Original Article

Discovery of Biomarkers Related to Colorectal Cancer by Systematic Proteomics Analysis and Experimental Procedures

Noosha Zia Jahromi[1], PhD; Amin Mahmoudi Mandani[1], MD; Nahid Sadeghi Moghadam Kaji[1], Msc; Hadis Abbasi Teshnizi[1], MSc; Farnaz Nourmohammadian Dehkordi[2], PhD; Maryam Yousefi[3,4], PhD; Hossein Sazegar[1], PhD

[1]Department of Biology, ShK.C., Islamic Azad University, Shahrekord, Iran;
[2]Department of Biology, Science and Research Branch, Islamic Azad University of Babol, Babol, Iran;
[3]Genetic and Biotechnology Institute, Isfahan, Iran;
[4]Department of Biology, Science and Research Branch, Islamic Azad University, Tehran, Iran

Correspondence:
Hossein Sazegar, PhD;
Department of Biology, ShK.C., Islamic Azad University, Shahrekord Branch, Postal code: 88137-33395, Shahrekord, Iran
Tel: +98 38 33361000
Email: sazegarhossein8@gmail.com
Received: 18 May 2025
Revised: 05 September 2025
Accepted: 06 October 2025

What's Known

- Colorectal cancer (CRC) is associated with mRNA processing and immune response-related pathways, which are crucial in tumor progression and metastasis.
- Regulatory molecules such as miRNAs and lncRNAs are pivotal players in CRC pathogenesis, but their roles in detailed mechanisms remain underexplored in certain key pathways.

What's New

- This study constructed a multi-layer network integrating DEPs, miRNAs, and lncRNAs, revealing novel molecules such as SLC9A3-AS1 and LINC00943, as potential diagnostic and therapeutic targets.
- It highlights specific splicing and immune-related mechanisms, providing actionable insights into the complex interactions driving CRC progression.

Abstract

Background: Long non-coding RNAs (lncRNAs) play critical roles in gene expression regulation and have emerged as potential biomarkers for cancer diagnosis and treatment. This study aimed to investigate the potential role of two lncRNAs, LINC00943 and SLC9A3-AS1, as candidate biomarkers for colorectal cancer (CRC) diagnosis and therapy.

Methods: Initially, we investigated differentially expressed proteins (DEPs) in CRC through a comprehensive bioinformatics analysis and constructed a multi-layer regulatory network to understand the complex interactions. Subsequently, tissue samples were collected from CRC patients at Milad Hospital in Isfahan, Iran, between June and December 2021. The expression levels of LINC00943 and SLC9A3-AS1 were evaluated using quantitative reverse transcription polymerase chain reaction (qRT-PCR). Statistical analyses were performed using GraphPad Prism version 9. Expression differences between tumor and adjacent non-tumor tissues were evaluated using paired Student's *t* test. Receiver operating characteristic (ROC) curve analysis was conducted to evaluate diagnostic performance. A P value less than 0.05 was considered statistically significant.

Results: *LINC00943* expression was significantly upregulated in tumor tissues compared to adjacent controls (P<0.0001), with strong diagnostic performance (AUC=0.8078, sensitivity=83.3%, specificity=76.7%). *SLC9A3-AS1* also showed increased expression (P=0.0299), but with limited diagnostic value (AUC=0.6124). Network analysis identified SERBP1, KHSRP, and HNRNPA1 as central hub proteins, while *miR-15b-5p* and *NEAT1* emerged as key regulatory elements.

Conclusion: These findings suggest that *LINC00943* could serve as a valuable biomarker for CRC diagnosis and treatment, providing insights into the regulatory mechanisms underlying CRC pathogenesis.

This manuscript has been released as a preprint at the Unraveling Candidate Biomarkers and Involved Biomedical Phenomena in Colorectal Cancer by Systematic Proteomics Analysis and Experimental Procedures. DOI: https://doi.org/10.21203/rs.3.rs-1974798/v1

Please cite this article as: Zia Jahromi N, Mahmoudi Mandani A, Sadeghi Moghadam Kaji N, Abbasi Teshnizi H, Nourmohammadian Dehkordi F, Yousefi M, Sazegar H. Discovery of Biomarkers Related to Colorectal Cancer by Systematic Proteomics Analysis and Experimental Procedures. Iran J Med Sci. doi: 10.30476/ijms.2025.107064.4143.

Keywords ● Systems biology ● Proteomics ● Colorectal neoplasms ● Biomarkers ● Bioinformatics

Introduction

Over the past decades, the number of colorectal cancer (CRC) patients has dramatically increased, making it the third leading cause of cancer-related deaths worldwide with over 930,000 fatalities annually.[1, 2] This alarming rise in mortality rates indicates that our current understanding of CRC pathogenesis remains incomplete. The predominant research focus on individual genes, proteins, or single signaling pathways may not sufficiently unravel the complex mechanisms underlying CRC. Consequently, there is a pressing need for more comprehensive approaches to investigate the molecular basis of CRC.

Recent advancements in systems biology have renewed interest in researching complex disorders by offering a holistic perspective on disease mechanisms. Systems biology integrates data from various biological layers, such as genomics, proteomics, and transcriptomics, to provide a comprehensive understanding of disease processes. This approach allows researchers to elucidate the intricate interactions and regulatory networks within biological systems, facilitating the identification of novel therapeutic targets and biomarkers.[3, 4]

One such modern approach involves constructing multi-layer networks (DEPs-miRNAs-lncRNAs), which can effectively identify key players in various regulatory layers. Multi-layer network construction involves integrating differentially expressed proteins (DEPs), microRNAs (miRNAs), and long non-coding RNAs (lncRNAs) to uncover the regulatory interactions that drive disease progression.[5] However, few studies have applied multi-layer network construction to CRC, particularly focusing on protein expression changes.

Based on this framework, our central hypothesis is that integrating DEPs with regulatory layers of miRNAs and lncRNAs will reveal critical molecular mechanisms driving CRC progression. Specifically, we propose that the lncRNAs, LINC00943 and SLC9A3-AS1, act as central regulatory hubs, modulating splicing and immune-related pathways, and thereby contributing to tumor development. This approach aims to provide a comprehensive view of the regulatory interactions and identify critical drivers involved in CRC pathogenesis.

