Journal of Advanced Biomedical Sciences

Journal of Advanced Biomedical Sciences
https://jabs.fums.ac.ir/
Online ISSN: 2783-1523

Fasa University of Medical Sciences

Bioinformatic Analysis of lncRNA H19 Reveals Its Regulatory Networks and Clinical Relevance in Gastric Cancer

Kianoush Mohammadi[1], Reza Safaralizadeh[1], Elham Safarzadeh[2]

1. Department of Animal Biology, Faculty of Natural Sciences, University of Tabriz, Tabriz, Iran
2. Cancer Immunology and Immunotherapy Research Center, Ardabil University of Medical Sciences, Ardabil, Iran

Article Info

Article Type:
Original Article

Article history:
Received
06 Sep 2025
Received in revised form
30 Sep 2025
Accepted
06 Oct 2025
Published online
10 Oct 2025

Publisher
Fasa University of Medical Sciences

Abstract

Background & Objectives: Gastric cancer (GC) continues to rank among the leading causes of cancer-related mortality worldwide, primarily because of late-stage diagnosis, marked molecular heterogeneity, and the emergence of therapeutic resistance. The long non-coding RNA (lncRNA) *H19* has been recognized as an oncogene in multiple malignancies; however, its precise molecular mechanisms and clinical significance in GC remain incompletely understood.

Materials & Methods: We conducted an integrative bioinformatics analysis of 431 TCGA-STAD (stomach adenocarcinoma) samples, integrating somatic mutation, RNA-seq, and clinical datasets. The study examined mutational landscapes, tumor mutational burden (TMB), and distinct mutational signatures. Patients were classified according to *H19* expression levels for subsequent differential expression, correlation, pathway enrichment, protein–protein interaction (PPI) network construction, and survival analyses.

Results: The most frequent mutations were identified in *TTN* (51%), *TP53* (46%), *MUC16* (31%), *ARID1A* (27%), and *LRP1B* (27%). Six distinct mutational signatures were detected, reflecting processes associated with aging, mismatch repair deficiency, POLE-driven hypermutation, and prior chemotherapy exposure. Stratification based on *H19* expression revealed 15,179 differentially expressed genes that were significantly enriched in pathways related to extracellular matrix organization, focal adhesion, and cell adhesion. *H19* exhibited strong positive correlations with *IGF2*, *TCF15*, and *miR-675*, suggesting a potential competing endogenous RNA (ceRNA) function**,** and negative correlations with *ATP4A* and *ATP4B*, indicating possible disruption of parietal cell activity. The hub genes identified within the PPI network included *GAPDH*, *COL1A1*, *TGFB1*, and *SIRT1*.

Conclusion: Collectively, these findings suggest that *H19* acts as a pivotal regulator in GC by modulating ceRNA networks, promoting extracellular matrix remodeling, and influencing oncogenic signaling cascades. Although its independent prognostic significance has yet to be fully established, this comprehensive systems-level analysis provides valuable insights and lays the groundwork for future experimental and clinical studies exploring H19 as a potential diagnostic biomarker and therapeutic target.

Keywords: Gastric cancer, Long non-coding RNA H19, Bioinformatics analysis, ceRNA network, Tumor heterogeneity

Cite this article: Mohammadi K, Safaralizadeh R, Safarzadeh E. Bioinformatic Analysis of lncRNA H19 Discovers its Regulatory Networks and Clinical Relevance in Gastric Cancer. J Adv Biomed Sci. 2025; 15(4): 381-395.
DOI: 10.18502/jabs.v15i4.19738

Corresponding Author: Reza Safaralizadeh, Department of Animal Biology, Faculty of Natural Sciences, University of Tabriz, Tabriz, Iran.
Email: safaralizadeh@tabrizu.ac.ir

Introduction

Gastric cancer (GC) is one of the leading causes of cancer-related mortality worldwide. Despite advances in surgery, chemotherapy,

and targeted therapy, patients with advanced disease experience dismal outcomes due to late presentation, tumor heterogeneity, and drug resistance. The mechanism of GC pathogenesis is initiated by genetic mutations, epigenetic dysregulation, and aberrant signaling pathways, ultimately resulting in tumor initiation and growth (1). Molecularly targeted agents such as HER2 (human epidermal growth factor receptor), VEGFR (Vascular Endothelial Growth Factor Receptor), CLDN18.2, and FGFR2b inhibitors have demonstrated efficacy in specific patient populations, while immune checkpoint inhibitors targeting PD-1/PD-L1 and CTLA-4 can induce durable responses in certain subsets of patients. Nevertheless, most patients do not achieve sustained benefit, underscoring the need for novel biomarkers to improve patient stratification and guide therapy (1).

Mutational signatures provide critical insights into the underlying biological processes of GC. These signatures reflect DNA damage and repair mechanisms, including mismatch repair deficiency, homologous recombination deficiency, and exposure to environmentally acquired carcinogens. Certain signatures are correlated with clinical behavior and molecular subtypes, thereby influencing treatment response. Tumor-specific, immunogenic signatures particularly predict responses to immune checkpoint blockade. Consequently, mutational signature analysis is emerging as an increasingly important precision medicine tool for GC (2).

In addition to mutations, long noncoding RNAs (lncRNAs) play a central role in cancer biology. Among them, lncRNA H19 has been identified as a multifunctional oncogene across numerous cancers. H19 overexpression is strongly associated with advanced tumor stage, lymph node metastasis, distant metastasis, and reduced overall survival in GC (3-7). Meta-analyses further support its prognostic value across cancer types, ethnic populations, and study designs, demonstrating its robustness as a biomarker (3).

Mechanistically, H19 promotes oncogenesis through multiple pathways. H19 functions **as** a tumor-suppressive microRNA sponge, a regulator of gene expression, and a binder of chromatin-modifying complexes (4, 8). In GC, H19 modulates proliferation, migration, invasion, and colony formation, while suppressing apoptosis in vitro and in vivo (5-7). It also enhances tumor development and metastatic potential in animal models (7). Oncogenic signaling regulates H19 expression. Specifically, H19 is transcriptionally controlled by c-Myc, thereby amplifying its tumor-promoting activity. Elevated H19 expression correlates once more with poor clinical outcomes, reaffirming its prognostic significance (6).

H19 is additionally involved in post-transcriptional regulation and signaling pathways. It serves as a precursor for miR-675 and modulates oncogenic signaling pathways such as PI3K/AKT (phosphatidylinositol 3'–kinase/ protein kinase B), MAPK (mitogen-activated protein kinase), and Wnt/β-catenin (7, 8). Bioinformatics analyses have identified ceRNA networks such as H19/miR-29a-3p/LOX and H19/miR-107/COL1A1 that are strongly associated with tumor aggressiveness and patient survival (9). These findings confirm that H19 regulates key transcriptional and post-transcriptional programs in GC.

Immune mechanisms further extend its functional repertoire. The H19/miR-378a-5p/SERPINH1 axis has been shown to modulate immune cell infiltration in GC by altering macrophage and T-cell function. Patients with high H19 or SERPINH1 expression exhibit poorer survival rates, indicating that H19 plays a pivotal role in shaping the tumor immune microenvironment (10).

In addition to differential expression, genetic polymorphisms of H19 contribute to susceptibility to GC. Specific SNPs (single nucleotide polymorphisms), including rs217727

and rs2839698, are associated with increased risk in the Chinese Han population, particularly within subgroups defined by age, gender, and lifestyle factors (11). A meta-analysis across multiple cancers further confirmed the association of certain H19 variants, predominantly in gastrointestinal cancers and Asian populations, highlighting ethnic- and tumor type-specific genetic predictors (12). These findings support the utility of H19 polymorphisms as potential genetic markers for risk assessment.

