Genomics & Informatics Vol. 8(1) 1-8, March 2010

Association Analysis of *SERPINB5* Polymorphisms with HBV Clearance and HCC Occurrence in a Korean Population

Jason Y. Kim[1], Tae Joon Park[1], Jin Sol Lee[1], Ji Yong Chun[1], Joon Seol Bae[1], Byung Lae Park[2], Hyun Sub Cheong[2], Hyo-Suk Lee[3], Yoon Jun Kim[3] and Hyoung Doo Shin[1,2]*

[1]Department of Life Science, Sogang University, Seoul 121-742, Korea, [2]Department of Genetic Epidemiology, SNP Genetics, Inc., Seoul 153-801, Korea, [3]Department of Internal Medicine and Liver Research Institute, Seoul National University, Seoul 110-744, Korea

Abstract

Serpin peptidase inhibitor, Clade B (ovalbumin), Member 5 (*SERPINB5*), also known as maspin, is a potent tumor suppressor gene. It has correlations with many tumor cells, from pancreas cancer to breast cancer, so it is possible that it may also affect liver cancer. There has also been a report that *SERPINB12*, a gene placed right next to *SERPINB5*, is expressed in liver. For this study, 32 polymorphisms were identified in *SERPINB5* by direct DNA sequencing, and 11 of them were selected to be tested with a larger scale subjects. The association of the 11 *SERPINB5* polymorphisms with Hepatitis B virus (HBV) clearance, hepatocellular carcinoma (HCC) occurrence and the onset age of HCC were analyzed. There were no significant associations found between 11 *SERPINB5* polymorphisms and HBV clearance. In the case of HCC occurrence, one of the haplotypes (*ht*) showed association with HCC occurrence (OR=2.26, p=0.005, P^{Cor}=0.05), albeit with a low statistical power (40.8%) and haplotype frequency (0.052). Further study with a bigger sample size will be needed to clearly verify the association between *ht5* and HCC occurrence.

Keywords: haplotypes, hepatocellular carcinoma, hepatitis B virus, polymorphism, serine protease inhibitor clade B

Introduction

Hepatitis B virus (HBV) infection is one of the more common diseases worldwide, with more than 360 million chronic carriers (Kane, 1998). It particularly affects Asia, Africa, Southern Europe and South America (Lin & Kao, 2008), and South Korea is no exception to the virus, as it is an endemic area of the Hepatitis B virus. There are several directions for clinical course of HBV infection, from spontaneous recovery to chronic infection, which may further develop to liver cirrhosis and hepatocellular carcinoma. Age at infection is a crucial factor on clinical outcome. Most HBV infection occurs at birth or within 5 years after the birth, and they also become chronic carriers. HBV causes not only Hepatitis B; chronic infection of the virus could also lead to liver cirrhosis (LC) and hepatocellular carcinoma (HCC). It has been noted that HBV carriers have as much as 100 times the risk of developing HCC than non-carriers (Koshy, 1998).

Hepatocellular carcinoma is a primary cancer type of liver. HCC usually develops after HBV infection or development of LC and is one of the most common types of tumors worldwide. It generally affects men more than women, and the age of onset is usually between 30s to 50s (Kumar V, 2003). Pathogenesis of HCC is similar to other types of cancers in that HCC develops when a mutation occurs at cellular machinery which results in faster replication of the cell and/or nullification of apoptosis. Chronic infection of Hepatitis B or C virus aids the development of HCC because the viruses cause the immune system to attack the body' s own liver cell. This cycle of constant damaging and repair can lead to mistakes which may then result in carcinogenesis. Recently, it has been suggested that genetic polymorphisms may be associated with the risk of HCC (Chun *et al*., 2009; Kim *et al*., 2006; Oh *et al*., 2008; Park *et al*., 2006; Park *et al*., 2007; Shin *et al*., 2003; Shin *et al*., 2007).

SERPINB5 is one of serpins and a tumor suppressor gene (Sager *et al*., 1997), also commonly known as maspin. Serpins are sets of proteins that are able to inhibit proteases and affect proteolytic cascades. However, *SERPINB5* is a special case, which acts as a tumor suppressor gene instead of affecting protein pathways. *SERPINB5* has previously shown the ability to suppress tumor cells in places such as bones, pancreas, and esophagus (Cai *et al*., 2009; Hall *et al*., 2008; Hong *et al*., 2009). From these findings, it can be inferred that *SERPINB5* has a potent ability to suppress tumor cells. It has also been reported that *SERPINB12*, a member of SERPIN family and a gene that is resided right next to *SERPINB5*, is expressed in liver (Askew *et al*., 2001). *SEPRINB12* also affects many tissues of body, including

*Corresponding author: E-mail hdshin@sogang.ac.kr
Tel +82-2-705-8615, Fax +82-2-2026-4299
Accepted 22 February 2010

pancreas and bone marrows, which are same parts that are affected by *SERPINB5*. Therefore, we can also infer that *SERPINB5* may affect liver as well. As mentioned above, genetic change may be a factor of HCC carcinogenesis, so we hypothesized that genetic polymorphisms of *SERPINB5* could influence HBV clearance or HCC development of HBV infected patients.

