

**Table S1. The demographic and clinicopathological characteristics of all donors**

Sample ID	Type	Gender	Age	HBV Subtype	HBV DNA level (IU/ml)	AFP	HBeAg (PEIU/ml)	Hepatic fibrosis	AJCC stage	Tumor size	Number of tumor nodule(s)	Recurrence	Antiviral therapy	TP53
647	HCC	male	53	C1	2.13×10E5	30.60	0.00	Y	I	5	single	Y	Y	N
654	HCC	male	60	C1	1.06×10E7	12.70	0.00	-	IIIC	4.5	single	Y	Y	N
678	HCC	male	54	C1	5.10×10E4	67.00	0.08	-	I	15	single	Y	N	N
687	HCC	male	26	B2	1.82×10E5	2374.00	0.00	Y	I	3.7	single	N	N	N
690	HCC	male	66	B2	2.53×10E4	221.77	0.00	-	IIIA	10	multiply	Y	Y	N
693	HCC	male	48	C1	0	13.60	0.00	Y	IIIA	5	multiply	Y	N	N
699	HCC	female	46	B2	1.38×10E8	3882.80	0.81	Y	II	5	multiply	Y	Y	N
700	HCC	male	56	B2	1.55×10E6	11.01	0.00	Y	I	6.5	single	N	Y	N
703	HCC	female	49	C1	1.26×10E7	1256.70	0.00	Y	I	2.5	single	N	Y	N
712	HCC	male	62	C1	1.65×10E6	1203.40	0.00	N	I	12.5	single	Y	Y	Y
715	HCC	male	48	C1	0	16	0.00	Y	I	3.2	single	N	N	Y
722	HCC	female	41	B2	0	56.30	0.00	-	IIIB	7.5	single	Y	N	Y
724	HCC	female	51	C1	1.11×10E6	535.40	0.07	Y	II	2.5	single	N	Y	N

						502.4								
725	HCC	male	48	C1	0	0	0.00	Y	II	5	single	Y	N	N
729	HCC	male	58	C1	2.33×10E4	10.00	0.00	Y	I	4	single	Y	Y	N
734	HCC	male	48	C1	1.60×10E5	4.70	0.00	-	I	8.5	single	N	Y	Y
						9441.								
745	HCC	male	36	C1	2.2×10E6	30	0.00	Y	IIIA	6	multiply	N	Y	Y
747	HCC	male	65	C1	1.32×10E4	5.10	43.21	Y	I	4	single	Y	Y	N
760	HCC	male	46	B2	0	82.50	0.00	N	IIIA	8	multiply	Y	Y	Y
764	HCC	male	61	C1	4.29×10E7	26.70	7.20	Y	I	3	single	Y	Y	N
765	HCC	male	69	B2	1.95×10E4	5.10	0.00	Y	I	5.2	single	Y	Y	N
766	HCC	male	65	C1	3.93×10E6	5.10	1.42	-	I	5	single	Y	Y	Y
						301.4								
768	HCC	male	64	C1	2.10*×10E7	0	0.00	Y	IIIA	5.2	multiply	Y	Y	N
775	HCC	male	58	B2	3.06*×10E6	8.20	0.00	-	I	6	single	Y	Y	N
777	HCC	male	55	C1	0	44.50	0.00	Y	I	11	single	N	N	N
						17401								
778	HCC	male	51	C1	1.16×10E4	.00	0.00	Y	IIIB	11	single	Y	Y	N
780	HCC	male	57	C1	0	15.20	0.28	Y	I	2.5	single	Y	Y	N
						119.7								
781	HCC	male	63	B4	1.79×10E5	0	0.00	Y	I	2.8	single	Y	Y	Y
						737.5								
783	HCC	male	56	B1	2.5×10E5	0	0.13	N	I	4.5	single	Y	N	Y
791	HCC	male	46	B2	1.02×10E7	11.40	0.00	Y	I	0.8	single	N	N	N
794	HCC	male	63	C1	5.89×10E4	72.10	0.00	N	I	6	single	Y	Y	N
						443.4								
795	HCC	male	59	C1	0	0	0.00	Y	I	3.5	single	N	N	N

801	HCC	fema le	55	B3	0	4.40 149.1	0.06	Y	II	2.8	single	Y	N	N
805	HCC	male fema	74	B2	2.59×10E4	0 46923	0.51	Y	II	4	single	Y	Y	Y
807	HCC	le	41	C1	9.31×10E5	.50 1532.	0.13	Y	I	0	single	Y	Y	N
809	HCC	male fema	44	B2	2.05×10E4	10 155.1	0.00	Y	I	14	single	Y	N	N
810	HCC	le fema	58	C1	0	0 2789.	0.00	Y	I	2	single	N	N	Y
812	HCC	le	50	B1	2.92×10E6	00	0.24	Y	I	2.5	single	Y	Y	N
813	HCC	male	55	C1	5.43×10E6	3.90	0.00	-	I	12	single	N	N	N
815	HCC	male	59	C1	0	12.60	0.00	Y	I	3.2	single	Y	Y	N
821	HCC	male	56	B2	2.06×10E5	5.90 12271	0.00	Y	II	13.5	single	Y	Y	N
822	HCC	male	67	B2	0	.50	0.00	N	I	6.4	single	Y	N	Y
828	HCC	male	54	C1	6.7×10E7	12.70 87500	43.58	N	II	5	single	N	N	N
829	HCC	male	42	C1	0	.00 50000	0.16	Y	II	7	single	N	Y	N
831	HCC	male	41	B2	1.37×10E6	.00 342.6	0.00	N	IIIB	12	single	Y	Y	N
846	HCC	male fema	44	B2	0	0 39447	0.00	Y	I	7.3	single	N	N	N
850	HCC	le	43	C1	3.19×10E7	.60	68.53	Y	II	5	single	Y	Y	N

