Identification of INHBA as a Potential Biomarker for Gastric Cancer by Comprehensive Analysis

Guang-jie LIU
Hebei Cancer Institute: Hebei Medical University Fourth Affiliated Hospital and Hebei Provincial Tumor Hospital

Xiaojie HU
Hebei General Hospital

Yuxuan Jia
Hebei Medical University

Bingjie Huo
Hebei Cancer Institute: Hebei Medical University Fourth Affiliated Hospital and Hebei Provincial Tumor Hospital

Lei Yang
Hebei Medical University School of Public Health

Fang Liu (liufangj@163.com)
Hebei Medical University Fourth Affiliated Hospital

Research

Keywords: INHBA, gastric cancer, prognosis, biomarkers, bioinformatics analysis

Posted Date: October 20th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-92167/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

Read Full License
Abstract

Background

Inhibin subunit beta A (INHBA) is a member of the TGF-beta (transforming growth factor-beta) superfamily proteins, which plays a fundamental role in various cancers. However, there is little systematical analysis on the exact role of INHBA in patients with gastric cancer (GC). Herein, we explored the exact role and the underlying mechanisms of INHBA regarding GC using multiple bioinformatic approaches.

Methods

The expression levels of INHBA in GC were analyzed in TIMER, GEPIA2, GEO, Oncomine and UALCAN databases. Protein and PCR test were performed to verify the expression states of INHBA in GC tissues. The correlation of INHBA and prognosis of GC was analyzed based on Kaplan Meier plotter database. Besides, the relationship between INHBA expression and immune infiltration levels and the type markers of immune cells in GC was explored in the TIMER database. What's more, we studied INHBA mutations, promoter methylation, functional enrichment analysis in GC patients based on cBioportal, MEXPRESS, Metascape and LinkedOmics databases. Besides, we performed immunohistochemistry (IHC) and polymerase chain reaction (qRT-PCR) verification in tissues from patients with gastric cancer.

Results

INHBA was elevated in GC and high expression level of INHBA in GC was significantly related to the unfavorable prognosis. Protein and PCR test verified the highly expression states of INHBA in GC tissues. Further analysis showed that INHBA was negatively correlated with B cell while positively correlated with the marker type of CD8+ T cells, macrophage, neutrophil and dendritic cell infiltration. High INHBA expression level had a poor prognosis in different enriched immune cells subgroups in GC. And there is week significant methylation level change between tumor and normal tissues. Moreover, INHBA mainly enriched cancer-related signaling pathways, including TGF-beta signaling pathway, ECM-receptor signaling pathway, PID ALK1/2 pathway, and AGE-RAGE signaling pathway.

Conclusions

The present study implies that INHBA may serve as a potential biomarkers for predicting prognosis in GC patients. High INHBA expression in GC may affect prognosis through immune infiltration.

Background

Thanks to the techniques in the diagnosis and treatment of gastric cancer (GC), the incidence of GC has decreased significantly, but it is still one of the most prevalent malignancy as well as the third leading cause of cancer-related mortality worldwide \(^{[1,2]}\). Most GC is induced by a complex interaction between epigenetic changes and environmental factors, such as Helicobacter pylori infection and trace element
concentration[3-5]. At present, surgery and chemotherapy are the main therapeutics for GC patients, however, the 5-year survival rates in GC patients remain frustrating due to many patients are still initially diagnosed at an advanced stage and relapse after treatment[6-8]. Therefore, more efforts should be made to seek more beneficial biomarkers for early diagnosis and targeted therapy.

Inhibin subunit beta A (INHBA) is a member of the TGF-beta (transforming growth factor-beta) superfamily proteins which exert a variety of biological functions, including immune response, sex determination, stem cell differentiation, developmental differentiation, and control of cellular migration and proliferation[9-12]. Nowadays, the emerging studies have shown that INHBA is aberrantly expressed in multiple tumor types, such as nasopharyngeal carcinoma[13], lung adenocarcinoma[14], ovarian cancer[15,16], colon cancer[17,18], esophageal squamous cell carcinoma[19], pancreatic cancer[20], breast cancer[21] and bladder cancer[22], and served as the prognostic factor in patients. Although Wang and Oshima et al. have shown that INHBA is highly expressed in GC tumor tissues and is a prognostic biomarker for patients with GC[23,24]. However, there is little systematical analysis on the exact role of INHBA in patients with GC. Consequently, it is necessary to explore the exact role and the underlying mechanisms of INHBA regarding GC.

The aim of our current research aimed to systematically assess the correlation of INHBA expression with GC survival, as well as the function and mechanism of INHBA in human GC. Typically, the mRNA expression of INHBA was detected in both gastric cancer tissues and normal tissues by using Tumor Immune Estimation Resource (TIMER), GEPIA2, GEO, Oncomine and UALCAN databases. Then, the significance of INHBA in predicting prognosis for GC was analyzed based on Kaplan Meier plotter database. The relationship between INHBA expression and immune infiltration levels in GC was explored in the TIMER database. Then, through MEXPRESS database, we explored whether INHBA expression is correlated with the changes of INHBA methylation in GC compared to normal sample. Besides, Gene Oncology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) were performed for INHBA to examine the underlying mechanism of action. These findings shed light on the important role of INHBA in GC and illustrated the potential mechanism related to immune infiltration in GC. It also provides a certain theoretical foundation for making early diagnosis, prognosis evaluation, and specific treatment for GC.