Due to its stability and detectability in body fluids, H19 also represents a promising non-invasive diagnostic marker. ROC curve analysis has demonstrated its ability to distinguish GC patients from healthy controls (5). Clinically, H19 consistently predicts poor prognosis and malignant disease progression, further solidifying its role as a diagnostic and predictive biomarker (3-7, 13). Therapeutic strategies aimed at silencing H19 expression have shown preclinical efficacy, positioning it as a potential therapeutic target (4, 7, 13).

Collectively, these data position H19 as a central oncogenic driver in GC, influencing tumorigenesis through genetic, transcriptional, post-transcriptional, and immunological mechanisms. Despite substantial evidence, the molecular programs regulating H19 expression and their intersection with the mutational landscape remain incompletely understood.

To address this knowledge gap, the present study integrates mutational, transcriptomic, and clinical data from The Cancer Genome Atlas (TCGA) stomach adenocarcinoma cohort. By combining mutational signature analysis, differential expression profiling, ceRNA network construction, and survival analysis, we aim to provide a systems-level understanding of H19 in GC. This comprehensive approach will clarify its biological roles and clinical significance, thereby enhancing its potential as a biomarker and therapeutic target in precision oncology.

Materials and Methods

Preprocessing and Data Acquisition

Somatic mutation, RNA-seq, and miRNA expression data of STAD were downloaded from The Cancer Genome Atlas (TCGA) using the TCGAbiolinks package. Mutation data in MAF format were filtered to retain only non-silent variants. For transcriptome profiling, raw counts of lncRNAs and miRNAs aligned using STAR were retrieved. Low-abundance transcripts (lncRNAs <1×10^6, miRNAs <1×10^5) were excluded, and gene expression data were normalized using DESeq2 variance-stabilizing transformation (VST) or log2 (TPM + 1) to ensure comparability across samples. Clinical annotations, including sex, age, stage, grade, and survival status, were obtained from TCGA and harmonized for consistency.

Mutation Landscape and Tumor Mutational Burden

The mutation landscape of TCGA-STAD was characterized using maftools. Recurrently mutated genes were identified and visualized via oncoplots and mutation summary dashboards. Tumor mutational burden (TMB) was estimated by dividing the number of somatic coding mutations by the estimated exome size in megabases (Mb). Mutational signature analysis was performed using NMF on the trinucleotide mutation matrix, enabling the derivation of six de novo mutational processes. The signatures were then cross-referenced with COSMIC reference signatures to infer underlying mutagenic mechanisms, including mismatch repair deficiency, POLE mutations, and chemotherapy-induced signatures.

Transcriptome-Wide Analyses

Differential gene expression analysis was conducted on RNA-seq data using DESeq2. Patients were stratified into high- and low-H19 expression groups based on the median H19 expression across tumor samples. This median-based stratification ensures balanced group sizes and allows assessment of relative differences

within tumors rather than comparisons between tumor and normal tissues. Notably, negative log2 fold-change values reflect relative expression differences between high- and low-H19 subgroups within tumors without contradicting the documented overexpression of H19 in cancer. Genes with an adjusted p-value <0.05 were considered statistically significant. Top-ranked gene heatmaps were visualized using pheatmap. Positively and negatively co-expressed genes of H19 were identified through correlation analyses employing Spearman's method.

To evaluate biological relevance, highly differentially expressed genes were subjected to enrichment analysis using clusterProfiler. Gene Ontology biological processes and Kyoto Encyclopedia of Genes and Genomes pathways were assessed with mapped org. Hs.eg.db annotations (14, 15). Furthermore, co-expression of H19 with candidate miRNAs (e.g., hsa-miR-675, hsa-miR-21, hsa-miR-200a) was determined based on correlation statistics. miRNA–lncRNA interactions were visualized using ggplot2 barplots.

Multivariate Analysis of H19 and Clinical Outcomes

Clinical and transcriptomic data from 360 GC patients were analyzed to investigate the association of H19 expression with cancer progression and mortality. Tumor grade and stage were dichotomized (Stage I–II versus III–IV; Grade G1–G2 versus G3) after standardizing H19 expression as z-scores. Covariates included age and sex. Multivariable logistic regression was applied to assess associations with stage and grade, while multivariable Cox proportional hazards models evaluated associations with mortality. Multicollinearity and proportional hazards assumptions were tested to validate model integrity.

Network and Survival Analyses

PPI networks were constructed for differentially expressed genes using the STRINGdb interface. Networks were subsequently converted to graph objects using igraph to facilitate hub gene identification based on degree centrality.

Survival analysis was performed by integrating H19 expression with TCGA clinical data. Patients were stratified into high- and low-H19 groups using a median split. Kaplan–Meier survival plots were generated with survminer, and statistical significance was assessed via log-rank tests. Univariate and multivariate Cox proportional hazards regression models (via the survival package) were employed to examine the prognostic impact of H19 expression while adjusting for clinical covariates, including age, sex, and tumor stage, as appropriate. All analyses were conducted in R (version 4.5.1) (Table 1).

Results

Mutation Landscape of TCGA-STAD Cohort

To characterize the genomic landscape of STAD, we analyzed somatic mutations in 431 TCGA-STAD samples. Across all samples, 137,650 non-silent mutations were identified, with missense mutations representing the most prevalent class. Although less frequent than missense mutations, insertion/deletion events, nonsense mutations, and splice-site mutations also contributed substantially to the overall mutational spectrum. Among SNPs, C>T transitions were the most common, consistent with previous reports attributing these substitutions to spontaneous 5-methylcytosine deamination and mismatch repair deficiency, which are dominant mutational pathways in GC (Figure 1).

Out of 431 samples, 392 (90.95%) harbored at least one non-silent mutation. The most frequently recurrently mutated genes included TTN (51.2%), TP53 (45.6%), MUC16 (31.0%), ARID1A (26.9%), and LRP1B (26.6%). Additional genes, including CSMD3, SYNE1, FAT4, FLG, and PCLO, exhibited mutation frequencies ranging from 19% to 24%.

Table 1. R/Bioconductor packages and their utilization in TCGA-STAD mutation and transcriptome dataset analysis.

Package	Ver.	Purpose in the Study	Ref.
TCGAbiolinks	2.36.0	Data acquisition from TCGA (mutations, RNA-seq, clinical)	(16)
maftools	2.24.0	Mutation visualization, mutational signatures, TMB	(17)
GenomicRanges	1.60.0	Handling genomic intervals	(18)
biomaRt	2.64.0	Gene annotation, ID mapping, sequence retrieval	(19)
DESeq2	1.48.0	Differential expression analysis	(20)
clusterProfiler	4.16.0	GO and KEGG enrichment analysis	(21)
org.Hs.eg.db	3.21.0	Gene annotation database for Homo sapiens	(22)
pheatmap	1.0.13	Heatmap visualization of gene expression data	(23)
ggplot2	3.5.2	Data visualization (barplots, correlation plots)	(24)
survival	3.8-3	Cox proportional hazards modeling	(25)
survminer	0.5.0	Kaplan–Meier survival curve visualization	(26)
STRINGdb	2.20.0	Protein–protein interaction network construction	(27)
igraph	2.1.4	Network analysis and visualization	(28)

