**Abstract:** Gastric cancer (GC) is a multifactorial disease, which can be triggered by numerous factors, including *Helicobacter pylori* infection, genetic susceptibility, and environmental factors. GC develops from the accumulation of multiple genetic and epigenetic alterations in oncogenes and tumor suppressor genes. In this article, we review the literature published in the past year regarding GC genomic and epigenetic characteristics associated with cancer biology. With advances in high-throughput sequencing techniques, many studies have undertaken the molecular profiling of GCs and a large number of non-coding transcripts from regions previously termed “junk DNA”. Long non-coding RNAs and circular RNAs are representative non-coding RNAs that fill a significant gap that was previously unknown or not well understood in the field of gastric carcinogenesis. The study of lncRNA has gone beyond the level of mechanistic studies to clinically relevant studies, indicating its usefulness as a biomarker. Understanding the molecular characteristics of GC is very important for deciding treatment strategies, such as target agents and immunotherapy. Therefore, research analyzing tumor behavior at the molecular level plays a pivotal role in dissecting GC heterogeneity. In this review, we summarize the most recent research updates on the role and function of the gastric microbiota in the process of carcinogenesis. While there is little evidence regarding the drivers or modifiers in the gastric microbiota that mediate the development of GC, many studies have identified gastric microbiota players associated with gastric carcinogenesis and have presented these as candidate targets for therapeutic interventions.

**Keywords:** Gastric cancer, Long non-coding RNA, Circular RNA, Gastric microbiota.

**INTRODUCTION**

Gastric cancer (GC) is one of the deadliest malignancies worldwide. While early GC shows excellent prognosis, advanced GC is associated with poor outcomes. GC development is a consequence of the complex interactions between microbial agents and environmental and host factors. Tumor heterogeneity is a major obstacle in making an accurate diagnosis and deciding the most optimal treatment strategy for GC. Due to the characteristic heterogeneity of GC, patients with the same stage of disease may show quite different clinical outcomes. In recent years, genomic analyses of GC have focused on molecular heterogeneity. Furthermore, the molecular characteristics of GC have been studied to identify tumor biology and predict an accurate prognosis. Accurate identification of novel targets and accurate prediction of prognosis are essential for precision treatment. In addition, a growing number of gastric microbiota studies using gastric tissues and oral cavity samples have provided evidence of dysbiosis in the development of GC.

Alterations in epigenetic regulation are widely involved in the various causes of GC. Among the mechanisms of epigenetic regulation, non-coding RNA (ncRNA) is known to be associated...
with GC in many ways. In the early and the mid-2000s, research on microRNAs exploded, and since then, studies on long non-coding RNAs (lncRNAs), ncRNAs 200 nt or more in size, have been extensively conducted. Recent studies have identified thousands of lncRNAs in mammalian genomes that regulate gene expression in different biological processes, and which play a critical role in the genetic, epigenetic, and post-transcriptional regulation of tumorigenesis. LncRNAs tend to broadly regulate DNA methylation, histone modification, and chromatin remodeling, which are major mechanisms of epigenetic regulation.