Methods

Study subjects

A total of 1,074 Korean subjects having either present or past evidence of HBV infection were prospectively enrolled from the outpatient clinic of the liver unit or from the Center for Health Promotion of Seoul National University Hospital between January 2001 and August 2003. Subjects were placed in two different groups according to serologic markers: the chronic carrier (CC) group, or the spontaneous recovery (SR) group. The CC and SR cohorts consisted of 639 and 435 subjects, respectively (Table 1). The diagnoses of the CC and SR subjects were established by repeated seropositivity for the hepatitis B surface antigen (HBsAg) (Enzygnost® HBsAg 5.0; Dade Behring, Marburg, Germany) over a 6-month period, and for both anti-HBs (Enzygnost® Anti-HBs II; Dade Behring, Marburg, Germany) and anti-HBc (AB-Corek; DiaSorin s.r.l., Saluggia, Italy) of the IgG type without HBsAg, respectively. Asymptomatic HBV carriers were also included in CC group. These patients usually have inactive liver disease on liver biopsy. However, it has been known that in these patients, HBV continues to replicate, albeit very low levels, some patients have residual liver disease, and HCC develop frequently. We excluded subjects who were positive for anti-HBs but not for anti-HBc, and those positive for anti-HCV or anti-HIV (GENEDIA®; Greencross Life Science Corp., Yongin-shi, Korea, HCV®3.2; Dong-A Pharmaceutical Co., Seoul, Korea). The patients who had any other types of liver disease such as autoimmune hepatitis, toxic hepatitis, primary biliary cirrhosis, or Budd-Chiari syndrome were also excluded. No patients had a previous history of immunosuppression or anti-viral treatment. Informed consent was obtained from each patient, and the Institutional Review Board of Human Research at Seoul National University Hospital approved the study protocol. The clinical parameters are summarized in Table 1.

Sequence analysis of the human *SERPINB5*

In order to discover genetic variants, we have sequenced exons and their flanking regions, including the promoter region (1.4 kb) in 24 Korean unrelated individual DNA samples using the ABI PRISM 3700 DNA analyzer (Applied Biosystems, Foster City, CA). Primer sets for the amplification and sequencing analysis of *SERPINB5* were designed based on GenBank sequences (Ref. Genome seq.; NC_000018.9). Sequence analysis was carried out using SeqMan® software.

Genotyping with fluorescence polarization detection

For genotyping of eleven polymorphic sites, amplifying primers and probes were designed for TaqMan® (An *et al.*, 2002). Primer Express (Applied Biosystems) was used to design both the PCR primers and the MGB TaqMan probes. One allelic probe was labeled with the FAM dye and the other with the fluorescent VIC dye. These probes are listed in Supplement Table 1. Typically, PCR was run in the TaqMan Universal Master

Table 1. Clinical profile of the study subjects

Clinical profile	SR	CC		CC	Total
		CH or LC	HCC		
No. of subjects	435	332	307	639	1,074
Age (mean (range))	54.5 (22～79)	49.6 (22～81)	58.2 (25～79)	53.8 (22～81)	54.1 (22～81)
Sex (male/female)	245/190	271/61	261/46	532/107	777/297
HBeAg (positive rate, %)	0	33.7	20.5	27.4	
HBeAb (positive rate, %)	0	30.7	45.3	37.7	
HBsAg (positive rate, %)	0	100	100	100	
HBsAb (positive rate, %)	99.5	0	0	0	
U albumin (positive rate, %)	0	6.6	13	9.7	
U blood (positive rate, %)	27.8	12	21.5	16.6	

SR, spontaneously recovered; CH, chronic hepatitis; CC, chronic carrier; LC, liver cirrhosis; HCC, hepatocellular carcinoma; U, Urine.

mix without UNG (Applied Biosystems) at primer concentration of 900 nM and TaqMan MGB-probe concentration of 200 nM. The reaction was performed in a 384-well format in a total reaction volume of 5 ul using 20 ng of genomic DNA. The plate was then placed in a thermal cycler (PE 9700, Applied Biosystems) and heated for 2 min at 50°C and for 10 min at 95°C, followed by 40 cycles of 95°C for 15 s and 60°C for 1 min. The TaqMan assay plate was then transferred to a Prism 7900HT instrument (Applied Biosystems) where the fluorescence intensity of each well was read. Fluorescence data files from each plate were analyzed by automated software (SDS 2.1). Primer sequences are listed in Supplement Table 2.

Statistics

We examined Lewontin's D' (|D'|) and LD coefficient r^2 between all pairs of biallelic loci (Hedrick, 1987). Haplotypes of each individual were inferred using the algorithm (PHASE) developed by Stephens *et al*. (Stephens *et al*., 2001), which uses a Bayesian approach incorporating a priori expectations of haplotypic structure from population genetic and coalescent theory. Genetic effects of inferred haplotypes were analyzed in the same way as SNPs. Logistic regression models were used for calculating odds ratios (95% confidential interval) and corresponding p-values controlling for age (continuous value) and sex (male=0, female=1) as covariates. In our analysis of HCC occurrence, LC (LC=1, no LC=0) and HBeAg (negative=0, blank=1, positive=2) were also used as covariates. Cox models were used for calculating relative hazards and P-values controlling for sex, adjusted age (0, <40; 1, 40~60; and 2, >60; classification factor), LC (Liver Cirrhosis) (LC=0, no LC=1; classification factor) and HBeAg (negative=0, positive=1; classification

