

ORIGINAL ARTICLE

A polymorphic minisatellite region of *BORIS* regulates gene expression and its rare variants correlate with lung cancer susceptibility

Se-Lyun Yoon^{1,7}, Yun-Gil Roh^{1,7}, In-Sun Chu², Jeonghoon Heo³, Seung Il Kim⁴, Heekyung Chang⁵, Tae-Hong Kang¹, Jin Woong Chung¹, Sang Seok Koh¹, Vladimir Larionov⁶ and Sun-Hee Leem¹

Aberrant expression of *BORIS/CTCF* (Brother of the Regulator of Imprinted Sites/CTCF-like protein) is reported in different malignancies. In this study, we characterized the entire promoter region of *BORIS/CTCF*, including the CpG islands, to assess the relationship between *BORIS* expression and lung cancer. To simplify the construction of luciferase reporter cassettes with various-sized portions of the upstream region, genomic copies of *BORIS* were isolated using TAR cloning technology. We analyzed three promoter blocks: the GATA/CCAAT box, the CpG islands and the minisatellite region *BORIS-MS2*. Polymorphic minisatellite sequences were isolated from genomic DNA prepared from the blood of controls and cases. Of the three promoter blocks, the GATA/CCAAT box was determined to be a critical element of the core promoter, while the CpG islands and the *BORIS-MS2* minisatellite region were found to act as regulators. Interestingly, the polymorphic minisatellite region *BORIS-MS2* was identified as a negative regulator that repressed the expression levels of luciferase reporter cassettes less effectively in cancer cells compared with normal cells. We also examined the association between the size of *BORIS-MS2* and lung cancer in a case-control study with 590 controls and 206 lung cancer cases. Rare alleles of *BORIS-MS2* were associated with a statistically significantly increased risk of lung cancer (odds ratio, 2.04; 95% confidence interval, 1.02–4.08; and $P = 0.039$). To conclude, our data provide information on the organization of the *BORIS* promoter region and gene regulation in normal and cancer cells. In addition, we propose that specific alleles of the *BORIS-MS2* region could be used to identify the risk for lung cancer.

Experimental & Molecular Medicine (2016) 48, e246; doi:10.1038/emm.2016.50; published online 15 July 2016

INTRODUCTION

The *BORIS/CTCF* (Brother of the Regulator of Imprinted Sites/CTCF-like protein) gene is classified as a member of the cancer-testis antigen (CTA) family; it is expressed in the testis during spermatogenesis,^{1,2} as well as in multiple cancers, including uterine (endometrial), breast, lung and gastric cancers.^{3–6} High levels of the BORIS protein in the leukocytes of patients with breast cancer were detected, which suggests that *BORIS* can be used as a valuable marker.⁷ Moreover, a CpG island in the promoter region of *BORIS* is known to be a key factor regulating gene expression,^{4,8} and *BORIS* expression in human cancer cells has been proposed to lead to epigenetic deregulation. Overexpression of *BORIS* has been correlated

with the hypomethylation of its promoter in prostate and ovarian cancers in several reports,^{9,10} and partial demethylation of the *BORIS* promoter has been detected in ovarian,¹¹ colon¹¹ and lung cancer⁸ and in leukemic cell lines.¹¹ However, in our previous study, no correlation between methylation status and *BORIS* gene expression in gastric cancer tissues was detected.⁶

To elucidate the regulation of *BORIS* during tumorigenesis, in a previous study, we characterized the entire genomic region of the *BORIS* locus including the promoter region in gastric cancer cells.¹² A CpG island (–1096 to –762 from the first ATG) and two minisatellites (variable number of tandem repeats; *BORIS-MS1* and *BORIS-MS2*) were identified through the characterization of the genomic DNA sequence upstream of

¹Department of Biological Science, Dong-A University, Busan, Korea; ²Korean Bioinformation Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon, Korea; ³Department of Molecular Biology and Immunology, Kosin University College of Medicine, Busan, Korea; ⁴Drug & Disease Target Team, Korea Basic Science Institute, Daejeon, Korea; ⁵Department of Pathology, Kosin University College of Medicine, Busan, Korea and ⁶Developmental Therapeutic Branch, National Cancer Institute, Bethesda, MD, USA

⁷These authors contributed equally to this work.

Correspondence: Professor S-H Leem, Department of Biological Science, Dong-A University, 37 Nakdong-Daero 550 Beon-gil, Saha-gu, Busan 49315, Korea.

E-mail: shleem@dau.ac.kr

Received 18 December 2015; revised 20 January 2016; accepted 21 January 2016

the gene-coding region.¹² In this study, we assessed the expression and methylation status of the *BORIS* gene in lung cancer tissues and re-analyzed the regulatory elements, that is, the CpG island and two minisatellites, in its promoter region to determine whether these elements are necessary for expression of a reporter gene in lung cancer cells. In addition, we described two well-positioned elements that contained the binding sequences for GATA-1¹³ and the CCAAT box.¹⁴

Lung cancer is the most commonly diagnosed cancer and the leading cause of death in males. In addition, it is the fourth most common cancer in females worldwide, based on GLOBOCAN 2008 estimates,¹⁵ and it is a major public health problem in Korea. Because some minisatellite alleles are associated with human diseases,^{16,17} we investigated the relationship between cancer predisposition and minisatellite variants in *BORIS*.^{6,12} and revealed increased susceptibility in young patients with breast cancer containing short rare minisatellite alleles of *BORIS-MS2*. To determine whether allelic variation in *BORIS* minisatellites influences susceptibility to lung cancer, a case-control study was performed using a PCR-based method. Here, we report that rare allelic variants of *BORIS-MS2* are correlated with lung cancer susceptibility in a Korean population. In addition, to assess whether this minisatellite region has a role in gene regulation, the transcriptional levels of a reporter gene linked to the minisatellites and driven by the *BORIS* promoter in lung cell lines were examined.

MATERIALS AND METHODS

Study population and genotyping assays for *BORIS-MS2*

To examine the minisatellite polymorphisms in *BORIS-MS2*, a case-control study was performed with 590 cancer-free controls and 206 lung cancer cases. The controls and cases had similar characteristics regarding sex and age (control average age, 60.5 years, range 30–83 years; patient average age, 63.0 years, range 32–82 years) (Table 1). Controls were selected from the Department of Preventive Medicine and Internal Medicine at Dong-A University hospitals between 1997 and 2004 (Busan, Korea).¹² The control group, with no personal history of cancer or current cancer, was recruited and completed an interview. The biospecimens from lung cancer patients in this study

were provided by the Chungbuk National University Hospital, a member of the National Biobank of Korea, which is supported by the Ministry of Health, Welfare and Family Affairs and Dong-A University Hospital in Busan, Korea. All samples from the National Biobank of Korea were obtained with informed consent and institutional review board-approved protocols. The Committee of Bioethics of Dong-A University approved the procedure and design of this study (#IRB-06-10-02 & IRB-07-10-7; Busan, Korea).

