NJSC «al-Farabi Kazakh National University»

UDC: 620.3; 616.34-006(043) On manuscript rights

AKIMNIYAZOVA AIGUL NURLANKYZY

The interaction of miRNAs with mRNAs of candidate genes of gastrointestinal tract cancer

6D074000-Nanomaterials and nanotechnology

Dissertation for Doctor of
Philosophy (PhD) degree

Scientific consultants:

Ivashchenko A.T., Doctor of Biological Sciences, professor

Uversky V.N., Doctor of Physico-mathematical sciences, PhD, professor of University of South Florida (Tampa, USA)

Republic of Kazakhstan
Almaty, 2023

TABLE OF CONTENTS

	SYMBOLS AND ABBREVIATIONS	4
	INTRODUCTION	5
1	**LITERATURE REVIEW**	11
1.1	The modern ideas about carcinogenesis of GI tract	11
1.1.1	The genesis of esophageal cancer	12
1.1.2	The genesis of gastric cancer	13
1.1.3	The genesis of small intestinal cancer	14
1.1.4	The genesis of colorectal cancer	16
1.2	Markers and diagnostics methods of GI tract cancer	18
1.3	The molecular biology of miRNA nanomolecules: structural features, biogenesis, and functions in human cells	20
1.3.1	miRNA nanomolecules' origin and biogenesis	22
1.3.2	Nanomolecular miRNA-mediated regulation of gene expression	24
1.3.3	The role of nanomolecular miRNAs in carcinogenesis	25
1.3.4	The prediction of miRNA nanomolecules binding sites in mRNAs of candidate genes	27
2	**MATERIALS AND METHODS**	30
2.1	Materials	30
2.2	Methods	30
2.3	Programs for processing data of miRNA	33
3	**RESULTS AND DISCUSSION**	35
3.1	Creation of databases on genes associated with carcinogenesis of the GI tract	35
3.2	The characteristics of the miRNA nanomolecules interactions with mRNAs of genes associated with carcinogenesis of the GI tract	37
3.2.1	The characteristics of the miRNA nanomolecules interactions with mRNAs of genes associated with esophageal cancer	37
3.2.1.1	The characteristics of the miRNA nanomolecules interactions with mRNAs of genes associated with esophageal adenocarcinoma	37
3.2.1.2	The characteristics of the miRNA nanomolecules interactions with mRNAs of genes associated with esophageal squamous cell carcinoma	44
3.2.2	The characteristics of the miRNA nanomolecules interactions with mRNAs of genes associated with gastric cancer	50
3.2.3	The characteristics of the miRNA nanomolecules interactions with mRNAs of genes associated with small intestinal cancer	69

3.2.4 The characteristics of the miRNA nanomolecules interactions with mRNAs of genes associated with colorectal cancer 80
CONCLUSION 96
REFERENCES 100
SUPPLEMENTARY MATERIALS A 122
SUPPLEMENTARY MATERIALS B 123
SUPPLEMENTARY MATERIALS C 125

SYMBOLS AND ABBREVIATIONS

mRNA - messenger RNA
miRNA - mRNA inhibitory RNA
RNA – ribonucleic acid
RISC - RNA inducing silencing complex
GI tract cancer – gastrointestinal tract cancer
EC – esophageal cancer
EAC – esophageal adenocarcinoma
ESCC - esophageal squamous cell carcinoma
GC – gastric cancer
SIC – small intestinal cancer
CRC – colorectal cancer
5'UTR – 5'-untranslated region
CDS – coding domain sequences
3'UTR – 3'-untranslated region
NCBI – National Center for Biotechnology Information
PUBMED – Biochemical Literature Citation and Abstracts
nt – nucleotide
TM – Tumor markers

INTRODUCTION

General description of work. This work is devoted to the study of the characteristics of the interaction of miRNA nanomolecules with mRNA of genes implicated in gastrointestinal tract cancer development and the search for markers based on associations of miRNA and candidate genes for the development of methods for the early diagnosis of these diseases in nanomedicine.

Relevance of the research topic.

Cancer continues to pose a substantial threat to public health worldwide, manifesting as a considerable factor contributing to global mortality. The year 2020 witnessed an alarming surge with an estimated 19.3 million new cases of cancer diagnosed, a staggering number that resulted in nearly 10 million deaths across the globe, emphasizing the enduring impact and severity of this disease on populations worldwide [1]. Moreover, cancer's global escalation is anticipated to persist and grow further in its impact, with an estimated 21.7 million new cases and 13 million deaths projected by 2030 [1]. Experts of the International Agency for Research on Cancer (IARC) have recorded a large number of deaths from colorectal cancer (9.4%), gastric cancer (7.7%) and esophageal cancer (5.5%) [1]. Despite substantial advancements in comprehending cancer's origins and concerted endeavors aimed at developing and enhancing diagnostic tools and treatment modalities, the prevalence of gastrointestinal cancer and its associated fatalities remain notably elevated. This elevated burden persists across both developing and developed countries, underscoring the continued challenges in reducing morbidity and mortality rates attributed to gastrointestinal cancers worldwide [2].

Esophageal cancer stands as a prevalent form of gastrointestinal malignancy, characterized by a bleak prognosis. It ranks among the most frequently diagnosed cancers affecting the gastrointestinal tract, often presenting challenging prognostic outcomes. Esophageal cancer primarily manifests in two predominant forms: esophageal adenocarcinoma (EAC) and esophageal squamous cell carcinoma (ESCC) [3]. These two distinct types represent the prevalent classifications under which esophageal cancer is commonly identified, each exhibiting unique characteristics and implications for diagnosis and treatment [3]. The observation of esophageal malignancy in patients with esophageal motility disorders was noted in the literature many years ago and continues to this day [4, 5-10]. The development of ESCC is believed to be associated with food stasis in the dilated lumen. This leads to contamination by bacteria, which eventually leads to chemical irritation of the esophageal lining. These patients can develop chronic retention esophagitis, dysplasia, and eventually cancer [4,11-13]. Gastroesophageal reflux disease leads to development of Barrett's esophagus and progression to esophageal adenocarcinoma [13-14].

Gastric cancer (GC) ranks as the second most prevalent cause of cancer-related mortality on a global scale and remains a growing concern. The primary cause of gastric ulcers is attributed to the presence of *H.pylori* bacteria [15]. Notably, *H. pylori*

infection has a clear association with the incidence of gastric cancer [16]. Given that gastric ulcers constitute open wounds, they are highly susceptible to bacterial infection. This leads to DNA mutations and harm to the cells comprising the gastric mucosa [17-18]. Consequently, the damaged gastric tissue undergoes replacement by intestinal or fibrous tissue as part of the healing process.

Tumors of the small intestine account for 1–5% of all tumors of the gastrointestinal tract [19]. Colorectal cancer ranks among the frequently occurring cancers on a global scale. An increased risk of colorectal cancer has been found in patients with inflammatory bowel disease, especially with prolonged and extensive ulcerative colitis. Studies indicate that individuals affected by inflammatory gastrointestinal conditions face elevated susceptibility to colorectal cancer [20-21] and heightened overall mortality linked to gastrointestinal tract tumors, such as small intestine cancer [22]. Exploring novel biomarkers that encapsulate various facets of this condition holds promise in facilitating the choice of suitable treatments and improving the monitoring of patients diagnosed with gastrointestinal tract cancer. Such advancements in biomarker discovery offer potential avenues for enhanced therapeutic strategies and more effective surveillance protocols for individuals grappling with gastrointestinal tract malignancies.

Tissue-specific and circulating miRNA nanomolecules stand out among various biomarkers, demonstrating robust potential for diagnosing gastrointestinal tract cancer. Their specificity to tissues and presence in circulation have positioned them as promising candidates in the realm of gastrointestinal cancer diagnosis, showing considerable effectiveness in identifying this specific type of cancer [23]. The miRNA (**m**RNA-**i**nhibitory **RNA**) are nanosized non-coding RNA molecules having a length of 6–9 nm. MiRNAs play a pivotal role in modulating gene expression and exhibit remarkable stability when present in blood, plasma/serum, urine, and various other bodily fluids, encapsulated within exosomes. These small regulatory molecules are consistently detectable and endure in a resilient form within the circulation and diverse bodily fluids, particularly within exosomes, fostering their potential as reliable biomarkers for diagnostic purposes in various medical contexts. Hence, the conventional definition of microRNA appears insufficient as it fails to align with existing evidence regarding the presence of numerous miRNAs contained within exosomes measuring 30–80 nm in size. These miRNAs function as post-transcriptional regulators, exerting control over gene expression, and individual instances have demonstrated their ability to modulate the expression of multiple genes simultaneously. miRNAs bind to nanosized mRNA molecules and block the synthesis of nanosized proteins on ribosomes consisting of two nanosized large and small subunits. This nanotechnological process (nanobiotechnology) is often used in modeling protein synthesis *in vitro*. Since their discovery, miRNAs have emerged as crucial regulators of intricate nanobiological mechanisms linked to diverse oncological conditions, impacting the regulation of numerous genes simultaneously. Changes in miRNA expression have emerged as a significant characteristic of cancer,

where several miRNAs serve as either tumor suppressors or oncogenes, contributing to cancer development when dysregulated [24]. These miRNAs possess the potential to influence intricate cellular functions like cell growth, the cell cycle, apoptosis, and invasion. Given their pivotal role in cancer, the investigation of miRNAs in gastrointestinal tract cancer has been prompted [25]. These processes are the foundation of nanomedicine.

The objective of the research: To reveal the quantitative characteristics of the interaction of miRNA nanomolecules with mRNA of genes involved in the development of gastrointestinal tract cancer.

Research objectives:

1. Create a database for miRNA nanomolecules and genes involved in the development of gastrointestinal tract cancer.
2. To study the characteristics of the interaction of miRNA nanomolecules with mRNA genes involved in the development of esophageal cancer.
3. To study the characteristics of the interaction of miRNA nanomolecules with mRNA genes involved in the development of gastric cancer.
4. To study the characteristics of the interaction of miRNA nanomolecules with mRNA genes involved in the development of small intestinal cancer.
5. To study the characteristics of the interaction of miRNA nanomolecules with mRNA genes involved in the development of colorectal cancer.
6. To establish the associations of miRNA nanomolecules and mRNA genes involved in the development of gastrointestinal tract cancer.

Research objects: the nucleotide sequences of 6274 miRNA nanomolecules and human genes involved in the development of gastrointestinal tract cancer.

The subject of research: Characteristics of interactions between miRNA nanomolecules and mRNA of genes involved in the development of gastrointestinal tract cancer**.**

Research methods. Methods of computer analysis and search for miRNA nanomolecules binding sites in mRNA of genes based on hydrogen bonds modeling using the MirTarget program.

The scientific novelty of the research. The quantitative characteristics of the interaction of 6274 miRNA nanomolecules with mRNA of genes associated with the development of gastrointestinal tract cancer have been established for the first time. Associations of miRNA nanomolecules and their target genes have been proposed for the development of the basis for the diagnosis of these diseases. The binding sites of miRNA nanomolecules in mRNA with a free energy of interaction equal to or more than -130 kJ/mol and binding sites with a cluster organization were found. Fully complementary interactions of miRNA with 5'UTR, CDS and 3'UTR of mRNA genes involved in the development of gastrointestinal tract cancer were revealed. Associations of miRNA nanomolecules and candidate genes can be used for the diagnosis of human gastrointestinal tract cancer.

The theoretical significance of the work. The study's theoretical importance

lies in establishing the quantitative traits of how nanoscale miRNAs interact with mRNA genes associated with gastrointestinal tract cancer development. These findings pave the way for advancements in nanomedicine, offering a new scientific approach to enhance early cancer diagnosis, thereby enabling the identification of patients with diverse types and subtypes of gastrointestinal tract cancer.

The practical significance of the research. Finding and using associations of nanomolecule miRNA~target gene applies to other cancers. The proposed methodology will be important for the diagnosis of polygenic diseases when it is necessary to identify the genes most responsible for the development of the disease and selectively change their expression. The use of computational methods in the search for miRNA target genes allows thousands of times to accelerate and reduce the cost of finding effective associations of nanomolecule miRNA~target gene. At the same time, the method allows one to take into account the likelihood of side effects of a particular miRNA nanomolecules on all potentially expressed genes of the human genome. The applied computing technologies will make it possible to develop test systems in nanomedicine for the early diagnosis of gastrointestinal tract cancer. The developing test systems will not only allow early diagnosis of diseases but also offer target genes for the development of therapetic methods.

The main provisions for the defense:

The human miRNA nanomolecules database and the database on 372 candidate genes associated with the development of cancer of the esophagus, stomach, small intestine, and large intestine, created in the format for the MirTarget program, can be used to develop test systems for early diagnosis of these diseases.

The human genome contains candidate gastrointestinal cancer genes, having miRNA nanomolecules binding sites with high nucleotide complementarity in the 5'UTR, CDS and 3'UTR of their mRNA. The miRNA nanomolecules binding sites in the CDS of candidate genes have high phylogenetic conservatism.

Most candidate genes are targeted by one or more miRNA nanomolecules, the binding sites separate from each other in the mRNA. In the mRNAs of other candidate genes, miRNA nanomolecules binding sites are located overlapping nucleotide sequences, forming clusters of binding sites for two or more miRNA nanomolecules. The clusters of miRNA binding sites are characterized by their compact arrangement and competition between miRNA nanomolecules, since their binding sites are located with overlapping nucleotides.

The quantitative characteristics of the interaction of miRNA nanomolecules with mRNA of candidate target genes indicate a potentially high dependence of the expression of these genes on the concentration of miRNA nanomolecules and mRNA. They serve as the basis for the selection of miRNA nanomolecules and mRNA associations for the diagnosis of gastrointestinal tract cancer.

Associations of miRNA nanomolecules with target genes for the following localizations of gastrointestinal tract cancer have been proposed.

For esophageal adenocarcinoma - ID00792.3p-miR and *ALDH1A2*; ID01778.3p-

miR, ID02084.3p-miR and *ARID1A;* ID02052.5p-miR, ID02064.5p-miR and *ARID1A, BARX1;* ID00296.3p-miR, ID02294.5p-miR, ID01702.3p-miR and *ARID1A, IGFBP2;* ID03332.3p-miR and *AR, FOXF1, FOXP1, GDF7;* ID00961.3p-miR and *BARX1.* For esophageal squamous cell carcinoma - ID01041.5p-miR and *ELN, SIX4, EGR1;* ID02761.3p-miR and *COL27A1;* ID01804.3p-miR and *CCNE1, EGR1, ELN;* ID00049.5p-miR and *C2orf40;* ID00966.5p-miR and *EGR1;* ID00602.3p-miR and *BMP7;* ID02668.5p-miR and *LAMC2;* ID01641.3p-miR and *ADAMTS2.*

For gastric cancer - miR-3120-3p and *DNM3OS;* miR-3661 and *PPP2CA;* miR-4253 and *EPHB2;* miR-4273 and *ZNF717;* miR-1273 family and *ATM, CASP10, CDH1, CDH17, HPSE, IGF1, KRAS, MACC1, MDM2, PRKAA1, TP53;* miR-619-5p and *CASP10, CEACAM5, CLDN18, CREB1, ERBB3, F2R, HPSE, IL10, IRF1, LEP, MACC1, MTHFR, PDCD4, STAT3;* miR-1183 and *THBS1;* miR-4767 and *SIRT1;* ID02051.3p-miR and *E2F1;* ID01190.5p-miR and *PIK3CA;* ID01816.3p-miR and *TERT;* ID00798.3p-miR and *MAPK1*; ID00920.5p-miR and *OGG1;* ID03491.3p-miR and *TWIST1;* ID01338.5p-miR and *TERT;* ID01662.3p-miR and *MYC;* ID01546.5p-miR, ID01293.5p-miR and *MAPK1;* ID01458.5p-miR and *VEGFC.*

For small intestinal cancer - ID02093.5p-miR and *GNAS;* miR-1273f and *KRAS;* ID00061.3p-miR, ID02260.5p-miR, ID03064.3p-miR, ID00296.3p-miR, ID01641.3p-miR and *ASXL1;* ID02457.3p-miR and *EGFR;* ID02052.5p-miR, ID02294.5p-miR, ID01041.5p-miR, ID01804.3p-miR ID01702.3p-miR, ID02064.5p-miR, ID02084.3p-miR and *LRP1B;* ID01190.5p-miR, ID01702.3p-miR, ID01895.5p-miR, ID00966.5p-miR, ID01804.3p-miR and *PIK3CA.*

For colorectal cancer - miR-548 family and *CDC27, ONECUT2;* miR-619-5p and *AGMAT, ERBB3;* miR-5096 and *PPP2R1B;* miR-574-5p, ID00470.5p-miR and *CCND1, IGF1, PPP2R1B, SMAD4, XRCC1, ZEB1;* miR-466, ID00436.3p-miR, ID01030.3p-miR and *FLT1, GSK3B, IRS1, MACC1;* miR-6089 and *TGFB1;* ID00296.3p-miR and *FMNL3, GSK3B, TGFB1;* ID03332.3p-miR and *MTA1, SIRT1, TCF4, TGFBR1;* miR-4665-3p and *STAT3;* miR-6789-5p and *HIF1A;* ID01453.3p-miR, ID01099.5p-miR, ID01868.3p-miR, and *ONECUT2;* ID01003.3p-miR and *MTA1;* ID01492.3p-miR and *SMAD2,* ID00156.5p-miR and *MSH6;* ID03064.3p-miR and *SMAD7.*

Connection of the research with the scientific project.

The dissertation work was carried out within the framework of the project "Development of test-systems for early diagnosis of cardiovascular, oncological and neurodegenerative diseases based on miRNA associations and their target genes" №AP05132460 MES RK.

Approbation of the research. The materials of the dissertation work were reported and discussed:

- at the international conference "Biological markers in fundamental and clinical medicine", Prague, Czech Republic, October 31-November 02, 2018;
- at the international congress “Biotechnology: state of the art and perspectives”, Moscow, Russia, February 25-27, 2019;

- at the XX International Scientific and Practical Conference "Chemistry, Physics, Biology, Mathematics: Theoretical and Applied Research", Russia, February 4, 2019;

- at the international scientific conference of students and young scientists "Farabi Alemi", Al-Farabi KazNU, Almaty, Kazakhstan, April 8-9, 2019;

- at the international conference "9th Moscow Conference on Computational Molecular Biology MCCMB'19", Moscow, Russia, July 27-30, 2019;

- at the 12th international multiconference “Bioinformatics of Genome Regulation and Structure/Systems Biology (BGRS/SB-2020)”, Novosibirsk, Russia, July 6-10, 2020.