872	HCC	male	44	B2	0	201.0	0	0.00	Y	I	5	single	N	N	N
878	HCC	male	54	C1	0	3909.	90	0.00	-	I	5.5	single	N	Y	N
B3	HB	male	32	NA	NA	4.70	-	-	-	autoimm une hepatitis acute	-	-	-	Y	A
B4	AHB	male	22	NA	0	3.41	0.23	-	-	hepatitis acute	-	-	-	Y	A
B5	AHB	male	34	NA	4.39×10E3	1.56	0.11	-	-	hepatitis acute	-	-	-	N	A
SD086	AHB	male	21	NA	0	NA	0.10	-	-	hepatitis B HBeAg- chronic	-	-	-	NA	A
SD077	CHB	male	30	NA	0	4.49	0.46	-	-	hepatitis B HBeAg- chronic	-	-	-	NA	A
SD083	CHB	male	43	NA	0	2.97	0.38	-	-	hepatitis B HBeAg- chronic	-	-	-	NA	A
SD090	CHB	male	41	NA	0	3.41	0.30	-	-	hepatitis B HBeAg- chronic	-	-	-	Y	A

									hepatitis B								N
SD079	CHB	male	28	NA	0	4.49	0.27	-	hepatitis B	-	-	-	-	Y		A	
									cirrhosis	-	-	-	-				
									HBeAg+								
									chronic								
									hepatitis								N
B2	CHB	male	35	NA	0	3.41	11.03	-	B	-	-	-	-	Y		A	
									HBeAg+								
									chronic								
									hepatitis								N
B6	CHB	male	40	C	1.22×10E6	1.56	778.39	-	B	-	-	-	-	N		A	
									HBeAg+								
									chronic								
									hepatitis								N
B7	CHB	fema le	22	B	1.09×10E6	59.04	32.19	-	B	-	-	-	-	N		A	
									HBeAg+								
									chronic								
									hepatitis								N
B8	CHB	male	27	C	2.13×10E5	1.56	256.79	-	B	-	-	-	-	Y		A	
									HBeAg+								
									chronic								
									hepatitis								N
B9	CHB	male	37	NA	5.52×10E8	NA	34.53	-	B	-	-	-	-	Y		A	
SD073	HB	male	41	NA	0	NA	NA	-	HBeAg+	-	-	-	-	NA		N	

									chronic hepatitis B HBeAg+ chronic hepatitis							A
SD075	CHB	male	34	NA	8.97×10E7	4.49	33.07	-	B HBeAg+ chronic hepatitis	-	-	-	NA			N A
SD081	CHB	male	23	NA	0	3.41	3.15	-	B HBeAg+ chronic hepatitis	-	-	-	Y			N A
SD084	CHB	male	37	NA	2.42×10E4	3.41	23.34	-	B HBeAg+ chronic hepatitis	-	-	-	NA			N A
SD088	CHB	male	32	NA	0	NA	308.91	-	B HBeAg+ chronic hepatitis	-	-	-	Y			N A

**Table S1 can be found in the website: <https://github.com/HJlab/HBV-DNA-integration>.**

**Table S2. List of primers in this study**

ID of breakpoint		Forward primer	Reverse primer	
766Ca	S289	TGTGCCTGGCTATGGTGGAGA	TGTCTCAGAGTTTACACGGTAGCAC	
724Ca	S299	AAGTAAGAGGTAGAAGGCTGGAAACA	TAAATGTCGTCTTTTGTGAAGGGTCTGT	
783Ca	S279	GGATGGGACCAGATTCCTGAGTT	GCAGGGACTTCTGTTCTCATCAT	
725Ca	S255	GCACTACCGCCACCACAGCCACAGC	TGAGCCGAGATCACGCCACT	
734Ca	S285	AAGTGCAGTCCCACAAAGC	GTGGCTTCAGCACAGGGATGG	
807Ca	S280	GTTGAAACACTGGTTGTAGGCGAGGTA	GATTGACCTGTATGTAAAGGAGAT	
813Ca	S292	TCCCATCCGAGCTTTCCTGTCCT	GATGAATAGGTGCAGGCATCTCACTC	
812Ca	S291	AGGCTGTGGGGTAACCCGAGG	GGGCGGGATGTGACCAGATGTT	
783Ca	S279	AGGGTGGGGTGGTGGCGGGGAGA	AGGGCAGAAAACAAATCCAGTCACAGT	For validation
775Ca	S290	GAGCCTGAGCCTCGGGCAGTGGG	GAGCCGAGATCACGCCACTG	of HBV size
734Ca	S285	AAATGGGCATCATAACAATCCCTACC	CAGAGGTTGCAGTGAGCCGAGAT	
764Ca	S301	AAATGATGTAATAGACCGTTGTAGACT	AACTCCAAACAGTTTTTGGTCCCT	
768Ca	S303	TGCATACACTGTTAGGGAACGAC	AGCAAATAACATTGGTAAGGGTG	
734Ca	S285	TGGGGTAGGAGTGATGGGTTTATG	CCGTGTTAGCCAGGATGGTCTCG	
766Ca	S289	TCCCTCCAGCCTGAGCAACAAGA	TTTCTGTGACCACCTGTTAT	
764Ca	S301	CTAAGCATCTCCACTTCTCATCA	GCACATTCTTCCAGCCCTACTTACTCT	
813Ca	S292	ATGTAAGCTTCAACTCAACTAGA	TCCTCAAGCAATGGGACCAACTA	
764Ca	S301	TGCCACATCAGAAAGCAGGGAAGA	GGGGCAGAGGATGTTGGTAATGTAAAA	
764Ca	S301	TTCCAATTCTTAAGTCTTC	GGGACTACAGGCACATGCCACCA	
807Ca	S280	TAATGGCCTGTTACAGTTGGTTGA	TTCAACAGAAGAACCACCT	