For genotyping, genomic DNA was isolated from peripheral leukocytes, which were isolated from 400 μ l of whole blood using a Blood and Cell Culture DNA Mini Kit (Qiagen, Valencia, CA, USA). The genotyping assay of the minisatellite polymorphism was described previously¹² with the PCR primer pair 5'-CTTGGGAGACCTGGGGG ATGAATAG-3' (forward) and 5'-GCACCCCATTCCTC-3' (reverse) for *BORIS-MS2*. The PCR products were separated on a 1.2% agarose gel at 80 V for 4 h and then stained with ethidium bromide.

RNA isolation and reverse transcription

Total RNA was isolated from tissues using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. The Qiagen RNeasy Mini kit (Qiagen) was used for isolation of total RNA from cell lines and lung cancer tissue specimens. For the reverse transcription reaction, a mixture of total RNA (3 μ g), an oligo(dT)₂₀ (50 nM) and 10 mM dNTP were incubated at 65 °C for 5 min. A mixture of 1 \times RT buffer, 25 mM MgCl₂, 0.1 M DTT and RNaseOUT were incubated at 42 °C for 2 min. After incubation, Invitrogen SuperScript III (200 U μ l⁻¹) was added. This mixture was then incubated at 42 °C for 50 min followed by 70 °C for 15 min. RNase H (2 U μ l⁻¹) was then added, and the sample was incubated at 37 °C for 20 min in a 9700 Thermocycler (Perkin-Elmer, Waltham, MA, USA) as described previously.¹²

Bisulfate sequencing of the *BORIS* promoter regions

Genomic DNA was subjected to bisulfite modification using the EZ DNA Methylation-Gold Kit (Zymo Research, Irvine, CA, USA). Each sequence was amplified using G-Taq polymerase (Cosmo Genetech, Seoul, Korea) as previously described.⁶ PCR fragments were gel extracted using a Gel Extraction kit (Qiagen) and ligated into the Mighty TA cloning vector (TaKaRa, Kusatsu, Shiga, Japan). Following transformation, plasmids from individual bacterial colonies were isolated and sequenced. Using bisulfate sequencing, the methylation status of the *BORIS* promoter region was determined and represented

Table 1 Age and sex distribution of cases and controls

Characteristic	Controls, N (%)			Lung cancer cases, N (%)		
	Males	Females	Total	Males	Females	Total
Age (years)						
30–39	7 (2.1)	21 (8.0)	28 (4.7)	2 (1.2)	2 (5.7)	4 (1.9)
40–49	21 (6.4)	43 (16.3)	64 (10.8)	10 (5.8)	4 (11.4)	14 (6.8)
50–59	71 (21.8)	71 (26.9)	142 (24.1)	37 (21.6)	7 (20.0)	44 (21.4)
60–69	153 (46.9)	81 (30.7)	234 (39.7)	84 (49.1)	11 (31.4)	95 (46.1)
70–79	68 (20.9)	42 (15.9)	110 (18.6)	34 (19.9)	9 (25.7)	43 (20.9)
≥80	6 (1.8)	6 (2.3)	12 (2.0)	4 (2.3)	2 (5.7)	6 (2.9)
Average	62.9	57.5	60.5	63.4	61.1	63.0
Median	64	56	62	64	64	64
N	326	264	590	171	35	206

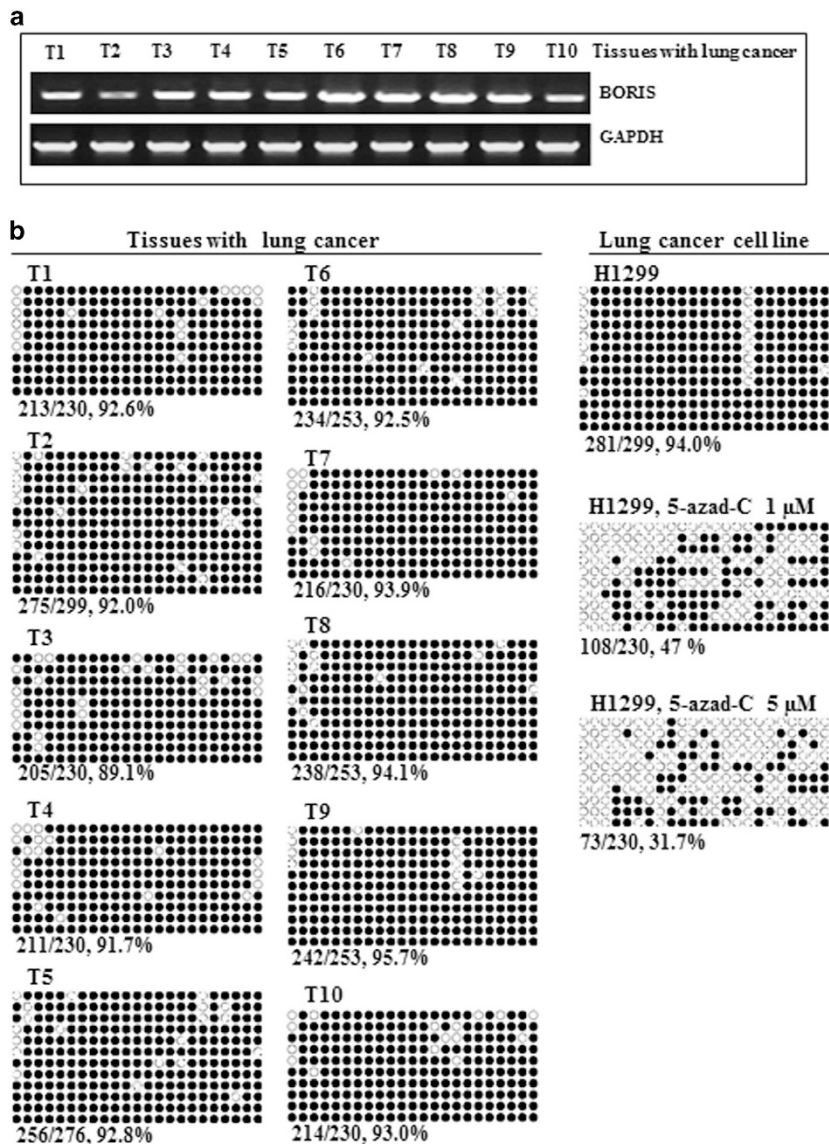


Figure 1 Gene expression and methylation status of the 5' flanking non-coding region of *BORIS*. (a) Gene expression of *BORIS* in tumor tissues (T1–T10) of lung cancer patients. (b) Results of bisulfite sequencing analysis of a CpG island in the 5' flanking region of *BORIS* in tumor tissues (T) and lung cancer cells (H1299 cells and H1299 cells treated with 1 and 5 μM 5-azad-C). Open circles (○) indicate non-methylated CpG; filled circles (●) indicate methylated CpG; percentages indicate the ratio of methylation.

as follows: open circles indicate non-methylated CpG and filled circles indicate methylated CpG (Figure 1).

