

ARTICLE

Open Access

# Identification of strong intron enhancer in the heparanase gene: effect of functional rs4693608 variant on HPSE enhancer activity in hematological and solid malignancies

Olga Ostrovsky<sup>1</sup>, Ania Hava Grushchenko-Polaq<sup>1</sup>, Katia Beider<sup>1</sup>, Margarita Mayorov<sup>1</sup>, Jonathan Canaani<sup>1</sup>, Avichai Shimoni<sup>1</sup>, Israel Vlodavsky<sup>2</sup> and Arnon Nagler<sup>1</sup>

## Abstract

Heparanase is an endo- $\beta$ -glucuronidase that specifically cleaves the saccharide chains of heparan sulfate (HS) proteoglycans and releases HS-bound cytokines, chemokines, and bioactive growth-promoting factors. Heparanase plays an important role in the nucleus as part of an active chromatin complex. Our previous studies revealed that rs4693608 correlates with heparanase levels and increased risk of acute and extensive chronic graft vs. host disease (GVHD). Discrepancy between recipient and donor in this SNP significantly affected the risk of acute GVHD. In the present study, we analyzed the HPSE gene region, including rs4693608, and demonstrated that this region exhibits SNPs-dependent enhancer activity. Analysis of nuclear proteins from normal leukocytes revealed their binding to DNA probe of both alleles with higher affinity to allele G. All malignant cell lines and leukemia samples disclosed a shift of the main bands in comparison to normal leukocytes. At least five additional shifted bands were bound to allele A while allele G probe was bound to only one main DNA/protein complex. Additional SNPs rs4693083, rs4693084, and rs4693609 were found in strong linkage disequilibrium (LD) with rs11099592 (exon 7). Only rs4693084 affected protein binding to DNA in cell lines and leukemia samples. As a result of the short distance between rs4693608 and rs4693084, both SNPs may be included in a common DNA/protein complex. DNA pull-down assay revealed that heparanase is involved in self-regulation by negative feedback in rs4693608-dependent manner. During carcinogenesis, heparanase self-regulation is discontinued and the helicase-like transcription factor begins to regulate this enhancer region. Altogether, our study elucidates conceivable mechanism(s) by which rs4693608 SNP regulates HPSE gene expression and the associated disease outcome.

## Introduction

Heparanase is an endo- $\beta$ -glucuronidase that specifically cleaves the saccharide chains of heparan sulfate (HS)

proteoglycans (HSPG), important components of the cell surface, basement membrane, and extracellular matrix (ECM). Cleavage of HS by active heparanase leads to loss of integrity of the basement membrane and ECM and release of HS-bound cytokines, chemokines, and growth-promoting factors<sup>1–6</sup>.

Heparanase plays critical roles in tumor metastasis and angiogenesis<sup>5</sup>. High expression of heparanase is

Correspondence: Olga Ostrovsky ([olga.ostrovsky@sheba.health.gov.il](mailto:olga.ostrovsky@sheba.health.gov.il))

<sup>1</sup>Department of Hematology and Bone Marrow Transplantation, Chaim Sheba Medical Center, Tel-Hashomer, Israel

<sup>2</sup>Cancer and Vascular Biology Research Center, Rappaport Faculty of Medicine, Technion, Haifa, Israel

© The Author(s) 2018



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

frequently observed in an increasing number of primary human tumors of all etiologies (carcinoma, sarcoma, and hematological malignancies), correlating with high vessel density and poor clinical outcome<sup>7–10</sup>. The enzyme is also involved in inflammation<sup>11</sup>, fibrosis<sup>12</sup>, diabetes<sup>13</sup>, and kidney dysfunction<sup>14</sup>.

Heparanase can trigger the phosphorylation of protein kinases, such as p38, Erk, and Akt, a process mediated by the C-terminal domain of the protein<sup>5,15</sup>. Moreover, heparanase plays a role in the nucleus as part of an active chromatin complex which regulates inducible gene transcription. Nuclear heparanase associates with euchromatin and regulates histone H3K4 and H3K9 methylation by binding to target gene regulatory regions in association with the demethylase LSD1<sup>16</sup>.

A number of functional single-nucleotide polymorphisms (SNPs) were identified in the HPSE gene<sup>17–19</sup>, rs4693608 polymorphism being the most prominent<sup>19,20</sup>. Combinations of rs4693608 and rs4364254 were found in significant correlation with heparanase levels in normal leukocytes and both before and after pre-transplantation conditioning<sup>19,21</sup>. In addition, a highly significant association was identified between rs4693608 and the risk of acute and extensive chronic graft vs. host disease (GVHD). Moreover, discrepancy between recipient and donor in this SNP significantly affected the risk of GVHD<sup>20</sup>. A positive association was found between recipients and donors rs4693608 discrepancy and the recovery time of neutrophils and platelets<sup>21</sup>. rs4693608 also affects HPSE gene expression in lipopolysaccharide (LPS)-treated mononuclear cells (MNC) from peripheral and cord blood<sup>21</sup>.

Recently, rs4693608 and rs4364254 were found to be a genetic risk factor for sinusoid obstruction syndrome/veno-occlusive disease<sup>22</sup>. These polymorphisms were also associated with different levels of HPSE mRNA within carotid plaque tissue<sup>23</sup>. Functional HPSE SNPs were found to play a role in gastric<sup>24,25</sup> and ovarian<sup>26</sup> cancer progression and patient survival. In addition, rs4693608 and rs6535455 correlate with bone disease and the outcome of multiple myeloma patients<sup>27</sup>.

In the present study, we analyzed 440 bp of the HPSE gene (intron 2) (chr4: 84,241,177–84,242,376) that include the most prominent rs4693608 SNP. Our results indicate that this region exhibits enhancer activity in both the sense and antisense directions. A–G alteration leads to a different enhancer activity as a result of a unique DNA–protein complex, which binds to DNA regulatory elements in the enhancer. The DNA–protein complex that binds to the G allele shifted similarly in all the examined cell lines and primary leukemic samples, but not in normal white blood cells. At least five different variations of DNA–protein complexes bound to allele A were identified. DNA pull-down assay of normal cell

samples revealed that heparanase is involved in self-regulation of the HPSE gene by negative feedback in rs4693608-dependent manner. During carcinogenesis, heparanase self-regulation is attenuated and the helicase-like transcription factor begins to regulate this enhancer region. Altogether, the present study reveals points to a possible mechanism by which rs4693608 affects HPSE gene expression and thereby dictates disease outcome.

## Results

### Characterization of SNPs in the cloned enhancer region

According to the UCSC Genome Browser website, the region that contains rs4693608 SNP is a putative strong HPSE enhancer. Detailed analysis of the enhancer region revealed additional three polymorphic SNPs, which were mapped in this area (Fig. 1). As shown in Fig. 1, the distance between rs4693083 and rs4693608 is 40 bp, whereas the interval between rs4693608 and rs4693084 is 17 bp. Additional rs4693609 SNP is located 199 bp upstream the rs4693084 (Fig. 1). Genotyping of the SNPs was performed among 148 healthy individuals and the results are summarized in Table 1. The SNP genotype frequencies did not significantly deviate from the Hardy–Weinberg distribution (data not shown). rs4693083, rs4693084, and rs4693609 were found in strong linkage disequilibrium (LD). Only three out of the 148 healthy individuals had a different genotype for rs4693609 SNP in comparison to rs4693083 and rs4693084 polymorphisms. However, all three SNPs exhibited a partial LD with the functional rs4693608 SNP. In addition, we found that rs11099592, located in exon 7, is in strong LD with each of the three SNPs (Table 1).

Our previous study identified the presence of two blocks of SNPs in the HPSE gene<sup>16</sup>. The first block includes polymorphisms which were mapped between intron 1 and

```

5'CTGGGCGACAAAGCGAGACACTGTCTAAAAAAGCTGGAAAT
TACCAGAGAACTCTAATGAAGAAAACATAGTGCAGTCTCTTAAGGTTCAAAATC
CATAGCAATTTGTTTACATGGAGTGTATCTGACAGTGGTGTITTTCTCTTCCATCATGG
GAGGTAGGATGGGTGGGAACCTTGACTTTTTTCGCTTGAATTTGGTACCATCAGAGC
TTCTGGGACCACAACATGTGCACCTTGAAAAATCCACCTGGTCAGGCACAGTGGCTCAC
GCTTGAATTCAGCTCTTTGGGAGGCCAAGCGAGATGGATAGCTTGAGCTCAGGAGTTT
GAGACTAGCCTGGACAACATAGTAAACCTGCCTCTACTAAAAATACGAAAACCTAGCTG
GGCGTGGTGGTGTGCCT 3'

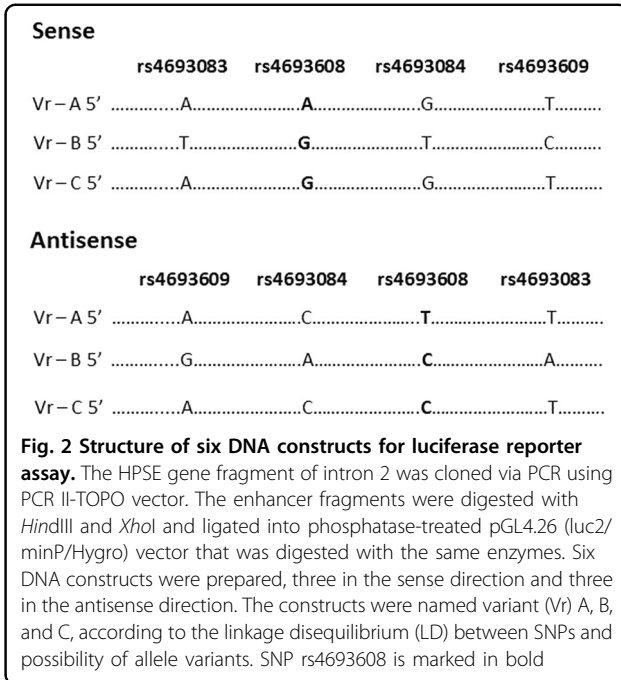
```

