Comprehensive analysis of histone modification-associated genes on differential gene expression and prognosis in gastric cancer

XIANGYU MENG, YAN ZHAO, JINGWEI LIAO, LU WANG, ZHE DONG, TAO ZHENG, XIAOHU GU and ZHICHAO ZHENG

Department of Gastric Surgery, Cancer Hospital of China Medical University/Liaoning Cancer Hospital and Institute, Shenyang, Liaoning 110042; Department of Anorectal Surgery, The First Hospital of China Medical University, Shenyang, Liaoning 110001; Department of Ultrasonography, The Fourth Affiliated Hospital of China Medical University, Shenyang, Liaoning 110032, P.R. China

Received January 31, 2019; Accepted June 27, 2019

DOI: 10.3892/etm.2019.7808

Abstract. Accumulating evidence suggests that the epigenetic alterations caused by histone modifications have important roles in the genesis of gastric cancer (GC), particularly the well-studied acetylation and methylation modifications. In the present study, a Bioinformatics analysis of the expression of histone modification-associated genes in GC and normal tissues was performed by using datasets from Oncomine, the Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA). The clinical data of GC patients were downloaded from TCGA to determine the association between histone modification-associated gene expression and clinicopathological parameters or survival of GC. Finally, lysine acetyltransferase 2A (KAT2A), nuclear receptor coactivator 1 (NCOA1), SMYD family member 5 (SMYD5), protein arginine methyltransferase 1 (PRMT1) and PRDF1-RIZ (PR)/Su(var)3-9, enhancer-of-zeste and trithorax (SET) domain 16 (PRDM16) were screened; KAT2A, SMYD5 and PRMT1 were upregulated, while PRDM16 expression was downregulated in GC. Analysis of the GEO and Oncomine datasets revealed that NCOA1 was upregulated, which was contrary to the result obtained with the TCGA stomach adenocarcinoma dataset. Aberrant expression of KAT2A, NCOA1, SMYD5 and PRMT1 in GC was more obvious in gastric intestinal-type adenocarcinoma; low NCOA1 expression was associated with better overall survival of GC patients (hazard ratio (HR)=0.690, 95% CI=0.570-0.840, P<0.001) and was an independent predictor for patients diagnosed with GC (HR=0.639, 95% CI=0.437-0.933, P=0.020). Correlation analysis and protein-protein interaction network analysis indicated a close association between ATAD2 and estrogen receptor 1 (ESR1), PRMT1, NCOA1 and KAT2A. In conclusion, differential expression of KAT2A, NCOA1, SMYD5, PRMT1 and PRDM16 was identified in GC vs. normal tissues, low NCOA1 expression was associated with poor survival of GC and ATAD2 may interact with ESR1 to regulate NCOA1 and PRMT1 in GC.

Introduction

Gastric carcinoma (GC) is the second most common human cancer type and a leading cause of cancer-associated mortality worldwide. Although the incidence has significantly declined due to recent advances in diagnostics and therapeutics, the survival rate remains poor. Multistep processes, including genetics, epigenetics and environmental factors, have pivotal roles in tumorigenesis and progression (1). The identification of novel biomarkers in the above processes for clinical applications is urgently required.

The histone core proteins (two of each H2A, H2B, H3 and H4) together with 146 bp DNA are wrapped around each other and form a nucleosome, which constitutes the basic units of chromatin (2). In dynamic and reversible processes, as the chromosomes are condensed or loosened, the N-terminus of the histone proteins may be altered by multiple covalent modifications, including acetylation, methylation, phosphorylation and ubiquitination, at the post-transcriptional level to regulate gene expression. Numerous studies on covalent modifications have focused on exploring the roles of acetylation and methylation (3,4). Dysregulations of these epigenetic modifications and associated gene expression causes may drive carcinogenesis. Downregulation of lysine acetyltransferase 5 (KAT5) (5) and upregulation of enhancer of zeste 2 polycomb repressive complex 2 subunit (EZH2) (6), protein arginine methyltransferase 1 (PRMT1) (7) and the lysine demethylase 1A (KDM1A) (8) have been reported to be significantly associated with poor clinicopathological features and survival of patients with GC and other cancer types. As one of the epigenetic mechanisms, histone modifications participate in
transcriptional regulation, DNA repair and condensation (9). Histone deacetylase 4 has been reported to facilitate GC progression by inhibiting p21 (10). P300 acetylates transcription factor (TF) STAT3 in histone H3 on lysine 56 to regulate gene expression (11). Furthermore, EZH2 recruits DNA methyltransferase to the promoter region of PcG target genes to downregulate PcG targets (12). In addition, lysine demethylase 6B interacts with NF-kB via demethylation of histone H3 trimethylated at lysine 27 (H3K27me3) at downstream gene promoters participating in wound healing (13). In the DNA damage repair process, KAT8 was also reported to be required (14). Alterations in histone modification levels and the expression of numerous genes encoding histone modification-associated enzymes have also been reported in various cancer types as epigenetic changes. For instance, SUV420H2 (KMT5C)-mediated histone H4 trimethylation on lysine 20 is important for epidermal homeostasis. SET domain bifurcated histone lysine methyltransferase 1, which performs H3K9 trimethylation (H3K9me3), is established as an oncogene in melanoma (15). Upregulation of JMJD3 in metastatic prostate cancer indicates its potential oncogenic role (16). Similarly, KDM1A and JMJD1C are upregulated in certain GC cells and tissues (17). Therefore, histone modification-associated changes may serve as biomarkers for early diagnosis, therapeutic targets and prognosis prediction of GC.