Table 2. Genotype and allele frequencies of 32 polymorphisms discovered in *SERPINB5*

rs#	Loci	Position	Amino acid change	Genotype (number of subjects)			Total number	MAF*	Heterozygosity	HWE**
rs10048322	*−8523A>G*	Promoter	-	A (21)	AG (3)	G (0)	24	0.021	0.117	0.744
rs11659680	*−8493C>T*	Promoter	-	C (21)	CT (2)	T (0)	23	0.313	0.083	0.827
rs62099999	***−8217A>G***	**Promoter**	**-**	**A (617)**	**AG (375)**	**G (76)**	**1,068**	**0.247**	**0.372**	**0.07**
rs55958273	***−8177C>T***	**Promoter**	**-**	**C (712)**	**CT (320)**	**T (40)**	**1,072**	**0.187**	**0.304**	**0.589**
Novel	*−7888C>T*	Promoter	-	C (23)	CT (1)	T (0)	24	0.5	0.041	0.917
Novel	*−7875A>T*	Promoter	-	A (9)	AT (15)	T (0)	24	0.5	0.43	0.026
rs3744941	*−7609C>T*	Promoter	-	C (6)	CT (12)	T (6)	24	0.083	0.5	1
rs17071138	***−7754T>C***	**Promoter**	**-**	**T (929)**	**CT (130)**	**C (7)**	**1,066**	**0.068**	**0.126**	**0.299**
rs57171297	***−7722A>G***	**Promoter**	**-**	**A (818)**	**AG (217)**	**G (11)**	**1,045**	**0.114**	**0.202**	**0.434**
rs3744940	***−7628C>T***	**Promoter**	**-**	**C (301)**	**CT (507)**	**T (247)**	**1,055**	**0.474**	**0.499**	**0.238**
rs3744942	*−7300A>G*	Intron2	-	A (6)	AG (12)	G (6)	24	0.063	0.5	1
rs2292294	***+36C>T***	**Exon3**	**Ala12Ala**	**C (811)**	**CT (246)**	**T (12)**	**1,069**	**0.126**	**0.221**	**0.162**
rs2292295	*+237C>T*	intron3	-	C (20)	CT (4)	T (0)	24	0.146	0.153	0.656
rs2271254	*+2727C>T*	Intron4	-	C (21)	CT (3)	T (0)	24	0.083	0.117	0.744
rs17071181	*+4838T>C*	Intron4	-	T (17)	CT (7)	C (0)	24	0.5	0.249	0.403
rs12454742	***+4945C>G***	**Exon5**	**Pro111Pro**	**C (328)**	**CG (533)**	**G (210)**	**1,071**	**0.445**	**0.494**	**0.805**
Novel	*+5177C>G*	Intron5	-	C (20)	CG (4)	G (0)	24	0.022	0.153	0.656
rs1858143	*+8374A>G*	Intron5	-	A (5)	AG (13)	G (5)	23	0.348	0.5	0.532
rs2289518	*+8610G>A*	Exon6	Lys170Lys	G (22)	AG (1)	A (0)	23	0.042	0.043	0.915
rs2289520	***+8659G>C***	**Exon6**	**Val187Leu**	**G (657)**	**CG (355)**	**C (54)**	**1,066**	**0.217**	**0.34**	**0.502**
rs2289521	*+8724T>C*	Intron6	-	T (10)	CT (10)	C (3)	23	0.438	0.454	0.842
Novel	*+15092C>T*	Intron7	-	C (22)	CT (2)	T (0)	24	0.417	0.08	0.831
rs1840557	*+18806A>G*	Intron7	-	A (7)	AG (13)	G (4)	24	0.438	0.492	0.622
rs1455557	*+18839C>T*	Intron7	-	C (8)	CT (12)	T (4)	24	0.438	0.486	0.889
rs1455556	***+19060C>T***	**Exon8**	**Ser298Ser**	**C (305)**	**CT (506)**	**T (261)**	**1,072**	**0.479**	**0.499**	**0.075**
rs1455555	***+19121A>G***	**Exon8**	**Val319Ile**	**A (354)**	**AG (499)**	**G (218)**	**1,071**	**0.437**	**0.492**	**83**
rs894	*+19936C>A*	Exon8	-	C (7)	AC (13)	A (4)	24	0.438	0.492	0.622
Novel	*+20018T>A*	Exon8	-	T (7)	AT (13)	A (4)	24	0.438	0.492	0.622
rs17071220	*+20242C>T*	Exon8	-	C (7)	CT (13)	T (4)	24	0.438	0.492	0.622
rs11664401	*+20484A>G*	Exon8	-	A (7)	AG (13)	G (4)	24	0.438	0.492	0.622
rs11152386	*+20573A>C*	Exon8	-	A (7)	AC (0)	C (4)	11	0.364	0.463	0.001
rs11542560	***+20621C>T***	**Exon8**	**-**	**C (305)**	**CT (488)**	**T (261)**	**1,054**	**0.479**	**0.499**	**0.019**