812Ca	S291	ATGCCACCAACAATACCTACAACCTGC	TCATTGTGGTGGTACAGTGTTTTA	
821Ca	S282	CACCAACCACCTCAGCCTGCCT	TGTCCAAGGAGGCGGAAGGTTAGGG	
764Ca	S301	CCTGACAAATAGTAGGTGCTC	TGGGAAGCCAAAGCCAGAGGAT	
678Ca	KMT2B,NM_014727,exon3	CACTTCCAGCATTTACAGGAGGA	AGGAGGACATTATTAATAGATGTC	
821Ca	KMT2B,NM_014727,exon3	TGTAGGCATAAATTGGTGTGTTTAC	ATCCTCCCCTGGAAAAAGCCCGGA	
821Ca	FAIM2,3'-UTR	TGGAACCTTTGTGTCTCCTCTGCC	TCCTCTGCCTGGATTGGAGA	
725Ca	KMT2B,NM_014727,exon3	GGTTTCACATTTCTGTCTTACTT	CTCCATGCCGGTGAGTG	
725Ca	KMT2B,NM_014727,exon3	GCTATTCAAGATCGATCATGGAG	TGGAGTGTGGATTCGCACTC	For validation
809Ca	KMT2B,NM_014727,exon3	AGGAGGCTGTAGGCATAAGTGGT	TGTAGGCATAAGTGGTGGCGGTA	of expressed
821Ca	KMT2B,NM_014727,exon3	CCACAGCTGCAGCCACCGCCCGTC	GCGGCCACCGCCCGTCTGTGCCTTC	HBx-human
725Ca	KMT2B,NM_014727,exon3	AGCTATTCAAGATCGATCATGGAG	ATGCCCCTATCTTATCAACACTT	chimeric
729Ca	OR52A5,NM_001005160,exon1	TGTAGGCATAAATTGGTCTGTTTAC	TTCACCATATTGTGAAAGTGTT	transcripts
729Ca	OR52A5,NM_001005160,exon1	TTCACCAGCACCATGCAACTTTT	TAAGAGTCACTAACGATCAAGT	
699Ca	ZNF669,NM_024804,exon1	CCCACCGGAACCTGCCCAAGGTC	AAAGGGCTCCGGCCGGCGGA	
828Ca	GFOD2,NM_030819,exon3	TTGCTGCCCCTTTTACACAATGTGG	GGGGAGCCGACCTCTATGGGCAGAAG	
794Ca	KMT2B,NM_014727,exon3	GCCACCGCCGTCACCACAGCAGTG	GTCTTTAATCCTGAGTGGCAAACCTC	

**Table S2 can be found in the website: <https://github.com/HJlab/HBV-DNA-integration>.**



**Table S3. The summary of selected HBV integrations validation**

Sample	Type	Reads number	Total Bases(M)	No.of the reads mapped to HBV	No. of the reads mapped to human	No. of effective reads	10 fold coverage for HBV	Average depth for HBV	No. of integration breakpoints(supporting reads $\geq 2$ )
647Ca	HCC	668157	66.82	37516	26108	627	87.81%	958.52	51
647P	non-HCC	73016	9.13	34977	22518	124	85.30%	1278.72	20
654Ca	HCC	124615	12.46	20712	8139	215	94.22%	739.74	15
654P	non-HCC	172239	21.53	174103	105066	2005	98.67%	5287.36	14
678Ca	HCC	2597797	259.78	6410	1230	108	94.02%	237.98	10
678P	non-HCC	145801	18.23	117943	76569	732	99.77%	4731.41	99
687Ca	HCC	638096	63.81	104936	9307	225	84.25%	2882.58	36
687P	non-HCC	94579	11.82	52899	34827	471	90.57%	1990.69	52
690Ca	HCC	10025998	1002.6	67257	19633	470	97.30%	2158.37	5
690P	non-HCC	107523	13.44	71218	44039	800	98.15%	2927.34	135
693Ca	HCC	2977808	297.78	383	12	5	57.45%	15.35	0
693P	non-HCC	727010	90.88	896354	535127	738	92.10%	6199.10	43
699Ca	HCC	13317133	1331.71	4439683	1678283	13225	100.00%	5711.39	185
699P	non-HCC	233444	23.34	124805	22675	414	84.96%	3631.40	24
700Ca	HCC	1649033	164.9	704181	726057	4790	97.47%	2389.13	27
700P	non-HCC	99507	12.44	52205	34019	415	96.74%	2159.83	43
703Ca	HCC	189188	18.92	6675	2551	84	96.73%	214.67	8
703P	non-HCC	203203	25.4	174786	122656	432	89.55%	5475.60	38
712Ca	HCC	785830	78.58	679335	274214	5268	83.91%	5540.75	76
712P	non-HCC	845661	84.57	612945	235154	3044	93.33%	6826.65	384
715Ca	HCC	3614467	361.45	13105	3287	256	88.69%	453.63	9

715P	non-HCC	94398	9.44	41227	15304	298	98.66%	1532.62	35
722Ca	HCC	4336406	433.64	26449	3164	317	97.31%	877.18	16
722P	non-HCC	145500	18.19	104848	73560	722	98.92%	4131.60	112
724Ca	HCC	690841	69.08	406078	174127	1198	90.05%	6356.59	83
724P	non-HCC	115898	11.59	83960	26073	1024	98.49%	2955.41	98
725Ca	HCC	5031075	503.11	86814	13403	1363	97.83%	3126.48	30
725P	non-HCC	4610734	461.07	473	15	11	51.62%	20.35	1
729Ca	HCC	3393426	339.34	20232	4994	341	96.22%	757.07	16
729P	non-HCC	86440	10.81	41407	25975	358	91.04%	1724.29	33
734Ca	HCC	4754346	475.43	5017493	1873308	31976	90.76%	6073.05	287
734P	non-HCC	103244	12.91	55751	32810	343	97.77%	2433.57	57
745Ca	HCC	1272614	127.26	829549	429920	1984	93.19%	6001.29	33
745P	non-HCC	2294386	229.44	1660363	1035770	2430	92.64%	6002.44	244
747Ca	HCC	7059377	705.94	3380	496	56	96.40%	134.05	3
747P	non-HCC	4073859	407.39	1297	21	19	92.66%	49.71	2
760Ca	HCC	166318	23.95	18867	18383	23	76.27%	676.42	3
760P	non-HCC	103841	10.38	68484	23585	1109	90.37%	2108.59	204
764Ca	HCC	953235	95.32	742642	375990	8500	84.94%	6131.12	207
764P	non-HCC	5455515	545.55	1987572	726731	4009	86.40%	5997.79	459
765Ca	HCC	5153617	515.36	7637	735	223	93.75%	243.78	16
765P	non-HCC	57875	7.23	13816	9259	108	97.61%	553.48	14
766Ca	HCC	4054126	405.41	3606008	1811064	7414	92.78%	6072.14	59
766P	non-HCC	142857	17.86	120702	64604	922	88.99%	4322.66	112
768Ca	HCC	1748573	174.86	1290072	614708	3688	84.20%	4933.59	40
768P	non-HCC	897700	89.77	604326	216010	1154	90.29%	6171.59	200
775Ca	HCC	667048	66.7	123664	65054	1337	97.55%	1215.57	38