Although a large number of studies have indicated that histone modifications and abnormal expression of associated genes have important roles in oncogenesis, only few studies have provided comprehensive analyses of the expression and prognostic value of associated genes in carcinomas, particularly in GC. The present study analyzed Oncomine, The Cancer Genome Atlas (TCGA) and Gene Ontology (GEO) datasets to perform comprehensive analyses on histone modifications and associated gene expression profiles, as well as their prognostic role in GC.

Materials and methods

Oncomine database analysis. The Oncomine database (http://www.oncomine.org) incorporates 715 datasets that include 35 cancer types and contain microarray data of 86,733 samples, supporting various methods of online statistical analysis (18). The differential expression of histone modification and associated genes (HMGs) was compared by using the Student’s t-test to generate a P-value. As cut-off values, the \(|\text{logFC}| > 1\) for the comparison between tumor and non-tumor bases. DEGs were screened out using the cutoffs of \(P<0.05\) and \(|\text{logFC}|>1\) for the comparison between tumor and non-tumor samples. A total of 1,311 upregulated and 384 downregulated expressed genes (DEGs) between the tumor and non-tumor groups were identified. Detailed information on the location and function of the HMGs is provided in supplementary Table SI. The differences in expression from HMGs between various carcinomas and paracancerous tissues were also indicated in the Oncomine analysis (Fig. S2).

Statistical analysis. The association between histone modifications and the expression of relevant genes as well as clinicopathological features was evaluated by the \(\chi^2\) test. An unpaired t-test was used to compare the differentially expressed genes (DEGs) between the tumor and non-tumor group. Spearman's test was used for correlation analysis. The Kaplan-Meier method was used to determine the patients' survival and differences between groups were assessed using a log-rank test. In TCGA dataset, statistically significant variables in the univariate analysis were included into the multivariate analysis in the Cox proportional hazards model and results were expressed as the HR with 95% CI. Data analysis was performed using SPSS software v.23.0 (IBM Corp.). The mRNA data downloaded from TCGA and GEO were analyzed using the edgeR package in R (v.3.5.1) to identify DEGs between GC and non-tumor tissues. The median value of mRNA expression was applied as a cut-off to stratify samples into high- or low-expression groups. The STRING v.10.5 online tool (https://string-db.org/) was used to analyze the interactions among differential proteins. All of the P-values reported were two-sided and \(P<0.05\) was considered to indicate statistical significance.

Results

Screening and identification of DEGs in the HMGs from Oncomine, GEO and TCGA datasets. The microarray expression profile dataset GSE79973 and mRNA expression data were respectively downloaded from the GEO and TCGA databases. DEGs were screened out using the cutoffs of \(P<0.05\) and \(|\text{logFC}|>1\) for the comparison between tumor and non-tumor samples. A total of 1,311 upregulated and 384 downregulated genes screened from GSE79973 (Fig. S1) and 8,159 upregulated and 3,758 downregulated genes were screened from TCGA (Fig. 1A). The HMGs were then retrieved and determined using the UALCAN website. A total of 30 genes involved the acetyl modification of histones (12 acetyltransferases and associated genes, 18 deacetylases and associated genes) and 61 genes participating in methyl modification (53 methyltransferases and associated genes, 8 demethylases and associated genes) were identified. Detailed information on the location and function of the HMGs is provided in supplementary Table SI. The differences in expression from HMGs between various carcinomas and paracancerous tissues were also indicated in the Oncomine analysis (Fig. S2).
To screen the HMGs that were differentially expressed in GC vs. non-cancer tissues among all Oncomine, GEO and TCGA datasets, Venn diagram analysis was used to obtain the intersection of the three datasets. As presented in Fig. 1B-E, the five genes KAT2A, nuclear receptor coactivator 1 (NCOA1), SMYD family member 5 (SMYD5), PRMT1 and PR/SET domain 16 (PRDM16) were identified. Detailed information on the location and function of these HMGs is provided in Table I.