CANCER GENOMICS AND BIOLOGY

Recently, the field of genome-wide molecular technology has rapidly developed, and the molecular characteristics of GC have been elucidated. Therefore, efforts have been made to classify GC into molecular subtypes for differentiated prognoses and precision treatments. In 2014, The Cancer Genome Atlas research network group divided GC into four molecular subtypes: Epstein-Barr virus (EBV), microsatellite instability (MSI), chromosomal instability (CIN), and genomically stable (GS). However, this is insufficient to explain the difference in prognosis among the four groups. As mentioned before, GC is a highly heterogeneous tumor. Zhang et al performed transcriptome-wide single-cell RNA sequencing (sc-RNA-seq) on 27,677 cells from nine tumor and three non-tumor samples to decipher gastric tumor heterogeneity. They classified five subgroups exhibiting distinct expression. Using a differentiation-related gene panel, the three subgroups showed various differentiated grades according to Lauren’s classification. The other subgroup displayed unique transcriptome features expressing fundic gland type GC, which is a rare tumor type. The last subgroup exhibited immune-related signature genes associated with EBV infection. Sundar et al studied intratumoral spatial heterogeneity in GC by annotating the superficial primary tumors, deep primary tumors, and tumors in matched lymph node metastases. In the Nanostring profiling (PanCancer Progression Panel, 770 genes) of 64 GCs, 43% of the genes showed differences between PT_sup and PT_deep and 38% between PT_sup and LN_met. In contrast, 16% of the genes were differentially expressed between PT_deep and LN_met. Next-generation sequencing data exhibited similar results to the NanoString data. Therefore, by performing genomic and transcriptomic profiling, superficial tumor profiles were shown to be significantly different from those of deep tumors and lymph node metastases. Studies compared the molecular characteristics of the adenocarcinoma of the gastroesophageal junction (AGEJ) with those of the esophageal adenocarcinoma (EAC) and gastric adenocarcinoma (GC). They classified AGEJ as EAC-like (31.2%) and GC-like (68.8%) based on the 400-gene classifier. The GC-like group exhibited significantly increased phosphoinositide 3-kinase-AKT signaling with ERBB2 inactivation. The EAC-like group displayed significantly different alternative splicing in the skipped exon of RPS24, a higher copy number amplification including ERBB2 amplification, and the activated protein expression of ERBB2 and EGFR. Therefore, AGEJ also presented heterogeneous entities of the EAC-like and GC-like groups with different molecular characteristics.

Several studies have used genome-wide expression profiling to predict the patients who are at a high risk for GC recurrence and metastasis. Lee et al analyzed genome-wide expression profiling and reported the use of a 12-gene panel for the prediction of peritoneal recurrence in patients with GC from two publicly available datasets (GSE15081 and GSE62254). The panel of 12 genes included: ZBTB1, CHCHD3, KLHL41, POPDC2, LTBP3, CAVIN2, STT3B, TXNDC16, PHYHD1, KCNJ6, PHYHD1, KCNJ6, SLITRK6, and LMBR1. They developed a logistic regression model to predict peritoneal recurrence in patients with GC, which showed an area under the curve (AUC) of 0.95 (95% CI: 0.89-0.98, p < 0.001). The same group also reported a 7-gene panel for predicting recurrence in patients with diffuse GC by analyzing genome-wide transcriptomic profiling data from three publicly available datasets (GSE62254, GSE13861, and TCGA-STAD). The 7-gene panel included HLF, CAV2, HACD1, MLF1, GC, VSNL1, and SERPINB5. The logistic regression model to predict tumor recurrence in patients with diffuse GC yielded an AUC value of 0.91 (95% CI 0.83-0.96, p < 0.001). These novel transcriptomic signatures for the prediction of high-risk patients with a chance of recurrence could be used in clinical decision making.
Lin et al. investigated the genomic and transcriptomic changes in AGEJ in Chinese patients. They showed that the major genomic changes in Chinese patients with AGEJ are focal copy number variations (CNVs) and COSMIC Signature 17-featured single nucleotide variations. In addition, tumor mutation burden and gene level CNVs were significantly correlated with the survival of these Chinese patients. Hao et al. profiled 40 AGEJ patients and classified patients into two groups: 20 short-span survivors (< 13 months) and 20 long-span survivors (>36 months). They performed whole-exome sequencing and RNA sequencing to identify the molecular determinants of prognosis. KMT2C alterations were enriched in the short-span survivors with high levels of intratumor heterogeneity, whereas APOBEC mutational signatures were observed in the long-span survivors. In addition, the short-span survivors showed loss of heterozygosity of chromosome 4 and decreased levels of B, CD8, and natural killer cells, and interferon-gamma responses, which together exhibited a “cold” tumor immune microenvironment. Therefore, the short-span survivors had decreased antitumor immunity and showed the worst prognosis.

