



## HMGB3 characterization in gastric cancer

Y. Gong<sup>1\*</sup>, Y. Cao<sup>2\*</sup>, L. Song<sup>3</sup>, J. Zhou<sup>4</sup>, C. Wang<sup>1</sup> and B. Wu<sup>1</sup>

<sup>1</sup>Gastrointestinal Department of Southern Building, PLA General Hospital, Beijing, China

<sup>2</sup>Changhai Hospital, Second Military Medical University, Shanghai, China

<sup>3</sup>Michael Smith Laboratories, East Mall, University of British Columbia, Vancouver, BC, Canada

<sup>4</sup>Department of Oncology, Provincial Hospital Affiliated to Shandong University, Jinan, Shandong, China

\*These authors contributed equally to this study.

Corresponding author: B. Wu

E-mail: benyanwu@hotmail.com

Genet. Mol. Res. 12 (4): 6032-6039 (2013)

Received March 27, 2013

Accepted August 10, 2013

Published December 2, 2013

DOI <http://dx.doi.org/10.4238/2013.December.2.1>

**ABSTRACT.** Gastric cancer is a major health problem worldwide; it is the second most common cause of cancer death in the world. Recent studies indicate that the high-mobility group (HMG) of chromosomal proteins is associated with cancer progression. However, HMGB3 has been little studied. We analyzed the co-expression network between HMGB3 and differentially-expressed genes in the GSE17187 database, identifying the relevant transcription factors, and the conserved domain of HMGB3 to understand the underlying regulation mechanisms involved in gastric cancer. Thirty-one relationships between 11 differentially-expressed genes were included in a co-expression network; many of these genes have been identified as related to cancer, including TBX5 and TFR2. Further analysis identified nine transcription factors, these being GATA3, MZF1, GATA1, GATA2, SRY, REL, NFYB, NFYC, and NFYA,

which could interact with HMGB3 to regulate target gene expression and consequently regulate gastric cancer cell proliferation, migration and invasion. The HMG-box domain was very similar in various species, with only a few amino acid changes, indicating conserved functions in HMG-box. This information helps to provide insight into the molecular mechanisms of HMGB3 in human gastric cancer.

**Key words:** HMGB3; Gastric cancer; HMG-box

## INTRODUCTION

Gastric cancer is a major health problem as it constitutes the second leading cause of cancer-related deaths (Saif et al., 2010). Recently, gastric cancer has attracted much attention from epidemiologic investigators, particularly with the emergence of *Helicobacter pylori* as a risk factor for the condition (Krejs, 2010). *H. pylori* infection may result in the development of gastric cancer through different mechanisms, including induction of hyperproliferation or of direct changes on the DNA of the host, such as oxidative damage, methylation, activation of oncogenes, inactivation of tumor suppressor genes, microsatellite instability, and chromosomal instability (Machado et al., 2010, Nobili et al., 2011). Hereditary gastric cancer is suggested to be associated with germline mutation of E-cadherin (Mayrbaeurl et al., 2010).

HMGB3, also known as HMG2a, is an X-linked member of the high-mobility group (HMG) superfamily of HMG proteins and is classified with HMGB1 and HMGB2 into the HMG-box subfamily (Nemeth et al., 2003). The 80% identity between Hmg-box proteins suggests similar functions at the molecular level. HMGB1 and HMGB2 can interact with DNA and subsequently bend linear DNA, thereby facilitating nucleoprotein complex formation through alteration of local chromatin architecture (Nemeth et al., 2005), and these proteins have been implicated in the development of both benign and malignant neoplasias. HMGB1 and HMGB2 can also directly interact with DNA-binding proteins (Nemeth et al., 2006), such as Rel and p53, and regulate cancer progression.

Recent reports have demonstrated the occurrence of HMGB3 gene rearrangements and truncated proteins (loss of exons 4 and/or 5) in benign mesenchymal tumors, lipomas, and natural killer lymphomas. HMGB3 silencing may activate the endogenous cyclin A gene to regulate cell cycle and inhibit tumor growth (Hayes et al., 2006). A further study has found that HMGB3 mRNA is also expressed in small cell and non-small cell lung carcinomas. HMGB3 may play a role as a receptor for advanced glycation endproduct (RAGE) activators such as HMGB1 to be involved in cancer (Sparvero et al., 2009). HMGB3 has been identified as being specifically upregulated in chromosomal instability-colorectal cancer samples, with HMGB3 upregulation also being reflected at the mRNA level (Albrethsen et al., 2010).