**Differential expression of KAT2A, NCOA1, SMYD5, PRMT1 and PRDM16 in GC vs. normal tissues.** KAT2A, SMYD5 and PRMT1 were upregulated in GC vs. normal tissues in all three datasets (Oncomine, GSE79973 and TCGA). In the GSE79973 and TCGA STAD datasets, PRDM16 was observed to be downregulated. Furthermore, in the GSE79973 dataset and the dataset DErrico Gastric (Oncomine dataset), NCOA1 was upregulated, and these results were contrary to those obtained with the TCGA STAD dataset (Table II). A 29% decline of NCOA1 mRNA expression was observed in TCGA STAD dataset. In the Wang Gastric and Cui Gastric datasets (Oncomine dataset), KAT2A was 1.756-fold (P=1.15x10^{-4}) and 1.461-fold (P=1.37x10^{-5}) increased, respectively, in GC vs. normal tissues. Furthermore, the FC of upregulated KAT2A in the GSE79973 dataset was 1.809 (P=1.89x10^{-5}). In the Wang Gastric dataset from Oncomine, SMYD5 and PRMT1, which in turn was also determined to be 1.470-fold (P=1.00x10^{-5}) increased in the Cui Gastric dataset, were 2.154-fold (P=6.18x10^{-4}) and 1.884-fold (P=8.01x10^{-4}) elevated, respectively. Similar to the results of the Oncomine analysis, screening of the GSE79973 dataset indicated that NCOA1, SMYD5 and PRMT1 were 1.347-fold (P=8.95x10^{-5}), 1.708-fold (P=7.39x10^{-4}) and 1.724-fold (P=2.00x10^{-4}) increased. Furthermore, comparisons based on different pathological classifications were performed to determine differential expression of KAT2A, NCOA1, SMYD5 and PRMT1, and in the datasets DErrico Gastric and Cho Gastric in the Oncomine analysis, the FC of KAT2A, NCOA1, SMYD5 and PRMT1 in gastric intestinal-type adenocarcinoma was higher than in gastric mixed and diffuse gastric adenocarcinoma compared to normal tissues (respective FCs and P-values: 2.132, 1.61x10^{-11}; 1.562, 7.00x10^{-3}; 4.106, 1.08x10^{-10}; 2.776, 2.62x10^{-11}). The box plots in Fig. 2A-I display the significant alterations in expression among the five genes in subgroup comparisons.

**Association of KAT2A, NCOA1, SMYD5, PRMT1 and PRDM16 expression with clinicopathological features.** To explore the potential clinical significance of the five DEGs in patients with GC, the clinical data of 443 patients were downloaded from TCGA and only 334 patients with complete clinical and gene data were retained after elimination of patients with missing data by processing in R. The cases were divided into a low-expression group and a high-expression group according to the median mRNA expression value of the five genes in tumor or non-tumor tissues. The associations between the expression of the five genes and clinicopathological characteristics are presented in Table III. Factors including age, sex, tumor size, AJCC stage and T/N/M stage were evaluated. Unfortunately, no significant associations were observed between the five aberrantly expressed genes and the above-mentioned clinical features.

---

**Figure 1.** (A) Volcano plot of differentially expressed mRNAs between gastric adenocarcinoma and para-carcinoma tissues. Red indicates high expression and green indicates low expression (|logFC|>1 and adjusted P<0.05). (B-E) Venn diagrams for differentially expressed (B) histone acetyltransferases, (C) deacetylases, (D) methyltransferases and (E) demethylases in the three datasets. FDR, false discovery rate; FC, fold change; TCGA, The Cancer Genome Atlas; GEO, Gene Expression Omnibus (GSE79973).
Prognostic value of KAT2A, NCOA1, SMYD5, PRMT1 and PRDM16 in GC. The prognostic value of the expression status of KAT2A, NCOA1, SMYD5, PRMT1 and PRDM16 was first examined in the TCGA dataset (Fig. 3A-E). Low mRNA expression of NCOA1 mRNA was observed to be linked to significantly better OS in GC (HR=0.667, 95% CI=0.473-0.943, P=0.022; Fig. 3B). However, the mRNA expression status of KAT2A, SMYD5, PRMT1 and PRDM16 mRNA was not associated with OS in GC (HR=1.146, 95% CI=0.811-1.620, P=0.440; HR=1.031, 95% CI=0.733-1.452, P=0.860; HR=1.024, 95% CI=0.725-1.447, P=0.892; HR=0.950, 95% CI=0.667-1.354, P=0.777; Fig. 3A and C-E, respectively). Kaplan-Meier plotter analysis was also used to verify the associations between the mRNA expression of five genes and the clinical survival outcome (Fig. 3F-J). In order to reduce the bias caused by the conversion of time (day or month), the time units were used as presented in the respective datasets for the Kaplan-Meier survival curves. As in the TCGA dataset, low NCOA1 expression in GC tissues was associated with better OS (HR=0.690, 95% CI=0.570-0.840, P<0.001), and the survival curve for the group with high PRMT1 mRNA expression in GC tissues did not exhibit any significant difference from that of the low expression group (HR=0.840, 95% CI=0.710-1.010, P=0.061; Fig. 3l). Contrary to the results obtained with the TCGA dataset, patients with high expression of KAT2A, SMYD5 and PRDM16 in their GC tissue exhibited significantly poorer OS than those with corresponding low expression according to Kaplan-Meier plotter analysis (Fig. 3F, H and J, respectively). The Cox proportional hazards model indicated that NCOA1 expression, age, N stage, M stage were independent predictors for survival in GC (HR=1.523, 95% CI=1.072-2.163, P=0.019; HR=1.939, 95% CI=1.291-2.914, P=0.001; HR=1.233, 95% CI=1.049-1.450, HR=2.419, 95% CI=1.423-4.115, respectively; Table IV).