Plasmid construction

To generate the luciferase reporter vectors with various fragments of the *BORIS* 5'-promoter region, fragments were amplified from a bacterial artificial chromosome clone (Supplementary Figure S1) containing the *BORIS* genomic sequence by PCR and inserted into the *KpnI/NheI* sites of the luciferase reporter vector pGL3-Basic (Promega, Madison, WI, USA). The primers and the plasmids in this study are listed in Supplementary Table S1. To characterize the effect of *BORIS*-MS2 on *BORIS* expression, part of the *BORIS* promoter p700 (–1108 to –425) was cloned and introduced into the *BglII/HindIII* sites of the luciferase reporter vector pGL3-Basic to generate the p700 fragment (Figure 2). Mutated constructs (GATA-1Δ#1, GATA-1Δ#2 and CCAATΔ) were prepared by substitution of the

GATA-1 site and CCAAT box in the p700 construct as shown in Figure 3. Two common (TR14 and TR15) alleles and five rare alleles (TR10, TR13, TR16, TR17 and TR18) were amplified from genomic DNA derived from controls and cases and inserted into the *KpnI/BglII* sites of the p700 fragment to generate the reporter plasmids p700 +TR10, p700+TR13, p700+TR14, p700+TR15, p700+TR16, p700 +TR17 and p700+TR18 (Figure 4). All constructs were confirmed by DNA sequencing.

Cells and luciferase assays

The following human cell lines were tested for the effect of *BORIS*-MS2 on *BORIS* expression: 293T/HEK293T (human embryonic kidney cell line obtained from the Korean Cell Line Bank (KCLB, Seoul, Korea), H1299 (lung cancer cell line from the American Type Culture Collection (ATCC, Manassas, VA, USA) and MDA-MB-231 (breast cancer cell line from KCLB). For the luciferase assay, cells

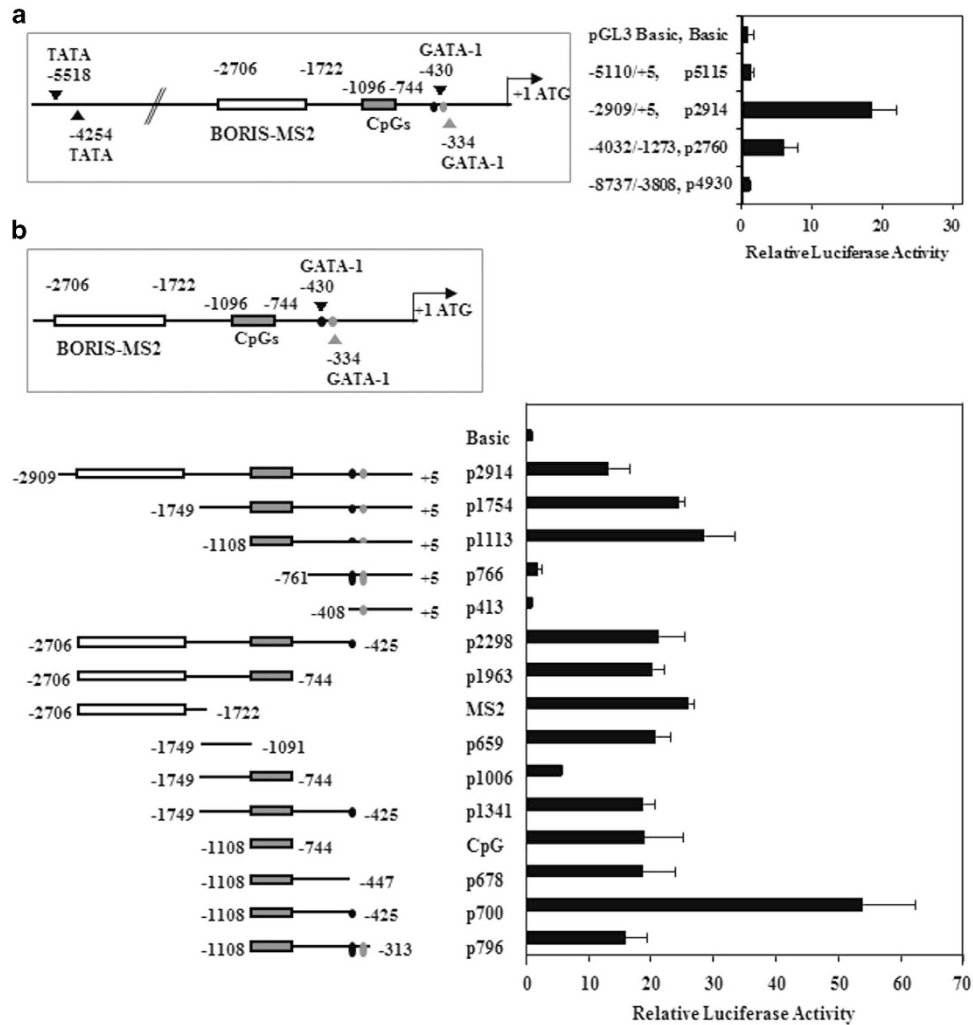


Figure 2 Activity analysis of the *BORIS* promoter. (a) Structure of the ~6-kb upstream region of *BORIS*. The figure in the box shows the putative binding sites for TATA and GATA-1, the minisatellite region (*BORIS*-MS2) and the CpG islands. The position of the CCAAT box is shown in the figure. The results of the luciferase assay with the empty control vector and the *BORIS* promoter vectors with four different regions are shown on the right. (b) Expression of *BORIS* promoter-luciferase reporters. The upper boxed panel shows the positions of *BORIS*-MS2, the CpG islands and the putative binding sites for GATA-1 in the p2914 vector, which represents the highest expression level in the above study. The structures of the *BORIS* luciferase reporter vectors are shown on the left. The results of the luciferase assay are shown on the right. Firefly luciferase activities of at least four independent experiments were calculated after normalization to *Renilla* luciferase values.

(1×10^5) were seeded in 12-well plates, cultured overnight and transfected with the *BORIS* promoter-luciferase plasmids (0.5 μ g per well) using the FuGENE6 transfection reagent (Roche Diagnostics, Indianapolis, IN, USA) at a ratio of DNA/FuGENE6 of 1:3. Analysis of the cells was performed using a dual-luciferase reporter assay system (Promega) 48 h after completion of the transfection procedure. Firefly luciferase activities were normalized to *Renilla* luciferase activity, and activity was expressed in relative luciferase units to reflect the promoter activity. Triplicate transfections of each construct were tested for each experiment, and the final results were calculated from four independent experiments.

Statistical analysis

Regression analyses were carried out to determine the odds ratios (ORs) of association between the control and case groups. ORs were predicted using the natural logarithm and its s.e. Where relevant,

a χ^2 -test was performed with one degree of freedom for statistical significance. Differences were regarded as significant with confidence intervals (CIs) of 95%. All tests were two-sided, with $P < 0.05$ considered statistically significant. Statistical analyses were performed using MS Excel with CHITEST and R statistical software (v2.5.1, www.r-project.org) with χ^2 -test for the χ^2 -calculation values. The Kaplan–Meier plot was used in the R program version 2.10.0 (<https://www.r-project.org/>).