Materials and Methods

Bioinformatics Study

Proteomics Dataset Acquisition: A proteomics dataset (PXD005693) generated by Costanza and colleagues was obtained from the ProteomeXchange database (http://proteomecentral.proteomexchange.org),[6] comprising protein expression profiles from six CRC tumor tissues and their corresponding non-tumor tissues.[7]

In-silico Protein Identification: To detect the proteins from the peptide spectrum, the target-decoy strategy was applied utilizing MaxQuant software (Max Planck Institute of Biochemistry, Germany; https://maxquant.org). The following parameters were set to optimize protein identification: 1. Methionine oxidation, N-terminal acetylation, and cysteine carbamidomethyl were selected as variable and fixed post-translational modifications, respectively, 2. Trypsin enzyme with up to two missed cleavages was selected as theoretical cleavage enzyme. 3. The maximum peptide length for identification was set to seven amino acids, 4. False discovery rate (FDR)<0.05 in PSM was used as the threshold for statistical significance in protein identification. 5. Match between runs was enabled to match undetected peptides among technical replicates.[7, 8]

Quality Control Evaluation: Two unsupervised methods, principal component analysis and hierarchical clustering, were used to evaluate the quality of the samples. The ggplot2 package (tidyverse, USA; https://ggplot2.tidyverse.org) in R software (R Foundation for Statistical Computing, Vienna, Austria; https://www.r-project.org) was also used to visualize each plot.[9]

Protein Quantification: The differentially expressed proteins were screened by Perseus software (Max Planck Institute of Biochemistry, Martinsried, Germany; https://maxquant.org). In this process, the following filters were applied: removal of contaminants, reverse hits, and proteins identified only by site. Log2 transformation was applied, and proteins with at least three valid intensity values in any condition were retained for further analysis. A *t* test of two samples based on Benjamini-Hocheberg as one of the strict FDR strategies was used, and Q value was set at 0.05. To make sure about the quality control of the dataset, PCA, as well as heatmap clustering using Euclidean correlation, were performed.[10]

Enrichment Analysis: To explore the functionals of identified DEPs, gene ontology in three sections (Biological Process (BP), Molecular Function (MF), and Cellular Component (CC) analysis was applied employing the ClueGO plugin in Cytoscape software (The Cytoscape Consortium, USA; https://cytoscape.org), and adj. P<0.05 was selected as the cut-off criteria. After that, the obtained results were summarized in parent terms utilizing the REVIGO database.[11, 12]

Pathway Enrichment Analysis: Reactome (Ontario Institute for Cancer Research, Toronto, Canada; https://reactome.org), an essential resource for understanding the functions and biological pathways of DEPs, was used for pathway enrichment analysis; each term with adj. P value<0.05 was considered the significantly altered pathway. For the next step, the EnrichR database was applied to visualize the obtained results.[13]

miRNAs and lncRNAs prediction: To predict the microRNA based on identified DEPs, miRTarBase section (v.) via the EnrichR database was applied, adjusted P<0.05 was used as a significant statistical criterion to select the most reliable microRNA. The miRNet database was employed in order to find lncRNAs associated with predicted microRNAs; only lncRNAs with adjusted P value<0.05 were considered potentially correlated with the predicted miRNAs.[14, 15]

Multi-Layer Network Construction: First, a protein-protein interaction (PPI) network was constructed among identified DEPs using STRING (https://string-db.org), and interaction edges were retrieved by the CluePedia plugin in Cytoscape. After creating PPI interaction, a multi-layer network was constructed amid curated lncRNAs, microRNA, and DEPs using Cytoscape software. After network construction, central molecule in each layer were identified based on topological parameters (degree, betweenness, and closeness) separately using the cytoHubba plugin.[16, 17]

Practical Study; Samples Collection

Samples were collected from patients at the time of initial diagnosis, before receiving any therapy. Patients with major comorbidities were excluded as much as possible, although some patient characteristics were beyond our control. Tissue sampling was conducted with the informed consent of patients with CRC at Milad hospital, Isfahan, Iran, between June and December 2021. Thirty tumor tissues and thirty adjacent normal tissues were preserved in RNAlater solution. All ethical aspects of the study were approved and supervised by the Ethical Committee of Milad Hospital, Isfahan. Demographic and clinical details of the CRC patients are provided in table 1.

The study protocol was reviewed and approved by the Ethics Committee of the Islamic Azad University of Shahrekord (IR.IAU.SHK.REC.1400.051). All analyses were conducted in accordance with the principles of the Declaration of Helsinki. The patient and his parents provided written informed consent to participate in this study.

Expression Assay

According to the protocol, RNA extraction from tissues was performed using Trizol (Invitrogen, USA). cDNA was then synthesized using first-strand cDNA synthesis kit (TaKaRa, Shiga, Japan), and SYBR green master mix (Amplicon, Odense, Denmark) was used for the qRT-PCR reaction. The primers used to check the expression of each gene are provided in table 2.

Statistical analysis was conducted using GraphPad Prism software version 9.0 (GraphPad Software, USA). Expression differences were assessed using t-test and one-way ANOVA. The data were analyzed using both the $2^{-\Delta\Delta CT}$ method and log2 fold change transformation, and statistical significance was set at P<0.05.

Results

Bioinformatics Results

The rapid growth of high-throughput technologies in biology has generated an increasing amount of data, which should be re-evaluated to gain valuable information. Accordingly, in this investigation, in order to provide a comprehensive understanding of the underlying mechanisms of CRC, a pre-existing proteomics dataset was assessed. Additionally,

Table 1: Demographic and clinical details of the colorectal cancer patients

Demographics		Colorectal cancer patients
Sample size		30
Sex	Male (%)	12 (40%)
	Female (%)	18 (60%)
Age range (year, mean±SD)		62.6±10.4
Tumor size (cm, mean±SD)		5±2.2
Grade	Grade I	22 (73.3%)
	Grade II	8 (26.7%)
Stage	Stage I	8 (26.7%)
	Stage II	20 (66.7%)
	Stage III	2 (6.6%)
Metastasis	Positive	11 (36.7%)
	Negative	19 (63.3%)

Table 2: The sequences of primers used in the study

LncRNAs	Primers
LINC00943	Forward: 5′ ACAGATACCAGATACGGGACA 3′ Reverse: 5′ GGAAGTGACTATTGCAACCCA 3′
SLC9A3-AS1	Forward: 5′ TGCGGGAGAGGAAGGCGAG 3′ Reverse: 5′ CCGCAACCTGAAAGGCACG 3′

a multi-layer network was constructed as a valuable analytical method to give a holistic insight into the regulator interaction of components. The selected dataset encompassed the colorectal non-cancerous and cancerous tissues reported by Costanza and colleagues, in the main article, the pivotal role of EXPEL as a novel method to isolate specific biological layers in order to be introduced as a diagnostic and prognostic biomarker was demonstrated.[6] To transform the RAW data into an informative protein table, MaxQuant (Max Planck Institute of Biochemistry, Martinsried, Germany; https://maxquant.org) was applied. Accordingly, 3024 proteins were identified. In the next step, to ensure the quality of samples, PCA, a predominant unsupervised method for quality assessment, was used; the analysis showed that most samples were segregated based on experimental groups (figure 1). Although a few samples that did not cluster according to group categorization, were eliminated from further analysis to improve data quality. In the following, 1148 differentially expressed proteins were identified utilizing Perseus software Max Planck Institute of Biochemistry, Martinsried, Germany; https://cox-labs.github.io/coxdocs/perseus_instructions.html), considering FDR<0.05 as the statistically significant threshold. Pathway enrichment analysis was carried out to decipher the biomedical phenomena of identified DEPs. The preliminary analysis of pathway enrichment results indicates that immune response, mRNA metabolism, and hemoglobin scavenger receptor-related pathways are the main enriched pathways identified based on DEPs. Furthermore, actin cytoskeleton-related pathways, such as Rho GTPase activity, were also identified that pertains to altered proteins (figure 2). To provide a comprehensive view of GO