**Fig. 1 Map of the HPSE enhancer region.** The size of analyzed enhancer region is 440 bp. It includes four polymorphic SNPs: rs4693083, rs4693608, rs4693084, and rs4693609. The map represents exact location of the SNPs and nucleotide substitution. The distance between rs4693083 and rs4693608 is 40 bp, whereas the interval between rs4693608 and rs4693084 is 17 bp. Additional rs4693609 SNP is located 199 bp upstream of rs4693084

**Table 1 Genotype and allele frequencies of the HPSE gene SNPs**

SNPs	Position	Genotypes and alleles	No	Frequencies, %
rs4693084	Intron 3	GG	91	61.5
		GT	50	33.8
		TT	7	4.7
		G	232	78.4
rs4693083	Intron 3	T	64	21.6
		AA	91	61.5
		AT	50	33.8
		TT	7	4.7
rs4693609	Intron 3	A	232	78.4
		T	64	21.6
		TT	88	59.5
		TC	53	35.8
		CC	7	4.7
rs4693608	Intron 3	T	229	77.4
		C	67	22.6
		AA	56	37.8
		AG	62	41.9
		GG	30	20.3
rs11099592	Exon 7	A	174	58.8
		G	122	41.2
		GG	91	61.5
		GA	49	33.1
		AA	8	5.4
		G	231	78
		A	65	22

exon 7. The second block starts at rs6856901 (3'UTR). However, the most significant rs4693608 was found in partial LD with all analyzed SNPs in spite of the presence of two blocks. rs4693083, rs4693084, and rs4693609 are located in the first block and were found in strong LD with rs11099592 (exon 7). Hence, these SNPs are in complete LD with rs6535455 (intron 4) and rs11099594 (intron 2, 510 bp downstream of rs4693608), which were previously found in strong LD with rs11099592<sup>19</sup>. Previous studies revealed an association between rs11099592 and the risk of acute lymphoblastic leukemia (ALL) development<sup>22</sup>, and poor outcome of gastric cancer<sup>24,25</sup>. As a result of strong LD between rs11099592 and the other SNPs located in introns 2 and 4, we assume that newly identified additional SNPs may also correlate with the malignant phenotype of these cancers.

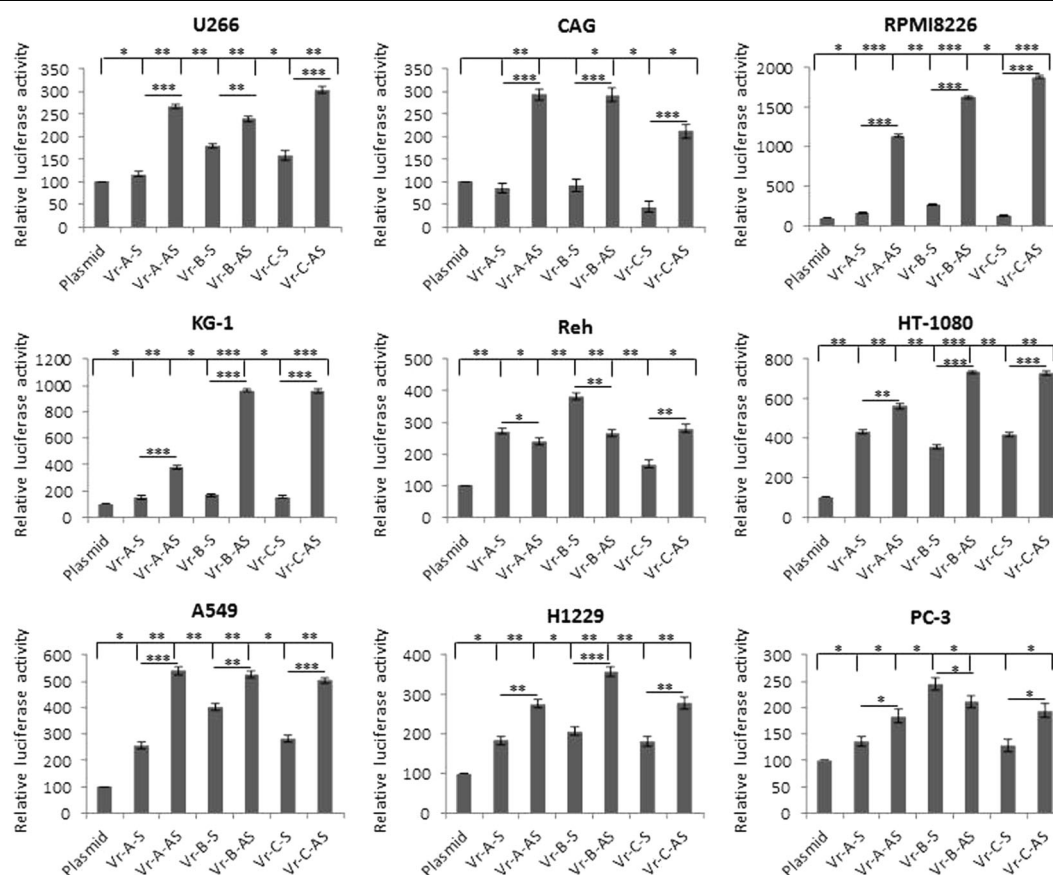


**Detection of enhancer activity in hematological and non-hematological cancer cell lines**

To determine the existence of the active enhancer of HPSE in intron 2 and the functional effects of enhancer SNPs we applied the luciferase reporter gene with minimal promoter to measure enhancer activity. Five hematological malignancies (U266, CAG, RPMI8226, KG-1, and Reh) and four solid tumor (HT-1080, A549, H1229, and PC-3) cell lines were transiently transfected with each of the six DNA constructs (Fig. 2) or an empty vector. Our results indicate that the 440 bp fragment of HPSE exhibits enhancer activity in both the sense and antisense directions of the HPSE gene (Fig. 3). Relative luciferase activity was higher in constructs with antisense direction of the enhancer in most of the tested cell lines. Moreover, constructs which included the allele G (Vr-B) revealed elevated levels of luciferase activity compared to constructs which included the A allele (Vr-A) (Table 2). Vr-C is a combination between Vr-A and Vr-B. It also includes the allele G of rs4693608; however, genotypes of additionally linked SNPs (rs4693083, rs4693084, and rs4693609) are identical to Vr-A. Enhancer activity of variant Vr-C was sometimes similar to that of Vr-A, and occasionally to Vr-B, depending on the cell line.

**The A-G substitution of rs4693608 alters nuclear protein binding to the enhancer region**

The nuclear extract of 12 hematological cancer cell lines (Reh, MOLT-3, CEM, Jurkat, U-937, KG-1, K-562, NK-92, NKL, SU-DHL4, CAG, and U266), 14 solid tumor cell lines (A549, H1229, HT-1080, 5637, PC-3, JAR,



**Fig. 3 Reporter activity of the HPSE gene enhancer containing various polymorphisms.** Six HPSE enhancer constructs were designed and inserted upstream of the luciferase gene in pGL4.26 vector with minimal promoter. The results were derived from five hematological malignancies (U266, CAG, RPMI8226, KG-1, and Reh) and four solid tumor cell lines (HT-1080, A549, H229, and PC-3). Luciferase activity of pGL4.26 without insert was used as a reference (100%). Data shown are of representative experiments performed in triplicates. Differences in luciferase activities among various constructs were evaluated by *t*-test. *p* value of  $\leq 0.05$  was regarded as statistically significant. \**p* value of  $\leq 0.05$ ; \*\**p* value of  $\leq 0.001$ ; \*\*\**p* value of  $\leq 0.0001$

COLO205, SK\_N\_SH, PANC-1, MCF-7, HeLa, T84, AGS, and Saos-2), and primary white blood cells from healthy individuals and patients with AML and ALL were incubated with biotin-labeled probes representing the enhancer region of either the A or G alleles and subjected to electrophoretic mobility shift assay. DNA–protein complexes appeared as a band shift pattern. The specificity of the protein–oligonucleotide interaction was shown for both alleles by the reduced intensity of the shifted complex upon addition of unlabeled oligonucleotide probe. Excess unlabeled control oligonucleotide probes compete for binding of the target proteins in control samples.