Investigating the basic mechanisms of GC tumorigenesis may help identify new therapeutic interventions. The telomerase catalytic subunit (TERT) is transcriptionally reactivated in 90% of cancers. Xing et al. identified that the transcription factor early B cell factor 1 (EBF1) as a TERT transcriptional repressor and the abolishment of EBF1 function caused TERT upregulation. Inactivation of EBF1 is silenced by DNA methyltransferase, polycomb-repressive complex 2 (PRC2), histone deacetylase activity, EBF1 DNA-binding domain, and more rarely, genomic deletions and rearrangements proximal to the TERT promoter. Chi et al. elucidated that the guanine nucleotide exchange factor for Rac1 and CDC42 (DOCK6) is a biomarker of GC prognosis. Expression of DOCK6 promotes GC cancer stem cell (CSC) properties and affects chemo- or radioresistance through Rac 1 activation. The yes-associated protein 1 (YAP1) oncogene is implicated in many human malignancies. Ajani et al. reported the expression of YAP1 and peritoneal carcinomatosis (PC) in GC cell heterogeneity using RNA-Seq and sc-RNA-Seq. YAP1 was upregulated in PC cells and conferred CSC characteristics. Sc-RNA-Seq results also showed that PC cells are highly heterogenous. Tseung et al. reported that PHD finger protein 8 (PHF8, KDM7B) was significantly associated with poor clinical outcomes in HER2-negative GC. In addition, PHF8 interacts with c-Jun on the PRKCA promoter. The depletion of PHF8 or PKCα upregulated PTEN expression and was rescued by the expression of PKC or active Src. Therefore, the PKC-Src-PTEN pathway regulated by PHF8/c-Jun is a potential therapeutic target for HER2-negative GC.

Although immunotherapy was not the first treatment strategy for GC, efforts to identify a favorable response to immunotherapy in GC patients have been continuously performed. Zhang et al. analyzed 468 tissue microarray specimens, 52 fresh GC tissues, and TCGA data of 298 GC patients. GC patients with high Interleukin-10+ tumor-associated macrophage infiltration showed poor prognosis. It exhibited an immune-evasive tumor microenvironment by regulatory T cell infiltration and CD8+ T cell dysfunction. Immune check point inhibitors (ICIs) have prominent efficacy in MSI or EBV molecular subtypes of GC, but show a markedly lower efficacy in GS and CIN GC subtypes. Derks et al. reported that the GS group enriched CD 4+ T cells, macrophages, and B cells, and 50% of the cancers of this group exhibited a tertiary lymphoid structure. In addition, between the CIN groups classified as “hot” CIN GC and “cold” GC, the “cold” CIN GC showed an expression of MYC gene and cell cycle pathways and the amplification of CCNE1 gene. Therefore, subgroups of GS and CIN GC could be candidates for immunotherapy.

MICROBIOTA

To date, there is no convincing evidence on whether the gastric microbiota plays an essential role in GC pathogenesis. However, the gastric microbiota is known to directly interact with gastric tissues and to affect gastric carcinogenesis. Therefore, many studies have explored the role of the gastric microbiota in gastric carcinogenesis. Zhang et al. investigated the gastric microbiota by 16s rRNA gene analysis in gastric mucosal samples from 47 patients, including those with superficial gastritis (SG), atrophic gastritis (AG), gastric intraepithelial neoplasia (GIN), and GC. They found no difference in the richness or diversity of the gastric microbiota.
across the various stages of gastric carcinogenesis. The genera *Slackia*, *Selenomonas*, *Bergeyella*, and *Capnocytophaga* were continuously enriched from SG to GC. In addition, *Parvimonas*, *Eikenella*, *Prevotella-2*, *Kroppenstedtia*, *Lentibacillus*, and *Oceanobacillus* were the most representative of GC patients\(^{19}\). Other studies also investigated 30 healthy controls, and 21 chronic gastritis (CG), 27 intestinal metaplasia (IM), 25 GIN, and 29 GC patients via 16S rRNA gene profiling. The bacterial diversity and abundance of the phyla *Armatimonadetes*, *Chloroflexi*, *Elusimicrobia*, *Nitrospirae*, *Planctomycetes*, *Verrucomicrobia*, and WS3 decreased from CG, IM, and GIN to GC. *Actinobacteria*, *Bacteroides*, *Firmicutes*, *Fusobacteria*, *SR1*, and *TM7* were enriched in the GIN and GC \(^{20}\). Kadeerhan et al\(^{21}\) reported a longitudinal study to investigate the alterations in the gastric microbiota with the development of gastric carcinogenesis. They assessed dynamic microbial changes in GC development by deep sequencing the 16R rRNA gene in a 4-year endoscopic follow-up cohort in China. They identified *Helicobacter*, *Bacillus*, *Capnocytophaga*, and *Prevotella* as associated with lesion progression-to-dysplasia (DYS)/GC. The panel including the four genera predicted patients’ progression to DYS/GC quite well. Therefore, gastric microbial dysbiosis is a potential predictive marker of lesion progression\(^{21}\).