Publications and personal contributions of the author. The main content of the dissertation is reflected in 13 printed works, including 2 articles in the international journals Q1-Q2 with an impact factor, cited in the Scopus and Web of Knowledge databases; 5 articles from the list of the Committee for Quality Assurance in Education and Science; 6 theses in the materials of international conferences. The author independently conducted an analysis of literature data on the topic of research, experimental research, and analysis of research results, writing and design of a dissertation manuscript. A copyright certificate for the MirTarSeq program (miRNA target sequence - oligopeptides and oligonucleotides) was obtained (Supplementary materials A). The certificate of implementation of the research results in the course «"Omics" technologies in medicine and pharmacy» of the master's educational program 7M101 "Public Health" was obtained (Supplementary materials B).

The structure of the dissertation. The dissertation is presented on 149 pages and consists of the following sections: symbols and abbreviations, introduction, literature review, materials and methods, results and discussion, conclusion, list of references from 296 titles; contains 19 tables, 25 figures, and 16 supplementary materials.

1 Literature review

1.1 The modern ideas about carcinogenesis of GI tract

Cancer is a disease characterized by the formation and growth of abnormal cell populations that threaten human life by disrupting vital body functions. Although it is known that ultimately cancer is caused by uncontrolled cell proliferation, the main biological cause of the disease remains poorly understood [26]. The advancement of tumors and the transformation of healthy cells are associated with genetic modifications stemming from disruptions in typical cellular processes caused by genetic mutations inherited or exposure to carcinogens [27-29]. These transformed cells gain the capacity to persist in the body by modifying signaling pathways. These alterations confer self-sufficiency in growth signals, insensitivity to growth-inhibiting signals, resistance to apoptosis-inducing signals, the potential for uncontrolled proliferation, the ability to promote blood vessel formation (angiogenesis), and the capability to invade surrounding tissues and metastasize. [30, 31]. The ability for invasive growth and metastasis is also one of the characteristic features of malignant neoplasms [32, 33]. The acquisition of such properties by a tumor cell presupposes the presence in it of several molecular genetic changes and features that distinguish it from the tissue from which the primary tumor originated [34]. Changes in intracellular signaling pathways, specifically the TGFb signaling pathway [35] and HGF/c-Met [36], are affected by these alterations. The potential for cancer cells to metastasize is largely determined by the mechanisms that suppress apoptosis, cytoskeleton remodeling, matrix metalloprotease (MMP) activity, and the expression of growth factors and their receptors [37].

Molecular changes responsible for the acquisition of the above properties can be used as tumor markers. Tumor markers (TM) are such biological features of a tumor that can themselves influence the clinical course and predict the outcome of cancer, as well as allow tracing the response to therapeutic intervention [38]. By the nature and methods of their detection, TM can be divided into several types. Genetic (mutations, changes in the copy number of genes, nanomolecule mRNA expression); epigenetic (changes in the DNA methylation profile); proteomic (changes in the level and profile of protein expression); metabolic (changes in the level and spectrum of low molecular weight metabolites), synthesis profile [39-41] and miRNA nanomolecules level [42-44]. The majority of circulating TM, due to insufficient sensitivity and specificity, is suitable for early detection of a number of tumors only in high-risk groups with a higher incidence of this disease. Despite this, tumor markers are widely used in the clinic for assessing prognosis, assessing the surgical radicality, monitoring, and obtaining additional information [39-41].

In practice, proteins are most often used as TM, usually lipo- and glycoproteins produced by tumor cells or by normal tissues surrounding the tumor, which have a pathogenetic significance and an increased content in the biological fluids of patients [41, 45]. Pretreatment levels of most TM correlate with tumor size, lymph node

involvement, presence or absence of metastases, histological type, degree of tumor differentiation, etc [41].

Regardless of the type of specific TM, its use in clinical practice should provide a proven improvement in the standard of living of patients after surgery, as well as increase their overall survival [46]. In addition, the content of TM in the analyzed sample should be determined without significant labor and resource costs and be distinguished by high sensitivity and specificity. However, to date, not a single tumor-specific marker has been found, and their sensitivity is low. TM used in clinical practice is few and is applicable only for a limited number of malignant neoplasms and clinical conditions. In most cases, to form an adequate strategy for the treatment of TM, it is necessary to use it in conjunction with standard imaging and biopsy methods, taking into account the main clinical and pathological parameters. Despite extensive technological advances in laboratory methods and the expansion of the corresponding peer-reviewed scientific literature, biomarkers still do not find wide clinical use; only a small number of biomarkers are currently validated and used [45].

1.1.1 The genesis of esophageal cancer

Despite advances in medical, endoscopic, and surgical therapy, the incidence of esophageal cancer (EC) has increased over the past few decades, and approximately 604 100 cases are diagnosed and 544 076 incidences of death registered worldwide in 2020 [47]. EC, is the seventh most common cancer and the sixth, most common cause of cancer death in the world [48]. EC is a malignant tumor that develops from the cells of the inner membrane - the mucous layer of an organ and spreads into the thickness of the wall, growing as the tumor grows along with the organ. Due to the fact, that the width of the lumen of the esophagus is small, the growth of the tumor gradually overlaps it, causing difficulty in the passage of food through the esophagus. The asymptomatic nature in the initial stages of EC often results in a delayed clinical presentation, contributing to a grim prognosis and limited efficacy of available treatments [49]. Efforts to identify reliable diagnostic or prognostic markers for clinical use have so far been unsuccessful. Thus, there's a critical need to develop a novel non-invasive biomarker for the early detection and prognosis of EC. Adenocarcinoma and squamous cell carcinoma are the most prevalent types of EC [50, 51]. These two forms of EC typically arise in distinct segments of the esophagus and are linked to different genetic alterations.

Squamous cell carcinoma is cancer that forms in the thin, flat cells that line the inside of the esophagus [52]. This cancer most commonly occurs in the upper and middle esophagus but can occur anywhere in the esophagus [53]. Squamous cell carcinoma is often diagnosed at a later stage. Although there are many treatments available, such as surgery, chemotherapy, and radiation therapy, the prognosis is poor [54].

Adenocarcinoma is cancer that begins in glandular cells. The glandular cells of the lining of the esophagus produce and secrete fluids such as mucus.

Adenocarcinoma usually begins in the lower esophagus, near the stomach [55]. Barrett's esophagus serves as the metaplastic precursor of adenocarcinoma [56]. A continuing improvement of our comprehension regarding the molecular basis of this condition is necessary. Cancer has recently shown altered miRNA expression to be an important characteristic [57]. The miRNA molecules are a diverse set of small, non-coding RNA sequences that exert negative regulation over protein-coding genes, demonstrating associations with a wide array of physiological and pathological processes, notably cancer [58, 59]. MiRNAs can influence cancer pathogenesis as oncogenes or tumor suppressors. Numerous studies have examined the relevance of miRNAs for EC due to recent observations about their functions and role in cancer.

1.1.2 The genesis of gastric cancer

Gastric cancer (GC), commonly known as stomach cancer, represents a significant global health challenge, contributing to over a million new cases and approximately 769 000 deaths in 2020. It ranks fifth in terms of incidence and fourth in mortality worldwide [60]. This cancer encompasses various histological types, including adenocarcinoma, lymphoma, carcinoid tumor, and stromal tumor, originating from different parts of the stomach. Adenocarcinoma, emerging from the stomach's mucous membrane, constitutes approximately 95% of gastric cancers [61]. Mucosal-associated lymphoid lymphoma (MALT) represents around 4% of gastric cancers, while carcinoid tumors from hormone-producing cells and stromal tumors in the gastrointestinal tract account for 3% and a lesser percentage, respectively [62]. Stromal tumors, originating from interstitial cells of Cajal within the stomach wall, are the least common. Gastric cancer is categorized as proximal or distal, depending on its location. Proximal gastric cancer is located near the gastroesophageal junction, while distal gastric cancer is situated in the lower stomach region [63]. There are two primary classification systems for GC: the Ming classification, which categorizes based on growth patterns, and the Lauren classification, which relies on dominant histological patterns. Among these, Lauren's classification is widely accepted due to its effectiveness in assessing the natural progression of GC [64]. According to Lauren's classification, GC is divided into two subtypes: the intestinal type and the diffuse type. The intestinal type is a well-differentiated tumor characterized by organized neoplastic cells forming glandular tubular structures. In contrast, the diffuse type is poorly differentiated, featuring individual cells that infiltrate and thicken the stomach wall [65]. These subtypes exhibit distinct morphological characteristics, pathogenesis, and genetic profiles. There have been documented cases of gastric cancer that do not clearly align with either histological type and instead display a mixed or intermediate pattern, combining features of both the intestinal and diffuse types [65]. Diffuse-type adenocarcinomas have a poorer prognosis compared to intestinal-type adenocarcinomas [64], which are primarily caused by an infection with *Helicobacter pylori*. In recent decades, there has been a decrease in the incidence of intestinal adenocarcinoma of the stomach, which could be linked to

improvements in the socio-economic situation, sanitation, food preservation, and a reduction in the occurrence of *H. pylori* [66, 67].

Numerous studies have demonstrated the correlation between the development of GC and *H. pylori* infection. Individuals infected with this bacterium are 2.5 times more likely to be diagnosed with GC [68, 69]. Since 1994, the International Agency for Research on Cancer has classified *H. pylori* as a Group I carcinogen due to its ability to cause mucosal atrophy, intestinal metaplasia, and epithelial dysplasia, which contribute to the development of cancer [68, 69]. However, it is worth noting that the proximal type of GC is not associated with *H. pylori* infection as stated in reference [67]. Eliminating *H. pylori* halts the formation of GC [67]. Nevertheless, after intestinal metaplasia emerges, *H. pylori* disruption only decelerates GC progression.

Alternatively, the gastric mucosa can host various microorganisms such as streptococci, staphylococci, micrococci, and *Candida*, all of which are implicated in the development of GC [70]. Both gastritis and *H. pylori* infection are acknowledged as precursors to the onset of GC [70]. However, relying solely on these indicators is insufficient for the early-stage diagnosis of GC in individuals at risk, as there is currently a lack of an effective tumor marker for this purpose.

Genomic classification of GC. At the genetic level, gastric cancer driver mutations have been identified in *TP53, PIK3CA, ARID1A,* and *RHOA* genes [71-74]. *RHOA* mutations are largely associated with diffuse-type gastric cancer. Genes regulating receptor tyrosine kinase or RAS signaling cascades are also frequently amplified in gastric cancer (for example, *ERBB2, FGFR2, MET, KRAS*), suggesting targetable opportunities [75], and anti-ERBB2 therapies are considered standard of care for ERBB2-positive gastric cancer. However, although clinical responses to fibroblast growth factor receptor (*FGFR*) inhibitors have been reported in *FGFR2*-amplified tumors [76], results from phase II/III trials evaluating receptor-tyrosine kinase-directed therapies in gastric cancer have been disappointing [77]. Such results suggest the existence of additional molecular heterogeneity in gastric cancer and the need for additional biomarkers.

1.1.3 The genesis of small intestinal cancer

The small intestine, the longest part of the digestive tract, is divided into the duodenum, the jejunum, and the ileum. In the small intestine, the main processes of absorption of nutrients take place [78]. Here, the chemical processing of food continues, the absorption of products of its breakdown. The endocrine function of the small intestine is important: the production by enteroendocrine cells (intestinal endocrinocytes) of biologically active substances (cholecystokinin, secretin, serotonin, enteroglucagon, gastrin, cholecystokinin, etc) [78]. Approximately 64% of all tumors of the small intestine are malignant, and approximately 40% of these tumors are adenocarcinomas [79]. Epidemiologically, small intestinal adenocarcinomas are strikingly similar to colon adenocarcinomas. For example, although adenocarcinomas of the small intestine are only 50 times less common than

adenocarcinomas of the colon, they have a similar geographical distribution with a predominance in Western countries [80]. In addition, they tend to occur in the same people with an increased risk of small intestinal adenocarcinoma in colorectal cancer survivors and vice versa [80].

Similar to colon adenocarcinomas, tumors in the small intestine typically arise from precancerous adenomas. These adenomas can develop sporadically or in the context of familial adenomatous polyposis [81]. As genetic mutations accumulate within these adenomas, they progress from a dysplastic stage to carcinoma and eventually evolve into invasive adenocarcinoma. Subsequently, these tumors have the potential to metastasize via the lymphatic vessels or the portal circulatory system to distant sites in the body, such as the liver, lungs, bones, and brain [81].

Small intestinal adenocarcinomas share similarities with colon cancer but tend to be distributed throughout the entire length of the small intestine. Research shows that approximately 55.7% of these tumors are located in the duodenum [82], which is the segment of the small intestine that first comes into contact with ingested chemicals and secretions from the bile ducts of the pancreas. This suggests that substances such as ingested chemicals and pancreatic secretions may be potential carcinogens, and also underlies the high incidence of duodenal cancer [78]. Animal studies suggest that bile may be carcinogenic, as evidenced by the reduction in the prevalence of experimentally induced small intestinal cancer through bile diversion [83].

Genetic analysis of sporadic small bowel adenocarcinomas unveils both shared features and dissimilarities in comparison to the pathogenesis of colorectal carcinoma. While mutations in the K-ras gene and the prevalent overexpression of p53 appear to be similarly common in both small intestinal and colorectal adenocarcinomas, distinct differences are noted. Notably, mutations in the *APC* tumor suppressor gene, typical of colorectal carcinoma, are not typically observed in the small intestine [84]. Additionally, the *SMAD4/DPC4* gene, commonly mutated in pancreatic and colorectal carcinomas, exhibits inactivity in small intestinal adenocarcinomas [84].

Sarcomas make up around 15% of malignant tumors in the small intestine [85]. While certain tumors display distinct histological markers of smooth muscle origination, many tumors exhibit incomplete expression of muscle-associated antigens [86]. Because they are mesenchymal neoplasms, GISTs are thought to be of Cajal's interstitial cell origin in the digestive tract [87]. Recent research has revealed that almost all GISTs express the c-kit proto-oncogene's growth factor receptor, which has tyrosine kinase activity, distinguishing them from true sarcomas [88]. Mutations occurring in c-kit lead to continuous activation of tyrosine kinase, triggering uncontrolled cell growth and proliferation. These mutations are detected in approximately 60% of GISTs and are considered crucial in the process of carcinogenesis [89].

Although most of the GIST found in the stomach, 30% of the GIST is found in the small intestine. These tumors are more evenly distributed in the small intestine than adenocarcinomas and tend to grow extra luminal [90]. Intestinal bleeding is a

common symptom because it is vascular lesions that usually ulcers. Compared to GISTs of the stomach, those of the small intestine tend to be more aggressive with a poorer prognosis [91]. Metastases frequently occur hematogenously, primarily affecting the liver and lungs [92]. Moreover, GISTs can penetrate directly into adjacent organs or spread through peritoneal dissemination. Although rare, lymphatic metastases are an indication of more widespread metastatic disease [92].

1.1.4 The genesis of colorectal cancer

Certainly, colorectal cancer is a significant health concern globally. In 2020 alone, it was estimated that more than 1.9 million new cases of colorectal cancer emerged, resulting in approximately 935,000 deaths. This cancer type accounts for around one in 10 cancer cases and deaths [93]. Despite being the third most prevalent cancer in terms of incidence, colorectal cancer holds the second position in mortality rates. Its prevalence continues to increase, making it a prominent component of contemporary oncological diseases, with a rising number of new diagnoses each year [94]. This oncological disease is characterized not only by different clinical manifestations due to tumor localization in the colorectum, molecular genetic heterogeneity of carcinomas but also by frequent relapses, metastases, and low 5-year survival rates [95]. Every second case of diagnosed cancer ends in death, but the duration of the disease, quality of life, asymptomatic manifestations have wide variability depending on the stage of detection and biological characteristics of the tumor [96-97]. The likelihood of developing CRC increases markedly after age 50, with 90% of new cases and 94% of deaths associated with CRC occurring over the age of 50 [98].

Colorectal cancer is most often divided into 2 forms: colon cancer and rectal cancer. Colorectal cancer is a rather heterogeneous group of tumors, which manifests itself both in clinical behavior and in the pathogenesis of the development of this malignant tumor [99].

In modern literature, based on the results of molecular genetic studies, there are 3 main types of carcinogenesis of colorectal cancer: 1) The transition from adenoma to carcinoma [100]; 2) hereditary non-polyposis colorectal cancer (HNPCC-type) (Lynch syndrome) [101]; 3) colitis-associated neoplasia [102].

According to the modern classification, hereditary forms are usually divided into polyposis and non-polyposis, depending on the number of polyps in the colon. Non-polyposis syndrome includes hereditary non-polyposis cancer, which is also known as "Lynch syndrome". This syndrome is characterized by an autosomal dominant mode of inheritance and is associated with an embryonic mutation in the DNA repair gene, including abnormalities in *MLH1, PMS1, PMS2, MSH2, MSH3,* and *MSH6* [103].

Allocate hyperplastic and adenomatous polyps. Adenomatous polyps characterized by glandular origin, they are tubular or papillary growths with stroma, high mitotic activity, and partial or complete loss of the ability to differentiate, which

leads to malignancy [104]. Polyposis syndromes include adenomatous polyposis (familial adenomatous polyposis (FAP), MUTYH-associated polyposis, polymerase-associated polyposis) and hamartomatous polyposis syndrome (juvenile polyposis syndrome, Pates-Yegers syndrome, PTEN hamartoma tumor syndrome) and hyperplastic polyposis [105, 106].

Another path of pathogenesis is based on the malignancy of the cells of the intestinal mucosa when affected by chronic colitis. For sporadic forms of colorectal cancer and forms associated with ulcerative colitis, an increase in the number of mutations leading to carcinogenesis is characteristic. Thus, it has been shown that the *KRAS* oncogene and tumor suppressor genes *APC*, *SMAD4,* and *TP53* are involved in the development of sporadic forms [107, 108]. It has been shown that the pathogenesis of CRC forms associated with chronic colitis is characterized by an orderly staged development from inflammation and hyperplastic processes to dysplasia, which subsequently transform into adenocarcinomas [109].