**Methods**

**Tissue collection**

All fresh specimens were collected from January 2018 to June 2019 at Hebei Medical University Fourth Affiliated Hospital. Resected tissues of gastric cancer and matched adjacent non-tumour gastric tissues (n=65) were snap-frozen in liquid nitrogen and stored at −80°C for qRT-PCR assay. None of the patients were treated by radiotherapy or chemotherapy before surgery.

**RNA extraction and quantitative real-time PCR**
RNA extraction was performed as described in the operating manual. Primers were purchased from GeneCopoeia (HQP017978, Eockville, MD, USA); the housekeeping gene GAPDH (ab9485; Bioworld, USA) served as an interval reference to normalize RNA abundance. The relative fold-changes in mRNA expression were expressed by the geometric mean and calculated using the $2^{-\Delta CT}$ method. All the reactions were conducted in triplicate and the values are shown as means±SD.

**Immunohistochemistry (IHC)**

INHBA expression was detected using IHC in tumour and adjacent normal tissues. Anti-INHBA monoclonal antibody was purchased from AB-clonal (1:60, a6614; ABclonal, Wuhan, China). All slides were interpreted by two pathologists independently who were blinded to the clinical information. The IHC staining score included the proportion of positively stained tumour cells and the staining intensity. The proportion of positively stained tumour cells was scored as follows: 0 (no tumour cells stained), 1 (<25% tumour cells stained), 2 (25%–50% tumour cells stained), 3 (50%–75% tumour cells stained), and 4 (75%–100% tumour cells stained). The grading of staining intensity was evaluated by the following criteria: 3 (brown, strong staining), 2 (yellow brown, moderate staining), 1 (light yellow, weak staining), and 0 (no staining). The final total staining score was calculated by multiplying the proportion of stained tumour cells and the staining intensity score. INHBA expression was scored and a score of $\leq 3$ indicated negative INHBA expression, while a score of $>3$ was considered as positive INHBA expression. A score of $\leq 6$ was designated as low expression and a score of $>6$ as high expression. The pathological diagnosis was made in accordance with the histological classification of tumours developed by the World Health Organization.

**Oncomine database analysis**

Oncomine ([www.oncomine.org](http://www.oncomine.org)) database contains 715 gene expression datasets and 867,33 cancers and normal samples, is also the biggest and user-friendly oncogene chip database and integrated data mining tool.[25] The DNA copy number and mRNA expression differences of INHBA gene between GC tumor and normal tissues were determined using the Oncomine database. In the present study, we drew on a series of GC studies, including Cho GC[26], DErrico GC[27], Deng GC[28], Cui GC[29], Chen GC[30], TCGA GC [31]and Wang GC[32] studies. The expression of INHBA was involved in evaluated in GC tissues in respect to its expression in normal tissues, and P<0.05 and foldchange of 1 as the cutoff criterion considered statistically significant.

**GEPIA2 database analysis**

GEPIA2 (Gene Expression Profiling Interactive Analysis, [http://gepia2.cancer-pku.cn/](http://gepia2.cancer-pku.cn/)) is an updated version of GEPIA for analyzing the transcript data of 198,619 isoforms and 84 cancer subtypes, including 9,736 tumor samples and 8,587 normal tissues samples from the TCGA and Genotype-Tissue Expression projects using a standard processing pipeline[33, 34]. In the present study, we used GEPIA2 to analyze the expression level of INHBA between GC tumor tissues and normal tissues.
UALCAN Database analysis

UALCAN (http://ualcan.path.uab.edu/) is a user-friendly web resource for analyzing cancer transcriptome data and in-depth analyses of gene expression, methylation information, and survival curves\[35\]. In the present study, we used UALCAN to evaluate the expression of INHBA in different tumor subgroups, such as tumor stages, grade, nodal metastasis status, gender, TP53 mutation status, race, age, historical subtypes, and HPV status.

TIMER database analysis

TIMER (https://cistrome.shinyapps.io/timer/) is an a comprehensive and user-friendly online tool to systematically investigate and visualize the correlation between immune infiltrates and a wide spectrum of factors, including gene expression, clinical outcomes and somatic mutations over 10,897 tumors from 32 cancer types\[36, 37\]. The differential expression of INHBA between tumor and normal tissues could be evaluated using Diff Exp module across all the TCGA database tumors and the results were shown with boxplots. The abundances of six immune infiltrates (CD8+, T cells, B cells, CD4+ T cells, Macrophages, Neutrophils and Dendritic cells) are assessed by statistical method, and we also analyze the comparison of tumor infiltration levels among tumors with different somatic copy number alterations for INHBA in GC. P-value < 0.05 was considered statistically significant.

