GENETIC VARIATIONS ON THE HEPATITIS B VIRUS ENV GENE AND THEIR ASSOCIATION WITH THE EVOLUTION TO CIRROSIS AND HEPATOCELLULAR CARCINOMA

Laura M. W. Reis (DO)[1], Juliana D. Siqueira[1], João Marcello Neto[1,2], Henrique Sérgio M. Coelho[2], Esmeralda A. J. M. Soares[1], Marcelo Alves Soares[1,3]

[1] Programa de Oncovirologia, Instituto Nacional de Câncer, Rio de Janeiro, RJ, Brazil; [2]Hospital Universitário Clementino Fraga Filho, Universidade Federal do Rio de Janeiro (HUCFF-UFRJ), Rio de Janeiro, RJ, Brazil; [3]Departamento de Genética, Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, RJ, Brazil

INTRODUCTION

Hepatitis B virus (HBV) infection is a pandemic that affects two billion people worldwide and about 350 million of them are chronic carriers. Up to 80% of cases of liver cancer are caused by chronic HBV infection. The HBV genome has four ORFs (pre-S/S, pre-C/C, X and Pol) and is approximately 3.2 kb in length (Figure 1). We have previously reported deletions in the pre-S/S region in minority HBV populations circulating in newly diagnosed individuals through cloning. Other studies have linked such deletions to a worse prognosis of liver disease.

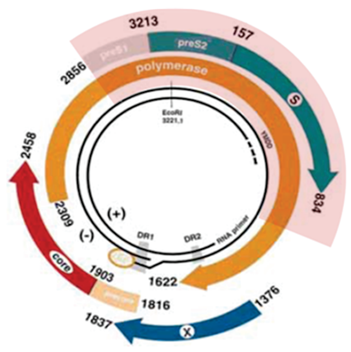

Figure 1. HBV genome organization. The highlighted region corresponds to the region analyzed in this study. (adapted from Jalali et al., 2006).

METHODOLOGY

This is a prospective project study that attempts to associate the presence of pre-S/S deletions with the evolution of liver disease. Patients with chronic HBV infection in different fibrosis stages and with hepatocellular carcinoma followed up at the HUCFF-UFRJ and at the Santa Casa from Joinville (SC) participated in the study. The project was approved by the Ethics Committee of INCA, HUCFF and UNIVILLE. A blood sample was collected and the serum was separated and stored in the -80ºC freezer. Total DNA was extracted from the serum and the pre-S/S ORF was amplified using specific primers. Amplicon libraries were prepared with Nextera XT kit and large-scale sequencing was performed in MiSeq. Geneious was used to assemble the reads and generate a consensus sequence for each patient, that was aligned with references from different HBV genotypes. Phylogenetic analyses were performed with MEGA6. Vaccine escape mutations were evaluated manually using Geneious. Eight different algorithms were tested to find virus deletions (Shorah, QuRE, Pindel, Subread, Samtools, SOAPdenovo, Def-GPU and Geneious R9).

RESULTS

The demographic information of the patients from HUCFF-UFRJ are shown in Table 1. Up to now, 33 HBV samples from HUCFF were collected, and the HBV amplification was successful in 17 cases. From the 14 samples received from Joinville, four were successful in HBV amplification.

Table 1. . Demografic information of patients from HUCFF-UFRJ.

