Targeted long-read sequencing reveals clonally expanded HBV-associated chromosomal translocations in patients with chronic hepatitis B

Nicholas van Buuren, Ricardo Ramirez, Cameron Soulette, Vithika Suri, Dong Han, Lindsey May, Scott Turner, P.C. Parvangada, Ross Martin, Henry L.Y. Chan, Patrick Marcellin, Maria Buti, Nam Bui, Neeru Bhardwaj, Anuj Gaggar, Li Li, Hongmei Mo, Becket Feierbach

Table of contents
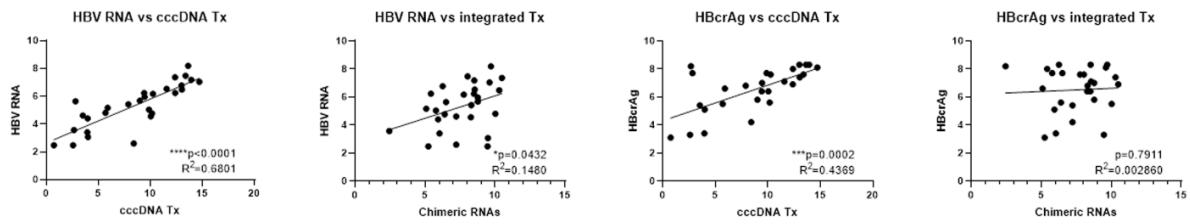
Fig. S1. Correlation of the HBV transcriptome to HBV RNA and HBcrAg. Normalized cccDNA transcription and normalized integrated transcription (chimeric and non-chimeric non-canonical combined) was correlated to both HBV RNA and HBcrAg levels quantified from matched serum samples.
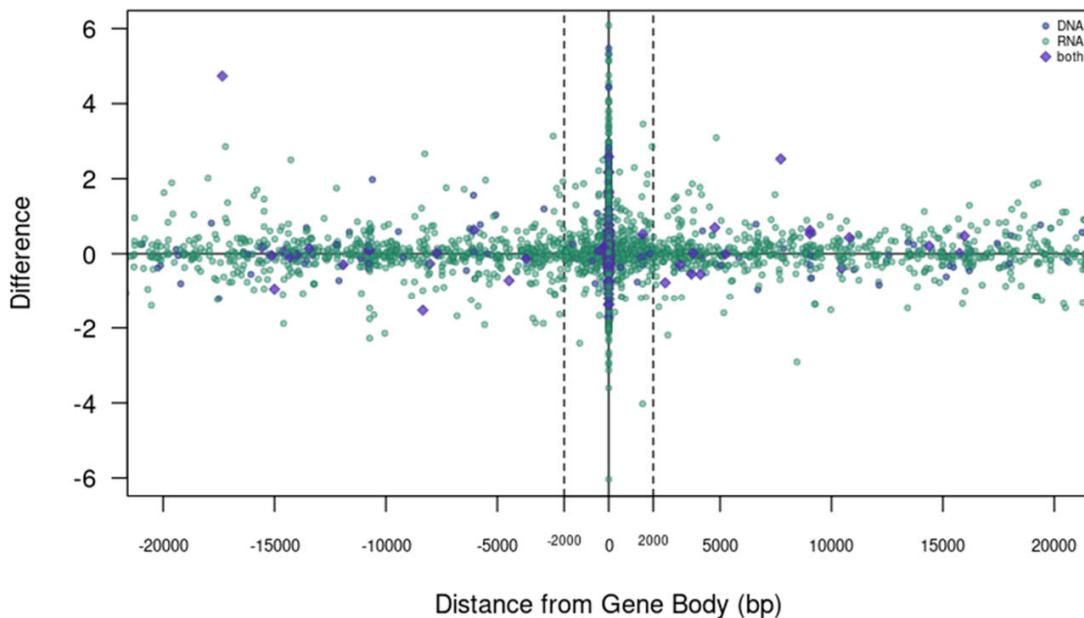
Fig. S2. Alteration of the host transcriptome by integrated HBV. Mean normalized gene expression across the entire sample collection was generated. We then plotted mean normalized changes in gene expression for genes in close proximity to each integration event we identified by either RNA-Seq (green circle), targeted PacBio or WGS (blue circle), or both (blue diamond). Transcriptional changes are mapped based on the distance in base pairs between the integration event and the associated gene body.