The aim of our study was to analyze the co-expression network between HMGB3 and differential expression genes (DEGs) and to identify the significant transcription factors (TFs) that regulate HMGB3 according to a high score for transcription factor binding site (TFBS), and the conserved domain in HMGB3 to explain its underlying regulatory mechanism in gastric cancer.

## MATERIAL AND METHODS

### Data source

The gastric cancer dataset GSE17187 was obtained from National Center for Biotechnology Information Gene Expression Omnibus (NCBI GEO, <http://www.ncbi.nlm.nih.gov/geo/>) which is based on the Affymetrix Human Genome U133 Plus 2.0 Array. Three node positive vs three node negative intestinal type gastric carcinomas were analyzed by the limma method (Diboun et al., 2006). The original expression datasets from all conditions were extracted into expression estimates, and the linear model was then constructed. Finally, 1395 differentially expressed genes (DEGs) only with a fold-change >1.5 and  $P < 0.05$  were selected.

### Co-expression network in gastric cancer

For demonstrating the potential relationship between HMGB3 and other DEGs, the Pearson correlation coefficient (PCC) was calculated for all pair-wise comparisons of gene-expression values. The correlation whose absolute PCC were greater than 0.95 were considered to be significant. Interactions between two DEGs with PCC greater than 0.6 were then selected as candidate relationships to construct the co-expression network.

Cytoscape plugin BiNGO (Maere et al., 2005) was used to perform Gene Ontology (GO) enrichment analysis. We selected the nodes in the co-expression network as the dataset and chose 0.05 as the significance level. Moreover, hypergeometric tests were used for statistical analysis and the Benjamini and Hochberg false discovery rate (FDR) procedure was used for multiple testing correction.

### TFSEARCH using the TRANSFAC database

TFSEARCH (<http://www.cbrc.jp/research/db/TFSEARCH.html>), which predicted the TF binding sites (TFBS) based on TRANSFAC data, is directly owed to the TRANSFAC databases (Wingender, 2008). The *H. sapiens* HMGB3 upstream 2000 was downloaded from UCSC (Sanborn et al., 2011). TFBS were then predicted using TFSEARCH with the score >90 set to vertebrates.

### HMGB3 domain analysis

To find the conserved domain in HMGB3, BLAST (Altschul et al., 1997) was used to find the homologous gene in vertebrates. Finally, CLUSTAL (Thompson et al., 1994) (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>), a multiple sequence alignment tool, was used to find the conserved domain in 5 conserved species.

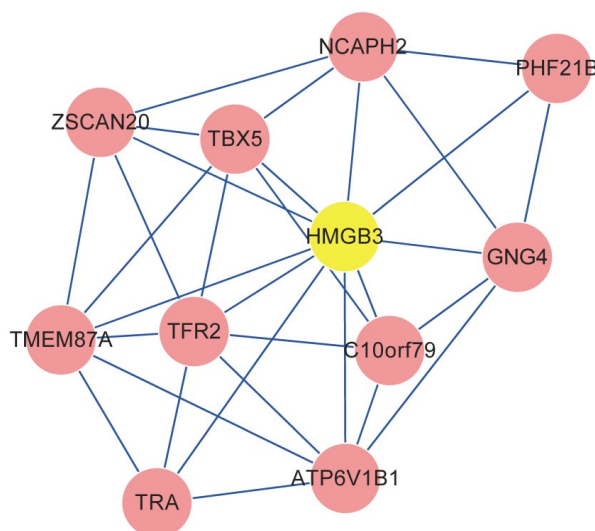
## RESULTS AND DISCUSSION

### Co-expression network

To determine the DEGs of gastric cancer, we obtained publicly available microarray

data sets GSE17187 (Ingold et al., 2010) from GEO. After microarray analysis, the DEGs with a fold change  $>1.5$  and  $P < 0.05$  were selected. A total of 1395 genes contained HMGB3 and were selected as DEGs.