775P	non-HCC	133742	13.37	103364	33428	1580	86.68%	3025.14	210
777Ca	HCC	2832504	283.25	4572	771	176	95.36%	169.37	12
777P	non-HCC	45719	4.57	10019	3294	49	97.97%	359.45	9
778Ca	HCC	4382025	438.2	3734	86	87	92.58%	140.04	5
778P	non-HCC	98284	12.29	57100	35832	92	98.51%	2561.49	12
780Ca	HCC	3843882	384.39	309815	28918	1416	87.50%	5088.26	70
780P	non-HCC	199278	24.91	159337	110602	1152	92.61%	5135.02	53
781Ca	HCC	5629418	562.94	9328	2163	52	79.38%	281.02	5
781P	non-HCC	117189	14.65	72362	46508	808	88.37%	2658.27	143
783Ca	HCC	2145890	214.59	1624423	686037	5224	89.67%	5788.20	50
783P	non-HCC	130136	13.01	87314	29868	1437	95.57%	2776.35	56
791Ca	HCC	936743	93.67	780886	290760	7744	83.18%	5398.23	106
791P	non-HCC	1181500	118.15	766533	338493	2437	93.46%	5481.28	196
794Ca	HCC	4371648	437.16	1473	264	112	34.89%	66.31	2
794P	non-HCC	3892826	389.28	1673	283	12	95.40%	65.09	2
795Ca	HCC	4423686	442.37	3210	60	55	73.52%	112.03	9
795P	non-HCC	3171253	317.13	6502	103	58	95.22%	241.07	10
801Ca	HCC	2528200	252.82	7760	1054	142	88.54%	230.87	4
801P	non-HCC	125688	12.57	98265	35108	397	94.72%	3182.07	14
805Ca	HCC	4356998	435.7	2898	117	36	80.38%	90.38	1
805P	non-HCC	71207	7.12	35595	12339	373	98.10%	1192.22	20
807Ca	HCC	4859423	485.94	4984036	2262670	24132	91.85%	6036.96	103
807P	non-HCC	1106755	110.68	771920	303933	2570	96.36%	7078.69	117
809Ca	HCC	3992046	399.2	7695	1923	62	91.91%	253.74	4
809P	non-HCC	1601193	160.12	352498	129377	852	85.47%	5452.72	95
810Ca	HCC	6375672	637.57	11341	1614	320	85.32%	451.47	12

810P	non-HCC	102569	10.26	59806	19512	743	93.16%	1986.17	60
812Ca	HCC	4488875	448.89	4877897	1865676	24891	89.33%	6686.43	206
812P	non-HCC	194506	19.45	182784	59777	1078	94.43%	4949.18	107
813Ca	HCC	3093593	309.36	2496034	1348623	3372	91.85%	5744.19	38
813P	non-HCC	2029750	202.98	1644861	815032	1561	89.25%	6276.42	140
815Ca	HCC	4552540	455.25	8567	208	187	93.84%	298.87	16
815P	non-HCC	3008051	300.81	810	14	11	92.40%	32.51	0
821Ca	HCC	2035333	203.53	1278206	868478	13198	93.96%	4144.41	72
821P	non-HCC	516390	51.64	208261	88384	1741	83.19%	4224.72	99
822Ca	HCC	6077320	607.73	9786	1783	186	89.98%	313.28	18
822P	non-HCC	92430	9.24	51670	16468	634	82.63%	1443.13	91
828Ca	HCC	64245	6.42	66268	27346	795	97.62%	2350.62	50
828P	non-HCC	378009	37.8	438623	122089	376	97.54%	6030.59	37
829Ca	HCC	4737653	473.77	9186	1197	77	93.45%	340.07	11
829P	non-HCC	4380724	438.07	2488	47	17	89.30%	94.13	3
831Ca	HCC	313803	31.38	176466	76114	2595	89.11%	3459.52	113
831P	non-HCC	105714	10.57	64226	21925	777	98.84%	2145.65	150
846Ca	HCC	3630552	363.06	25680	228	88	94.75%	858.31	7
846P	non-HCC	58982	5.9	7112	791	44	86.82%	220.50	5
850Ca	HCC	2185791	218.58	2516173	953350	6585	92.58%	6806.26	88
850P	non-HCC	843673	84.37	1069545	324251	1394	90.09%	6535.76	142
872Ca	HCC	3491418	349.14	625989	402203	2527	83.33%	5239.90	98
872P	non-HCC	148757	14.88	120198	41998	1159	93.42%	3643.43	119
878Ca	HCC	4608399	460.84	1800	23	15	82.97%	72.60	2
878P	non-HCC	46942	4.69	3555	1223	23	95.69%	114.69	1
080962T	HCC	7429565	743.0	1419882	157520	1377	93.77%	16729.30	77

080962L	non-HCC	8577432	857.7	1951638	259328	445	95.19%	16478.16	35
080949T	HCC	11068656	1106.9	5581530	779131	4014	89.44%	15500.71	189
080949L	non-HCC	15478750	1547.9	5604738	955139	1218	92.80%	15014.69	82
081222T	HCC	6135145	613.5	1178391	187114	2211	90.83%	4810.10	234
081222L	non-HCC	7045542	704.6	146407	16743	110	91.07%	14703.18	5
084999T	HCC	8980233	898.0	309115	53648	311	93.20%	4714.33	11
084999L	non-HCC	8564748	856.5	167797	9093	152	95.11%	6938.17	19
085227T	HCC	7376810	737.7	5960	883	56	99.30%	8508.92	4
085227L	non-HCC	8710803	871.1	247218	7305	184	89.67%	149.68	11
B3	Autoimmune hepatitis HBeAg+	11870174	1068.3	1322	54	3	97.42%	51.54	0
B2	chronic hepatitis B	10384651	934.6	1623	84	6	87.66%	42.31	0
B4	Acute hepatitis B	11132275	1001.9	7503	1631	16	95.81%	226.77	1
B5	Acute hepatitis B HBeAg+	6828173	614.5	3536	121	12	92.74%	99.91	2
B6	chronic hepatitis B HBeAg+	8439440	759.5	37589	5247	42	99.20%	1178.64	2
B7	chronic hepatitis B	10844949	976.0	154263	27105	91	90.31%	3489.90	6