Microsatellite and chromosomal instability play a role in the pathogenesis of colorectal cancer. Sporadic forms of colorectal cancer are mostly caused by chromosomal instabilities, resulting in the loss of heterozygosity and the inactivation of the *APC, TP53, DCC*, and *KRAS* genes [110, 111].

A rather low percentage of tumors of the colorectum are congenital and result from mutations in the *APC* gene and the *MUTYH* gene. Such variants include familial adenomatous polyposis and MUTYH-associated polyposis, which occur in 2% of colorectal cancers [112]. According to the literature, microsatellite instability occurs and plays an important role in the development of cancer of both the right and left colon [113].

Particular attention in modern science is given to the molecular aspects of the development of colorectal cancer. Disruption in the system of suppressor genes, oncogenes, and autophagy regulator genes leads to the development of this pathology. Joint mutation in the *KRAS* gene (Kirstenrat sarcoma viral oncogene homolog) together with a disturbance in the *BRAF* gene (Rafmurine sarcoma viral oncogene homolog B1) led to an increase in the incidence of cancer in patients [114, 115].

Various signaling pathways can play a role in the development of the oncological process: MAPK/ERK and AKT/m-TOR, etc. [116, 117]. Each pathway contributes to the processes of survival, adaptation, and activation of cellular components and molecules. The mitogen-activated protein kinases (MAPK)cascade plays a significant role in the survival, spread, and resistance to chemotherapy of human cancer cells [117]. The MAPK/ERK pathway receives signals from many internal factors, such as abnormalities in DNA, metabolic stress, changes in protein concentration [118].

Mutations in the genes encoding this pathway can lead to a change in the protein concentration in the tumor environment and thereby activate tumor proliferation [119, 120]. According to the literature, these mutations can occur in the genes of membrane receptors (*EGFR* receptor for epithelial growth factor) [121, 122] in the genes of

signal transduction (*RAS*) [123, 124], and other kinases that play a role in the regulation of the MAPK/ERK pathway [125].

1.2 Markers and diagnostics methods of GI tract cancer

Numerous diagnostic methods are essential for identifying the tumor process, determining its stage, and selecting treatment tactics for patients with oncological diseases. In certain cases, multiple studies utilizing different methods are necessary for differential diagnosis. Firstly, understanding the principles, capabilities, and limitations of each method is crucial for ensuring effective diagnosis and treatment of cancer patients. Personal patient characteristics and disease progression must also be considered when selecting diagnostic methods and research tactics [38, 41]. Therefore, the choice of diagnostic method and research tactics is a fundamental component of cancer patient treatment. Analyzing the results and determining the presence of a tumor can provide insight into the type of tumor, the stage of the tumor process, and the involvement of adjacent anatomical structures in the affected organ's pathological process.Recently, practical oncology has employed biochemical markers of tumor growth due to advancements in studying the molecular biological and biochemical processes of tumor formation. [38, 39, 41].

The group of molecular markers used in the diagnosis of tumors can be divided into biochemical, immunohistochemical, and molecular genetic [126].

Tumor markers are chemicals produced by tumor cells that can be found in the blood. However, normal cells in the body can also trigger tumor markers to increase their levels, even in non-cancerous cases. This hinders the effectiveness of cancer marker tests for diagnosing cancer and such tests are only reliable in rare cases. [39, 127].

1. Cancer-embryonic antigen (CEA) is a tumor marker, first isolated in 1965 by P. Gold and S. Freedman from fetal cells and cancer cells of the colon [128]. Currently, its role in adults is unknown. It is understood that CEA molecules are actively produced by gastrointestinal epithelial cells during child development and indirectly regulate cell division [129, 130].

A blood test for CEA is employed both for cancer diagnosis and to distinguish between benign and malignant neoplasms. An increase in CEA levels to critical levels indicates the presence of malignant neoplasms in the stomach, colon, and rectum due to carcinogenesis [131, 132].

However, CEA levels may also indicate the presence of benign tumors or inflammation in the body and do not necessarily indicate cancer. Elevated CEA levels can occasionally be detected in common acute respiratory illnesses and frequently during pregnancy through a blood test for tumor markers [133]. Smokers tend to have slightly higher levels of CEA compared to non-smokers [134]. The CEA marker of intestinal tumors is found in high concentrations in several autoimmune processes, such as rheumatoid arthritis [135]. Thus, elevated levels of CEA do not provide definitive evidence of malignant neoplasms. Most likely, this method serves to predict the likelihood of malignant neoplasm development and requires further

research and analysis. Therefore, it is frequently employed to ascertain the possibility of their manifestation, rather than for diagnostic purposes [135].

Quantitative analysis of CEA is primarily recommended for high-risk patients predisposed to oncologic diseases. They are also prescribed to determine the localization of cancer in the presence of precancerous pathology and a previously established diagnosis [136]. Precise interpretation of CEA levels aids in oncologic diagnosis and the successful management of cancer patients. Subjectivity is excluded from the evaluation of CEA levels. Levels within the normal range don't generate reasons for concern, while those above 10 ng/mL indicate extensive disease and over 20 ng/mL, metastatic disease [137].

2. CA 19-9 was first discovered in 1979 by Koprowski and his colleagues through the use of monoclonal antibodies to isolate tumor-associated antigens in colorectal carcinoma. In 1981, it was discovered that pancreatic cancer cells also produce CA 19-9 [138]. CA 19-9 is a glycoprotein that is secreted by developing neoplasm cells and is associated with Lewis antigens located on the surface of red blood cells [139]. This protein is found in the body of all healthy individuals. Its significant increase indicates the development of pathologies in the esophagus, stomach, colon, gallbladder, and rectum [140-144]. Analyzing the CA 19-9 cancer antigen is prescribed to diagnose pathological conditions in these organs. It reveals malignant and benign neoplasms, inflammatory processes, and the effects of drug therapy [145]. Therefore, the CA 19-9 biomarker is utilized in conjunction with CEA and CA 242 for diagnosing and monitoring the development of cancer in the gastrointestinal organs [141].

3. The CA 242 antigen is synthesized in large quantities by the mucosal epithelium lining the digestive organs: stomach, colorectum, and the pancreas [146, 147]. It contains oligosaccharide structures present on highly glycosylated high molecular weight mucins. However, its chemical structure has not yet been fully identified, although the antigen was discovered in the early 1990s [148]. The CA 242 tumor marker presents a distinguishing trait by revealing the onset of malignant neoplasms within the gastrointestinal tract's organs at early stages of the disease's development [147]. Elevated levels during therapy signify either inefficacy of the employed strategy or a probable disease recurrence. Thus, CA 242 can predict a relapse six months prior to the first clinical signs [147]. When CA 242 rates are high, a combined study with CEA is necessary. This allows early detection of colon and rectal carcinogenesis. A marked increase in both indicators, along with ultrasound diagnostics, is enough to diagnose rectal cancer [39, 127]. This tumor marker shows greater sensitivity than other gastrointestinal tract tumor markers. Several studies suggest that utilizing CEA and CA242 in combination has high specificity and sensitivity. Attalah et al. found that combining both markers increased the overall sensitivity to 72% [149, 150].

4. Gastric tumor markers include the CA 72-4 tumor antigen. Its concentration is determined in the diagnosis of gastric, pancreatic, and colon cancers [151, 152]. A

joint test of CA 72-4 level with such tumor markers as CA 125, CA 19-9 and CEA is of diagnostic value. Its level also increases in breast fibroadenoma, benign ovarian neoplasms, as well as in pregnant women [152]. CA 72-4 is used in combination with CEA and CA 19-9 for monitoring gastric cancer. The sensitivity of the CA 72-4 test for gastric cancer sensitivity of 33.3% and specificity of 92.8%. The level of increase correlates with the staging of the process; after removal of the tumor, it returns to normal in 3-4 weeks [151].

5. Alpha-fetoprotein (AFP) is a transport protein that exhibits an affinity for various nanoscale substances, including bilirubin, fatty acids, and steroid hormones [153]. Typically, it is present in the blood at low concentrations across all ages, serving a protective function by aiding in the release of enzymes that combat malignant tumors [154-156]. During pregnancy in women, the levels of this glycoprotein tend to rise.

Most biochemical tumor markers for gastric cancer detected in clinical practice are found in the blood, namely CEA, CA-19-9, and CA-72-4, rather than in the gastric mucosa. While this makes material collection for research more straightforward, it is known that these markers have low sensitivity and specificity [157, 158] and can be found in high concentrations in the blood for benign stomach neoplasms, inflammatory diseases, or tumors in other locations.

The best way to use tumor markers in diagnosing cancer has not been determined and the use of some tumor marker tests is controversial.

1.3 The molecular biology of miRNA nanomolecules: structural features, biogenesis, and functions in human cells

Biological nanostructures encompass a variety of molecules like proteins, nucleic acids (DNA, RNA), and polysaccharides. These constituents are integral to building an intracellular cytoskeleton, extracellular matrix, membrane channels, receptors, and transporters. They also contribute to intracellular signaling systems, molecular machineries for synthesizing, packaging, and utilizing proteins and nucleic acids, energy production, intracellular transport, and cell movement. Proteins and supramolecular complexes typically range in size from 1 to 100 nm, while a DNA helix has a 2 nm diameter that can extend several centimeters in length. Cytoskeleton filaments form complexes 7–25 nm thick and spanning several microns in length. Pore-forming protein complexes can reach diameters of 120 nm. Extracellular structures, such as exosomes facilitating material exchange between cells, exhibit sizes from 30–80 nm, and plasma lipoprotein particles responsible for lipid transport measure 8–25 nm in diameter (Table 1). Viruses represent the sole nanoscale form of living matter, varying in size from 25–300 nm.

Table 1 - The size of different biological nanoobjects.

Object	Size (nm)
Antibody	10
Coronavirus	40-60
Cell membrane	10-12
Liposome	From 30
Exosomes	30-80
Hemoglobin	5
low density lipoproteins (LDL)	17–25
high density lipoproteins (HDL)	8-12
miRNA	6-9

The miRNA nanomolecules was first described in 1993 by V. Ambros, R. Lee, and R. Feinbaum [159]. They worked on the task to find out why some *Caenorhabditis elegans* roundworms cannot properly pass the larval stage. The research revealed that a decrease in quantities of a specific protein, lin-14, was crucial for the worms to progress from the initial larval period to the second. A thorough analysis of the mRNA encoding the lin-14 protein has revealed the presence of complementary repeats to the *lin-4* gene sequence towards the end of the lin-14 mRNA [159]. The research further demonstrated that lin-4 produces small RNAs that are not translated into proteins [159]. The researchers hypothesized that the mRNA of *lin-4* binds to and down-regulates the mRNA of lin-14. The interaction between lin-4 and lin-14 was the first report that miRNA lin-4 regulates target mRNA lin-14 [159, 160]. However, the critical roles of mRNA were obscure until the detection of let-7, another miRNA nanomolecule found in numerous organisms [161]. This led to the discovery of a substantial group of small RNAs, comparable to those found in *C. elegans, Drosophila melanogaster,* and humans. Subsequent to this, the study of miRNA has become a significant focus in biology and medicine. Research indicates that miRNA nanomolecules exist and play a role in all biological and metabolic processes in eukaryotes [161].

In 2001, it was demonstrated that these molecules are a part of a sizable group of non-coding RNAs that pertain to proteins. Mature miRNA nanomolecules have a conservative structure and consist of 16-27 nucleotides [162, 163].

From a physical dimension's perspective, miRNAs are nanoscale, non-coding RNA molecules, having a contour length of 5–9 nm (Table 1) [164]. The miRNAs can be encoded in any region of the genome. Most of the miRNA nanomolecules genes (61%) are located in the intron regions of the genes; however, miRNA genes can be located in exons or intergenic regions. The miRNAs are transcribed from genes called miRNA genes, RNA polymerase II, some are used by polymerase III [159].

The miRNA nanomolecules are present in numerous animals and plants. The most recent version of the annotated miRNA database miRBase

(http://www.mirbase.org) includes data on over 25,000 potential miRNAs found in 271 biological species, among them 2,400 human miRNAs, with roughly 1,000 of these having been validated with a high level of reliability [165, 166]. The study investigated the structure features of miRNA nanomolecules and their genes, as well as the mechanisms involved in their expression and maturation [165, 166]. The functions of miRNA nanomolecules in the cell and the associated spectrum of protein factors were also determined. Currently, researchers worldwide are examining the structural characteristics of miRNA nanomolecules and their genes, alongside assessing the mechanisms behind their expression and maturation. Additionally, scientists are scrutinizing the main roles of miRNA nanomolecules in cells, as well as exploring the range of protein factors connected to them [165, 166].

1.3.1 miRNA nanomolecules' origin and biogenesis

Among the classes of short regulatory RNAs, including small nucleolar RNA (snRNA), piwi-interacting RNA (piRNA), endogenous small interfering RNA (endo-siRNA), short hairpin RNA (shRNA), and transcriptional initiation RNA (tiRNA), miRNA stands out due to distinct structural characteristics in both its mature molecule and precursor forms, as well as its specific maturation process [167-173]. Initially, the products of transcriptional processing of miRNA nanomolecule-encoding genes were considered the earliest identified mRNAs [159, 160]. miRNA nanomolecules are not solely derived from genes encoding miRNAs; up to half of mammalian miRNAs originate from RNA transcripts of other genes [174]. Introns primarily serve as the source of "intragenic" miRNAs, but miRNA genes can also be computationally predicted within exons and regulatory regions of genes, such as the 5'- and 3'-untranslated regions [175]. Moreover, other RNA types, like snRNA, tRNA mobile elements within the genome, and products of processed pseudogenes, can also act as sources for miRNA nanomolecules [176-181].

Many miRNA nanomolecules are highly conserved, and their homologs are found in species that are phylogenetically distant from each other [182]. Evolutionarily the most rigidly fixed is the sequence of functional regions of miRNA nanomolecules [182, 183]. Regulatory regions of mRNA containing miRNA nanomolecules binding sites are also under evolutionary conservation [183, 184]. The presence in most animals of the homologs of the main proteins involved in the maturation and functioning of miRNA nanomolecules suggests a common phylogenetic origin of miRNA-mediated regulation in animals [185]. Maturation of miRNA is a series of sequential cleavages of immature precursor molecules [186]. Nowadays the types of miRNA nanomolecules maturation divided into the canonical pathway and several exceptions that form a heterogeneous group of non-canonical pathways for miRNA [187].

The biogenesis of miRNA nanomolecules is a complex and sequential process involving several enzymes, in contrast to tRNA biogenesis. Initially, a large RNA is transcribed from the gene or intron, consisting of hundreds to thousands of

nucleotides. This RNA contains a hairpin, which is a double-stranded RNA section with partially unpaired nucleotides in the middle. The miRNA nanomolecules of the future are located within this hairpin [188].

Proteins called Pasha and Drosha are present in the nucleus for the production of miRNA nanomolecules. These proteins form a complex that processes miRNA by removing unnecessary segments [188]. After successful processing, a transport complex sends a small pre-miRNA (60-90 nucleotides) to the cytoplasm, near where proteins are produced with exportin-5/Ran GTPase [189]. The protein Dicer removes any remaining fragments, resulting in mature miRNA nanomolecules. The outcome is a duplex composed of two 22-nucleotide miRNA strands (miRNA/miRNA) [189].

One of the two chains is loaded into the RNA-induced silencing complex (RISC), specifically the RNA inducible miRNA-off gene ribonucleoprotein complex, where the miRNA and its target mRNA interact [189].

The Argonaut (Ago) proteins are crucial to the functioning of the RISC complex, with four domains including a PAZ domain that interacts with a site on the 3' end of miRNAs. Together, they bind mature miRNA nanomolecules and orient them suitably for interacting with the mRNA target [189].

Turning off a gene can be achieved through mRNA degradation at full complementarity or by preventing its translation if there is complete complementarity.

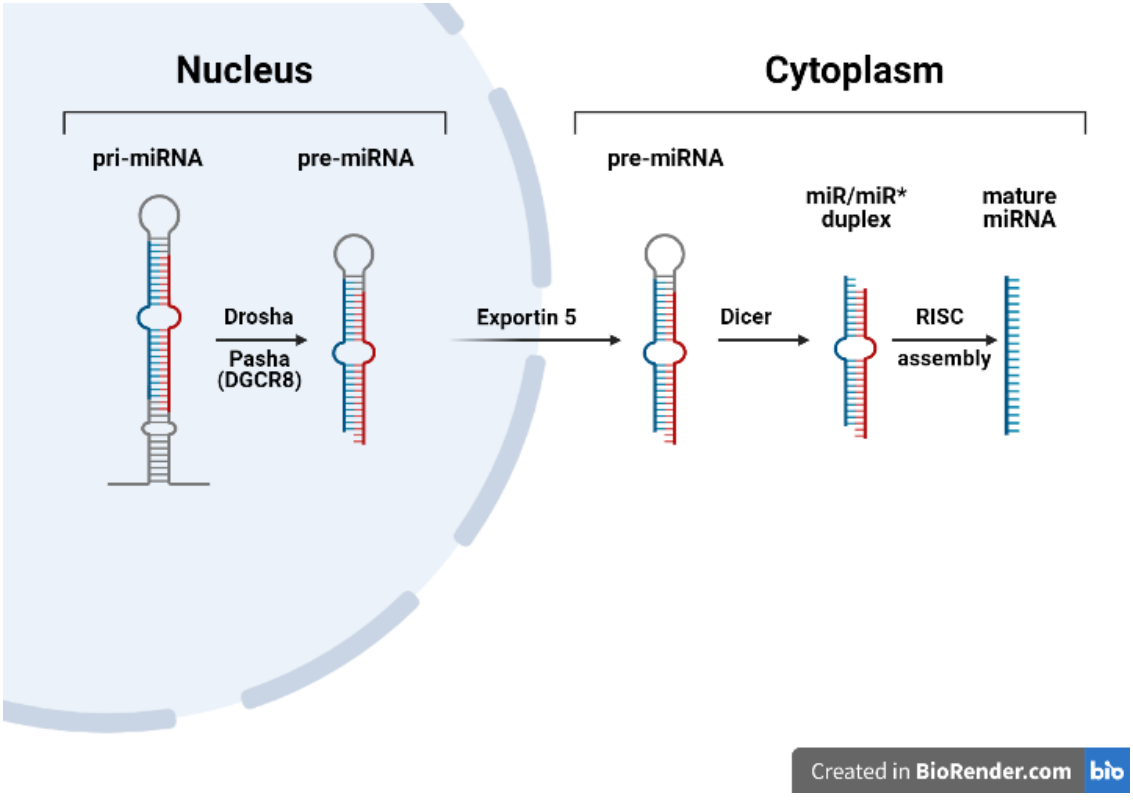


Figure 1 - miRNA biogenesis. Figure adapted from [190] by Biorender.com.