Kaplan-Meier plotter database

Receiver operating characteristic (ROC) curve conducted by pROC\[38\] package in R software to explore the sensitivity and specificity for distinguishing GC patients from healthy individuals. Kaplan-Meier plotter (http://kmplot.com/) is an online database containing microarray gene expression data and survival information extrcated from Gene Expression Omnibus (GEO) and TCGA database, which contain gene expression data and survival data of 1065 GC patients\[39\]. And the valid number of GC patients were included in this study were after exclusive the patients, which missing expression values or don't have complete clinical data including survival time and status. In order to investigate the underlying prognostic value of INHBA, we performed Overall Survival (OS), First Progression Survival (FPS), Post Progression Survival (PPS) evaluated using Kaplan-Meier plotter database based on median expression (high vs. low) and assessed using Kaplan-Meier survival plot, with a hazard ratio with 95% confidence intervals and log rank p-value. Furthermore, the correlation between INHBA expression and differently kinds of clinicopathological characteristic also using this database, such as gender, human epidermal growth factor receptor 2 (HER2) status and lauren classification. Correlations between INHBA and prognosis in GC patients enriched immune infiltrates also evaluated using Kaplan-Meier plotter database.

cBioportal database

cBioportal (http://www.cbioportal.org/) is an interactive open-source platform, containing 245 cancer studies, which provides large scale cancer genomics data sets to visualize, analyze and download\[40, 41\]. The frequency of INHBA alterations (amplification, deep deletion, and missense mutations) in GC patients
were assessed using the cBioportal for Cancer Genomics database and TCGA. What's more, we using the Kaplan-Meier analysis in cBioportal to analyze the effect of INHBA expression dysregulation on the patients' overall survival and disease-free survival.

**MEXPRESS database analysis**

MEXPRESS ([https://mexpress.be/](https://mexpress.be/)) is an intuitive and user-friendly online web tool for the visualization of TCGA gene expression (normalized RNASEqV2 value), DNA methylation and clinical data, as well as the correlation between them on a interested single gene level\(^{[42, 43]}\). By defaults, the samples are sorted by the expression (from low to high) of the gene that was entered. The Pearson correlation is used to calculate the difference between expression value and methylation data. In the present study, we evaluated the correlation between INHBA expression and promoter methylation in GC.

**Functional enrichment analysis**

Metascape ([http://metascape.org](http://metascape.org)) is a new, free and user-friendly gene list analysis online tool to perform a functional enrichment analysis, which including cellular component (CC), biological process (BP), molecular function (MF), KEGG pathway analysis and protein-protein interaction analysis\(^{[44]}\). In the present study, Metascape was used to perform GO and KEGG pathway analysis of INHBA and neighboring genes significantly associated with INHBA. \(P\)-value < 0.05 was considered as significant.

**LinkedOmics analysis**

The LinkedOmics database ([http://www.linkedomics.orglogin.php](http://www.linkedomics.orglogin.php)) is a comprehensive and unique online tool to access, analyze and compare disseminating data from large-scale cancer omics projects within and across all 32 TCGA cancer type\(^{[45]}\). Linkedomics three analytical modules were applied to explore attributes that are associated with entered gene, perform functional enrichment analysis, and compare integrated association results. In the present study, we used the database to explore the genes differentially expressed related to INHBA in the TCGA STAD cohort. Then GSEA was utilized to perform analyses of Ge Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis.

**Statistical analysis**

R (version 3.6.5, United States) and GraphPad Prism version 8 (GraphPad Software, La Jolla, CA, United States) was utilized for statistical analysis in this study. Continuous variables were analyzed by the Student's \(t\) test. \(P < 0.05\) was considered as statistically significant.

**Results**

**INHBA expression in patients with GC**

The comparison of INHBA mRNA expression among various cancer types and normal tissues in TIMER database revealed significantly higher INHBA expression in bladder urothelial carcinoma (BLCA), breast
invasive carcinoma (BRCA), cholangiocarcinoma (CHOL), colon adenocarcinoma (COAD), esophageal carcinoma (ESCA), head and neck squamous cell carcinoma (HNSC), kidney chromophobe (KICH), kidney renal clear cell carcinoma (KIRC), rectum adenocarcinoma (READ) and stomach adenocarcinoma (STAD). By contrast, INHBA expression was significantly lower than adjacent normal tissues in kidney renal papillary cell carcinoma (KIRP), lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) (Figure 1A). And then we compared the expression level of INHBA in GC tumor tissues with this in normal tissues by using Oncomine and GEPIA databases. The results uncovered that expression of INHBA was significantly higher in GC tissues when compare with normal tissues ($P<0.05$) (Figure 1B, C).