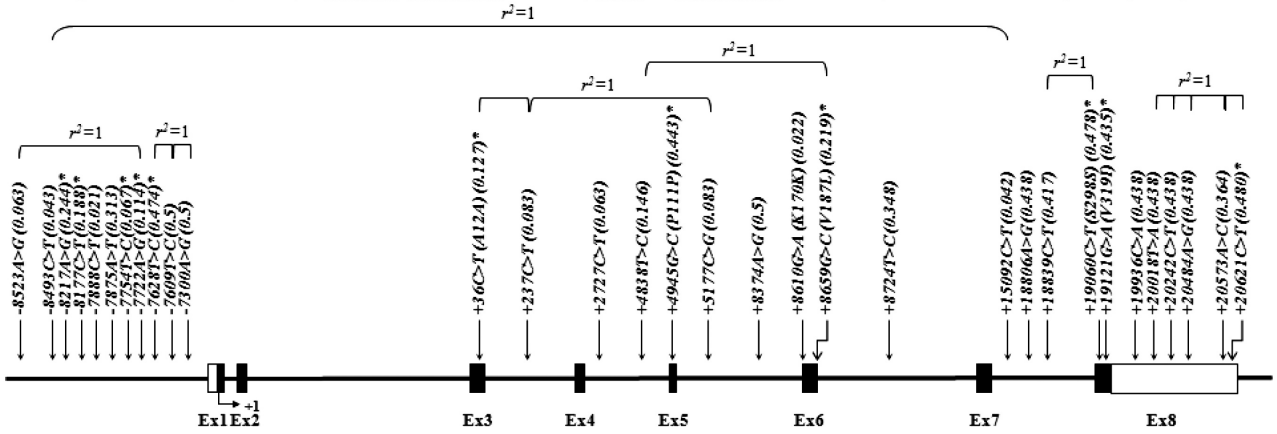
*Minor allele frequency. **p values of deviation from Hardy-Weinberg equilibrium in the Korean population. Bold letters indicate SNPs selected for a larger scale genotyping.

factor). In order to achieve the optimal correction for multiple testing of single-nucleotide polymorphisms (SNPs) in linkage disequilibrium (LD) with each other, the effective number of independent marker loci (9.997) in *SERPINB5* was calculated using the software, on the basis of the spectral decomposition (SpD) of matrices of pair-wise LD between SNPs (http://genepi.qimr.edu.au/general/daleN/SNPSpD/). Statistical power is calculated with PGA matlab application (Menashe *et al*., 2008). PGA is an application specifically designed to calculate statistical power and other values of case-control genetic association studies. For the present study, a co-dominant (1df) model with relative risk 1.3, disease prevalence value 7.1% (Lee *et al*., 1998; Vildozola Gonzales & Salinas, 2009), EDF (Effective Degree of Freedom) 2, and alpha error level 5% were used to calculate the statistical power.

Results

There were 32 observed polymorphisms in *SERPINB5*, and 11 of them were selected to be genotyped in a larger samples (n=1,074) based on their minor allele frequencies (MAF), linkage disequilibriums (LD), and locations. The SNPs selected are *rs62099999, rs55958273, rs17071138, rs57171297, rs3744940* in promoter region, *rs2292294* in Exon 3, *rs12454742* in Exon 5, *rs2289520* in Exon 6, and *rs1455556, rs1455555, rs11542560* in Exon 8 (Fig. 1A and Table 2). Some of the polymorphisms located in exon induced amino acid change, and they were *rs2292294* (alanine at 12 to alanine), *rs12454742* (proline at 111 into proline), *rs2289520* (valine at 187 into leucine), *rs1455556* (serine at 298 into serine), and *rs1455555* (valine at 319 into isoleucine) (Table 2). MAFs of each polymorphism are also shown on Fig. 1A and Table 2, and the statistical powers for

A. Map of *SERPINB5* (serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 5) on chromosome18q21.3 (30 kb)



B. Haplotypes in *SERPINB5*

Hap.	rs62099999	rs55958273	rs17071138	rs57171297	rs3744940	rs2292294	rs12454742	rs2289520	rs1455556	rs1455555	rs11542560	Freq.
ht1	A	C	T	A	T	C	G	G	T	A	T	0.240
ht2	G	C	T	A	C	C	C	G	C	G	C	0.179
ht3	A	C	T	A	T	C	G	G	C	G	C	0.156
ht4	A	C	T	A	C	C	C	G	C	A	C	0.070
ht5	A	T	T	G	C	T	C	C	T	A	T	0.052
ht6	A	T	T	G	C	T	C	T	T	A	T	0.050
Others.	.	.	.	.	.	.	.	.	.	.	.	0.254

C. LDs among *SERPINB5* polymorphisms

	SNPs	rs62099999	rs55958273	rs17071138	rs57171297	rs3744940	rs2292294	rs12454742	rs2289520	rs1455556	rs1455555	rs11542560
|D'|	*rs62099999*	-	1	1	0.878	0.967	0.718	0.920	0.558	0.637	0.694	0.503
	rs55958273	0.075	-	1	0.934	0.992	0.813	0.870	0.850	0.767	0.721	0.683
	rs17071138	0.024	0.314	-	1	1	1	0.728	0.777	0.621	0.560	0.432
	rs57171297	0.033	0.519	0.010	-	1	0.859	0.981	0.850	0.740	0.685	0.727
	rs3744940	0.274	0.207	0.067	0.119	-	0.887	0.908	0.867	0.153	0.054	0.115
r^2	*rs2292294*	0.024	0.419	0.010	0.705	0.104	-	1	1	0.960	0.959	0.939
	rs12454742	0.222	0.139	0.031	0.099	0.727	0.116	-	1	0.202	0.084	0.173
	rs2289520	0.028	0.589	0.154	0.343	0.193	0.517	0.225	-	0.759	0.853	0.662
	rs1455556	0.122	0.147	0.030	0.079	0.023	0.145	0.036	0.176	-	0.998	0.922
	rs1455555	0.205	0.093	0.018	0.048	0.002	0.103	0.004	0.158	0.708	-	0.913
	rs11542560	0.076	0.117	0.015	0.077	0.013	0.139	0.026	0.135	0.844	0.597	-