1.3.2 Nanomolecular miRNA-mediated regulation of gene expression

The miRNA nanomolecules have diverse regulatory mechanisms that support various cellular processes. They can regulate gene expression via different pathways and modes [191]. One such mode involves miRNA nanomolecules binding to the complementary mRNA nanomolecules, resulting in mRNA degradation by Ago2-protein through endonuclease restriction. Otherwise, if miRNA nanomolecules are not complementary to mRNA binding sites, they inhibit translation initiation or elongation. In this case, mRNA exonuclease cleavage processes occur by exploiting P-bodies, which are concerned with mRNA degradation [191]. Experimental data confirms nine mechanisms of miRNA nanomolecule action once it binds to the target gene's mRNA [191].

1) The suppression initiation might be caused by AGO proteins binding to the mRNA cap sequence, which physically cancels out the eIF4E initiator factor [192]. However, this recently discovered explanation demonstrated that a domain analogous to eIF4E in AGO proteins cannot engage in cap-binding because it is possessed by GW182, an AGO protein's acting partner; this AGO protein prevents interaction with the cap sequence [192].

2) The inhibition of the 60S joining. Ribosome 60S binds to ribosome 40S, preventing the formation of ribosome 80S and reducing the number of assembled ribosomes [193].

Additionally, it inhibits the elongation stage of mRNA, decreasing the number of mRNA-associated ribosomes and resulting in complete inhibition of translation elongation, leading to mRNAs harboring a single 80S ribosome. This suppression of translation occurs in cap-independent mRNA [191].

4) The premature termination. The drop-off of the elongating ribosome from miRNA nanomolecules. The translation of mRNA was performed with skipping of stop-codons and it is observed the dissociation of polyribosomes [194].

5) Targeting by mi-RISC complex causes simultaneous degradation of the peptide, as this may be due to uptake of a not yet detected protease [191].

6) Isolation of mRNA in P-bodies. P-bodies are separate foci formed in the cytoplasm of a eukaryotic cell, consisting of many enzymes involved in the mRNA cycle, as a result of phase division [195]. The P-body components include Argonaute proteins, GW182, the CCR4-CAF1-NOT deadenylase complex, the decapping DCP1-DCP2 complex, and XRN1 [196]. They are required for deadenylation and rapid degradation of target mRNAs. All of these proteins are localized in P bodies. Argonaute proteins interact with GW182, DCP1, DCP2, and RCK/p54 and can be present as complexes in P bodies [196]. Additionally, these proteins, along with miRNAs and miRNA forms, are present in P bodies, providing clear evidence of their involvement in RNA silencing [196].

7) The degradation of mRNA. mRNA degradation can occur by inhibiting translation with or without subsequent destruction of the mRNA. When protein levels decrease by 33%, miRNA silencing leads to mRNA cleavage as the primary

outcome. The mRNA degradation process begins with deadenylation, decapping, or 5'- and 3'-mRNA degradation [193, 194, 197].

8) mRNA Cleavage: This mechanism involves the nearly complete complementarity of miRNA with the mRNA of the target gene, resulting in mRNA cleavage. This process resembles the mRNA cleavage mechanism induced by small interfering RNA (siRNA). While mRNA cleavage is common in plants, it's relatively rare in animals, but when it occurs, it leads to the generation of RNA fragments [193].

9) Translation Inhibition and Gene Silencing: miRNA can inhibit translation by inducing chromatin reorganization. For this mechanism, full complementarity between specific miRNA and mRNA promoter sites is required. Additionally, miRNA may potentially increase methylation of mRNA target gene promoters, leading to gene silencing [192].

Regarding the influence of miRNA nanomolecules on mRNA located in different regions:

The known mechanisms you've collected mainly focus on miRNA action on mRNA sites situated in the 3'UTR. However, there is still a lack of exploration regarding how miRNA affects mRNA sites located in the 5'UTR and CDS. The question arises: Do these mechanisms differ when miRNA targets sites in the 5'UTR and CDS compared to the 3'UTR?

So far, it's been observed that sites situated in the CDS have a more pronounced degrading effect compared to sites located in the 3'UTR [193]. However, the specific mechanisms and differences in action between these regions, especially the 5'UTR and CDS, require further investigation and exploration. Understanding these distinctions could provide valuable insights into the diverse regulatory roles and potential therapeutic applications of miRNA in modulating gene expression.

1.3.3 The role of nanomolecular miRNAs in carcinogenesis

In recent times, the role of miRNA nanomolecules in modulating oncological diseases has garnered significant attention. Enhanced methodologies for isolating and scrutinizing miRNA expression have enabled the identification of these molecules not only within cells but also across a diverse range of biological fluids [198]. Despite limited understanding surrounding the mechanisms governing the release of miRNA into biological fluids and their extracellular functions, an increasing body of research is establishing connections between circulating miRNAs and the detection of diverse pathological conditions [198].

MiRNAs wield influence over gene expression and are notably detectable in a remarkably stable state in various biological fluids, including whole blood, plasma/serum [199], urine [200, 201], and other bodily fluids [202]. Operating as post-transcriptional regulators, miRNA nanomolecules efficiently oversee gene expression. Remarkably, individual miRNAs can regulate multiple genes, while conversely, individual genes may be under the control of multiple miRNAs [203-

205]. Hence, since their discovery a little over two decades ago, miRNA nanomolecules have earned recognition as pivotal regulators implicated in intricate biological processes associated with a spectrum of oncological pathologies.

An indication of the relationship between miRNA nanomolecules and cancer is that numerous miRNA genes are found in genomic regions where rearrangements are frequently linked to the development of oncological diseases. Thus, the study on the genomic localization of 186 miRNA genes revealed that 98 (52.5%) of them are located in the "fragile" regions of chromosomes, which are prone to mutagenesis and genetic material reorganization [206]. Additionally, a high occurrence of alterations in the copy number of miRNA genes was observed in non-small-cell lung cancer, ovarian cancer, breast cancer, and melanoma [206-208].

In B-cell chronic lymphocytic leukemia, the overexpression of BCL-2 results from the reduced expression of miR-15a and miR-16-1, primarily due to the deletion of the 13q14 region [209]. This discovery establishes that miRNA nanomolecules can serve as a leading factor in promoting the development of cancer. The term onco-miRNA (oncomiR) is used to identify oncogenic miRNAs (Figure 2) [210]. The example of such onco-miR nanomolecules is miR-21, miR-155, and the miR-17-92 cluster, that confirmed in animal models of lymphoproliferative diseases, with an increase in expression of miR-155 in the absence of other permissive mutations was sufficient for the formation of a malignant phenotype of lymphoid cells [210-227].

Extensive evidence has been gathered to date demonstrating significant changes in miRNA expression profiles within tumor cells [228-231]. The miRNAs regulate gene expression in cancer development and control apoptosis and tumor cell proliferation through pathways such as *p53, HER2, PTEN/PI3K/Akt, mTOR*, and *ERK/MAPK*. They also affect tumor growth response to DNA damage and repair, as well as tumor cell interaction with the microenvironment [232-243].

Tumor cells from various locations exhibit different profiles of miRNA nanomolecule expression. A classifier comprised of 48 miRNA nanomolecules enables the prediction of the tissue of origin for tumor tissue samples in a group of 336 primary and metastatic tumors, with respective accuracies of 86% and 77% [244]. Another study developed a classifier for 21 types of cancer using clinical data and information on miRNA nanomolecule expression within The Cancer Genome Atlas (TCGA) tumor genome research project. The classifier accurately diagnosed cancer with a minimum accuracy of 97% and 80% in the training and validation sets, respectively [245]. The biogenesis of miRNA molecules changes in cancer, with distinct miRNA isoform spectra varying significantly according to cancer location [198]. Further, miRNA expression profiles differ among histological tumor subtypes, such as gastric [246] and colorectal [247] cancers. Changes in individual miRNA molecule expression often coincide with a global decrease in miRNA expression in tumor cells [248, 249].

The miRNA nanomolecules exert a negative influence on gene expression by either repressing translation or leading to the degradation of mRNAs, thereby

influencing various cellular processes [187, 250]. Currently, "miRBase" lists 2578 mature Homo sapiens miRNAs, while an article by Londin et al. indicates 3707 mature miRNAs [251]. Computational estimations regarding target mRNAs suggest that over 60% of all mammalian protein-coding genes are conserved targets of miRNAs [252]. Each individual miRNA is found to have target sites in hundreds of genes. These miRNAs demonstrate tissue-specific expression patterns and display effects dependent on their concentration in organs and tissues affected by pathological conditions [253].

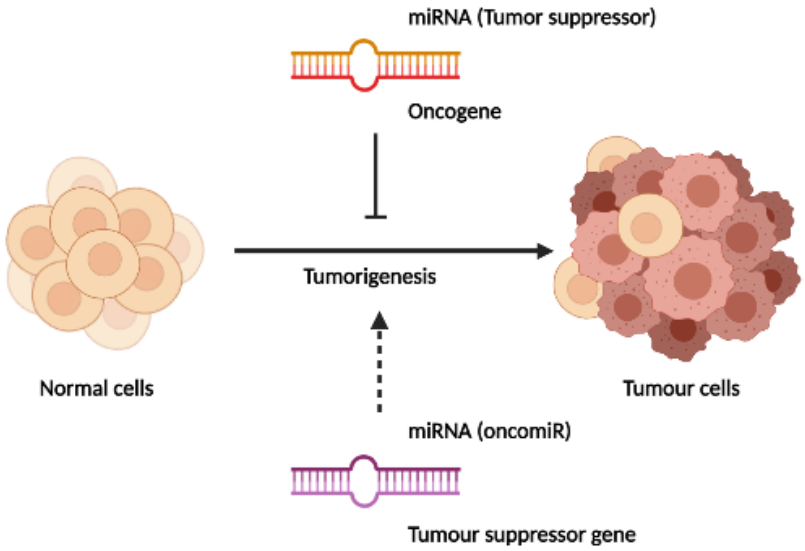


Figure 2 - Roles of miRNAs in cancer. Depending on the target gene, miRNAs suppress or stimulate cancer-related genes. Figure modified from [254], created by BioRender.

Since then, miRNA nanomolecules have garnered significant attention from scientists due to their potential role as diagnostic and prognostic biomarkers in various diseases [255]. They have also been identified as potential gene-specific therapeutic targets in disease modeling [256]. Particularly in oncological diseases, miRNAs are viewed as disease-specific biomarkers, exhibiting distinct groups of circulating miRNA nanomolecules [257].

1.3.4 The prediction of miRNA nanomolecules binding sites in mRNAs of candidate genes

Characterizing the effect of miRNA nanomolecules on their target mRNAs is challenging due to the each miRNA has multiple target mRNAs and vice versa. As a result, accurately identifying interactions between miRNA nanomolecules and mRNAs remains problematic. Typically, this issue is addressed by predicting interactions and conducting experimental verification of these miRNA-mRNA interactions. Bioinformatics tools and novel experimental approaches have emerged

to address this challenge. Bioinformatics tools aim to predict effective associations between miRNA and mRNA for experimental validation. This knowledge is essential for elucidating the mechanisms of miRNA action in both normal physiological processes and disease. For this purpose, several tools have been developed to predict miRNA targeting. Over the last few decades, various approaches have been used to recognize miRNA-mRNA interactions. However, the complexity of algorithms and parameters used by each tool presents a challenge for scientists with limited or no experience in bioinformatics when selecting suitable miRNA for subsequent validation.

Many programs have been developed to predict potential miRNA binding sites in mRNA using the application of canonical interaction rules often coupled with additional parameters, including complementarity, free binding energy, and conservativity [258, 259]. Depending on the specific algorithm, the priority of these parameters may vary. For example, TargetScan, PicTar, and miRanda use sequence conservation as the main parameter, while PITA and RNA22 rely primarily on estimates of the free energy of miRNA-mRNA pairing [259, 260]. All of these software packages are permeated by the same problem: they all consider interactions only in the “seed” region and not along the entire length of the target mRNA, and this explains the excessive false positive results.

Considering the interaction of miRNA nanomolecules with mRNA nanomolecules along the entire miRNA nucleotide sequence, rather than focusing solely on the "seed" region, is essential due to various factors. One of the key reasons is the high specificity characterizing the interaction between miRNA and mRNA.Thus, the full length of miRNA nanomolecules must be taken into account. It is same as applying primers, of at least 20 nucleotides length in a polymerase chain reaction, when it is necessary for amplification to specifically choose only one nucleotide sequence among numerous nucleotide sequences. Another reason for using the full miRNA nucleotide sequence is that during the evolution process, a part of miRNA nanomolecules other than "seed" will vary if it is not a critical site in interactions of miRNA with mRNA. Confirmation of the need to maintain the entire length of miRNA is the high conservatism of miRNA in organisms that have diverged over tens of millions of years of species evolution. According to miRBase the nucleotide sequences were identical in miR-200c-5p of human, mouse, rat, miR-216a of human, bull, mouse, miR-574-5p and miR-574-3p of human, mouse, rat, pig, etc [261]. One of the first misconceptions in the study of miRNA was the assumption that miRNA binds only (or predominantly) in the 3'UTR mRNA of human genes. However, miRNAs do not have the property of distinguishing binding sites in 5'UTR, CDS, and 3'UTR. The miRNA interacts with mRNA based on the physicochemical properties of these molecules. Therefore, the interaction site can be located in any mRNA region, and as yet unknown prohibitions on the location of such sites throughout the mRNA nucleotide sequence. The conditions for the successful interaction of miRNA nanomolecules with mRNA are the energy characteristics and

conformational properties of this interaction. The proposed assumption of preferential miRNA nanomolecules binding in the 3'UTR contradicts several established properties of miRNA and mRNA. The proposed assumption of preferential miRNA nanomolecules binding in the 3'UTR contradicts several established properties of miRNA and mRNA. Considering that known miRNAs (miRBase) have differences in GC-content in the same range as in human genes, it was logical to assume that the probability of binding miRNA in 5'UTR, CDS, and 3'UTR of mRNA will correlate with the GC-content of genes and corresponding miRNAs. In numerous publications, sites of miRNA nanomolecules interaction with mRNA were studied only in the 3'UTR. This is because practically all programs for detecting miRNA binding sites are predicted only in this mRNA region. Several publications show the interaction of miRNA with mRNA in 5'UTR and CDS [262-265].

2 MATERIALS AND METHODS

2.1 Materials

The data used in this study encompassed mRNA nucleotide sequences of human (Homo sapiens) involved in the development of esophageal, gastric, small intestinal, and colorectal cancers, obtained from GenBank (http://www.ncbi.nlm.nih.gov). Information regarding the genes' roles and functions was sourced from GenBank databases and related publications.

The 2567 mature miRNA nanomolecule sequences were sourced from the miRBase database (http://mirbase.org). Over the years, the identification of miRNA nanomolecules has progressively increased, leading to updates in the miRBase database. In 2011, the database contained approximately 1500 miRNAs, which expanded to around 2000 miRNAs in 2012 and further increased to 2500 miRNAs in 2013. Subsequently, in 2015, Londin et al. identified an additional 3707 novel miRNAs [251]. The expanding pool of novel miRNA nanomolecules suggests a mix of evolutionarily conserved and species- or tissue-specific miRNAs. Moreover, these miRNAs exhibit varying expression levels across different tissues, emphasizing their distinct tissue- and cell-specific functional impacts.

2.2 Methods

Using the MirTarget program, we have found binding sites in 5'UTR, CDS, and 3'UTR of many human, animal and plant genes.

The following examples of search for binding sites by the MirTarget program and other programs show the great advantage of the MirTarget program. The result of the TargetScan program almost coincidentally almost coincides with the result of the MirTarget program because this binding site has a complete complementarity of interaction between miRNA and miRNA of the *CYP2W1* gene. The lack of G-C bond (marked in red) in the TargetScan interaction scheme is a program error. Who can forbid these nucleotides to form hydrogen bonds?

The hydrogen bonds that the MirTarget program takes into account and according to which the binding sites for miR-5095, miR-619-5p are reliable, are indicated in red. TargetScan Program considers these sites "reliable" with an obvious lack of selectivity. The term "seed" is unreasonably used as the main factor ensuring the binding of miRNA with the mRNA of the target gene. Because of this misconception, thousands of publications deal with false-positive candidate target genes. Because of this misconception, over the twenty years of studying miRNA, no adequate method for diagnosing diseases using miRNA has been created. The data on the operation of programs based on "seed" is given in several tables below. These results clearly show the inadequacy of the prediction of binding sites.

The DIANA-MicroT-CDS program does not provide information on the interaction of miR-5095, miR-1273a, miR-1273e with mRNA of the *ATM* gene. The TargetScan program identified a reliable site for miR-5095, but did not consider the complete interaction with the *ATM* gene's mRNA due to a non-canonical A-C pair. The second site is a false-positive, containing hydrogen-free pairs (A and G, G and

G) that violate stacking interactions and do not form hydrogen bonds. Furthermore, the program does not consider the interaction of eight nucleotides. The first binding site for miR-619-5p is found without accounting for three nucleotides, which significantly disrupt stacking interactions in the scheme. Moreover, the second binding site for miR-619-5p cannot be located based on the presented scheme. The first reliable site for miR-619-5p TargetScan's binding site is identified, although it does not account for the interaction of seven nucleotides. The second site includes hydrogen-free couples A, G and G, G, yet it neglects the interaction of nine nucleotides. The third binding site, despite indicating seven interacting nucleotides as per the scheme, does not consider the interaction of 15 nucleotides. TargetScan software does not identify miR-1273a. The length of miR-1273e is attributed to the presence of a non-canonical A-C pairing.

