Jentashapir J Cell Mol Biol. 2024 September; 15(3): e150493. https://doi.org/10.5812/jjcmb-150493.

Published online: 2024 September 23. **Research Article**

Exploring the Role of *Escherichia coli* K-12 in Reducing MIR17HG Expression as a Poor Prognostic Biomarker in Colorectal Cancer

Zeinab Rohani [1], Hossein Sazegar [1,*], Ebrahim Rahimi [2]

[1] Department of Biology, Faculty of Basic Sciences, Shahrekord Branch, Islamic Azad University, Shahrekord, Iran
[2] Department of Food Hygiene, Faculty of Veterinary Medicine, Shahrekord Branch, Islamic Azad University, Shahrekord, Iran

[*] *Corresponding author:* Department of Biology, Faculty of Basic Sciences, Shahrekord Branch, Islamic Azad University, Shahrekord, Iran. Email: h.sazegar@iaushk.ac.ir

Received 2024 July 3; **Revised** 2024 July 30; **Accepted** 2024 July 31.

Abstract

Background: Long noncoding RNAs (lncRNAs) play significant roles in various cellular processes, and alterations in their expression levels can contribute to the pathogenesis of colorectal cancer (CRC).
Objectives: This study aims to identify lncRNAs highly associated with poor prognosis in CRC and determine those that exhibit significant expression changes under the influence of *Escherichia coli* K-12.
Methods: Potentially susceptible lncRNAs to expression modulation in the presence of *E. coli* K-12 were identified by analyzing GSE50040 datasets. Data from the cancer genome Atlas (TCGA) were utilized to assess *E. coli* K-12-affected lncRNAs with the most significant impact on CRC pathogenesis. The association between the candidate lncRNA and patient prognosis was investigated using clinical data. The co-expression network was employed to identify pathways related to the identified lncRNA via the MsigDB database. To validate the in silico findings, CRC and adjacent normal samples were examined using the RT-qPCR method.
Results: Cox regression analysis demonstrated that *MIR17HG* is a strong biomarker associated with poor prognosis in CRC patients. Increased expression of *MIR17HG* in cancer samples was correlated with key pathways involving cell proliferation, anti-apoptosis, and metastasis. RT-qPCR results showed that the expression level of MIR17HG in CRC samples was significantly higher than in normal samples. Further analysis revealed that MIR17HG expression is susceptible to suppression by *E. coli* K-12.
Conclusions: High expression of *MIR17HG* in cancer samples is associated with an increased probability of mortality in CRC patients. Our study highlights the potential of *E. coli* K-12 to reduce CRC malignancy by downregulating *MIR17HG* expression.

Keywords: LncRNA, Prognosis, Biomarker, *Escherichia coli*,Colorectal Cancer

1. Background

Colorectal cancer (CRC) is one of the leading causes of cancer-related deaths worldwide, accounting for approximately 10% of all cancer cases (1). This disease poses a significant threat to millions of people globally, with metastatic disease occurring in around 20% of patients (2). At the molecular level, colorectal carcinogenesis involves the activation of oncogenes and the inactivation of tumor suppressor genes (3). Genes with highly differential expression patterns in CRC can serve as molecular markers, complementing histopathological factors in diagnosis and guiding therapeutic strategies for personalized patient care (4).

Noncoding RNAs (ncRNAs) are functional RNA molecules that are not translated into proteins, comprising 98% of the human genome (5). Long noncoding RNAs (lncRNAs) are a subclass of ncRNAs with more than 200 nucleotides and are evolutionarily conserved (6). Advances in understanding the molecular mechanisms of lncRNAs have shown that these molecules can function as either tumor suppressors or promoters in carcinogenesis (7). Abnormal lncRNA expression has been reported in various solid tumors (8, 9). As a result, lncRNAs are emerging as attractive targets for therapeutic interventions against cancer and as potential diagnostic and prognostic biomarkers in tumors (10-12).