When analyzed the expression level of INHBA in, respectively, Cho GC, DErrico GC, Deng GC, Cui GC, Chen GC, TCGA GC and Wang GC, GSE81948, GSE54129 and GSE13911 datasets, scatter plot showing the expression level of INHBA was fundamentally upregulated in GC patients tumor tissues compared with adjacent normal tissues (Figure 2). And previous studies confirmed that INHBA protein was highly expressed in GC tumor tissues by using immunohistochemistry and western blotting $^{[23, 46, 47]}$. Further analysis of diverse clinical pathological characteristics of 415 GC samples in the UALCAN database indicated higher transcriptional level of INHBA. In stage subgroup (normal-vs-Stage I, normal-vs-Stage II, normal-vs-Stage III and normal-vs-Stage IV) analysis, tumor grade subtype (normal-vs-grade 1, normal-vs-grade 2 and normal-vs-grade 3), nodal metastasis status subgroup (normal-vs-N0, normal-vs-N1, normal-vs-N2 and normal-vs-N3), gender subgroup (normal-vs-male and normal-vs-female), TP53 mutation status subgroup (normal-vs-TP53 mutant and normal-vs-nonmutant), race subgroup (normal-vs-Caucasian, normal-vs-African American and normal-vs-Asian), HPV status subgroup (normal-vs-with HPV infection and normal-vs-without HPV infection status) analysis the INHBA expression was fundamentally higher in GC patients, as well as in age subgroup analysis (Figure 3). These findings suggest that INHBA expression levels can serve as a potential diagnostic biomarker in GC. And we validated the protein and mRNA expression of INHBA in GC tissues, the results showed that INHBA was significantly upregulated in GC tissues compared to normal tissues (Figure 4).

**Diagnostic value of INHBA in patients with GC**

According to the difference of INHBA in GC, we get further to explore the diagnostic value of INHBA for distinguishing GC patients from healthy individuals, we performed ROC curve based on the data from Cho GC, DErrico GC, Deng GC, Cui GC, Chen GC, TCGA GC and Wang GC, GSE81948, GSE54129 and GSE13911 datasets, respectively. The results showed that INHBA had high diagnostic value for distinguishing GC patients from healthy individuals (Cho GC, AUC=0.670, 95 CI% [0.524-0.788], Figure 5A; DErrico GC, AUC=0.652, 95 CI% [0.517-0.759], Figure 5B; Deng GC, AUC=0.817, 95 CI% [0.768-0.864], Figure 5C; Cui GC, AUC=0.838, 95 CI% [0.776-0.901], Figure 5D; Chen GC, AUC=0.942, 95 CI% [0.874-0.975], Figure 5E; Wang GC, AUC=0.836, 95 CI% [0.636-0.917], Figure 5F; TCGA GC, AUC=0.756, 95 CI% [0.711-0.801], Figure 5G; GSE81948 GC, AUC=1, 95 CI% [1-1], Figure 5H; GSE54129 GC, AUC=0.991, 95 CI% [0.924-0.996], Figure 4I; GSE13911 GC, AUC=0.970, 95 CI% [0.877-0.973], Figure 5J).

**The correlation between the expression level of INHBA and prognosis in patients with GC**
According to the difference of INHBA in GC, we get further analyzed the correlation between INHBA expression and prognosis in patients with GC, so it is necessary to clarify whether INHBA act as the promoter or suppressor of GC. Kaplan Meier plotter was utilized to investigate the relationship between the expression of INHBA and the prognosis, including OS, FPS and PPS, in patients with GC. The results showed that the high expression of INHBA in GC was significantly related to the worse OS (HR=1.31; 95% CI [1.08-1.58]; \( P=0.015 \)), FPS (HR=1.32; 95% CI [1.08-1.61]; \( P=0.007 \)) and PPS (1.43; 95% CI [1.12-1.82]; \( P=0.004 \)), respectively (Figure 6A, 6B, 6I, 6J, 6Q, 6R). We further analyzed the correlation between the expression levels of INHBA and clinicopathological subtypes, such as gender, HER2 status and Lauren classification. In gender subtypes, high expression of INHBA in male patients with GC had poor OS (HR=1.34; 95% CI [1.06-1.69]; \( P=0.016 \)), FPS (HR=1.37; 95% CI [1.07-1.73]; \( P=0.010 \)) and PPS (1.62; 95% CI [1.22-2.16]; \( P=0.00085 \)), respectively (Figure 6A, 6D, 6I, 6L, 6Q, 6T). In HER2 status subtypes, high expression of INHBA in HER2-positive patients with GC had worse OS (HR=1.78; 95% CI [1.36-2.34]; \( P<0.011 \)), FPS (HR=2.15; 95% CI [1.53-3.01]; \( P<0.011 \)) and PPS (HR=1.94; 95% CI [1.35-2.79]; \( P<0.011 \)), respectively (Figure 6A, 6E, 6I, 6M, 6Q, 6U). In Lauren classification subtypes, high expression of INHBA in intestinal patients with GC had worse OS (HR=1.65; 95% CI [1.18-2.32]; \( P=0.004 \)) and FPS (HR=1.60; 95% CI [1.09-2.35]; \( P=0.016 \)) (Figure 6A, 6G, 6I, 6O), however, high expression of INHBA in diffuse patients with GC had worse PPS (HR=1.69; 95% CI [1.11-2.57]; \( P=0.014 \)) (Figure, 6Q, 6X). These results suggested that INHBA may serve as a potential biomarker for specific subtypes gastric cancer.