Analysis of healthy control samples demonstrated gel shift bands for both allelic probes (Fig. 4a). The affinity of the complex to the G allele probe was higher in comparison to the A allele probe. In all analyzed cancer cell lines the band of DNA–protein complex was shifted significantly more in comparison to normal samples (Figs. 4a

and 5), likely due to additional proteins recruited to the complex.

EMSA analysis of 26 different cell lines revealed that a single DNA/protein complex binds to allele G, while multiple complexes bind to allele A (Fig. 5a, b). At least five different shifted bands were detected after incubation with allele A probe, reflecting the complexity of the DNA/protein interactions. Some of the cell lines exhibited the same shifted bands pattern while other revealed individual shifted bands (Fig. 5a, b). Analysis of AML and ALL patient samples disclosed a pattern of shifted bands similar to cancer cell lines, but the affinity of the DNA/protein complexes was significantly higher in the majority of the leukemia samples (Fig. 5c).

In addition, we analyzed the effect of G-CSF and LPS on the DNA/protein complexes. Nuclear protein extracts from total leukocytes of healthy donors before and after 5 days of G-CSF treatment and prior to graft collection were obtained. EMSA revealed that G-CSF led to an

**Table 2 SNPs effect on the reporter activity of the HPSE gene enhancer**

Cell lines	SNP variants (sense)	p value	SNP variants (antisense)	p value
U226	A-B	0.0003	A-B	0.006
	A-C	0.008	A-C	0.004
	B-C	0.04	B-C	0.0004
CAG	A-B	NS	A-B	NS
	A-C	0.01	A-C	0.002
	B-C	0.01	B-C	0.003
RPMI8226	A-B	0.0003	A-B	$5.8 \times 10^{-6}$
	A-C	0.02	A-C	$3.8 \times 10^{-7}$
	B-C	0.0001	B-C	$0.7 \times 10^{-5}$
KG-1	A-B	NS	A-B	$3.1 \times 10^{-6}$
	A-C	NS	A-C	$9.1 \times 10^{-7}$
	B-C	NS	B-C	NS
REH	A-B	0.0003	A-B	NS
	A-C	0.0005	A-C	0.02
	B-C	$2.48 \times 10^{-5}$	B-C	NS
HT-1080	A-B	0.001	A-B	0.0001
	A-C	NS	A-C	$9.98 \times 10^{-5}$
	B-C	0.002	B-C	NS
A459	A-B	0.0001	A-B	NS
	A-C	0.06	A-C	0.03
	B-C	0.0003	B-C	NS
H1229	A-B	0.05	A-B	0.001
	A-C	NS	A-C	NS
	B-C	0.04	B-C	0.004
PC-3	A-B	0.0004	A-B	0.06
	A-C	NS	A-C	NS
	B-C	0.0003	B-C	NS

elevation of nuclear protein levels including the amount of DNA/protein complexes for both alleles (Fig. 4b). In contrast, LPS treatment resulted in disappearance of DNA/protein complexes (Fig. 4c). No difference was found in affinity and shifted bands pattern between DNA/

proteins complexes in normal total leukocytes versus normal mononuclear cells (MNCs)<sup>28</sup>.

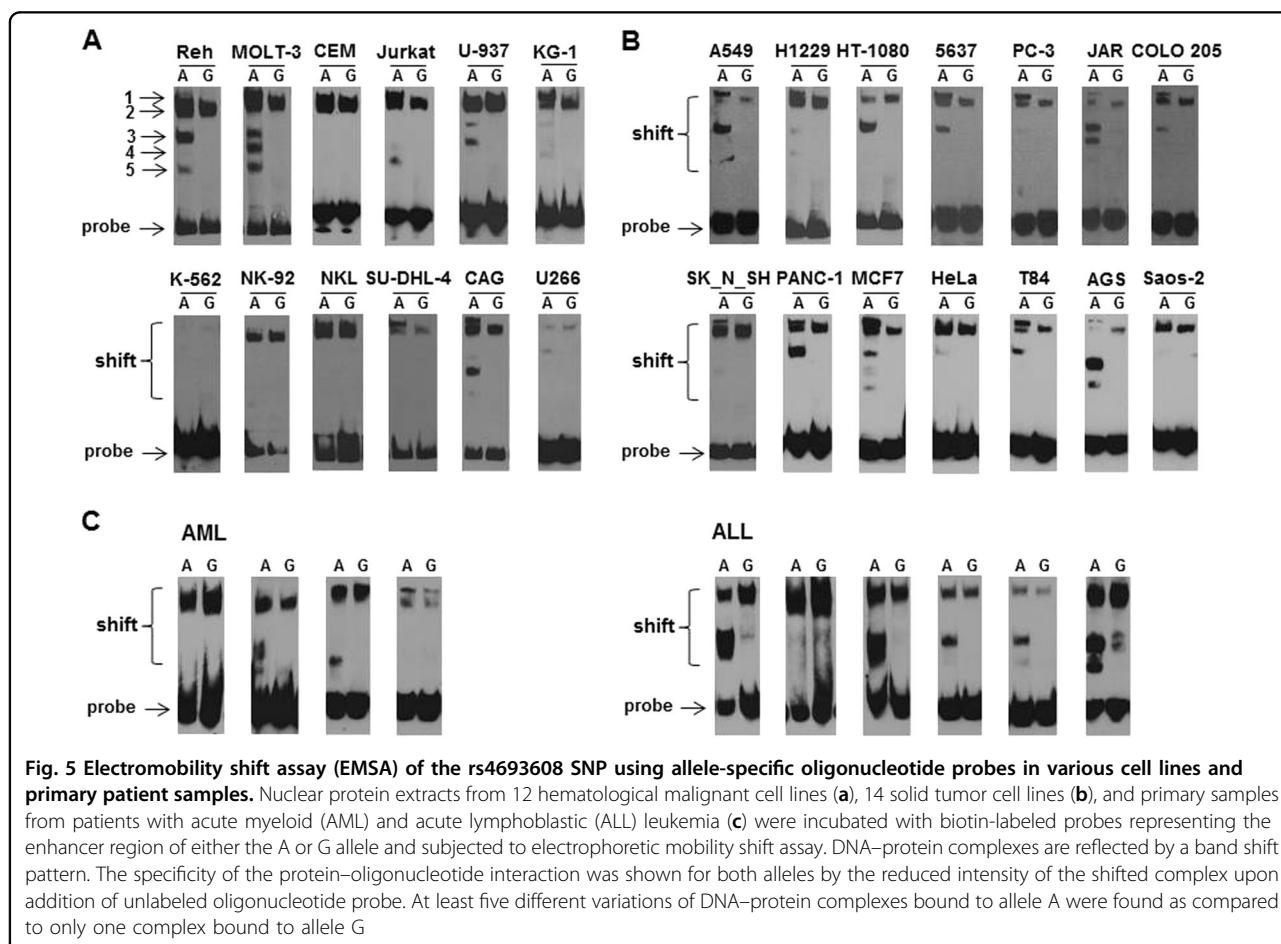
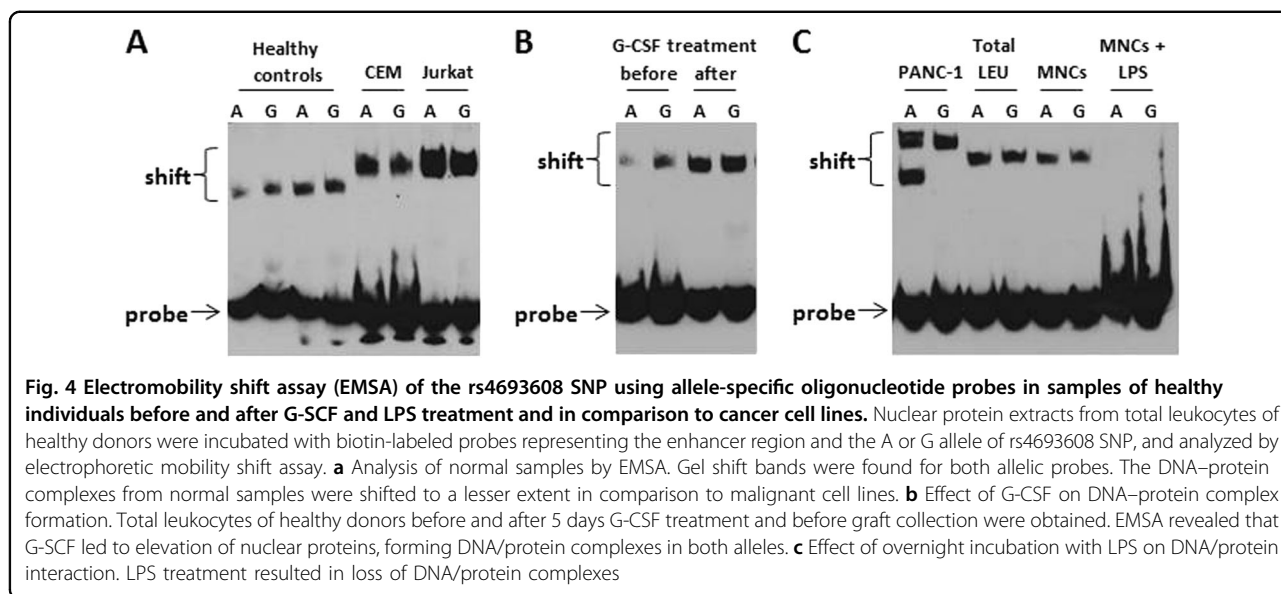
#### Analysis of nuclear protein binding to the DNA in the enhancer region containing rs4693084, rs4693083, and rs4693609 SNPs

The distance between rs4693608 and the other three SNPs is short. Therefore, these SNPs together with rs4693608 may be part of a large DNA/protein complex. The effect of additional three SNPs was analyzed. For this purpose, three allele-specific biotin-labeled probes were generated and EMSA was performed in hematological malignant cell lines and leukemia samples. The assay revealed that only rs4693084 affected protein binding to DNA in the SNP region in both the cell lines and primary patient samples. In multiple myeloma (MM) cell lines, DNA/protein complex formation was observed only for the G allele. Analysis of ALL cell lines and primary samples demonstrated gel shifted bands for both allelic probes. The affinity of the complex to the G allele probe was higher in comparison to the A allele probe (Fig. 6). No nuclear protein binding was found with a probe that included rs4693083 in MM and ALL cell lines. No differences in protein binding between alleles of rs4693083 and rs4693609 were detected (Fig. 6).