Bacteria have the ability to target both primary tumors and metastases, making them useful for tumor-selective drug delivery (13-15). *Escherichia coli* K-12 is a well-known facultative anaerobic bacterium and human

intestinal commensal (16). Due to its increased abundance in necrotic areas rather than the peripheral proliferative regions of tumors, the hypoxic areas of solid tumors could be targeted for treatments using this bacterium (17). Although *E. coli* K-12 has been reported to inhibit CRC progression through altered expression of protein-coding transcripts (18), no studies have yet examined its influence on ncRNAs. Given the role of oncogenic lncRNAs in cancer development, we aimed to investigate the impact of *E. coli* K-12 on downregulating the lncRNA *MIR17HG*.

In the present study, we evaluate the potential of *E. coli* K-12 to reduce *MIR17HG* expression, an oncogenic lncRNA with high expression in CRC patients. Furthermore, the Cox regression test was used to assess the effect of *MIR17HG* on CRC patient survival.

2. Objectives

This study aims to investigate the impact of *E. coli* K-12 on reducing CRC malignancy by altering *MIR17HG* expression. To achieve this, data from the cancer genome Atlas (TCGA) and the GEO database were used to identify the relationship between *MIR17HG* and *E. coli* K-12 in CRC. Additionally, this study assesses the effect of *MIR17HG* on CRC patient mortality through Cox regression analysis.

3. Methods

3.1. Data Sources, Normalization, and Differential Expression

To identify noncoding RNAs (ncRNAs) associated with the development and malignancy of CRC, transcriptomic data (RNAseq) was obtained from the TCGA database. For this purpose, the raw format (HTseq-Counts) of CRC RNAseq data was downloaded using the TCGAbiolinks package (19). The CPM (count per million) criterion of less than 10 was used to filter out genes with zero or near-zero expression using the edgeR package (20). The data were then normalized by the TMM (trimmed mean of M values) method and transformed into logarithmic form (base 2) using the limma package and the voom method (21). Subsequent analyses were conducted using the normalized expression matrix.

The cancer genome Atlas data were divided into cancer and normal groups, and the differential expression between these two groups was calculated using the linear model method. To investigate the relationship between the candidate lncRNA and *E. coli* K-12 in CRC, the GSE50040 dataset was analyzed from the GEO database. The raw expression profile data related to the study were downloaded and processed using the Affy package. Afterward, data normalization and logarithmic transformation (base 2) were carried out using the limma package and the RMA method (22). Samples from GSE50040 were divided into two groups, and the linear model method was applied to calculate expression differences between the treatment group with *E. coli* K-12 and the control group.

3.2. Clinical Data Processing and Survival Testing

The latest update of CRC clinical data was downloaded from the TCGA database to examine the role of *MIR17HG* in predicting patient survival. The normalized expression matrix was merged with the clinical data to assess the association between *MIR17HG* expression and patient prognosis using the Z-score scale. A Cox regression test was then performed to calculate the relationship between the expression level of *MIR17HG* and the patients' mortality rate. Finally, the Kaplan-Meier (K-M) curve was used to validate the results, with the median expression of *MIR17HG* set as the cut-off to divide the cancer samples into high and low expression groups.

3.3. Co-expression Network and Pathways

The RNAseq data from the TCGA database was utilized to investigate gene expression changes in CRC. Pathways related to *MIR17HG* were identified through the co-expression network and the normalized expression matrix. For this purpose, a Pearson correlation test was conducted between the expression level of *MIR17HG* and all genes in the expression matrix. Genes with $R > 0.5$ and $P < 0.01$ were selected as co-expressed genes. Pathways associated with these co-expressed genes were then identified using the MsigDB database through the Enrichr tool (https://maayanlab.cloud/Enrichr/).