**Immune infiltrates in correlation with INHBA in GC**

Previous studies had shown that tumor infiltration was significantly related to the progression and prognosis of GC\(^{[48-50]}\). So, we utilized TIMER database to investigate whether the expression levels of INHBA in GC tumor was correlated with immune infiltration. The results showed that INHBA was negatively correlated with B cell while positively correlated with Macrophage, Neutrophil and Dendritic cell infiltration (\( P<0.05 \), Figure 7A). Cumulative survival analysis uncovered that Macrophage of immune infiltrates statistically significant (\( P<0.05 \)) of INHBA in GC indicating that Macrophage negatively affecting the prognosis, it is worthy to get further investigate (Figure 7B). Finally, somatic copy number alterations are characterized by GISTIC 2.0, including deep deletion (-2), arm-level deletion (-1), diploid/normal (0), arm-level gain (1), and high amplification (2). Box plots are presented to show the distributions of each immune subset at each copy number status with INHBA in GC (Figure 7C).

**Relationship between INHBA expression and immune cell type markers in GC**

We feather analyzed the correlation between the expression of INHBA and different immune cells type markers in GC based on TIMER database. The results showed that INHBA in GC was positively correlated with CD38 in B cells (Table 1). INHBA in GC was also positively correlated with CD8A in CD8+ T cells. Similarly, INHBA in GC was positively correlated with MPO, FCGR3B, FPR1, CSF3R in Neutrophils and CD209 in dendritic cells (Table 1). INHBA in GC was positively correlated with CD68, CD84, CD163, MS4A4A in macrophages (Table 1). These results further confirmed that INHBA in GC were correlated to immune infiltration.


Table 1. Correlation analysis between INHBA and immune cell type markers in TCGA-STAD cohort via TIMER database.

<table>
<thead>
<tr>
<th>Cell type</th>
<th>Gene marker</th>
<th>None</th>
<th></th>
<th>None</th>
<th></th>
<th>None</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Cor</td>
<td>P</td>
<td>Cor</td>
<td>P</td>
<td>Cor</td>
<td>P</td>
</tr>
<tr>
<td>B cells</td>
<td>CD19</td>
<td>-0.023</td>
<td>6.04E-01</td>
<td>-0.041</td>
<td>4.28E-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FCRL2</td>
<td>0.021</td>
<td>6.66E-01</td>
<td>-0.005</td>
<td>9.20E-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CD38</td>
<td>0.136</td>
<td>5.00E-03</td>
<td>0.091</td>
<td>7.56E-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MS4A1</td>
<td>-0.021</td>
<td>6.74E-01</td>
<td>-0.061</td>
<td>2.38E-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD8+ T cells</td>
<td>CD8A</td>
<td>0.120</td>
<td>1.50E-02</td>
<td>0.083</td>
<td>1.09E-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CD8B</td>
<td>0.037</td>
<td>4.51E-01</td>
<td>0.017</td>
<td>7.47E-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neutrophils</td>
<td>MPO</td>
<td>0.223</td>
<td>4.50E-06</td>
<td>0.215</td>
<td>2.38E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FCGR3B</td>
<td>0.153</td>
<td>1.73E-03</td>
<td>0.130</td>
<td>1.16E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FPR1</td>
<td>0.334</td>
<td>5.68E-13</td>
<td>0.328</td>
<td>5.92E-11</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CSF3R</td>
<td>0.287</td>
<td>2.52E-09</td>
<td>0.263</td>
<td>1.94E-07</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S100A12</td>
<td>0.064</td>
<td>1.96E-01</td>
<td>0.030</td>
<td>2.66E-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macrophages</td>
<td>CD68</td>
<td>0.209</td>
<td>1.76E-05</td>
<td>0.196</td>
<td>1.27E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CD84</td>
<td>0.208</td>
<td>1.98E-05</td>
<td>0.193</td>
<td>1.54E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CD163</td>
<td>0.312</td>
<td>8.29E-11</td>
<td>0.281</td>
<td>2.67E-08</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MS4A4A</td>
<td>0.310</td>
<td>1.10E-10</td>
<td>0.292</td>
<td>6.76E-09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dendritic cells</td>
<td>CD209</td>
<td>0.137</td>
<td>9.14E-03</td>
<td>0.101</td>
<td>5.00E-02</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CD1C</td>
<td>0.005</td>
<td>9.24E-01</td>
<td>-0.018</td>
<td>8.01E-01</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

STAD, stomach adenocarcinoma; Cor, r value of Spearman’s correlation; Purity, correlation adjusted by purity; Bold represents \( P<0.05 \).