#### Identification of proteins that bind to the enhancer region of HPSE in healthy individuals and cancer cell lines

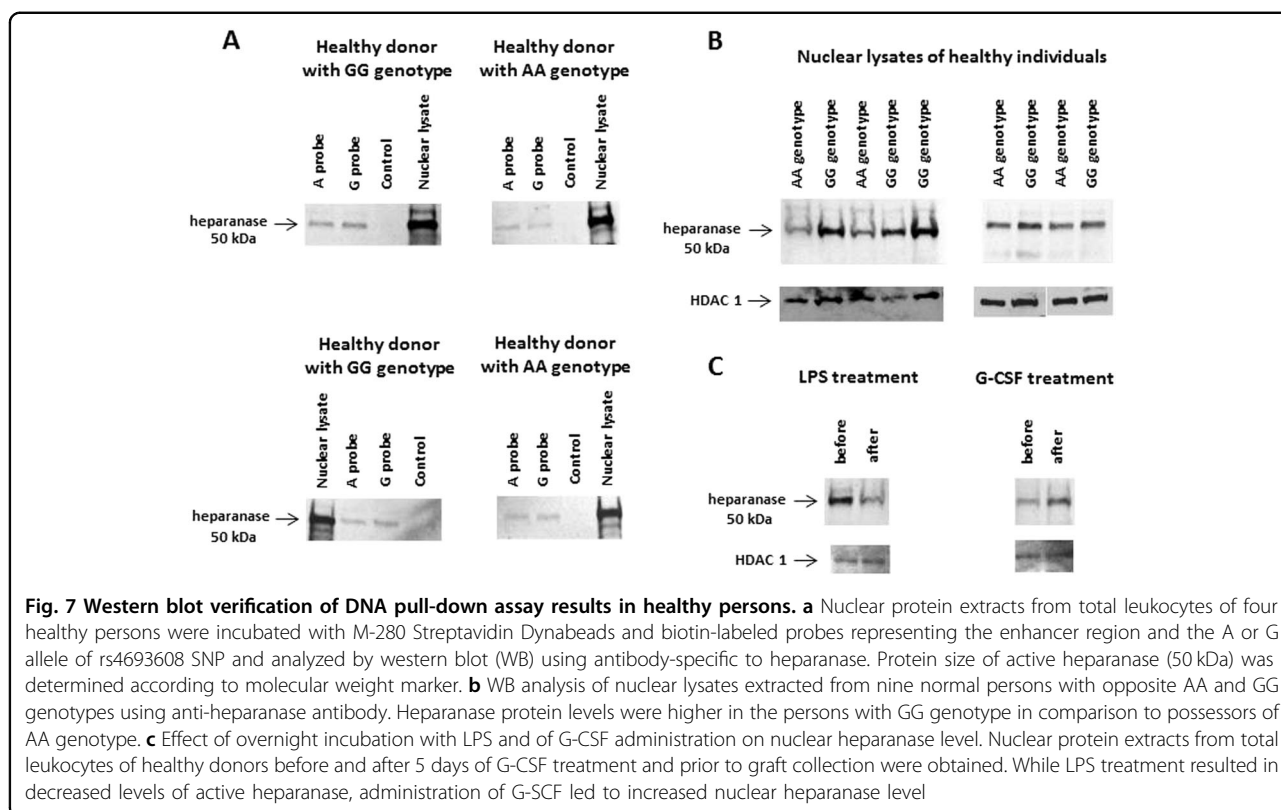
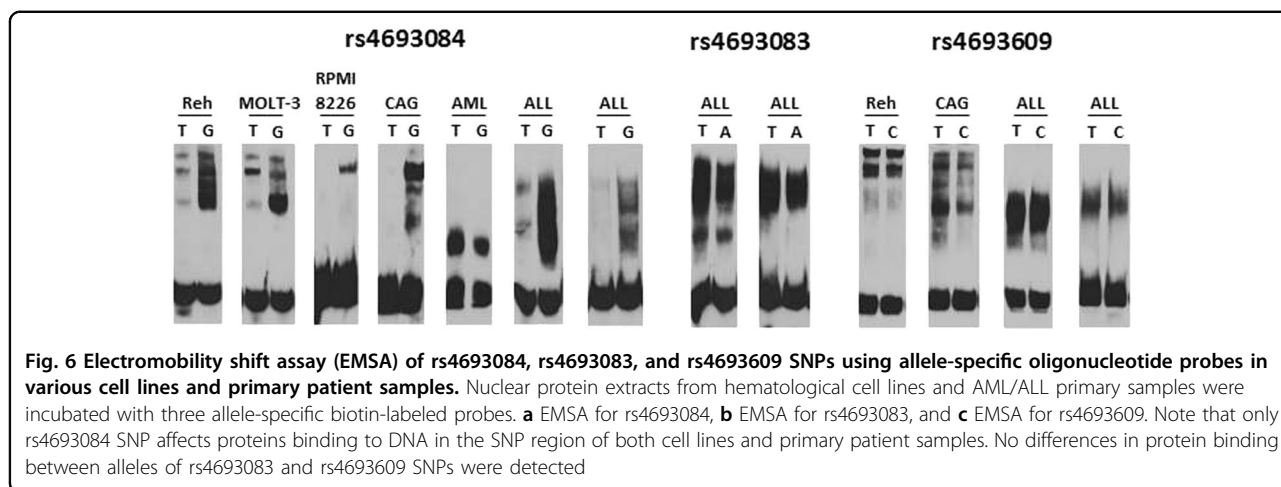
We applied DNA pull-down assay to identify proteins that were bound to the enhancer region. 5'-biotinylated A and G probes that included the rs4693608 SNP were incubated with M-280 Streptavidin Dynabeads and with nuclear extracts from total leukocytes of healthy individuals, Reh, H1229, and PC-3 cell lines. Specific proteins bound to the probes were separated by gradient Mini-PROTEAN<sup>®</sup> TGX<sup>™</sup> Gels and compared with those processed without biotinylated probes. Unique proteins present in the samples were analyzed by mass spectrometry, taking into account the EMSA results. In normal samples we selected proteins that bind to both the A and G alleles but did not appear in control samples and the three cancer cell lines. In contrast, for analysis of malignant cells we chose proteins that were bound to both probes in the three cancer cell lines but were not detected in control samples and normal leukocytes (Supplementary Tables 1 and 2). Proteins that were discovered in the DNA pull-down assay were verified by western blot analyses (Figs. 7 and 8).

Mass spectrometry results of normal samples (Supplementary Table 1) revealed that heparanase itself may bind to the enhancer region. Western blot analysis confirmed this observation (Fig. 7a). Heparanase binds to both A and G probes in the DNA-protein complex. Moreover, the amount of heparanase bound to the G probe was higher



than that bound to the A probe. This corresponds to the EMSA results in which the affinity of the complex to the G allele probe was higher in comparison to the A allele