3.4. Tissues Samples, RNA Isolation, cDNA Synthesis, and RT-qPCR

Forty CRC samples and an equal number of adjacent normal tissue samples were collected from the repository of the Iranian Tumor Bank. The review board of Imam Khomeini Hospital, where the samples originated, approved all bioethical aspects based on the regulations of the Iranian Ministry of Health, Treatment, and Medical Education. Ethical approval was granted under access number IR.IAU.PS.REC.1401.319. RNA extraction from the samples was performed using the TRIzol method (Sigma, USA) according to the manufacturer's instructions. cDNA synthesis was then conducted using the TaKaRa kit, following the company's guidelines. To examine the expression level

Table 1. The Specific Primer Sequences Used in This Study

Target Gene	Primer Name	Sequence
MIR17HG	Forward	5'TCAGGAGTTCGAGACCAACC3'
MIR17HG	Reverse	5'TGCCTCAGCCTCCAGAGTAG3'
GAPDH	Forward	5'ACAGTCAGCCGCATCTTCT3'
GAPDH	Reverse	5'CCCAATACGACCAAATCC3'

of the candidate lncRNA in CRC and adjacent normal tissues, specific primers for *MIR17HG* as the target gene and *GAPDH* as the internal control were designed using the Primer-Blast tool (NCBI). The primer sequences are summarized in Table 1. The expression level of *MIR17HG* was quantitatively assessed using the RT-qPCR method in both cancerous and normal tissue samples. The expression of *MIR17HG* was calculated using the $2^{-\Delta Ct}$ method.

3.5. Statistics and Software

All pre-processing, differential expression, co-expression network, and survival data analyses were conducted using the R programming language (v4.0.2), while GraphPad software (v8) was employed for generating and displaying graphs. The linear model method was applied to calculate the differences in gene expression between groups. The significance of the association between *MIR17HG* expression and patient prognosis was assessed using the log-rank test. Cytoscape (v4) was used to visualize the co-expression network of genes correlated with *MIR17HG*.

4. Results

4.1. Expression Changes of MIR17HG by Escherichia coli K-12 in CRC

The lncRNA with significant expression changes under the influence of *E. coli* K-12 in CRC was identified based on the GSE50040 data analysis. Additionally, TCGA data was used to assess alterations in candidate lncRNA expression in CRC and normal samples. The analysis of expression differences revealed that the expression of *MIR17HG* was significantly increased in CRC samples compared to normal samples, with logFC > 1 and FDR < 0.01 (Figure 1A). In contrast, the expression of this lncRNA was downregulated in the presence of *E. coli* K-12, with logFC < -1 and FDR < 0.01 (Figure 1B). Furthermore, RT-qPCR results confirmed the in-silico findings, showing a significant increase in *MIR17HG* expression in CRC samples (Figure 1C). Enrichr analysis indicates that the dysregulation of mTOR, NF-kB, and BCR signaling pathways is highly associated with *MIR17HG*. These findings suggest that *E. coli* K-12 may impair tumor survival by reducing the overexpression of *MIR17HG*.

4.2. Relationship of MIR17HG and the Mortality Rate

The impact of *E. coli* K-12 on the candidate lncRNA and its association with malignancy and patient mortality was assessed by analyzing the relationship between *MIR17HG* expression levels and patient prognosis using TCGA clinical data. The Cox regression analysis revealed that *MIR17HG* (HR = 1.33, log-rank < 0.0001) is significantly associated with poor survival in CRC patients. Additionally, Kaplan-Meier analysis showed that increased *MIR17HG* expression is correlated with a higher mortality rate among patients (Figure 1D). These findings suggest that *MIR17HG* expression can serve as a prognostic biomarker in CRC.

4.3. Correlating of MIR17HG with Genes Related to the Main Pathways of Cancer

A co-expression network was used to identify pathways related to *MIR17HG*. The analysis revealed that *MIR17HG* expression levels are significantly correlated with 264 other genes, with a correlation coefficient greater than 0.5 and P < 0.01 (Figure 2A). These genes are involved in pathways such as TNF-alpha signaling via NF-kB, KRAS signaling, G2-M checkpoint, mitotic spindle, and E2F targets (Figure 2B, FDR < 0.01). These findings indicate that *MIR17HG* may contribute to the pathogenesis of CRC through these pathways.