Fig. 1. Gene maps and haplotypes of the *SERPINB5* Gene. (A) A map of *SERPINB5* on chromosome18q21.3. Coding exons, UTR, and SNPs are shown. Coding exons are marked by shaded blocks while UTRs are marked by white blocks. SNPs marked with asterisks (*) indicate that they were genotyped in the larger population. (B) Haplotypes of *SERPINB5* in the Korean population. Only those with frequencies over 0.05 are shown. (C) LD coefficients (|D'| and r^2) among selected SNPs based on the genotypes of whole study subjects in this study (n=1,074).

Table 3. Association analysis of *SERPINB5* SNPs and haplotypes with clearance of HBV infection, HCC occurrence and the onset age of HCC

rs#	HBV Clearance					HCC Occurrence						Onset age of HCC				
	MAF		OR (95%CI)	p	Statistical power (%)*	MAF		OR (95%CI)	p	P^{Cor}**	Statistical power (%)*	n/event	x^2	p	P^{Cor}**	RH
	CC (n=639)	SR (n=435)				HCC (n=307)	CH/LC (n=332)									
rs62099999	0.256	0.233	1.15 (0.93～1.41)	0.19	92.8	0.27	0.243	1.06 (0.79～1.43)	0.68	1	85.9	635/299	0.49	0.48	1	0.94
rs55958273	0.185	0.189	0.98 (0.78～1.23)	0.86	86.2	0.18	0.19	0.87 (0.62～1.24)	0.44	1	74.9	637/299	0.51	0.48	1	0.93
rs17071138	0.061	0.077	0.77 (0.54～1.09)	0.14	44	0.057	0.064	0.58 (0.34～0.99)	0.05	0.46	32.8	633/298	1.75	0.19	1	0.8
rs57171297	0.115	0.113	1.01 (0.76～1.35)	0.95	70	0.117	0.113	1.30 (0.83～2.05)	0.26	1	58.8	611/284	1.1	0.3	1	1.15
rs3744940	0.464	0.49	0.90 (0.75～1.08)	0.24	96.9	0.448	0.478	0.99 (0.75～1.29)	0.91	1	91.5	622/293	0.94	0.33	1	1.08
rs2292294	0.133	0.116	1.21 (0.91～1.60)	0.2	75.6	0.136	0.131	1.25 (0.83～1.89)	0.29	1	64.7	635/299	0.35	0.56	1	1.08
rs12454742	0.438	0.455	0.91 (0.76～1.10)	0.33	96.9	0.413	0.461	0.89 (0.68～1.16)	0.37	1	91.3	638/299	0.11	0.92	1	1.01
rs2289520	0.218	0.215	1.03 (0.83～1.28)	0.8	90.1	0.216	0.22	0.93 (0.67～1.29)	0.66	1	80.5	632/298	0.06	0.8	1	0.98
rs1455556	0.483	0.475	0.99 (0.83～1.18)	0.9	96.9	0.462	0.502	0.89 (0.69～1.16)	0.4	1	91.6	637/299	0.04	0.85	1	1.02
rs1455555	0.432	0.444	0.99 (0.83～1.18)	0.87	96.8	0.44	0.424	1.00 (0.77～1.31)	0.97	1	91.5	636/300	0.18	0.67	1	0.966
rs11542560	0.482	0.476	0.98 (0.82～1.17)	0.84	96.9	0.459	0.505	0.91 (0.70～1.19)	0.49	1	91.6	622/296	0.18	0.67	1	1.04
ht1	0.236	0.245	0.89 (0.72～1.10)	0.29	91.5	0.215	0.256	0.82 (0.61～1.12)	0.22	1	80.4	639/300	0	0.98	1	1
ht2	0.183	0.172	1.11 (0.88～1.39)	0.39	85.9	0.192	0.178	0.98 (0.70～1.37)	0.9	1	77	639/300	1.06	0.3	1	0.9
ht3	0.159	0.15	1.09 (0.85～1.41)	0.49	81.8	0.164	0.152	1.18 (0.82～1.70)	0.38	1	71.7	639/300	0.63	0.43	1	1.09
ht4	0.07	0.069	1.02 (0.72～1.44)	0.92	49.4	0.085	0.057	1.48 (0.88～2.48)	0.14	1	46.3	639/300	0.28	0.59	1	1.08
ht5	0.056	0.046	1.16 (0.78～1.72)	0.46	40.8	0.073	0.041	2.26 (1.28～3.99)	**0.005**	**0.05**	40.8	639/300	4.38	**0.05**	0.50	1.37
ht6	0.053	0.045	1.23 (0.79～1.90)	0.36	38.9	0.039	0.063	0.63 (0.33～1.23)	0.18	1	23	639/300	1.3	0.25	1	0.78