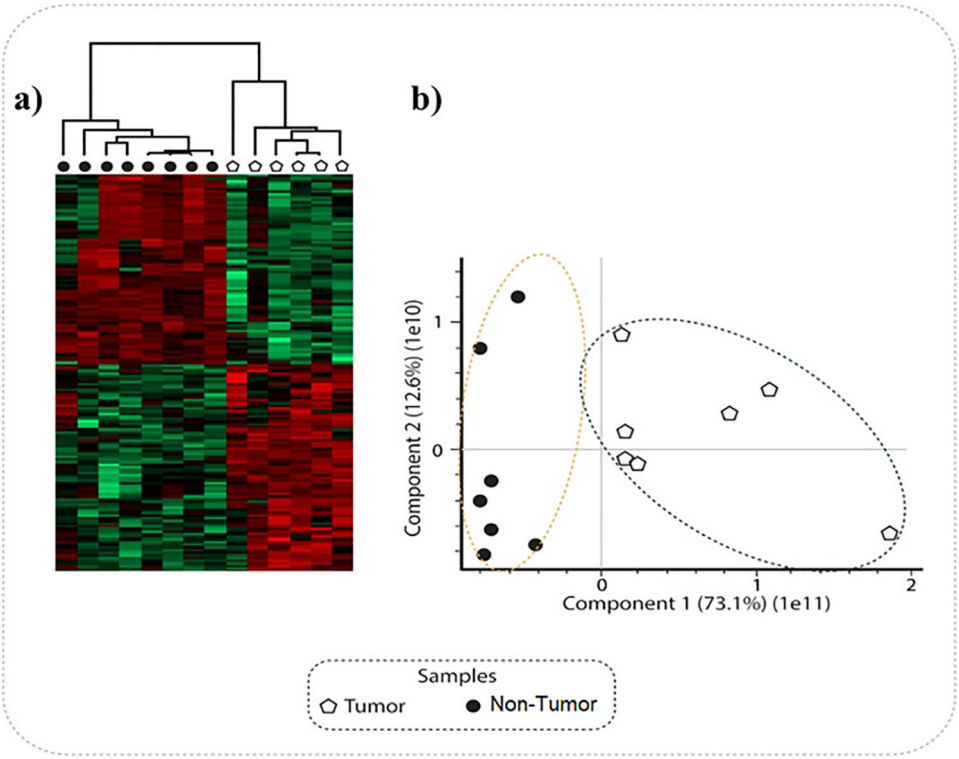

Figure 1: The figure illustrates the assessment of dataset quality using unsupervised clustering methods. a) A heatmap of gene expression data with hierarchical clustering is shown. The color gradient from red to green indicates relative gene expression levels, where red denotes high expression, green denotes low expression, and black represents intermediate levels. The dendrogram above the heatmap illustrates the clustering relationships among samples. Tumor samples are marked with open circles, while non-tumor samples are indicated by filled circles, confirming distinct grouping between the two conditions. b) A principal component analysis (PCA) scatter plot demonstrates the distribution of samples along Component 1 (73.1% variance) and Component 2 (12.6% variance). Tumor and non-tumor samples are represented by open and filled circles, respectively. Ellipses highlight the clustering of each group, demonstrating clear separation and supporting the integrity of the dataset.

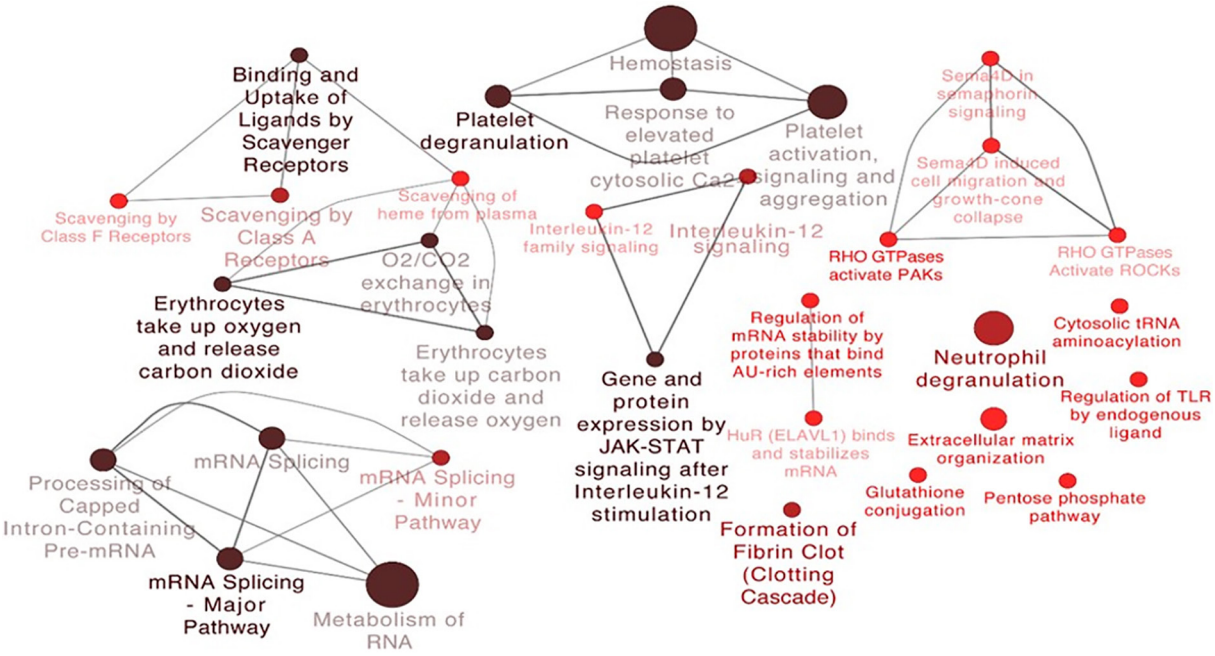

Figure 2: The figure demonstrates the functional enrichment analysis of differentially expressed proteins (DEPs) using cytoscape software. Cytoscape software together with the ClueGO plugin were employed to perform functional enrichment analysis based on the Reactome database. The resulting network diagram illustrates key biological processes and pathways associated with mRNA processing and immune response. Nodes represent enriched pathways, and edges indicate functional relationships between them. The color of each node reflects statistical significance, with an adjusted P value threshold of <0.05.

categorization, DEPs were examined concerning BP, MF, and CC. In this connection, GO analysis showed that the biological process (BP) of the DEPs was principally involved in two main classifications, such as “mRNA metabolism”. This can be illustrated by the “regulation of mRNA metabolic process” and “mRNA metabolic process”. The second classification of BP concerns the immune response function, such as “Neutrophil activation” and “platelet degranulation”. The most striking finding from pathway and biological process analysis was that all terms consistently indicated the alteration of mRNA metabolism and immune system-related activity. Interestingly, considerably enriched terms are directly attributed to scavenger receptor functions.

Notably, in agreement with the above findings, the majority of identified terms in the MF category are related to “mRNA binding”. Another primary enriched term is “cell adhesion molecule binding”, which aligns with the previous pathway analysis findings. Furthermore, the only significantly enriched term in the CC category was “extracellular exosome” (figure 3).