5. Discussion

Escherichia coli K-12 is an ideal anticancer bacterial strain due to its ability to specifically accumulate within tumors while being rapidly cleared from the rest of the body (13, 17). Recent evidence suggests that this bacterium can inhibit the development of CRC by affecting the transcriptome of cancer cells (18). In this study, we investigated the role of *E. coli* K-12 in altering

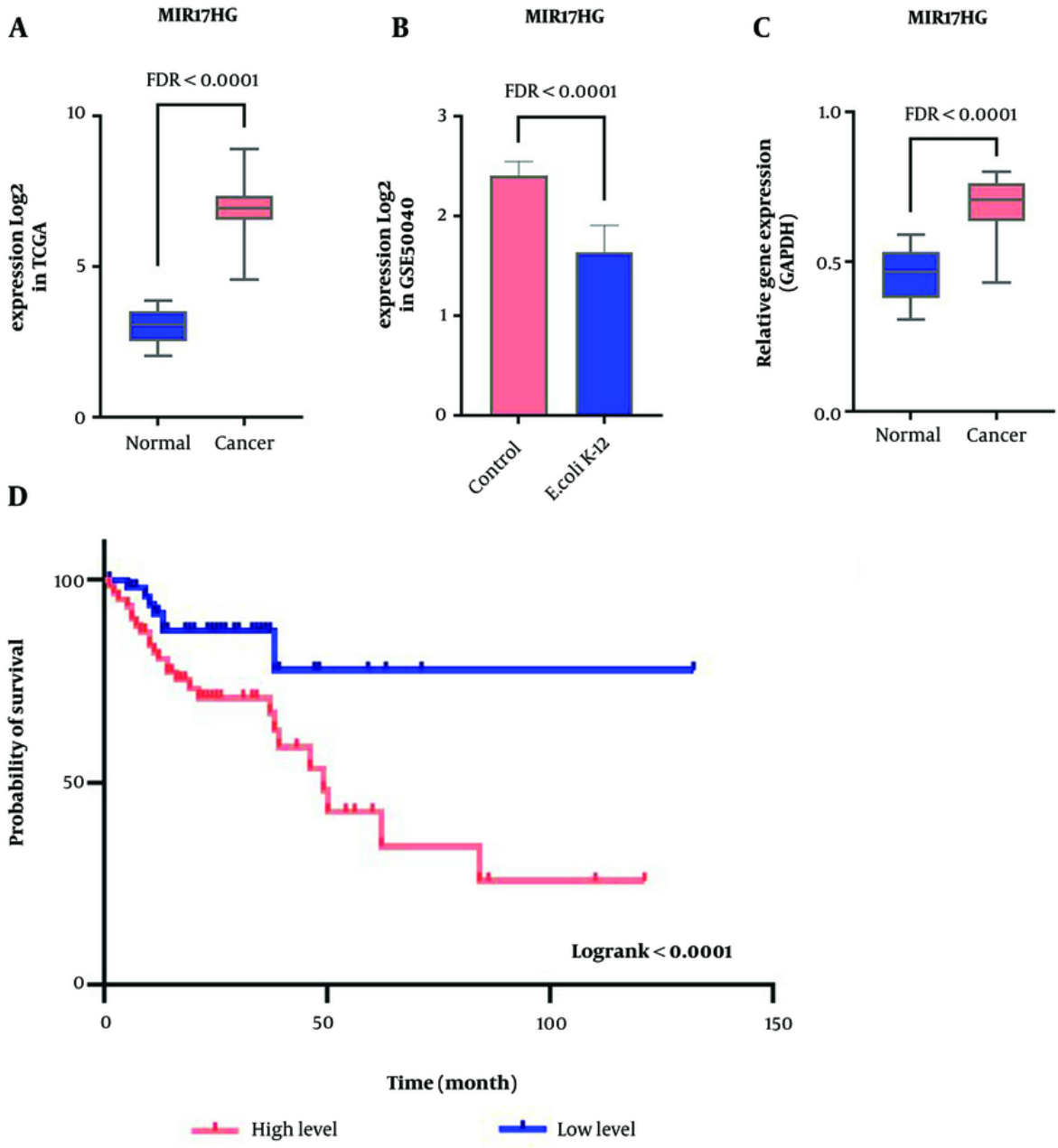

Figure 1. The high MIR17HG expression level as a poor prognosis biomarker downregulated by *Escherichia coli* K-12 in CRC. A, significant expression changes of *MIR17HG* in tumor samples compared to normal based on TCGA data; B, the *E. coli* K-12 impact on decreasing *MIR17HG* expression level; C, tthe RT-qPCR results for the expression of *MIR17HG* in forty CRC samples and adjacent normals; D, the association of *MIR17HG* expression with survival rate is shown through the Kaplan-Meier plot in CRC.