To find the relationships between HMGB3 and other DEGs, the co-expressed value ( $PCC > 0.95$ ) was chosen as the threshold. Finally, 31 relationships between 11 DEGs were used to construct a co-expression network (Figure 1). Among them, T-box transcription factor 5 (TBX5) is a member of a phylogenetically conserved family of genes that share a common DNA-binding domain, the T-box. TBX5 has been found to be silenced or downregulated in colon cancer cell lines (Yu et al., 2010). Loss of gene expression is associated with promoter methylation. Re-expression of TBX5 in silenced colon cancer cell lines suppresses colony formation, cell proliferation and cell migration and induces apoptosis. TBX5 inactivated by promoter methylation may serve as a potential biomarker for the prognosis of this malignancy. In addition, TFR2 encodes a single-pass type II membrane protein, which is a member of the transferrin receptor-like family. This protein mediates cellular uptake of transferrin-bound iron, and may be involved in iron metabolism, hepatocyte function and erythrocyte differentiation. A recent study shows that intravenous administration of ascorbic acid can be used as a way to selectively kill cancer cells. The apoptotic effect of ascorbic acid on human gastric cancer cells is iron-mediated oxidative stress, which increases upregulated expression of transferrin receptor (Ha et al., 2009). Therefore, we suggest that HMGB3 may also be involved in gastric cancer cell proliferation, migration, and apoptosis by interacting with those DEGs.



**Figure 1.** Co-expression network of HMGB3. In the graph, yellow point stands for the HMGB3. All the points are the DEGs.

BINGO analysis was performed to find the significant GO terms in the network with  $P < 0.05$ . Finally, cardiac left ventricle formation, cardiac left ventricle morphogenesis, cardiac chamber formation, cardiac ventricle formation, His-Purkinje system development, and others were enriched in the network. All details are shown in Table 1.

**Table 1.** Gene Ontology (GO) term enrichment analysis.

GO-ID	P	corr P	Description
3218	9.78E-04	4.91E-02	Cardiac left ventricle formation
3214	9.78E-04	4.91E-02	Cardiac left ventricle morphogenesis
3207	1.47E-03	4.91E-02	Cardiac chamber formation
3211	1.47E-03	4.91E-02	Cardiac ventricle formation
3164	1.47E-03	4.91E-02	His-Purkinje system development
3166	1.47E-03	4.91E-02	Bundle of His development
3161	1.47E-03	4.91E-02	Cardiac conduction system development
35117	1.96E-03	4.91E-02	Embryonic arm morphogenesis
35140	1.96E-03	4.91E-02	Arm morphogenesis
60044	1.96E-03	4.91E-02	Negative regulation of cardiac muscle cell proliferation

### HMGB3 TF predicted

Through TFSEARCH analysis with a threshold score of 90 predicted in vertebrates, we found 9 TFs (namely GATA3, MZF1, GATA1, GATA2, SRY, REL, NFYB, NFYC, and NFYA) occurring in humans. Further text mining was carried out to find article numbers and to see whether these TFs play a role in cancer, especially gastric cancer, using “TFs” and “cancers” as keywords. The detailed results are shown in Table 2.

**Table 2.** PubMed search results with articles numbers.

TFs	Cancer	Gastric cancer
GATA3	200	3
MZF1	23	1
GATA1	374	0
GATA2	132	0
SRY	186	5
REL	824	3
NFYB	6	0
NFYC	3	0
NFYA	17	0

TFs = transcription factors.

GATA3 belongs to the GATA family of TFs. This protein contains two GATA-type zinc fingers and is an important regulator of T-cell development. GATA-3 is considered to be involved in Th2 response in T cells (Yang et al., 2010) and an inhibitor of IFN- $\gamma$  production. The expression of GATA-3 is markedly higher in gastric cancer patients compared to healthy individuals. Furthermore, there are higher levels of GATA-3 expression in natural killer cells treated with transforming growth factor (TGF- $\beta$ ). This result supports the hypothesis that TGF- $\beta$  is responsible for the suppression of natural killer cells from gastric cancer patients and that upregulation of the interferon (IFN)- $\gamma$  inhibitory transcription factor GATA-3 may mediate the suppression (Lindgren et al., 2011).

The SRY gene, also known as testis-determining factor (TDF), encodes a transcription factor that is a member of the high mobility group (HMG)-box family of DNA-binding proteins. A recent study reports that one SRY-box factor gene, namely the SOX gene, plays critical roles in cell fate determination, differentiation and proliferation. Exogenous expression of SOX2 suppresses cell proliferation in gastric epithelial cell lines through cell-cycle arrest and apoptosis. Downregulated expression of SOX2 may be related to gastric carcinogenesis and poor prognosis (Tsukamoto et al., 2005; Otsubo et al., 2008).