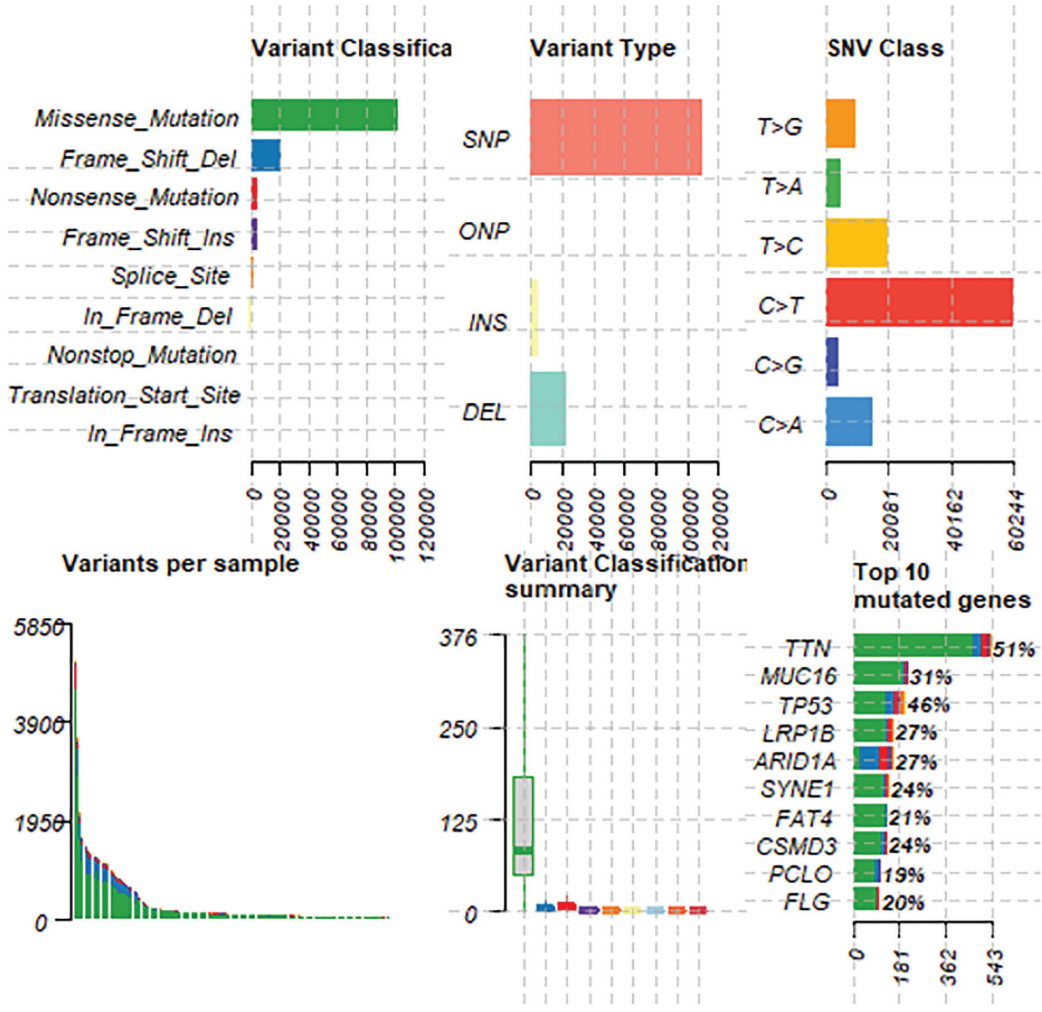


Figure 1. Landscape of somatic mutations in the TCGA-STAD cohort. Panels display (top left) distribution of variant classifications, (top center) variant types, and (top right) single nucleotide variant (SNV) classes. The bottom panel shows variants per sample, a summary of classification totals, and the 10 most frequently recurrently mutated genes (TTN, TP53, MUC16, LRP1B, ARID1A, SYNE1, FAT4, CSMD3, PCLO, and FLG) with mutation rates from 20% to 51%.

Other recurrently affected genes included HMCN1, ACVR2A, ZFHX4, DNAH5, OBSCN, and RYR2 (15–18%). These data highlight the importance of recurrent mutations in structural and chromatin-modifying genes in STAD pathogenesis (Figure 2).

Tumor mutational burden (TMB) analysis revealed marked heterogeneity. The mean TMB was 6.39 mutations per megabase (Mb), ranging from 0.02 to 117. The median TMB was 1.94, with approximately 10% of samples (44 cases) classified as high TMB, which may have implications for immunotherapy response. The distribution of TMB indicated that most tumors fell within the low- to moderate-TMB category, with only a small fraction exhibiting hypermutated profiles (Figure 3).

Mutational signature analysis identified six processes contributing to the STAD mutational landscape. Signature 1 resembled SBS1, reflecting age-related mutagenesis due to 5-methylcytosine deamination. Signature 2 resembled SBS40a, of unknown etiology, while Signature 3 resembled SBS21, associated with mismatch repair deficiency. Signatures 4 and 5 corresponded to SBS10b (linked to POLE mutations) and SBS15 (another signature indicative of mismatch

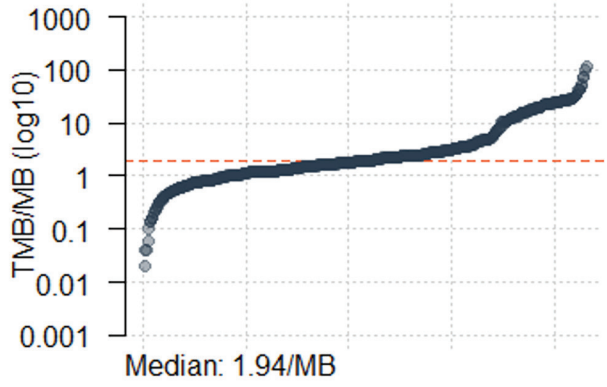


Figure 3. Tumor mutational burden (TMB) across TCGA-STAD samples. Each dot represents a tumor sample, with mutation rates in units of mutations per megabase (Mb) on a log10 axis. The median TMB was 1.94 mutations/Mb (dashed red line), with between-patient heterogeneity, a subgroup with hypermutated profiles.

repair deficiency), respectively. Finally, Signature 6 resembled SBS17b, associated with reactive oxygen species and exposure to 5-FU chemotherapy. Collectively, these findings underscore the roles of aging, DNA repair deficiency, and chemotherapy-induced stress in the mutational biology of GC (Figure 4).

Clinical Features and Expression Data

RNA-seq expression data were available for 448 samples, including 412 primary tumors and 36 solid tissue normal samples.

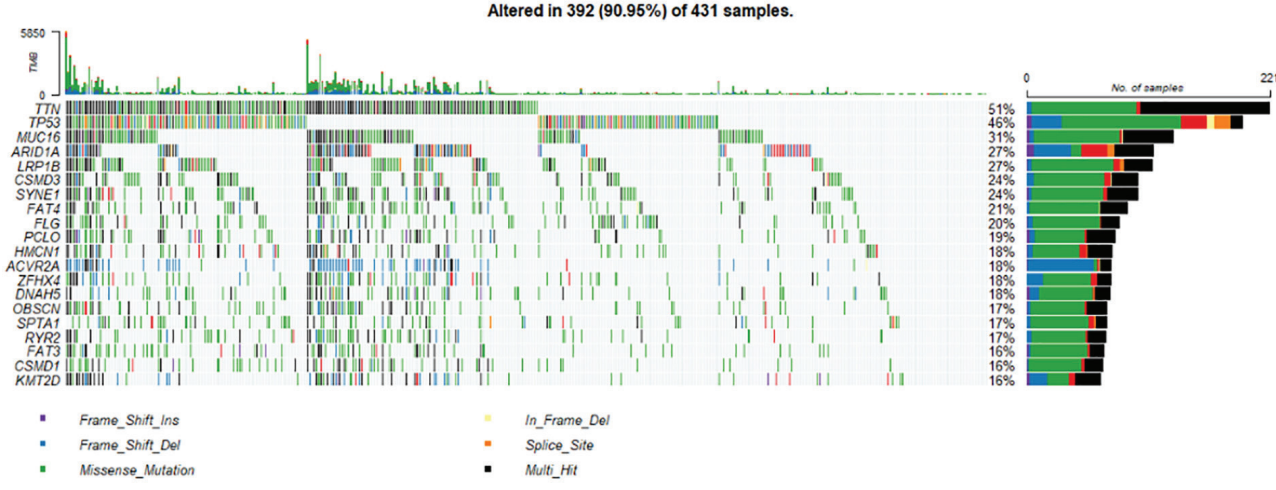


Figure 2. Oncoplot of the mutational profile of TCGA-STAD. There were 392 of 431 samples (90.95%) with at least one non-silent mutation. The most commonly mutated genes were TTN (51%), TP53 (46%), MUC16 (31%), ARID1A (27%), and LRP1B (27%). Mutation types are colored according to their categories (missense, frameshift insertions/deletions, splice-site, in-frame deletions, and multi-hit events), demarcating heterogeneity in mutational profile between patients.