the expression of *MIR17HG*, a key lncRNA in CRC carcinogenesis.

It has been reported that lncRNA *MIR17HG* plays a significant role in CRC liver metastasis (CRLM) by promoting aerobic glycolysis in CRC patients. *MIR17HG* transcription is elevated due to glycolysis-accelerated lactate accumulation via the p38/Elk-1 pathway (23). Our findings showed that the expression levels of *MIR17HG* in

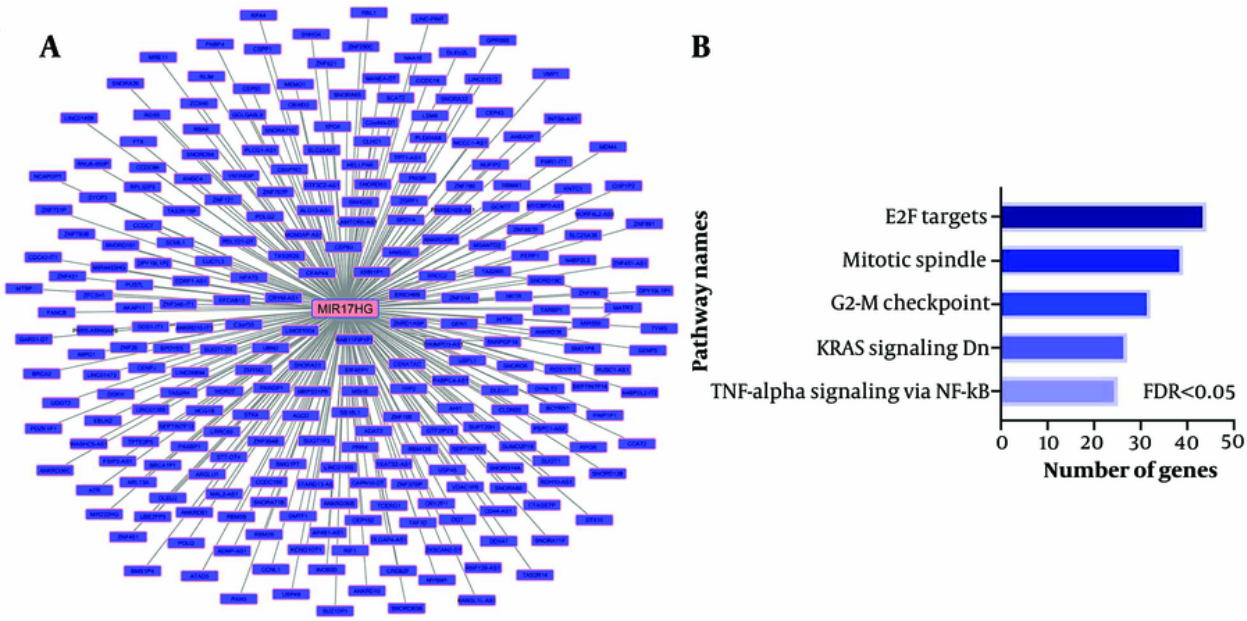

Figure 2. Co-expression of *MIR17HG* with genes related to cancer development pathways. A, the co-expression network of all genes that had an expression correlation with *MIR17HG* at R > 0.5 and P < 0.01. The Pearson correlation test was performed between *MIR17HG* expression and all genes in TCGA colorectal cancer samples; B, enrichment results of all correlated genes with *MIR17HG* demonstrate their significant role in cancer progression pathways.