RESULTS

BORIS expression and methylation status in lung cancer tissues

BORIS gene expression was examined in 20 cell lines, which included normal fibroblast, osteosarcoma, melanoma, breast cancer, prostate cancer, colon cancer, lung cancer and gastric cancer cell lines from a previous study.⁶ Compared with the

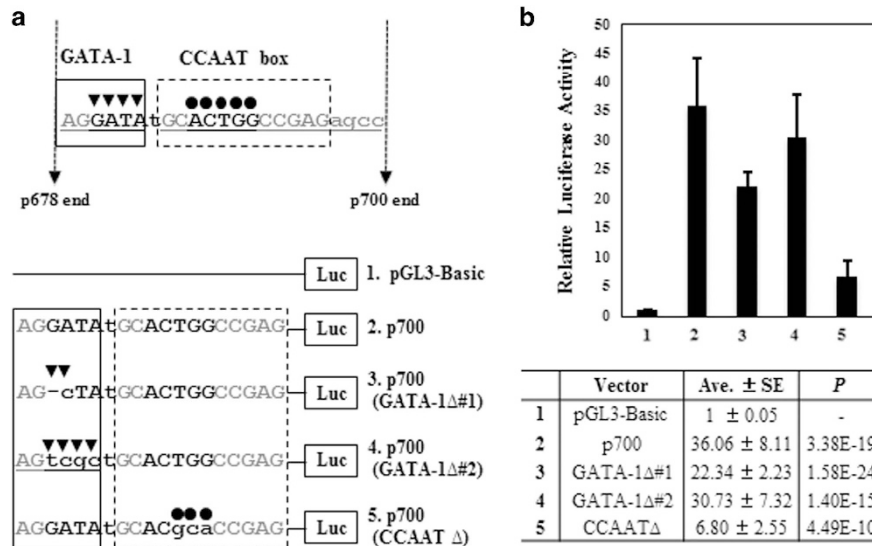


Figure 3 Analysis of the core promoter of *BORIS*. (a) The upper panel shows the positions of putative binding sites for GATA-1 and the CCAAT box in the p700 vector, which had the highest expression in Figure 2b. This region represents the difference between the p678 and p700 vectors. The wild-type construct (p700) contained putative binding sites for GATA-1 and the CCAAT box, whereas the mutant constructs (GATA-1 Δ #1, GATA-1 Δ #2 and CCAAT Δ) were mutated at the GATA-1 site or the CCAAT box in the p700 vector. (b) Luciferase activity of each vector is shown in the upper panel. The lower table shows the relative luciferase activity compared with control vector as the mean \pm s.e. and the *P*-values of at least four independent experiments.

normal cells, various cancer cells, including osteosarcoma, melanoma, breast cancer, prostate cancer, colon cancer, lung cancer and gastric cancer showed abnormal expression of *BORIS*. Among the lung cancer cell lines (A549, H358, H460, H1299), H1299 cells showed the highest expression.⁶ Then, we investigated *BORIS* gene expression in lung tumor tissues (Figure 1). In 10 lung tumor tissues, the *BORIS* levels were different; the lowest levels were observed in 2 cancer tissues (T2 and T10), moderately increased levels were observed in 4 cancer tissues (T1, T3, T4 and T5) and the highest levels were detected in 4 cancer tissues (T6, T7, T8 and T9) at the same level as the H1299 cancer cell line (Figure 1a). To analyze the possible relationship between *BORIS* gene activity and epigenetic changes in lung cancer, we assessed DNA methylation levels in these tumor tissues. As previously reported in gastric cancer tissues,⁶ methylation (89.1–95.7%) of the *BORIS* promoter region in lung cancer tissues was almost the same despite the different degree of *BORIS* expression (Figure 1b). The lung cancer cell line H1299 also showed similar levels of methylation (94%), but cells treated with 5-azad-C (1 and 5 μ M) showed decreased methylation levels (31.7–47%). We next extensively analyzed the *BORIS* promoter region to elucidate the regulation of this gene during tumorigenesis.

Cloning of the 5'-regulatory region of the *BORIS* gene and characterization of transcriptional activities of the promoter
We cloned the genomic fragment containing the entire *BORIS* gene region (Supplementary Figure S1) from human genomic DNA using the TAR cloning method.¹⁸ Two clones containing a \sim 70-kb genomic fragment were isolated and found to carry the entire *BORIS* gene (Supplementary Figures S1B and C).

The YAC clone was identified by screening a mini-YAC library with PCR using a set of diagnostic primers (Supplementary Table S1). The YAC includes the *BORIS*-coding region, the \sim 30 kb 5' upstream region and the \sim 8 kb 3' flanking regions. TAR cloning of *BORIS* in yeast greatly simplifies PCR amplification of the promoter regions enriched by repeated and CG-rich sequences, which are poorly amplified from human genomic DNA. We constructed the various reporter vectors (see Materials and Methods, Supplementary Table S1, Figures 2,3,4) using this clone (*BORIS* I, Supplementary Figures S1B and C),

Transient transfection of 293T cells with the cloned -2909/+5 genomic fragment (p2914) using the luciferase reporter construct resulted in a significant ($P < 0.001$) increase in promoter activity, which was > 18 -fold greater than the activity observed with the pGL3-Basic vector (Figure 2a). In contrast, a larger 5115 bp genomic fragment (p5115; -5110/+5) of the *BORIS* 5'-regulatory region cloned into the pGL3-Basic plasmid and transfected in 293T cells did not show any promoter activity (Figure 2a).

To determine the core *BORIS* promoter, we performed 5'-deletion analysis to determine the minimal region required for *BORIS* promoter activity (Figure 2b). The promoter activity of each of the deletion constructs was assessed by measuring luciferase activities in transiently transfected 293T cells. The results showed that luciferase activity from the shorter 700-bp promoter fragment (-1108/-425, p700) had approximately threefold greater activity than the full-length fragment (p2914); this p700 vector contained the CpG and GATA-1 regions. This observation suggests the possible existence of a strong suppressor element between -2909 and -1108 nt of the