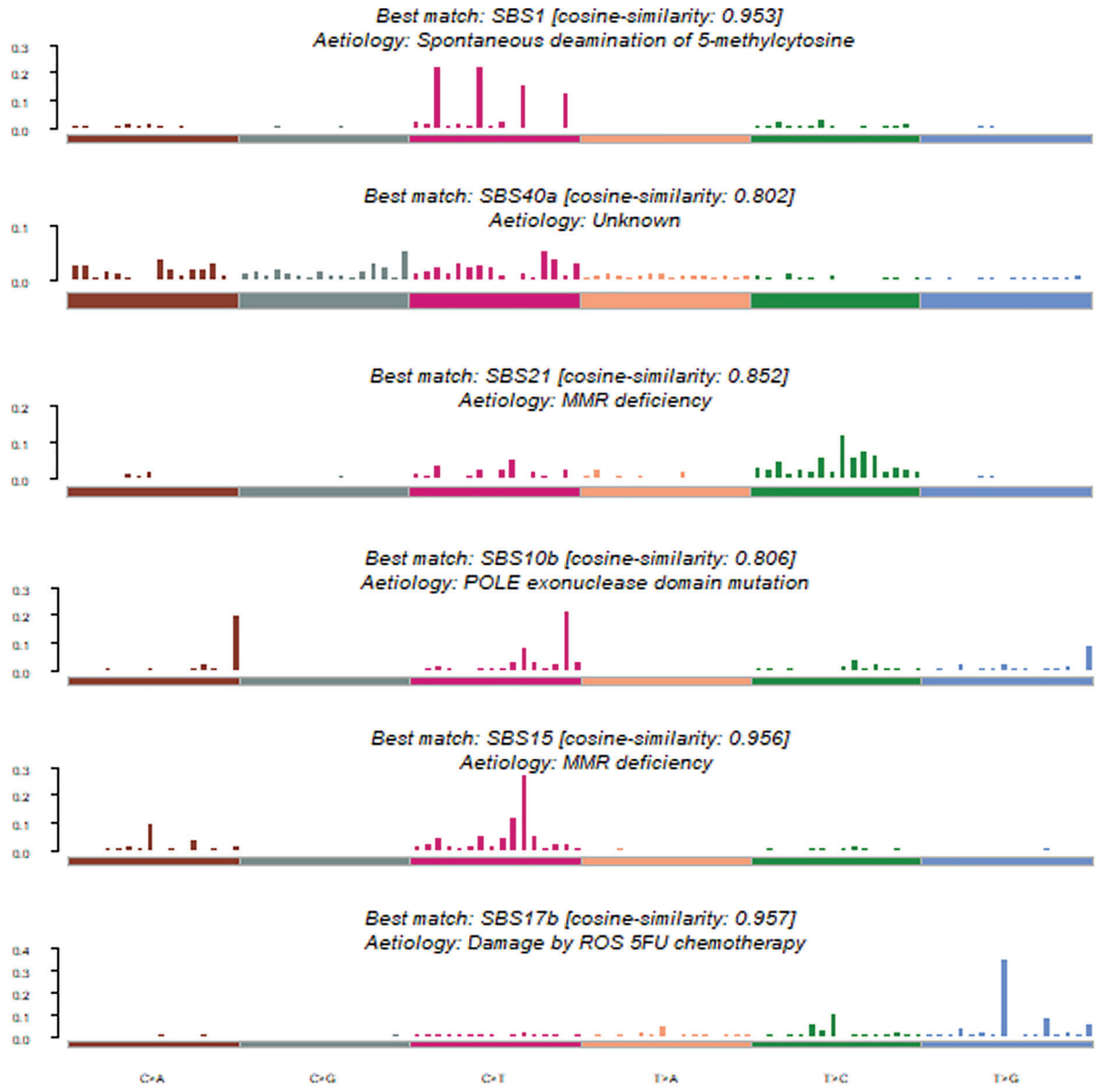


Figure 4. Mutational signature analysis of TCGA-STAD identified six processes, which are largely associated with age, DNA repair deficiency, POLE mutation, and chemotherapy-induced damage.

After quality control and preprocessing, 60,659 genes were retained for analysis. The patient cohort comprised 290 males (64.7%) and 158 females (35.3%), with a median age at diagnosis of 66.3 years. Survival follow-up data revealed that 274 patients (61.2%) were alive at last follow-up, whereas 174 (38.8%) had died. Tumors were frequently diagnosed at advanced stages, particularly IIB, IIIA, IIIB, and IV, and tumor grading was predominantly moderately and poorly differentiated (G2 and G3), with a minority of well-differentiated (G1) tumors. None of the RNA-seq subset harbored mutations in canonical driver genes such as TP53, PIK3CA, KRAS, ARID1A, or RHOA. Given that RNA-seq is less sensitive for variant calling than DNA-based approaches, this absence likely reflects dataset characteristics rather than true mutation

Table 2. Extremely significantly differentially expressed genes chosen between high and low H19 expression groups in TCGA-STAD, with log2 fold change, statistical significance, and expression direction.

Gene	log2FC	p-value	adj p-value	Direction
H19	-4.89	2.37E-216	1.19E-211	Down
IGF2	-4.20	1.97E-109	4.95E-105	Down
ATP4A	5.73	2.02E-67	3.38E-63	Up
KRT4	6.08	9.75E-62	1.22E-57	Up
ATP4B	5.27	2.61E-60	2.62E-56	Up

status, suggesting that expression variation was the dominant feature observed in this cohort.

H19-Associated Transcriptomic Reprogramming

To evaluate expression changes, patients were stratified into high- and low-H19 expression groups. DESeq2 differential analysis identified 15,179 significantly altered genes (adjusted $p < 0.05$), including 7,471 upregulated and 7,708 downregulated genes in the high-expression group compared to the low-expression group. Notably, H19 exhibited overall higher expression in GC relative to normal tissues, consistent with prior reports. However, when tumor samples were stratified into high- and low-expression groups based on the median, DESeq2 modeling revealed a relative decrease in H19 levels within the high-expression subgroup. This does not contradict the tumor-overexpression of H19; rather, it reflects differences that emerge when comparing subgroups within tumors rather than tumors versus normal tissue. In contrast, genes such as ATP4A, ATP4B, and keratins such as KRT4 were markedly upregulated. These observations illustrate the complexity of H19-associated transcriptional reprogramming in GC, suggesting that its regulatory impact may be heterogeneous across patient subgroups rather than uniform.