Sample ID	Timepoint	Baseline HBeAg	BL HBsAg [log10 IU/mL]	BL HBV DNA [log10 IU/mL]	BL ALT [IU]	RNAseq	WGS	Targeted DNA-Seq	Targeted Iso-Seq	mIF	HBcrAg, HBV RNA
G149.01	Baseline	Reactive	3.46488283	7.2571452	176	Y				Y	Y
G149.02	Baseline	Reactive	4.81276139	8.9055	166	Y	Y	Y	Y	Y	Y
G149.03	Baseline	Reactive	Screen Fail			Y		Y	Y	Y	
G149.04	Baseline	Non-Reactive	3.28319264	7.3346429	168	Y		Y	Y	Y	Y
G149.05	Baseline	Reactive	3.62439246	7.240933	95	Y		Y	Y	Y	Y
G149.06	Baseline	Reactive	3.83696748	4.3756088	93	Y		Y	Y	Y	Y
G149.07	Baseline	Reactive	3.74267408	7.5421324	200	Y		Y	Y	Y	Y
G149.07w	Week 96	Reactive	3.74267408	7.5421324	200	Y	Y		Y		Y
G149.08	Baseline	Reactive	3.87090347	7.45	347	Y		Y	Y	Y	Y
G149.08w	Week 96	Reactive	3.87090347	7.45	347	Y		Y	Y		Y
G149.09	Baseline	Non-Reactive	3.61206721	3.7856145	6	Y		Y	Y	Y	Y
G149.10w	Week 96	Reactive	4.4397409	8.9307598	308	Y	Y	Y		Y	Y
G149.11w	Week 96	Reactive	4.20127693	9	103	Y	Y	Y	Y	Y	Y
G149.12w	Week 96	Reactive	4.39777234	8.8883413	80	Y	Y	Y	Y		Y
G149.13w	Week 96	Reactive	5.39750551	9	94	Y	Y	Y			Y
G149.14w	Week 96	Reactive	3.97588058	7.4883617	38	Y	Y	Y	Y		Y
G149.15	Baseline	Reactive	Screen Fail			Y		Y	Y	Y	
G149.16	Baseline	Non-Reactive	3.30193672	4.7099988	116	Y		Y	Y	Y	Y
G149.17	Baseline	Reactive	4.22557159	8.6958641	72	Y		Y	Y	Y	Y
G149.18	Baseline	Reactive	4.01051493	7.4819323	108	Y		Y	Y	Y	
G149.19	Baseline	Reactive	4.40114196	8.4330947	236	Y		Y	Y	Y	Y
G149.20	Baseline	Reactive	4.39103573	8.2204619	49	Y		Y	Y	Y	Y
G149.21	Baseline	Non-Reactive	Screen Fail			Y					
G149.22	Baseline	Non-Reactive	3.74952375	6.8673464	341	Y	Y	Y	Y		Y
G149.22w	Week 96	Non-Reactive	3.74952375	6.8673464	341	Y		Y	Y		
G149.23	Baseline	Reactive	4.17991441	8.6892629	74	Y					
G149.24	Baseline	Reactive	3.48306069	6.4765896	66	Y	Y				Y
G149.24w	Week 96	Reactive	3.48306069	6.4765896	66	Y		Y	Y		Y
G149.25	Baseline	Reactive	3.90770581	7.1989915	68	Y	Y	Y	Y	Y	Y
G149.25w	Week 96	Reactive	3.90770581	7.1989915	68	Y					Y
G149.26	Baseline	Reactive	2.35472293	5.170247	47	Y	Y	Y	Y	Y	Y
G149.27	Baseline	Reactive	4.14684062	7.9665435	31	Y	Y	Y	Y	Y	Y
G149.27w	Week 96	Reactive	4.14684062	7.9665435	31	Y		Y	Y		Y
G149.28	Baseline	Reactive	3.44964499	6.6301	77	Y	Y	Y	Y	Y	Y
G149.29	Baseline	Reactive	Screen Fail			Y					
G149.30	Baseline	Non-Reactive	4.49466125	3.9545802	91	Y		Y	Y		Y
G149.31	Baseline	Reactive	Screen Fail			Y					Y
G149.32	Baseline	Reactive	Screen Fail			Y				Y	
G149.33	Baseline	Non-Reactive	Screen Fail			Y	Y	Y	Y	Y	Y
G149.34	Baseline	Non-Reactive	2.10839587	5.2207381	188	Y		Y	Y	Y	Y
G149.35	Baseline	Non-Reactive	3.89993863	7.2218982	67	Y		Y	Y	Y	Y
G149.35w	Week 96	Non-Reactive	3.89993863	7.2218982	67	Y		Y	Y		
G149.36	Baseline	Non-Reactive	3.59898409	6.0190685	42	Y		Y	Y	Y	Y
G149.37	Baseline	Non-Reactive	3.87786239	7.5634164	191	Y		Y	Y	Y	Y
G149.38	Baseline	Non-Reactive	3.4107047	4.1411361	27	Y			Y	Y	Y
G149.39	Baseline	Non-	3.21170508	6.6413879	147	Y		Y	Y	Y	Y