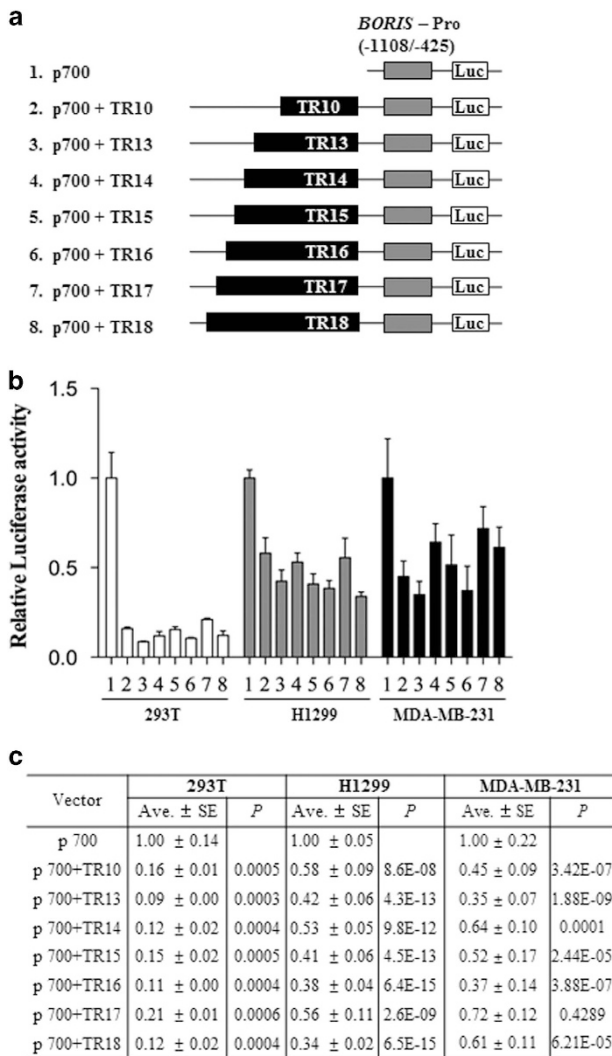


Figure 4 Effect of allelic types of *BORIS-MS2* in *BORIS* promoter luciferase constructs. **(a)** The structures of the p700 (#1) and TR (*BORIS-MS2*) reporter constructs. The gray square indicates the *BORIS* p700 promoter region, and the open square represents the open reading frame of luciferase. The black squares include the *BORIS-MS2* polymorphic regions; seven different sizes of *BORIS-MS2* (10–18 copies) were inserted in the p700 plasmid. **(b)** The effects of *BORIS-MS2* on gene expression in the luciferase reporter system. Three different cell lines (293T, H1299 and MDA-MB-231) were transfected with eight different vectors. **(c)** The table shows the relative luciferase activity compared with the p700 vector as the mean \pm s.e. and the *P*-values of four independent experiments.

BORIS 5'-regulatory region, that is, the minisatellite region (Figure 2b).

A previous study¹¹ reported three transcription start sites in the *BORIS* sequence—1447, 899 and 658 bp upstream of the first ATG codon. This observation suggests the presence of three alternative promoter regions that may have cell type-specific activity. We constructed three expression cassettes (PrA, PrB and PrC) containing the predicted promoters¹¹ and compared levels of luciferase expression in HEC293 cells (Supplementary Figure S2). The p700 vector containing the

CpG and GATA-1 regions showed the highest promoter activity, which was over 60-fold greater than the activity observed with the pGL3-Basic vector (Supplementary Figure S2).

To obtain information on putative regulatory sites in the p700 promoter region of the *BORIS* gene, we analyzed the DNA fragment (–1108/–425). As shown in Figure 2, a deletion of 23 bp from the 3'-end of the p700 vector (vector p678) resulted in a dramatic reduction in promoter activity. The deleted region (–447/–425 region, Figure 2) includes the GATA-1-binding site and the CCAAT box.¹¹ To assess the involvement of these putative regulatory sites in regulating the promoter activity of the *BORIS* gene, we introduced specific mutations in the core sequences of these sites and examined the effect of these mutations on promoter activity (Figure 3a). The results show that mutating these elements led to a significant ($P < 0.01$) reduction in promoter activity compared with the p700 promoter (Figure 3b). These results suggest that the region including the GATA-1 site and the CCAAT box has a role in activating *BORIS* expression.

Effect of minisatellites (*BORIS-MS2*) on *BORIS* expression

The comparative promoter activities of p2914 and p700 in Figure 2b suggests the presence of a suppressor element located in the –2909 to –1108 region; this region corresponds to the minisatellites of *BORIS-MS2*. Therefore, we constructed seven different vectors (Figures 4a; p700+TRs) consisting of the p700 fragment and the *BORIS-MS2* minisatellites (TR10, TR13, TR14, TR15, TR16, TR17 and TR18), which were genotyped using genomic DNA from lung cancer cases (Figure 5). For all seven constructs harboring minisatellites, the luciferase activities were significantly reduced in 293T cells compared with the p700 reporter vector (5- to 10-fold reductions) (Figure 4b). Notably, when the same *BORIS-MS2*-containing constructs were transfected into lung cancer cells (H1299) or breast cancer cells (MDA-MB-231), we observed only a approximately twofold reduction in promoter activity (Figures 4b and c). This result suggests that the constructs containing the minisatellites showed different promoter activity in normal and cancer cell lines. In addition, we also constructed seven variants of the p2298 vector, which included the p700 promoter region, the region between the p700 promoter and *BORIS-MS2*, and different fragments of *BORIS-MS2* (10–18 copies of a 56 bp repeat) (Supplementary Figure S3). These constructs were also transfected into three different cell lines, and similar results were observed. Therefore, *BORIS-MS2* minisatellites may have a role in aberrant activation of *BORIS* during carcinogenesis. Next, we investigated a possible link between the minisatellite variants and lung cancer in a Korean population.

Analysis of minisatellite variants in lung cancer patients

Several lung cell lines and many lung tumor specimens express high levels of *BORIS*.^{6,8} Because *BORIS* is a potential new marker for the detection of lung cancer, we investigated whether *BORIS-MS2* may correlate with lung cancer development. In a previous study, *BORIS-MS2* was found to possess

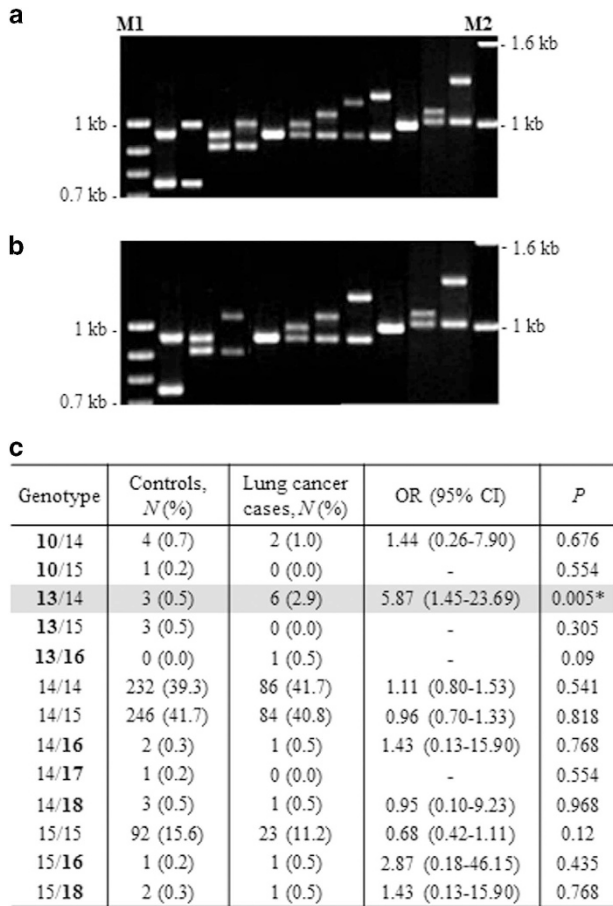


Figure 5 Haplotyping of *BORIS-MS2* in cancer-free controls and lung cancer patients. **(a)** Electrophoretic patterns of PCR products of *BORIS-MS2* in controls. Seven *BORIS-MS2* alleles and 12 haplotype patterns were detected in DNA from 590 cancer-free female controls. **(b)** Electrophoretic patterns of PCR products of *BORIS-MS2* in lung cancer patients. Six *BORIS-MS2* alleles and 10 haplotype patterns were detected in the DNA from 206 patients with lung cancer. The first and last lanes correspond to a 100-bp (M1; Invitrogen) and a 1-kb size marker (M2; Invitrogen). **(c)** Allelic genotypes and frequency in controls and cases with lung cancer. Bold numbers represent the rare alleles of *BORIS-MS2*. * $P < 0.05$.