**Prognostic analysis of INHBA expression in GC based on immune cells**

We have confirmed that the INHBA expression was correlated with the immune infiltration in GC, and the expression of INHBA was also related to the poor prognosis of the patients with GC. Thus, we speculated that expression of INHBA in GC affected the prognosis partly due to immune infiltration. Then we performed prognostic analysis based on the expression levels of INHBA of GC in related immune cells subgroup via Kaplan Meier plotter database. The results uncovered that the high expression of INHBA of GC in enriched CD4+ T cells (HR=2.12; 95% CI [1.17-3.84]; \( P=0.011 \)), CD8+ T cells (HR=1.97; 95% CI [1.12-
3.18; \( P=0.0046 \) and Macrophage (HR=1.79; 95% CI [1.05-3.03]; \( P=0.029 \)) cohort had worse overall survival respectively (Figure 8B, 8C, 8D), while B cells cohort had no statistical significance (Figure 7A). The high expression of INHBA of GC in enriched B cells (HR=2.62; 95% CI [1-6.85]; \( P=0.042 \)), CD4+ T cells (HR=4.08; 95% CI [1.35-12.32]; \( P=0.0073 \)), CD8+ T cells (HR=6.91; 95% CI [1.57-30.38]; \( P=0.0034 \)) and Macrophage (HR=3.24; 95% CI [0.94-11.12]; \( P=0.048 \)) cohort had worse relapse free survival respectively (Figure 8E, 8F, 8G, 8H). The above results suggested that high INHBA expression in GC may affect prognosis partly because of immune infiltration.

**Genetic alteration differences of INHBA in GC**

In order to explore the sequence alterations of INHBA in GC, we then used cBioPortal database to investigate the sorts and frequency of INHBA modification in GC from STAD patients sequencing data in the TCGA. As shown in Figure 9A 69 of 1590 (4%) GC patients were altered. Further study suggested that mRNA upregulation and mutation are the most common types of INHBA in patients with GC (Figure 9B). Besides, the results of Kaplan-Meier plotter and log-rank test demonstrated no significantly statistical difference in overall survival (OS) and disease-free survival (DFS) in cases with and without INHBA alterations (\( P \)-value was 0.972 and 0.524, respectively. Figure 9C, 9D)

**The correlation between INHBA expression and methylation around the promoter region**

Previous studies had shown that DNA promoter methylation is a meaningful pattern which affect tumorigenesis of tumors\([51-53]\). To explore the correlation between INHBA expression and DNA methylation, then the methylation levels of INHBA in GC were performed by using of MEXPRESS database. Figure 10 shows that default MEXPRESS plot for INHBA in gastric cancer in the samples sorted based on the INHBA expression value. The results showed week significant methylation level change between tumor and normal tissues, which indicates that INHBA expression might not controlled by DNA methylation.

**Functional enrichment analysis of genes co-expressed with GC**

In order to uncovered the potential function of the INHBA, we performed protein-protein interaction (PPI) network, GO function and KEGG pathway enrichment analysis via Metascape database. The PPI network as shown in Figure 11A and 11B. As Figure 11C-E shows that the most significant enriched GO terms were regulation transmembrane receptor protein serine/threonine kinase signaling pathway, BMP signaling pathway, SMAD protein signal transduction, nodal signaling pathway, cell proliferation and metabolic process. The most enriched KEGG pathways were TGF-beta signaling pathway and PID ALK1/2 pathway.

**GO function and KEGG pathway enrichment analysis of co-expression genes correlated with INHBA in GC**

LinkedOmics was utilized to uncovered mRNA sequencing information from GC patients in the TCGA-STAD cohort. Spearman's test was conducted to analyze correlations between INHBA and genes
differentially expressed in GC (red represents positively related genes and green represents negatively related genes) (Figure 12A). The top 50 genes that were positively and negatively correlated with INHBA were shown in heat maps (Figure 12B-C). Significantly GO and KEGG functional enrichment analysis were conducted by gene set enrichment analysis (GSEA) suggested that these genes differentially expressed in correlation with INHBA in GC were mainly enriched collagen metabolic process, extracellular structure organization, cellular response to vascular endothelial growth factor stimulus, connective tissue development and DNA replication, and so on biological process (Figure 12D). Essentially molecular functions and cellular component were collagen binding, extracellular matrix structural constituent, growth factor, Wnt-protein binding and SMAD binding, collagen trimer, endoplasmic reticulum lumen and so on (Figure 12E-F). KEGG pathway analysis showed that cancer-related signaling pathways were enriched, including TGF-beta signaling pathway, ECM-receptor signaling pathway, AGE-RAGE signaling pathway and so on (Figure 12G).

Discussion

GC is a heterogeneous malignancy worldwide with high recurrence probability and unfavorable prognosis[8, 54]. Currently, effective molecules have been identified in GC diagnosis and therapy by mechanistic analysis, such as CASC2[55], PROX1[56], HOXA11[57], circLMTK2[58] and VCAM-AS1[59], however, the 5-year survival rates in GC patients remain frustrating. Therefore, it is imperative to search a promising biomarkers and therapeutic targets relevant to GC diagnosis and treatment. As an important member of the TGF-β, INHBA is associated with tumorigenesis and tumor progression in multiple types of solid tumors. Si et al. found that high expression of INHBA is an adverse prognostic factor for de novo acute myeloid leukemia[60]. Wamsley et al. demonstrates that INHBA mRNA and protein expression are elevated in human non-small cell lung cancer (NSCLC) and INHBA is a critical autocrine factor that maintains mesenchymal properties of cancer-initiating cells in NSCLC[61]. Seder et al. found INHBA is upregulated in lung adenocarcinoma (AD) relative to controls, and its overexpression is associated with worse survival in stage I AD patients and overexpression of INHBA may be affected by promoter methylation[14]. Studies carried out by Okano et al[62], and Li et al[63]. showed that INHBA expression was significantly higher in colorectal cancer (CRC) compared with corresponding normal tissues and INHBA is useful as a predictive biomarker for prognosis in CRC patients. In addition, several studies investigated the role of epigenetic regulation of INHBA gene expression in CRC[64] and esophageal cancer[65], and the results indicted that overexpression of INHBA may be affected by promoter methylation. Hofland et al[66] showed that INHBA upregulated in prostate cancer (PC) and which is controlled by androgens in PC models and regulates local androgen production. Liu et al[67]. showed that INHBA and GJB2 as key regulators for the transition of benign ductal carcinoma in situ to aggressive phenotype and might serve as promising prognosis biomarkers.