		Reactive								
G149.40	Baseline	Non-Reactive	3.5251835	7.2571741	103	Y		Y	Y		Y
G149.41	Baseline	Non-Reactive	3.15934157	5.1170922	80	Y					Y
G149.42	Baseline	Reactive	3.48829966	5.7893205	74	Y		Y	Y		Y
G149.43	Baseline	Reactive	3.14243327	7.1225373	41	Y		Y	Y		Y
G149.43w	Week 96	Reactive	3.14243327	7.1225373	41	Y		Y	Y		Y
G149.44	Baseline	Reactive	3.23251584	6.776101	38	Y			Y		Y
G149.45	Baseline	Non-Reactive	3.84555319	7.7313037	180	Y		Y	Y	Y	Y
G149.46	Baseline	Reactive	2.01986371	4.6085688	152	Y	Y	Y	Y	Y	

Table S1: List of Samples and Baseline clinical characteristics. All patients who were analyzed in this manuscript have been assigned a random patient number for tracking. Assigned patient IDs are also used in our companion manuscript [1]. The assays performed on each sample are denoted with a "Y". Baseline clinical characteristics including HBV DNA, HBsAg, HBeAg status and ALT are listed for each patient.

Patient	Timepoint	Host_on_target_reads	HBV_reads	HBV_chimeric_reads	Junctions	Translocations	Mean_reads_per_junction	Median_reads_per_junction	DR1_Junctions
02	baseline	1356	1715	76	9	1	8.2	5	2
03	baseline	39738	6414	3541	27	4	168	99	8
04	baseline	284	60	6	1	0	2	2	0
05	baseline	54001	34228	4932	63	1	83.9	46	7
06	baseline	88540	2898	2554	21	0	134.7	151	7
07	baseline	158	17	4	0	0	NaN		0
08	baseline	295	142	7	1	0	2	2	1
08	week_96	5503	134	131	7	0	18.6	18	3
09	baseline	191	9	8	3	0	2	2	1
10	week_96	29899	1177	1176	32	1	43.6	9	3
11	week_96	42899	1091	1087	31	1	45.8	14	10
12	week_96	1526	119	111	5	0	22.8	6	2
13	week_96	40362	1817	1763	55	2	41.8	13	10
14	week_96	24978	1286	1143	23	0	54.5	42	7
15	baseline	50112	15959	1947	53	4	41.8	14	8
16	baseline	98092	1405	1194	11	0	115.7	90	3
17	baseline	60212	73266	2965	81	0	33.2	5	9
18	baseline	56904	6779	2557	34	5	101.1	53.5	6
19	baseline	208	101	11	3	0	2	2	1
20	baseline	123	26	6	2	0	2	2	2
22	baseline	784	64	62	5	0	12.4	10	1
22	week_96	7999	156	95	10	0	10.2	6.5	1
24	week_96	15050	211	211	7	0	31.9	13	3
25	baseline	1190	361	153	7	0	21.9	19	3
26	baseline	50079	399	366	6	0	79.3	25.5	1
27	baseline	21243	3319	730	22	0	32.7	18	7
27	week_96	6653	169	167	13	4	15.9	10	4
28	baseline	2441	138	103	3	0	33.7	21	1
30	baseline	83019	8496	8171	66	3	140.6	105	33
33	baseline	37441	6202	6036	22	4	329.5	75.5	6
34	baseline	52557	6973	2485	37	0	82.2	33	20
35	baseline	19359	1792	1189	20	1	79.2	68	8
35	week_96	5417	251	250	18	0	14.8	14	11
36	baseline	270	20	14	4	0	3	2	0
37	baseline	180	59	2	0	0	NaN		0
39	baseline	7461	168	167	7	0	29.7	27	2
40	baseline	11516	1048	774	23	0	37.4	24	9
42	baseline	4605	324	131	11	0	12.1	7	1
43	baseline	8945	528	147	3	0	48.3	7	0
43	week_96	9616	160	109	6	0	28.2	18.5	2
45	baseline	9258	1055	499	25	0	19.6	17	10
46	baseline	52147	1553	1495	28	2	85.2	21	2

Table S2: Targeted DNA Sequencing Summary. Each row has data from each liver biopsy. Junctions and translocations are based on a 2 read minimum. Junctions are classified as DR1 if they are within 20 bp of the DR1 sequence.