genetically variable alleles.^{6,12} In this study, we investigated whether these alleles are linked to activation of *BORIS* during tumorigenesis. For this purpose, the frequency of polymorphic *BORIS-MS2* alleles was compared between controls and patients with lung cancer in a Korean population (Table 1). This analysis included PCR amplification of the minisatellite from human genomic DNA samples isolated from 590 cancer-free controls and 206 patients with lung cancer (Table 1). *BORIS-MS2* had 12 types of haploid patterns in controls (Figure 5a) and 10 types of haploid patterns in cancer cases (Figure 5b). Figure 5c summarizes the frequency of haploid patterns for *BORIS-MS2* between the controls and the lung cancer groups. Among the identified haploid patterns, the 13/14 genotype of *BORIS-MS2* (corresponding to 13 copies of the satellite repeat on 1 chromosome and 14 copies of the satellite repeat on another chromosome) was significantly more

frequent in cases (2.9%) than in controls (0.5%) ($P < 0.005$), and for the frequency of the allelic distribution (Table 2) 13 repeated alleles of *BORIS-MS2* were significantly more frequent in cases than in controls (0.5 vs 1.7%). Analysis of these data revealed a significant association between 13 tandem repeat alleles and the odds for lung cancer (*BORIS-MS2* and lung cancer OR: 3.38, CI: 1.13–10.12; $P = 0.021$).

For further analysis, each *BORIS-MS2* allele was divided into common or rare alleles (the frequency for rare alleles was $< 1\%$) based on their frequency in the control population.¹² Seven alleles of *BORIS-MS2* were grouped into 2 common alleles (14 and 15 repeats) and 5 rare alleles (10, 13, 16, 17 and 18 repeats) in this study (Table 3). There was a significant difference in the frequency of rare alleles found in controls and cancer cases (OR: 2.04, CI: 1.02–4.08; $P = 0.039$). Specifically, the frequency of rare alleles in female cases with lung cancer showed a statistically significant difference between controls and cases (OR = 3.14; $P = 0.047$). Furthermore, we divided the rare alleles into short (10 and 13) and long (16, 17 and 18) groups based on their tandem repeat lengths (Table 4). The analysis of the short rare alleles (10 and 13 repeats) group, comparing controls and cases, also showed an association, with a relative lung cancer OR of 2.37 ($P = 0.049$). There was a more significant difference in the frequency of short rare alleles between male controls and male cases (OR = 3.86; $P = 0.041$) (Table 4).

We used clinicopathological information obtained in 2009 from the Korea University Anam Hospital in Korea¹⁹ and between 2002 and 2007 from Dong-A University and Chungbuk National University Hospital in Korea (Supplementary Table S2). Tumors, nodes and distant metastasis stages were analyzed according to the World Health Organization (WHO) system. Lung tumors were categorized into the appropriate class, and we then assessed the frequency of each class in the entire cancer group and the rare allele group. There was no association between short rare alleles and cancer classification corresponding to stage, T stage, N stage, M stage or cell type.

DISCUSSION

BORIS is a CTA gene family member, and its transcription is abnormally activated in various tumors and cancer cells.^{2,4–6,20,21} It was also reported that the expression of *BORIS* is predominantly controlled by DNA methylation¹ and that *BORIS* mediates the epigenetic regulation of other genes (*NY-ESO-1*, *NOTCH3*, *hTERT* and so on) in cancer cells.^{8,22,23} In our previous study, however, no correlation between methylation status and *BORIS* gene expression was detected in tissues from gastric cancer.⁶ Similar data were obtained in this work based on analysis of tissues from lung cancer. Based on these results, we suggest that the expression of *BORIS* is regulated by more complex epigenetic mechanisms.

To elucidate the regulation of *BORIS* beyond DNA methylation, we physically characterized the entire *BORIS* locus. A CpG island (–1096 to –762 from the first ATG codon) and two minisatellites (variable number of tandem repeats; *BORIS-MS1* and *BORIS-MS2*) were identified through the

Table 2 Comparison of allelic sizes and frequency of *Boris-MS2*

TR	Size (bp)	Controls, N (frequency)			Lung cancer cases, N (frequency)			OR (95% CI)	P-value
		Male N = 652	Female N = 528	Total N = 1180	Male N = 342	Female N = 70	Total N = 412		
10	785	1 (0.002)	4 (0.008)	5 (0.004)	1 (0.003)	1 (0.014)	2 (0.005)	1.15 (0.22–5.93)	0.871
13	935	2 (0.003)	4 (0.008)	6 (0.005)	5 (0.015)	2 (0.029)	7 (0.017)	3.38 (1.13–10.12)	0.021*
14	990	392 (0.601)	331 (0.627)	723 (0.613)	219 (0.640)	47 (0.671)	266 (0.646)	1.15 (0.91–1.45)	0.236
15	1050	250 (0.383)	187 (0.354)	437 (0.370)	113 (0.330)	19 (0.271)	132 (0.320)	0.80 (0.63–1.01)	0.069
16	1100	2 (0.003)	1 (0.002)	3 (0.003)	3 (0.009)	—	3 (0.007)	2.88 (0.58–14.31)	0.177
17	1160	1 (0.002)	—	1 (0.001)	—	—	—	—	0.554
18	1220	4 (0.006)	1 (0.002)	5 (0.004)	1 (0.003)	1 (0.014)	2 (0.005)	1.15 (0.22–5.93)	0.871

Abbreviations: CI, confidence interval; OR, odds ratio; TR, tandem repeat.
* $P < 0.05$.

Table 3 The frequency of rare *Boris-MS2* alleles in controls and lung cancer cases

	No. of alleles	Common alleles			Rare alleles						OR (95% CI)	P-value
		14	15	Total	10	13	16	17	18	Total		
Controls												
Male	652 (%)	392 (60.1)	250 (38.4)	642 (98.5)	1	2	2	1	4	10 (1.5)	1 (reference)	—
Female	528 (%)	331 (62.7)	187 (35.4)	518 (98.1)	4	4	1	0	1	10 (1.9)	1 (reference)	—
Total	1180 (%)	723 (61.3)	437 (37.0)	1160 (98.3)	5	6	3	1	5	20 (1.7)	1 (reference)	—
Lung cancer cases												
Male	342 (%)	219 (64.0)	113 (33.0)	332 (97.0)	1	5	3	0	1	10 (3.0)	1.93 (0.80–4.69)	0.138
Female	70 (%)	47 (67.2)	19 (27.1)	66 (94.3)	1	2	0	0	1	4 (5.7)	3.14 (0.96–10.29)	0.047*
Total	412 (%)	266 (64.6)	132 (32.0)	398 (96.6)	2	7	3	0	2	14 (3.4)	2.04 (1.02–4.08)	0.039*

Abbreviations: CI, confidence interval; OR, odds ratio.
* $P < 0.05$.