Bromodomain protein ATPase family AAA domain-containing protein 2 (ATAD2) is associated with the HMGs. ATAD2, a member of the AAA + ATPase family of proteins, contains a bromodomain and its functions are linked to genome regulation and histone modification. A previous study by our group on hepatocellular carcinoma suggested that ATAD2 was overexpressed in tumor tissue and associated with poor survival (21). To explore the association between ATAD2 and the five genes screened out from datasets in the present study, the STRING v.10.5 online tool (https://string-db.org/) was used to display the interactions among them. As presented in Fig. 4A, protein-protein interaction network highlighted the association between ATAD2 and estrogen receptor 1 (ESR1), PRMT1, NCOA1 and KAT2A. Associations between clinicopathological features and ATAD2 or ESR1 mRNA expression were also assessed (Table SII) and no statistical significances were observed. Spearman's correlation analysis revealed that ATAD2 expression was positively correlated with PRMT1.

Table I. Basic characteristics and function of five histone modification enzymes and associated genes.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Location</th>
<th>Exon</th>
<th>Protein mass (kDa)</th>
<th>Encoding protein and biological function</th>
</tr>
</thead>
<tbody>
<tr>
<td>KAT2A</td>
<td>17q21.2</td>
<td>18</td>
<td>93.94</td>
<td>Also known as GCN5, KAT2A is a HAT that functions primarily as a transcriptional activator. It also functions as a repressor of NF-κB by promoting ubiquitination of the NF-κB subunit RELA in a HAT-independent manner</td>
</tr>
<tr>
<td>NCOA1</td>
<td>2p23.3</td>
<td>24</td>
<td>156.8</td>
<td>A transcriptional coactivator for steroid and nuclear hormone receptors. A member of the p160/steroid receptor coactivator family, which has histone acetyltransferase activity and contains a nuclear localization signal, as well as basic helix-loop-helix and PAS domains. NCOA1 also binds nuclear receptors directly and stimulates transcriptional activities in a hormone-dependent fashion.</td>
</tr>
<tr>
<td>SMYD5</td>
<td>2p13.2</td>
<td>13</td>
<td>34.42</td>
<td>-</td>
</tr>
<tr>
<td>PRMT1</td>
<td>19q13.33</td>
<td>13</td>
<td>39.61</td>
<td>A member of the PRMT family, which is involved in post-translational modification of target proteins in numerous biological processes. As a type I PRMT, it is responsible for the majority of cellular arginine methylation activity. Increased expression of this gene may have a role in numerous types of cancer.</td>
</tr>
<tr>
<td>PRDM16</td>
<td>1p36.32</td>
<td>18</td>
<td>-</td>
<td>A zinc finger transcription factor containing an N-terminal PR domain. Translocation results in the overexpression of a truncated version of this protein that lacks the PR domain, which may have an important role in the pathogenesis of myelodysplastic syndrome and acute myelocytic leukemia.</td>
</tr>
</tbody>
</table>

KAT2A, lysine acetyltransferase 2A; NCOA1, nuclear receptor coactivator 1; arginine methyltransferase 1; SMYD5, SMYD family member 5; SET, su(var)3-9, enhancer-of-zeste, trithorax; PR domain, positive regulatory domain I element-BF1 and RIZ homology domain; PRDM16, PR/SET domain 16; ESR1, estrogen receptor 1.
Table II. Differential expression of five histone modification enzymes and associated genes between different types of gastric cancer and normal tissues.

