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	61.8	16.0	31.5	32.0	20.5	0.0	42.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	N/A	27.5	19.5	22.5	30.5	0.0
PLC/PRF/5 Chromosome 17	65.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	73.8	17.5	31.5	33.5	17.5	0.0	48.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	84.7	36.0	16.5	31.5	16.0	0.0
HepG2215 Chromosome 1	30.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	25.0	33.5	12.0	18.5	36.0	30.5	43.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	12.0	25.0	20.0	23.0	32.0	0.0
HepG2215 Chromosome 2	46.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	93.0	16.0	31.0	15.5	37.5	0.0	59.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	91.0	19.5	25.5	33.5	21.5	0.0
HepG2215 Chromosome 4	45.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	24.0	20.5	24.0	21.5	34.0	0.0	39.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	27.7	34.5	20.5	18.5	26.5	0.0
HepG2215 Chromosome 7	49.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	25.5	14.0	35.0	25.5	49.0	46.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	24.0	35.0	19.5	26.5	19.0	0.0
HepG2215 Chromosome 10	12.0	0.0	99.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	10.5	45.0	0.0	12.0	43.0	0.0	25.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	31.0	34.5	0.0	25.0	40.5	0.0
HepG2215 Chromosome X	45.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	24.0	22.5	22.5	31.0	0.0	54.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	21.0	29.5	24.5	25.0	0.0
HepG2215 Chromosome M	49.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	81.0	25.5	14.0	35.0	25.5	0.0	46.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	90.5	35.0	19.5	26.5	19.0	0.0
No.7Tumor Chromosome 6	27.0	0.0	63.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	44.0	36.5	11.5	15.5	36.5	0.0	49.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	25.5	21.5	28.0	25.0	0.0
No.7Tumor Chromosome 6	27.0	0.0	63.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.0	36.5	11.5	15.5	36.5	0.0	49.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	25.5	21.5	28.0	25.0	0.0
No.7ADJ Chromosome 3	35.5	0.0	0.0	0.0	49.5	0.0	0.0	0.0	0.0	0.0	3.0	40.0	19.0	16.5	24.5	0.0	49.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	42.5	25.0	27.0	22.0	26.0	0.0
No.8Tumor Chromosome 2	39.9	0.0	0.0	86.0	0.0	0.0	0.0	0.0	0.0	0.0	41.0	28.5	23.3	16.6	31.6	0.0	50.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	70.7	19.5	27.0	23.0	30.5	0.0
No.8Tumor Chromosome 18	41.5	41.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	12.5	32.5	20.0	21.5	26.0	0.0	49.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	23.0	25.5	20.9	29.1	24.0	2.0
No.8ADJ Chromosome 5	41.2	0.0	0.0	0.0	95.5	0.0	0.0	0.0	0.0	0.0	N/A	31.2	15.1	26.1	27.6	0.0	49.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	25.5	22.0	27.5	25.0	0.0
No.8ADJ Chromosome 7	30.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.0	30.8	15.4	15.4	38.5	0.0	43.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.0	33.0	24.5	18.5	24.0	0.0
No.8ADJ Chromosome 7	31.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	N/A	40.0	15.4	16.4	28.2	0.0	49.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	26.0	22.5	27.0	24.5	0.0
No.8ADJ Chromosome 7	32.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	26.0	47.5	15.0	17.0	20.5	0.0	40.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	N/A	38.0	11.0	29.0	22.0	0.0
No.8ADJ Chromosome 10	33.0	44.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	40.0	33.0	19.0	14.0	34.0	0.0	48.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.0	26.5	20.9	27.6	25.0	0.0
No.8ADJ Chromosome 14	28.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	42.0	39.5	11.0	17.0	32.5	0.0	46.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	N/A	26.5	20.5	26.0	27.0	0.0
No.8ADJ Chromosome 14	28.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.0	39.5	11.0	17.0	32.5	0.0	46.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.0	26.5	20.5	26.0	27.0	0.0
No.8ADJ Chromosome 15	55.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	14.0	15.5	31.5	24.0	29.0	0.0	56.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	9.3	26.0	25.5	31.0	15.5	2.0
No.9Tumor Chromosome 6	60.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	57.2	19.5	27.0	33.5	20.0	0.0	53.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	56.3	24.0	30.0	23.0	23.0	0.0
No.9Tumor Chromosome 7	41.5	0.0	91.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.0	23.5	19.5	22.0	35.0	0.0	48.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	31.0	20.5	27.5	21.0	0.0
No.9Tumor Chromosome 7	42.0	0.0	91.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	20.0	35.0	22.5	19.5	23.0	0.0	45.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	14.0	25.0	22.0	23.5	28.5	1.0
No.9Tumor Chromosome 9	36.8	0.0	0.0	0.0	0.0	0.0	0.0	99.2	0.0	0.0	25.0	29.6	12.8	24.0	32.8	0.8	49.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	42.7	33.0	18.5	30.5	18.0	0.0
No.9Tumor Chromosome 9 random	36.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	18.0	32.0	14.0	22.0	32.0	0.0	49.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	27.0	30.0	18.0	31.0	21.0	0.0
No.9Tumor Chromosome 9 random	36.5	0.0	0.0	0.0	0.0	0.0	0.0	97.5	0.0	0.0	28.0	28.5	14.5	22.0	35.0	0.0	51.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	N/A	32.0	20.5	31.0	16.5	0.0
No.9Tumor Chromosome 18	37.0	0.0	0.0	0.0	0.0	0.0	0.0	99.4	0.0	0.0	27.0	29.9	16.2	20.8	33.1	0.0	51.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	31.0	19.5	32.0	17.5	0.0
No.9ADJ Chromosome 2	40.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	93.0	29.5	20.5	19.5	30.5	0.0	46.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	9.0	18.0	23.3	23.3	35.3	0.0
No.9ADJ Chromosome 3	50.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	15.0	27.5	26.5	24.0	22.0	0.0	36.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	34.0	20.0	16.0	30.0	0.0
No.9ADJ Chromosome 15	54.0	94.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	28.5	26.0	28.0	17.5	0.0	50.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	24.2	28.3	22.2	25.3	0.0
No.10Tumor Chromosome 7	24.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	N/A	35.0	17.5	7.0	40.5	0.0	50.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	79.0	17.0	28.0	22.0	33.0	0.0
No.10Tumor Chromosome 12	47.5	0.0	0.0	99.5	0.0	0.0	0.0	0.0	0.0	0.0	7.0	30.5	18.5	29.0	22.0	0.0	53.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	17.0	26.5	20.5	33.0	19.0	1.0
No.10Tumor Chromosome 12	41.5	0.0	0.0	82.5	0.0	0.0	0.0	0.0	0.0	0.0	3.5	38.5	20.0	21.5	20.0	0.0	48.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.0	26.5	20.5	28.0	24.5	0.5
No.10Tumor Chromosome 12	40.4	0.0	0.0	94.3	0.0	0.0	0.0	0.0	0.0	0.0	1.0	23.4	25.5	14.9	36.2	0.0	56.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.0	19.0	33.5	23.0	24.5	0.0
No.10Tumor Chromosome 12	41.5	0.0	0.0	82.5	0.0	0.0	0.0	0.0	0.0	0.0	4.0	38.5	20.0	21.5	20.0	0.0	48.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.3	26.0	20.0	28.5	25.5	0.0
No.10Tumor Chromosome 19	46.9	96.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	89.0	21.1	23.8	23.1	32.0	0.0	48.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	98.7	17.0	26.5	21.5	34.5	0.5
No.10Tumor Chromosome 19	47.0	29.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	20.5	24.5	22.5	32.5	0.0	49.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	16.0	28.5	21.0	34.5	0.0
No.10Tumor Chromosome 19	50.5	97.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	89.5	20.0	25.0	25.5	29.5	0.0	48.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	96.5	17.0	27.0	21.5	34.5	0.0
No.10ADJ Chromosome 1	69.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	N/A	16.0	62.5	7.0	14.5	0.0	48.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.5	27.0	20.5	27.5	25.0	0.0
No.10ADJ Chromosome 17	29.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.0	35.5	14.5	15.0	35.0	0.0	47.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.0	27.0	21.0	26.0	26.0	0.0

Supplementary Table 3 Chromosome ideograms (PLC/PRF/5 and HepG2215)

cell lines	Band	Position	*Giemsa stain results	Level
PLC/PRF/5	3q22.1	chr3:129200001-133700000	gpos25	1
PLC/PRF/5	3q22.1	chr3:129200001-133700000	gpos25	1
PLC/PRF/5	4q34.3	chr4:177500001-183200000	gpos100	4
PLC/PRF/5	5p15.33	chr5:1-4500000	gneg	0
PLC/PRF/5	8p12	chr8:28800001-36500000	gpos75	3
PLC/PRF/5	10q11.21	chr10:42300001-46100000	gneg	0
PLC/PRF/5	11q13.1	chr11:63400001-65900000	gneg	0
PLC/PRF/5	11q13.1	chr11:63400001-65900000	gneg	0
PLC/PRF/5	12q24.11	chr12:109000001-111700000	gneg	0
PLC/PRF/5	12q24.11	chr12:109000001-111700000	gneg	0
PLC/PRF/5	16q22.1	chr16:66700001-70800000	gneg	0
PLC/PRF/5	17q25.3	chr17:75300001-81195210	gneg	0
PLC/PRF/5	17q25.3	chr17:75300001-81195210	gneg	0
HepG2215	1p31.1	chr1:69700001-84900000	gpos100	4
HepG2215	2p12	chr2:75000001-83300000	gpos100	4
HepG2215	4q34.3	chr4:177500001-183200000	gpos100	4
HepG2215	7q11.22	chr7:67000001-72200000	gpos50	2
HepG2215	10q21.1	chr10:52900001-61200000	gpos100	4
HepG2215	Xq23	chrX:108700001-116500000	gpos75	3
HepG2215	M9	chrM:9,084-9,653	N/A	

Giemsa stain negative (gneg)	0
Giemsa stain 25 (gpos25)	1
Giemsa stain 50 (gpos50)	2
Giemsa stain 75 (gpos75)	3
Giemsa stain 100 (gpos100)	4

---

* Furey TS and Haussler D. Integration of the cytogenetic map with the draft human genome sequence, Hum Molec Genet. 2003;12(9):1037-44.

* BAC Resource Consortium, Cheung VG, Nowak N, Jang W, Kirsch IR, Zhao S, Chen XN, Furey TS, Kim UJ, Kuo WL et al.. Integration of cytogenetic landmarks into the draft sequence of the human genome, Nature. 2001;409:953-98.

Supplementary Information

Figure legends for Supplementary Figures

Supplementary Fig 1 CGI mapping of the HBV genome. (*A*) Overview of the MCAM analysis. MCAM analysis was performed to detect methylated genes in the human PLC/PRF/5 cell line and in six paired specimens of primary HBV-HCC and adjacent tissues. (*B*) Gene-tree view of the MCAM analysis results. A set of 714 probes (514 unique genes) was selected, and hierarchical clustering was performed. Each row represents a single probe. Compared with the DNA methylation of CGIs in the healthy peripheral blood leukocytes of volunteers or the non-cancerous tissues, levels of DNA methylation were not remarkable in the PLC/PRF/5 cells and the cancerous tissues obtained from HBV-HCC patients.

Supplementary Fig 2 CGI mapping of the HBV genome. (*A*) Structure of the HBV genome. The CGI is shown as a dark box. (*B*) DNA sequence of the HBV genome genotype A. The CGI is shown.

Supplementary Fig 3 HBV genotypes. HBV genotypes were derived from Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Duran C, Field M, Heled J, Kearse M, Markowitz S, Moir R, Stones-Havas S, Sturrock S, Thierer T, Wilson A (2011) Geneious v5.5, which is available at http://www.geneious.com.

Supplementary Fig 4 FISH analysis of HBV integration. A FISH analysis was used to detect HBV DNA and demonstrate its presence in the genome of PLC/PRF/5 cells. (*A*) The positions of the original FISH probes 1–12 are shown. These probe designs were based on the reference HBV DNA sequence. FISH probes 6–7 correspond to the *HBx* gene. (*B*) When anti-DIG-HRP, biotin-tyramide, and Alexa488 (TSA-Bio) were used, the fluorescent signals with the best signal-to-noise ratio were detected. (*C*)-(*E*) The fluorescent signal was detected in the nuclei using probes that detect the *HBx* gene ((*C*) probes 5 and 6; (*D*) probe 5; (*E*) probe 6).