To provide a holistic insight into the potential interaction of DEPs, a PPI network was constructed using the STRING database as the interaction source. As a further step, given the regulatory essence of miRNAs and lncRNAs as practical elements in the expression of proteins, we were interested in making an integrative network between DEPs and their related regulatory elements; hence, a multi-layer network comprising proteins, miRNAs, and lncRNAs was constructed, which was evaluated in terms of well-known graph theory topological concepts (Degree, Betweenness, and Closeness) to recognize the key-drivers in each layer as an indispensable part of a biological network (figure 4). The top 10 nodes attributed to each predicted layer were considered hub molecules based on their highest topological score. The topmost identified hub-DEPs were KHSRP, HNRNPA1, SERBP1, HSPA8, CALR, SRF2, SERBP1, in particular, can be considered a decisive molecule in CRC development and warrants further investigation, which is inspected as a decisive molecule in CRC development.[18] Furthermore, the top regulatory miRNAs in the created multi-layer network comprised miR-15b-5p, miR-424-5p, miR-1301-3p, miR-16-5p, and miR-497-5p. Given their nature, they may serve as non-invasive biosignatures for early recognition of these kinds of complex conditions.[19] For example, miR-15b-5p has been identified as a biomarker in CRC through both systematic and experimental studies.[20]

Concerning the fact that lncRNAs, as one of the main regulatory biological layers, play an influential part in the alteration of protein expression levels through the impact on transcriptional and post-translational regulation,[21] they were therefore included

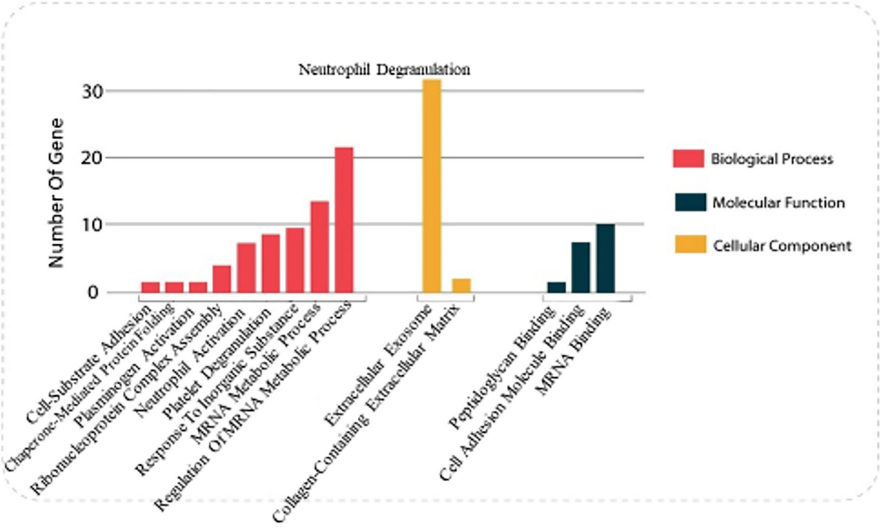

Figure 3: Gene Ontology (GO) enrichment analysis was performed on differentially expressed proteins (DEPs), categorized into three GO domains: Biological Process (red), Molecular Function (blue), and Cellular Component (yellow). The bar chart displays the number of genes associated with each GO category, as indicated on the x-axis. The y-axis represents the number of genes identified per category.

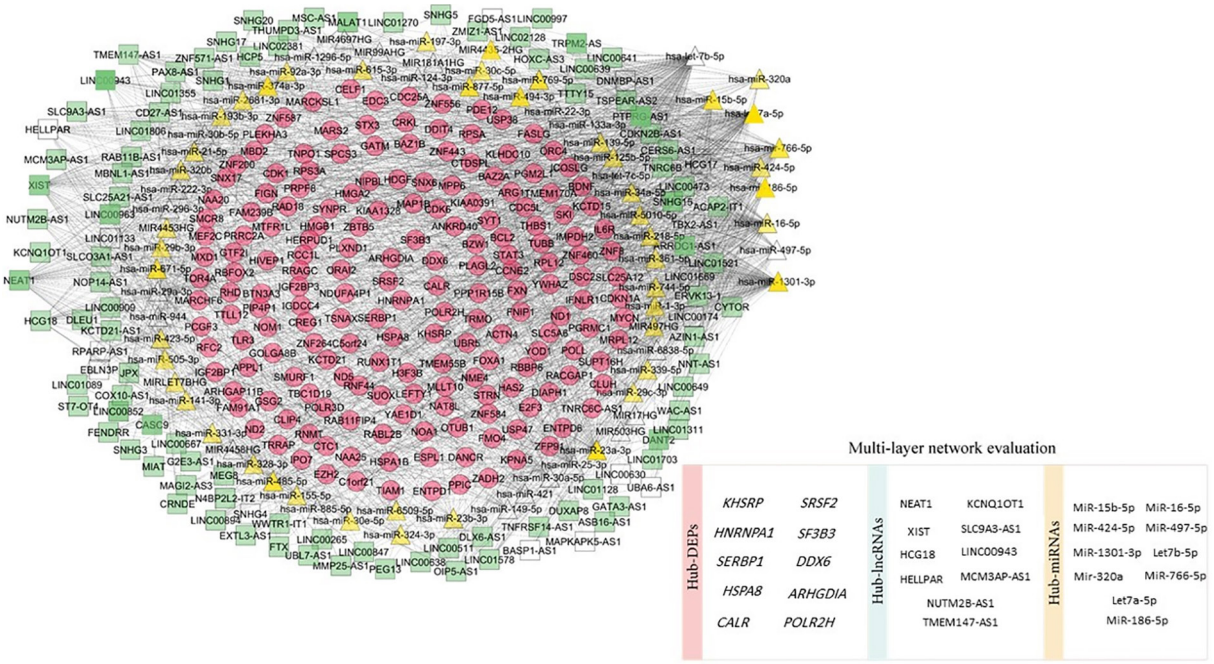

Figure 4: The figure illustrates the integrative multi-layer network construction describing interactions among differentially expressed proteins (DEPs), predicted microRNAs (miRNAs), and long non-coding RNAs (lncRNAs). The network was generated using the CluePedia plugin in Cytoscape. Nodes are color-coded by biomolecule type: blue for lncRNAs, yellow for miRNAs, and red for proteins. The circular layout highlights the complexity and connectivity of the regulatory network. Hub elements from each layer, identified based on centrality and interaction density, are listed alongside the network, revealing key regulatory components potentially involved in the studied biological context.

as another component of the multi-layer network. According to the network evaluation approach, several top lncRNAs were identified, including *NEAT1*, *KCNQ1OT1*, *XIST*, *HCG18*, *SLC9A3-AS1*, *LINC00943*, and *MCM3AP-AS1*.

In accordance with the centrality score, it could be seen that the top lncRNAs have previously been investigated in the progression of CRC. Nonetheless, the other elicited lncRNAs should be further examined in CRC.

Experimental Results

The results of this study showed that the expression of *LINC00943* in the tumor tissues was increased compared to that in the adjacent non-tumor tissues (P<0.0001), as analyzed using GraphPad Prism software. In addition, SLC9A3-AS1 expression was also significantly upregulated in tumor tissues compared to that in the adjacent non-tumor tissues (P=0.0299).