Patient	Timepoint	Contig Name	junction1.chr	junction1.pos	hbv.pos1	hbv.orientation1	junction2.chr	junction2.pos	hbv.pos2	hbv.orientation2	nread	longest_HBV_segment	DR1_associated
20	baseline	G149.20b_1_45	chr9	102458696	3161	+	chr2	237697410	3161	+	45	3196	no
36	baseline	G149.36b_6_196	chr12	63260437	3096	+	chr12	63260477	3096	+	196	3189	no
37	baseline	G149.37b_9_81	chr2	87522200	3104	+	chr2	87522176	3104	+	81	3186	no
18	baseline	G149.18b_7_175	chr5	85680722	2655	+	chr14	56123995	2655	+	175	3174	no
35	baseline	G149.35b_6_73	chr13	65471786	3108	+	chr13	65471793	3108	+	73	3160	no
39	baseline	G149.39b_3_89	chr6	34649285	3179	+	chr6	34649288	3179	+	89	3144	no
14	week_96	G149.14w_3_105	chr8	23122421	3195	+	chr8	23122034	697	+	105	3127	no
43	week_96	G149.43w_1_61	chr5	23826792	3207	+	chr5	23826829	3207	+	61	3122	no
33	baseline	G149.33s_6_223	chr11	42817010	3281	+	chr11	42817001	3281	+	223	3058	no
37	baseline	G149.37b_2_212	chr15	65230302	20	+	chr15	65230253	20	+	212	2967	no
37	baseline	G149.37b_25_30	chr20	60864930	3352	+	chr4	11309	3352	+	30	2932	no
10	week_96	G149.10w_5_91	chr5	140479240	53	+	chr5	140479192	53	+	91	2925	no
37	baseline	G149.37b_5_116	chr3	31908648	3298	+	chr3	31908924	3298	+	116	2906	no
30	baseline	G149.30b_39_32	chr22	15858702	3491	+	chr14	19263655	3491	+	32	2809	no
35	baseline	G149.35b_1_208	chr1	103667777	3484	+	chr18	43043065	3484	+	208	2767	no
03	baseline	G149.03b_3_338	chr2	81712955	345	+	chr2	81712954	345	+	338	2756	no
37	baseline	G149.37b_16_50	chr19	40755964	3599	+	chr19	40755966	3599	+	50	2691	no
34	baseline	G149.34b_1_764	chr2	133541783	326	+	chr2	133541829	326	+	764	2652	no
11	week_96	G149.11w_1_343	chr1	161220160	3703	+	chr1	161220160	3703	+	343	2614	no
07	baseline	G149.07b_6_102	chr2	215419885	3806	+	chr2	215419845	3806	+	102	2551	no
09	baseline	G149.09b_2_335	chr1	239097738	556	+	chr1	239095040	556	+	335	2541	no
30	baseline	G149.30b_17_210	chr5	175927691	3989	+	chr5	176100288	3989	+	210	2379	no
05	baseline	G149.05b_17_130	chr4	154557661	647	+	chr4	154643761	647	+	130	2378	no
08	week_96	G149.08w_4_49	chr9	93853280	983	+	chr9	93853242	983	+	49	2373	no
16	baseline	G149.16b_1_470	chr14	87159892	2701	+	chr14	87159908	3996	+	470	2348	no
35	baseline	G149.35b_2_188	chr2	185912056	742	+	chr2	185912068	742	+	188	2330	no
13	week_96	G149.13w_9_61	chr21	9140276	3160	+	chr21	9140105	3160	+	61	2326	no
15	baseline	G149.15b_16_34	chr6	69427905	4990	+	chr6	69427888	4015	+	34	2323	no
34	baseline	G149.34b_3_234	chr20	11172144	3941	+	chr20	11172150	3941	+	234	2314	no
27	week_96	G149.27w_3_46	chr3	93641001	271	+	chr3	92844993	271	+	46	2295	no
40	baseline	G149.40b_4_39	chr5	10282320	3947	+	chr5	10282382	3947	+	39	2262	no
09	baseline	G149.09b_3_301	chr11	58957376	728	+	chr11	58957378	728	+	301	2180	no
35	week_96	G149.35w_2_77	chr14	68693871	5539	-	chr8	5375799	5539	-	77	2154	no
05	baseline	G149.05b_12_158	chr5	8399908	3370	+	chr5	8399881	897	+	158	2062	no
34	baseline	G149.34b_2_320	chr12	57700786	3883	+	chr12	57700786	1615	+	320	1934	no
11	week_96	G149.11w_4_93	chr16	87882777	3288	+	chr16	87738547	3288	+	93	1866	no
13	week_96	G149.13w_3_120	chr22	11814038	2701	+	chr13	82424331	2701	+	120	1795	no
15	baseline	G149.15b_12_48	chr15	101962915	4333	+	chr5	11547	4333	+	48	1770	no