In the study, the changes of INHBA mRNA in GC were analyzed in TIMER, GEPIA2, GEO, Oncomine and UALCAN databases, meanwhile, we get further evaluated the expression of INHBA in different tumor subgroups, including tumor stages, grade, nodal metastasis status, gender, TP53 mutation status, race,
age, historical subtypes, and HPV status, the results showed that INHBA mRNA expression level is higher in GC compared with normal tissue. Our validated experiments and previous studies suggested that INHBA protein was highly expressed in GC tumor tissues by using immunohistochemistry and western blotting\textsuperscript{[23, 46, 47]}, which strongly indicated that INHBA was fundamentally upregulated in GC. In order to uncover the causes of upregulated INHBA in GC, we analyzed the methylation levels of INHBA in GC performed by using of MEXPRESS database. The results showed week significant methylation level change between tumor and normal tissues, which indicates that INHBA expression might not controlled by promoter methylation. ROC analysis showed that INHBA have high diagnostic value for distinguishing GC patients from healthy individuals. The Kaplan Meier plotter was used to explore the correlation between the expression level of INHBA and the prognosis in patients with GC. The results showed that the high expression level of INHBA in GC was significantly related to the unfavorable prognosis in GC. In addition, clinicopathological subtypes analysis suggested that INHBA may serve as a potential prognostic biomarker for specific subtypes GC, including male, HER2 positive and intestinal classification.

Previous studies had shown that tumor infiltration was significantly related to the progression and prognosis of GC\textsuperscript{[48-50]}. Therefore, TIMER database was utilized to investigate the relationship between INHBA expression level and immune infiltration in GC. The results showed that INHBA was negatively correlated with B cell while positively correlated with Macrophage, Neutrophil and Dendritic cell infiltration. The immune cell type markers in GC was further explored, the results showed that INHBA in GC was positively correlated with CD8A in CD8+ T cells and MPO, FCGR3B, FPR1, CSF3R in Neutrophils. Similarly, INHBA in GC was also positively correlated with CD68, CD84, CD163, MS4A4A in macrophages and CD209 in dendritic cells. These results further confirmed that INHBA in GC were correlated to immune infiltration. Prognostic analysis of INHBA expression level in GC based on immune cells subgroup was performed, the results uncovered that the high expression of INHBA in GC had a worse prognosis in enriched B cells, CD4+ T cells, CD8+ T cells and Macrophage. The above results strongly suggested that high INHBA expression in GC may affect prognosis partly because of immune infiltration.

In order to further analyze the role of INHBA in regulating the progression and immune microenvironment in GC, the genes co-expressed with INHBA in GC were conducted using LinkedOmics database. Further, GO functional and KEGG enrichment analysis were examined using LinkedOmics and Metascape databases. The results showed that INHBA mainly enriched cancer-related signaling pathways, including TGF-beta signaling pathway, ECM-receptor signaling pathway, PID ALK1/2 pathway, and AGE-RAGE signaling pathway. Which consistent with Chen et al\textsuperscript{[68]} analysis. They showed that INHBA expression elevated in GC tissues and activated GC cells migration, invasion and proliferation abilities. In addition, INHBA gene silencing inhibited the progression of GC by inactivating the TGF-β pathway.

**Conclusions**

Our findings suggest that INHBA is highly expressed in GC and may serve as a potential biomarkers for predicting prognosis. It may theoretically affect the prognosis of GC through immune infiltration.
mechanism, which may provide a new insight for future in-depth study.

**Abbreviations**


**Declarations**

**Acknowledgments**

There was no funding body having role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Availability of data and materials**

The datasets supporting the conclusions of this article are included within the article and additional files.

**Authors’ contributions**

GJ L and F L performed and analyzed the data and wrote the manuscript. XJ H, YX J and BJ H collect the data form the data set. L Y helped analyzing the data and critically revised the manuscript. All authors approved the final manuscript and consented for publication.

**Ethics approval and consent to participate**

This study was approved by the ethics committee of Hebei Medical University Fourth Affiliated Hospital (2019054), and informed consent was obtained from all patients.

**Competing interests**

The authors declare that they have no competing interests.