TargetScan identified binding sites solely for miR-1273e and miR-5095 within the mRNA of the *IL11* gene, but did not consider nucleotide interactivity throughout the entire lengths of miR-1273e and miR-5095 (see Figure 3). Neither the DIANA-MicroT-CDS nor the RNA22 programs detected binding sites for the four miRNAs in the *IL11* gene's mRNA. TargetScan only identified false positive binding sites for miR-1273a and miR-3198 in the mRNA of the *CBL* gene. However, it failed to take into account the interaction of nucleotides throughout the length of miR-1273a and miR-3198 (Figure 4). In contrast, RNA22 did not detect any binding sites for the four miRNAs in the mRNA of the *ATM* gene (Figure 5). The programs DIANA-MicroT-CDS, TargetScan, and RNA22 were unable to uncover binding sites for miR-1285-3p and miR-1273c in the mRNA of the STMN1 gene (Figure 6). Some programs may not find information about miRNA or genes when introduced, but they can still identify a reliable site on the ATM gene mRNA. However, they may not consider the entire length of the interaction between the molecules.

Gene *IL11*		
miRNA	MirTarget	TargetScan
miR-1273d	5'-ACUGCA**A**CCUCCACCUCC**C**GGGUUC-3' ||||||||||| ||||| ||||||| 3'-UGACGU**C**GGAGUUGGAGU**A**CCCAAG-5'	Does not determine
miR-1273e	5'-UCCACCUCC**C**GGGUUCAAGCAA-3' |||||||||||||||||||| 3'-AGGUGAAGG**A**CCCAAGUUCGUU-5'	5'...CUCCACCUCCCGGGUUCAAGCAA... ||| ||||||| 3' AGGUGAAGGACCCAAGUUCGUU
miR-1273f	5'-CACUGCAACCUCCA**C**CUCC-3' ||||||||||||||||| 3'-GUGACGUUGGAGGU**A**GAGG-5'	Does not determine
miR-5095	5'-C**AU**GGUGG**C**UCACGCCUGUAA-3' ||||||||||||||||||| 3'-G**CG**CCACC**A**AGUGCGGACAUU-5'	5' ...GGCAUGGUGGCUCACGCCUGUAA... ||||| ||||||| 3' GCGCCACCAAGUGCGGACAUU

Figure 3 - Comparison of miR-1273d,e,f binding sites with mRNA of *IL11* gene in miRTarget and TargetScan programs.

CBL		
miRNA	MirTarget	TargetScan

miR-1273a	5'GAGAUGGAGUCUCGCUGUGUCGCCC-3'
	  |||| |||||||||| ||||||||
	3'UUCUUUCUCAGAACGAAACAGCGGG-5'
	
	MirTarget program evaluates this site for miR-1273a as inadequate because ΔG/ΔGm is less than 50%.
		5' ...CUGGACCUUUCCAGCUGUCGCCA.
		  ||| |||||||
		3'UUCUUUCUCAGAACGAA—ACAGCGGG
		
		5'...GAUGGAGUCUCGCUGUGUCGCCC.
		  |||||| |||||||
		3' UUCUUUCUCAGAACGAAACAGCGGG

miR-3198	5'UCUCCAUUCUCCAUGGCCCCAC-3'
	  |||||||||||| ||||||||
	3'AGAGGUAAGGGGUCCUGAGGUG-5'
	
	MirTarget program evaluates this site for miR-3198 as inadequate because ΔG/ΔGm is less than 50%.
		5'...CAGCUAACACUGGUCACUCCAAA.
		  ||||||
		3' AGAGGUAAGGGGUCCUGAGGUG
		
		5' ..AGAGAAGUGAGCAGAGACUCCAC.
		  |||||||
		3' ACCUUAGGAUCUGAGGUC

Figure 4 - Comparison of miR-1273a and miR-3198 binding sites with mRNA of *CBL* gene in miRTarget and TargetScan programs.

ATM		
miRNA	MirTarget	DIANA-MicroT-CDS
miR-5095	5' - CGCAGCGGCUCACGCCUGUAA - 3' ||||||||||||||||||| 3' - GCGCCACCAAGUGCGGACAUU - 5'	Does not determine

miR-619-5p	5' - GGCUCACGCCUGUAAUCCCAGC - 3'
	  |||||||||||||||||||
	3' - CCGAGUACGGACAUUAGGGUCG - 5'
	
	MirTarget evaluates this site to be inadequate for miR-619-5p since ΔG/ΔGm is below 50%.
		5'GCAUGGUGGCGGGCA CCUGUAAUCCCAGC 3'
		  ||| || ||||||||||||
		3' CCGA GUACGGACAUUAGGGUCG 5'
		
		5' UGAGUGUCUGUUUAUGUCAUAU UCCCAG 3'
		  |.|.|||.| ||||||
		3' CCGAGUACGGACAUUAGGGUCG 5'

miR-1273a	5' - GAGACAGAGUCUUGCUCUGUCACCC - 3' |||| |||||||||||||||||| 3' - UUCUUUCUCAGAACGAAACAGCGGG - 5'	Does not determine
miR-1273e	5' - UCUGCCUCCUGGGUUCAAGCAA - 3' ||||||||||||||||||| 3' - AGGUGAAGGACCCAAGUUCGUU - 5'	Does not determine
miRNA	TargetScan	RNA22

miR-5095	5' ...GGCGCAGCGGCUCACGCCUGUAA...
	  |||||||||||
	3' GCGCCACCAAGUGCGGACAUU
	5' ...GCAUGGUGGCGGGCACCUGUAAU...
	  ||||| ||||||
	3' GCGCCACCAAGUGCGGACAUU
		Does not determine

miR-619-5p	5' ...CGGCUCACGCCUGUAAUCCCAGC...
	  ||||||||||||||
	3' CCGAGUACGGACAUUAGGGUCG
	
	5' ...UGGCGGGCACCUGUAAUCCCAGC...
	  |||||||||||||
	3' CCGAGUACGGACAUUAGGGUCG
	
	5' ...CUUAGGAAAAUGUUCAUCCCAGC...
	  |||||||
	3' CCGAGUACGGACAUUAGGGUCG
		Does not determine

miR-1273a		Does not determine
miR-1273e	5' ...CUCUGCCUCCUGGGUUCAAGCAA... ||||||||||||||| 3' AGGUGAAGGACCCAAGUUCGUU	Does not determine

Figure 5 - Comparison of miR-5095, miR-619-5p, and miR-1273a,e binding sites with mRNA of *ATM* gene in miRTarget, DIANA-MicroT-CDS, TargetScan and RNA22 programs.

Gene *STMN1*				
miRNA	MirTarget	DIANA-MicroT-CDS	TargetScan	RNA22
miR-1285-3p	5'-GAGUCUCACUCUGUCGCCCAGG-3' |||||||||||||||||||| 3'-UCCAGAGUGAAACAACGGGUCU-5'	Does not determine	Does not determine	Does not determine
miR-1273c	5'-GGCAGAGUCUCACUCUGUCGCC-3' |||||||||||||||||||| 3'-CUGUCCCAGAGCAAAACAGCGG-5'	Does not determine	Does not determine	Does not determine

Figure 6 - Comparison of miR-1285-3p and miR-1273c binding sites with mRNA of *STMN1* gene in miRTarget, DIANA-MicroT-CDS, TargetScan and RNA22 programs.

From the above analysis of the comparison of various programs for searching binding sites, it is clear the prevalence of the MirTarget program.

2.3 Programs for processing data of miRNA

For analysis of the obtained results, there were used programs, which has been specially designed and created by A.Yu. Pyrkova and O.A. Berillo, the staff of our laboratory.

Lextractor.

In the initial stage, distinct textual files were generated, containing nucleotide sequences and specific parameters for each examined gene and miRNA. For binding site prediction programs to function effectively, a text format of data is necessary, available in files with mir and gene extensions. The nucleotide sequence, gene name, and chromosomal location of a mature miRNA were transcribed into the mir extension file. The mRNA nucleotide sequence for each gene was obtained from the GenBank database. The mRNA information and exon-intron organization were saved in gbk documents. These files underwent processing via Lextractor programs to produce text in the FASTA format, including the mRNA nucleotide sequence with identified boundaries of the 5'UTR, CDS, and 3'UTR regions. The resulting file names include the gene extension.

MirTarget Program.

The process of searching for target genes for miRNAs remains consistent when using the MirTarget program, which was developed in our laboratory.

This program identifies the binding sites of miRNAs in the mRNA of candidate genes, including the distribution of binding sites in the 5'-untranslated region (5'UTR), protein-coding region (CDS), and 3'-untranslated region (3'UTR). Additionally, it calculates the free energy of hybridization (ΔG, kJ/mol) and provides a scheme of the miRNA-mRNA nucleotide interaction.

The ΔG/ΔGm (%) value was calculated for each binding site, where ΔGm represents the free binding energy of miRNAs with a fully complementary nucleotide sequence. The miRNAs-mRNAs binding sites with a ΔG/ΔGm value of 90% or higher were selectively chosen. The language used is formal, with a clear and concise structure. The position of binding sites is provided from the first nucleotide in the

5'UTR of mRNA. The MirTarget program includes a unique feature that considers nucleotide interactions between miRNAs and target gene mRNAs. Notably, interactions are accounted for with adenine (A) and uracil (U), guanine (G) and cytosine (C), G-U, and A and C through hydrogen bonding. The distance between A and C is equal to that of G-C, A-U, and G-U nucleotides.

MirTarSeq.

The miRTarSeq is a program designed to identify oligonucleotides that encode oligopeptides based on the nucleotide sequences found in miRNA sites of mRNA genes' 5'UTR, CDS, and 3'UTR. These fragments are identified with 15 additional nucleotides situated to the left and right of the binding site, an essential factor in analyzing evolutionary conservatism. This program operates when files containing binding site schemes acquired from MiRTarget's site search (scheme.mres) and gene files (Gene.gene) are present.

Upon processing the data, the program produces a table detailing the information for each site. mRNA sequence of the gene of interest, specific miRNA (identified by its ID number and origin), the nucleotide position where the binding site initiates (nt), the calculated ΔG value (kJ/mol) and ΔG/ΔGm ratio (%), the length of the miRNA (nt), and the nucleotide sequence of the binding site with flanking nucleotides on either side. Additionally, the encoded oligopeptide within the binding site, and the oligopeptides found within the three open reading frames.

The program offers data on the flanking nucleotide sequences of binding sites in the 5'UTR, CDS, and 3'UTR. Binding sites in the CDS have an advantage because the program converts their nucleotide sequences into amino acid sequences encoded by them while considering various open reading frames. Using this program allows for the study of evolutionary conservation and variability, specifically analyzing changes in nucleotide sequences during the process of evolution. The results of the program analyze orthologous gene properties, which are genes that originated from a common predecessor and were vertically transmitted during evolution across various species. This analysis aids in selecting experimental animals for in vivo studies, focusing on genes implicated in a variety of diseases. Orthologous sequences provide evidence that certain sequences are maintained throughout evolutionary processes despite speciation. The conservation of specific protein sequences indicates the presence of a conservative gene, as information regarding amino acid sequences is typically transmitted from one generation to the next.

3 RESULTS AND DISCUSSION

3.1 Creation of databases on genes associated with carcinogenesis of the GI tract

To create a reliable gene database, the information available from the NCBI (National Center for Biotechnology and Information) was obtained and the gene was verified by PubMed, identified by name, and entered into the appropriate information retrieval system to determine its relationship to the desired disease and add to our database. Data validation was carried out by identifying the association of these genes with the disease under study and searching publications over the past seven years. It should be noted that almost all genes have several synonymous names, which significantly complicates the identification of genes according to the classification established by the NCBI. It is difficult to determine which type of cancer is the cause of this cancer because there is no single system for classifying cancer.

Gene database in Excel format and as a text file. It provides information about the gene and its synonymous names, the full name of the gene, a brief description of its function, and the source of information (PMID). As a result of the search for genes that regulate the development of the gastrointestinal tract, 38 genes responded to esophageal adenocarcinoma, 66 to esophageal squamous cell carcinoma, 215 to gastric cancer, 40 to the small intestinal cancer, and 135 to colorectal cancer, summarized in Table C1.

Analysis of research on the role of genes in the development of gastrointestinal tract cancer indicates a recent increase in publications on this topic. These studies have confirmed the presence of numerous previously known and newly discovered genes, highlighting the promising potential of this field and the effectiveness of a targeted database in generating precise insights into the genes involved in the progression of gastrointestinal tract cancer.

For computer analysis, the nucleotide sequences of miRNA nanomolecules and mRNA of *H. sapiens* GI tract cancer genes were obtained from miRBase and Genbank databases also the information about novel 3707 miRNAs taken from the publication of Londin et al [248]. These sequences were analyzed through the created in our laboratory Lextractor, MirTarget, and MirTarSeq programs.

3.2 The characteristics of the miRNA nanomolecules interactions with mRNAs of genes associated with carcinogenesis of the GI tract

It was searched for miRNA nanomolecules binding sites in mRNA of 372 human genes responsible for the development of GI tract cancer. As a result of the study it was found that 305 genes out of 372 genes have binding sites with high affinity for miRNA.

Analysis of the information on the involvement of genes responsible for the development of GI tract cancer shows that the number of publications on this problem has increased in recent years. These publications confirmed the involvement of many previously established genes in the development of these diseases and revealed new genes. This indicates the promise of the direction we have chosen and

the possibility of obtaining, on a large factual basis, more reliable results on the participation of genes in the development of various types of GI tract cancer carcinogenesis.

Previously, we used only the degree of complementarity of miRNA nanomolecules binding to mRNA ΔG/ΔGm (%) for the selection of miRNA nanomolecules that effectively bind to mRNA. However, this criterion did not take into account the length of the miRNA, on which the free energy also varied, which was different for different miRNA lengths. With the same ΔG/ΔGm value for miRNA nanomolecules of 17 nt and 25 nt, the binding energy of miRNA 25 nt was 1.47 times higher with respect to the absolute value of the energy of miRNA nanomolecules with 17 nucleotides. Therefore, the criterion ΔG/ΔGm was introduced for miRNA of different length. This value makes it possible to reduce the number of false positive miRNAs with a length of less than 20 nt.

The obtained results are sorted according to the value of ΔG/ΔGm (%), where ΔG is the Gibbs free energy and ΔGm is the free energy of miRNA nanomolecules binding with strand complementarity of 100%. In addition, miRNA-mRNA binding sites with ΔG/ΔGm values less than the corresponding length were excluded from the results. The maximum standard value for the most common length (22 nt corresponds to 90%) was used as the basis for this free energy binding correspondence (Table 2).

Table 2 - $\Delta G/\Delta G_m$ value, taking into account the length of miRNA.

Length of miRNA (nt)	ΔG/ΔGm (%)
17	98
18	96
19	94
20	92
21	91
22	90
23	89
24	88
25	87
26	86

When ΔG is equal to 85% or less, miRNA of different families differ by more than three nucleotides, and this circumstance makes it necessary to change the accepted classification of miRNA nanomolecules by families when families differ within three nucleotides at a length of 22 nt. The obtained results were analyzed taking into account the new criterion.

In order to identify miRNA nanomolecules targeting genes responsible for the development of gastrointestinal tract cancer, a search for binding sites in 372 mRNA nanomolecules of human genes was performed. The study revealed that 305 out of

372 genes possess binding sites with high affinity for miRNA. Based on the computer analysis results, tables were generated to show the acceptable energy range for miRNA nanomolecule interaction with mRNA, with Table 3 providing the initial data. The tables list the names of the genes responsible for GI tract carcinoma regulation and the miRNA nanomolecules that bind to them, along with their respective interaction energy values.

The tables below display gene names that regulate GI tract carcinogenesis (first column), followed by miRNAs that bind to these genes (second column). The third column indicates the position of binding sites, while the fourth column notes the location of these sites. The fifth and sixth columns detail the interaction energy value, and the seventh column denotes the length of the binding site.

The selected key genes in GI tract carcinogenesis fall into the following gene categories and groupings:

- Esophageal cancer
- Gastric cancer
- Small intestinal cancer
- Colorectal cancer

3.2.1 The characteristics of the miRNA nanomolecules interactions with mRNAs of genes associated with esophageal cancer

3.2.1.1 The characteristics of the miRNA nanomolecules interactions with mRNAs of genes associated with esophageal adenocarcinoma

The examination revealed that among the 38 candidate genes, 32 were the targets of miRNA nanomolecules. The distribution of miRNA binding sites within the mRNA sequences varied across these candidate genes. While multiple miRNA nanomolecules exhibited independent binding sites in distinct regions of the mRNA, some miRNAs demonstrated overlapping nucleotide sequences at these sites, leading to the formation of clusters. Tables 3 and C2 outline the properties of miRNA interactions in the 5'UTR mRNA genes implicated in EAC development.

For instance, within the ALDH1A2 gene's mRNA in the 5'UTR, interactions occurred with ID00792.3p-miR and ID00744.3p-miR, having miRNA ΔG values of -132 kJ/mol and -125 kJ/mol, respectively (Table C2). Furthermore, the mRNA of the ARID1A gene in the 5'UTR contained three clusters of miRNA binding sites (Table 3). The first cluster spanned from 136 nt to 192 nt, exhibiting an average free energy of interaction at -124 kJ/mol. The second cluster spanned from 206 nt to 228 nt, and the third from 295 nt to 331 nt, totaling 37 nucleotides in length with an average ΔG value of -111 kJ/mol. The ARID1A gene plays a role as a tumor suppressor and holds significance in the carcinogenesis process of various organs [266]. Loss of ARID1A protein has been identified in Barrett's esophagus, a precancerous lesion of esophageal adenocarcinoma, and the rate of loss was higher in lesions showing more severe dysplasia [267].

Table 3 - The characteristics of miRNA interactions with 5'UTR mRNAs of EAC candidate genes [268].