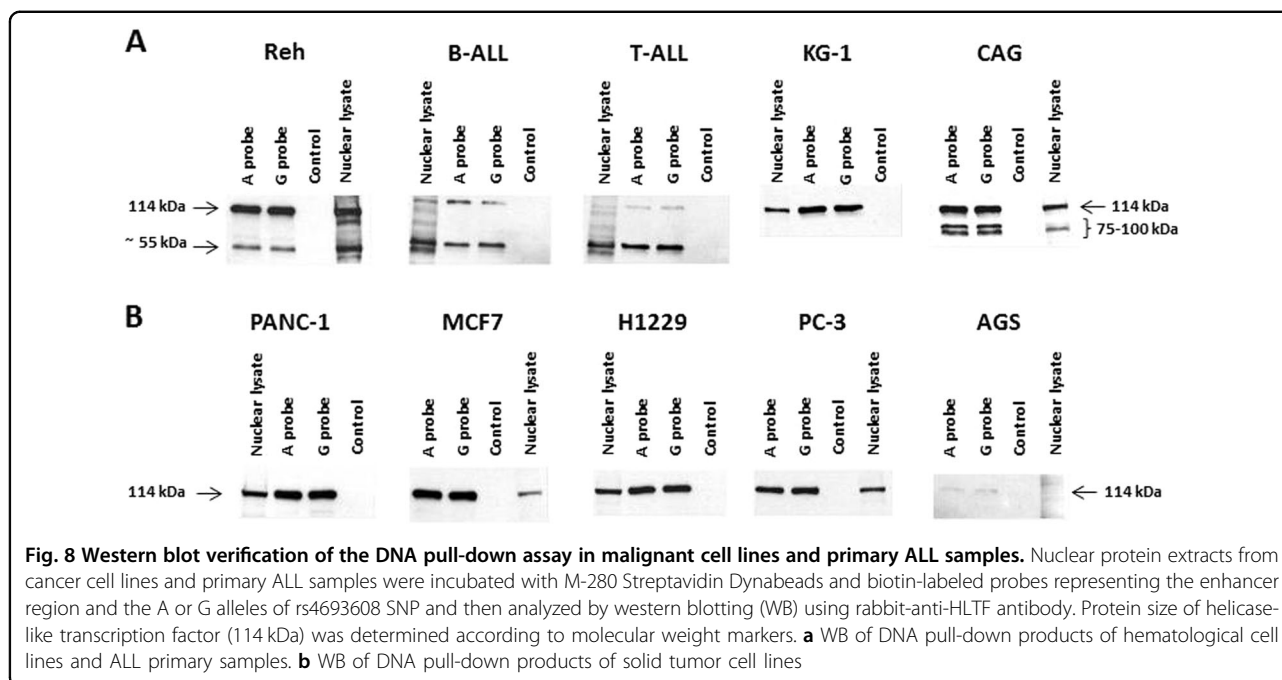
probe (Fig. 4). Western blot analysis of nuclear extracts from healthy individuals with AA and GG genotypes showed that heparanase levels were higher in individuals



with GG genotype in comparison to possessors of the AA genotype (Fig. 7b). Nuclear heparanase appeared as a 50 kDa protein corresponding to the active form of the enzyme.

Western blot analysis of MNCs pre-treated with LPS or G-CSF revealed a decrease in the amount of nuclear heparanase in response to LPS treatment as opposed to increased level of nuclear heparanase following treatment with G-CSF (Fig. 7c). These observations coincide with the EMSA results (Fig. 4b, c).

Unlike the above results, DNA pull-down analysis followed by western blot verification revealed that heparanase was not detected in the nuclear fraction of malignant cell lines and primary ALL samples. Instead, the helicase-like transcription factor binds to the complex (Supplementary Table 2). According to the EMSA results (Fig. 5), the protein that was bound to the A and G alleles is of molecular weight (MW) higher than that of heparanase. Notably, the MW of HLTF is 114 kDa and it binds to both alleles in each of the malignant samples. Whereas



in solid tumor cell lines (PANC I, PC-3, H1229, MCF-7, and AGS), one specific 114 kDa band was noted (Fig. 8b), in hematological cell lines and primary ALL samples additional specific bands were observed (Fig. 8a). It is conceivable that alternative splice forms of HLTF may also bind to this enhancer region. Notably, the amount of HLTF protein that was bound to the enhancer probe in AGS cells line was low (Fig. 8b), corresponding to the EMSA results in this cell line (Fig. 5b).

## Discussion

SNPs located in gene regulatory elements often facilitate the progression of diverse human diseases. Genetic variations contribute to disease development by alterations in transcription factor binding, enhancer activity, long-range enhancer–promoter interactions, posttranslational histone modifications, and/or RNA polymerase function. Identification of causality between non-coding genetic variants, gene regulation, and disease progression and outcome is essential for the development of novel pathways and therapies for disease treatment<sup>29</sup>.

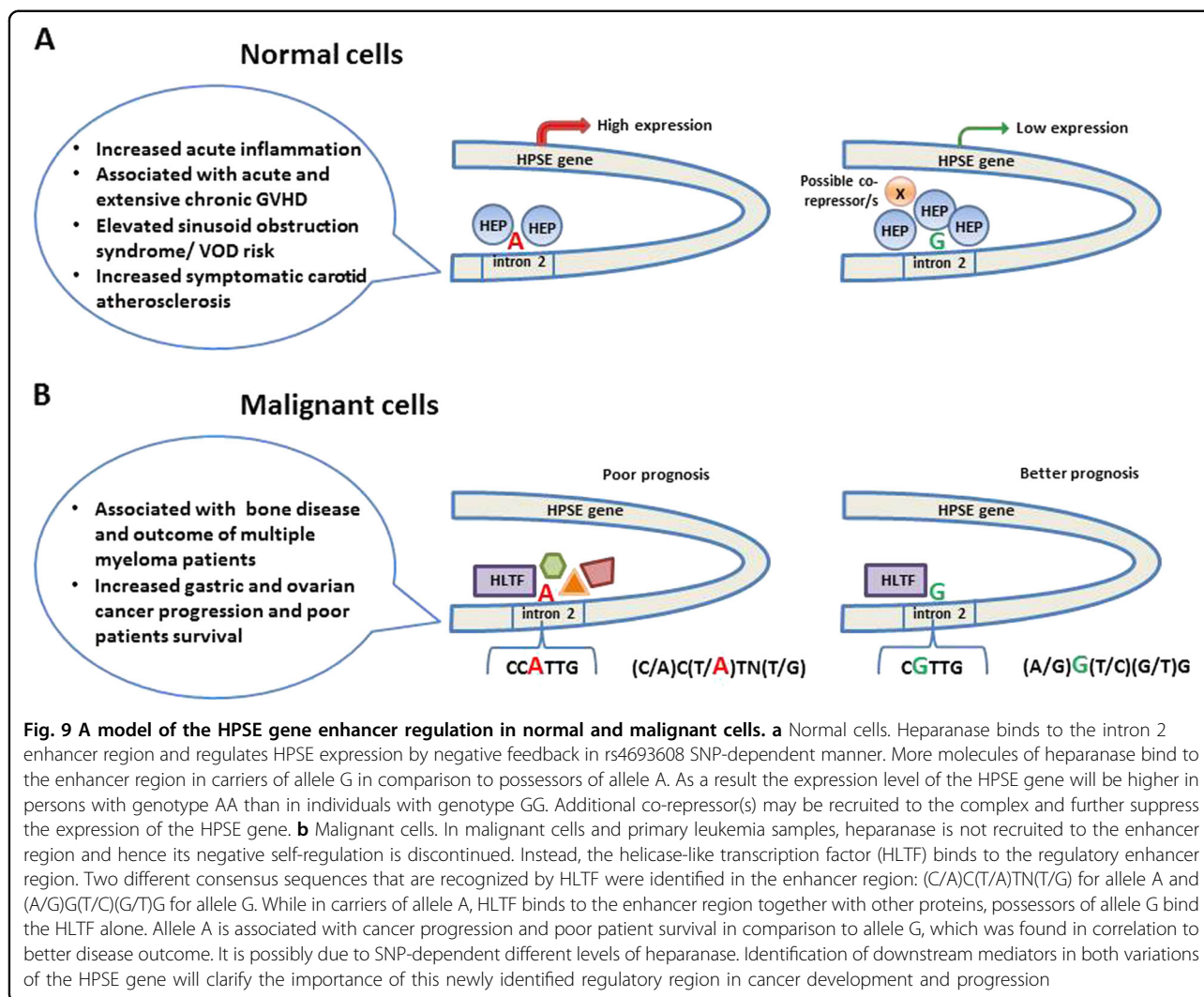
The main objective of our study was to understand how rs4693608 affects heparanase expression and thereby the development of various pathological processes. Using the UCSC Genome Browser website, we identified a potential strong enhancer in the rs4693608 region and then demonstrated that this region exhibits SNPs-dependent enhancer activity. Analysis of normal blood samples revealed binding of DNA/protein complex to both alleles with higher affinity to allele G. All the examined malignant cell lines and primary leukemia samples disclosed a

more prominent shift of the main bands in comparison to normal samples. This result is likely due to formation of DNA/protein complexes with additional proteins that were bound to the probes. At least five extra shifted bands were incorporated to allele A probe while allele G probe was bound to only one main DNA/protein complex. Despite a common band pattern, each malignant sample revealed specific shifted bands.

DNA pull-down assay followed by western blot verification showed that heparanase binds to the enhancer region of intron 2 and regulates HPSE gene expression via negative feedback in rs4693608 SNP-dependent manner (Fig. 9). This may explain our previously published results of inverse correlation between HPSE gene SNPs and heparanase expression<sup>19</sup>. Thus, an inverse correlation between HPSE mRNA and plasma heparanase levels was associated with the same SNPs. For example, the rare genotype GG of rs4693608 was associated with low HPSE mRNA expression level and high plasma levels of active heparanase. In contrast, the frequent AA genotype of this SNP was associated with high HPSE mRNA expression and low level of plasma heparanase<sup>19</sup>. The present western blot analyses revealed that the level of nuclear heparanase was higher in possessors of the GG genotype in comparison to carriers of the AA genotype (Fig. 7b).