**Author details**

1Department of Thoracic Surgery, Hebei Medical University Fourth Affiliated Hospital, Shijiazhuang, Hebei, 050001, China. 2Department of General Surgery, Hebei Provincial People’s Hospital, Shijiazhuang, Hebei, 050055, China. 3Department of Clinical Medicine, school of Basic Medicine, Hebei Medical University, Shijiazhuang, Hebei, 050017, China. 4Department of Chinese Medicine, Hebei Medical University Fourth Affiliated Hospital, Shijiazhuang, Hebei, 050001, China. 5Department of Epidemiology and Statistics, School of Public Health, Hebei Medical University, Shijiazhuang, Hebei, 050017, China.
References


Figures
Figure 1

INHBA expression levels in different human cancers. (A) INHBA expression profile across all tumor samples and normal tissues analyzed by TIMER database. (B) increased or decreased INHBA in data sets of different cancers compared with normal tissues in the Oncomine database. (C) the expression level of INHBA in GC tumor tissues and paired normal tissues determined by GEPIA database. *: P<0.05, **: P<0.01, ***: P<0.001, NS: P>0.05.
Figure 2

Relative expression level of INHBA between the GC tissues and non-cancerous tissues in ten cohorts determined by Oncomine and GEO databases. *: P<0.05, **: P<0.01, ***: P<0.001. (A-G) Determined by Oncomine database. (H-J) Determined by GEO database.
Boxplot showing relative expression of INHBA in subgroups of patients with gastric cancer (UALCAN). (A) Transcriptional expression level of INHBA between the GC tissues and non-cancerous tissues. (B-J) Boxplot showing tumor stages, grade, nodal metastasis status, gender, TP53 mutation status, race, age, historical subtypes, and HPV status with INHBA expression in gastric cancer (GC), respectively. (N0: No regional lymph node metastasis; N1: metastases in 1 to 3 axillary lymph nodes; N2: metastases in 4 to 9 axillary lymph nodes; N3: metastases in 10 or more axillary lymph nodes. *: P<0.05, **: P<0.01, ***: P<0.001, NS: P>0.05.)
The protein and mRNA expression of paired gastric cancer. (A-C) Comparison of INHBA protein expression in 62 paired normal gastric tissues and gastric cancer tissues. (D) Relative mRNA expression of INHBA in 65 paired normal gastric tissues and cancer tissues. ***, represents $P < 0.001$.
Figura 5

ROC curve of INHBA for distinguishing GC patients from healthy individuals based on the data from Cho GC, DErrico GC, Deng GC, Cui GC, Chen GC, TCGA GC and Wang GC, GSE81948, GSE54129 and GSE13911 cohort, respectively.
Figure 6

The correlation of INHBA expression with survival analysis in gastric cancer. Kaplan-Meier plotter was applied to evaluate the prognostic value of INHBA. Survival analysis in patients within gender, HER2 status and Lauren classification. (A-H: overall survival, OS; I-P: first progression survival, FPS; Q-X: post progression survival, PPS).
Figure 7

Immune infiltrates in correlation with INHBA in B cells, CD4+ T cells, CD8+ T cells, Neutrophils, Macrophage and Dendritic cells of gastric cancer (TIMER). (A): Correlation between INHBA expression and abundance of immune infiltrates. (B): Clinical outcome and abundance of immune infiltrates of INHBA expression. (C): Correlation between somatic copy number alterations and abundance of immune infiltrates of INHBA expression. *: P<0.05, **: P<0.01, ***: P<0.001.
Comparison of Kaplan-Meier curves of the high and low expression of INHBA in GC based on immune cells subgroup. (A-D) Relationship between INHBA expression of different immune cells subgroup and overall survival in STAD. (E-H) Relationship between INHBA expression of different immune cells subgroup and Relapse Free Survival in STAD. Notes: *: P<0.05, **: P<0.01, ns: P>0.05.
Figure 9

Visual summary of INHBA alterations in gastric cancer (Cbioportal databases). (A) OncoPrint of INHBA genetic alterations in GC. (B) The genetic alteration type and frequency of INHBA were studied in various gastric cancer. (C) Overall survival. (D) Progression-free survival.
**Figure 10**

Visualization of the TCGA data for INHBA in gastric cancer using MEXPRESS database, the samples were ordered by their expression value. The view showed the correlation between INHBA expression and promoter region and clinical features, as well as CNVs with the Pearson correlation coefficients on the right side. The height of the dark green line represents the INHBA expression value (normalized RNASeqV2 in TCGA) and the beta value for the Infinium 450k probes.
Figure 11

The protein-protein interaction network and functional enrichment analysis of INHBA (Metascape database). (A-B) Protein-protein interaction network and MCODE components identified in the gene lists. (C-D) Bar graph of enriched terms, colored by p-values. (E) Network of enriched terms colored by cluster ID, where nodes that share the same cluster ID are typically close to each other.
Figure 12

Genes differentially expressed in correlation with INHBA in gastric cancer and functional enrichment analysis (LinkedOmics database). (A) Volcano plot of correlation between INHBA and genes differentially expressed in GC using T-test. (B-C) The heatmaps showing genes positively and negatively correlated with INHBA in GC. Red indicates positively correlated while blue indicates negatively correlated genes. (D) Biological processes. (E) Molecular function. (F) Cellular components. (G) KEGG pathway analysis.