B8	HBeAg+ chronic hepatitis B	11120378	1000.8	55670	13838	136	95.81%	1730.84	7
B9	HBeAg+ chronic hepatitis B	12374761	1113.7	1780819	372108	1620	90.77%	5365.43	129
SD073	HBeAg+ chronic hepatitis B	13834259	1245.1	3350	47	4	85.83%	115.42	0
SD075	HBeAg+ chronic hepatitis B	14580127	1312.2	463166	82014	383	85.20%	5113.60	22
SD077	HBeAg- chronic hepatitis B	13544777	1219.0	6432	1395	32	86.10%	182.81	2
SD079	Hepatitis B cirrhosis	13610019	1224.9	10352	1914	37	89.63%	285.05	5
SD081	HBeAg+ chronic hepatitis B	10548500	949.4	721	36	7	79.71%	24.92	0
SD083	HBeAg- chronic hepatitis B	11870955	1068.4	495	22	2	70.06%	16.17	0
SD084	HBeAg+ chronic hepatitis B	11337884	1020.4	8289	1465	19	95.75%	262.02	1

SD086	Acute hepatitis B HBeAg+	12353976	1111.9	909	10	2	72.28%	34.16	0
SD088	chronic hepatitis B HBeAg-	6599489	594.0	4983	831	22	90.19%	135.69	0
SD090	chronic hepatitis B	6514011	586.3	229	7	3	29.44%	7.45	0

**Table S3 can be found in the website: <https://github.com/HJlab/HBV-DNA-integration>.**

**Table S4. Summary of sequencing statistics in all samples**

<b>Category</b>	<b>HCC(n=54)</b>	<b>non-HCC(n=54)</b>	<b>HB(n=18)</b>
<b>Total reads</b>	206,817,190	93,042,772	197,788,798
<b>Average reads</b>	3,829,948	1,723,014	10,988,267
<b>Total effective reads</b>	185,736	45,166	2,437
<b>Average effective reads</b>	3,440	836	135
<b>Total detected breakpoints</b>	2,870	4,466	177
<b>Average breakpoints</b>	53	83	10

**Table S4 can be found in the website: <https://github.com/HJlab/HBV-DNA-integration>.**



**Table S5. 75 randomly selected integration breakpoints for validation by PCR and Sanger sequencing**

Sample	Sample type	Human location	Nearest gene (distance)	Location type	HBV location	HBV domain	Orientation	Supporting reads number	Human location	Sanger validation
647Ca	HC C	chr22:4908 2817	FAM19A5	intronic	1917	precore/core protein Core and e antigen	hum hbv +-	26	chr22	no
654Ca	HC C	chr14:7898 1628	NRXN3	intronic	1828	X protein precore/core protein	hum hbv ++	22	chr14	yes
654Ca	HC C	chr2:17127 4896	MYO3B	intronic	1765	X protein	hum hbv +-	29	chr2	yes
678Ca	HC C	chr19:3621 2805	MLL4	intronic	2553	Polymerase	hum hbv ++	6	chr19	yes
690Ca	HC C	chr19:3621 2890	MLL4	intronic	359	S protein Polymerase large S protein middle S protein	hum hbv +-	118	chr19	yes
699Ca	HC C	chr1:24726 7270	ZNF669	exonic	1703	X protein	hum hbv --	208	chr1	yes
699Ca	HC C	chr11:4613 2644	PHF21A	intronic	398	S protein Polymerase large S protein middle S protein	hum hbv +-	772	chr11	yes
700Ca	HC C	chr19:1863 4749	ELL(dist=1812), FKBP8(dist=7819)	intergenic	1674	X protein	hum hbv --	570	chr19	no
700Ca	HC C	chr5:12967 75	TERT(dist=1613), MIR4457(dist=12650)	intergenic	116	Polymerase large S protein middle S protein	hum hbv ++	2668	chr5	yes
715Ca	HC C	chr14:4146 2121	FBXO33(dist=1560417), LRFN5(dist=614643)	intergenic	2888	Polymerase large S protein	hum hbv +-	14	chr14	yes
715Ca	HC	chr15:4967	FAM227B	intronic	636	S protein Polymerase large S	hum hbv +-	10	chr15	no

	C	6946				protein middle S protein					
715Ca	HC	chr5:12958	TERT	upstream	506	S protein Polymerase large S	hum hbv +-	82	chr5	yes	
	C	01				protein middle S protein					
715Ca	HC	chr5:12958	TERT	upstream	1897	precore/core protein	hum hbv -+	19	chr5	yes	
	C	01									
722Ca	HC	chr17:3823	THRA	intronic	573	S protein Polymerase large S	hum hbv -+	46	chr17	no	
	C	9834				protein middle S protein					
724Ca	HC	chr10:7049	CCAR1	intronic	2340	precore/core protein Core and e	hum hbv -+	90	chr10	yes	
	C	2716				antigen Polymerase					
724Ca	HC	chr11:2464	LUZP2	intronic	2496	Polymerase	hum hbv +-	60	chr11	yes	
	C	9990									
724Ca	HC	chr11:2465	LUZP2	intronic	1840	precore/core protein	hum hbv ++	43	chr11	yes	
	C	4658									
725Ca	HC	chr19:3621	MLL4	exonic	2266	precore/core protein Core and e antigen	hum hbv ++	208	chr19	yes	
	C	2668									
725Ca	HC	chr19:3621	MLL4	exonic	2220	precore/core protein Core and e antigen	hum hbv --	122	chr19	yes	
	C	2696									
725Ca	HC	chr4:10376	UBE2D3	intronic	1765	X protein	hum hbv +-	60	chr4	yes	
	C	2567									
729Ca	HC	chr11:5152	OR52A5	exonic	1809	X protein	hum hbv +-	22	chr11	yes	
	C	963									
729Ca	HC	chr8:71126	NCOA2	intronic	2991	Polymerase large S protein	hum hbv +-	20	chr8	yes	
	C	926									
734Ca	HC	chr15:9045	C15orf38,	intronic	1829	X protein precore/core protein	hum hbv ++	1169	chr15	yes	
	C	4143	C15orf38-AP3S2								
734Ca	HC	chr20:3056	XKR7	intronic	1557	X protein Polymerase	hum hbv +-	320	chr20	no	