Logistic regression models were used for calculating odds ratios (95% confidential interval) and corresponding p-values for each SNP sites and haplotypes controlling age, and sex as covariables using SAS. p-values of co-dominant models are also given. Age (continuous value), and sex (male=0, female=1) were adjusted by including in logistic analysis as co-variables. All patients included in study were HBsAg-positive (chronic hepatitis). *Statistical power is calculated with PGA application. **To achieve the optimal correction for multiple testing of single-nucleotide polymorphisms (SNPs) in linkage disequilibrium (LD) with each other, the effective number of independent marker loci (9.997) in SERPINB5 was calculated using the software SNPSpD (http://genepi.qimr.edu.au/general/daleN/SNPSpD/), on the basis of the spectral decomposition (SpD) of matrices of pair-wise LD between SNPs. Bold value indicates the case of P smaller than 0.05. MAF stands for Minor Allele Frequency. Cox models (co-dominant model) were used for calculating relative hazards and p-values for SNPs and haplotypes controlling for sex, adjusted age (age＜40, adjage=0; 40＜=age＜60, adjage=1; age＞60, adjage=2), LC (LC=0, no LC=1), and HBeAg (negative=0, positive=1) by SAS. All patients included in this Table were HBsAg-positive (chronic HBV).

each polymorphism are shown in Table 3. The frequencies of haplotypes are listed in Fig. 1B, and haplotype p-values and association analyses of HBV clearance and HCC occurrence are listed in Table 3. There were six major haplotypes with MAF bigger than 0.05, which accounted for 74.6% of total distribution. LDs among the polymorphisms of *SERPINB5* are listed in Fig. 1C.

Association analyses results of *SERPINB5* for HBV clearance are shown in Table 3. There was no significant association found with SNPs and haplotypes in the case of HBV clearance ($p > 0.05$). The results for HCC occurrence initially had two associations found, with the *rs17071138* (p=0.05, OR=0.58, 95% CI (0.34～0.99)) and one of the haplotypes, *SEPIRNB5 ht5* (*[A-T-T-G-C-T-C-C-C-T-A-T]*) (p=0.005, OR=2.26, 95% CI (1.28～3.99)). The initial association disappeared for *rs17071138* when the p-value was adjusted for the multiple testing (correction value=9.997), but *ht5* retained the association ever after the correction (P^{Cor}=0.05). However, the statistical power and MAF of *ht5* was low for the data to be completely trustworthy (statistical power=40.8%, MAF=0.052).

To analyze the role of *SERPINB5* polymorphisms in the onset age of HCC, Cox relative hazards analysis for age of HCC occurrence was run for CC groups. There were no significant associations observed after correction for the multiple testing (Table 3).

Discussion

SERPINB5 has been associated with prostate cancer, breast cancer (Shao *et al*., 2008), pancreas cancer (Kashima *et al*., 2008; Maass *et al*., 2001) and various types of carcinomas (Bal *et al*., 2008; Blandamura *et al*., 2006; Cai *et al*., 2009). Also, *SERPINB12*, a member of SERPIN family and a gene placed next to *SERPINB5*, is expressed in liver, among various parts of body (Askew *et al*., 2001). The gene has been discovered in relatively recent time (Khalkhali-Ellis, 2006), so there are still unknown facts about *SERPINB5*, which implies that *SERPINB5* may have unidentified function and effects helpful for human population. From this reasoning, the gene was selected to examine whether it has relations with HBV and HCC.

Polymorphisms of *SERPINB5* were screened and 11 variants were selected to test their relations with HBV clearance and HCC occurrence. There were no significant associations found between HBV clearance and the genetic polymorphisms. *SERPINB5* is primarily known as a tumor suppressor gene, so the result is reasonable in that HBV infection is just a precursor of HCC. In the case of HCC occurrence analysis, *ht5* showed some relations in co-dominant model (p=0.005, P^{Cor}=0.05, OR=2.26, 95% CI (1.28～3.99)). By looking at this result only, it would mean that *ht5* has association with HCC occurrence. However, although *ht5* showed association with HCC occurrence, it had a low statistical power (40.8%) and haplotype frequency (0.052), meaning the susceptible population was of a small size and unreliable. Therefore, it is still unsure whether *ht5* really had an association with HCC occurrence or was just a false positive association, thus the association may need to be studied again in the future with bigger sample number, preferably with a higher number of subjects that exhibit *ht5* haplotypes. Although *SEPRINB5* functions as a tumor suppressor gene in other parts of body as mentioned above, association analyses show that it does not play a significant role in the case of HBV clearance or HCC occurrence overall, at least in the genetic polymorphisms we studied.

In summary, 11 SNPs of *SEPRINB5* that were tested did not have any significant relations with HBV clearance. In the case of HCC occurrence, one of the haplotypes, *ht5* showed a minor association with HCC occurrence, but there is a possibility that it might have shown the association due to the small sample size and low frequency value. Therefore, this may need to be further studied in the future with a bigger number of subjects. *SERPINB5* still has potential to be a tumor suppressor gene for other parts of the body, so future research may focus on relations between *SERPINB5* and other types of cancers. However, it seems that *SERPINB5* does not influence development of HCC or have association with HBV clearance, at least in the studied genetic polymorphisms, although this does not rule out *SERPINB5* from the candidate genes of HBV clearance or HCC occurrence altogether, because there are still many unknown aspects of genes, and many current researches are focusing on those aspects, such as copy number variations and methylations of DNA. Therefore, it is premature to say that *SERPINB5* does not influence HBV clearance or HCC occurrence on any level.

Acknowledgements

We greatly acknowledge and thank the participants of this study and their families who took parts in the HBV cohort study by Seoul National University. This work was supported by a grant from Korea Science and Engineering Foundation (KOSEF), funded by the Korea government (MEST) (No. 2009-0080157) as well as the National R&D Program for Cancer Control, Ministry for Health, Welfare and Family affairs, Republic of Korea (0920280).