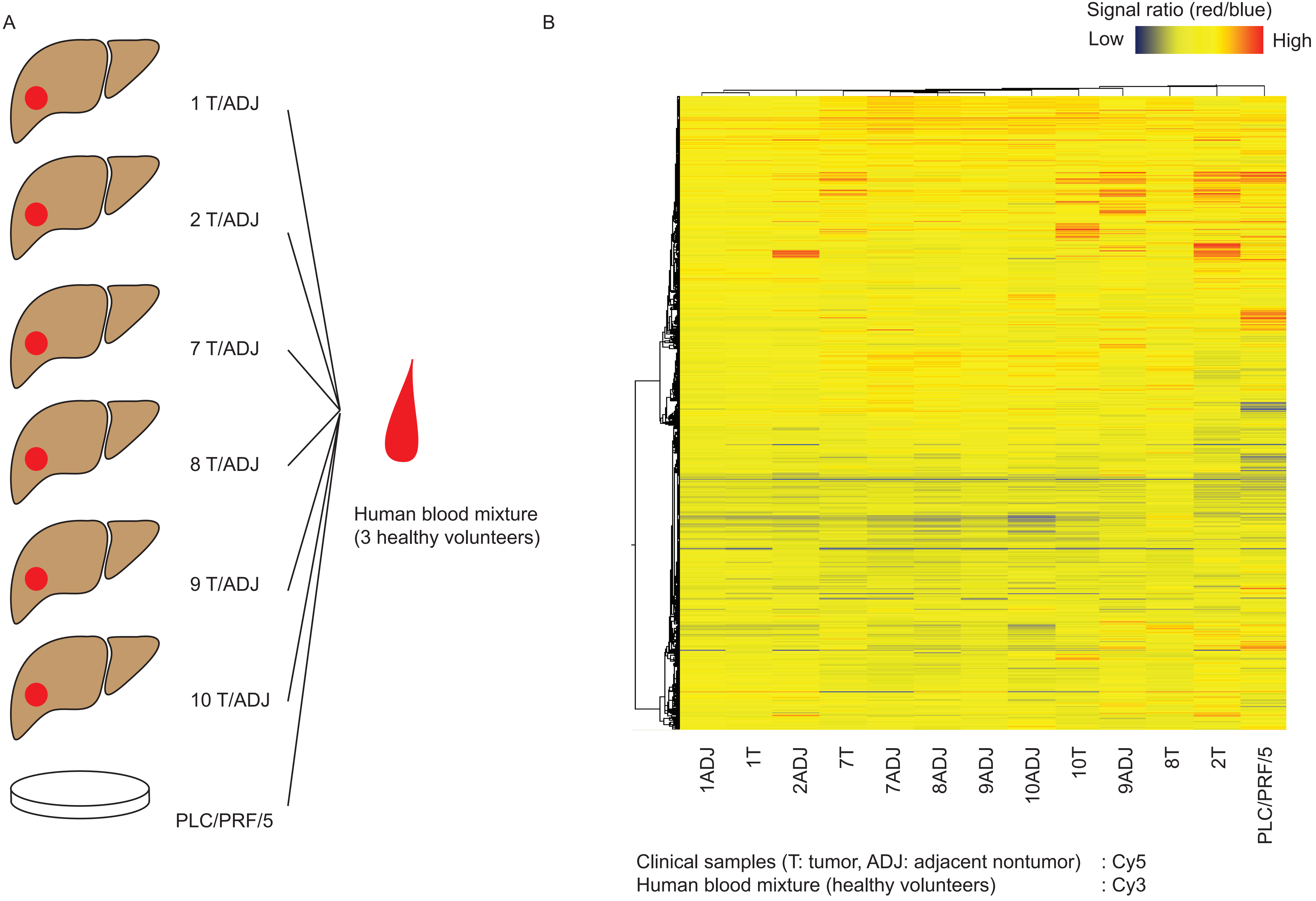
Supplementary Fig 5 Identification of the host genome integration sites for *HBx* using an Alu-PCR method. Chromosomal locations of viral-host junctions were analyzed using the Alu-PCR method. The 200-bp promoter region and 13-bp segment of the *HBx* gene body from the *HBx* primer position are shown. *HBx* integration was detected in host Chromosome 5.

Supplementary Fig 6 Analysis of the HBV DNA integration site sequences using NGS. (*A*) NGS-based structural methylation analysis using the G-NaVI method. (*B*) HBV specific genome capture. A total of 12,391 custom baits was synthesized based on the sequences of the HBV genotypes A to J and on those sequences present in the HBV-transformed PLC/PRF/5 cells that were not related to the sequence of the human genome. (*C*) Distribution of the numbers of library reads and read lengths from four paired clinical samples, PLC/PRF/5 cells, and HepG2.2.15 cells. Average read length was 333.14 bp with a modal length of approximately 500 bp.

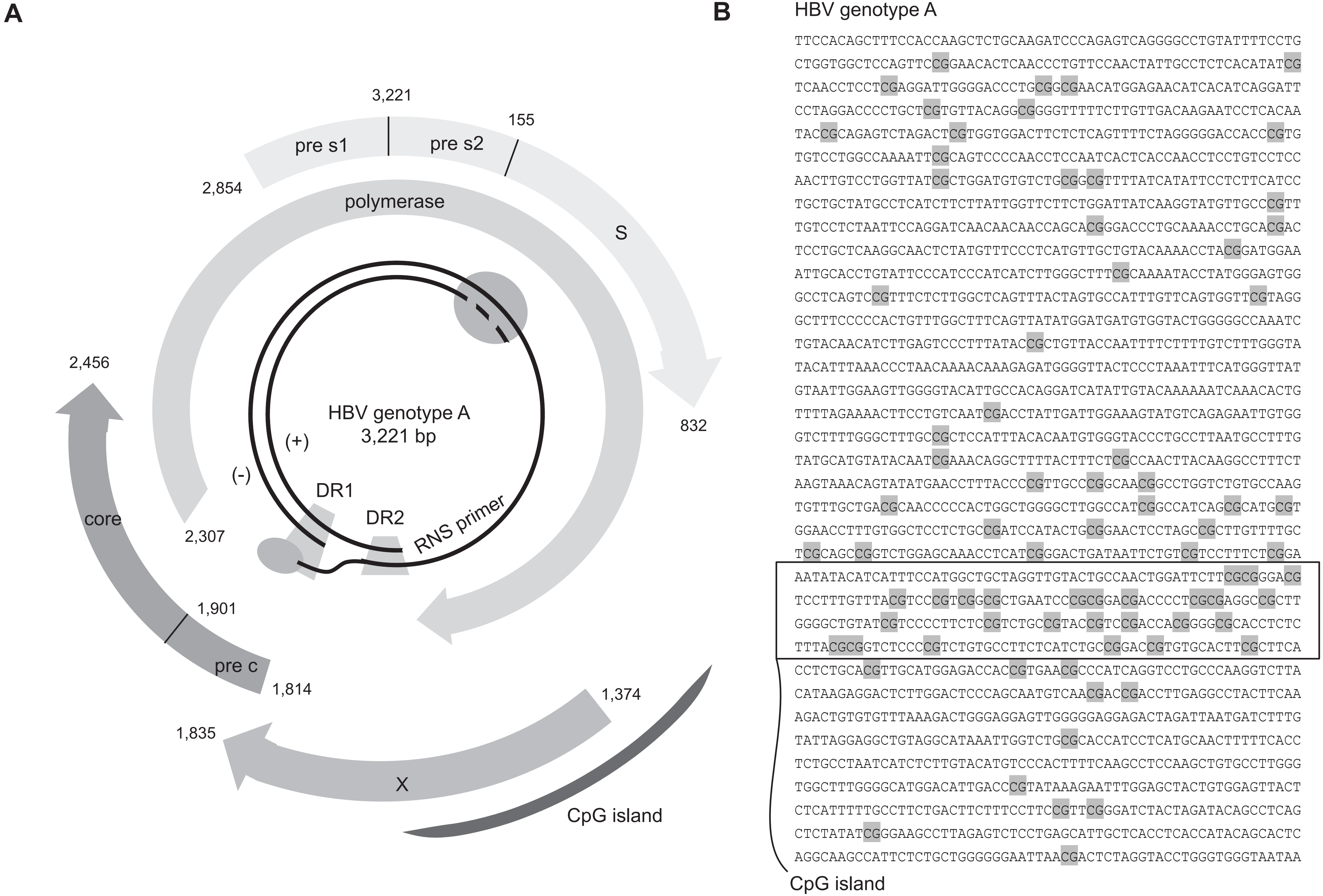
Supplementary Fig 7 The methylation levels of the HBV genome and the PLC/PRF/5 genome based on the chromosomal integration site. DNA methylation of the integrated HBV genome as well as the flanking human genome was analyzed by bisulfite pyrosequencing. Pyrograms of integration site at the Chromosomes 3 and 4 (*A*), Chromosomes 5, 8, and 10 (*B*), Chromosomes 11 and 12 (*C*), and Chromosomes 16 and 17 (*D*) are shown. Varying levels of methylation of the HBV sequences integrated into the genome of PLC/PRF/5 cells were detected.

Supplementary Fig 8 Distribution of the integration sites in the HBV genome and human chromosomes represented by Circos plots of the tumor (T) and adjacent non-tumor (ADJ) sample genomes. HBV DNA integration was analyzed using the G-NaVI method in four surgical specimen pairs of HCC and adjacent non-tumor tissues. Integration sites were rarely detected in exonic regions of the DNA from HBV-HCC samples. The HBV genes (PreC, Precore; C, Core; PreS, Presurface; S, Surface; X, X) and the 24 human chromosomes are shown.