36	baseline	G149.36b_3_353	chr13	52161786	29	+	chr13	52161857	4692	+	353	1646	no
27	week_96	G149.27w_4_31	chr8	91819367	4668	+	chr4	181452947	4668	+	31	1636	no
46	baseline	G149.46s_2_409	chr13	114092825	4196	+	chr13	114092832	4196	+	409	1635	no
46	baseline	G149.46s_8_36	chr7	70841818	4793	+	chr7	70841768	4793	+	36	1542	no
10	week_96	G149.10w_8_37	chr9	122018023	2920	+	chr9	122018010	2920	+	37	1504	no
11	week_96	G149.11w_7_45	chr14	39201164	4917	+	chr14	39201199	4917	+	45	1395	no
19	baseline	G149.19b_3_84	chr16	86153991	4545	+	chr16	86475206	4545	+	84	1378	no
08	week_96	G149.08w_3_58	chr6	59083306	4042	+	chr6	59562852	4042	+	58	1310	no
08	baseline	G149.08b_1_63	chr7	53195657	3268	+	chr7	53198515	3268	+	63	1203	no
13	week_96	G149.13w_13_28	chr5	41879500	815	+	chr5	49095039	815	+	28	1164	no
46	baseline	G149.46s_1_433	chr14	52346943	2972	+	chr11	20832401	2972	+	433	1068	no
46	baseline	G149.46s_4_121	chr4	80258396	1092	+	chr15	53446255	1092	+	121	1020	no
43	baseline	G149.43b_1_135	chr12	118854879	1805	+	chr12	10808327	1805	+	135	1014	no
42	baseline	G149.42b_2_79	chr21	29739443	2504	+	chr21	28660789	2504	+	79	998	no
37	baseline	G149.37b_27_25	chr7	159335248	2055	+	chr4	156030739	2055	+	25	988	no
13	week_96	G149.13w_1_431	chr13	74177474	4386	+	chr8	23386540	4386	+	431	914	no
10	week_96	G149.10w_4_97	chr13	30591762	4856	+	chr13	30593423	4856	+	97	879	no
26	baseline	G149.26b_1_343	chr21	27051754	2107	+	chr21	27051740	2107	+	343	878	no
18	baseline	G149.18b_6_188	chr14	33014204	5388	+	chr3	112749272	5388	+	188	848	no
08	week_96	G149.08w_2_98	chr1	27717275	740	+	chr1	27716744	740	+	98	806	no
37	baseline	G149.37b_8_82	chr4	45313312	5479	+	chr12	91452904	5479	+	82	708	no
18	baseline	G149.18b_18_38	chr18	64533376	3633	+	chr18	64538528	4459	+	38	700	no
05	baseline	G149.05b_21_78	chr2	133861800	4451	+	chr7	110526066	2515	+	78	563	no
05	baseline	G149.05b_24_62	chr2	131905900	2865	+	chr2	131902406	2564	+	62	536	no
03	baseline	G149.03b_6_183	chr4	106610790	6023	+	chr2	133822409	6023	+	183	319	no
03	baseline	G149.03b_14_41	chr15	58209264	5454	+	chr9	100580827	5454	+	41	193	no
03	baseline	G149.03b_13_42	chr10	190123047	1679	+	chr19	41917526	1679	+	42	81	no
36	baseline	G149.36b_5_258	chr6	28495808	3118	+	chr6	28495814	3118	+	258	3246	yes
30	baseline	G149.30b_31_81	chr10	76460121	3148	+	chr10	76378422	3148	+	81	3220	yes
30	baseline	G149.30b_4_383	chr13	81021129	3158	+	chr13	81021141	3158	+	383	3210	yes
40	baseline	G149.40b_2_251	chr4	102790638	3135	+	chr4	102790667	3135	+	251	3199	yes
36	baseline	G149.36b_2_662	chr6	98308243	6326	-	chr13	64517465	6326	-	662	3187	yes
35	week_96	G149.35w_6_56	chr5	87225234	3129	+	chr5	87225265	3129	+	56	3171	yes
09	baseline	G149.09b_5_227	chr8	72950119	3153	+	chr8	72950155	3153	+	227	3167	yes
30	baseline	G149.30b_9_281	chr2	140527979	3148	+	chr2	140527961	3148	+	281	3147	yes
35	week_96	G149.35w_11_25	chr3	93470501	3128	+	chr3	93470721	3128	+	25	3125	yes
30	baseline	G149.30b_1_493	chr15	80761591	3146	+	chr15	80761591	3146	+	493	3103	yes
30	baseline	G149.30b_2_446	chr12	91483509	3144	+	chr8	35361927	3144	+	446	3079	yes
30	baseline	G149.30b_29_102	chr6	138184361	3144	+	chr6	138184332	3144	+	102	3000	yes
19	baseline	G149.19b_6_62	chr2	57614481	3135	+	chr2	57614449	3135	+	62	2950	yes
30	baseline	G149.30b_24_146	chr2	173274069	3140	+	chr10	25998476	3140	+	146	2894	yes
30	baseline	G149.30b_6_313	chr21	22161356	300	+	chr21	22161321	300	+	313	2843	yes
33	baseline	G149.33s_10_81	chr3	146226370	347	+	chr3	146226370	347	+	81	2795	yes
40	baseline	G149.40b_5_35	chr10	38141077	432	+	chr10	38141080	432	+	35	2709	yes
36	baseline	G149.36b_8_62	chr8	76870024	3155	+	chr8	76870049	3155	+	62	2547	yes
11	week_96	G149.11w_2_134	chr9	96478424	3149	+	chr9	96478419	3149	+	134	2430	yes