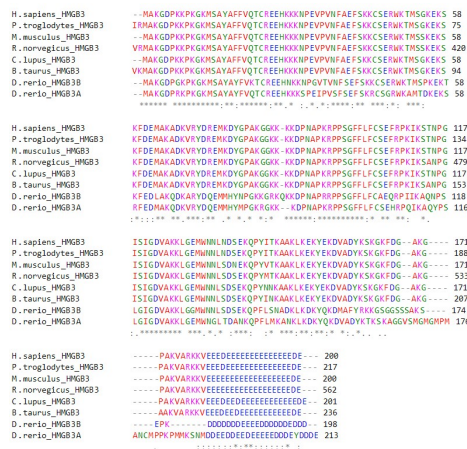
The REL gene encodes c-Rel, a transcription factor that is a member of the Rel/NFKB family, which also includes RELA, RELB, NFKB1, and NFKB2. They are assembled by dimerization of two of five subunits: p65 (RelA), c-Rel, RelB, p50/NF-κB1, and p52/NF-κB2. These proteins are related through a highly conserved N-terminal region termed the Rel domain, which is responsible for DNA binding, dimerization, nuclear localization, and binding to the NFKB inhibitor (Belguise and Sonenshein, 2007). The first clue linking NF-κB to cancer is the realization that c-rel, which is the cellular homolog of the v-rel oncogene, encodes a NF-κB subunit and that all of these proteins share the same DNA binding domain, the Rel homology domain. Not surprisingly, over-expression of normal Rel proteins promotes oncogenic transformation (Maeda and Omata, 2008).

Myeloid zinc finger 1 (MZF1) is a member of the SCAN domain family TFs that form dimers through their highly conserved SCAN motifs. Studies have shown that overexpression of MZF1 induces migration and invasion in colorectal and cervical cancer cells. In addition, MZF1 binds to the Axl promoter, transactivates promoter activity, and enhances Axl-mRNA and protein expression, which is correlated with poor prognosis in gastric cancer patients. Taken together, MZF1 induces invasion and metastasis in colorectal, cervical cancer, and gastric cancer at least in part by regulating Axl gene expression (Mudduluru et al., 2010).

Therefore, we predict that HMGB3 may interact with those TFs to bind to target gene promoters and also to regulate gastric cancer cell proliferation, migration and invasion.

### HMGB3 domain analysis

Using Blast analysis in the NCBI (Altschul et al., 1997), we screened the homologous genes of HMGB3 in vertebrates. The results showed that 8 related HMGB3 genes of 7 species were selected. In addition, ClustalW analysis was applied to find the conserved domain in these 8 HMGB3 genes of different species. The HMG-box (conserved domains) from CDD (Marchler-Bauer et al., 2011) was found in protein sequences by rpsblast searching. The results depicted that the HMG-box domain was very similar in these different species, only with few amino acids changes. This result also indicated that there may be a conserved function in HMG-box, which is found in a variety of eukaryotic chromosomal proteins and TFs (Figure 2).