Code	Metavir	Sex	Age	HBV DNA (UI/mL)	HBeAg	Anti-HBe	HBsAg	Anti-HBs	Point B	Point C	AST	ALT	Outros
HU01	F0	F	60	335	N	P	P	N			30	48	GRAVES DISEASE
HU02	F4	M	53	16	N	P	P	N	---	---	75	73	TREATMENT
HU03	F1	F	33	650.089.261	P	N	P	N			31	50	
HU04	F2	M	35	315	N	P	N	P			29	53	
HU05	F3	M	53	IND	N	P	P	N	---	---	---	---	TREATMENT
HU06	F0	F	38	5.154	N	P	P	N	✔	✔	18	17	
HU07	F1	M	56	70	N	P	P	N			24	35	
HU08	F0	F	24	2.609	N	P	P	N	✔		10	11	
HU09	F1	F	47	<10	N	P	N	P	✔	✔	14	8	
HU10	F4	F	58	11	N	P	P	N			24	32	
HU11	F0	F	32	757	P	N	P	N	✔		28	50	
HU12	F0	M	32	1.260	P	N	P	N			13	19	
HU13	F0	M	58	89	N	P	P	N			26	50	
HU14	F2	F	57	388	N	P	P	N			20	27	
HU15	F1	M	62	<10	N	P	N	P			17	26	
HU16	F1	F	42	IND	N	P	P	N			30	40	
HU17	F1	F	47	10.160	N	P	P	N			21	31	
HU18	F0	F	37	245	N	P	P	N	✔	✔	22	40	
HU19	F3	M	42	169	N	P	P	N			27	57	
HU20	F1	F	46	1294	N	P	P	N			24	40	
HU21	F2	M	77	4.793	N	P	P	N	---	---	388	546	TREATMENT
HU22	F0	F	32	17.940	N	P	P	N	✔		19	34	
HU23	F0	F	33	236	---	---	---	---	✔		---	---	
HU24	F0	M	63	404	---	---	---	---			17	13	
HU25	F0	F	40	3.823	N	P	P	N			25	31	
HU26	F1	F	56	149	N	P	N	P	✔	✔	24	38	
HU27	F4	M	38	-----	---	---	P	N			37	78	
HU28	F1	M	30	18.079	N	P	P	N			34	43	
HU29	F0	M	60	589	N	P	P	N			27	23	
HU30	F0	F	61	2.013	N	P	P	N			20	17	
HU31	F0	F	61	200	---	---	P	N			20	26	
HU32	F0	F	36	8.590	P	N	P	N			21	18	
HU34	F1	M	66	<10	N	P	N	P	---	---	14	30	

From the 33 patients at HUCFF-UFRJ, 54% were women and the mean age was 47 years (minimum of 24 and maximum of 77 years). Three patients have high levels of TGO and/or TGP (HU02, HU21 and HU27). HU02 patient had high AST and ALT and have clinical cirrhosis, but the HBV DNA was low. Five patients (HU04, HU09, HU15, HU26 and HU34) were HBsAg- negative and positive for anti-HBs, which indicates a previous infection, and all had low HBV DNA. Only four patients (HU03, HU11, HU12 e HU32) were HBeAg-positive, which indicates chronic infection. This can happen when the circulating HBV carry a pre-C mutation which suppresses the production of the HBeAg protein.

Four patients (HU23, HU24, HU27 and HU31) did not have any serology data on the medical records, and patient HU27 did not have data on HBV DNA levels.

Six viruses have been sequenced so far; three sequences belonged to genotype D, one to A and two to F (Figure 2). In two samples, two vaccinal escape mutants were found, one in each sample. No antiviral resistance mutations were found. Using reads generated in silico from a viral population with deletions in low proportions, Geneious was the unique program capable of identifying deletions. Using the same parameters in Geneious, we tested our six sequenced samples and no deletions were found so far. Also using Geneious, we were able to identify the diversity of the pre-S/S region and the Pol region at the quasispecies level.