06	baseline	G149.06b_5_228	chr14	88063561	752	+	chr14	88063572	752	+	228	2390	yes
33	baseline	G149.33s_4_423	chr4	2437317	774	+	chr4	2437319	774	+	423	2346	yes
03	baseline	G149.03b_1_818	chr18	70081843	818	+	chr18	70081799	818	+	818	2303	yes
15	baseline	G149.15b_1_266	chr3	75819257	6353	-	chr3	75824810	846	+	266	2287	yes
04	baseline	G149.04b_3_53	chr5	123183957	3164	+	chr5	123223038	1338	-	53	2236	yes
06	baseline	G149.06b_8_157	chr16	10383788	2334	+	chr16	10383705	4065	+	157	2219	yes
06	baseline	G149.06b_13_81	chr15	42302293	1023	+	chr15	42236594	1023	+	81	2113	yes
17	baseline	G149.17b_11_94	chr3	61251902	2248	+	chr3	62169081	1098	+	94	2028	yes
19	baseline	G149.19b_2_101	chr1	30996821	4301	-	chr1	30996831	998	+	101	1605	yes
13	week_96	G149.13w_6_74	chr17	33269621	3134	+	chr17	33269732	3134	+	74	1567	yes
07	baseline	G149.07b_23_40	chr13	106052772	1665	+	chrY	56763004	1665	+	40	1468	yes
19	baseline	G149.19b_12_26	chr8	50734875	1959	+	chr8	136770787	1959	+	26	1183	yes
18	baseline	G149.18b_8_168	chr8	24615941	2848	+	chr11	60999863	4717	+	168	645	yes
08	week_96	G149.08w_8_35	chr2	135809986	3133	+	chr2	135809971	1803	+	35	630	yes

Table S3: Summary of high confidence consensus sequences. Integrations with at least 25 reads and both HBV-host junctions were used to create consensus sequences. The top contigs are modeled in Figure 2. HBV positions are based on a 2X-length HBV reference sequence that begins with the start codon of Core. Contigs are classified as "DR1_associated" if either junction if HBV is within 20 bp of DR1.

Supplementary reference

[1]   van Buuren N, R. Ramirez, S. Turner, D. Chen, V. Suri, A. Aggarwal, C. Moon, S. Kim, D. Kornyeyev, N. Bui, N. Bhardwaj, H.L.Y. Chan, P. Marcellin, M. Buti, J. Wallin, A. Gaggar, S.P. Fletcher, L. Diehl, L. Li, H. Mo and B. Feierbach. Characterization of the Liver Immune Microenvironment in Chronic HBV Infected Patient Liver Biopsies. JHep Reports 2021;0.