Supplementary Figure 1



Supplementary Figure 2

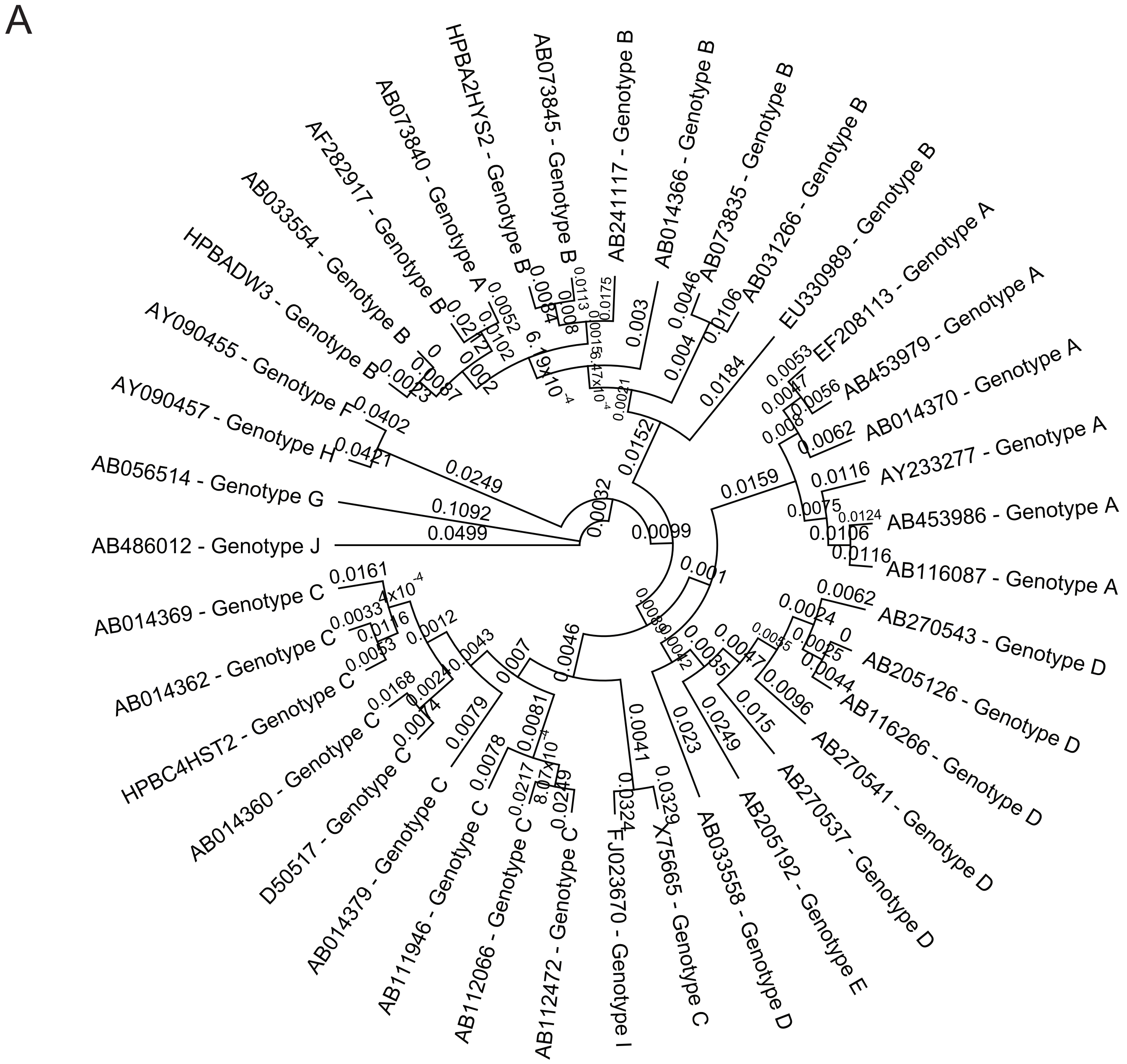


HBV genotype A

```
TTCCACAGCTTTCCACCAAGCTCTGCAAGATCCCAGAGTCAGGGGCCTGTATTTTCCTG
CTGGTGGCTCCAGTTCCGGAACACTCAACCCTGTTCCAACTATTGCCTCTCACATATCG
TCAACCTCCTCGAGGATTGGGGACCCTGCGGCGAACATGGAGAACATCACATCAGGATT
CCTAGGACCCCTGCTCGTGTTACAGGCGGGGTTTTTCTTGTTGACAAGAATCCTCACAA
TACCGCAGAGTCTAGACTCGTGGTGGACTTCTCTCAGTTTTCTAGGGGGACCACCCGTG
TGTCCTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAACCTCCTGTCCTCC
AACTTGTCCTGGTTATCGCTGGATGTGTCTGCGGCGTTTTATCATATTCCTCTTCATCC
TGCTGCTATGCCTCATCTTCTTATTGGTTCTTCTGGATTATCAAGGTATGTTGCCCGTT
TGTCCTCTAATTCCAGGATCAACAACAACCAGCACGGGACCCTGCAAAACCTGCACGAC
TCCTGCTCAAGGCAACTCTATGTTTCCCTCATGTTGCTGTACAAAACCTACGGATGGAA
ATTGCACCTGTATTCCCATCCCATCATCTTGGGCTTTCGCAAAATACCTATGGGAGTGG
GCCTCAGTCCGTTTCTCTTGGCTCAGTTTACTAGTGCCATTTGTTCAGTGGTTCGTAGG
GCTTTCCCCCACTGTTTGGCTTTCAGTTATATGGATGATGTGGTACTGGGGGCCAAATC
TGTACAACATCTTGAGTCCCTTTATACCGCTGTTACCAATTTTCTTTTGTCTTTGGGTA
TACATTTAAACCCTAACAAAACAAAGAGATGGGGTTACTCCCTAAATTTCATGGGTTAT
GTAATTGGAAGTTGGGGTACATTGCCACAGGATCATATTGTACAAAAAATCAAACACTG
TTTTAGAAAACTTCCTGTCAATCGACCTATTGATTGGAAAGTATGTCAGAGAATTGTGG
GTCTTTTGGGCTTTGCCGCTCCATTTACACAATGTGGGTACCCTGCCTTAATGCCTTTG
TATGCATGTATACAATCGAAACAGGCTTTTACTTTCTCGCCAACTTACAAGGCCTTTCT
AAGTAAACAGTATATGAACCTTTACCCCGTTGCCCGGCAACGGCCTGGTCTGTGCCAAG
TGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGCCATCGGCCATCAGCGCATGCGT
GGAACCTTTGTGGCTCCTCTGCCGATCCATACTGCGGAACTCCTAGCCGCTTGTTTTGC
TCGCAGCCGGTCTGGAGCAAACCTCATCGGGACTGATAATTCTGTCGTCCTTTCTCGGA
AATATACATCATTTCCATGGCTGCTAGGTTGTACTGCCAACTGGATTCTTCGCGGGACG
TCCTTTGTTTACGTCCCGTCGGCGCTGAATCCCGCGGACGACCCCTCGCGAGGCCGCTT
GGGGCTGTATCGTCCCCTTCTCCGTCTGCCGTACCGTCCGACCACGGGGCGCACCTCTC
TTTACGCGGTCTCCCCGTCTGTGCCTTCTCATCTGCCGGACCGTGTGCACTTCGCTTCA
CCTCTGCACGTTGCATGGAGACCACCGTGAACGCCCATCAGGTCCTGCCCAAGGTCTTA
CATAAGAGGACTCTTGGACTCCCAGCAATGTCAACGACCGACCTTGAGGCCTACTTCAA
AGACTGTGTGTTTAAAGACTGGGAGGAGTTGGGGGAGGAGACTAGATTAATGATCTTTG
TATTAGGAGGCTGTAGGCATAAATTGGTCTGCGCACCATCCTCATGCAACTTTTTCACC
TCTGCCTAATCATCTCTTGTACATGTCCCACTTTTCAAGCCTCCAAGCTGTGCCTTGGG
TGGCTTTGGGGCATGGACATTGACCCGTATAAAGAATTTGGAGCTACTGTGGAGTTACT
CTCATTTTTGCCTTCTGACTTCTTTCCTTCCGTTCGGGATCTACTAGATACAGCCTCAG
CTCTATATCGGGAAGCCTTAGAGTCTCCTGAGCATTGCTCACCTCACCATACAGCACTC
AGGCAAGCCATTCTCTGCTGGGGGGAATTAACGACTCTAGGTACCTGGGTGGGTAATAA
```

CpG island

Supplementary Figure 3



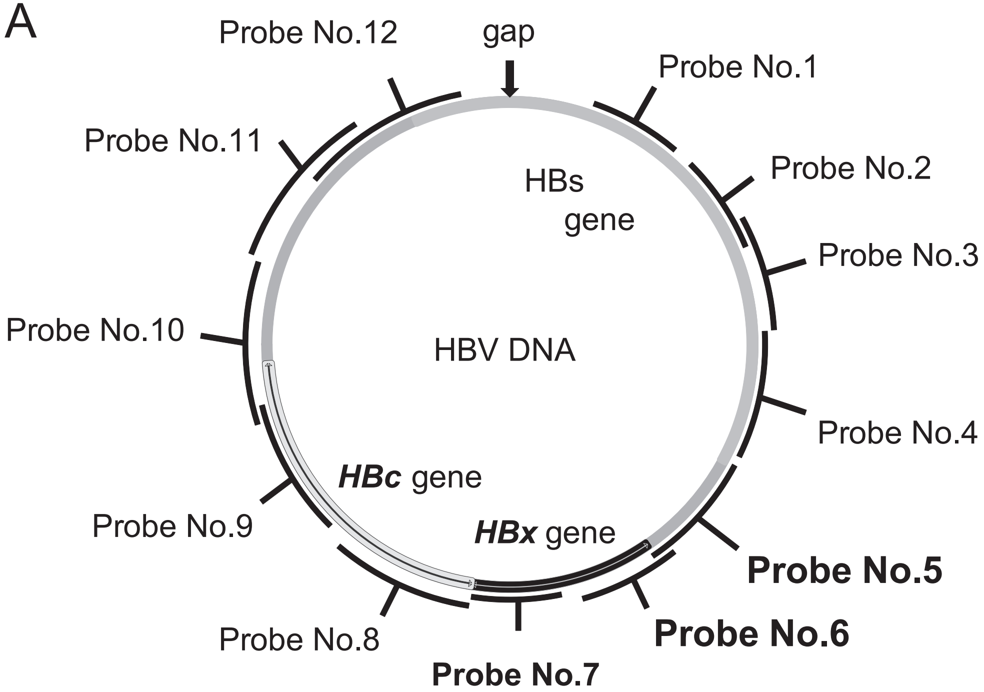
Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Duran C, Field M, Heled J, Kearse M, Markowitz S, Moir R Stones-Havas S, Sturrock S, Thierer T, Wilson A (2011) Geneious v5.5, Available from http://www.geneious.com

B

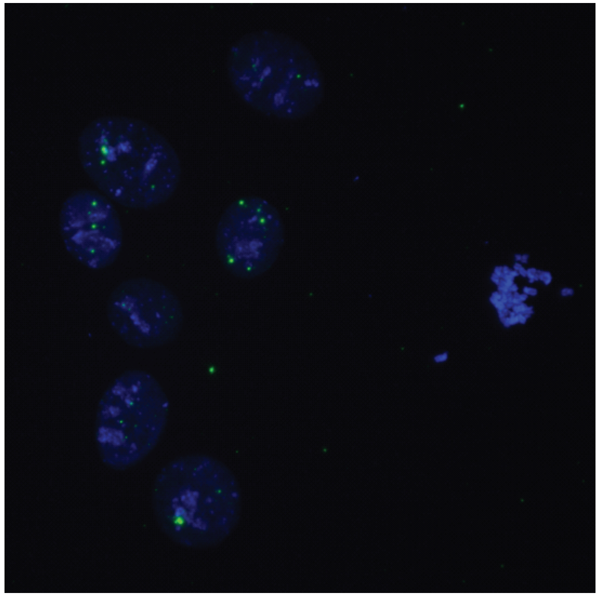
Accession No.	Subtype	Country
AB014370	Ae	Japan
EF208113	Ae	Germany
AB453979	Ae	Japan
AY233277	Aa	South Africa
AB116087	Aa	India
AB453986	Aa	Japan
AB014366	B1(Bj)	Japan
D23678	B1(Bj)	Japan
AB073845	B1(Bj)	Japan
AB073840	B2(Ba)	Taiwan
AF282917	B2(Ba)	China
D00331	B3	Indonesia
AB033554	B3	Indonesia
AB031266	B4	Vietnam
AB073835	B4	Vietnam
AB241117	B5	Philippines
EU330989	B6	China
AB111946	C1	Vietnam
AB112066	C1	Myanmar
AB112472	C1	Thailand
AB014360	C2	Japan
AB014379	C2	Japan
AB014362	C2	Japan
AB014369	C2	Japan
D50517	C2	Japan
D23681	C2	Japan
X75665	C3	Polynesia
AB270541	D1	Mongolia
AB270543	D1	Mongolia
AB205126	D2	Japan
AB116266	D2	Japan
AB270537	D3	Mongolia
AB033558	D4	Japan
AB205192	E	Ghana
AY090455	F	Nicaragua
AB056514	G	USA
AY090457	H	Nicaragua
FJ023670	I	Laos
AB486012	J	Japan:Ryuku

Supplementary Figure 4

A

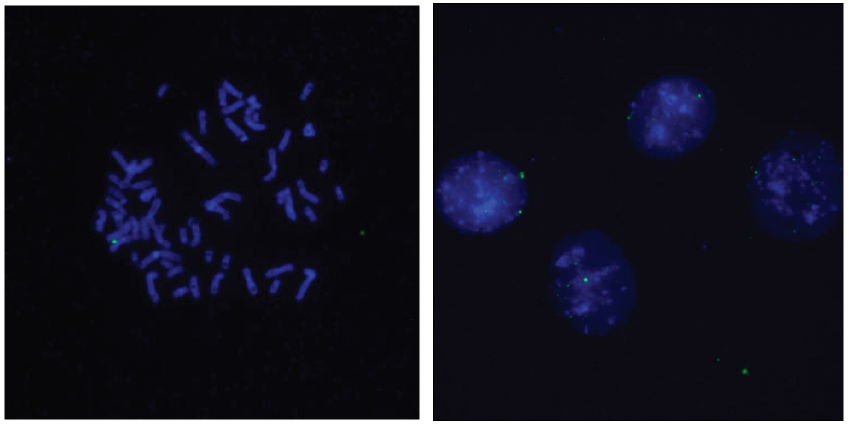


B



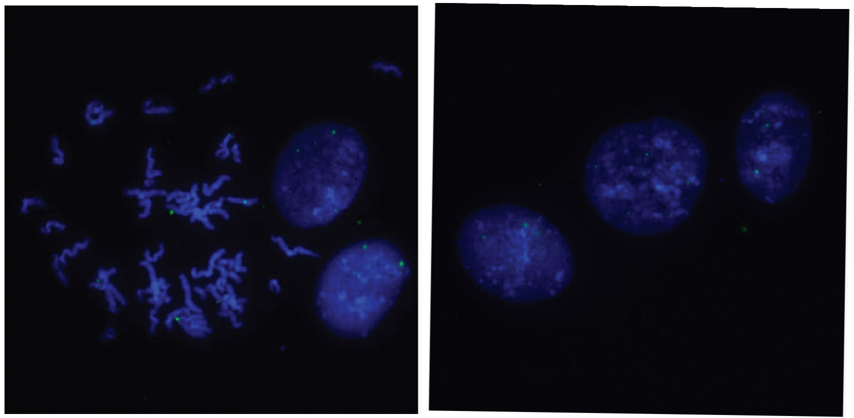
No.1 – No.12 dig probe
(anti dig-HRP, biotin-tyramide + Alexa 488-streptavidin)