<table>
<thead>
<tr>
<th>Gene/comparison of groups</th>
<th>Up/downregulation</th>
<th>Fold change</th>
<th>t-test</th>
<th>P-value</th>
<th>Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>KAT2A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gastric cancer vs. Normal</td>
<td>↑</td>
<td>1.756</td>
<td>4.356</td>
<td>1.15×10^{-4}</td>
<td>Wang Gastric</td>
</tr>
<tr>
<td></td>
<td>↑</td>
<td>1.461</td>
<td>4.374</td>
<td>1.37×10^{-5}</td>
<td>Cui Gastric</td>
</tr>
<tr>
<td></td>
<td>↑</td>
<td>1.809</td>
<td>6.698</td>
<td>1.89×10^{-5}</td>
<td>GSE79973</td>
</tr>
<tr>
<td>Gastric intestinal type adenocarcinoma vs. Normal</td>
<td>↑</td>
<td>2.132</td>
<td>8.265</td>
<td>1.61×10^{-11}</td>
<td>DErrico Gastric</td>
</tr>
<tr>
<td>Gastric mixed adenocarcinoma vs. Normal</td>
<td>↑</td>
<td>1.636</td>
<td>5.839</td>
<td>7.92×10^{-6}</td>
<td></td>
</tr>
<tr>
<td>Diffuse gastric adenocarcinoma vs. Normal</td>
<td>↑</td>
<td>1.848</td>
<td>4.568</td>
<td>7.12×10^{-4}</td>
<td></td>
</tr>
<tr>
<td>Gastric adenocarcinoma vs. Normal</td>
<td>↑</td>
<td>1.998</td>
<td>-</td>
<td>3.10×10^{-14}</td>
<td>TCGA STAD</td>
</tr>
<tr>
<td>NCOA1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gastric cancer vs. Normal</td>
<td>↑</td>
<td>1.347</td>
<td>3.283</td>
<td>8.95×10^{-3}</td>
<td>GSE79973</td>
</tr>
<tr>
<td>Gastric mixed adenocarcinoma vs. Normal</td>
<td>↑</td>
<td>1.562</td>
<td>3.675</td>
<td>7.00×10^{-3}</td>
<td>DErrico Gastric</td>
</tr>
<tr>
<td>Gastric adenocarcinoma vs. Normal</td>
<td>↓</td>
<td>1.509</td>
<td>-</td>
<td>2.65×10^{-12}</td>
<td>TCGA STAD</td>
</tr>
<tr>
<td>SMYD5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gastric cancer vs. Normal</td>
<td>↑</td>
<td>2.154</td>
<td>3.660</td>
<td>6.18×10^{-4}</td>
<td>Wang Gastric</td>
</tr>
<tr>
<td>Gastric intestinal type adenocarcinoma vs. Normal</td>
<td>↑</td>
<td>1.708</td>
<td>7.350</td>
<td>7.39×10^{-6}</td>
<td>GSE79973</td>
</tr>
<tr>
<td>Gastric mixed adenocarcinoma vs. Normal</td>
<td>↑</td>
<td>4.106</td>
<td>7.776</td>
<td>1.08×10^{-10}</td>
<td>DErrico Gastric</td>
</tr>
<tr>
<td>Diffuse gastric adenocarcinoma vs. Normal</td>
<td>↑</td>
<td>1.404</td>
<td>6.754</td>
<td>1.58×10^{-9}</td>
<td>Chen Gastric</td>
</tr>
<tr>
<td>Gastric adenocarcinoma vs. Normal</td>
<td>↑</td>
<td>1.215</td>
<td>3.231</td>
<td>1.00×10^{-3}</td>
<td>Cho Gastric</td>
</tr>
<tr>
<td>Dištice gastric adenocarcinoma vs. Normal</td>
<td>↑</td>
<td>3.901</td>
<td>8.553</td>
<td>2.62×10^{-11}</td>
<td>DErrico Gastric</td>
</tr>
<tr>
<td>Gastric adenocarcinoma vs. Normal</td>
<td>↑</td>
<td>1.263</td>
<td>4.691</td>
<td>1.87×10^{-5}</td>
<td>Chen Gastric</td>
</tr>
<tr>
<td>PRMT1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gastric cancer vs. Normal</td>
<td>↑</td>
<td>1.884</td>
<td>3.608</td>
<td>8.01×10^{-4}</td>
<td>Wang Gastric</td>
</tr>
<tr>
<td>Gastric intestinal type adenocarcinoma vs. Normal</td>
<td>↑</td>
<td>1.470</td>
<td>3.114</td>
<td>1.00×10^{-3}</td>
<td>Cui Gastric</td>
</tr>
<tr>
<td>Gastric mixed adenocarcinoma vs. Normal</td>
<td>↑</td>
<td>1.724</td>
<td>8.310</td>
<td>2.00×10^{-6}</td>
<td>GSE79973</td>
</tr>
<tr>
<td>Diffuse gastric adenocarcinoma vs. Normal</td>
<td>↑</td>
<td>2.776</td>
<td>8.515</td>
<td>2.62×10^{-11}</td>
<td>DErrico Gastric</td>
</tr>
<tr>
<td>Gastric adenocarcinoma vs. Normal</td>
<td>↑</td>
<td>1.473</td>
<td>6.843</td>
<td>2.16×10^{-9}</td>
<td>Chen Gastric</td>
</tr>
<tr>
<td>PRDM1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gastric cancer vs. Normal</td>
<td>↑</td>
<td>1.574</td>
<td>3.539</td>
<td>5.42×10^{-3}</td>
<td>GSE79973</td>
</tr>
<tr>
<td>Gastric adenocarcinoma vs. Normal</td>
<td>↓</td>
<td>1.832</td>
<td>-</td>
<td>6.87×10^{-5}</td>
<td>TCGA STAD</td>
</tr>
</tbody>
</table>