Our previously published study showed that HPSE gene expression was elevated after LPS treatment in both cord blood and adult MNCs in rs4693608 SNP-dependent manner<sup>21</sup>. Our present EMSA and western blot results obtained with LPS-treated MNCs support such a negative feedback regulation of heparanase expression. LPS





treatment led to disappearance of DNA/protein complexes and a decrease in nuclear heparanase. Further research is needed to clarify how A–G substitution influences on a negative feedback of heparanase regulation. We assume that additional co-repressor(s) may take part in this regulation and that binding of heparanase, in *cis* or *trans*, to the enhancer region could explain the effect of rs4693608 SNP on HPSE gene expression. These results are important to better understand the involvement of heparanase and rs4693608 SNP in acute inflammation, acute and extensive chronic GVHD, risk of sinusoid obstruction syndrome/VOD following allogeneic stem cell transplantation, and symptomatic carotid atherosclerosis (Fig. 9)<sup>20–23</sup>.

Our results revealed that in malignant cell lines and primary leukemia samples, heparanase discontinues self-regulation of the enhancer region. Instead of heparanase, the helicase-like transcription factor (HLTF) binds to the regulatory region. HLTF is a

member of the yeast mating SWItch/sucrose non-fermenting (SWI/SNF) family of proteins involved in chromatin remodeling. HLTF plays a significant role in gene transcription, DNA repair, ubiquitination, and support of genome stability<sup>30,31</sup>. HLTF appears to function as a tumor suppressor, supported by the detection of HLTF promoter hypermethylation in cancer tissues and cell lines<sup>31,32</sup>. Recent studies, however, reveal increased expression of HLTF in transformed cells and cancer specimens, suggesting that HLTF could be associated with carcinogenesis and may function as an oncogene<sup>32</sup>. Two different consensus sequences recognized by HLTF were proposed: (C/A)C(T/A)TN(T/G) and (A/G)G(T/C)(G/T)G<sup>30</sup>. Analysis of the HPSE enhancer region included in rs4693608 SNP indicated that in the case of allele A, HLTF recognizes the first sequence CCATTG. A–G alteration does not result in loss of the transcription site but rather a change to the second variant of the HLTF recognition site (CGTTG),

except the first nucleotide. Further studies will clarify if these recognition sites of HLTF may activate various downstream processes. Our EMSA results disclosed that not only HLTF binds to the HPSE enhancer region in allele A, but other proteins may form a common DNA–protein complex. In the case of allele G only HLTF binds to the HPSE gene enhancer region (Fig. 9). It is therefore important to understand how different types of DNA–protein binding complexes influence heparanase expression and affect its involvement in cancer progression. Future identification of particular proteins bound to allele A of rs4693608 will likely reveal proteins specific to each cell type and disease situation. ChIP analysis applying specific ChIP-graded antibodies will enable to further elucidate the mode of interaction between DNA and proteins, involved in regulating the HPSE intron enhancer in specific cells and processes.

Enhancers are regulatory DNA sequences widely scattered throughout the genome and involved in gene activation independent of their orientation or distance. Enhancers determine the precise tissue-specific gene expression pattern during normal development and are regarded as key players in directing disease transcriptional program during pathogenesis<sup>33,34</sup>.

According to the UCSC Genome Browser website, HPSE enhancer region reveals all signs of active enhancer. Histone H3 lysine 4 monomethylation (H3K4me1), dimethylation (H3K4me2), and histone H3 lysine 27 acetylation (H3K27ac) sites were detected around rs4693608 in K-562 cell line, CD14+ and CD20+ immune cells, endothelial cells (Huvéc), and adult normal human dermal fibroblasts (NHDF-Ad).

Our previous study revealed strong correlation between rs4693608 and heparanase expression among healthy individuals<sup>19</sup>. The present study indicates that this is likely mediated by the enhancer activity. In a previously published study, Parish et al. demonstrated a novel role for nuclear heparanase as a direct epigenetic mediator of histone H3 methylation across immune responsive genes in human T cells<sup>13,16</sup>. Heparanase binds to chromatin and acts as part of an active multi-subunit complex, which includes RNA polymerase II and the lysine-specific demethylase 1 (LSD1). This complex was identified in the promoter and 5' transcribed regions of inducible immune genes. In addition, heparanase is associated with differentiation and development of activated T cells and with genes affect metabolic and biosynthetic processes in resting T cells. Chromatin-anchored heparanase also targets a distinct cohort of microRNA promoters previously found to play a key role in T cell functions<sup>13,16</sup>. It is not clear if the newly identified enhancer affects expression of other genes. An indirect epigenetic effect of rs4693608 and enhancer activity on heparanase function is a likely possibility.

According to MeDIP-seq (methylated DNA immunoprecipitation and sequencing) of the UCSC Genome Browser and MethPrimer prediction tools, allele G of rs4693608 is part of a putative CpG island. G–A alteration leads to loss of the CpG site. CpG islands are the basis of DNA modifications such as methylation and hydroxymethylation. In mammals, the majority of cytosines (70–80%) in CpG dinucleotides are methylated in somatic cells. DNA methylation has been evidently linked to transcriptional regulations. Numerous studies focus on differentially methylated regions in complex diseases. The role of DNA methylation in cancer etiology and progression is well established<sup>13,35</sup>. Our EMSA results indicate that while allele G of rs4693608 is susceptible to DNA methylation, allele A may escape this modification. Thus, at least five different shifted bands were found after incubation with allele A probe in various malignant cell lines and leukemia primary samples, only one main shifted band was detected after incubation with allele G probe.

In our study we analyzed three additional SNPs located in the enhancer. Only rs4693084 affected protein binding to DNA in the SNP region of cell lines and leukemia primary samples. In addition, all three SNPs were found in strong LD with rs11099592. Previous studies revealed association between rs11099592 and the risk of ALL development. Additional correlation was found with the outcome of gastric cancer<sup>18,24,25</sup>. Given the strong LD between rs11099592 and rs4693084, disease associations may result from different enhancer activity. The distance between rs4693608 and rs4693084 is only 17 bp, and under appropriate conditions both SNPs may be part of a common DNA/protein complex.

Analysis of the functional effect of SNPs on luciferase activity revealed that constructs that included the G allele (Vr-B) exhibited higher levels of luciferase in comparison to constructs which included the A allele. Taken into account the EMSA and western blot results in malignant cells, we assume that the activity of HPSE enhancer depends on the quality and quantity of DNA/protein complexes. When the number of nuclear proteins bound to the enhancer is low, the enhancer activity is high.

Remarkably, SNPs appear to affect binding of transcription factors to the enhancers and thereby modulate the expression of target genes. Importantly, regions termed variant enhancer loci (VEL) regulate the complex interactions between genetic and epigenetic elements that redefine which enhancers are turned on or off in cancer<sup>36,37</sup>.

The present study identified the HPSE gene enhancer, located in intron 2, which is involved in self-regulation of heparanase by a negative feedback exerted in rs4693608-dependent manner. During carcinogenesis, self-regulation of heparanase is attenuated and the helicase-like

**Table 3 Primer sequences for SNPs detection and generation of allele-specific probes**

Name of SNP or probe	Modification	Sequence of primers	PCR/probe product size, bp
rs4693083 SNP	T → A	Forward: 5' CTGGGCGCAAAGCGAGACA 3' Reverse: 5' TGACCAGGGTGGATTTTTTC 3'	285
rs4693084 SNP	T → G	T allele forward: 5' TTGACTTTTTTCGCTTT 3' G allele forward: 5' TTGACTTTTTTCGCTTG 3' Reverse: 5' ACAGGCACACACCACCACG 3'	238
rs4693609 SNP	C → T	Forward: 5' TGTTTTCCTCTTGCCATCAT 3' T allele reverse: 5' CACACACCACCACGCCAG 3' G allele reverse: 5' CACACACCACCACGCCAA 3'	279
Enhancer cloning		Forward: 5' CTGGGCGCAAAGCGAGACA 3' Reverse: 5' ACAGGCACACACCACCACG 3'	440
rs4693608 probe	A → G	A allele forward: 5' [Btn]GGGTGGGAACCAATTGACTTTTTTCGC 3' A allele reverse: 5' [Btn]GCGAAAAAAGTCAATGGTCCCACCC 3' G allele forward: 5' [Btn]GGGTGGGAACCGTTGACTTTTTTCGC 3' G allele reverse: 5' [Btn]GCGAAAAAAGTCAACGGTCCCACCC 3'	30
rs4693083 probe	T → A	T allele forward: 5' [Btn]TCTGCAGTGTGGTGTTCCTCTTGCCATC 3' T allele reverse: 5' [Btn]GATGGCAAGAGGAAAACACCACACTGCAGA 3' A allele forward: 5' [Btn]TCTGCAGTGTGGTGAATTCCTCTTGCCATC 3' A allele reverse: 5' [Btn]GATGGCAAGAGGAAATCACCACACTGCAGA 3'	30
rs4693084 probe	T → G	T allele forward: 5' [Btn]ACTTTTTTCGCTTTGAATTTGGTACCATC 3' T allele reverse: 5' [Btn]GATGGTACCAAATCAAGCGAAAAAAGT 3' G allele forward: 5' [Btn]ACTTTTTTCGCTTGAATTTGGTACCATC 3' G allele reverse: 5' [Btn]GATGGTACCAAATCAAGCGAAAAAAGT 3'	30
rs4693609 probe	C → T	C allele forward: 5' [Btn]TACGAAAACCTTAGCTGGGCGTGGTGGTGT 3' C allele reverse: 5' [Btn]ACACCACCACGCCACGCTAAGTTTCGTA 3' T allele forward: 5' [Btn]TACGAAAACCTTAGTTGGGCGTGGTGGTGT 3' T allele reverse: 5' [Btn]ACACCACCACGCCATCTAAGTTTCGTA 3'	30

transcription factor starts to regulate this enhancer region, again in rs4693608-dependent manner. Analysis of disease-specific cell lines or primary samples allows precise identification of tissue-specific proteins that bind to the HPSE enhancer region. This knowledge could potentially enable modification of the enhancer activity in a way that will modulate the expression of HPSE and thereby influence downstream processes and disease progression.