Table 4 The frequency of short rare *Boris-MS2* alleles in controls and lung cancer cases

	No. of alleles	Common+long rare alleles						Short rare alleles			OR (95% CI)	P-value
		14	15	16	17	18	Total	10	13	Total		
Controls												
Male	652 (%)	392 (60.1)	250 (38.3)	2 (0.3)	1 (0.2)	4 (0.6)	649 (99.5)	1	2	3 (0.5)	1 (reference)	—
Female	528 (%)	331 (62.7)	187 (35.4)	1 (0.2)	0 (0.0)	1 (0.2)	520 (98.5)	4	4	8 (1.5)	1 (reference)	—
Total	1180 (%)	723 (61.3)	437 (37.0)	3 (0.3)	1 (0.1)	5 (0.4)	1169 (99.1)	5	6	11 (0.9)	1 (reference)	—
Lung cancer cases												
Male	342 (%)	219 (64.0)	113 (33.0)	3 (0.9)	0 (0.0)	1 (0.3)	336 (98.2)	1	5	6 (1.8)	3.86 (0.96–15.54)	0.041*
Female	70 (%)	47 (67.2)	19 (27.1)	0 (0.0)	0 (0.0)	1 (1.4)	67 (95.7)	1	2	3 (4.3)	2.91 (0.75–11.24)	0.105
Total	412 (%)	266 (64.6)	132 (32.0)	3 (0.7)	0 (0.0)	2 (0.5)	403 (97.8)	2	7	9 (2.2)	2.37 (0.98–5.77)	0.049*

Abbreviations: CI, confidence interval; OR, odds ratio.
* $P < 0.05$.

characterization of the genomic DNA sequence upstream of the gene.¹² In this study, we examined the ~5.5-kb region upstream of the start codon of *BORIS*. A set of luciferase reporter cassettes carrying fragments from the upstream region

of various sizes were constructed and tested (Figure 2). The highest luciferase activity was detected with a 700-bp promoter fragment (–1108/–425, p700) which contained CpG islands, a GATA-1 site and CCAAT box regions. This finding is in

agreement with a previous study,¹¹ where three transcription start sites were mapped within the same 5'-upstream region. However, our analysis revealed a potential suppressor between -2909 and -1108 nt of the *BORIS* 5'-regulatory region containing the minisatellite region of *BORIS-MS2* (Figure 2b). This finding suggests that the transcriptional regulation of *BORIS* is more complicated and that other types of regulation may exist in addition to DNA methylation. Moreover, there are also several reports indicating that minisatellite polymorphisms can influence gene expression.^{16,17,24-28}

In separate experiments, we investigated how polymorphic variants of *BORIS-MS2* minisatellite affect the transcriptional activity of *BORIS* in cell lines using a luciferase reporter system (Figure 4). In 293T normal cell lines, all seven minisatellite constructs (TR10, 13, 14, 15, 16, 17 and 18) significantly decreased (5–10 times reduction) the activity of the p700 *BORIS* promoter, but there was no clear relationship between the tandem repeat copy number and level of gene expression. In principle, this result may be suggestive of the restriction of *in vivo* experiments. However, when the same constructs were transfected in lung cancer cells (H1299) or breast cancer cells (MDA-MB-231), the promoter activity was reduced only approximately two times compared with the control. This result suggests that the expression cassettes carrying *BORIS-MS2* are controlled in a different way in normal and cancer cell lines. Similar results for transcriptional regulation by minisatellite fragments were previously reported with *H-Ras*, *IL-1 α* , *IGF2*, *MUC6* and *TERT* genes.^{16,17,24-26} Minisatellites of *IL-1 α* , *MUC6* and *TERT* may control gene expression in cancer cell lines, although these are located in the intronic regions of genes.^{16,17} We suggest that *BORIS-MS2*, located in the promoter region, differentially regulates the gene transcription in cancer cells and normal cells. This suggests a potential association between *BORIS-MS2* and cancer. For example, specific alleles of *BORIS-MS2* may affect susceptibility to lung cancer. In this study, we found a significantly elevated frequency of rare alleles of the minisatellite in lung cancer cases compared with cancer-free controls. In addition, short rare minisatellite alleles were associated with higher susceptibility in lung cancer patients. A similar result has been shown in a report examining the relationship between short rare alleles of *BORIS-MS2* and breast cancer in a younger group.¹² Other studies on genetic susceptibility due to specific minisatellites have also been reported.^{17,29-31}

It is commonly accepted that genomic instability has a crucial role in the accumulation of genetic alterations that are responsible for cancer cell development.³² It has also been reported that genomic instability in minisatellite regions is related to loss of heterozygosity and rearrangements in cancer tissues.^{33,34} Thus, the *BORIS-MS2* minisatellites may be involved in tumorigenesis not only via gene regulation but also via genomic instability. Further analysis is required to elucidate the correlation between rare minisatellite alleles and the susceptibility to lung cancer observed in this study.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We gratefully acknowledge the patients, families and their caregivers for their willing participation in this project and for providing consent for the use of their information in this study. The study design and procedures were approved by the Committee of Bioethics at Dong-A University (#IRB-06-10-02 & IRB-07-10-7; Busan, Korea). This research was supported by the Mid-Career Researcher Program through the National Research Foundation of Korea (NRF) grant (NRF-2013R1A2A2A04008115) funded by the Korea government (MEST). This work was supported by the Brain Busan 21 Project in 2015. SIK was supported by the Bio-Synergy Research Project (NRF-2014M3A9C4066461) of the Ministry of Science, ICT and Future Planning through the National Research Foundation and the Korea Basic Science Institute research program (D35402).