	C	4086									
734Ca	HC	chr20:3056									
	C	4095	XKR7	intronic	1483	X protein Polymerase	hum hbv +-	329	chr20	yes	
734Ca	HC	chr20:3056									
	C	4095	XKR7	intronic	1554	X protein Polymerase	hum hbv +-	602	chr20	yes	
734Ca	HC	chr5:18047									
	C	7469	BTNL9	intronic	1386	X protein Polymerase	hum hbv +-	950	chr5	yes	
734Ca	HC	chr9:28517									
	C	157	LINGO2	intronic	1992	precore/core protein Core and e antigen	hum hbv --	655	chr9	yes	
734Ca	HC	chr9:28517									
	C	157	LINGO2	intronic	2064	precore/core protein Core and e antigen	hum hbv --	954	chr9	yes	
745Ca	HC	chr1:24174									
	C	1597	KMO	intronic	521	S protein Polymerase large S protein middle S protein	hum hbv ++	732	chr1	yes	
745Ca	HC	chr8:12952	MIR1208(dist=362474),								
	C	4908	LOC728724(dist=703805)	intergenic	1612	X protein Polymerase	hum hbv ++	646	chr8	yes	
764Ca	HC	chr1:96406									
	C	69	SLC25A33	intronic	1820	X protein precore/core protein	hum hbv ++	51	chr1	yes	
764Ca	HC	chr12:2069									
	C	1805	PDE3A	intronic	2713	Polymerase	hum hbv ++	401	chr12	yes	
764Ca	HC	chr12:2250									
	C	290	CACNA1C	intronic	434	S protein Polymerase large S protein middle S protein	hum hbv --	382	chr12	yes	
764Ca	HC	chr12:3098									
	C	319	TEAD4	intronic	2634	Polymerase	hum hbv +-	501	chr12	yes	
764Ca	HC	chr12:7291									
	C	403	CLSTN3	intronic	2263	precore/core protein Core and e antigen	hum hbv --	249	chr12	yes	
764Ca	HC	chr16:6767									
			CTCF(dist=1799),	intergenic	1461	X protein Polymerase	hum hbv --	17	chr16	no	

	C	4887	RLTPR(dist=4143)								
765Ca	HC	chr5:12951	TERT	upstream	1818	X protein precore/core protein	hum hbv +-	47	chr5	yes	
	C	73									
766Ca	HC	chr14:8101	CEP128	intronic	1821	X protein precore/core protein	hum hbv +-	529	chr14	yes	
	C	4676									
766Ca	HC	chr5:12974	TERT(dist=2265),	intergenic	1693	X protein	hum hbv --	861	chr5	yes	
	C	27	MIR4457(dist=11998)								
766Ca	HC	chr9:28282	LINGO2	intronic	45	Polymerase large S protein middle S protein	hum hbv --	623	chr9	yes	
	C	905									
768Ca	HC	chr5:11282	CTNND2	intronic	1495	X protein Polymerase	hum hbv +-	476	chr5	yes	
	C	976									
775Ca	HC	chr19:3621	MLL4	intronic	1554	X protein Polymerase	hum hbv ++	982	chr19	yes	
	C	2884									
775Ca	HC	chr19:3866	SIPA1L3	intronic	1061	Polymerase	hum hbv +-	13	chr19	no	
	C	4508									
777Ca	HC	chr17:2108	DHRS7B	intronic	389	S protein Polymerase large S protein middle S protein	hum hbv --	50	chr17	no	
	C	2020									
780Ca	HC	chr5:13978	ANKHD1,	intronic	3100	Polymerase large S protein	hum hbv ++	19	chr5	yes	
	C	4760	ANKHD1-EIF4EBP3								
780Ca	HC	chr9:75563	ALDH1A1	intronic	1309	Polymerase	hum hbv --	87	chr9	yes	
	C	798									
781Ca	HC	chr2:19897	PLCL1	intronic	1371	Polymerase	hum hbv --	27	chr2	yes	
	C	3674									
783Ca	HC	chr6:15447	IPCEF1	UTR3	1432	X protein Polymerase	hum hbv +-	517	chr6	yes	
	C	7145									
791Ca	HC	chr19:3621	MLL4	exonic	2196	precore/core protein Core and e antigen	hum hbv +-	139	chr19	yes	

	C	2659										
791Ca	HC	chr19:3621	MLL4	intronic	1800	X protein	hum hbv --	326	chr19	yes		
	C	3045										
794Ca	HC	chr19:3621	MLL4	exonic	2511	Polymerase	hum hbv ++	33	chr19	yes		
	C	2554										
807Ca	HC	chr10:8854	BMPR1A	intronic	234	S protein Polymerase large S protein middle S protein	hum hbv ++	3273	chr10	yes		
	C	1372										
807Ca	HC	chr22:4908	FAM19A5	intronic	1917	precore/core protein Core and e antigen	hum hbv +-	1075	chr22	no		
	C	2817										
809Ca	HC	chr19:3621	MLL4	exonic	1731	X protein	hum hbv +-	38	chr19	yes		
	C	2519										
810Ca	HC	chr19:4877	ZNF114	intronic	1774	X protein	hum hbv +-	17	chr19	yes		
	C	8613										
810Ca	HC	chr5:12957	TERT	upstream	1728	X protein	hum hbv +-	83	chr5	yes		
	C	00										
812Ca	HC	chr16:6767	CTCF(dist=1799), RLTPR(dist=4143)	intergenic	1418	X protein Polymerase	hum hbv --	686	chr16	no		
	C	4887										
812Ca	HC	chr19:3621	MLL4	exonic	1808	X protein	hum hbv --	12	chr19	no		
	C	2557										
815Ca	HC	chr19:3620	ZBTB32	UTR3	2750	Polymerase	hum hbv ++	7	chr19	yes		
	C	7835										
821Ca	HC	chr12:1133	RPH3A	intronic	2189	precore/core protein Core and e antigen	hum hbv --	1237	chr12	yes		
	C	19972										
821Ca	HC	chr12:5026	FAIM2	UTR3	1204	Polymerase	hum hbv --	1182	chr12	yes		
	C	1607										
821Ca	HC	chr12:8843	C12orf29	intronic	2902	Polymerase large S protein	hum hbv --	752	chr12	yes		