C



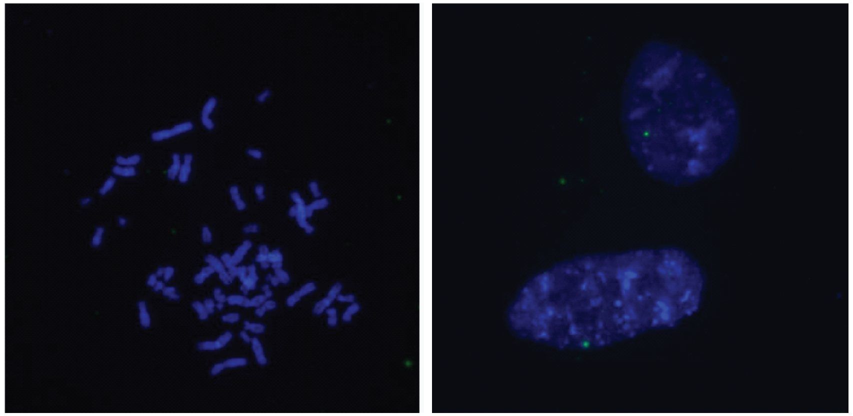
Metaphase Interphase
No.5 + 6 dig probe (anti dig-HRP, biotin-tyramide + Alexa 488-streptavidin)

D



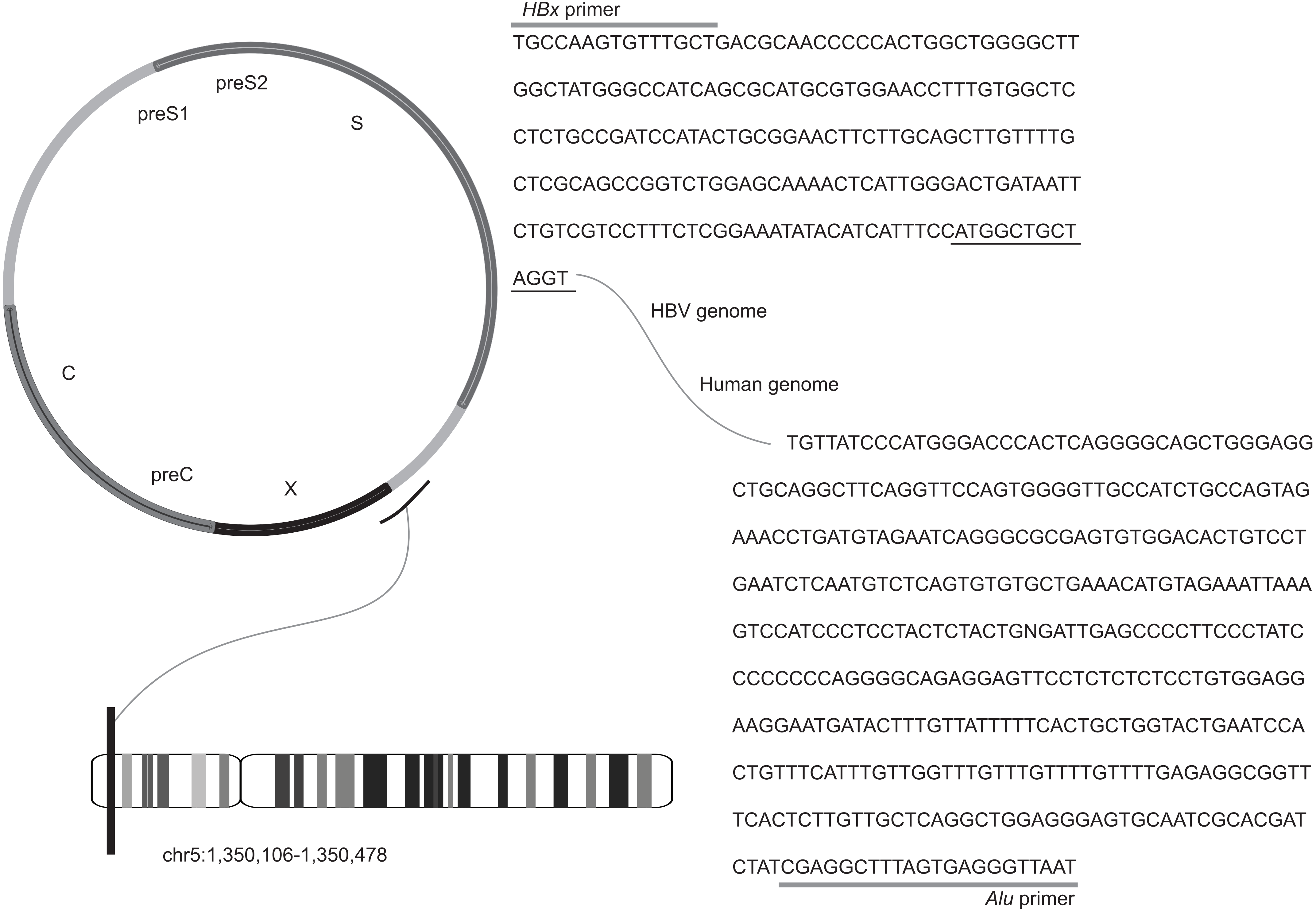
Metaphase Interphase
No.5 dig probe (anti dig-HRP, biotin-tyramide + Alexa 488-streptavidin)

E



Metaphase Interphase
No.6 dig probe (anti dig-HRP, biotin-tyramide + Alexa 488-streptavidin)

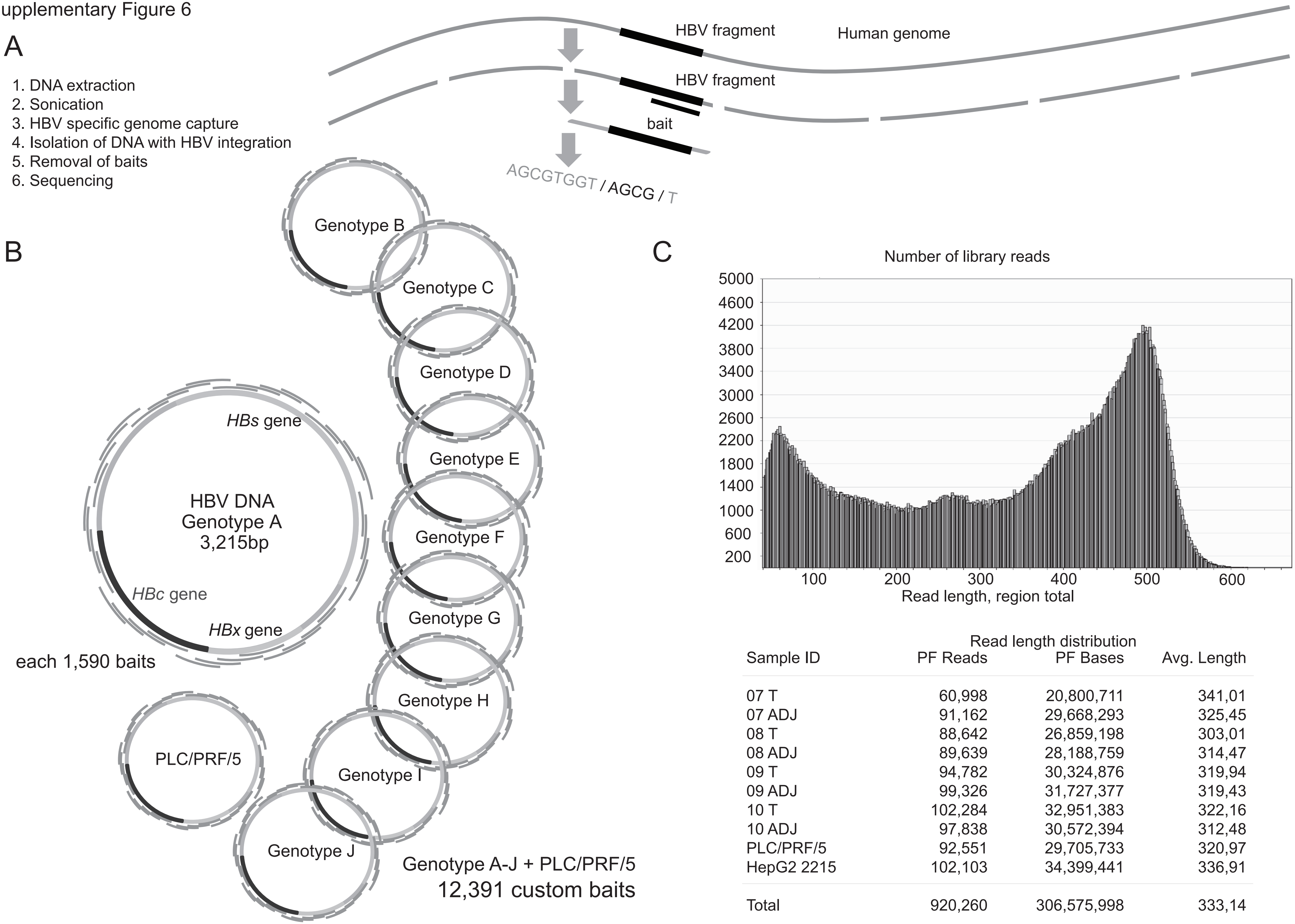
Supplementary Figure 5



TGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTT
GGCTATGGGCCATCAGCGCATGCGTGGAACCTTTGTGGCTC
CTCTGCCGATCCATACTGCGGAACTTCTTGCAGCTTGTTTTG
CTCGCAGCCGGTCTGGAGCAAAACTCATTGGGACTGATAATT
CTGTCGTCCTTTCTCGGAAATATACATCATTTCCATGGCTGCT
AGGT

TGTTATCCCATGGGACCCACTCAGGGGCAGCTGGGAGG
CTGCAGGCTTCAGGTTCCAGTGGGGTTGCCATCTGCCAGTAG
AAACCTGATGTAGAATCAGGGCGCGAGTGTGGACACTGTCCT
GAATCTCAATGTCTCAGTGTGTGCTGAAACATGTAGAAATTAAA
GTCCATCCCTCCTACTCTACTGNGATTGAGCCCCTTCCCTATC
CCCCCCCAGGGGCAGAGGAGTTCCTCTCTCTCCTGTGGAGG
AAGGAATGATACTTTGTTATTTTTCACTGCTGGTACTGAATCCA
CTGTTTCATTTGTTGGTTTGTTTGTTTTGTTTTGAGAGGCGGTT
TCACTCTTGTTGCTCAGGCTGGAGGGAGTGCAATCGCACGAT
CTATCGAGGCTTTAGTGAGGGTTAAT

Supplementary Figure 6



1. DNA extraction
2. Sonication
3. HBV specific genome capture
4. Isolation of DNA with HBV integration
5. Removal of baits
6. Sequencing