Correlation analysis revealed strong positive correlations between H19 and genes including IGF2, TCF15, MFAP2, and C11orf95, suggesting potential co-regulatory interactions. Genes such as AKR1B10, CYSTM1, CYP2C18, and CA2 were negatively correlated, indicating opposing regulatory pathways. Spearman correlation coefficients ranging from +0.53 to –0.45 demonstrated robust gene-level correlations with H19 expression (Table 2, Figure 5).

Pathway Enrichment

Gene Ontology (GO) enrichment analysis of differentially expressed genes identified 1,030 significantly enriched biological processes. Key terms included extracellular matrix organization, extracellular structure organization, cell-cell adhesion, and cell-substrate adhesion, indicating that H19-driven gene expression changes profoundly influence the tumor microenvironment and tissue architecture (Figure 6A).

KEGG pathway analysis revealed 133 significantly enriched pathways. Among the most significant were cytoskeletal regulation in muscle cells, ECM (enrichment of extracellular matrix)-receptor interaction, protein digestion and absorption, focal adhesion, and cell adhesion molecules. These findings suggest novel roles in structural remodeling, intercellular communication, and signal transduction in H19-mediated GC biology (Table 3, Figure 6B).

Association of H19 Expression with Clinical Parameters and Survival

Multivariate analyses revealed that H19 expression was not significantly associated with tumor grade or stage. Cox regression analysis indicated that H19 expression did not independently predict overall survival. In contrast, advanced patient age, higher tumor stage (III–IV), and higher tumor grade (G3) were strongly associated with poor prognosis, whereas sex showed no significant effect.

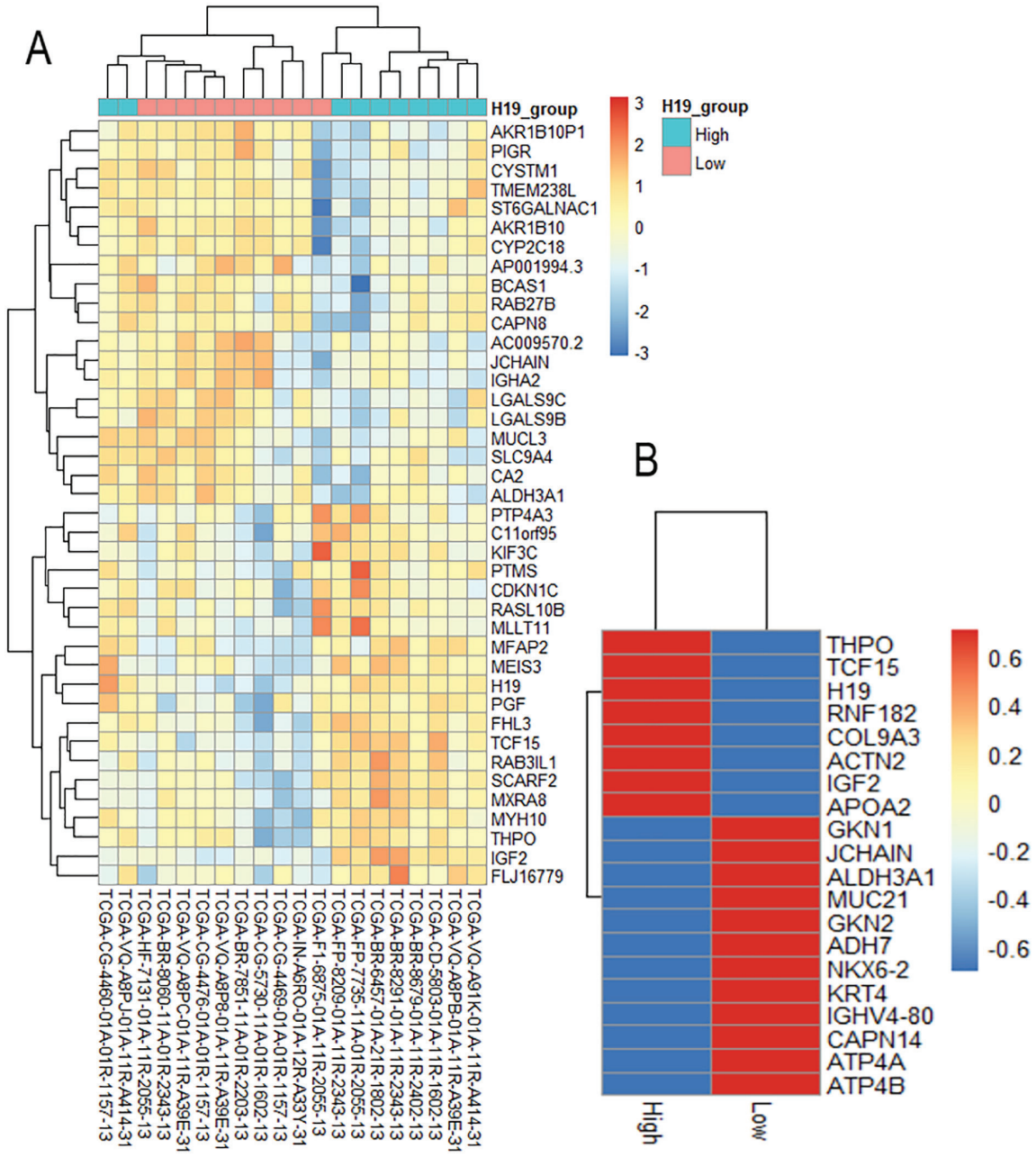


Figure 5. Heatmap visualization of H19-associated expression patterns in TCGA-STAD. (**A**) Genes that are differentially expressed between the high and low H19 expression categories, with evident patient clustering by H19 expression status. (**B**) Highest positive and negative correlated genes with H19, illustrating strong co-expression (e.g., IGF2, TCF15) and inverse correlations (e.g., ATP4A, ATP4B, KRT4). Gradients of color represent scaled expression values (row Z-scores).

The proportional hazards assumption was satisfied, confirming the validity of the Cox regression results (Table 4).

Integration of H19 with miRNA Networks and Clinical Outcome

Given its potential regulatory interactions with microRNAs, we examined correlations with candidate miRNA partners. H19 displayed the strongest positive correlation with hsa-miR-675 ($\rho = 0.64$, $p < 0.001$), followed by hsa-miR-21 and hsa-miR-216a. Conversely, it was negatively correlated with

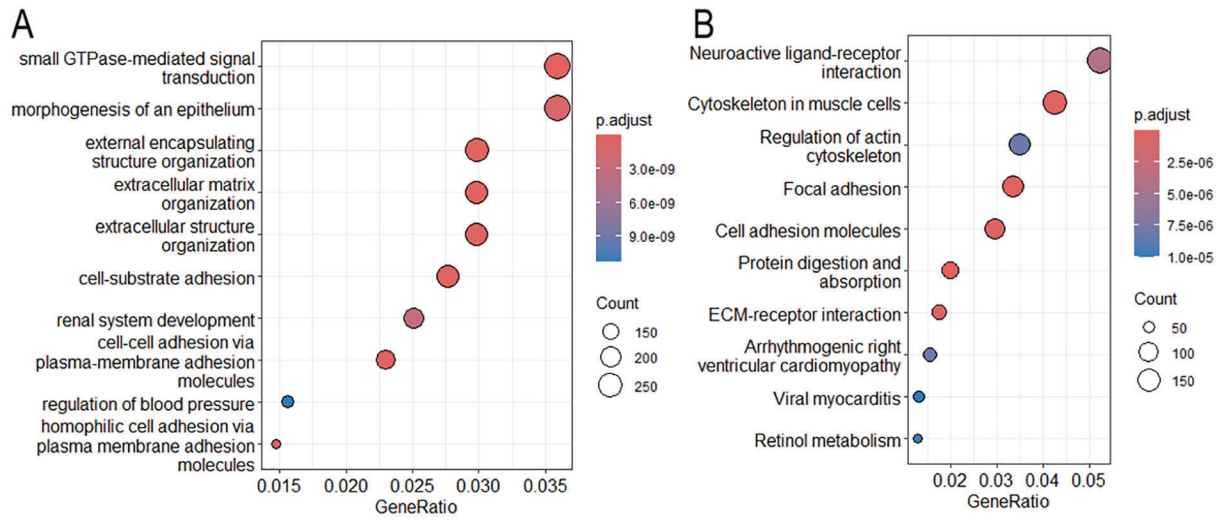


Figure 6. GO (**A**) and KEGG (**B**) enrichment analyses of H19-associated genes indicate pathways for extracellular matrix organization, cell adhesion, and structural remodeling in TCGA-STAD.