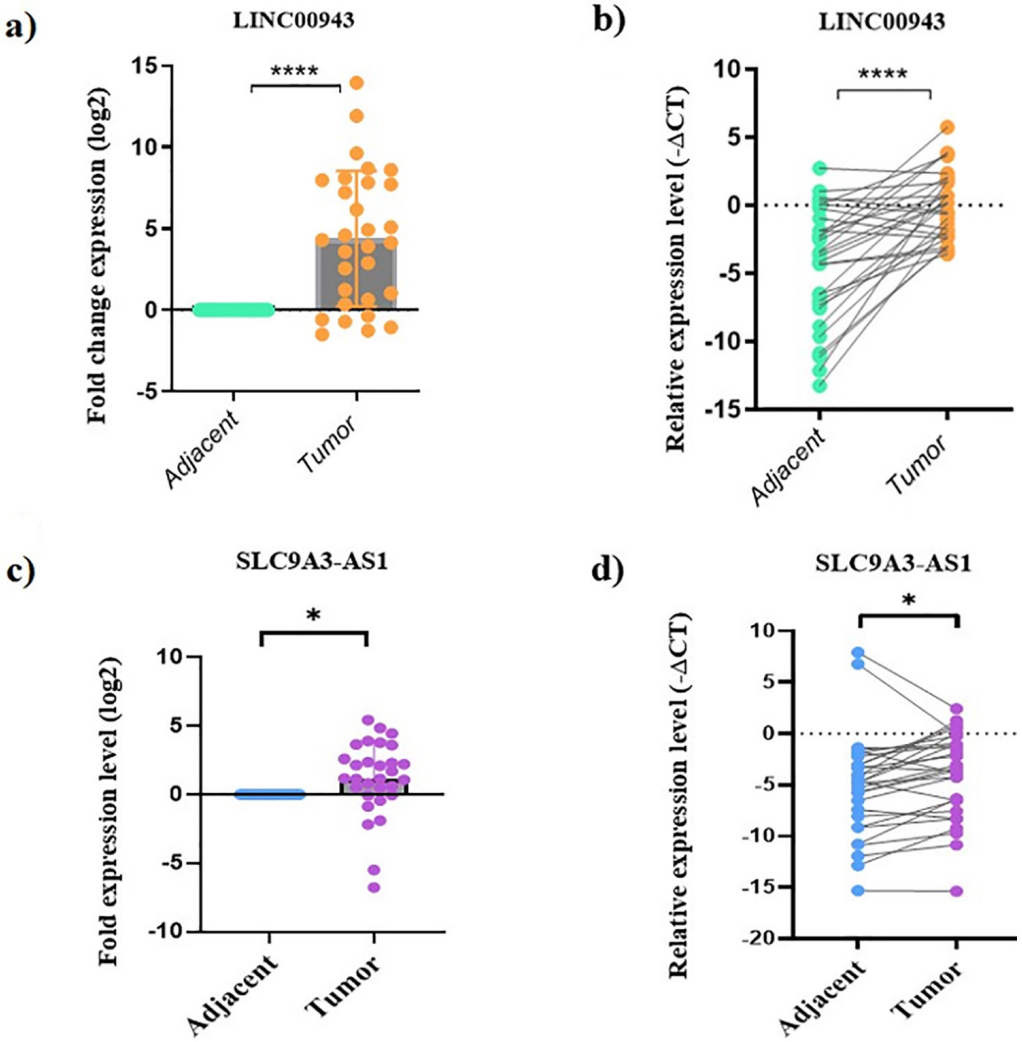

Figure 5: The figure presents the relative expression analysis of long non-coding RNAs (lncRNAs) in colorectal cancer tissues compared to adjacent normal tissues. a, b) The fold change (log_2) and relative expression (ΔCT) of *LINC00943*, showing significantly elevated expression in tumor tissues (****P<0.0001). c, d) The fold change (log_2) and relative expression (ΔCT) of *SLC9A3-AS1*, also showing increased expression in tumor tissues (*P=0.0299). Each dot represents an individual patient sample, with lines connecting paired adjacent and tumor tissues.

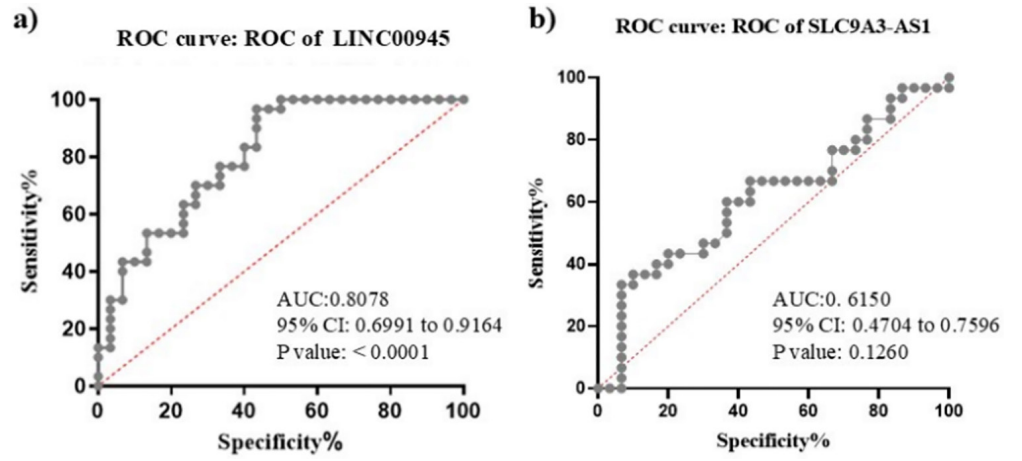

Figure 6: The figure presents the Receiver Operating Characteristic (ROC) curve analysis evaluating the diagnostic performance of two long non-coding RNAs (lncRNAs) in distinguishing colorectal tumor tissues from adjacent normal samples. a) The ROC curve for *LINC00943* shows strong discriminatory power with an AUC of 0.8078 (95% CI: 0.6991–0.9164, P<0.0001), indicating high sensitivity and specificity. b) The ROC curve for *SLC9A3-AS1* displays moderate diagnostic potential with an AUC of 0.6150 (95% CI: 0.4704–0.7596, P=0.1260).

Table 3: RNA expression analysis in clinical pathology

Clinical feature	Category	SLC9A3-AS1 (mean±SD)	LINC00943 (mean±SD)
Metastasis	Positive	-3.593±3.898	3.412±3.220
	Negative	-4.147±4.430	4.953±4.598
	P value#	0.7333	0.3365
Grade	Grade I	-4.394±4.516	4.254±4.402
	Grade II	-2.706±4.145	4.754±3.633
	P value#	0.3371	0.7764
Stage	Stage I	-3.949±5.355	4.768±4.584
	Stage II	-3.710±3.884	4.144±4.218
	Stage III	-6.268±2.845	5.305±3.422
	P value##	0.7259	0.8966

#Unpaired *t* test; ##Ordinary One Way ANOVA; A significance level of P<0.05 was considered statistically significant.

Figure 5 shows the expression changes of *LINC00943* and *SLC9A3-AS1* in tumor versus adjacent non-tumor tissues. *LINC00943* was evaluated for its diagnostic utility in differentiating CRC tumor samples from adjacent controls. The transcript levels of LINC00943 showed significant diagnostic power, with an AUC of 0.8078. Additionally, measuring the AUC related to *SLC9A3-AS1* showed no significant diagnostic power. Figure 6 provides details about ROC analysis.

Table 3 shows the details of the expression profile of these RNAs in relation to clinicopathological. However, there is no significant relationship between the expression profile of these RNAs and clinicopathological features. This lack of significant association may be due to the limited sample size. Statistical analysis was performed using GraphPad Prism software, and differences between groups were assessed using a t-test and ordinary one-way ANOVA.