KAT2A, lysine acetyltransferase 2A; NCOA1, nuclear receptor coactivator 1; PRMT1, protein arginine methyltransferase 1; SMYD5, SMYD family member 5; PRDM16, PRDF1-RIZ/Su(var)3-9, enhancer-of-zeste and trithorax domain 16; '/', '-', not reported; TCGA STAD, The Cancer Genome Atlas stomach adenocarcinoma. Oncomine dataset: Wang Gastric, Cui Gastric, DErrico Gastric, Chen Gastric, Cho Gastric.

(R^2=0.123, Spearman rho=0.356, P=1.19×10^{-12}; Fig. 4B) and negatively correlated with ESR1 (R^2=0.177, Spearman rho=-0.449, P=5.30×10^{-20}; Fig. 4C) and NCOA1 (R^2=0.022, Spearman rho=-0.160, P=0.002; Fig. 4D).
Figure 2. Box plots that represent the mRNA expression levels of histone modification-associated genes in different types of gastric cancer. (A-G) Datasets of Oncomine; (A) KAT2A: 0, normal tissues; 1, diffuse gastric adenocarcinoma; 2, gastric adenocarcinoma; 3, gastric intestinal type adenocarcinoma; 4, gastric mixed adenocarcinoma; (B) KAT2A: 1, gastric mucosa; 2, gastric tissue; 3, gastric cancer; (C) NCOA1: 0, normal tissue; 1, gastric mucosa; 2, diffuse gastric adenocarcinoma; 3, gastric intestinal type adenocarcinoma; 4, gastric mixed adenocarcinoma; (D) PRMT1: 0, normal tissues; 1, diffuse gastric adenocarcinoma; 2, gastric adenocarcinoma; 3, gastric intestinal type adenocarcinoma; 4, gastric mix adenocarcinoma; (E) PRMT1: 1, gastric mucosa; 2, gastric tissue; 3, gastric cancer; (F) SMYD5: 0, normal tissues; 1, diffuse gastric adenocarcinoma; 2, gastric adenocarcinoma; 3, gastric intestinal type adenocarcinoma; 4, gastric mix adenocarcinoma; (G) SMYD5: 1, gastric mucosa; 2, gastric tissue; 3, gastric cancer. (H) Datasets of GSE79973. (I) Dataset of TCGA.

* P<0.05; ** P<0.01 and *** P<0.001 vs. non-tumor. TCGA, The Cancer Genome Atlas; KAT2A, lysine acetyltransferase 2A; NCOA1, nuclear receptor coactivator 1; PRMT1, protein arginine methyltransferase 1; SMYD5, SMYD family member 5; PRDM16, PRDF1-RIZ/Su(var)3-9, enhancer-of-zeste and trithorax domain 16.

Figure 3. Kaplan-Meier survival curves from TCGA and Kaplan-Meier plotter analyses, depicting the survival of gastric cancer patients according to the expression levels of histone modification-associated genes. (A-E) Kaplan-Meier survival curves from TCGA datasets; (A) KAT2A, (B) NCOA1, (C) SMYD5, (D) PRMT1, (E) PRDM16; (F-J) Kaplan-Meier plotter analysis for five genes (F) KAT2A, (G) NCOA1; (H) SMYD5; (I) PRMT1; (J) PRDM16. TCGA, The Cancer Genome Atlas; KAT2A, lysine acetyltransferase 2A; NCOA1, nuclear receptor coactivator 1; PRMT1, protein arginine methyltransferase 1; SMYD5, SMYD family member 5; PRDM16, PRDF1-RIZ/Su(var)3-9, enhancer-of-zeste and trithorax domain 16; HR, hazard ratio; The values in brackets are the 95% CI.
The maintenance of stable and ordered chromatin during dynamic packaging is vital to normal cellular homeostasis. Warped histones and DNA are subject to covalent post-translational modifications in order to influence the number of chromatin-associated cellular events, including transcription, replication recombination and DNA repair (22).