Read length distribution

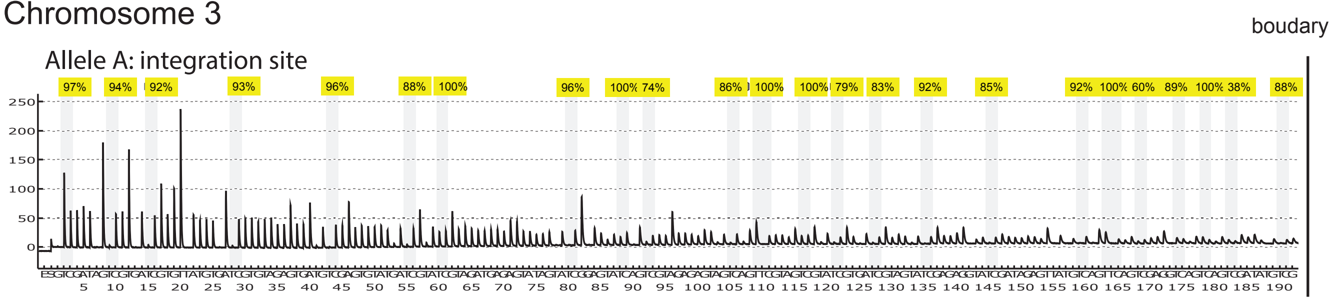
Sample ID	PF Reads	PF Bases	Avg. Length
07 T	60,998	20,800,711	341,01
07 ADJ	91,162	29,668,293	325,45
08 T	88,642	26,859,198	303,01
08 ADJ	89,639	28,188,759	314,47
09 T	94,782	30,324,876	319,94
09 ADJ	99,326	31,727,377	319,43
10 T	102,284	32,951,383	322,16
10 ADJ	97,838	30,572,394	312,48
PLC/PRF/5	92,551	29,705,733	320,97
HepG2 2215	102,103	34,399,441	336,91
Total	920,260	306,575,998	333,14

Supplementary Figure 7A PLC/PRF/5

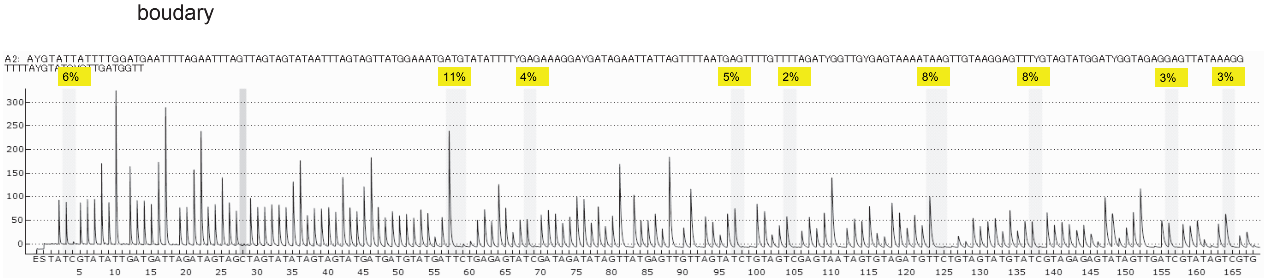
Chromosome 3

▭ HBV Genome ▨ Human Genome ▭ Primers — Sequencing 🟨 CG

Allele A: integration site



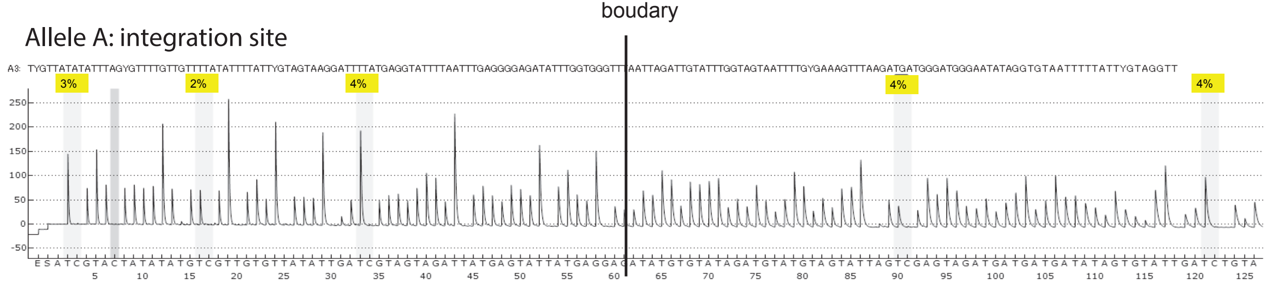
Allele A: integration site		Allele B: non-integration site	
Sense	AGTTTTTTTATGTAAGATTTTGGGTAGG	Sense	GGGAGGGAATATATTTAAGTAAATTAAAGT
Anti-sense	CTTTTATAACAAATCAAACCACCTTTTA	Anti-sense	ACCTTAACCCCTCCTTATAAACTC
Sequencing	GTAAGATTTTGGGTAGGA	Sequencing	ATATTTAAGTAAATTAAAGTGGAA



Allele A: integration site		Allele B: non-integration site	
Sense	AGAGTTTGGAAAGATAAATTTTGATT	Sense	GTTTAGAGTTTGGAAAGATAAATTTTGA
Anti-sense	ACCCCCACTAACTAAAACTTAACTATA	Anti-sense	ACATACCACAAAATTCCCTATTATTTCT
Sequencing	TGGAAAGATAAATTTTGATTTTATT	Sequencing	GTTTGGAAAGATAAATTTTGAT

Chromosome 4

Allele A: integration site



Allele A: integration site		Allele B: non-integration site	
Sence	GGGAGTAATGGAATGATAGAATAGTTATG	Sence	GGGAGTAATGGAATGATAGAATAGTTATG
Anti-sence	ACTCTATATTTCCCTCATATTACTATACAA	Anti-sence	ATCTCCCCTCAAATTAAAATACCTC
Sequencing	ATTAATTGTTTTATGAAGGGT	Sequencing	ATTAATTGTTTTATGAAGGGT

Allele A: integration site

```
GGAGGCTTGAACAGTGGGACATGTACAAGAGATGATTAGGCAGAGGTGAAAAAGTTGCATGATGATG
GTGCGCAGACCAATTTATGCCTACAGCCTCCTAATATTTAACCTAATCTCCTCCCCCAACTCCTCCC
AGTCTTTAAACACACAGTCTTTGAAGTAAGCCTCAAGGTCGGTCGTTGACATTGCTGGGAGTCCAAG
AGTCCTCTTATGTAAGACCTTGGGCAGGATCCGATGGGCGTTCACGGTGGTCTCCATGCAACGTGCA
GAGGTGAAGCGAAGTGCACACGGACCGGCAGATGAGAAGGCACAGACGGGGAGACCGCGTAAAGAGA
GGTGCGCCCCGTGGTCGGACGGTACGGCAGACGGAGAAGGGGACGATAGAGTCCCAAGCGGCCCCGC
GAGGGGTCGTCCGCGGGATTCAGCGCCGACGGGACGTAGACAAAGGACAAAGGCGTGCCACCAAGTC
TGTCTTTCCTCTCTTTTAAAAGGTGGTTTGATTTGCCATAAAAGACACACATTTGACATTTTTGAAG
AATACATTTTCTTCTCTATTTTTTTTGTAGCTCACAAGGAGGGGTCNAAGGTCNGGGGGTGGGGTGA
NCCCTCGGGGANTGTGTACATGTGAGGCCACCAAGGAGTGTGGAGAAATACANTTTTTTTGTTGGG
AGTTAATTAA
```

Allele A: integration site

```
TATCCCTCCCTAATCCTAGTATTTAGTATCCTATTGATGTATTTTTATGTGAACCAATCTTTATCAT
GATGGTTGCAAAATAATGAAGTTAATAATAAAGTTGGATTTTCCAACTCTTAACANCCCTACTATAT
TTATTAGTTGGTACTCAACATTCTACTGTAATCGACAGTCCTCCCTTTTCTCTTGTTTATTTATTAT
TATGGGTATGAACTCCGGATTCCTGTTTTCCAAGTTTAAAAATAGTAAGCTGACCTTTAAAAGTATT
TTTCTCTACTTGTTCTATTTTTAAAATCTACTTATTGTTGTGCTAGTACCACATGCACTTTGTTCTT
TCTCAAGACAGAGTCTATCGTTTAGAGTCTGGAAAGATAAATCTTGACCCTATTACGTATTATTCTG
GATGAATTTTAGAATCCAGTTAGCAGCACAACCTAGCAGCCATGGAAATGATGTATATTTCCGAGAA
AGGACGACAGAATTATCAGTCCCAATGAGTTTTGCTCCAGACCGGCTGCGAGCAAAACAAGCTGCAA
GGAGTTCCGCAGTATGGATCGGCAGAGGAGCCACAAAGGTTCCACGCATGCGCTGATGGCCCATAGC
CAAGCCCCAGCCAGTGGGGGTTGCGTCAGCAAACACTTGGCACAGGCCAGGCCGTTGCCGGGCAACG
GGGTAAAGGTTCATATACTGTTTACTTAGAAAGGCCTTGTAAGTTGGCGAGAAAGTAAAAGCCTGTT
TTGCTTGTGTACATGCATATAAAGGCATTAAGGCAGGGTAACCACATTGTGTAAAGGGCGCGGCAAA
GCCCAAAAGACCCACAATTCTTTGACATACTTTCCAATCAATAGGTCGATTGACAGGAAGTTTCCTA
AAACAG
```

Allele A: integration site

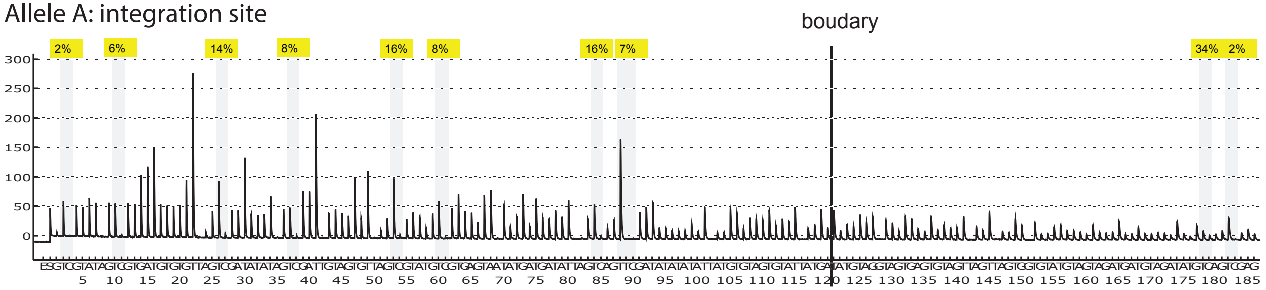
```
TAACATTTTCATTTAAATGAAGAAAAGAAACTTTGCAGCGATAAATATGTTTGGGAGCAATGGAATG
ACAGAATAGTCATGAACATCACTGCCTACACGATTCTTAGTTAATTTATTCTCACTCTTTCTATTAT
AGTCAATAACAAAAAAAATGTTTAAGCCTACCTATCAACTGCTTCATGAAGGGTTCGTTACATACTT
AGCGTTCTGCTGTTCCATATTTTATTCGTAGCAAGGATTCTATGAGGTACTCTAACCTGAGGGGAGA
CACCTGGTGGGTTTAATTAGATTGCATCTGGCAGTAATTTTGCGAAAGCCCAAGATGATGGGATGGG
AATACAGGTGCAATTTCCATCCGTAGGTTTTGTACAGCAACATGAGGGAAACATAGAGTTGCCTTGA
GCAGGAGTCGTGCAGGTTTTGCAGGGTCCCGTACTGGTTGTTGTGGATCCTGGAATTAGAGGACAAA
CGGGCAACATACCTTGATAGTCCAGAAGAACCAATAAGAAGATGAGGCATAGCAGCAGGATGAAGAG
GAATATGATAAAACGCCCGCAGACACATCCAGCGATAACCAGGACAAATTGGAGGACAGGAGGTTGGT
GAGTGATTGGAGGTTGGGGACTGCGAATTTTGGCCAAGACACACGGGTGATCCCCCTAGAAAATTGA
GAGAAGTCCACCACGAGTCTAGACTCTGTGGTATTGTGAGGATTCTTGTCAACAAGAAAAACCCCGC
CTGTAACACGAGCAGGGGTCCTAGGAATCCTGATGTAATGTTCTCCATGTTCGGTGCAGGGTCCCCA
GTCCTCGAGGAGATTGACGAGATGTGAGAGGCAATATTCTGAGCAGGGTGTACTGTTCCTGAACTGG
AGCCACCAGACTTCCTTATAATATTTATTATTAAAATATGATAGTATATGACTAT
```

Supplementary Figure 7B PLC/PRF/5

☐ HBV Genome ▨ Human Genome ☐ Primers — Sequencing 🟨 CG

Chromosome 5

Allele A: integration site



Allele A: integration site

Sense	TGGGGTTTGGTTATGGGTTATTA
Anti-sense	AAATTCAAAACAATATCCACACT
Sequencing	GTTTGGTTATGGGTTATTAG