tumor samples significantly increased compared to normal tissues. Additionally, our survival analysis revealed that elevated *MIR17HG* expression is strongly associated with poor prognosis and CRC malignancy in patients. Furthermore, in silico results indicated the potential of *E. coli* K-12 in reducing *MIR17HG* expression in CRC cells.

The co-expression network in our study demonstrated that genes co-expressed with *MIR17HG* were associated with pathways such as TNF-alpha signaling via NF-kB, KRAS signaling, G2-M checkpoint, mitotic spindle, and E2F targets. Previous research shows that these pathways are related to cancer cell proliferation, spindle abnormalities, genomic instability, inhibition of apoptosis, cancer cell migration, invasion, and malignancy through their effects on immune cell infiltration (24-27). These findings suggest that *MIR17HG* plays a critical role in CRC pathogenesis through these pathways.

In summary, both in silico and ex vivo results of this study demonstrated that MIR17HG expression significantly increases in CRC, likely contributing to cancer development and malignancy. Additionally, we concluded that *E. coli* K-12 may modulate MIR17HG expression, potentially inhibiting cancer progression. However, further in vitro and in vivo studies are required to confirm the role of *E. coli* K-12 in reducing cancer progression by downregulating *MIR17HG*.

5.1. Conclusions

According to our findings, *MIR17HG* exhibits oncogenic properties and can serve as a biomarker for CRC prognosis. Moreover, *E. coli* K-12 may play an inhibitory role in CRC development by reducing *MIR17HG* expression. Therefore, *E. coli* K-12 could be beneficial in decreasing the malignancy of CRC through its impact on *MIR17HG*.

Footnotes

Authors' Contribution: The study design was performed by H.S. and Z.R; data analysis and experiments were done by Z.R; interpretations of data were performed by Z.R., H.S., and E.R.; bioinformatics analysis was performed by Z.R.; manuscript writing was performed by Z.R.; the manuscript was finally approved by Z.R., H.S. and E.R.

Conflict of Interests Statement: There are no potential conflicts of interest to declare by the authors.

Data Availability: The dataset presented in the study is available on request from the corresponding author during submission or after its publication.

Ethical Approval: The review board of Imam Khomeini Hospital thoroughly examined and confirmed all bioethical issues by the criteria set forth by the Ministry of Health and Medical Education of Iran. This study obtained approval from the Biomedical Ethics Committee of Islamic Azad University with the ethics code IR.IAU.PS.REC.1401.319 .

Funding/Support: There is no funding or support to declare by the authors.