Gene; NX	miRNA	Position, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
ARID1A; 23.2	ID02106.3p-miR	136	-123	89	23
	ID01778.3p-miR	140	-134	90	24
	ID00296.3p-miR	141, 166	-138	88	25
	miR-6081	143	-125	89	24
	ID01702.3p-miR	147	-134	89	24
	ID00465.5p-miR	148	-113	93	20
	ID01377.3p-miR	152	-117	92	20
	ID02592.5p-miR	164, 167	-123	89	23
	miR-3960	167	-115	92	20
	ID03065.3p-miR	172	-115	92	21
	ID02751.3p-miR	206	-125	92	23
	miR-483-3p	208	-104	91	21
	ID01257.3p-miR	295	-113	93	20
	ID00414.3p-miR	303	-108	93	20
	ID02428.3p-miR	310	-113	91	22
BARX1; 0.3	ID00961.3p-miR	105	-132	93	23
	ID00756.3p-miR	126	-123	89	23
	ID00457.3p-miR	129	-125	92	22
	ID02064.5p-miR	134	-129	90	23
CDK9; 20.3	ID02823.3p-miR	25, 48	-117, -119	90, 92	22
	ID02094.3p-miR	62	-113	90	22
	ID01158.3p-miR	86	-119	98	20
DKK3; 10.0	ID03462.5p-miR	73	-123	91	22
	ID01632.5p-miR	93	-125	92	23
FOXF1; 6.6	ID00278.3p-miR	1	-123	89	23
	ID01310.3p-miR	6, 12	-119	90	22
	ID03332.3p-miR	14	-132	89	24
FOXP1; 3.6	ID00387.3p-miR	108	-125	89	23
	ID03332.3p-miR	130	-134	90	24
GDF7; 0.3	ID00009.3p-miR	546	-115	92	20
	ID01910.3p-miR	546	-115	92	20
	ID02781.3p-miR	547	-121	97	20
IGFBP2; 49.0	ID02036.3p-miR	60	-117	93	20
	ID01293.5p-miR	65	-123	91	22
	ID01638.3p-miR	84	-113	95	20
	ID03120.3p-miR	86	-117	92	21
SEPP1; 20.5	miR-5095	116	-110	95	21
	miR-619-5p	122	-115	95	22

The cell cycle is a highly conserved and highly regulated biological system that controls cell proliferation and differentiation. Changes in regulatory proteins, such as

cyclins, Cyclin-dependent Kinases (CDKs), and CDK inhibitors, are associated with the loss of normal cell cycle regulation. This is a hallmark of several cancer types, including EAC [269]. The *CDK9* gene is a candidate gene that we have identified specifically for EAC. The ID02823.3p-miR and ID02094.3p-miR cluster together to form binding sites on mRNA of the *CDK9* gene. These sites are located at positions 25 to 83 and have a total length of 59 nucleotides. The average interaction energy measures -116 kJ/mol.

The mRNAs for *AR, ERBB2, ERBB3, FKBP5, FOXM1, ROCK2,* and *SOX2* genes in 5'UTR were targeted by single miRNAs (see Table C2). Nuclear *AR* and high *FKBP5* expression have been found to be associated with reduced survival in the EAC [270]. Six miRNAs target the mRNA of the *GDF7* gene, of which three form a cluster of binding sites (see Table 3). This cluster of binding sites includes all three miRNAs at positions ranging from 546 to 567 nucleotides, with a length of 22 nucleotides. The combined length of the three sites is 60 nucleotides. The presence of a cluster of binding sites within the 5'UTR of the *GDF7* gene suggests that it has an increased capacity for compaction, resulting in competition among miRNAs for these binding sites. The average binding energy for these three miRNAs was -117 kJ/mol.

Du et al [271] determined that *BTG3* gene expression was significantly suppressed in EAC tissues compared to adjacent normal tissues. *BTG3* is a direct transcriptional target gene for p53 [272], which affects its role in tumor suppression and DNA damage response. *BTG3* mRNA targeted by three miRNAs, and these binding sites did not overlap, and each of them could independently bind to mRNA. The mRNA of the *DKK3* gene had a cluster of binding sites from 73 to 115 nucleotides with an average ΔG = -124 kJ/mol.

The average free energy of interaction of all miRNA binding sites in the 5'UTR is -120 kJ/mol. The number of binding sites with a ΔG value above -130 kJ/mol is eight, including the associations of ID00792.3p-miR and *ALDH1A2*; ID00296.3p-miR, ID01778.3p-miR, ID01702.3p-miR and *ARID1A;* ID00961.3p-miR and *BARX1;* ID03332.3p-miR and *FOXF1, FOXP1.* The association of these miRNAs with their target genes are recommended as markers for the diagnosis of EAC.

Table 4 and A3 show the results of miRNA binding to mRNA of candidate EAC genes in CDS. The *AR* gene mRNA is targeted by eight miRNAs with binding sites in the CDS. Five of them formed a binding sites cluster in positions from 2484 nt to 2539 nt with average interaction energy of -127 kJ/mol with the length of 56 nt. The total length of all 24 sites was 552 nt. The formation of a cluster of binding sites for *AR* gene in the CDS indicates a greater ability of this gene for compaction, which causes the competition of miRNAs data for the binding site. The average binding energy for 5 miRNAs was -127 kJ/mol.

The mRNA of the *ARID1A* gene was a target of 28 miRNA in the CDS. Eight of them form a cluster of binding sites in position from 410 nt to 441 nt with an average ΔG equal to -129 kJ/mol. The second cluster of binding sites formed by ID01704.5p-miR, ID02761.3p-miR, ID01810.3p-miR, and ID02761.3p-miR takes place from 471

nt to 516 nt and the average free energy of interaction of these miRNAs equal to -128 kJ/mol.

Table 4 - The characteristics of miRNA interactions with CDS mRNAs of EAC candidate genes [268].

Gene; NX	miRNA	Position, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
1	2	3	4	5	6
AR; 2.2	ID03332.3p-miR	2484÷2511 (10)	-134	90	24
	ID01310.3p-miR	2485÷2512 (10)	-121	92	22
	ID02761.3p-miR	2514	-134	90	24
	ID00278.3p-miR	2515, 2518	-123, -125	89, 91	23
	ID01310.3p-miR	2518	-123	94	22
ARID1A; 23.2	miR-3960	390	-115	92	20
	ID02052.5p-miR	410	-132	89	24
	ID00522.5p-miR	410	-127	91	23
	ID02187.5p-miR	411	-127	92	23
	ID02692.5p-miR	413	-127	90	23
	ID00457.3p-miR	415	-125	92	22
	ID02064.5p-miR	417	-134	93	23
	ID02084.3p-miR	418	-138	92	24
	ID02538.3p-miR	420	-125	94	22
	ID01704.5p-miR	471	-123	89	23
	ID02761.3p-miR	487	-140	94	24
	ID01810.3p-miR	490	-115	89	23
	ID02761.3p-miR	493	-134	90	24
	ID00756.3p-miR	843	-125	91	23
	ID02294.5p-miR	851	-129	88	24
	ID00061.3p-miR	852	-125	91	22
	ID03167.3p-miR	1399	-123	91	23
	ID01610.5p-miR	1404	-110	91	21
	ID01388.5p-miR	4587	-119	89	23
	ID02687.5p-miR	4587	-119	92	23
	ID01565.5p-miR	4916	-115	93	21
	miR-6768-5p	4928	-115	89	23
ERBB2; 40.9	ID00692.3p-miR	3815	-113	90	22
	miR-4734	3817	-119	90	22
FOXF1; 6.6	ID02761.3p-miR	80	-132	89	24
	ID00278.3p-miR	81, 84	-123, -125	89, 91	23
	ID00811.3p-miR	87	-125	88	24
GDF7; 0.3	ID01508.5p-miR	699	-129	90	23
	ID03332.3p-miR	700, 703	-134	90	24
	ID01310.3p-miR	701	-121	92	22
	ID02761.3p-miR	703	-132	89	24
	ID02692.3p-miR	705	-127	91	23

Table 4 (continued).

1	2	3	4	5	6
	miR-1913	708	-115	90	22
IGFBP2; 49.0	ID01662.3p-miR	139	-129	91	23
	ID00071.3p-miR	142	-121	97	20
	ID00296.3p-miR	142	-140	89	25
	ID01702.3p-miR	142, 146	-134, -136	89, 90	24
	ID00522.5p-miR	143	-125	89	23
	ID01403.5p-miR	144	-125	92	23
	ID03367.5p-miR	145	-119	95	20
	ID02294.5p-miR	145, 153	-129, -132	88, 90	24
	ID02429.3p-miR	148	-123	91	23
MUC1; 51.8	ID00645.5p-miR	507	-110	93	20
	miR-6752-5p	510	-119	90	22
OXTR; 1.1	ID00009.3p-miR	1066	-115	92	20
	ID02770.5p-miR	1066	-115	92	20
	ID02036.3p-miR	1068	-115	92	20
	ID01310.3p-miR	1328	-123	94	22
PARP1; 18.1	ID00550.5p-miR	1265	-113	90	23
	ID01616.3p-miR	1275	-119	90	23

ID00756.3p-miR, ID02294.5p-miR and ID00061.3p-miR form a cluster of binding sites in *ARID1A* gene in position from 843 nt to 873 nt with an average ΔG = -127 kJ/mol. The mRNA of the *BARX1* gene was the target of five miRNAs (Table C3), two of which form a cluster of binding sites from 311 to 336 nucleotides with an average free hybridization energy of -127 kJ/mol. The mRNA of *BTG3, FOXM1, IGFBP2, OXTR, ROCK2* and *TBX5* in CDS have binding sites for individual miRNAs. ID01508.5p-miR, ID00102.3p-miR, ID01419.3p-miR, and ID01346.3p-miR target mRNA of the *GDF7* gene (Table C3). The mRNA of the *PAPR1* gene has a cluster of binding sites at positions from 1265 nt to 1297 nt with an average ΔG value of 116 kJ/mol.

The average free energy of binding of all miRNAs to mRNA in CDS was equal to -123 kJ/mol. Eight miRNAs bind to mRNAs of six target genes, forming 27 associations with a free interaction energy of more than -130 kJ/mol (Table 4 and A3), which allows us to recommend these miRNAs as markers for the diagnosis of EAC.

16 mRNA genes were targeted by miRNA in the 3'UTR (Table 5 and A4). The mRNA of the *FKBP5* gene was the target of five miRNAs (Table 5), forming clusters binding sites. miR-1273a, ID00367.5p-miR and miR-1273g-3p formed a binding sites cluster at positions from 4855 nt to 4897 nt with average free interaction energy of -113 kJ/mol. The second cluster was formed by ID01360.3p-miR, ID00367.5p-miR, and miR-1273g-3p at the 6364 nt position at 6394 nts with an average ΔG = -109 kJ/mol. The mRNAs *AR, FOXP1* and *RUNX3* were targets of single miRNAs (Table C4).

It is noteworthy that ID00037.3p-miR and ID00125.3p-miR occupy the same binding site, starting from 1671 nt in mRNA of the *SOX2* gene (Table C4). At the same ΔG/ΔGm values, the free interaction energy ID00037.3p-miR (ΔG = -121 kJ/mol) is higher than the ΔG ID00125.3p-miR value in mRNA of the *SOX2* gene.

Table 5 - The characteristics of miRNA interactions with 3'UTR mRNAs of EAC candidate genes [268].

Gene; NX	miRNA	Position, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
1	2	3	4	5	6
CTSE; 0.6	miR-574-5p	1474	-108	89	23
	ID00470.5p-miR	1482	-108	89	23
FKBP5; 19.6	miR-1273a	4855	-115	87	25
	ID00367.5p-miR	4870	-110	90	22
	miR-1273g-3p	4877	-113	96	21
	ID01360.3p-miR	6364	-104	91	21
	ID00367.5p-miR	6367	-110	90	22
	miR-1273g-3p	6374	-113	96	21
FOXP1; 3.6	miR-466	5950	-104	89	23
	miR-466	5952÷5968 (4)	-104	89	23
	ID00436.3p-miR	5952, 5970	-104	89	23
	ID01727.5p-miR	5953, 5955	-104, -106	89, 91	23
	ID02882.3p-miR	5960	-110	93	21
LEP; 0.0	ID03149.5p-miR	3087	-113	90	22
	ID01263.5p-miR	3092	-113	90	22
	miR-5095	3094	-110	95	21
	miR-619-5p	3100	-119	98	22
OXTR; 1.1	miR-574-5p	2501÷2513 (7)	-108÷-113	89	23
	ID00470.5p-miR	2505, 2511	-108, -113	89, 93	23
SEPP1; 20.0	miR-548aj-3p	1536	-89	91	21
	miR-548ae	1536	-98	96	21
	miR-548x-3p	1537	-89	93	20
SMAD4; 5.4	ID01838.5p-miR	4291	-113	90	24
	miR-1273g-3p	4312	-110	95	21
	ID01656.3p-miR	4342	-115	89	23
	miR-1273d	4346	-119	87	25
	ID01404.5p-miR	4349	-115	93	23
	ID02732.3p-miR	7721	-123	91	23
	ID00470.5p-miR	7744÷7752 (5)	-108	89	23
	miR-574-5p	7740÷7756 (8)	-108÷-110	89÷93	23
	ID00106.5p-miR	7825	-106	91	22
	miR-574-5p	7830	-108	89	23
TP53; 14.3	ID00548.3p-miR	1393	-115	89	23
	ID02379.3p-miR	1397	-119	89	24
	miR-1273c	2297	-110	91	22

Table 5 (continued).

1	2	3	4	5	6
	miR-1285-3p	2301	-110	95	22
	miR-1273g-3p	2317	-106	91	21
	miR-1273h-5p	2351	-106	91	21
	ID01838.5p-miR	2459	-115	92	24
	ID00785.5p-miR	2520	-113	90	23

The miR-548aj-3p, ae, x-3p formed a binding sites cluster in the mRNA of the *SEPP1* gene. The mRNA of the *SMAD4* gene is a target of 11 miRNAs in the 3'UTR, which form two clusters of binding sites. miR-1273g-3p, ID01656.3p-miR, miR-1273d, and ID01404.5p-miR form a cluster from 4291 nt to 4371 nt with an average value ΔG = -114 kJ/mol. The mRNA of the *LEP* gene has a binding sites of six miRNAs in the 3'UTR, four of which form a cluster of binding sites from 3087 nt to 3121 nt with an average value of ΔG = -114 kJ/mol.

The average free energy of binding of all miRNAs to mRNAs in the 3'UTR was -110 kJ/mol. The number of associations with a free binding energy of more than -120 kJ/mol is three. The association of these miRNAs with their target genes recommended as markers for the diagnosis of EAC.

Figure 7 shows the schemes of binding of miRNAs with mRNAs of EAC candidate genes. The interaction of nucleotides occurs along the entire length, except for the absence of hydrogen bonding between purines (A, G) or pyrimidines (C, U). Interactions of non-canonical pairs of nucleotides A-C and G-U are accounted by MirTarget program.

Gene; miRNA; The start of site; Region; The characteristics of binding	Gene; miRNA; The start of site; Region; The characteristics of binding

ARID1A; ID00414.3p-miR; 303 nt;
5'UTR; -108 kJ/mole; 93%; 20 nt
5' - GGG**G**CGGGGUGGGAGGGGG**G** - 3'
  | |||||||||||||||||
3' - CAC**U**GCCCCACCCUCCCCC**U** - 5'
	CDK9; ID02823.3p-miR; 25 nt;
	5'UTR; -117 kJ/mole; 90%; 22 nt
	5' - **C**GCG**G**CCGCGGAGGG**G**CCUGGA - 3'
	  ||||||||||| ||||||||||
	3' - **A**CGC**U**GGCGCCGCCC**U**GGACCU - 5'

AR; ID01261.5p-miR; 1744 nt;
CDS; -113 kJ/mole; 95%; 20 nt
5' - GAGCGAGGGAGGCCUCGGGG - 3'
  ||| ||||||||||||||||
3' - CUCCCUCCCUCCGGAGCCCC - 5'
	GDF7; ID01419.3p-miR; 876 nt;
	CDS; -115 kJ/mole; 89%; 23 nt
	5' - CGGGAG**G**GCUCCGGCCGGG**G**CAG - 3'
	  |||||||||||| |||| |||||
	3' - GCCCUC**U**CGAGGAUGGCAC**U**GUC - 5'

TP53; ID01838.5p-miR; 2460 nt;
3'UTR; -110 kJ/mole; 88%; 24 nt
5' - **G**AUGG**G**GUCUCACAGUGUUGC**C**CA - 3'
  || |||||||||| ||||||||||
3' - **U**UUCC**U**CAGAGUGACACAACG**A**GU - 5'
	ROCK2; ID01836.5p-miR; 6497 nt;
	3'UTR; -115 kJ/mole; 92%; 23 nt
	5' - G**C**CACUGCACUCCAGCCUGGGCA - 3'
	  ||||||||||||| ||||||||
	3' - A**A**GUGACGUGAGGUAGGACCCGU - 5'

Note: in bold type highlighted the non-canonical A-C and G-U pairs of nucleotides

Figure 7 - The schemes of miRNA interaction with mRNA of some genes participating in the development of EAC [268]

One of the misconceptions in many studies that miRNAs binds only (of predominantly) in 3'UTR of mRNAs [273]. However, miRNAs do not have the ability to distinguish binding sites in 5'UTR, CDS and 3'UTR. The miRNA interacts with mRNA based on the physico-chemical properties of these molecules. Consequently, the interaction site can be localized in any region of mRNA and prohibitions on the location of such sites in the nucleotide sequence of the mRNA are still unknown. The conditions for the successful interaction of miRNA with mRNA are the energy characteristics and conformational properties of this interaction.

The given results demonstrate the important role of non-canonical pairs of nucleotides in the interaction of miRNA with mRNA of candidate genes involved in the development of EAC. For example, in the interaction of ID01838.5p-miR with mRNA of *TP53* gene, one pair of A-C and two pairs of G-U are formed. ID01419.3p-miR has binding sites in mRNA of *GDF7* gene and forms two pairs of G-U.

The average free energy of miRNA binding in mRNA of genes involved in EAC development is greater in 5'UTR and CDS compared to 3'UTR, which suggests preferential binding of miRNA to 5'UTR and CDS of the studied genes.

The associations of miR-619-5p and *FKBP5, ERBB3;* ID03332.3p-miR and *AR, FOXF1, FOXP1, GDF7;* ID02761.3p-miR and *AR, ARID1A, FOXF1, GDF7*; miR-6081 and *ARID1A* chosen as recommended associations for EAC diagnosis. mRNA of these genes binds with miRNA with high affinity and, therefore, if the concentrations of the corresponding miRNAs are the same or exceed the concentrations of mRNAs, the effect of mRNA translation inhibition will be significant. Given that miRNAs are endogenous regulators of gene expression, these methods are quite feasible. The obtained results may be able to provide insights into pathogenesis mechanisms and pave the way for the development of new diagnostic markers and therapeutic targets for patients with esophageal adenocarcinoma.