## Materials and methods

### Samples and cell lines

Human total leukocytes from healthy individuals and patients with AML and ALL as well as malignant cell lines were applied in the study. All subjects gave their written

informed consent. Thirteen hematological malignant cell lines (K-562, KG-1, U-937, Reh, MOLT-3, Jurkat, CEM, SU-DHL4, CAG, RPMI8226, U266, NK-92, and NKL) and 15 solid tumor cell lines (HT-1080, JAR, A549, H1229, PC-3, 5637, SK\_N\_SH, MCF-7, HeLa, COLO205, T84, AGS, Saos-2, PANC-1, and MDA-MB435) were analyzed. The cells were maintained in RPMI1640 and DMEM media (Gibco by Life Technologies, Inc., Paisley, UK) supplemented with 10% fetal bovine serum (FBS), 1 mM L-glutamine, 100 U/ml penicillin, and 0.01 mg/ml streptomycin (Biological Industries, Israel) in a humidified atmosphere of 5% CO<sub>2</sub> at 37 °C. All cell lines were analyzed by STR DNA fingerprinting using AmpFISTR Identifier Kit (Applied Biosystem) and tested for mycoplasma contamination.

### SNPs analysis of the enhancer region

The enhancer region includes three additional polymorphic SNPs besides rs4693608. Genotypes of rs4693084, rs4693609, rs4693608, and rs11099592 were performed by allele-specific amplification. PCR-RFLPs-based assay was used for rs4693083 detection. The PCR fragments were amplified from genomic DNA (the Wizard<sup>®</sup> Genomic DNA Purification kit (Promega, Madison, WI), using the forward and reverse primers (Table 3). The PCR reactions were performed as described<sup>17–19</sup>. Annealing temperature was 56 °C for 45 s for rs4693083 and rs4693609, 52 °C for 30 s for rs4693084. Genotyping of rs4693083 was performed by digestion of PCR fragments with *HphI* restriction endonuclease (New England Biolabs).

### DNA constructs

The HPSE gene fragment of intron 2 (440 bp), which includes rs4693083, rs4693608, rs4693084, and rs4693609 SNPs, was cloned via PCR using PCR II-TOPO vector (Invitrogen by Life Technologies, Carlsbad, CA). The forward and reverse primers for PCR amplification of the enhancer region are shown in Table 1. The enhancer fragments were digested with *HindIII* and *XhoI* and ligated into phosphatase-treated pGL4.26 (*luc2/minP/Hygro*) vector (Promega, Madison, WI), which was digested with the same enzymes. JM109 competent cells (Promega, Madison, WI) were used for cloning. Six DNA constructs were prepared, three in sense direction and three in antisense direction (Fig. 2). The direction of all constructs was confirmed by DNA sequencing.

### Luciferase reporter assay

All cell lines (U266, CAG, RPMI8226, KG-1, Reh, HT-1080, A549, H1229, and PC-3) were seeded in six-well culture dishes according to Lonza instruction (AMAXA Biosystems, Lonza, Germany). For the luciferase assay, cells were transfected with 2 µg of each DNA construct using the Ingenio Electroporation Kit (Mirus Bio, Madison, WI) and Nucleofector<sup>TM</sup> (AMAXA biosystems, Lonza, Germany).

Twenty-four hours after electroporation, the cells were lysed in passive lysis buffer (Promega, Madison, WI). The luciferase assay system (Promega, Madison, WI) and 20/20<sup>n</sup> Luminometer (Turner BioSystems, Sunnyvale, CA) were used for the firefly luciferase activity measurement. Two methods were applied for normalization. The first was total protein quantification according to the Bradford method (Bio-Rad, Hercules, CA). pGL4.73 (*hRluc/SV40*) vector was used as an internal control vector (Promega, Madison, WI) in the second method. Luciferase activity of pGL4.26 without insert was defined as a 100%. The data shown are representative examples of at least three experiments, which were performed in triplicates.

Differences in luciferase activities among constructs were determined by *t*-test. A *p* value of ≤0.05 was considered as statistically significant.

### Electromobility shift assay (EMSA)

The primer sequences used for generating allele-specific biotin-labeled probes are shown in Table 3. Altered bases are marked in bold. Oligonucleotide probes were generated by heating the complementary oligonucleotides in annealing buffer (10 mM Tris, pH 7.5–8.0, 50 mM NaCl, 1 mM EDTA) at 95 °C for 5 min and then slowly cooled to room temperature. Nuclear protein extracts derived from cell lines and primary normal and leukemia samples were obtained using a Nuclear Extraction Kit (Millipore, Temecula, CA). EMSA was performed using Gelshift Chemiluminescent EMSA Kit (Active Motif, Rixensart, Belgium) according to the manufacturer's protocol. Twenty fmol of biotin-labeled oligonucleotides were incubated for 20 min with nuclear extracts (8 µg). Competitive binding was performed by including 4 pmol unlabeled oligonucleotides in control reactions. Biotin-labeled oligonucleotide-nuclear extract complexes were separated by 6% non-denaturing polyacrylamide gel in 0.5× TBE cooled buffer, at 100 V for 1.5 h and transferred onto positively charged nylon membranes BrightStar-Plus (Ambion by Life Technologies, Austin, TX) by Mini-Trans-Blot Electrophoretic Transfer Cell. Blots were visualized by chemiluminescence. The data shown are representative examples of at least three experiments

### Pull-down assay

The pull-down assay was performed using Dynabeads M-280 Streptavidin (Invitrogen by Thermo Fisher Scientific, Norway). An aliquot of 20 µl of beads was washed in PBS-T (0.1% Triton X-100) and incubated with 500 ng of each biotinylated DNA probe overnight at 4 °C. The beads were washed in PBS-T and then in pull-down wash buffer containing 20% glycerol, 0.1 M KCl, 0.2 mM EDTA, 20 mM Hepes-KOH pH 7.9, 1 mM DTT, 0.1% Triton X-100, 0.1× Halt Protease inhibitor Cocktail (Thermo Scientific, Rockford, IL, USA), 0.1 mM PMSF. Nuclear protein extract (500 µg) was incubated with beads for 15 min at 4 °C in pull-down incubation buffer (5% glycerol, 150 mM KCl, 0.05 mM EDTA, 0.025% Triton X-100, 5 mM Hepes-KOH pH 7.9, 1 mM DTT, 0.1× Halt Protease inhibitor Cocktail, 0.1 mM PMSF). The beads were washed in pull-down wash buffer, eluted in sample buffer, and boiled at 95 °C for 5 min.

### Mass spectrometry

Proteins, which were eluted by the DNA pull-down assay, were run into a Mini-PROTEAN<sup>®</sup> TGX<sup>TM</sup> 4%–20% precast Gels (BIO-RAD Laboratories, Inc. USA) and stained with Imperial<sup>TM</sup> Protein Stain (Thermo Scientific,

Rockford, IL, USA). The stained gel pieces were digested with trypsin, and then analyzed by LC-MS/MS on Q-Exactive plus (Thermo) and identified by Discoverer software (Discoverer 1.4) with two search algorithms: Sequest (Thermo) and Mascot (Matrix Science) against the human proteome from the Uniprot database, and a decoy database. All the identified peptides were filtered with high confidence, top rank, mass accuracy, and a minimum of two peptides. High-confidence peptides have passed the 1% FDR (false discovery rate) threshold. Semi quantitation was done by calculating the peak area of each peptide. The area of the protein is the average of the three most intense peptides from each protein.

### Western blotting

Pulled down proteins were separated on Mini-PROTEAN<sup>®</sup> TGX<sup>™</sup> 4–20% precast gels (BIO-RAD Laboratories, Inc. USA) and then transferred onto a 0.2 µm nitrocellulose membrane by Trans-Blot<sup>®</sup> Turbo<sup>™</sup> Transfer Pack (BIO-RAD Laboratories, Inc. USA) and Transfer Blot Turbo Transfer System (BIO-RAD). Each membrane was blocked with 5% BSA in PBS-T. A rabbit-anti-HLTF antibody (Abcam, ab17984) and rabbit-anti-heparanase antibody (InSight Pharmaceuticals, Rehovot, Israel) were applied for the assay. After washing, membrane was incubated with HRP-conjugated anti-rabbit secondary antibody. Detection was performed by applying enhanced chemiluminescence (EZ-ECL, Biological Industries, Israel).