Allele B: non-integration site

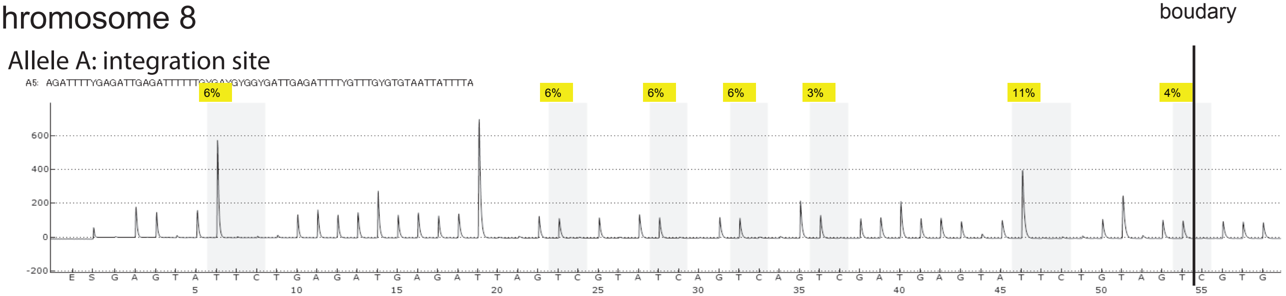
Sense	AGGGGTTTAATTTTAGTAGAGTAGG
Anti-sense	TCAAATCCCAATAAAATTACCATCTAC
Sequencing	TGAGATTTAGGATAGTGTTTA

Allele A: integration site

GCTGTTTTAGGAAACTTCCTGTCAATCGACCTATTGATTGGAAAGTATGTCAAAGAATTGTGGGTCT
TTTGGGCTTTGCCGCGCCCTTTACACAATGTGGTTACCCTGCCTTAATGCCTTTATATGCATGTATA
CAAGCAAAACAGGCTTTTACTTTCTCGCCAACTTACAAGGCCTTTCTAAGTAAACAGTATATGAACC
TTTACCCCGTTGCCCGGCAACGGCCTGGCCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTG
GGGCTTGGCTATGGGCCATCAGCGCATGCGTGGAACCTGTGTGGCTCCTCTGCCGATCCATACTGCG
GAACTCCTTGCAGCTTGTTTTGCTCGCAGCCGGTCTGGAGCAAAACTCATTGGGACTGATAATTCTG
TCGTCCTTTCTCGGAAATATACATCATTTCCATGGCTGCTAGGTTGTTATCCCATGGGACCCACTGC
AGGGGCAGCTGGGAGGCTGCAGGCTTCAGGTCCCAGTGGGGTTGCCATCTGCCAGTAGAAACCTGAT
GTAGAATCAGGGCGCGAGTGTGGACACTGTCCTGAATCTCAATGTCTCAGTGTGTGCTGAAACATGT
AGAAATTAAAGTCCATCCCTCCTACTCTACTGGGATTGAGCCCCTTCCCTACTCCCCCCCCACGGGC
AGAGGAGTTCCTCTCACTCCTGTGGAGGAAGGAATGATACTTTGNTTATTTTTCTACTGCTGGTACT
GAATCCACTGTTTCATTTGTTGGTTTGTTTGT

Chromosome 8

Allele A: integration site



Allele A: integration site

Sense	GGATTAAAGATAGGGATAGTAGAGGA
Anti-sense	ACAAATATTTACTAAACCATACCCAATTAT
Sequencing	ATGAGTTTAAGGAATATTAATATTG

Allele B: non-integration site

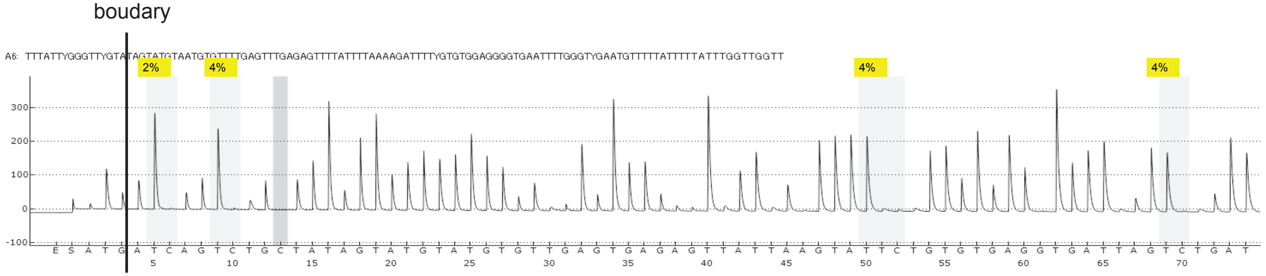
no CG site in human genome

Allele A: integration site

TTCTCTTTTCATTTACAGTGAGAGGGCCTACAAATTGCTGACACCTATTAATAATGTCCTCGTGTAA
ATGGATTTTAGGAAAGGAAGGCGTTTGCCATTCAGGATTAAAGATAGGGACAGTAGAGGAATAAAGC
CCAGTAAAATTTCCCACCTTATGAGTCCAAGGAATACTAACATTGAGATTCCCGAGATTGAGATCTT
CTGCGACGCGGCGATTGAGATCTTCGTCTGCGTGTAATTACTTTATGCATTTGATCTTTATGTTTAA
TTCTTTAACCTCCATAGCATTTGTTTTATTGTGATATAAGATGTAGATCTAATTTTTAAAAACTAAA
AAAAAAGTAATTACCCAAGTGTAATTAACTAATTCATTCCTTGTCATTCTTTAAAAAAGCCTAGCAA
AATAACTGGGTATGGTTTAGTAAATATTTGTTAAATAAATGAGTTTCATTAGACATCTCCAAATCTC
AGAAATTCCCTGTTGTTTTAATGAAGGTTTCAGTTTCCTTACCCCCTTTTCTAATTTCGCAAATTCT

Chromosome 10

Allele A: integration site



Allele A: integration site

Sense	TGATGATTGTATTYGGGTTTATGGATTA
Anti-sense	ATTCAAACAATCCAAACTAAAACTTCA
Sequencing	ATTTTATTTTATTTTATTAGATGAT

Allele B: non-integration site

no CG site in human genome

Allele A: integration site

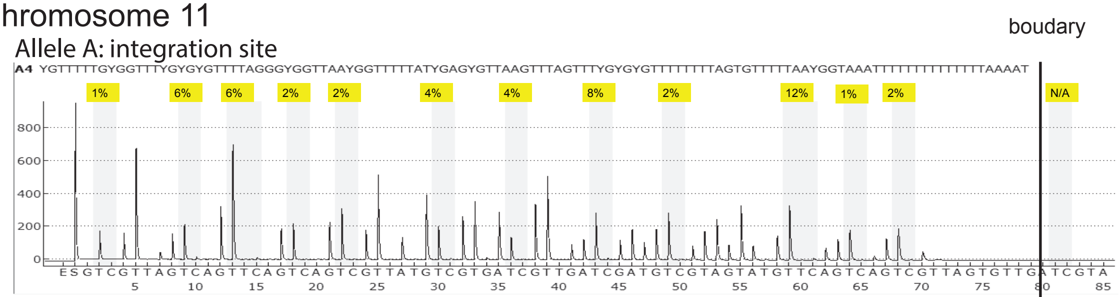
CCATTCAATTCCGTTCAATGtTTCCATTaGATTCCATcTGATGaTGATTCCATTCGATTCCATTTgA
TGATGATTCCATGCGATTCCATTAGATGATGACTCCTTTCATTTCCATTCGATGATGATTCCATTCG
TTTCCATCTGATGATGATTCCATTCGATTCCGTTCAATGATTATTCCATTCGAGTCCATTCGATGAT
TCCATTCGATTCTATTCGATGATGATTGCATTCGGGTCCATGGATT**ATTCCATTCCATTCCATTAGA**
TGATTCCATTCGGGTCCGCACAGTATGCAATGTGCCTTGAGCCTGAGAGCTCCACCCCAAAAGACCT
CCGTGTGGAGGGGTGAACCCTGGGCCGAATGCTCCCACTCCTACCTGGTTGGCTTGTGGCCAGTGGT
CCTTTGTGGGGTTGAAGTCCCAGTCTGGATTGTTTGAATTGGCTCCGAATGCTGGGTCCAACTGATG
ATCGGGAAAGAATCCCAGAGGGTTGGGGACAGACAGATTCGTCCCCATGCCTTTGCGAGGTTTTGGT
AACCAACCTCCCATGATGTAGCTCTTGTTCCCAAGAATATGGTGACCCGCAAAATGAGGCGCTACGT

Supplementary Figure 7C PLC/PRF/5

☐ HBV Genome ▨ Human Genome ☐ Primers — Sequencing 🟨 CG

Chromosome 11

Allele A: integration site

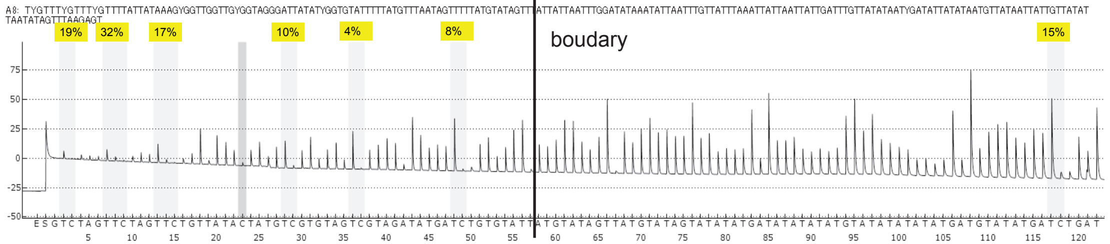


Allele A: integration site

Sense	GGGAGGGGTTGAATTTGGT
Anti-sense	AACCCACAAATTATTAACACCTAT
Sequencing	GATGTATTGAGTTGTAGTTTAGTTA

Allele B: non-integration site

Sense	GGGGYGGGGTTGAATTTGTT
Anti-sense	CCCCAAAACCAAATCCTAATTACT
Sequencing	GATGTATTGAGTTGTAGTTTAGTTA



Allele A: integration site

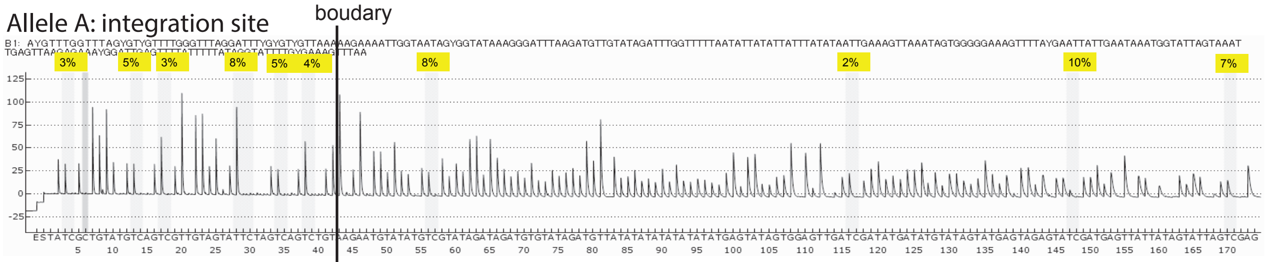
Sense	GGTTGGAAGGYGTGGTTTTAGT
Anti-sense	ACTAACACCCRCCATAAACACTTACAT
Sequencing	GGAAGGYGTGGTTTTAGTT

Allele B: non-integration site

N/A

Chromosome 12

Allele A: integration site

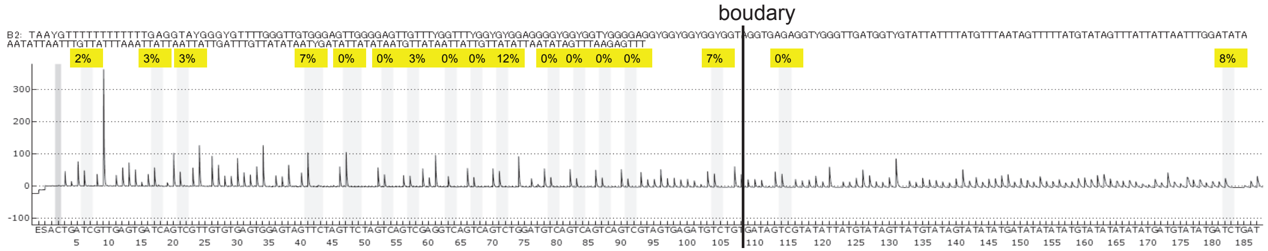


Allele A: integration site

Sense	GAGTTGGGTAAATTGAGTTATAGAAGT
Anti-sense	CCTATATTCCCATCCCATCATCT
Sequencing	AATTGAGTTATAGAAGTTGG