Table 3. Top enriched Gene Ontology biological processes and KEGG pathways of H19 expression in TCGA-STAD, listing the number of DEGs and adjusted p-values

Pathway	DEGs	Adjusted p-value
Extracellular matrix organization	242	0
Cell-cell adhesion	187	0
ECM-receptor interaction	66	0
Focal adhesion	126	0
Cell adhesion molecules	111	0

Table 4. Multivariable regression analysis of H19 expression and clinical outcomes.

Outcome	Predictor	OR / HR	95% CI	p-value
Stage (III–IV vs I–II)	H19_z	0.998	0.726–1.265	0.989
Grade (G3 vs G1–G2)	H19_z	0.903	0.728–1.108	0.326
Overall Survival	H19_z	1.037	0.896–1.201	0.623
Overall Survival	Stage III–IV	2.193	1.513–3.180	<0.001
Overall Survival	Grade G3	1.517	1.066–2.158	0.021
Overall Survival	Age	1.000	1.000–1.000	<0.001

hsa-miR-153 and hsa-miR-200a**,** suggesting that H19 may function within a ceRNA network that either promotes or represses the expression of specific miRNAs (Figure 7A).

Candidate miRNA expression across 245 samples was highly heterogeneous. For instance, hsa-miR-675 expression ranged from 0 to 14,326, with a median of 34, whereas hsa-miR-21 expression was substantially higher (median 777,220). These extensive expression ranges underscore the heterogeneity of miRNA regulation among patients.

Survival analysis of 390 patients stratified by H19 expression revealed a trend toward differential outcomes, although it did not reach statistical significance (log-rank p = 0.061). Univariate Cox regression further confirmed that H19 expression was not an independent predictor of overall survival (HR = 1.03, 95% CI: 0.89–1.19, p = 0.713). Thus, although H19 participates in transcriptional and post-transcriptional networks, its prognostic relevance remains uncertain (Figure 7B).

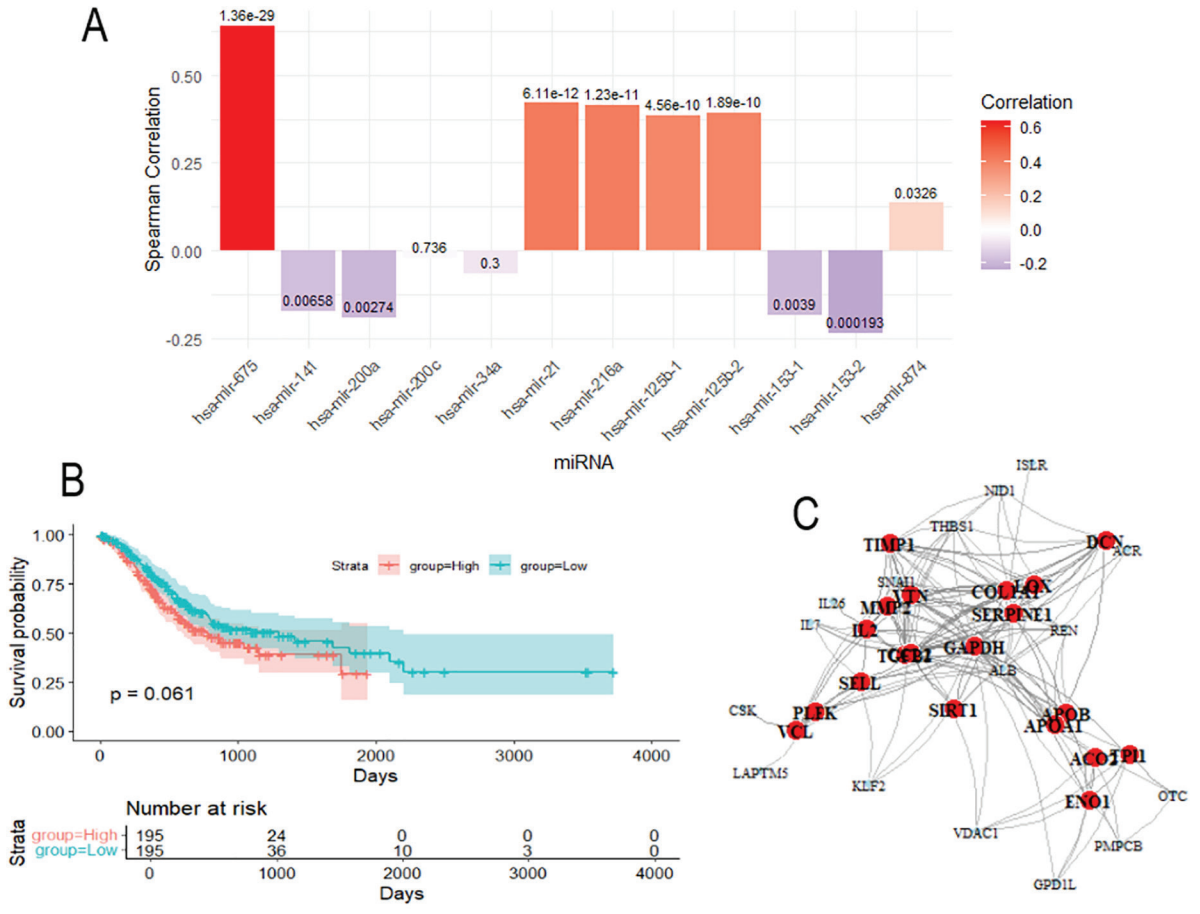


Figure 7. Molecular characteristics of H19 in TCGA-STAD. (A) Correlation analysis with high positive correlation with hsa-miR-675 and correlations with hsa-miR-153 and hsa-miR-200a. (B) Kaplan–Meier survival plots for comparison of high vs. low H19 expression groups, no significant difference observed (p = 0.061). (C) PPI network with hub genes (e.g., GAPDH, COL1A1, SIRT1, TGFB1) as representative of the central regulatory pathways.

Finally, PPI analysis of differentially expressed mRNAs predicted a complex network comprising 2,184 nodes and 9,461 edges**.** The hub genes included GAPDH, COL1A1, CCL2, SIRT1, and TGFB1, which serve as central regulators of metabolism, extracellular matrix signaling, immune modulation, and cell cycle control. These findings suggest that H19-dependent transcriptional programs extend beyond individual targets to encompass global regulatory networks in GC (Figure 7C).

Discussion

This study comprehensively characterized lncRNA H19 in GC using TCGA-STAD data, revealing recurrent mutations in TTN, TP53, MUC16, ARID1A, and LRP1B, as well as six mutational processes associated with aging, mismatch repair deficiency, POLE hypermutation, and chemotherapy-induced lesions. Stratification by H19 expression identified over 15,000 differentially expressed genes enriched in extracellular matrix organization, focal adhesion, and cell adhesion pathways. Co-expression analysis demonstrated positive correlations with IGF2 and TCF15 and inverse correlations with ATP4A, ATP4B, and KRT4, while strong correlation with miR-675 further supported H19 as a ceRNA hub. Although its pronounced influence on transcriptional networks

rendered it an unlikely independent prognostic factor in GC, these findings underscore its context-dependent functional role.