Discussion

This study provides an integrative, multi-layered analysis of CRC pathogenesis by combining protein expression data with regulatory elements such as miRNAs and lncRNAs. By comparing tumor with non-tumor tissues, we identified DEPs and constructed a multi-layer regulatory network. The overall results revealed two dominant biological themes, mRNA processing and immune response, that consistently appeared across enrichment and network analyses. Moreover, several hub molecules, including splicing regulators and immune-related proteins, emerged as potential biomarkers or therapeutic targets.

Cancer remains a major global health challenge and a leading cause of mortality worldwide.[22] Hence, despite the extensive survey, the exact mechanisms involved are hitherto rudimentary. However, considering the upward mobility of state-of-the-art equipment in the biological subject alongside the myriad generated big data in this area, re-analyzing the pre-existing datasets and integrating the kinds of the biological layer could be a promising approach to catch a comprehensive insight into the underlying mechanisms of complex disorders.[23, 24]

Accordingly in this study, a protein expression profile of individuals' colorectal tumor tissue compared to non-tumor tissue reported by Costanza and colleagues, was analyzed after validating the satisfying quality of the dataset; a set of differentially expressed proteins was detected to perform valuable further analysis.[6] Given the crucial role of integrative measures in examining complex disorders, an inclusive map comprised of DEPs, miRNAs, and lncRNAs was created to facilitate our understanding of the pathogenesis of CRC advancement, which could provide an opportunity to announce novel drug targets and remedial approaches. In the first step, various enrichment measures were carried out to overview the altered biomedical phenomena comprehensively. In accordance with the achieved findings, all results in these methods consistently revealed the critical roles of mRNA processing and immune response in the pathogenesis of CRC. Accordingly, a large body of evidence has indicated the pivotal function of mRNA processing, such as alternative splicing (AS), in CRC progression. It is also worth mentioning that AS events could be considered potential therapeutic targets for CRC. In this regard, the increased expression of COL6A3 as a survival element in CRC has been recently reported.[25] Remarkably, our results in network analysis revealed SF3B3 as the highest-scoring central protein known as the main splicing element in the mRNA splicing pathway, which has not yet been investigated in CRC advancement. However, a recent study has merely highlighted SF3B3 as an age-associated factor in CRC advancement.[26] Among all the identified hub proteins, other splicing elements

such as *HNRNPA1* and *SRSF2* were identified.[27] Accordingly, Shilo and colleagues indicated that SRSF2 is a well-known mutated protein in some sort of cancer. Moreover, it is shown that the alternative splicing in *SRSF2* is directly associated with tumor development;[28] however, detailed function of the *SRSF2* in CRC has not been previously shown. Further research should investigate splicing procedures, specifically hub proteins in CRC development, based on attained findings.

In line with previous studies, our results underscored the role of immune response pathways in CRC advancement.[29-31] Among the enriched related terms, a wide variety of pathways can be seen which are practically linked to innate immune systems. For instance, platelet degranulation, which is examined in a recent study that demonstrated that depending on the ambiance of a function such as a bloodstream or local, could be a promotive or prophylactic factor in CRC development.[32] Notably, notwithstanding the pathway as mentioned earlier, our result revealed the alteration of neutrophil and interleukin terms, which is generally considered a pioneer factor in the inflammation response against infectious state and homeostasis regulation. Moreover, inconsistent with the pivotal role of the immune response, such as announced terms, GO analysis indicated coherent findings based on neutrophil and platelet degranulation, which can describe the inflammatory nature of cancer and their critical role in the wound healing process. Strictly speaking, it has recently been considered that platelets play a predominant role in immune response regulation. More broadly, further research might be required to provide a comprehensive overview to introduce potential pertinent biomarkers.[33]

The proportion of enriched terms among all observations is related to scavenger receptor function that is directly linked to innate immune response. Accordingly, our results revealed scavenger receptor class A (SR-A1) as one of the essential pathways, which also accords with the earlier observations showing that how the suppression of SR-A1 can mitigate lung cancer metastasis.[34] This finding corroborates the ideas of Ryu and colleagues suggested that all sorts of scavenger receptors such as A to E are principally related to an array of cancer.[35] However, it seems that further research is required to better understand their roles, so that they can be considered both rational therapeutic targets and potential biomarkers.

To put it more simply, our functional analysis favors the modification of two main categories: mRNA processing and immune response and its proper functions such as neutrophil and platelet degranulation. Furthermore, scavenger receptors could point to metastasis and wound healing processes in CRC. The engagement of immune systems may provide a novel insight into the immunological aspect of CRC pathogenesis and should be focused on further investigation.

In the following step, based on the constructed multi-layer network, topological parameters were screened to identify the central molecules in each predicted layer to decipher the critical driver of pathogenesis in CRC. Hence, given the pivotal role of regulatory factors such as miRNAs and lncRNAs, an interacting multi-layer map could be a reasonable measure to decipher the complex nature of cellular regulatory processes. Despite the introduced hub DEPs, which showed that a considerable part of them is in line with our functional analysis, we recognized hub regulatory elements. Interestingly, the majority of central miRNAs have been previously examined as decisive elements either in CRC progression or metastasis and apoptotic function, *miR-15b-5p*, which has recently been introduced as a promotive factor for growth and apoptosis in breast and CRC.[36] Consequently, other identified hub-miRNAs, including *miR-16-5p*, *miR-424-5p*, *miR-497-5p*, and *let-7b-5p*, have been shown in the initial investigation that agrees with their influential role as a progressive factor in various sorts of cancer.[37, 38] Although it could be noted that some of the recognized hubs have not been directly investigated in CRC advancement, likewise *miR-1301-3p*. Therefore, since miRNAs could be an effective leader in certain kinds of specific biological processes, focusing on them can facilitate our understanding of underlying pathogenesis mechanisms, and also, provide potential biomarkers.

In line with our network findings, we identified *SLC9A3-AS1* and *LINC00943* as potential hub lncRNAs in CRC. *SLC9A3-AS1* is an antisense RNA 1 involved in several orders through the interaction with proteins and other regulatory factors including miRNAs in facilitating nasopharyngeal carcinoma.[39] Moreover, recent clinical evidence has strongly supported the role of *SLC9A3-AS1* as a circulating biomarker. In a cohort of 130 CRC patients versus 96 healthy controls, Xu and colleagues (2025) demonstrated that serum *SLC9A3-AS1* expression was significantly elevated, with high diagnostic accuracy in ROC curve analysis. Functional assays further confirmed that *SLC9A3-AS1* promotes proliferation, migration, and invasion of CRC cells by modulation of the *miR-486* axis, highlighting its potential as a clinically actionable

biomarker and therapeutic target.[40] This aligns with our computational predictions and emphasizes the translational relevance of *SLC9A3-AS1*.