Allele B: non-integration site

Sense	TTTTAGAGTTGGGTAAATTGAGTTATAGAA
Anti-sense	AAATACAACRCCTCCTCCCCTTAAA
Sequencing	GGGTAAATTGAGTTATAGAAG



Allele A: integration site

Sense	GTTTTGATTGGTTGGTTTGAAGT
Anti-sense	ACTAACACCCCCCATAAACACTTACAT
Sequencing	ATTGGTTGGTTTGAAGTA

Allele B: non-integration site

Sense	TTTTGATTGGTTGGTTTGAAGTA
Anti-sense	CTCTCAAAACTAAATAAACTAAACCACAAA
Sequencing	ATTGGTTGGTTTGAAGTATA

Allele A: integration site

```
TCGTGGGACGGCAGTGCGTGCGGGAAGGCACGGCGAGGCCGGGCCGGGGGTGGGGAGCGGCGGGTCG
GGGCGGGGCTGAACCTGGCTGGGCCCGCCTTCCCGCCATGANTCCCCGGCCTTCGGGCCGCCCACTC
GAGCCACCTCTGTCCTCTCACCCAGTCCAGTCGAGTTGGCCTCCCGAGATGCACTGCGTTGTAGTCC
AGCTACGCCTCTGCGGCTCCGCGCGCTCTAGGGCGGCTAACGGCCCCCATCGAGCGCCAAGTTCAGC
TTCGCGCGCTCCCCTTAGTGTTTTTAACGGCAAACTCCTTCCTTTCCTAAAATTCATTTACATGAGG
ATGTTATTAATAGGTGTCAACAATTTGTGGGCCCTCTCACGGTAAATGAAAAACAAAGATTGAACTT
AATTATGCCTGCTAGGTTTTATCCTAACTCCACTAAATATTTGCCTCTAGACAAAGGAATTAAGCCT
TATTATCCTGAACATGTGGTTAATCATTACTTCCAATCTCGGCACTATTTACATAAGCTTTGGAAGG
CGGGTATTTTATATAAGAGAGAAACATCCCACAAAGCCTCATTCTTTGGGTCACCATATTCTTGGGA
ACAAGAGCTACATCATGGGGCTTTCTTGGACGGTCCCTCTCGAGTGGGGGAAGAATCAGTCCACCAG
CAATCCACTGGGATTTTTTCCCGACCACCAGTTGGATCCAGCATTCGGAGCCAACTCAAACAATCCA
```

Allele A: integration site

```
GAGCCGGTTTATTGGGTGGAAGGGACGAAGGAGAGGAACCCAAAGCCCGGGTACTTGTCTCAGAGGC
CCCAAAGCCAAATCCTGATTGCTCAGCGGGCTGGAAGGCGTGGCCTCAGCCCCGCCCCGCCCCGCCT
CACCACAAAGCGGTTGGTTGCGGCAGGGACTACACCGGTGTATCCCTATGCCTAACAGCCTCCATGT
ATAGTTTATCATTAACCTGGACATAAACATTAATCTGTTACCTAAACCATTAATTATTGATTTGTTA
CATAACCGACATTACACAATGTCATAATTATTGCTACATTAATACAGTTCAAGAGTCCACTTATGTA
AGTGTTCATGGCGGGTGCCAGTTCCATGCGTCGTGCAGAGGTGAAGCAAAGTGTACACGGACCGGCA
GATGAGCAGGAACAGACGGGGAAACCGCGTAAAGAGAGGTGCGCCCCGTGGTCGGACGGTACGGCAG
ACGAAGAAGGGGACGATAGAGTCCCAAGCGGCCCCGGGAGGGGTCGTCCGCAGGATTCAGCGCCGAC
```

Allele A: integration site

```
TGTACTATGAGCTAAGTACTATCGTGATCCCCATTTCTTAGAAAAACTGAGGCTCAGAGCGGGGAAT
CCCCTTGCTTGTCAAGCTCAAAATCACGACTGTCAGATTGAGGTCGCCGATCCCTCGGGGCAAGACT
CTCAGAGCTGGGTAAACTGAGCCACAGAAGCTGGACGCCTGGCTCAGCGTCGTCCTGGGTCTAGGAC
CTCGCGCCGCCAAAAAGAAAATTGGTAACAGCGGTATAAAGGGACTCAAGATGTTGTACAGATTTGG
CCCCCAATACCACATCATCCATATAACTGAAAGCCAAACAGTGGGGGAAAGCCCTACGAACCACTGA
ACAAATGGCACTAGTAAACTGAGCCAAGAGAAACGGACTGAGCCCCACTCCCATAGGTATTTTGCGA
AAGCCCAAGATGATGGGATGGGAATACAGGTGCAATTTCCATCCGTAGGTTTTGTACAGCAACATGA
GGGAAACATAGAGCCATGGAAATGATGTATATTTCCGAGAAAGGACGACAGAATTATCAGTCCCGAT
```

Allele A: integration site

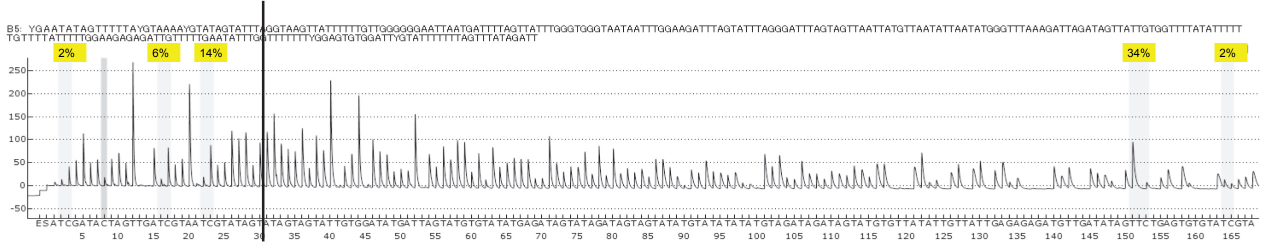
```
GGCTCGCGCGCTCTGATTGGCTGGCCTGAAGTACAACGCCTCCTCCCCTTGAGGCACGGGCGCTCTG
GGTTGTGGGAGTTGGGGAGCTGCTCCGGCTTCGGCGCGGAGGGGCGGCGGCCGGGGAGGCGGCGGCG
GCGGCAGGTGAGAGGCCGGGCTGATGGTCGCACCATTTTATGCCTAACAGCCTCCATGTATAGTTTA
TCATTAACCTGGACATAAACATTAATCTGTTACCTAAACCATTAATTATTGATTTGTTACATAACCG
ACATTACACAATGTCATAATTATTGCTACATTAATACAGTTCAAGAGTCCACTTATGTAAGTGTTCA
TGGCGGGTGCCAGTTCCATGCGTCGTGCAGAGGTGAAGCAAAGTGTACACGGACCGGCAGATGAGCA
GGAACAGACGGGGAAACCGCGTAAAGAGAGGTGCGNCCCGTGGTCGGACGGTACGGCAGACGAAGAA
```

Supplementary Figure 7D   PLC/PRF/5

☐ HBV Genome   ▨ Human Genome   ☐ Primers   — Sequencing   🟨 CG

Chromosome 16

Allele A: integration site



Allele A: integration site

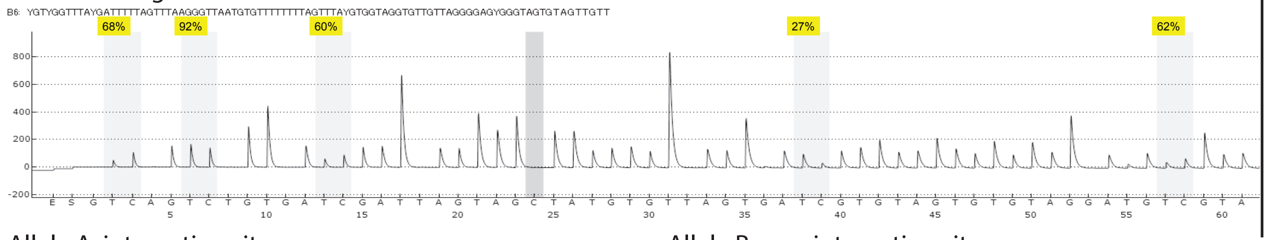
Sense	TGTGTGGTATTTTGTATGGTAGTT
Anti-sense	AAAAAATTCTTCTTCTAAAAAACCTAC
Sequencing	TTTGTATGGTAGTTTTAGTAAA

Allele B: non-integration site

N/A

Chromosome 17

Allele A: integration site



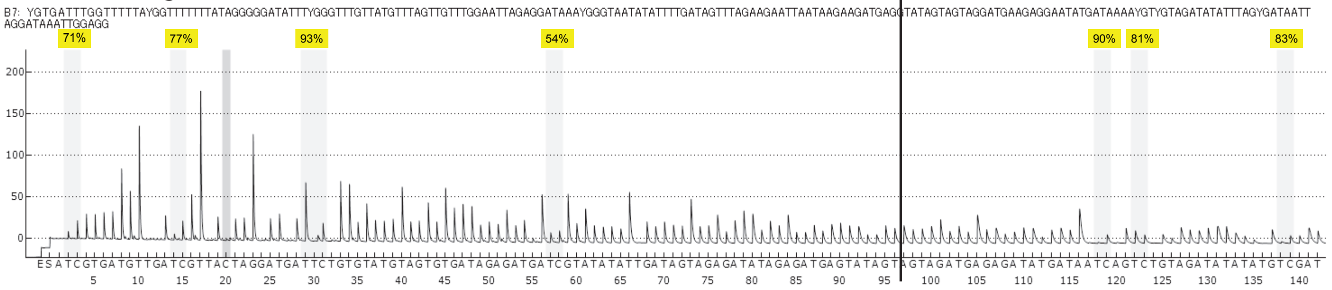
Allele A: integration site

Sense	GATGGTATTGTAGGATGTTGAGAGTAGT
Anti-sense	CCCACCCAAATAACTAAAATCATTAATTCC
Sequencing	TTTTAGAGTTTTATTTTAGGTGGA

Allele B: non-integration site

Sense	AGTTTAAGGGTTAATGTGTTTTTTTTAGT
Anti-sense	ACTATATAACTTATCACCCTCCATAAC
Sequencing	GTGTTGTTAGGGGAG

Allele A: integration site



Allele A: integration site

Sense	GGATTGAGTGTGATAATAGGAGAGTA
Anti-sense	ACCAACCTCCTATCCTCCAATTTA
Sequencing	GGGGAGGGAGGTTTA

Allele B: non-integration site

Sense	TGGGATTGAGTGTGATAATAGGA
Anti-sense	CCCAAAAAAAATACTAACATAAAACAA
Sequencing	GTATAGGGGGAGGGA

Allele A: integration site

GTATGAGTTATTATAATTAGTTTTAAATTAGTATGTTGAAGTTTTATTTTTATTATTTTCGTGTTAT
TGTATTTAGAGATAGGGTTTTTAAAGAGGTAATTAAAGTTAAATGAGGTTATTGGGGTGGGTTTTGA
TTTAGTTTGATTGGTTTCGTTATAAGAAGAGATATATAGAGGGAAGATTTTGTGAAGATATAGGGAG
AAGGCGGTTGTAGATAAGTTAAGGAGAGAGGTTTTTAGAGGGAATTAATTTTTCGGTATTTTGATTT
TGGTTTTTATGTTTTTAGAATTTTGAGAAAATATATTTGTTTTTTAAGTTATTTAG**TGTGTGGTAT**
TTTGTATGGTAGTTTTAGTAAACGAATATAGTTTTTACGTAAAACGTATAGTATTTAGGTAAGTTAT
TTTTTGTTGGGGGGAATTAATGATTTTAGTTATTTGGGTGGGTAATAATTTGGAAGATTTAGTATTT
AGGGATTTAGTAGTTAATTATGTTAATATTAATATGGGTTTAAAGATTAGATAGTTATTGTGGTTTT
ATATTTTTTGTTTTATTTTTGGAAGAGAGATTGTTTTTGAATATTTGGTTTTTTTCGGAGTGTGGAT
TCGTATTTTTTTTAGTTTATAGATTATTAAATGTTTTTATTTTATTAATATTTTCGGAAATTATTGTT
GTTAGACGACGAGATCGAG**GTAGGTTTTTTAGAAGAAGAATTTTTT**CGTTTCGTAGACGAAGATTTT
AATCGTCGCGTCGTAGAAGATTTTAATTTCGGGAATTTTAATGTTAGTATTTTTTGGATTTATAAGG