Support

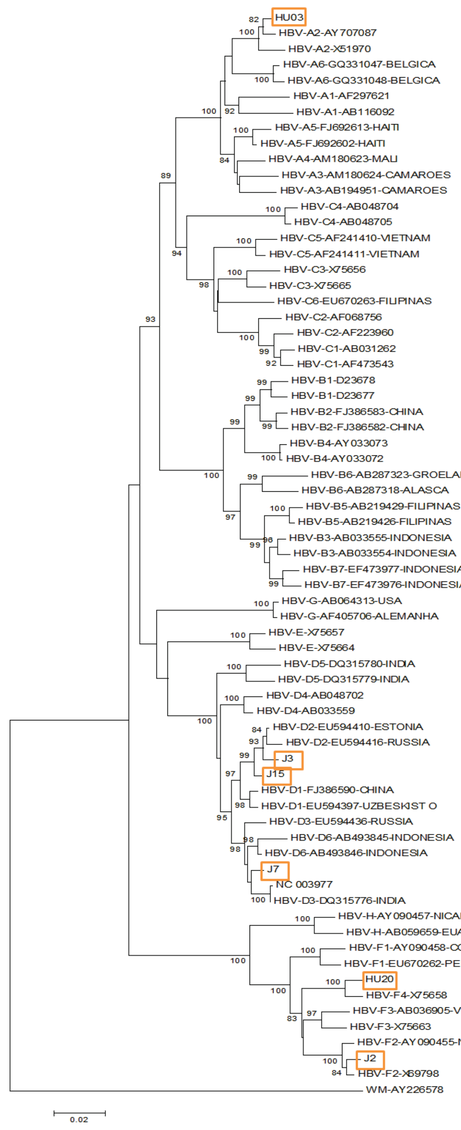

Figure 2. Phylogenetic neigbor-joining tree of the HBV pre-S/S sequences studies. The samples from HUCFF-UFRJ are depicted with HU initials, and those from Joinville begins with J. All sequences generated in the study are boxed. Reference sequences have their GenBank accession number, while sample WM-AY226578 was used as outgroup.

The F141L e W4P/R mutations, which are related to the worse prognosis or severity of liver disease, were also analyzed but were not found among our samples.

We were also able to identify the synonymous and non-synonymous changes in the pre-S/S region of the HBV virus quasispecies, and those above 2% in frequency are shown (Table 2).

Table 2. Diversity of the pre-S/S region of the HBV showing synonymous and nonsynonymous changes.**and non-sinonimous changes.**

Sample	Protein codon	Sequence	Variant	Frequency (%)	Coverage	aa change
J2	L28	A	G	2.6	3.485	S - G
	L35	A	G	2.1	3.674	N -D
	L301	T	C	5.5	4.019	C - R
	L303	T	C	4.4	3.958	C - R
	L310 (S147)	T	C	4.7	3.932	C - R
J3	L44	G	A	5.1	4.509	A - T
	L64	T	C	4.2	5.139	L- S
	L74	A	G	2.6	5.515	I - V
	L97	C	G	47.1	6.244	
	L245	A	C	48.8	7.652	
	L278	C	T	10.5	7.419	T - I
	L373	T	C	2.6	2.911	
J7	L11	T	C	5.4	3.278	L- P
	L16	T	C	39.2	3.512	
	L19	T	C	9.3	3.759	L- S
	L26	T	C	39.6	4.091	
	L59	G	T	35.8	5.003	
	L65	T	G	35.9	5.201	
	L68	C	T	33.9	5.090	
	L84	C	T	34.7	5.307	
	L87	T	G	30.1	5.005	
	L91	G	A	33.9	5.427	G - R
	L95	T	C	37.2	5.373	
	L101	T	C	39.0	5.809	
	L103	A	G	37.0	5.884	N - D
	L145	C	T	35.7	6.535	
	L149	T	C	36.6	7.185	
	L150	C	T	36.4	7.161	S - L
	L155	T	C	34.6	6.122	L- S
	L157	C	T	34.4	6.230	T - I
	L188	T	C	2.5	5.723	I -T
	L191	A	G	33.7	5.912	I - M
	L208	CG	AC	35.2	6.456 - 6.461	T - N
	L230	A	G	34.0	5.982	
	L245	A	C	32.3	6.104	
	L255	T	C	31.9	5.939	I - T
	L257	T	C	33.5	5.834	L- S
	L273	C	T	30.7	5.645	
	L346 (S183)	G	T	29.5	4.899	C - F
	L349	TC	CT	29.2 - 29.3	4.878 - 4.896	S - L
	L351	T	C	28.3	4.460	L- P
	L353	C	T	27.0	4.337	A - V
	L365	C	G	27.6	3.974	A - G
	L369	AA	TG	24.6 - 24.7	3.973 - 3.982	N - C
	L370	A	T	25.2	3.960	N -I
	L371	T	C	25.7	3.951	I - T
	L376	T	C	26.9	4.065	L- S
	L377	C	A	26.9	3.999	P - Q
	L380	C	T	23.6	3.800	P - L
J15	L27	C	T	10.0	2.921	
	L275	G	A	8.4	7.003	G - R
	L289	A	T	2.3	6.968	T - S
		C	G	18.6	6.993	T - S
HU03	L7	G	A	2.6	1.601	V - I
HU20	L239	T	C	3.0	6.132	

Protein codons represent those of the L (large) or S (small) protein (in parenthesis). Only S protein codons that are related to escape mutation sare depicted.
aa – aminoacid; L – large surface protein; S – small surface protein.

Projeto Gráfico: Setor de Edição e Informação Técnico-Científica / INCA