These results both corroborate and extend previous findings. The high mutation frequency of TP53, TTN, and MUC16 aligns with prior genomic studies. The frequent mutations of TTN, MUC16, and LRP1B were confirmed by Zhou et al. (2023) in TCGA-STAD (29), while TP53 mutations are established drivers of GC (30). TTN mutations have been linked to elevated tumor mutational burden (TMB) and enhanced immunotherapy response (Jia et al., 2019) (31), consistent with our observations. MUC16 mutations, observed in nearly one-third of tumors, have been associated with high TMB, improved prognosis, and increased immune infiltration (32). LRP1B mutations, in contrast, exhibit context-dependent prognostic effects: although linked to poor survival in GC with enteroblastic differentiation (33), our larger cohort suggests nuanced, subtype-specific outcomes.

Mutational signature analysis provides mechanistic insight into GC. SBS1, reflecting spontaneous 5-methylcytosine deamination, emerged as a dominant process, consistent with Silveira et al. (2024), who identified it as a principal mutational force in cancers (34). Signatures indicative of mismatch repair deficiency corroborate Meier et al. (2018), who demonstrated a direct link between MMR loss and hypermutation (35). POLE mutations, although rare, align with prior reports associating them with ultrahigh TMB, elevated immune infiltration, and favorable prognosis (36). Collectively, these consistencies validate both our mutational analysis and the utility of mutational signatures as mechanistic biomarkers.

Our transcriptomic findings are concordant with previously reported patterns in GC. ECM remodeling and adhesion pathways reflects earlier bioinformatic studies highlighting COL1A1, COL1A2, and COL3A1 as central regulators of ECM organization (37). Disruption of ECM–receptor interactions has been shown to promote invasion and metastasis (38), while Zhang et al. (2022) demonstrated that CPNE8 facilitates metastasis via focal adhesion signaling (39). Our results support these observations and suggest that H19 may contribute to GC progression by modulating ECM remodeling and adhesion dynamics.

Although H19 is actively involved in transcriptional and post-transcriptional regulatory circuits, its expression alone does not serve as an independent prognostic indicator in GC. Conventional clinical parameters, including age, tumor grade, and stage, remain more predictive of patient outcomes. Despite previous studies, such as Peng et al. (2017), linking high H19 expression to poorer survival across multiple cancer types (40), our analysis did not validate H19 as a standalone prognostic factor in GC. This apparent discrepancy may reflect tumor heterogeneity, subtype-specific effects, cohort composition, or treatment variability, suggesting that H19's prognostic relevance is context-dependent and may be better interpreted in combination with other markers or molecular subgroups.

The co-expression of H19 with IGF2 is consistent with evidence that loss of imprinting at the Igf2/H19 locus drives proliferative signaling (41). The positive correlation with miR-675 aligns with studies in breast and GC demonstrating H19's role as a miR-675 precursor, promoting proliferation and invasion (42). Conversely, the inverse correlation with ATP4A/ATP4B represents a novel finding, consistent with Chen et al. (2022), who identified ATP4A downregulation as a potential GC diagnostic biomarker (43). These data suggest that H19 may contribute to malignant transformation through suppression of parietal cell-associated genes.

The principal strengths of this study include its integrative approach, combining genomic, transcriptomic, and clinical analyses. The large sample size, rigorous statistical methodology,

and network and enrichment analyses enhance reliability and illuminate the global regulatory influence of H19. Importantly, the study situates H19 within the context of mutational signatures and immune modulation, illustrating its multifunctional role in GC biology.

Nonetheless, limitations should be acknowledged. The retrospective nature of TCGA data limits clinical interpretability due to incomplete information on therapy or comorbidities. The findings are largely correlative, without experimental validation to establish causality. Survival analyses may be influenced by unmeasured confounders, and the absence of independent cohort validation restricts generalizability. Functional studies are required to confirm the described regulatory interactions, particularly the antagonism between H19 and ATP4A/ATP4B.

Collectively, these results highlight potential regulatory roles for H19 in GC and generate hypotheses regarding its involvement in invasion, immune modulation, and chemoresistance. However, the retrospective design, lack of detailed clinical data, and correlative approach preclude definitive conclusions.

Future research should focus on experimental validation of H19-associated pathways, particularly its interactions with ATP4A/ATP4B and ECM-related genes. Clinical investigations should assess whether H19 serves as a marker of immunotherapy responsiveness, based on its associations with TMB and ceRNA networks. Therapeutic targeting of H19, including disruption of its ceRNA interactions with miR-675 and miR-21, may offer strategies to overcome chemoresistance (44). Finally, integration of H19 expression into biomarker panels such as MSI, PD-L1, and EBV status (45) could enhance patient stratification and refine precision oncology approaches in GC.

Conclusion

This study provides hypothesis-generating insights into the role of H19 within the genomic and transcriptomic landscape of GC. While it highlights potential functions in microenvironment modulation and transcriptional reprogramming, definitive prognostic or therapeutic implications cannot be established. Prospective studies and functional validation are required to confirm these observations and determine their clinical utility.

Acknowledgments

During the preparation of this work, the authors utilized ChatGPT (OpenAI) to organize and refine the manuscript text. Following this, the authors carefully reviewed and edited the content and assume full responsibility for the publication.

Conflict of Interest

The authors declare no conflict of interest.

Funding

No financial support was provided for this work.

Code of Ethics

N/A-2025-001 — This study used publicly available TCGA data and did not require ethical approval. General code assigned by University of Tabriz.

Ethical Consideration

Ethical approval was not required. All analyses were conducted in accordance with ethical guidelines for the use of publicly available data.

Authors Contribution

Kianoush Mohammadi performed the analyses and drafted the manuscript. Reza Safaralizadeh conceived and designed the study, and Elham Safarzadeh supervised the work and revised the manuscript. All authors read and approved the final manuscript.