In addition, bacteria in the oral cavity may contribute to gastric carcinogenesis and could therefore serve as new diagnostic markers for GC. Huang et al\(^{22}\) investigated the salivary microbiota in patients at different histological stages of gastric carcinogenesis by 16S rRNA gene sequencing. A total of 293 patients were divided between the following groups: 101 SG, 93 AG, and 99 GC patients. *Corynebacterium* and *Streptococcus* were enriched in GC compared to SG and AG. The abundance of *Haemophilus*, *Neisseria*, *Parvimonas*, *Peptostreptococcus*, *Porphyromonas*, *Prevotella*, *Haemophilus*, and *Neisseria* was decreased in GC patients. The proportion of unclassified *Streptophyta* and *Streptococcus* was higher in GC\(^{22}\). Wu et al\(^{23}\) also reported the impact of oral bacteria on GC development in 27 SG and 16 GC patients by 16S rDNA deep sequencing. They assessed the bacteria present in the paired gastric mucosa and tongue coating samples. The co-occurrence of bacteria between the tongue coating and gastric mucosa significantly differed between SG and GC patients. In addition, the core shared oral bacteria in the gastric mucosa were associated with *Helicobacter pylori* infection status. Therefore, oral microbes may be a primary driver of *H. pylori*-induced gastric microbial dysbiosis in patients with GC\(^{23}\).

Guo et al\(^{24}\) reported that the successful eradication of *H. pylori* could restore the gastric microbiota to a status similar to that of *H. pylori* non-infected subjects. They investigated alterations in paired gastric biopsies and stool samples from 59 subjects with successful *H. pylori* eradication and 57 subjects with failed *H. pylori* eradication, relative to 49 *H. pylori*-negative subjects, using deep sequencing of the microbial 16S rRNA gene. Successful *H. pylori* eradication increased the microbial richness and reversed the microbial dysbiosis. In addition, *Bifidobacterium* levels in fecal microbiota increased after a successful *H. pylori* eradication\(^{24}\).

There is currently no solid evidence that gastric microbiota other than *H. pylori* are directly involved in gastric carcinogenesis. Gantuya et al\(^{25}\) evaluated the gastric microbiota of 48 GC and 120 non-cancer patients [20 normal gastric mucosa (control), 20 CG, 40 AG, and 40 IM patients] by 16S rRNA gene amplicon sequencing. Alpha diversity was the highest in the control group, followed by the IM and cancer groups. The CG and AG groups exhibited the least diversity. *Lactobacilli* and *Enterococci* were enriched in patients with GC without *H. pylori* infection, while *Carnobacterium*, *Glutamicibacter*, *Paeniglutamicibacter*, *Fusobacterium*, and *Parvimonas* were associated with GC, regardless of *H. pylori* infection. They elucidated the role of non-*H. pylori* gastric microbiota in Mongolian GC carcinogenesis\(^{25}\).