3.2.1.2 The characteristics of the miRNAs interactions with mRNAs of genes associated with esophageal squamous cell carcinoma

Of the 66 candidate genes involved in the development of esophageal squamous cell carcinoma, 50 mRNA genes are targeted by 246 miRNAs (Tables 6-8, A5-A7).

The average free energy of binding of all miRNAs with mRNAs of all candidate genes in the 5'UTR region was equal to -121 kJ/mol. The number of miRNA associations with mRNA having free energy of interaction greater than -130 kJ/mol is 11 from 147 associations of miRNA-mRNA in the 5'UTR. mRNA of *BMI1* gene has clusters of binding sites in position from 160 nt to 195 nt and in position from 223 nt to272 nt. ID01481.5p-miR and ID02388.3p-miR take the binding site from 233 nt to 254 nt in mRNA of *BMP7* gene (Table C5). The biggest cluster of binding sites was identified in mRNA of *CCNE1* gene (Table 6), in position from 69 nt to 123 nt with a whole length equal to 55 nt and an average ΔG= -127 kJ/mol. The whole length of all four binding sites was equal to 268 nt, which is four times more than the length of a cluster. The formation of a cluster of four binding sites in the 5'UTR of *CCNE1* gene

shows the ability of a given gene to compaction, which serves as the emergence of given miRNAs competition for the binding site. That is, from here it can be concluded that ID00296.3p-miR with a free energy of interaction equal to -140 kJ/mol will occupy this binding site and will not allow other miRNAs to interact with this mRNA. Each miRNA has targeted various affinities and expression levels that essentially work as competitive inhibitors reducing free miRNA.

Table 6 - The characteristics of miRNA interactions with 5'UTR mRNAs of ESCC candidate genes [275].

Gene; NX	miRNA	Position, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
1	2	3	4	5	6
AURKA; 7.7	miR-5095	420	-108	93	21
	miR-619-5p	426	-119	98	22
BMI1; 16.0	ID03488.3p-miR	160	-119	89	23
	ID03028.3p-miR	173	-123	91	23
	ID01282.3p-miR	223, 228	-117, -121	89, 92	23
	ID02333.3p-miR	251	-117	90	22
BMP7; 17.2	ID01481.5p-miR	233	-117	92	22
	ID02388.3p-miR	234	-119	92	21
	ID01353.3p-miR	349	-119	90	22
	ID00602.3p-miR	353	-134	89	24
C2orf40; 16.5	ID03277.5p-miR	7	-108	93	20
	ID00049.5p-miR	12	-140	93	24
CCNE1; 2.5	ID02256.3p-miR	69	-125	91	22
	ID00296.3p-miR	71	-140	89	25
	ID01873.3p-miR	74	-121	92	21
	ID00061.3p-miR	80, 95	-125, -129	91, 94	22
	ID01804.3p-miR	80	-132	90	23
	ID03367.5p-miR	80, 83	-115	92	20
	ID02064.5p-miR	100	-132	91	23
	ID02084.3p-miR	100	-134	89	24
	ID02429.3p-miR	101	-123	91	23
EGR1; 76.5	miR-4787-5p	12	-121	90	22
	ID01041.5p-miR	13	-132	90	24
	ID01804.3p-miR	13	-132	90	23
	ID02187.5p-miR	14	-125	91	23
	ID00457.3p-miR	15	-123	91	22
ITGA3; 12.8	ID02648.5p-miR	71	-123	89	23
	ID03489.5p-miR	77	-125	89	24
MAL; 447.4	ID01730.3p-miR	70	-113	90	22
MMP2; 40.2	ID00278.3p-miR	110	-123	89	23
	ID01310.3p-miR	113	-121	92	22
	ID03037.3p-miR	115	-121	90	22

Table 6 (continued).

1	2	3	4	5	6
	ID03280.3p-miR	115	-121	90	22
	ID03345.5p-miR	124	-127	90	24
	ID03368.3p-miR	125	-117	89	23
	ID02014.5p-miR	264	-117	89	23
SASH1; 35.6	ID01607.3p-miR	403	-117	89	23
	ID02116.3p-miR	412	-125	91	24

The experimentally verified interaction between miR-619-5p and mRNA of *AURKA* gene shown by Heng et al [274]. The MirTarget program confirmed the presence of their interaction in 5'UTR (Table 6) with ΔG/ΔGm value equal to 98%, and identified more than three alternative miRNAs that can bind to the 5'UTR of the *AURKA* gene mRNA (Table 6 and Table C5).

mRNAs of *ADH1B* was targeted by segregate miRNAs (Table C5). mRNA of *MAL* gene (table 6) having the high value of NX=447,4 can interact only with one miRNA (ID01730.3p-miR). The mRNA of *EGR1* gene has five binding sites in 5'UTR that form a cluster from 12nt to 36 nt with average ΔG value equal to -128kJ/mol. mRNA of *ITGA3* gene have cluster of binding sites from 71 nt to 100 nt with an average ΔG = -124 kJ/mol (Table C5).

The analogous binding sites were identified in CDS (Table 7 and A6). The average free energy of interaction of all miRNAs with mRNAs in CDS was equal to -117 kJ/mol. The number of miRNA associations with mRNA having free energy of interaction greater than -130 kJ/mol is 21.

In mRNA of *ADAMTS2* gene it was identified the cluster of binding sites in CDS from 153 nt to 209 nt with a whole length equal to 57 nt and average ΔG = -130 kJ/mol. The highest ΔG value was detected in associations ofID00296.3p-miR and ID01702.3p-miR (ΔG = -142 kJ/mol) with mRNA of *ADAMTS2* gene. Another cluster located from 217 nt to 244 nt with an ΔG value equal to -122 kJ/mol. ID03270.3p-miR, miR-1226-5p and miR-4687-3p forms cluster of binding sites from 153 nt to 191 nt in mRNA of *ALDH2* gene. In mRNA of *CCND1* gene identified a cluster of binding sites from 1021 nt to 1047 nt with average ΔG value equal to -115 kJ/mol. ID02888.5p-miR and ID00848.5p-miR form a cluster of binding sites in mRNA of *COL7A1* gene in position from 4669 nt to 4699 nt with a length 31 nt and average ΔG = -121 kJ/mol (Table 7). mRNA of *ELN* gene had binding sites for 12 miRNAs in CDS, nine of them form a cluster of binding sites in position from 1949 nt to 1980 nt with a whole length 32 nt and average ΔG value equal to -124 kJ/mol.

Table 7 - The characteristics of miRNA interactions with CDS mRNAs of ESCC candidate genes [275, 276]

Gene; NX	miRNA	Position, nt	Energy, kJ/mol	Score, %	Length, nt
1	2	3	4	5	6
ADAMTS2; 2.2	ID02294.5p-miR	153÷177 (3)	-129 ÷ -136	88÷93	24
	ID00296.3p-miR	175, 178	-140, -142	89, 91	25
	ID01641.3p-miR	178	-132	89	24
	ID01873.3p-miR	179	-121	92	21
	ID00061.3p-miR	181	-125	91	22
	ID01702.3p-miR	181	-142	94	24
	ID02064.5p-miR	182, 186	-129	90	23
	miR-3960	183	-117	93	20
	ID02429.3p-miR	184	-121	89	23
	ID03151.3p-miR	185	-115	93	20
	ID02036.3p-miR	186	-117	93	20
	ID02084.3p-miR	186	-134	89	24
	ID02084.3p-miR	217	-134	89	24
	ID01315.3p-miR	218	-115	92	20
	ID01157.5p-miR	219	-119	95	20
	ID02538.3p-miR	219	-123	92	22
	ID01377.3p-miR	220	-123	97	20
	ID03404.5p-miR	223	-119	90	22
ALDH2; 84.0	ID03270.3p-miR	153	-108	91	21
	miR-1226-5p	156	-127	86	26
	miR-4687-3p	171	-110	91	21
CCND1; 41.1	miR-877-3p	1021	-108	93	21
	ID03238.3p-miR	1022	-119	92	23
	ID01352.3p-miR	1024	-115	90	23
	ID00777.3p-miR	1025	-113	90	23
COL7A1; 11.3	ID02888.5p-miR	4669	-123	91	23
	ID00848.5p-miR	4677	-119	89	23
	miR-483-3p	5556	-104	91	21
	ID00597.3p-miR	5567	-110	90	23
COL27A1; 3.7	ID02915.3p-miR	20	-115	89	23
	ID02611.3p-miR	35	-125	91	22
	ID02761.3p-miR	36, 39	-132, -134	89, 90	24
	ID01643.3p-miR	38	-119	90	22
	ID02585.5p-miR	40	-127	88	24
	ID03409.5p-miR	41	-115	93	20
EGR1; 76.5	ID03387.3p-miR	467	-127	90	24
	ID02200.3p-miR	470	-121	92	22
	ID03445.3p-miR	470	-121	89	23
	ID01859.5p-miR	475	-117	90	22
ELN; 24.3	ID00517.5p-miR	1949	-115	92	20

Table 7 (continued).

1	2	3	4	5	6
	ID00722.5p-miR	1953	-113	93	20
	ID02052.5p-miR	1953	-134	90	24
	ID01041.5p-miR	1954	-136	93	24
	ID01804.3p-miR	1954	-132	90	23
	ID01873.3p-miR	1954	-123	94	21
	ID01445.3p-miR	1956	-115	92	20
	ID03151.3p-miR	1956	-115	93	20
	ID00061.3p-miR	1959	-125	91	22
HOXA10; 0.3	ID00558.5p-miR	422	-113	90	22
	ID01541.3p-miR	424	-121	89	23
	ID03305.5p-miR	441	-113	93	20
IGFBP3; 11.7	ID00267.3p-miR	423	-113	91	21
	ID02982.3p-miR	435	-123	100	21
	ID02191.5p-miR	436	-110	91	21
SIX4; 0.0	ID02296.5p-miR	339	-117	95	20
	ID01041.5p-miR	341	-132	90	24
	miR-3960	343	-115	92	20
	ID02692.5p-miR	344	-129	91	23
	ID01702.3p-miR	345	-134	89	24
	ID02296.5p-miR	345	-115	93	20
	ID02084.3p-miR	346	-138	92	24
	ID02538.3p-miR	348	-121	90	22
	ID01879.5p-miR	352	-127	94	22

The mRNAs of certain genes from the matrix metalloproteinase family, like *MMP3* and *MMP11*, were targeted by individual miRNAs. Conversely, *MMP2* displayed a cluster of binding sites located within its CDS spanning from 1681 nt to 1712 nt (refer to Table C6). Similarly, the mRNA of *MMP9* was targeted by ID01341.5p-miR and ID01104.5p-miR within its CDS region. The mRNA of *IGFBP3* gene revealed a binding site with a ΔG/ΔGm value of 100%, forming part of a larger cluster of binding sites positioned from 423 nt to 456 nt with an average ΔG of -115 kJ/mol. In the case of the ITGB4 gene, clusters of binding sites were identified spanning from 2335 nt to 2360 nt and from 4431 nt to 4470 nt, exhibiting an average ΔG value of -114 kJ/mol (refer to Table C6). Additionally, the mRNA of the *SIX4* gene contained a cluster of binding sites located from 339 nt to 373 nt.

Considering the average free energy of binding between miRNAs and all mRNAs within the 3'UTR, it was found to be -111 kJ/mol (as documented in Table 8 and Table C7).

Table 8 - The characteristics of miRNA interactions with 3'UTR mRNAs of ESCC candidate genes [276].

Gene; NX	miRNA	Position, nt	Energy, kJ/mol	Score, %	Length, nt
1	2	3	4	5	6
CCND1; 41.1	miR-574-5p	2593÷2597 (3)	-108÷-113	89÷93	23
	ID00470.5p-miR	2595, 2597	-108	89	23
	ID01543.3p-miR	2963	-117	92	22
	ID00529.5p-miR	2972	-113	95	21
COL1A1; 26.9	miR-328-5p	5014	-125	91	23
	ID02647.3p-miR	5023	-121	93	21
ELN; 24.3	miR-6734-3p	2592	-113	90	23
	miR-6809-3p	2592	-102	91	21
	miR-877-3p	2592	-106	91	21
GCOM1; 22.4	miR-1273a	2456	-119	90	25
	miR-1273g-3p	2478	-113	96	21
MTHFR; 3.5	ID00345.3p-miR	6281	-119	89	23
	miR-5585-3p	6300	-108	93	22
	ID01811.5p-miR	6844	-117	93	22
	miR-5095	6855	-110	95	21
	miR-619-5p	6861	-115	95	22
ITGA3; 12.8	ID00264.5p-miR	3814	-123	89	24
	ID01921.5p-miR	3820	-104	91	21
SASH1; 24.5	ID03224.5p-miR	5498	-117	89	23
	ID01360.3p-miR	5502	-106	93	21
	miR-1273g-3p	5511	-108	93	21
	ID01656.3p-miR	5541	-115	89	23
	miR-7851-3p	5547	-108	91	22
	ID01103.3p-miR	5707	-129	90	24
	miR-3929	5711	-110	90	23
SIX4; 0.0	miR-1273a	4029	-115	87	25
	ID01838.5p-miR	4030	-113	90	24
	miR-5684	4045	-98	92	20
	miR-1273g-3p	4051	-113	96	21
	ID02017.3p-miR	4287	-115	90	22
	miR-1972	4291	-113	91	22
SOX2; 16.9	ID00037.3p-miR	1671	-121	90	23
	ID00125.3p-miR	1671	-113	90	22
TP53; 14.3	ID00548.3p-miR	1393	-115	89	23
	ID02379.3p-miR	1397	-119	89	24
	miR-1273c	2297	-110	91	22
	miR-1285-3p	2301	-110	95	22
	miR-1273g-3p	2317	-106	91	21
	miR-1273h-5p	2351	-106	91	21
	ID01838.5p-miR	2459	-115	92	24

Among the various miRNA associations with mRNAs, three demonstrated an interaction free energy greater than -125 kJ/mol. ID00470.5p-miR and miR-574-5p exhibited multiple binding sites in the mRNA of the *CCND1* gene, specifically positioned from 2595 nt to 2619 nt. The mRNA of the ITGA3 gene also displayed a cluster of binding sites, positioned from 3814 nt to 3841 nt, with an average ΔG of -114 kJ/mol.

Additionally, ID00125.3p-miR and ID00037.3p-miR shared the same binding site within the mRNA of the *SOX2* gene. Another notable association was between ID00548.3p-miR and ID02379.3p-miR, forming a cluster of binding sites from 1393 nt to 1420 nt within the mRNA of TP53, demonstrating an average ΔG of -117 kJ/mol.

The mRNA of the *MTHFR* gene contains two distinct clusters of miRNA binding sites. The first cluster spans from 6281 nt to 6321 nt, while the second begins at 6844 nt, with a total length of 39 nt and an average ΔG value of -114 kJ/mol.

A total of 223 potential binding sites for 203 miRNAs have been identified, influencing 47 out of 66 genes implicated in the development of ESCC. These binding sites are organized in clusters located within the 5'UTR, CDS, and 3'UTR regions of the mRNA. Notably, miRNA and mRNA associations exhibiting an interaction free energy equal to or greater than -130 kJ/mol could potentially serve as early diagnostic markers for this disease.

In examining the average free energy of miRNA-mRNA binding, it was observed that in the 5'UTR and CDS regions, it is higher compared to the 3'UTR, implying a preference for miRNA binding to the 5'UTR and CDS of the studied genes associated with ESCC. Moreover, specific regions within the clusters of miRNA binding sites feature overlapping sequences, resulting in a compact arrangement of binding sites in the mRNA. This compression significantly reduces the overall number of binding sites available in the mRNA.

The existence of overlapping miRNA binding sites creates a competitive environment among miRNAs for each binding site. In such scenarios, the RISC complex interacting with mRNA possessing stronger free energy will potentially hinder the binding of other RISC complexes with miRNAs exhibiting weaker interactions with mRNA.

Associations such as miR-619-5p and *AURKA, MTHFR*; ID00296.3p-miR and *ADAMTS2, C2orf40, CCNE1*; ID01041.5p-miR and *EGR1, ELN, SIX4;* ID01702.3p-miR and *ADAMTS2, SIX4*; ID01804.3p-miR and *CCNE1, EGR1, ELN* are suggested for use in the diagnosis and treatment of ESCC.

3.2.2 The characteristics of the miRNAs interactions with mRNAs of genes associated with gastric cancer

The study involved investigating the interaction characteristics between 6274 miRNAs and the mRNA of 215 candidate genes associated with gastric cancer (as detailed in Table C1). Table 9 specifically outlines the quantitative features of

miRNA binding in the 5'UTR of mRNA genes displaying clusters of miRNA binding sites.

Table 9 - The characteristics of the miRNA interactions with 5′UTR mRNAs of gastric cancer candidate genes [277].

Gene; NX	miRNA	Start of site, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
1	2	3	4	5	6
ARID1A; 19.5	ID02106.3p-miR	136	-123	89	23
	ID01778.3p-miR	140	-134	90	24
	ID00296.3p-miR	141,166	-138	88	25
	miR-6081	143	-125	89	24
	ID01702.3p-miR	147	-134	89	24
	ID00465.5p-miR	148	-113	93	20
	ID01377.3p-miR	152	-117	92	20
	ID02592.5p-miR	164, 167	-123	89	23
	miR-3960	167	-115	92	20
	ID03065.3p-miR	172	-115	92	21
E2F1; 3.4	ID02574.3p-miR	84	-115	93	20
	ID02052.5p-miR	85	-149	100	24
	ID01873.3p-miR	87	-123	94	21
	miR-3960	90	-115	92	20
	ID00071.3p-miR	90	-117	92	20
	ID00722.5p-miR	90	-113	93	20
	ID02064.5p-miR	95	-132	91	23
ODC1; 19.2	ID00756.3p-miR	9	-129	94	23
	ID01804.3p-miR	13	-132	90	23
	ID02187.5p-miR	14	-123	89	23
	ID00457.3p-miR	15, 21	-123, -125	91, 92	22
	ID02084.3p-miR	17	-140	93	24
	ID02064.5p-miR	17÷23(3)	-129÷ -140	90÷97	23
	miR-3960	18÷21(3)	-115	92	20
	ID01652.3p-miR	19	-125	89	23
	ID02538.3p-miR	19	-123	92	22
	ID01702.3p-miR	19, 22	-134, -142	89, 94	24
	ID02229.3p-miR	21	-121	92	21
	ID02499.3p-miR	21	-119	92	21
	ID01157.5p-miR	22	-117	93	20
	ID01377.3p-miR	23	-121	95	20
	ID00061.3p-miR	24	-125	91	22
PIK3CA; 10.6	ID01190.5p-miR	4	-140	92	24
	ID01702.3p-miR	4	-140	93	24

Table 9 (continued).