References

1. Xi Y, Xu P. Global colorectal cancer burden in 2020 and projections to 2040. *Transl Oncol*. 2021;**14**(10):101174. [PubMed ID: 34243011]. [PubMed Central ID: PMC8273208]. https://doi.org/10.1016/j.tranon.2021.101174.
2. Ciardiello F, Ciardiello D, Martini G, Napolitano S, Tabernero J, Cervantes A. Clinical management of metastatic colorectal cancer in the era of precision medicine. *CA Cancer J Clin*. 2022;**72**(4):372-401. [PubMed ID: 35472088]. https://doi.org/10.3322/caac.21728.
3. Xu H, Liu L, Li W, Zou D, Yu J, Wang L, et al. Transcription factors in colorectal cancer: molecular mechanism and therapeutic implications. *Oncogene*. 2021;**40**(9):1555-69. [PubMed ID: 33323976]. https://doi.org/10.1038/s41388-020-01587-3.
4. Kheirelseid EA, Miller N, Chang KH, Nugent M, Kerin MJ. Clinical applications of gene expression in colorectal cancer. *J Gastrointest Oncol*. 2013;**4**(2):144-57. [PubMed ID: 23730510]. [PubMed Central ID: PMC3635192]. https://doi.org/10.3978/j.issn.2078-6891.2013.010.
5. Fang XY, Pan HF, Leng RX, Ye DQ. Long noncoding RNAs: novel insights into gastric cancer. *Cancer Lett*. 2015;**356**(2 Pt B):357-66. [PubMed ID: 25444905]. https://doi.org/10.1016/j.canlet.2014.11.005.
6. Schmitt AM, Chang HY. Long Noncoding RNAs in Cancer Pathways. *Cancer Cell*. 2016;**29**(4):452-63. [PubMed ID: 27070700]. [PubMed Central ID: PMC4831138]. https://doi.org/10.1016/j.ccell.2016.03.010.
7. Huarte M. The emerging role of lncRNAs in cancer. *Nat Med*. 2015;**21**(11):1253-61. [PubMed ID: 26540387]. https://doi.org/10.1038/nm.3981.
8. Jiang MC, Ni JJ, Cui WY, Wang BY, Zhuo W. Emerging roles of lncRNA in cancer and therapeutic opportunities. *Am J Cancer Res*. 2019;**9**(7):1354-66. [PubMed ID: 31392074]. [PubMed Central ID: PMC6682721].
9. Zhuo W, Liu Y, Li S, Guo D, Sun Q, Jin J, et al. Long Noncoding RNA GMAN, Up-regulated in Gastric Cancer Tissues, Is Associated With Metastasis in Patients and Promotes Translation of Ephrin A1 by Competitively Binding GMAN-AS. *Gastroenterology*. 2019;**156**(3):676-691 e11. [PubMed ID: 30445010]. https://doi.org/10.1053/j.gastro.2018.10.054.
10. Olivero CE, Martinez-Terroba E, Zimmer J, Liao C, Tesfaye E, Hooshdaran N, et al. p53 Activates the Long Noncoding RNA Pvt1b to Inhibit Myc and Suppress Tumorigenesis. *Mol Cell*. 2020;**77**(4):761-774 e8. [PubMed ID: 31973890]. [PubMed Central ID: PMC7184554]. https://doi.org/10.1016/j.molcel.2019.12.014.
11. Hu Q, Ye Y, Chan LC, Li Y, Liang K, Lin A, et al. Oncogenic lncRNA downregulates cancer cell antigen presentation and intrinsic tumor suppression. *Nat Immunol*. 2019;**20**(7):835-51. [PubMed ID: 31160797]. [PubMed Central ID: PMC6619502]. https://doi.org/10.1038/s41590-019-0400-7.
12. Yao J, Xiao G, Kong D, Ye C, Chen R, Li L, et al. The Long Noncoding RNA TTTY15, Which Is Located on the Y Chromosome, Promotes Prostate Cancer Progression by Sponging let-7. *Eur Urol*. 2019;**76**(3):315-26. [PubMed ID: 30527798]. https://doi.org/10.1016/j.eururo.2018.11.012.
13. Weibel S, Stritzker J, Eck M, Goebel W, Szalay AA. Colonization of experimental murine breast tumours by Escherichia coli K-12 significantly alters the tumour microenvironment. *Cell Microbiol*. 2008;**10**(6):1235-48. [PubMed ID: 18208564]. https://doi.org/10.1111/j.1462-5822.2008.01122.x.
14. Min JJ, Nguyen VH, Kim HJ, Hong Y, Choy HE. Quantitative bioluminescence imaging of tumor-targeting bacteria in living animals. *Nat Protoc*. 2008;**3**(4):629-36. [PubMed ID: 18388945]. https://doi.org/10.1038/nprot.2008.32.