Similarly, *LINC00943*, though has not yet been investigated in colorectal cancer, was reported to play a pivotal oncogenic role in gastric cancer. Xu and colleagues (2021) demonstrated that *LINC00943* regulates cancer cell proliferation and chemosensitivity by the *hsa-miR-101-3p* axis, thereby modulating tumor progression and therapeutic response. This evidence suggests that *LINC00943* may function as a critical regulator across multiple gastrointestinal malignancies. Taken together, these findings support the plausibility of our network-based prediction that *LINC00943* could act as a novel lncRNA biomarker in CRC.[41]

Moving on to the constructed interacting network, other predicted regulatory factors were also associated with lncRNAs, which are known to participate in pre- and post-translational regulation and may affect miRNAs. Among these, hub-lncRNAs such as *NUTM2B-AS1* and *MCM3AP-AS1* (*C1orf132*) have not yet been directly investigated in CRC progression, highlighting novel candidates for future experimental validation. Their potential involvement in tumorigenesis emphasizes the importance of expanding lncRNA-focused studies in CRC.

However, this study was limited by its relatively small sample size and lack of functional assays to directly assess the mechanistic role of *SLC9A3-AS1* in CRC. Therefore, further investigations incorporating larger cohorts and experimental validation of regulatory interactions are warranted to deepen our understanding of CRC pathogenesis and to evaluate the therapeutic potential of these lncRNAs.

Conclusion

Our analyses consistently highlighted two regarding the involvement of two main pathways, such as mRNA processing and immune response-related pathways. Additionally, based on a constructed multi-layer network, we introduced several hub molecules in regulatory layers, including *SLC9A3-AS1*, *LINC00943*, and *miR-1301-3p*, which may serve as promising candidates for future investigations. These molecules could provide diagnostic value and offer mechanistic insights into the underlying molecular basis of CRC.

Acknowledgment

The authors are deeply grateful to all patients who contributed to this study.

Authors' Contribution

N.ZJ: Conceptualization, designing the study, data acquisition, and analysis; and drafting the manuscript; A.MM: Conceptualization, designing the study, data acquisition, and analysis; and drafting the manuscript; N.SMK: Data acquisition and validation, and reviewing the manuscript; H.AT: Data interpretation, analysis, and reviewing the manuscript; F.N: Data acquisition and validation, and reviewing the manuscript; M.Y: Data interpretation, analysis, and reviewing the manuscript; H.S: Data interpretation, ensuring integrity of the work, and reviewing the manuscript; All authors approved the submission of the manuscript and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

AI Declaration

The authors declare that they have not used any AI tools or technologies to prepare this manuscript.

Conflict of Interest: None declared.

References

1 Rawla P, Sunkara T, Barsouk A. Epidemiology of colorectal cancer: incidence, mortality, survival, and risk factors. Prz Gastroenterol. 2019;14:89-103. doi: 10.5114/pg.2018.81072. PubMed PMID: 31616522; PubMed Central PMCID: PMC6791134.

2 Zhang X, Fan H, Han S, Zhang T, Sun Y, Yang L, et al. Global burden of colon and rectal cancer and attributable risk factors in 204 countries and territories from 1990 to 2021. BMC Gastroenterol. 2025;25:332. doi: 10.1186/s12876-025-03948-2. PubMed PMID: 40316922; PubMed Central PMCID: PMC12048922.

3 Zhang Y, Wei Y, Jiang B, Chen L, Bai H, Zhu X, et al. Scavenger Receptor A1 Prevents Metastasis of Non-Small Cell Lung Cancer via Suppression of Macrophage Serum Amyloid A1. Cancer Res. 2017;77:1586-98. doi: 10.1158/0008-5472.CAN-16-1569. PubMed PMID: 28202524.

4 Yang X. Multitissue Multiomics Systems Biology to Dissect Complex Diseases. Trends Mol Med. 2020;26:718-28. doi: 10.1016/j.molmed.2020.04.006. PubMed PMID: 32439301; PubMed Central PMCID: PMC7395877.

5 Yousefi M, Hashemi MS, Peymani M, Ghaedi K, Irani S, Etemadifar M. Elucidating the role of a shared lncRNA-miRNA-mRNA network in exacerbating Parkinson's disease symptoms in the context of COVID-19 infection. Gene Expr. 2024;23:83-97. doi: 10.14218/GE.2023.00103.
6 Costanza B, Turtoi A, Bellahcene A, Hirano T, Peulen O, Blomme A, et al. Innovative methodology for the identification of soluble biomarkers in fresh tissues. Oncotarget. 2018;9:10665-80. doi: 10.18632/oncotarget.24366. PubMed PMID: 29535834; PubMed Central PMCID: PMC5828218.
7 Prianichnikov N, Koch H, Koch S, Lubeck M, Heilig R, Brehmer S, et al. MaxQuant Software for Ion Mobility Enhanced Shotgun Proteomics. Mol Cell Proteomics. 2020;19:1058-69. doi: 10.1074/mcp.TIR119.001720. PubMed PMID: 32156793; PubMed Central PMCID: PMC7261821.
8 Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat Protoc. 2016;11:2301-19. doi: 10.1038/nprot.2016.136. PubMed PMID: 27809316.
9 Kim M, Tagkopoulos I. Data integration and predictive modeling methods for multi-omics datasets. Mol Omics. 2018;14:8-25. doi: 10.1039/c7mo00051k. PubMed PMID: 29725673.
10 Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat Methods. 2016;13:731-40. doi: 10.1038/nmeth.3901. PubMed PMID: 27348712.
11 Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498-504. doi: 10.1101/gr.1239303. PubMed PMID: 14597658; PubMed Central PMCID: PMC403769.
12 Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics. 2009;25:1091-3. doi: 10.1093/bioinformatics/btp101. PubMed PMID: 19237447; PubMed Central PMCID: PMC2666812.
13 Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 2016;44:W90-7. doi: 10.1093/nar/gkw377. PubMed PMID: 27141961; PubMed Central PMCID: PMC4987924.
14 Chang L, Zhou G, Soufan O, Xia J. miRNet 2.0: network-based visual analytics for miRNA functional analysis and systems biology. Nucleic Acids Res. 2020;48:W244-W51. doi: 10.1093/nar/gkaa467. PubMed PMID: 32484539; PubMed Central PMCID: PMC7319552.
15 Huang HY, Lin YC, Li J, Huang KY, Shrestha S, Hong HC, et al. miRTarBase 2020: updates to the experimentally validated microRNA-target interaction database. Nucleic Acids Res. 2020;48:D148-D54. doi: 10.1093/nar/gkz896. PubMed PMID: 31647101; PubMed Central PMCID: PMC7145596.
16 Bindea G, Galon J, Mlecnik B. CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. Bioinformatics. 2013;29:661-3. doi: 10.1093/bioinformatics/btt019. PubMed PMID: 23325622; PubMed Central PMCID: PMC3582273.
17 Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol. 2014;8 Suppl 4:S11. doi: 10.1186/1752-0509-8-S4-S11. PubMed PMID: 25521941; PubMed Central PMCID: PMC4290687.
18 Gao Y, Nan X, Shi X, Mu X, Liu B, Zhu H, et al. SREBP1 promotes the invasion of colorectal cancer accompanied upregulation of MMP7 expression and NF-kappaB pathway activation. BMC Cancer. 2019;19:685. doi: 10.1186/s12885-019-5904-x. PubMed PMID: 31299935; PubMed Central PMCID: PMC6626379.
19 Wan TM, Iyer DN, Ng L. Roles of microRNAs as non-invasive biomarker and therapeutic target in colorectal cancer. Histol Histopathol. 2020;35:225-37. doi: 10.14670/HH-18-171. PubMed PMID: 31617575.
20 Zhao C, Zhao Q, Zhang C, Wang G, Yao Y, Huang X, et al. miR-15b-5p resensitizes colon cancer cells to 5-fluorouracil by promoting apoptosis via the NF-kappaB/XIAP axis. Sci Rep. 2017;7:4194. doi: 10.1038/s41598-017-04172-z. PubMed PMID: 28646148; PubMed Central PMCID: PMC5482850.
21 Fang Y, Fullwood MJ. Roles, Functions, and Mechanisms of Long Non-coding RNAs in Cancer. Genomics Proteomics Bioinformatics. 2016;14:42-54. doi: 10.1016/j.gpb.2015.09.006. PubMed PMID: 26883671; PubMed Central PMCID: PMC4792843.
22 Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN

Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021;71:209-49. doi: 10.3322/caac.21660. PubMed PMID: 33538338.
23 Martens L, Vizcaino JA. A Golden Age for Working with Public Proteomics Data. Trends Biochem Sci. 2017;42:333-41. doi: 10.1016/j.tibs.2017.01.001. PubMed PMID: 28118949; PubMed Central PMCID: PMC5414595.
24 Zheng C, Zhang Z. New era of cancer immunology driven by big data. Med Rev (2021). 2023;3:449-51. doi: 10.1515/mr-2023-0070. PubMed PMID: 38282800; PubMed Central PMCID: PMC10808867.
25 Lian H, Wang A, Shen Y, Wang Q, Zhou Z, Zhang R, et al. Identification of novel alternative splicing isoform biomarkers and their association with overall survival in colorectal cancer. BMC Gastroenterol. 2020;20:171. doi: 10.1186/s12876-020-01288-x. PubMed PMID: 32503434; PubMed Central PMCID: PMC7275609.
26 Gong Y, Liu Y, Wang T, Li Z, Gao L, Chen H, et al. Age-Associated Proteomic Signatures and Potential Clinically Actionable Targets of Colorectal Cancer. Mol Cell Proteomics. 2021;20:100115. doi: 10.1016/j.mcpro.2021.100115. PubMed PMID: 34129943; PubMed Central PMCID: PMC8441843.
27 Lan Z, Yao X, Sun K, Li A, Liu S, Wang X. The Interaction Between lncRNA SNHG6 and hnRNPA1 Contributes to the Growth of Colorectal Cancer by Enhancing Aerobic Glycolysis Through the Regulation of Alternative Splicing of PKM. Front Oncol. 2020;10:363. doi: 10.3389/fonc.2020.00363. PubMed PMID: 32296635; PubMed Central PMCID: PMC7136466.
28 Shilo A, Siegfried Z, Karni R. The role of splicing factors in deregulation of alternative splicing during oncogenesis and tumor progression. Mol Cell Oncol. 2015;2:e970955. doi: 10.4161/23723548.2014.970955. PubMed PMID: 27308389; PubMed Central PMCID: PMC4905244.
29 Grivennikov SI, Greten FR, Karin M. Immunity, inflammation, and cancer. Cell. 2010;140:883-99. doi: 10.1016/j.cell.2010.01.025. PubMed PMID: 20303878; PubMed Central PMCID: PMC2866629.
30 Ganesh K, Stadler ZK, Cercek A, Mendelsohn RB, Shia J, Segal NH, et al. Immunotherapy in colorectal cancer: rationale, challenges and potential. Nat Rev Gastroenterol Hepatol. 2019;16:361-75. doi: 10.1038/s41575-019-0126-x. PubMed PMID: 30886395; PubMed Central PMCID: PMC7295073.
31 Brennan CA, Garrett WS. Gut Microbiota, Inflammation, and Colorectal Cancer. Annu Rev Microbiol. 2016;70:395-411. doi: 10.1146/annurev-micro-102215-095513. PubMed PMID: 27607555; PubMed Central PMCID: PMC5541233.
32 Plantureux L, Mege D, Crescence L, Carminita E, Robert S, Cointe S, et al. The Interaction of Platelets with Colorectal Cancer Cells Inhibits Tumor Growth but Promotes Metastasis. Cancer Res. 2020;80:291-303. doi: 10.1158/0008-5472.CAN-19-1181. PubMed PMID: 31727628.
33 Chen Y, Zhong H, Zhao Y, Luo X, Gao W. Role of platelet biomarkers in inflammatory response. Biomark Res. 2020;8:28. doi: 10.1186/s40364-020-00207-2. PubMed PMID: 32774856; PubMed Central PMCID: PMC7397646.
34 Ryu S, Howland A, Song B, Youn C, Song PI. Scavenger Receptor Class A to E Involved in Various Cancers. Chonnam Med J. 2020;56:1-5. doi: 10.4068/cmj.2020.56.1.1. PubMed PMID: 32021835; PubMed Central PMCID: PMC6976765.
35 Wu B, Liu G, Jin Y, Yang T, Zhang D, Ding L, et al. miR-15b-5p Promotes Growth and Metastasis in Breast Cancer by Targeting HPSE2. Front Oncol. 2020;10:108. doi: 10.3389/fonc.2020.00108. PubMed PMID: 32175269; PubMed Central PMCID: PMC7054484.
36 Chen W, Lin G, Yao Y, Chen J, Shui H, Yang Q, et al. MicroRNA hsa-let-7e-5p as a potential prognosis marker for rectal carcinoma with liver metastases. Oncol Lett. 2018;15:6913-24. doi: 10.3892/ol.2018.8181. PubMed PMID: 29731866; PubMed Central PMCID: PMC5920543.
37 Xu Y, Shen L, Li F, Yang J, Wan X, Ouyang M. microRNA-16-5p-containing exosomes derived from bone marrow-derived mesenchymal stem cells inhibit proliferation, migration, and invasion, while promoting apoptosis of colorectal cancer cells by downregulating ITGA2. J Cell Physiol. 2019;234:21380-94. doi: 10.1002/jcp.28747. PubMed PMID: 31102273.
38 De Falco V, Napolitano S, Esposito D, Guerrera LP, Ciardiello D, Formisano L, et al. Comprehensive Review on the Clinical Relevance of Long Non-Coding RNAs in Cutaneous Melanoma. Int J Mol Sci. 2021;22. doi: 10.3390/ijms22031166. PubMed PMID: 33503876; PubMed Central PMCID: PMC7865742.
39 Li J, Li D, Zhang X, Li C, Zhu F. Long noncoding RNA SLC9A3 AS1 increases E2F6 expression by sponging microRNA 486 5p

and thus facilitates the oncogenesis of nasopharyngeal carcinoma. Oncol Rep. 2021;46. doi: 10.3892/or.2021.8116. PubMed PMID: 34165171; PubMed Central PMCID: PMC8218295.

40 Xu X, Qing H, Jiang C, Zhao X, Wei J. Influence of the lncRNA SLC9A3-AS1 on colon cancer and the biological activities of colon cancer cells. Discov Oncol. 2025;16:358. doi: 10.1007/s12672-025-02134-w. PubMed PMID: 40106182; PubMed Central PMCID: PMC11923312.

41 Xu Y, Ji T, An N, Wang X, Zhang H, Xu F. LINC00943 is correlated with gastric cancer and regulates cancer cell proliferation and chemosensitivity via hsa-miR-101-3p. Int J Clin Oncol. 2021;26:1650-60. doi: 10.1007/s10147-021-01945-5. PubMed PMID: 34086111.