Allele A: integration site

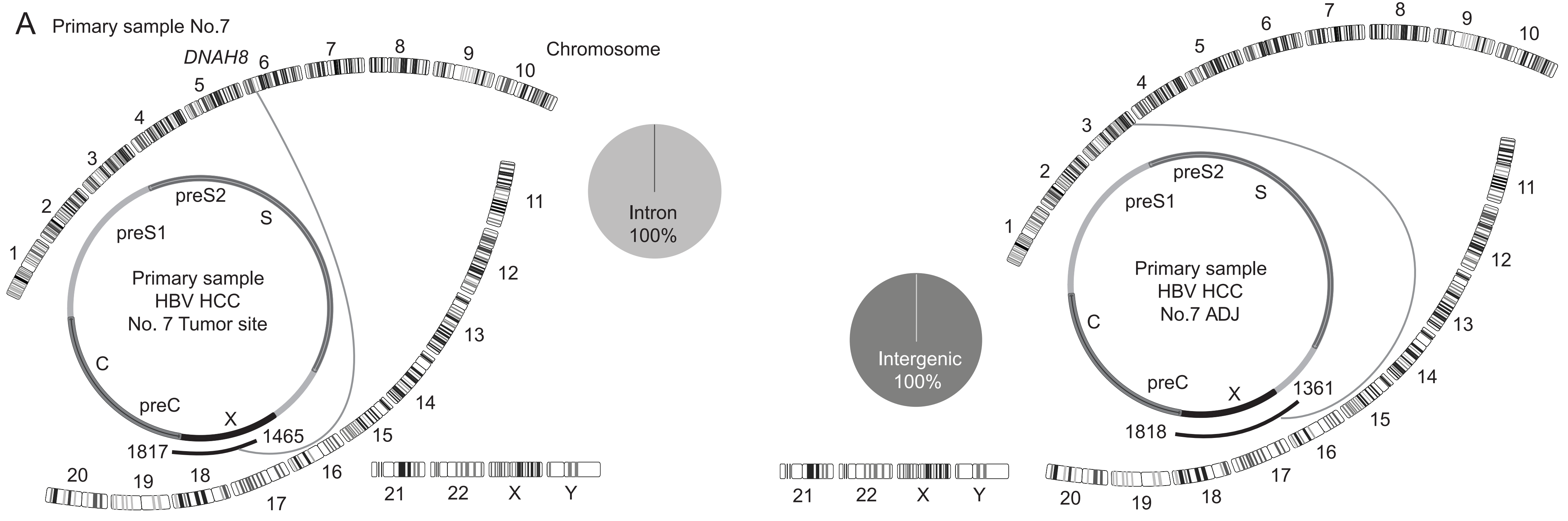
ATTAGAAATGTTTATGTATTTGTTGATTAGGATTTTTATTTTGAGGTGTGTGTAGAGTGATTTAAAG
TTTTTTYGTTYGTGTYGTGATTTTGGGTTTTGGAGTTTATTTTTTTTTTTTTAGAAAAGTTTGTTTTG
TAGGTATTTGTTTTGAGAGTATTYGATGGTATTGTAGGATGTTGAGAGTAGTTGAGGTAGGAGGGTY
GGYGGTTGTTTTTTTTTAGAGTTTTATTTTAGGTGGAYGTYGGTTTAYGATTTTTAGTTTAAGGGTT
AATGTGTTTTTTTTAGTTTAYGTGGTAGGTGTTGTTAGGGGAGYGGGTAGTGTAGTTGTTGGATAGT
AGGGGGTGTTTTTTTAATTTAGGTTGTTTTTTTTTTTTTGTTATATAGTTGTATGTTTTTTATTATA
TAGTATTTAGGTAAGTAATTTTTTGTTGGGGGGAATTAATGATTTTAGTTATTTGGGTGGGTAATAA
TTTGGAAGATTTAGTATTTAGGGATTTAGTAGTTAATTATGTGAATATTAATATGGGTTTAAAGATT
AGATAATTATTGTGGTTTTATATTTTTTGTTTTATTTTTGGAAGAGAAATAGTGTTTGAGTATTTGG
TTTTTTTTGGAGTGTGGATTYGTATTTTTTTTAGTATATAGATTATTAAATGTTTTTATTTTATTAAT
ATTTTYGGAAATTATTGTTGTTAGAYGAYGAGATYGAGGTAGGTTTTTTAGAAGAAGAATTTTTTYG
TTTYGTAGAYGAAGATTTTAATYGTYGYGTYGTAGAAGATTTTAATTTYGGGAATTTTAATGTTAGT
ATTTTTTGGATTTATAAGGTGGGAAATTTTAT

Allele A: integration site

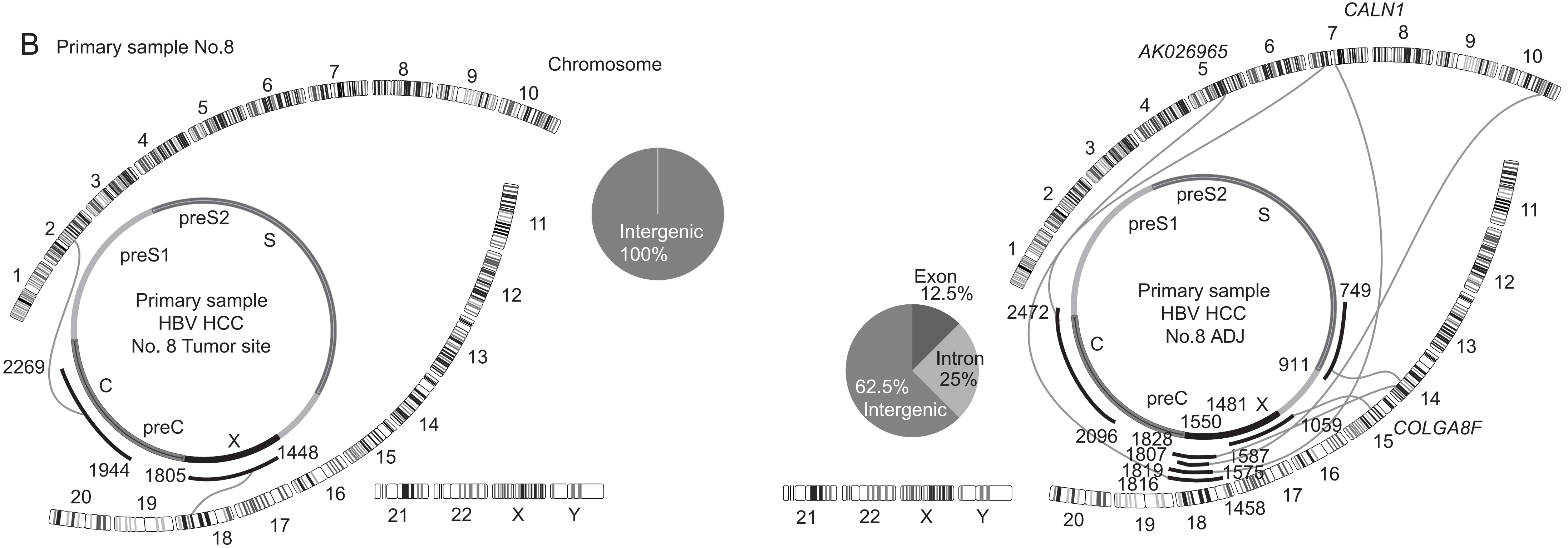
GCCACAGGGGTCCTCACGCACCCAAGATCAACCCTTCTCAAATGCATCTAGTTATTTCCAAGTCTCT
TGGGAACAAGGCTTTTCCCACCACTGTCAGCACGGCACCCAGGCCTTTAGTGCAGACGGTGACAGTG
AATGTGACTTGACTCCCCCCACCCCGAATCTCCAGAAGCTGCGTGCACCTTTCCAGGGTCCCCACGC
CTGGGAGGCTCACAGGCCACGCCCACCCGTGTCTCCCCAGCCTGCGTCAACCCTCACCTGCTAAAGT
CTCCTAATGGACACCCCCACCCCTGAGCTCCTGGAGCTCAGCCCTGGGGGGTCACCGTGGCCCCGGG
GCTGCCTTTGTGGGACTGAGTGTGACAACAGGAGAGCAATGGGGTACAGGGGGAGGGAGGCCCACGT
GATCTGGTTCCCACGGTCCCTCCACAGGGGGACACTCCGGGCCTGCTATGTCCAGCTGCCTGGAATT
AGAGGACAAACGGGCAACATACCTTGATAGTCCAGAAGAACCAATAAGAAGATGAGGCATAGCAGCA
GGATGAAGAGGAATATGATAAAACGCCGCAGACACATCCAGCGATAACCAGGACAAATTGGAGGACA
GGAGGTTGGTGAGTGATTGGAGGTTGGGGACTGCGAATTTTGGCCAAGACACACGGGTGATCCCCCT
AGAAAATTGAGAGAAGTCCACCACGAGTCTAGACTCTGTGGTATTGTGAGGATTCTTGTCAACAAGA
AAAACCCCGCCTGTAACACGAGCAGGGGTCCTAGGAATCCTGATGTAATGTTCTCCATGTTCGGTGC
AGGGTCCCCAGTCCTCGAGGAGATTGACGAGATGTGAGAGGCAATATTCTGAGCAGGGTTTACGTT

Supplementary Figure 8

A Primary sample No.7

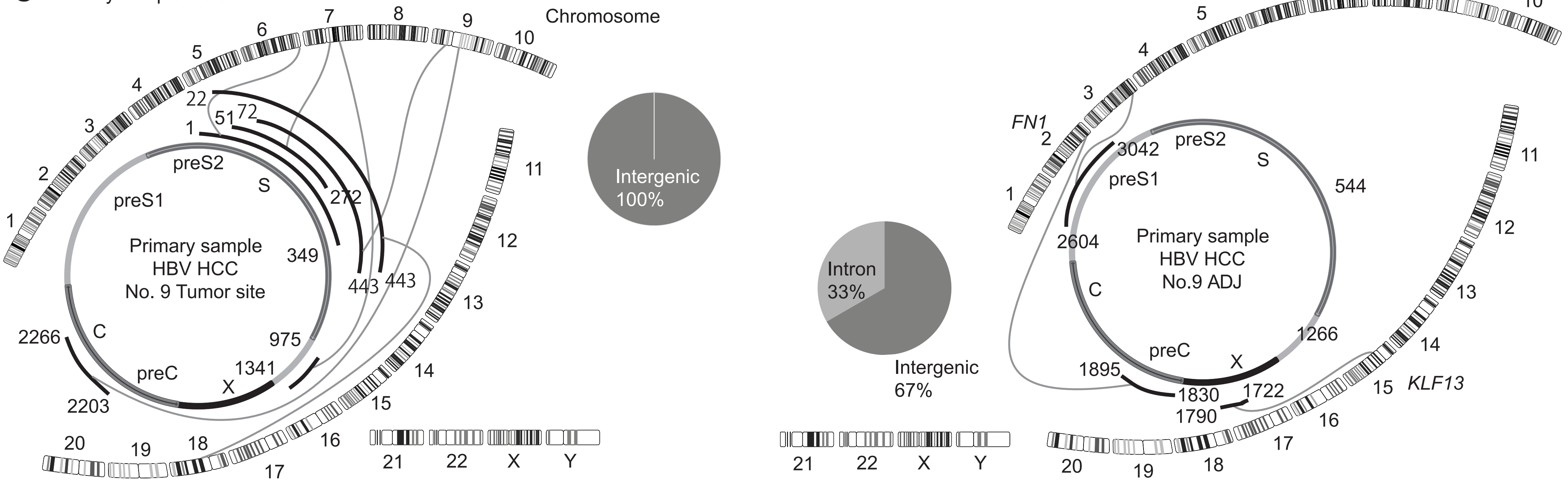


B Primary sample No.8



Supplementary Figure 8

C Primary sample No.9



D Primary sample No.10