15. Lin D, Feng X, Mai B, Li X, Wang F, Liu J, et al. Bacterial-based cancer therapy: An emerging toolbox for targeted drug/gene delivery. *Biomaterials*. 2021;**277**:121124. [PubMed ID: 34534860]. https://doi.org/10.1016/j.biomaterials.2021.121124.
16. Koli P, Sudan S, Fitzgerald D, Adhya S, Kar S. Conversion of commensal Escherichia coli K-12 to an invasive form via expression of a mutant histone-like protein. *mBio*. 2011;**2**(5). [PubMed ID: 21896677]. [PubMed Central ID: PMC3172693]. https://doi.org/10.1128/mBio.00182-11.
17. Jiang SN, Phan TX, Nam TK, Nguyen VH, Kim HS, Bom HS, et al. Inhibition of tumor growth and metastasis by a combination of Escherichia coli-mediated cytolytic therapy and radiotherapy. *Mol Ther*. 2010;**18**(3):635-42. [PubMed ID: 20051939]. [PubMed Central ID: PMC2839435]. https://doi.org/10.1038/mt.2009.295.
18. Rohani Z, Sazegar H, Rahimi E. Unlocking the potential of Escherichia coli K-12: A novel approach for malignancy reduction in colorectal cancer through gene expression modulation. *Gene*. 2024;**906**:148266. [PubMed ID: 38342251]. https://doi.org/10.1016/j.gene.2024.148266.
19. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res*. 2016;**44**(8). e71. [PubMed ID: 26704973]. [PubMed Central ID: PMC4856967]. https://doi.org/10.1093/nar/gkv1507.
20. Law CW, Alhamdoosh M, Su S, Dong X, Tian L, Smyth GK, et al. RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. *F1000Res*. 2016;**5**. [PubMed ID: 27441086]. [PubMed Central ID: PMC4937821]. https://doi.org/10.12688/f1000research.9005.3.
21. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol*. 2010;**11**(3):R25. [PubMed ID: 20196867]. [PubMed Central ID: PMC2864565]. https://doi.org/10.1186/gb-2010-11-3-r25.
22. Smyth GK. limma: Linear Models for Microarray Data. In: Gentleman R, Carey VJ, Huber W, Irizarry RA, Dudoit S, editors. *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. New York, NY: Springer New York; 2005. p. 397-420. https://doi.org/10.1007/0-387-29362-0_23.
23. Zhao S, Guan B, Mi Y, Shi D, Wei P, Gu Y, et al. LncRNA MIR17HG promotes colorectal cancer liver metastasis by mediating a glycolysis-associated positive feedback circuit. *Oncogene*. 2021;**40**(28):4709-24. https://doi.org/10.1038/s41388-021-01859-6.
24. Zinatizadeh MR, Schock B, Chalbatani GM, Zarandi PK, Jalali SA, Miri SR. The Nuclear Factor Kappa B (NF-kB) signaling in cancer development and immune diseases. *Genes Dis*. 2021;**8**(3):287-97. [PubMed ID: 33997176]. [PubMed Central ID: PMC8093649]. https://doi.org/10.1016/j.gendis.2020.06.005.
25. Ternet C, Kiel C. Signaling pathways in intestinal homeostasis and colorectal cancer: KRAS at centre stage. *Cell Commun Signal*. 2021;**19**(1):31. [PubMed ID: 33691728]. [PubMed Central ID: PMC7945333]. https://doi.org/10.1186/s12964-021-00712-3.

26. Wang Z, Zheng Z, Wang B, Zhan C, Yuan X, Lin X, et al. Characterization of a G2M checkpoint-related gene model and subtypes associated with immunotherapy response for clear cell renal cell carcinoma. *Heliyon*. 2024;**10**(7). e29289. [PubMed ID: 38617927]. [PubMed Central ID: PMC11015143]. https://doi.org/10.1016/j.heliyon.2024.e29289.

27. Zhang X, Ni Z, Duan Z, Xin Z, Wang H, Tan J, et al. Overexpression of E2F mRNAs Associated with Gastric Cancer Progression Identified by the Transcription Factor and miRNA Co-Regulatory Network Analysis. *Plos One*. 2015;**10**(2). https://doi.org/10.1371/journal.pone.0116979.