	C	1764									
821Ca	HC	chr19:3621									
	C	2538	MLL4	exonic	1547	X protein Polymerase	hum hbv ++	768	chr19	yes	
821Ca	HC	chr19:3621									
	C	2557	MLL4	exonic	1731	X protein	hum hbv --	720	chr19	yes	
821Ca	HC	chr19:3621									
	C	2557	MLL4	exonic	1805	X protein	hum hbv --	1156	chr19	yes	
828Ca	HC	chr10:8851	LDB3(dist=16914),								
	C	2738	BMPRI1A(dist=3658)	intergenic	1837	X protein precore/core protein	hum hbv ++	129	chr10	yes	
828Ca	HC	chr10:8851	LDB3(dist=16924),								
	C	2748	BMPRI1A(dist=3648)	intergenic	2325	precore/core protein Core and e antigen Polymerase	hum hbv --	95	chr10	yes	
828Ca	HC	chr16:6770									
	C	9426	GFOD2	exonic	1041	Polymerase	hum hbv +-	47	chr16	yes	
831Ca	HC	chr4:15840									
	C	797	CD38	intronic	2895	Polymerase large S protein	hum hbv ++	21	chr4	yes	
831Ca	HC	chr4:15840									
	C	843	CD38	intronic	453	S protein Polymerase large S protein middle S protein	hum hbv --	864	chr4	yes	
850Ca	HC	chr11:1202									
	C	10369	ARHGEF12	intronic	396	S protein Polymerase large S protein middle S protein	hum hbv ++	4	chr11	yes	
850Ca	HC	chr5:12961									
	C	39	TERT	upstream	986	Polymerase	hum hbv ++	752	chr5	yes	
872Ca	HC	chr5:12966	TERT(dist=1516),								
	C	78	MIR4457(dist=12747)	intergenic	509	S protein Polymerase large S protein middle S protein	hum hbv +-	462	chr5	yes	
872Ca	HC	chr5:12966	TERT(dist=1516),								
	C	78	MIR4457(dist=12747)	intergenic	581	S protein Polymerase large S protein middle S protein	hum hbv +-	632	chr5	yes	

**Table S5 can be found in the website: <https://github.com/HJlab/HBV-DNA-integration>.**

**Table S6 and S7 are so big that can't be presented in this file.**

**Table S6 and S7 can be found in the website: <https://github.com/HJlab/HBV-DNA-integration>.**

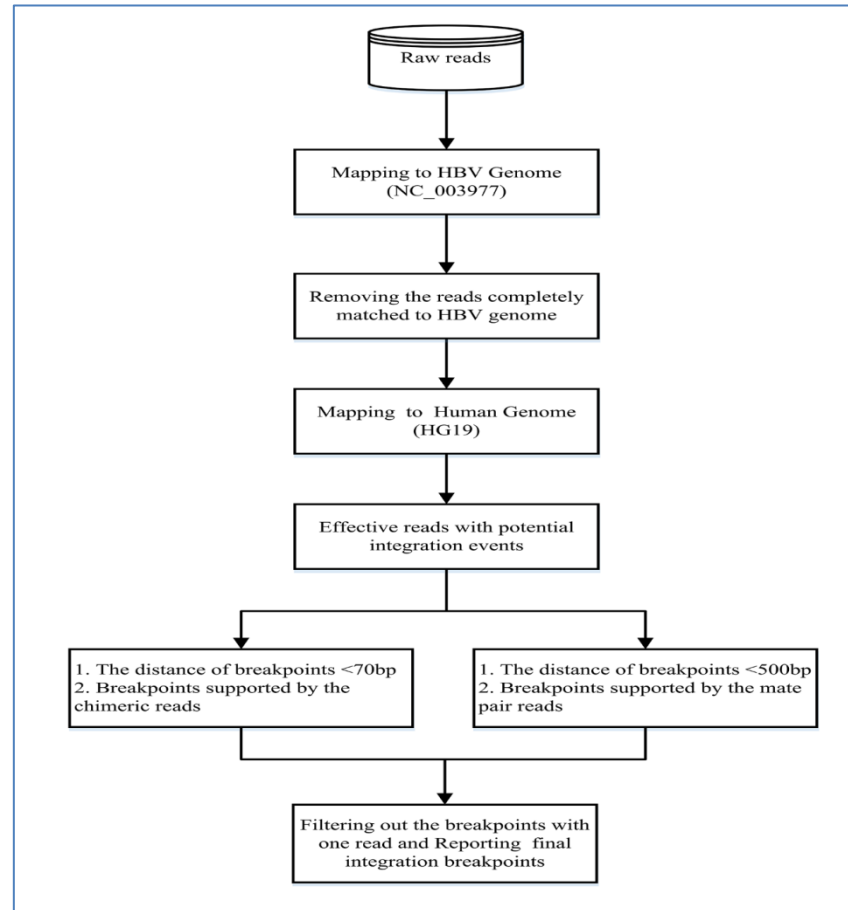


Fig. S1. Flow chat of data analysis including mapping to the human and HBV genome, filtering of effective reads and definition of integration breakpoints.