**EPGENETICS INCLUDING THE ROLE OF LncRNA**

The levels of the lncRNA HOTAIR are initially increased in the tissues of metastasized breast cancer patients, and this lncRNA is known to epigenetically interfere with the function of genes that inhibit metastasis by affecting the PRC2 complex\(^{26}\). In addition, HOTAIR is known to be involved in GC and progression, and alternatively, unknown mechanisms have also been reported in breast cancer. In our study, HOTAIR was found to promote the progression of GC\(^{27}\) and the methylation of the PCDH10 gene, a major tumor suppressor gene in GC, and was also found to sponge miR-148b\(^{28}\).
GC development or progression are known to involve multiple mechanisms rather than a single mechanism. It has been reported that IncRNA expression is extensively altered in the pathogenic course of GC. Furthermore, several lines of scientific evidence support that the degree of the abnormal expression of IncRNAs in GC tissues is associated with cancer invasion, lymph node metastasis, recurrence, and survival\(^1\). To date, many studies on IncRNAs have mainly focused on the mechanism of GC; however, more recently, research on the use of IncRNAs as biomarkers has also started to gain pace. The application of non-coding RNA detection in circulating blood for GC screening has also attracted much attention. Compared to proteins, IncRNAs are generally tissue-specific, stable in serum, and not easily degraded by RNases. These characteristics provide a theoretical basis for considering IncRNAs as biomarkers of cancer. In this respect, remarkable research results were published in 2020. The GC-associated long noncoding RNA1 (IncRNA-GC1) was reported to promote gastric carcinogenesis and may act as a modular scaffold of WDR5 and KAT2A complexes to specify the histone modification pattern\(^{29}\). Guo et al\(^{30}\) demonstrated that exosomal IncRNA-GC1 could be used as a noninvasive biomarker to monitor early-stage GC or disease progression, using a relatively large, well-designed clinical sample, including 164 test-phase patients and 622 validation patients\(^{30}\). The importance of this study is that it demonstrated exosomal IncRNA-GC1 to exhibit better diagnostic performance compared to CEA, CA72-4, and CA19-9, which are currently used as biomarkers for GC prognosis. LncRNA-GC1 levels achieved better diagnostic efficiency in distinguishing between patients with early GC, chronic atrophic gastritis, or intestinal metaplasia in this study. Indeed, this should be confirmed through a different set of large-scale clinical studies and comprehensive meta-analysis; however, it is still a valuable study showing the clinical utility of IncRNAs in GC. For a biomarker to become a clinically meaningful biomarker of GC, it must show adequate accuracy for advanced GC as well as early GC. Previously, there have been several reports of IncRNA biomarkers in plasma/serum or exosomes for early GC, but Zhou et al\(^{31}\) reported that C5orf66-AS1 levels were reduced in GC tissues, serum, and cell lines. In this study, the expression of C5orf66-AS1 was significantly reduced in the serum of patients with gastric dysplasia and early GC. The area under the curve (AUC) for distinguishing GC from non-GC was 0.688, with a sensitivity and specificity of 77.5% and 53.6%, respectively. For the diagnosis of early GC, the AUC reached 0.749 with a sensitivity and specificity of 94.9% and 48.2%, respectively\(^{31}\). Like previous studies, this study showed insufficient specificity, but this topic may be worth taking forward in future studies.

Circular RNAs (circRNAs) are regulatory IncRNAs that form a closed continuous loop. CircRNAs in blood or body fluids can be used as biomarkers in various diseases. CircRNA is derived from exonic or intronic sequences and is generated by the back-splicing of the mRNA precursor. Compared to canonical linear RNAs, circRNAs are resistant to exonuclease digestion because they can form a closed loop. A meta-analysis of 15 studies examining whether circRNAs are useful as markers for determining the clinicopathological progression or prognosis of GC showed that circRNAs can be used as biomarkers for this purpose\(^{32}\). CircRNAs can also act as sponges of microRNAs, transcriptional factors, and RNA-binding proteins, which means they have a very similar regulatory mechanism to IncRNAs, and as is known in the case of IncRNAs, small peptides derived from the ncRNA sequence may exhibit regulatory action in humans. Recently, the circRNA circMAPK1 (hsa_circ_0004872) was reported to be involved in the MAPK pathway involved in GC. In this study, the expression of circMAPK1 was reduced in GC tissues compared to the surrounding normal tissues, and it was reported that circMAPK1 encodes a protein composed of 109 amino acids and that this protein acts as a tumor-suppressor\(^{33}\).

CONCLUSIONS

Despite remarkable advances in understanding the mechanisms involved in GC development and progression, the clinical application of IncRNAs in the diagnosis and treatment of GC has started to gain attention only recently. Given the large number of new findings in GC research, it is important to identify the most clinically relevant targets among various genetic and epigenetic pathways. More frequent information exchange between basic researchers and clinical scientists could be the cornerstone of this development.
Conflict of interest
The authors declare that they have no conflict of interest.

REFERENCES