- 1 Loukinov DI, Pugacheva E, Vatolin S, Pack SD, Moon H, Chernukhin I *et al*. BORIS, a novel male germ-line-specific protein associated with epigenetic reprogramming events, shares the same 11-zinc-finger domain with CTCF, the insulator protein involved in reading imprinting marks in the soma. *Proc Natl Acad Sci USA* 2002; **99**: 6806–6811.
- 2 Simpson AJ, Caballero OL, Jungbluth A, Chen YT, Old LJ. Cancer/testis antigens, gametogenesis and cancer. *Nat Rev Cancer* 2005; **5**: 615–625.
- 3 D'Arcy V, Pore N, Docquier F, Abdullaev ZK, Chernukhin I, Kita GX *et al*. BORIS, a paralogue of the transcription factor, CTCF, is aberrantly expressed in breast tumours. *Br J Cancer* 2008; **98**: 571–579.
- 4 Vatolin S, Abdullaev Z, Pack SD, Flanagan PT, Custer M, Loukinov DI *et al*. Conditional expression of the CTCF-paralogous transcriptional factor BORIS in normal cells results in demethylation and derepression of MAGE-A1 and reactivation of other cancer-testis genes. *Cancer Res* 2005; **65**: 7751–7762.
- 5 Risinger JI, Chandramouli GV, Maxwell GL, Custer M, Pack S, Loukinov D *et al*. Global expression analysis of cancer/testis genes in uterine cancers reveals a high incidence of BORIS expression. *Clin Cancer Res* 2007; **13**: 1713–1719.
- 6 Yoon SL, Roh YG, Lee SH, Kim SH, Kim MC, Kim SJ *et al*. Analysis of promoter methylation and polymorphic minisatellites of BORIS and lack of association with gastric cancer. *DNA Cell Biol* 2011; **30**: 691–698.
- 7 D'Arcy V, Abdullaev ZK, Pore N, Docquier F, Torrano V, Chernukhin I *et al*. The potential of BORIS detected in the leukocytes of breast cancer patients as an early marker of tumorigenesis. *Clin Cancer Res* 2006; **12**: 5978–5986.
- 8 Hong JA, Kang Y, Abdullaev Z, Flanagan PT, Pack SD, Fischette MR *et al*. Reciprocal binding of CTCF and BORIS to the NY-ESO-1 promoter coincides with derepression of this cancer-testis gene in lung cancer cells. *Cancer Res* 2005; **65**: 7763–7774.
- 9 Hoffmann MJ, Muller M, Engers R, Schulz WA. Epigenetic control of CTCFL/BORIS and OCT4 expression in urogenital malignancies. *Biochem Pharmacol* 2006; **72**: 1577–1588.
- 10 Woloszynska-Read A, James SR, Link PA, Yu J, Odunsi K, Karpf AR. DNA methylation-dependent regulation of BORIS/CTCF expression in ovarian cancer. *Cancer Immunol* 2007; **7**: 21.
- 11 Renaud S, Pugacheva EM, Delgado MD, Braunschweig R, Abdullaev Z, Loukinov D *et al*. Expression of the CTCF-paralogous cancer-testis gene, brother of the regulator of imprinted sites (BORIS), is regulated by three alternative promoters modulated by CpG methylation and by CTCF and p53 transcription factors. *Nucleic Acids Res* 2007; **35**: 7372–7388.
- 12 Yoon SL, Kim DC, Cho SH, Lee SY, Chu IS, Heo J *et al*. Susceptibility for breast cancer in young patients with short rare minisatellite alleles of BORIS. *BMB Rep* 2010; **43**: 698–703.
- 13 Bresnick EH, Katsumura KR, Lee HY, Johnson KD, Perkins AS. Master regulatory GATA transcription factors: mechanistic principles and emerging links to hematologic malignancies. *Nucleic Acids Res* 2012; **40**: 5819–5831.

- 14 Dolfini D, Mantovani R. Targeting the Y/CCAAT box in cancer: YB-1 (YBX1) or NF-Y? *Cell Death Differ* 2013; **20**: 676–685.
- 15 Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin* 2011; **61**: 69–90.
- 16 Kwon JA, Lee SY, Ahn EK, Seol SY, Kim MC, Kim SJ *et al*. Short rare MUC6 minisatellites-5 alleles influence susceptibility to gastric carcinoma by regulating gene. *Hum Mutat* 2010; **31**: 942–949.
- 17 Yoon SL, Jung SI, Do EJ, Lee SR, Lee SY, Chu IS *et al*. Short rare hTERT-VNTR2-2nd alleles are associated with prostate cancer susceptibility and influence gene expression. *BMC Cancer* 2010; **10**: 393.
- 18 Kouprina N, Larionov V. Selective isolation of genomic loci from complex genomes by transformation-associated recombination cloning in the yeast *Saccharomyces cerevisiae*. *Nat Protoc* 2008; **3**: 371–377.
- 19 Sung JS, Whang YM, Park KH, Ryu JS, Choi JG, Seo JH *et al*. No association between promoter polymorphism of STK11 gene and lung cancer risk in the Korean population. *Cancer Res Treat* 2009; **41**: 211–217.
- 20 Cheema Z, Hari-Gupta Y, Kita GX, Farrar D, Seddon I, Corr J *et al*. Expression of the cancer-testis antigen BORIS correlates with prostate cancer. *Prostate* 2014; **74**: 164–176.
- 21 Martin-Kleiner I. BORIS in human cancers—a review. *Eur J Cancer* 2012; **48**: 929–935.
- 22 Zampieri M, Ciccarone F, Palermo R, Cialfi S, Passananti C, Chiaretti S *et al*. The epigenetic factor BORIS/CTCF regulates the NOTCH3 gene expression in cancer cells. *Biochim Biophys Acta* 2014; **1839**: 813–825.
- 23 Renaud S, Loukinov D, Alberti L, Vostrov A, Kwon YW, Bosman FT *et al*. BORIS/CTCF-mediated transcriptional regulation of the hTERT telomerase gene in testicular and ovarian tumor cells. *Nucleic Acids Res* 2011; **39**: 862–873.
- 24 Trepicchio WL, Krontiris TG. Members of the rel/NF-kappa B family of transcriptional regulatory proteins bind the HRAS1 minisatellite DNA sequence. *Nucleic Acids Res* 1992; **20**: 2427–2434.
- 25 Paquette J, Giannoukakis N, Polychronakos C, Vafiadis P, Deal C. The INS 5' variable number of tandem repeats is associated with IGF2 expression in humans. *J Biol Chem* 1998; **273**: 14158–14164.
- 26 Bailly S, Israel N, Fay M, Gougerot-Pocidalo MA, Duff GW. An intronic polymorphic repeat sequence modulates interleukin-1 alpha gene regulation. *Mol Immunol* 1996; **33**: 999–1006.
- 27 Fiskerstrand CE, Lovejoy EA, Quinn JP. An intronic polymorphic domain often associated with susceptibility to affective disorders has allele dependent differential enhancer activity in embryonic stem cells. *FEBS Lett* 1999; **458**: 171–174.
- 28 Nakamura Y, Koyama K, Matsushima M. VNTR (variable number of tandem repeat) sequences as transcriptional, translational, or functional regulators. *J Hum Genet* 1998; **43**: 149–152.
- 29 Ahn EK, Kim WJ, Kwon JA, Choi PJ, Kim WJ, Sunwoo Y *et al*. Variants of MUC5B minisatellites and the susceptibility of bladder cancer. *DNA Cell Biol* 2009; **28**: 169–176.
- 30 Ahn MH, Bae KB, Kwon JA, Choi HJ, Lee SR, Kim SH *et al*. Association of MUC6-minisatellite variants with susceptibility to rectal carcinoma. *Mol Biol Rep* 2013; **40**: 303–308.
- 31 Seol SY, Lee SY, Kim YD, Do EJ, Kwon JA, Kim SI *et al*. Minisatellite polymorphisms of the SLC6A19: susceptibility in hypertension. *Biochem Biophys Res Commun* 2008; **374**: 714–719.
- 32 Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. *Nature* 1998; **396**: 643–649.
- 33 Jeong YH, Kim MC, Ahn EK, Seol SY, Do EJ, Choi HJ *et al*. Rare exonic minisatellite alleles in MUC2 influence susceptibility to gastric carcinoma. *PLoS ONE* 2007; **2**: e1163.
- 34 Szutorisz H, Palmqvist R, Roos G, Stenling R, Schorderet DF, Reddel R *et al*. Rearrangements of minisatellites in the human telomerase reverse transcriptase gene are not correlated with its expression in colon carcinomas. *Oncogene* 2001; **20**: 2600–2605.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Supplementary Information accompanies the paper on Experimental & Molecular Medicine website (<http://www.nature.com/emm>)