**Figure 2.** ClustalW analysis part results. This figure displays the conserved domain of 8 HMGB3 related genes.



## REFERENCE

- Albrethsen J, Knol JC, Piersma SR, Pham TV, et al. (2010). Subnuclear proteomics in colorectal cancer. *Mol. Cell Proteomics* 9: 988-1005
- Altschul SF, Madden TL, Schaffer AA, Zhang J, et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25: 3389-3402.
- Belguise K and Sonenshein GE (2007). PKC $\theta$  promotes c-Rel-driven mammary tumorigenesis in mice and humans by repressing estrogen receptor alpha synthesis. *J. Clin. Invest.* 117: 4009-4021.
- Diboun I, Wernisch L, Orengo CA and Koltzenburg M (2006). Microarray analysis after RNA amplification can detect pronounced differences in gene expression using limma. *BMC Genomics* 7: 252.
- Ha YM, Park MK, Kim HJ, Seo HG, et al. (2009). High concentrations of ascorbic acid induces apoptosis of human gastric cancer cell by p38-MAP kinase-dependent up-regulation of transferrin receptor. *Cancer Lett.* 277: 48-54.
- Hayes DC, Secrist H, Bangur CS, Wang T, et al. (2006). Multigene real-time PCR detection of circulating tumor cells in peripheral blood of lung cancer patients. *Anticancer Res.* 26: 1567-1575.
- Ingold B, Simon E, Ungethum U, Kuban RJ, et al. (2010). Vascular CXCR4 expression - a novel antiangiogenic target in gastric cancer? *PLoS One* 5: e10087.
- Krejs GJ (2010). Gastric cancer: epidemiology and risk factors. *Dig. Dis.* 28: 600-603.
- Lindgren A, Yun CH, Sjolting A, Berggren C, et al. (2011). Impaired IFN-gamma production after stimulation with bacterial components by natural killer cells from gastric cancer patients. *Exp. Cell Res.* 317: 849-858.
- Machado AM, Figueiredo C, Seruca R and Rasmussen LJ (2010). Helicobacter pylori infection generates genetic instability in gastric cells. *Biochim. Biophys. Acta* 1806: 58-65.
- Maeda S and Omata M (2008). Inflammation and cancer: role of nuclear factor-kappaB activation. *Cancer Sci.* 99: 836-842.
- Maere S, Heymans K and Kuiper M (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* 21: 3448-3449.
- Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, et al. (2011). CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* 39: D225-D229.
- Mayrbaeurl B, Keller G, Schauer W, Burgstaller S, et al. (2010). Germline mutation of the E-cadherin gene in three sibling cases with advanced gastric cancer: clinical consequences for the other family members. *Eur. J. Gastroenterol. Hepatol.* 22: 306-310.
- Mudduluru G, Vajkoczy P and Allgayer H (2010). Myeloid zinc finger 1 induces migration, invasion, and in vivo metastasis through Axl gene expression in solid cancer. *Mol. Cancer Res.* 8: 159-169.
- Nemeth MJ, Curtis DJ, Kirby MR, Garrett-Beal LJ, et al. (2003). Hmgb3: an HMG-box family member expressed in primitive hematopoietic cells that inhibits myeloid and B-cell differentiation. *Blood* 102: 1298-1306.
- Nemeth MJ, Cline AP, Anderson SM, Garrett-Beal LJ, et al. (2005). Hmgb3 deficiency deregulates proliferation and differentiation of common lymphoid and myeloid progenitors. *Blood* 105: 627-634.
- Nemeth MJ, Kirby MR and Bodine DM (2006). Hmgb3 regulates the balance between hematopoietic stem cell self-renewal and differentiation. *Proc. Natl. Acad. Sci U. S. A.* 103: 13783-13788.
- Nobili S, Bruno L, Landini I, Napoli C, et al. (2011). Genomic and genetic alterations influence the progression of gastric cancer. *World J. Gastroenterol.* 17: 290-299.
- Otsubo T, Akiyama Y, Yanagihara K and Yuasa Y (2008). SOX2 is frequently downregulated in gastric cancers and inhibits cell growth through cell-cycle arrest and apoptosis. *Br. J. Cancer* 98: 824-831.
- Saif MW, Makrilia N, Zalonis A, Merikas M, et al. (2010). Gastric cancer in the elderly: an overview. *Eur. J. Surg. Oncol.* 36: 709-717.
- Sanborn JZ, Benz SC, Craft B, Szeto C, et al. (2011). The UCSC Cancer Genomics Browser: update 2011. *Nucleic Acids Res.* 39: D951-D959.
- Sparvero LJ, Asafu-Adjei D, Kang R, Tang D, et al. (2009). RAGE (Receptor for Advanced Glycation Endproducts), RAGE ligands, and their role in cancer and inflammation. *J. Transl. Med.* 7: 17.
- Thompson JD, Higgins DG and Gibson TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22: 4673-4680.
- Tsukamoto T, Mizoshita T, Mihara M, Tanaka H, et al. (2005). Sox2 expression in human stomach adenocarcinomas with gastric and gastric-and-intestinal-mixed phenotypes. *Histopathology* 46: 649-658.
- Wingender E (2008). The TRANSFAC project as an example of framework technology that supports the analysis of genomic regulation. *Brief. Bioinform.* 9: 326-332.

- Yang P, Qiu G, Wang S, Su Z, et al. (2010). The mutations of Th1 cell-specific T-box transcription factor may be associated with a predominant Th2 phenotype in gastric cancers. *Int. J. Immunogenet.* 37: 111-115.
- Yu J, Ma X, Cheung KF, Li X, et al. (2010). Epigenetic inactivation of T-box transcription factor 5, a novel tumor suppressor gene, is associated with colon cancer. *Oncogene* 29: 6464-6474.