References

An, P., Nelson, G.W., Wang, L., Donfield, S., Goedert, J.J., Phair, J., Vlahov, D., Buchbinder, S., Farrar, W.L., Modi, W., O'Brien, S.J., and Winkler, C.A. (2002). Modulating influence on HIV/AIDS by interacting RANTES gene variants. *Proc. Natl. Acad. Sci. USA* 99, 10002-10007.

Askew, Y.S., Pak, S.C., Luke, C.J., Askew, D.J., Cataltepe, S., Mills, D.R., Kato, H., Lehoczky, J., Dewar, K., Birren, B., and Silverman, G.A. (2001). SERPINB12 is a novel member of the human ov-serpin family that is widely expressed and inhibits trypsin-like serine proteinases. *J. Biol. Chem.* 276, 49320-19330.

Bal, N., Kocer, N.E., Ertorer, M.E., Canpolat, E.T., and Kayaselcuk, F. (2008). Maspin, E-selectin, and P-selectin expressions in papillary thyroid carcinomas and their correlation with prognostic parameters. *Pathol. Res. Pract.* 204, 743-750.

Blandamura, S., Giacomelli, L., Leo, G., Segato, P., and Ninfo, V. (2006). Nuclear maspin detection in renal cell tumours: possible diagnostic role and correlation with p53 status. *Histopathology* 49, 274-282.

Cai, Z., Zhou, Y., Lei, T., Chiu, J.F., and He, Q.Y. (2009). Mammary serine protease inhibitor inhibits epithelial growth factor-induced epithelial-mesenchymal transition of esophageal carcinoma cells. *Cancer* 115, 36-48.

Chun, J.Y., Park, B.L., Cheong, H.S., Kim, J.Y., Park, T.J., Lee, J.S., Lee, H.S., Kim Y.j., and Shin, H.D. (2009). Identification of polymorphisms in CYP2E1 gene and association analysis among chronic hbv patients. *Genomics Inform.* 7, 187-194.

Hall, D.C., Johnson-Pais, T.L., Grubbs, B., Bernal, R., Leach, R.J., and Padalecki, S.S. (2008). Maspin reduces prostate cancer metastasis to bone. *Urol. Oncol.* 26, 652-658.

Hedrick, P.W. (1987). Gametic disequilibrium measures: proceed with caution. *Genetics* 117, 331-341.

Hong, S.N., Lee, J.K., Choe, W.H., Ha, H.Y., Park, K., Sung, I.K., Lee, K.T., Kim, J.J., and Rhee, J.C. (2009). The effect of aberrant maspin expression on the invasive ability of pancreatic ductal adenocarcinoma cells. *Oncol. Rep.* 21, 425-430.

Kane, M.A. (1998). World-wide epidemiology of hepatitis B. *Soz Praventivmed* 43 Suppl 1:S24-26, S98-100.

Kashima, K., Ohike, N., Mukai, S., Sato, M., Takahashi, M., and Morohoshi, T. (2008). Expression of the tumor suppressor gene maspin and its significance in intraductal papillary mucinous neoplasms of the pancreas. *Hepatobiliary Pancreat Dis. Int.* 7, 86-90.

Khalkhali-Ellis, Z. (2006). Maspin: the new frontier. *Clin. Cancer Res.* 12, 7279-7283.

Kim, Y.J., Yoon, J.H., Kim, C.Y., Kim, L.H., Park, B.L., Shin, H.D., and Lee, H.S. (2006). IGF2 polymorphisms are associated with hepatitis B virus clearance and hepatocellular carcinoma. *Biochem. Biophys. Res. Commun.* 346, 38-44.

Koshy, R. (1998). Hepatitis B Virus. Imperial College Press, London.

Kumar, V.F.N. (2003). Robbins & Cotran Pathologic Basis of Disease. Saunders.

Lee, M.S., Kim, D.H., Kim, H., Lee, H.S., Kim, C.Y., Park, T.S., Yoo, K.Y., Park, B.J., and Ahn, Y.O. (1998). Hepatitis B vaccination and reduced risk of primary liver cancer among male adults: a cohort study in Korea. *Int. J. Epidemiol.* 27, 316-319.

Lin, C.L., and Kao, J.H. (2008). Hepatitis B viral factors and clinical outcomes of chronic hepatitis B. *J. Biomed. Sci.* 15, 137-145.

Maass, N., Hojo, T., Ueding, M., Luttges, J., Kloppel, G., Jonat, W., and Nagasaki, K. (2001). Expression of the tumor suppressor gene Maspin in human pancreatic cancers. *Clin. Cancer Res.* 7, 812-817.

Menashe, I., Rosenberg, P.S., and Chen, B.E. (2008). PGA: power calculator for case-control genetic association analyses. *BMC Genet.* 9, 36.

Oh, A.R., Lee, S.K., Kim, M.H., Cheong, J.Y., Cho, S.W., Yang, K.S., and Kwack, K.B. (2008). Polymorphisms in RAS guanyl-releasing protein 3 are associated with chronic liver disease and hepatocellular carcinoma in a Korean population. *Genomics Inform.* 6, 181-191.