References

1 Baccili Cury Megid T, Farooq AR, Wang X, Elimova E. Gastric cancer: molecular mechanisms, novel targets, and immunotherapies: from bench to clinical therapeutics. Cancers (Basel). 2023;15(20):5075.
2 Pužar Dominkuš P, Hudler P. Mutational Signatures in Gastric Cancer and Their Clinical Implications. Cancers (Basel). 2023;15(15):3788.
3 Jing W, Zhu M, Zhang X-w, Pan Z-y, Gao S-s, Zhou H, et al. The significance of long noncoding RNA H19 in predicting progression and metastasis of cancers: a meta-analysis. Biomed Res Int. 2016;2016(1):5902678.
4 Yang J, Qi M, Fei X, Wang X, Wang K. LncRNA H19: A novel oncogene in multiple cancers. Int J Biol Sci. 2021;17(12):3188.
5 Chen J, Wang Y, Zhang X, Lv J, Li Y, Liu X, et al. H19 serves as a diagnostic biomarker and up-regulation of H19 expression contributes to poor prognosis in patients with gastric cancer. Neoplasma. 2016;63(2):223–30.
6 Zhang E-B, Han L, Yin D-D, Kong R, De W, Chen J. c-Myc-induced, long, noncoding H19 affects cell proliferation and predicts a poor prognosis in patients with gastric cancer. Med Oncol. 2014;31(5):914.
7 Li H, Yu B, Li J, Su L, Yan M, Zhu Z, et al. Overexpression of lncRNA H19 enhances carcinogenesis and metastasis of gastric cancer. Oncotarget. 2014;5(8):2318.
8 Li A, Mallik S, Luo H, Jia P, Lee D-F, Zhao Z. H19, a long non-coding RNA, mediates transcription factors and target genes through interference of microRNAs in pan-cancer. Mol Ther Nucleic Acids. 2020;21:180–91.
9 Tang S, Liao K, Shi Y, Tang T, Cui B, Huang Z. Bioinformatics analysis of potential Key lncRNA-miRNA-mRNA molecules as prognostic markers and important ceRNA axes in gastric cancer. Am J Cancer Res. 2022;12(5):2397.
10 Li J, Han T, Wang X, Wang Y, Chen X, Chen W, et al. H19 may regulate the immune cell infiltration in carcinogenesis of gastric cancer through miR-378a-5p/SERPINH1 signaling. World J Surg Oncol. 2022;20(1):295.
11 Yang C, Tang R, Ma X, Wang Y, Luo D, Xu Z, et al. Tag SNPs in long non-coding RNA H19 contribute to susceptibility to gastric cancer in the Chinese Han population. Oncotarget. 2015;6(17):15311.
12 Yang M, Zhang M, Wang Q, Guo X, Geng P, Gu J, et al. Six polymorphisms in the lncRNA H19 gene and the risk of cancer: a systematic review and meta-analysis. BMC Cancer. 2023;23(1):688.
13 Hosseini SA, Haddadi MH, Fathizadeh H, Nemati F, Aznaveh HM, Taraj F, et al. Long non-coding RNAs and gastric cancer: an update of potential biomarkers and therapeutic applications. Biomed Pharmacother. 2023;163:114407.
14 Mohammadi K, Safaralizadeh R. Cigarette Smoke-Induced Transcriptomic Alterations and Angiogenesis in Non-Small Cell Lung Cancer: An Integrative Analysis. J Adv Biomed Sci. 2025;15(3):260-273.
15 Mohammadi K, Safaralizadeh R. Biotechnological Insights into KAT2A-Mediated Epigenetic Regulation in Colorectal Cancer: A Transcriptomic and Functional Genomics Approach. Iran J Biotechnol. 2025;23(3):99–108.
16 Mounir M, Lucchetta M, Silva TC, Olsen C, Bontempi G, Chen X, et al. New functionalities in the TCGAbiolinks package for the study and integration of cancer data from GDC and GTEx. PLoS Comput Biol. 2019;15(3):e1006701.
17 Mayakonda A, Lin D-C, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res. 2018;28(11):1747–56.
18 Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, et al. Software for computing and annotating genomic ranges. PLoS Comput Biol. 2013;9(8):e1003118.
19 Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, et al. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. Bioinformatics. 2005;21(16):3439–40.
20 Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.
21 Xu S, Hu E, Cai Y, Xie Z, Luo X, Zhan L, et al. Using clusterProfiler to characterize multiomics data. Nat Protoc. 2024;19(11):3292–320.
22 Carlson M, Falcon S, Pages H, Li N. org. Hs. eg. db: Genome wide annotation for Human. R package version. 2019;3(2):3.
23 Kolde R. Pheatmap: pretty heatmaps. R package version. 2019;1(2):726.
24 Wickham H. Data analysis. ggplot2: elegant graphics for data analysis: Springer; 2016. p. 189–201.
25 Therneau T. A package for survival analysis in S. R package version. 2015;2(7):2014.
26 Kassambara A, Kosinski M, Biecek P. survminer: Drawing Survival Curves using 'ggplot2'. R package version 0.5.1 [Internet]. 2025. Available from: https://CRAN.R-project.org/package=survminer.
27 Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, et al. The STRING

database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 2023;51(D1):D638–D46.
28 Csardi G, Nepusz T. The igraph software. Complex Syst. 2006;1695:1–9.
29 Zhou Q, Yuan Y, Lu H, Li X, Liu Z, Gan J, et al. Cancer functional states-based molecular subtypes of gastric cancer. J Transl Med. 2023;21(1):80.
30 Ghojazadeh M, Somi MH, Naseri A, Salehi-Pourmehr H, Hassannezhad S, Olia AH, et al. Systematic review and meta-analysis of TP53, HER2/ERBB2, KRAS, APC, and PIK3CA genes expression pattern in gastric cancer. Middle East J Dig Dis. 2022;14(3):335.
31 Jia Q, Wang J, He N, He J, Zhu B. Titin mutation associated with responsiveness to checkpoint blockades in solid tumors. JCI Insight. 2019;4(10):e127901.
32 Li X, Pasche B, Zhang W, Chen K. Association of MUC16 mutation with tumor mutation load and outcomes in patients with gastric cancer. JAMA Oncol. 2018;4(12):1691–8.
33 Zan L, Shen L, Zhang X, Gao N, Tian B, Geng X, et al. Clinicopathological analysis of gastric adenocarcinoma with elevated serum alpha-fetoprotein and enteroblastic differentiation. Zhonghua Zhong Liu Za Zhi. 2024;46(7):686–95.
34 Silveira AB, Houy A, Ganier O, Özemek B, Vanhuele S, Vincent-Salomon A, et al. Base-excision repair pathway shapes 5-methylcytosine deamination signatures in pan-cancer genomes. Nat Commun. 2024;15(1):9864.
35 Meier B, Volkova NV, Hong Y, Schofield P, Campbell PJ, Gerstung M, et al. Mutational signatures of DNA mismatch repair deficiency in C. elegans and human cancers. Genome Res. 2018;28(5):666–75.
36 Zhu M, Cui H, Zhang L, Zhao K, Jia X, Jin H. Assessment of POLE and POLD1 mutations as prognosis and immunotherapy biomarkers for stomach adenocarcinoma. Transl Cancer Res. 2022;11(1):193.
37 Hao S, Lv J, Yang Q, Wang A, Li Z, Guo Y, et al. Identification of key genes and circular RNAs in human gastric cancer. Med Sci Monit. 2019;25:2488.
38 Hu K, Chen F. Identification of significant pathways in gastric cancer based on protein-protein interaction networks and cluster analysis. Genet Mol Biol. 2012;35:701–8.
39 Zhang P, Cao X, Guan M, Li D, Xiang H, Peng Q, et al. CPNE8 promotes gastric cancer metastasis by modulating focal adhesion pathway and tumor microenvironment. Int J Biol Sci. 2022;18(13):4932.
40 Peng L, Yuan X-Q, Liu Z-Y, Li W-L, Zhang C-Y, Zhang Y-Q, et al. High lncRNA H19 expression as prognostic indicator: data mining in female cancers and pooling analysis in non-female cancers. Oncotarget. 2016;8(1):1655.
41 Nordin M, Bergman D, Halje M, Engström W, Ward A. Epigenetic regulation of the Igf2/H19 gene cluster. Cell Prolif. 2014;47(3):189–99.
42 Müller V, Oliveira-Ferrer L, Steinbach B, Pantel K, Schwarzenbach H. Interplay of lncRNA H19/miR-675 and lncRNA NEAT1/miR-204 in breast cancer. Mol Oncol. 2019;13(5):1137–49.
43 Chen Q, Wang Y, Liu Y, Xi B. ESRRG, ATP4A, and ATP4B as diagnostic biomarkers for gastric cancer: a bioinformatic analysis based on machine learning. Front Physiol. 2022;13:905523.
44 Solomon Y, Berhan A, Almaw A, Ersino T, Damtie S, Kiros T, et al. Long non-coding RNA as potential diagnostic markers for acute myeloid leukemia: a systematic review and meta-analysis. Cancer Med. 2024;13(11):e7376.
45 Röcken C. Predictive biomarkers in gastric cancer. J Cancer Res Clin Oncol. 2023;149(1):467–81.