### Acknowledgements

We thank Anna and Sagie Rodin (Technion) for bioinformatic assistance. This study was supported by research grants awarded to A.N. by Research grant from The Varda and Boaz Dotan Research Center in Hemato-Oncology Tel Aviv University; TRANSCNA (grant 11679/16) and to I.V. by the Israel Science Foundation (grant 601/14); the United States-Israel Binational Science Foundation (BSF); and the Israel Cancer Research Fund (ICRF); I.V. is a research professor of the ICRF.

### Authors' contributions

O.O. designed and performed experiments, analyzed and interpreted data, and wrote the manuscript; A.H.G.-P. and K.B. performed experiments; M.M. collected samples; J.C. and A.S. co-wrote the manuscript; I.V. co-directed the study and co-wrote the manuscript; A.N. directed the study and co-wrote the manuscript.

### Author details

<sup>1</sup>Department of Hematology and Bone Marrow Transplantation, Chaim Sheba Medical Center, Tel-Hashomer, Israel. <sup>2</sup>Cancer and Vascular Biology Research Center, Rappaport Faculty of Medicine, Technion, Haifa, Israel

### Conflict of interest

The authors declare that they have no conflict of interest.

### Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Supplementary Information** accompanies this paper at (<https://doi.org/10.1038/s41389-018-0060-8>).

Received: 5 September 2017 Revised: 3 May 2018 Accepted: 20 May 2018  
Published online: 29 June 2018

### References

- Vlodavsky, I. & Friedmann, Y. Molecular properties and involvement of heparanase in cancer metastasis and angiogenesis. *J. Clin. Invest.* **108**, 341–347 (2001).
- Sanderson, R. D., Yang, Y., Suva, L. J. & Kelly, T. Heparan sulfate proteoglycans and heparanase - partners in osteolytic tumor growth and metastasis. *Matrix Biol.* **23**, 341–352 (2004).
- Vlodavsky, I., Miao, H. Q., Medalion, B., Danagher, P. & Ron, D. Involvement of heparan sulfate and related molecules in sequestration and growth promoting activity of fibroblast growth factor. *Cancer Met. Rev.* **15**, 177–186 (1996).
- Parish, C. R., Freeman, C. & Hulett, M. D. Heparanase: a key enzyme involved in cell invasion. *Biochim. Biophys. Acta* **1471**, M99–M108 (2001).
- Ilan, N., Elkin, M. & Vlodavsky, I. Regulation, function and clinical significance of heparanase in cancer metastasis and angiogenesis. *Int. J. Biochem. Cell Biol.* **38**, 2018–2039 (2006).
- Vlodavsky, I. et al. Significance of heparanase in cancer and inflammation. *Cancer Microenviron.* **5**, 115–132 (2012).
- Lerner, I. et al. Function of heparanase in prostate tumorigenesis: potential for therapy. *Clin. Cancer Res.* **14**, 668–676 (2008).
- Gohji, K. et al. Heparanase protein and gene expression in bladder cancer. *J. Urol.* **166**, 1286–1290 (2001).
- Boyango, I. et al. Heparanase cooperates with Ras to drive breast and skin tumorigenesis. *Cancer Res.* **74**, 4504–4514 (2014).
- Ma, X. M. et al. Heparanase promotes human gastric cancer cell migration and invasion by increasing Src and p38 phosphorylation expression. *Int. J. Clin. Exp. Pathol.* **7**, 5609–5621 (2014).
- Vlodavsky, I. et al. Heparanase: from basic research to therapeutic applications in cancer and inflammation. *Drug Resist. Updat.* **29**, 54–75 (2016).
- Secchi, M. F. et al. Recent data concerning heparanase: focus of fibrosis, inflammation and cancer. *Biomol. Concepts* **6**, 415–442 (2015).
- Parish, C. R. et al. Unexpected new roles for heparanase in type 1 diabetes and immune gene regulation. *Matrix Biol.* **32**, 228–233 (2013).
- Rabelink, T. J. et al. Heparanase: role in cell survival, extracellular matrix, remodeling and the development of kidney disease. *Nat. Rev. Nephrol.* **13**, 201–212 (2017).
- Fux, L. et al. Structure-function approach identifies a C-terminal domain that mediates heparanase signaling. *Cancer Res.* **69**, 1758–1767 (2009).
- He, Y. Q. et al. The endoglycosidase heparanase enters the nucleus of T lymphocytes and modulates H3 methylation at actively transcribed genes via the interplay with key chromatin modifying enzymes. *Transcription* **3**, 130–145 (2012).
- Ostrovsky, O. et al. Characterization of HPSE gene single nucleotide polymorphisms in Jewish populations of Israel. *Acta Haematol.* **177**, 57–64 (2007).
- Ostrovsky, O. et al. Association of heparanase gene (HPSE) single nucleotide polymorphisms with hematological malignancies. *Leukemia* **21**, 2296–2303 (2007).
- Ostrovsky, O. et al. Inverse correlation between HPSE gene single nucleotide polymorphisms (SNPs) and heparanase expression: possibility of multiple levels of heparanase regulation. *J. Leukoc. Biol.* **86**, 445–455 (2009).
- Ostrovsky, O., Shimoni, A., Rand, A., Vlodavsky, I. & Nagler, A. Genetic variations in the heparanase gene (HPSE) associate with increased risk of GVHD following allogeneic stem cell transplantation: effect of discrepancy between recipients and donors. *Blood* **115**, 2319–2328 (2010).
- Ostrovsky, O. et al. Modification of heparanase gene expression in response to conditioning and LPS treatment: strong correlation to rs4693608 SNP. *J. Leukoc. Biol.* **95**, 677–688 (2014).
- Seifert, C., Wittig, S., Arndt, C. & Gruhn, B. Heparanase polymorphisms: influence on incidence of hepatic sinusoidal obstruction syndrome in children undergoing allogeneic hematopoietic stem cell transplantation. *J. Cancer Res. Clin. Oncol.* **141**, 877–885 (2015).
- Osterholm, C. et al. Increased expression of heparanase in symptomatic carotid atherosclerosis. *Atherosclerosis* **226**, 67–73 (2013).

24. Yue, Z. et al. Association of heparanase gene (HPSE-1) single nucleotide polymorphisms with gastric cancer. *J. Surg. Oncol.* **102**, 68–72 (2010).
25. Li, A. L. et al. Polymorphisms and haplotype in heparanase gene associations with the progression and prognosis of gastric cancer in a northern Chinese population. *PLoS ONE* **7**, e30277 (2012).
26. Ralph, S. et al. Heparanase gene haplotype (CGC) is associated with stage of disease in patients with ovarian carcinoma. *Cancer Sci.* **98**, 844–849 (2007).
27. Andersen, N. F. et al. Polymorphisms in the heparanase gene in multiple myeloma association with bone morbidity and survival. *Eur. J. Haematol.* **94**, 60–66 (2015).
28. Ostrovsky, O. et al. Functional rs4693608 SNP is located into strong HPSE gene intron enhancer and thereby influences to the ability of nuclear proteins binding and activity of this enhancer in broad number of hematological cell lines and ALL/AML primary samples. *Blood* **128**, 3934 (2016).
29. Vockley, C. M., Barrera, A. & Reddy, T. E. Decoding the role of regulatory element polymorphisms in complex disease. *Curr. Opin. Gen. Dev.* **43**, 38–45 (2017).
30. Debaue, G., Capouillez, A., Belayew, A. & Saussez, S. The helicase-like transcription factor and its implication in cancer progression. *Cell. Mol. Life Sci.* **65**, 591–604 (2008).
31. Dhont, L., Mascaux, C. & Belayew, A. The helicase-like transcription factor (HLTF) in cancer: loss of function or oncomorphic conversion of a tumor suppressor? *Cell. Mol. Life Sci.* **73**, 129–145 (2016).
32. Cho, S. H., Cinghu, S., Yu, J.-R. & Park, W.-Y. Helicase-like transcription factor confers radiation resistance in cervical cancer through enhancing the DNA damage repair capacity. *J. Cancer Res. Clin. Oncol.* **137**, 629–637 (2011).
33. Luo, Z. & Lin, C. Enhancer, epigenetics, and human disease. *Curr. Opin. Gen. Dev.* **36**, 27–33 (2016).
34. Heinz, S., Romanoski, C. E., Benner, C. & Glass, C. K. The selection and function of cell type-specific enhancers. *Mol. Cell. Biol.* **16**, 144–155 (2015).
35. Zhou, D. et al. Polymorphisms involving gain or loss of CpG sites are significantly enriched in trait-associated SNPs. *Oncotarget* **6**, 39995–40004 (2015).
36. Akhtar-Zaidi, B. et al. Epigenomic enhancer profiling defines a signature of colon cancer. *Science* **336**, 736–739 (2012).
37. Kurdistan, S. K. Enhancer dysfunction: how the main regulators of gene expression contribute to cancer. *Genome Biol.* **13**, 156–158 (2012).