Park, B.L., Kim, Y.J., Cheong, H.S., Kim, L.H., Choi, Y.H., Lee, H.S., and Shin, H.D. (2006). Association of common promoter polymorphisms of MCP1 with hepatitis B virus clearance. *Exp. Mol. Med.* 38, 694-702.

Park, B.L., Kim, Y.J., Cheong, H.S., Lee, S.O., Han, C.S., Yoon, J.H., Park, J.H., Chang, H.S., Park, C.S., Lee, H.S., and Shin, H.D. (2007). HDAC10 promoter polymorphism associated with development of HCC among chronic HBV patients. *Biochem. Biophys. Res. Commun.* 363, 776-781.

Sager, R., Sheng, S., Pemberton, P., and Hendrix, M.J. (1997). Maspin. A tumor suppressing serpin. *Adv. Exp. Med. Biol.* 425, 77-88.

Shao, L.J., Shi, H.Y., Ayala, G., Rowley, D., and Zhang, M. (2008). Haploinsufficiency of the maspin tumor suppressor gene leads to hyperplastic lesions in prostate. *Cancer Res.* 68, 5143-5151.

Shin, H.D., Park, B.L., Cheong, H.S., Yoon, J.H., Kim, Y.J., and Lee, H.S. (2007). SPP1 polymorphisms associated with HBV clearance and HCC occurrence. *Int. J. Epidemiol.* 36, 1001-1008.

Shin, H.D., Park, B.L., Kim, L.H., Jung, J.H., Kim, J.Y., Yoon, J.H., Kim, Y.J., and Lee, H.S. (2003). Interleukin 10 haplotype associated with increased risk of hepatocellular carcinoma. *Hum. Mol. Genet.* 12, 901-906.

Stephens, M., Smith, N.J., and Donnelly, P. (2001). A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* 68, 978-989.

Vildozola Gonzales, H., and Salinas, J.L. (2009). Natural history of chronic hepatitis B virus infection. *Rev. Gastroenterol Peru* 29, 147-157.

Supplementary Table 1. Sequences of amplifying, Taqman probe, and extension primer for *SERPINB5* SNP genotyping

rs#	Sequence			
	Forward	Reverse	Probe-1 (VIC)	Probe-2 (FAM)
rs62099999	tcgtctcagctttccaagtaactg	ggtgagaccccatctctacaaaaa	catcaccacactaggct	atcaccacgctaggct
rs55958273	gcatcaccacactaggctattgttt	ccaggagtttgagaccaacct	aacatggtgagaccc	acatggtaagaccc
rs17071138	gtgtctgagaaatttgtagtgttactatcatca	cggcccacgcagga	aaggtgaaatattcg	aaggtgaagtattcg
rs5717297	ccttccggtcctgcgt	gctgtgagttacatgcatacgtaca	ccgagaggattgccgt	cgagagggttgccgt
rs3744940	cttttggaagctgtgcagaca	gcagcttgtccacaattgaaagaaa	tcaggctgaagttactgt	aggctgaagttgctgt
rs2292294	gccctgcaactagcaaattcg	agtggctccttttcacatagttgtt	acagatcaacggcaaaa	cagatcaacagcaaaa
rs12454742	tcaggagttcatcagctctacga	ccaatttatctttgaagtcaacagtttcca	ctttgcatacggtctct	tttgcatagggtctct
rs2289520	ggcaagtggatgaagaaattttctga	ggctccttttactgctacatgct	catacctttgttgactctgaa	atacctttgttgagtctgaa
rs1455556	gtctggaaaatctagggctgaaaca	ctcccttggtctctgacattcc	agatgtgtcttcgctgaa	agatgtgtcttcactgaa
rs1455555	gtcagagaccaagggagtgg	ggaatccccaccatcttcagttatt	ccctatcaaatgttgtccaca	cctatcaaatgttatccaca
rs11542560	cctgcattgtaaataggttcttcttgttc	acggcttggttctatttcaaaatagtct	cataggagaaattca	cataggaaaaattca

Supplementary Table 2. Primer sequences for *SERPINB5* SNP screening

	Forward	Reverse
Frag-1	tttgctcaggatttctcaagc	tcttcgtggagcctgttctt
Frag-2	tgaatgaagtttggcacttacc	cctgtaatcccagcatttgg
Frag-3	gtaactgggaccacaggcat	gttgtctgcacagcttccaa
Frag-4	gcagctacctctcttggcat	taccccaccttgcttacctg
Frag-5	tgtaactcacagcccccttcc	cagtgttcctggtgctcaaa
Frag-6	ttcagacatttgcctggaag	acaacaatcgcttgaacctg
Frag-7	caaagactcatccagcctga	tgtaagacagctcattgtctatgg
Frag-8	cccataagagaattgggcat	gctcgtcagagagcaagaca
Frag-9	ttctcttcctttcttccctgc	tagctttcctcgctgggtta
Frag-10	tggatctgagtggatgatgaa	agcacctgccttctccttac
Frag-11	cagatgatgaacatggaggc	agctccctcactggtaaacaa
Frag-12	cactattacactccagcccg	tatggaatccccaccatctt
Frag-13	aagatgattgatcccaaggc	aactaccccggacaatttga
Frag-14	tggatgccgatttctgtaaa	cccagatctgatccatgaag
Frag-15	attctcgcttccctgaaaga	cggctattgtgagaatccct
Frag-16	gaaatgcaagaccccaagag	cacggcttggttctatttca
Frag-17	ttgtgacattccttctccca	agtcagagctgagggctagg