Dysregulation of the epigenetic mechanisms that govern transcriptional regulation resembles steps in the oncogenic process, causing inappropriate activation of oncogenes or the inhibition of tumor suppressors and leading to carcinogenesis. Accumulating evidence has demonstrated that epigenetic alterations caused by histone modifications also have important roles in gastric carcinogenesis, particularly the well-studied acetylation and methylation modifications (23).
modifications-associated enzymes and encoding genes are among the most frequent abnormal targets in aberrant histone modifications. Genes that encode histone modification enzymes, including the CREB binding protein (CBP), p300, KAT5, KDM1A and JMJD1C, have been reported to be aberrantly expressed in GC and are significantly correlated with poor survival (5,6,17,24-26). Possibly due to lack of data at the genomic level, comprehensive molecular characterization of histone modification regulations in GC has rarely been performed. To the best of our knowledge, the present study was the first to elucidate histone modifications, associated gene expression profiles and prognostic roles of key genes by using Bioinformatics analysis of datasets to explore the implication of deregulated histone modification in the initiation and development of GC and the underlying mechanisms.

The present results suggested that KAT2A, NCOA1, SMYD5, PRMT1 and PRDM16 were differentially expressed in GC vs. non-cancer tissues. Among them, KAT2A, SMYD5 and PRMT1 were highly upregulated in GC compared with normal tissues. The expression of PRDM16 was observed to be downregulated in GC. However, the opposite result was identified when detecting NCOA1 expression. Analysis results from a low NCOA1 expression by our analysis results in this article appeared to contradict previous evidence; however, the GC dataset from Frycz et al (32) indicated that the mRNA expression of NCOA1 (P=0.00021) was significantly upregulated in GC compared with non-cancer tissues.
reduced in the tumoral mucosa compared with that in the adjacent healthy mucosa. Decreased levels of NCOA1 mRNA in GC tissue may be due to upregulation of cytochrome P450 family 19 subfamily A member 1 mRNA in the tumoral gastric mucosa, which causes dysregulation of two 17β-estradiol (E2) synthesis routes (the sulfatase and aromatase signaling pathways), resulting in E2 deficiency and inhibition of NCOA1 expression in E2-dependent methods (33). Furthermore, low expression of NCOA1 mRNA has been observed in bladder cancer urothelium samples (34), and the results of the present study were consistent with these results. NCOA1 not only acts as a coactivator, but has also been indicated to possess histone acetyltransferase activity. Sheppard et al (35) reported that the recruitment and induction of NCOA1 in the H3 sequence are each critical for the NCOA1-CBP interaction, which is necessary for ER function. Although the FC of NCOA1 expression was <2 in the GEO, TCGA and Oncomine datasets, the present results demonstrated that, no matter whether the expression of NCOA1 was upregulated or downregulated, aberrant expression of NCOA1 is significantly associated with poor prognosis and is an independent predictor for GC from our analysis (TCGA data: HR=0.639, 95% CI=0.437-0.933, P=0.020). This evidence suggested that weakly DEGs may also have important functions and roles in the occurrence and development of tumors. For instance, by quantifying the adenosomatous polyposis coli (APC) gene, Yan et al (36) determined that a weak reduction of APC expression was closely associated with the occurrence of hereditary colorectal tumors.

Unfortunately, no associations were observed between the aberrant expression of five specific genes (KAT2A, NCOA1, SMYD5, PRMT1 and PRDM16) and clinical features when their potential clinical significance in patients with GC was explored in the TCGA dataset. Although age may serve as an independent prognostic factor according to the multivariate analysis, it was not a significant predictor for survival of patients whose genes were differentially expressed when stratified by age (data not presented). In addition, in a report on 97 patients with stage III-IVa/b head and neck squamous cell carcinoma, similar results were observed, in that NCOA1 expression was not significantly associated with any clinicopathological features, including sex, tumor site, T classification or nodal status (37). However, the majority of previous studies on certain malignant tumors indeed suggested that the aberrant expression of KAT2A, NCOA1, SMYD5, PRMT1 and PRDM16 is an independent prognostic factor for survival of patients whose genes were differentially expressed when stratified by age. For instance, CIN in colorectal cancer was reported to be correlated with its upstream and downstream effects. In glioma cells, hypoxia decreased the gene expression of NCOA1 (45). It is worth mentioning that these genes also have important roles in drug resistance (47,48).

Underlying genetic alterations, including promoter methylation, copy number variations (CNVs) and chromosomal instability (CIN), have been reported to regulate histone modification-associated gene expression. For instance, low PRDM16 expression levels in non-small cell lung cancer and esophageal cancer were reported to be correlated with its promoter methylation (49,50). A high frequency of CNVs at lp36.32 harboring the PRDM16 gene was observed in GC, suggesting that changes in gene CNVs also have vital roles in regulating gene expression (51). In addition, Burghel et al (52) further demonstrated that PRDM16 was highly expressed in gained focal minimal common regions caused by CIN in colorectal cancer. Besides genetic variations, post-transcriptional regulations also affected the functioning of associated enzymes. For instance, NCOA1 was confirmed as a target of miR-223-3p and demonstrated to have low expression levels (53,54). Evidence has indicated that changes to the microenvironment resulting from hypoxia, which is insufficient to maintain cellular function, were associated with cancer pathology. In glioma cells, hypoxia decreased the gene expression of NCOA1 (55). It is known that Helicobacter pylori is responsible for gastric inflammation and gastric malignancy, which causes general inflammatory stress within the gastric mucosa, activating multiple oncogenic pathways and inducing epigenetic alterations, including histone modifications (56). These results suggested that factors upstream of histone modification-associated genes are important for the regulation of the expression of these genes.