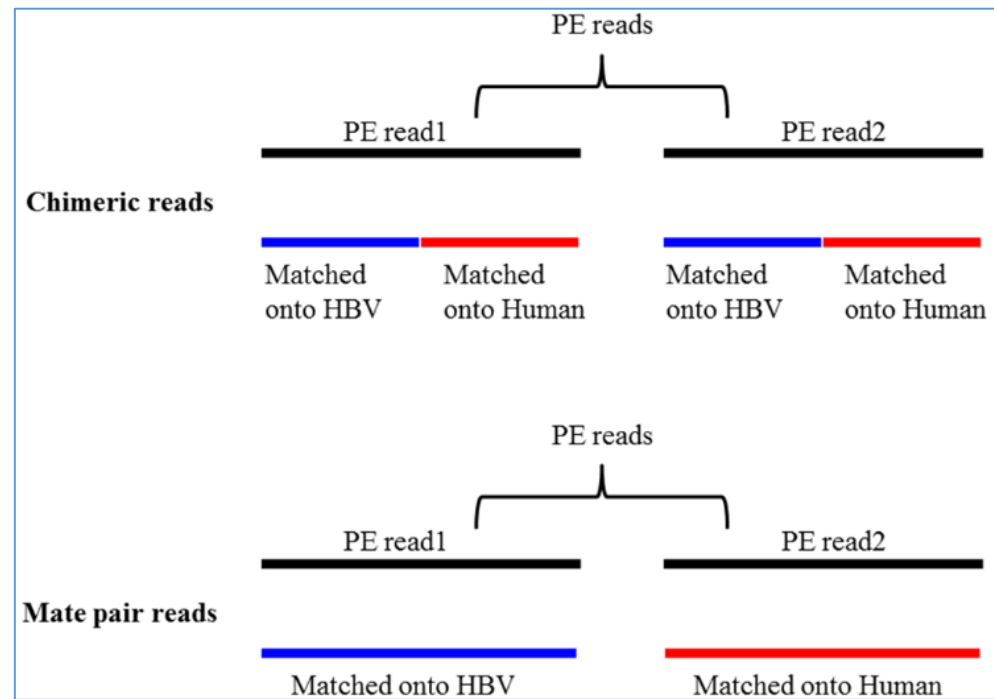


Fig. S2. Chimeric reads and mate pair reads analyzed for detecting HBV integration breakpoints. If the distance of breakpoints supported by the chimeric reads was less than 70bp, the upstream breakpoint was defined as the final HBV integration breakpoint. If the distance of breakpoints supported by the mate pair reads was less than 500bp, the upstream breakpoint was defined as the final HBV integration breakpoint.

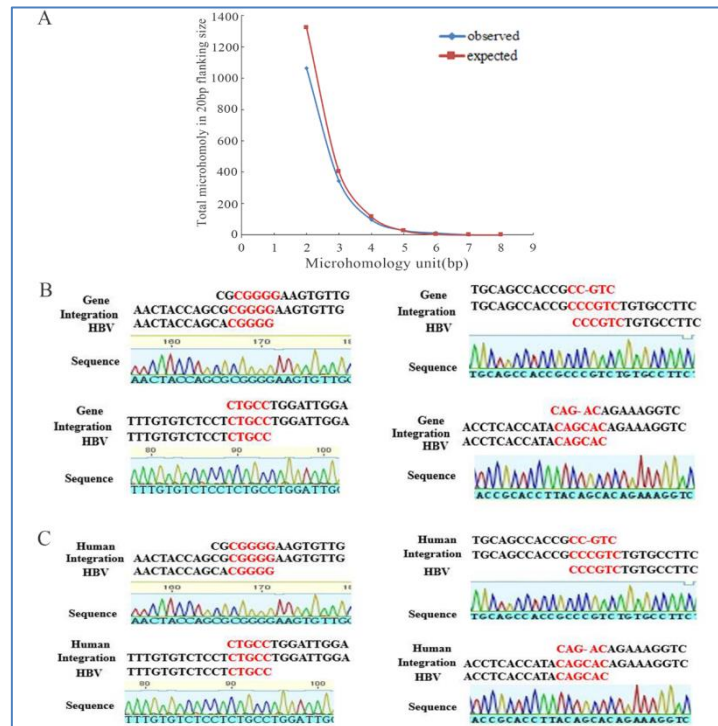


Fig. S3. Microhomologies analysis of the regions flanking HBV integration sites. (a) Comparison between observed and expected integrations with microhomologies in flanking regions of different sizes. (b) Alignment of the sequence around the integration site between the human genome and the HBV genome. Validation of microhomologies near the human-HBV integration breakpoints using sanger sequencing.

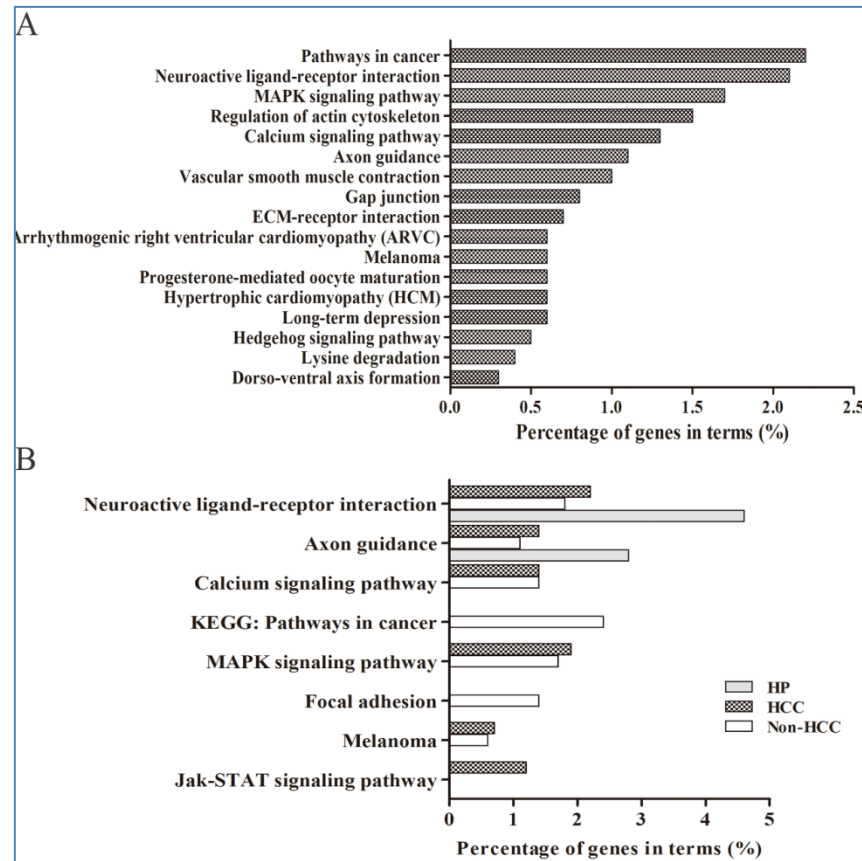


Fig. S4. KEGG analysis of integration related genes (ITGs). (a) KEGG analysis of ITGs from total samples. (b) KEGG analysis of ITGs from HP, non-HCC and HCC samples.