Aberrant gene expression caused by upstream factors initiates cascade reactions, resulting in normal cells transforming into cancer cells and other types of malignant behavior. For instance, deletion of the KAT2A gene induces apoptosis in acute myeloid leukemia (57). Similarly, an increased apoptotic rate was observed in prostate cancer cells when PRDM16 expression was downregulated (58). Regarding cell
regeneration and differentiation, histone modification-associated genes were also suggested to have important roles. PRMT1 methylates arginine on substrates Six1 or Eya1, and has been indicated to regulate muscle stem cell regeneration and differentiation (59). In embryonic stem cell differentiation, SMYD5 increases H4K20me3 and H3K9me3 levels and maintains chromosome integrity to ensure accurate differentiation (60). The functional basis of invasion and metastasis is the epithelial-to-mesenchymal transition (EMT). Previous studies have demonstrated that PRMT1, as a regulator, is closely associated with EMT. Katsumo et al (61) reported that PRMT1 is an essential mediator of transforming growth factor-β signaling, regulating the EMT and epithelial cell stemness by methylating SMAD7. Dysregulation of histone modification-associated genes not only promotes the acquisition of malignant biological phenotypes of cancer cells, but also has important effects on immunity. KAT2A, which is recruited by nuclear factor of activated T cells during the activation of T-cell receptor signaling pathways, methylates H3K9 of the interleukin-2 gene promoter to regulate T-cell activation and CD4+ T-cell differentiation into type 1 T-helper cells (Th1)/Th17 (62). In addition, methylation catalyzed by PRMT1 has been indicated to be required for pre-B-cell development and mature B-cell activation, together with B-cell translocation gene (63).

The present analysis on GC also indicated that, besides differential expression of NCOA1, our results on low PRDM16 and high PRMT1 levels were inconsistent with the corresponding contents of previous studies (64,65). The paradoxical results may be explained by the complexity of heterogeneous factors in GC. Furthermore, the varying expression profiles of KAT2A, NCOA1, SMYD5 and PRMT1 among different pathological classifications of gastric adenocarcinoma also demonstrated the presence of heterogeneity. A previous study by our group and other studies suggested that aberrant ATAD2 expression is associated with histone modification, hinting at a close correlation between ATAD2 and histone modification-associated genes (66). As speculated, the correlation and STRING interaction analysis indicated that ATAD2 may interact with ESR1 to regulate NCOA1 and PRMT1 in GC. Previous studies revealed that endogenous ATAD2 acts as a co-activator for ESR1 to activate downstream target gene expression, together with hormone-induced ESR1 recruiting to target genes at chromatin (67). NCOA1 also interacts with nuclear hormone receptors, including ESR1 (26). Of note, a negative correlation between NCOA1 and PRMT1 expression was observed in the present study, and the cross-talk among histone modification-associated genes was also previously reported (68). Certain limitations of the present study should be emphasized. Only a Bioinformatics analysis of external data was performed, and the expression of the five genes identified and interactions between NCOA1 and other genes, including the protein levels of ATAD2, were not verified through experimental methods as part of the present study. As a next step, verification in clinical samples from our center and assessment of potential interaction mechanism among ATAD2, ESR1 and NCOA1 will be performed. Hence, the results of such future experiments are to be anticipated.

In conclusion, the present study performed a comprehensive Bioinformatics analysis of the expression of histone modification-associated genes in GC and their association with prognosis. KAT2A, NCOA1, SMYD5, PRMT1 and PRDM16 were screened out and aberrant expression profiles were compared between GC and non-cancer tissues. Low NCOA1 expression was a closely associated with poor prognosis and was identified to be an independent predictor for GC. ATAD2 may interact with ESR1 to regulate NCOA1 and PRMT1 in GC. Due to the heterogeneity in GC, well-designed studies with larger sample sizes are required in the future.

Acknowledgements
Not applicable.

Funding
This study was supported by the Science Foundation of Liaoning Province (grant no. 20170540567).

Availability of data and materials
All datasets used and/or analyzed during the present study are available from the corresponding author on reasonable request.

Authors’ contributions
XM, JL, LW, TZ, XG, ZD and YZ analyzed the data. XM, ZZ and JL designed the study and prepared the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate
Not applicable.

Patient consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References

25.

20.

18.

17.

11.

10.

9.

8.

7.