1	2	3	4	5	6
	ID01895.5p-miR	4	-132	89	24
	ID01641.3p-miR	5	-127	86	24
	ID00966.5p-miR	6	-136	90	24
	ID00030.3p-miR	7	-121	90	22
	ID02294.5p-miR	7	-125	86	24
	ID01804.3p-miR	8	-134	91	23
	ID01873.3p-miR	8	-123	94	21
	ID02064.5p-miR	9	-127	88	23
	ID02084.3p-miR	9	-129	86	24
TBC1D9; 10.0	ID01895.5p-miR	125	-134	90	24
	ID02187.5p-miR	130	-127	92	23
	ID03229.5p-miR	130, 133	-121	90	22
	ID01041.5p-miR	132	-132	90	24
	ID01702.3p-miR	136	-140	93	24
	ID02084.3p-miR	137	-136	90	24
	ID00457.3p-miR	138	-125	92	22

For instance, the *ARID1A* gene encodes a protein that is part of the large ATP-dependent chromatin-remodeling complex SNF/SWI. This complex is crucial for the transcriptional activation of genes typically repressed by chromatin formation. Within the mRNA of the *ARID1A* gene, a cluster of ten miRNA binding sites is positioned from 136 nt to 192 nt.

The collective length of these miRNA binding sites sums up to 272 nt, compacted within a 57-nt cluster of BS, indicating approximately 4.8-fold compaction of binding sites in this mRNA segment. This leads to competition among miRNAs when they bind to this cluster. Preferential binding occurs for miRNAs interacting with the mRNA at a higher free energy and being present in higher concentrations compared to competing miRNAs.

The average free energy of interaction among all miRNAs in the 5'UTR mRNA of candidate genes (as shown in Table 9 and Table C8) stands at approximately -126 kJ/mol. Based on this observation, a proposed criterion for effective miRNA interaction with 5'UTR mRNA in the context of diagnosing gastric cancer is a value of -130 kJ/mol or higher. Regarding the hypothesis for diagnosing gastric cancer, it's suggested that associations exist between ID01778.3p-miR, ID00296.3p-miR, ID01702.3p-miR, and the *ARID1A* gene. Consequently, when developing a diagnostic chip, it's crucial to incorporate all nine miRNAs to account for changes in their concentrations. This factor significantly influences the competitive interactions they engage in with mRNA.

The nucleotide sequences of the miRNA binding site cluster at the *ARID1A* gene in various species—*G. gorilla, H. sapiens, M. mulatta, M. nemestrina, N. leucogenys, P. anubis, P. paniscus, and P. troglodytes*—are represented in Figure 8. Notably, the

flanking oligonucleotides from the 5' end (GGGAGCAGCUG) and 3' end (GAGCCUGAGCCGG) remain identical across all these subjects.

```
GGGAGCAGCUGAGCCGCCGGCGCCUCGGCCGCCGCCGCCGCCUCCUCCUCCUCCGCCGCCGCCAGCCCGGAGCCUGAGCCGG hsa
GGGAGCAGCUGAGCCGCCGGCGCCUCGGCCGCCGCCGCC---UCCUCCUCCUCCGCCGCCGCCAGCCCGGAGCCUGAGCCGG ggo
GGGAGCAGCUGAGCCGCCGGCGCCUCGGCCGCCGCCGCC---UCCUCCUCCUCCGCCGCCGCCAGCCCGGAGCCUGAGCCGG ptr
GGGAGCAGCUGAGCCGCCGGCGCCUCGGCCGCCGCCGCC---UCCUCCUCCUCCGCCGCCGCCAGCCCGGAGCCUGAGCCGG ppa
GGGAGCAGCUGAGCCGCCGGCGCCUCGGCCGCCGCCGCC---UCCUCCUCCUCCGCCGCCGCCAGCCCGGAGCCUGAGCCGG pan
GGGAGCAGCUGAGCCGCCGGCGCCUCGGCCGCCGCCGCC---UCCUCCUCCUCCGCCGCCGCCAGCCCGGAGCCUGAGCCGG nle
GGGAGCAGCUGAGCCGCCGGCGCCUCGGCCGCCGCCGCC---UCCUCCUCCUCCGCCGCCGCCAGCCCGGAGCCUGAGCCGG mml
GGGAGCAGCUGAGCCGCCGGCGCCUCGGCCGCCGCCGCC---UCCUCCUCCUCCGCCGCCGCCAGCCCGGAGCCUGAGCCGG mne
```

Figure 8 - The nucleotide sequences of mRNA regions of orthologous *ARID1A* genes containing clusters of binding sites of ten miRNAs [277].

These findings suggest the emergence and persistence of these miRNA and target gene associations across many millions of years among these species. This longevity indicates the conservation of these specific associations over evolutionary timescales. Such persistence provides an opportunity to investigate the role of the candidate *ARID1A* gene in these studied subjects as potential models for understanding gastric cancer. Comparing these sequences across different species allows researchers to explore the significance and potential implications of these associations in the context of gastric cancer evolution and pathology.

The transcription factor E2F1 plays a role in the development of several oncological diseases, including gastric cancer. Within the *E2F1* mRNA, there exists a 35 nucleotide-long cluster encompassing seven miRNA binding sites (as detailed in Table 9). Among these sites, two exhibit ΔG values exceeding -130 kJ/mol, and the ID02052.5p-miR pair displays complete complementarity with high free energy. The cumulative length of the binding sites for these seven miRNAs is 148 nucleotides, which is 4.2 times the length of the cluster. These associations between miRNAs and the E2F1 gene hold potential as markers for diagnosing gastric cancer.

One way to validate the binding between miRNA and target genes is by examining the presence of these binding sites in orthologous genes. Figure 9 illustrates the nucleotide sequences of the binding site cluster within the mRNA of orthologous *E2F1* genes. Notably, the flanking oligonucleotides from the 5' end (GGCCUGCCGC) and 3' end (GGCCGCGCGG) remain identical across all subjects. Substitutions of C for U generally do not significantly affect the free energy of interaction between miRNA and mRNA. This is because when C is substituted with U, non-canonical pairs of G-U are formed instead of G-C, which may have a minimal impact on the overall binding energy between the miRNA and mRNA.



Figure 9 - The nucleotide sequences of mRNA regions of orthologous *E2F1* genes containing clusters of binding sites of three miRNAs [277].

The mRNA of the *ODC1* gene displays a cluster comprising 15 miRNA binding sites, ranging from 9 to 46 nt (as depicted in Table 9). However, this 38-nt cluster length is notably 10.6 times shorter than the combined length of all binding sites. Consequently, only one miRNA among the 15 present within this cluster can bind effectively to the mRNA. Specifically, the miRNA capable of binding might be ID01804.3p-miR, ID02084.3p-miR, ID02064.5p-miR, or ID01702.3p-miR, given their potential interactions with a ΔG value exceeding -130 kJ/mol. If other miRNAs are present at significantly higher concentrations compared to these four miRNAs, they will ultimately influence the translation of the *ODC1* gene's mRNA.

Similarly, the mRNA of the *PIK3CA* gene demonstrates a cluster containing 12 miRNA binding sites spanning a length of 33 nucleotides. This length is 8.5 times shorter than the total combined length of these miRNAs' binding sites, which sums up to 282 nucleotides. This high compaction of binding sites creates intense competition among these miRNAs. Six out of the 12 miRNAs exhibit free interaction energy values exceeding -130 kJ/mol. Consequently, it's advisable to consider these six miRNAs' association with the *PIK3CA* gene as potential markers for diagnosing gastric cancer.

Furthermore, the mRNA of the *TBC1D9* gene also presents a cluster of seven miRNAs within a 37-nucleotide length. This length is five times shorter than the combined length of these miRNAs' binding sites, allowing only one miRNA to effectively bind. In orthologous genes of *TBC1D9*, a CGC insert was observed in humans within the clusters of miRNA binding sites (as shown in Figure 10).

```
CGCGGGCCACCCCCUGAAGUCGCCUGCCGCCGCCGCCGCCGCACCUAGCGGACGGG hsa
CGCGGGCCACCCCCUGAAGUCGCCUGCCGCCGCCGCCGCC---ACCUAGCGGACGGG ggo
CGCGGGCCACCCCCUGAAGUCGCCUGCCGCCGCCGCCGCC---ACCUAGCGGACGGG ptr
CGCGGGCCACCCCCUGAAGUCGCCUGCCGCCGCCGCCGCC---ACCUAGCGGACGGG ppa
CCCGGGCCGCCCCCUGAAGUCGCCUGCCGCCGCCGCCGCC---ACCUAGCGGACGGG pab
CGCGGGCCGCCCCCUGAAGUCGCCUGCCGCCGCCGCCGCC---ACCUAGCGGACGGG nle
CGCAGGCCGCCCCCUGAAGUCGCCUGCCGCCGCCGCCGCC---ACCUAGCGGACGGG pan
CGCAGGCCGCCCCCUGAAGUCGCCUGCCGCCGCCGCCGCC---ACCUAGCGGACGGG mml
CGCAGGCCGCCCCCUGAAGUCGCCUGCCGCCGCCGCCGCC---ACCUAGCGGACGGG mne
CGCAGGCCGCCCCCUGAAGUCGCCUGCCGCCGCCGCCGCC---ACCUAGCGGACGGG mfa
```

Figure 10 - The nucleotide sequences of mRNA regions of orthologous *TBC1D9* genes containing clusters of seven miRNA binding sites [277].

While the 3'-end flanking oligonucleotides (AGCGGACGGG) remain identical across all subjects, the differences in the 5'-end flanking oligonucleotides suggest distinct functional characteristics of the binding site cluster and the surrounding nucleotide sequences among these species.

To accurately evaluate the outcome of miRNA and mRNA interaction, it's crucial to understand both the quantitative characteristics of their interaction and the ratio between the concentrations of these molecules. This information is vital for predicting their interactions by analyzing the impact of the inhibitor (miRNA) on the translation process's activity. When a miRNA, possessing a single binding site, interacts with mRNA, the result can vary depending on the duration of their interaction. It might be reversible if the interaction is short-lived, irreversible if the bond lasts relatively longer, or it could even lead to mRNA degradation. The consequence of this interaction relies on the ratio between the concentrations of miRNA and mRNA. An increase in miRNA concentration will proportionally decrease protein synthesis. For instance, at a miRNA to mRNA concentration ratio of 1:10, protein synthesis might reduce by 10%. Conversely, at a ratio of 10:1, protein synthesis could be suppressed by 90%. Therefore, determining the ratio of miRNA and mRNA gene concentrations using chips becomes essential.

Within the coding sequence (CDS) of mRNAs associated with gastric cancer candidate genes, clusters of binding sites for multiple miRNAs are observed (as detailed in Table 10 and Table C9). Specifically, in the CDS of the *ARID1A* gene, three clusters of binding sites within the protein coding region are identified (as shown in Table 10). The first cluster, spanning from 410 nt to 442 nt, encompasses binding sites for eight miRNAs. Notably, the cluster's length is 32 nucleotides, significantly shorter (5.8 times less) than the combined lengths of all binding sites (184 nt). This compression implies that only one miRNA can effectively bind within this cluster. Among the eight miRNAs, three bind with a ΔG value equal to or greater than -130 kJ/mol, indicating robust interaction of these miRNAs' association members with the *ARID1A* gene.

Table 10 - The characteristics of miRNA interactions with CDS mRNAs of gastric cancer candidate genes [277].

Gene; NX	miRNA	Start of site, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
ARID1A; 19.5	ID02052.5p-miR	410	-132	89	24
	ID00522.5p-miR	410	-127	91	23
	ID02187.5p-miR	411	-127	92	23
	ID02692.5p-miR	413	-127	90	23
	ID00457.3p-miR	415	-125	92	22
	ID02064.5p-miR	417	-134	93	23
	ID02084.3p-miR	418	-138	92	24
	ID02538.3p-miR	420	-125	94	22
	ID01704.5p-miR	471	-123	89	23
	ID02761.3p-miR	487, 493	-134, -140	90, 94	24
	ID00756.3p-miR	843	-125	91	23
	ID02294.5p-miR	851	-129	88	24
	ID00061.3p-miR	852	-125	91	22
DNM3OS; 5.3	miR-4506	3149	-100	94	20
	miR-3120-3p	5958	-113	100	21
	miR-3120-5p	5996	-115	100	21
E2F1; 3.4	ID02051.3p-miR	291	-153	100	24
	ID03448.3p-miR	292	-123	91	22
	ID01157.5p-miR	295	-117	93	20
EPHB2; 2.3	ID00714.3p-miR	155	-127	91	24
	ID01403.5p-miR	169	-121	89	23
	miR-4253	1089	-102	100	18
MAPK1; 17.2	ID03332.3p-miR	243	-134	90	24
	ID01310.3p-miR	244	-121	92	22
	ID00798.3p-miR	246	-136	91	24
	ID01546.5p-miR	246	-132	90	24
PPP2CA; 21.8	miR-3661	462	-123	100	22
TERT; 0.4	ID01098.3p-miR	3224	-125	89	24
	ID01338.5p-miR	3236	-132	91	24
	ID01816.3p-miR	3237	-138	92	24
VEGFC; 0.4	ID02052.5p-miR	498	-132	89	24
	ID02187.5p-miR	500	-123	89	23
	ID01041.5p-miR	501	-132	90	24
	ID01873.3p-miR	501	-123	94	21
	ID00457.3p-miR	503	-127	94	22
	miR-3960	504	-115	92	20
	ID02064.5p-miR	505	-129	90	23
ZNF717; 14.0	miR-4273	1639	-96	100	18
	miR-4273	1891	-91	96	18

The nucleotide sequences forming the cluster of binding sites exhibit identical

characteristics among the mRNA orthologous genes in *G. gorilla, H. sapiens, M. mulatta, M. nemestrina, N. leucogenys, P. anubis, P. paniscus, and P. troglodytes* (as depicted in Figure 11). Both the 5'-end and 3'-end flanking oligonucleotides remain consistent across all these objects. Additionally, the oligopeptide LGNPPPPPPS, encoded by these binding site clusters, showcases identical attributes in all ARID1A orthologous proteins.

The strategic placement of miRNA binding sites in both the 5'UTR and at the CDS leads to an early halt in protein synthesis. This mechanism conserves energy compared to terminating protein synthesis at the end, enhancing the likelihood of forming an incomplete or abortive protein product. This strategic placement seems to aid in regulating protein synthesis efficiently, influencing the process at its initial stages and potentially preventing the unnecessary expenditure of resources in synthesizing proteins that might not fulfill their intended function.

```
GCCGCCAGCAGCCUGGGCAACCCGCCGCCGCCGCCGCCCUCGGAGCUGAAGAAAGC hsa
GCCGCCAGCAGCCUGGGCAACCCGCCGCCGCCGCCGCCCUCGGAGCUGAAGAAAGC ggo
GCCGCCAGCAGCCUGGGCAACCCGCCGCCGCCGCCGCCCUCGGAGCUGAAGAAAGC pab
GCCGCCAGCAGCCUGGGCAACCCGCCGCCGCCGCCGCCCUCGGAGCUGAAGAAAGC ptr
GCCGCCAGCAGCCUGGGCAACCCGCCGCCGCCGCCGCCCUCGGAGCUGAAGAAAGC mml
GCCGCCAGCAGCCUGGGCAACCCGCCGCCGCCGCCGCCCUCGGAGCUGAAGAAAGC mne
GCCGCCAGCAGCCUGGGCAACCCGCCGCCGCCGCCGCCCUCGGAGCUGAAGAAAGC mfa
GCCGCCAGCAGCCUGGGCAACCCGCCGCCGCCGCCGCCCUCGGAGCUGAAGAAAGC nle
```

Figure 11 - The nucleotide sequences of mRNA regions of orthologous *ARID1A* genes containing clusters of eight miRNA binding sites [277].

The mRNA of the *ARID1A* gene reveals a second cluster hosting binding sites for ID01704.5p-miR and ID02761.3p-miR, consisting of two binding sites with ΔG values exceeding -130 kJ/mol. Additionally, a third cluster incorporates binding sites for three miRNAs with a free interaction energy of miRNA above -125 kJ/mol. This emphasizes the substantial reliance of ARID1A protein synthesis on the expression of these 13 miRNAs binding in the 5'UTR mRNA. Given the pivotal role of this protein within cells, maintaining a significantly lower miRNA concentration than mRNA is crucial for synthesizing the requisite amount of protein.

Research by Wang and colleagues highlighted the oncogenic role of *DNM3OS*, an lncRNA implicated as a promising prognostic factor and potential therapeutic target in gastric cancer patients [278]. Within the *DNM3OS* gene, miR-3120-3p and miR-3120-5p form fully complementary binding sites in the lncRNA, while miR-4506 also establishes binding sites in the CDS region [261].

E2F1, a pivotal transcription factor within the E2F family, orchestrates critical functions in cell cycle regulation and tumor suppressor protein activity. Notably, the mRNA of the *E2F1* gene houses a cluster of binding sites for three miRNAs, among which ID02051.3p-miR exhibits complementarity across all nucleotides with a high

ΔG value (as indicated in Table 10). Consequently, the association involving ID02051.3p-miR binding in the CDS mRNA of this gene is suggested for diagnosing gastric cancer. Figure 12 depicts the nucleotide sequences of mRNA regions in orthologous *E2F1* genes encompassing clusters of binding sites for three miRNAs. Across various species, conserved flanking nucleotide sequences at both the 5'-end and 3'-end are evident. Despite substitutions in certain positions across different species, the encoding of specific oligopeptides remains consistent due to the replacement of canonical pairs with non-canonical pairs, signifying that single nucleotide polymorphisms scarcely alter the affinity of miRNA to the target gene's mRNA.



Figure 12 - The nucleotide sequences of CDS mRNA regions of orthologous *E2F1* genes containing clusters of binding sites of three miRNAs [277].

The heightened expression of *EPHB2* in tumor tissues, observed in both gene and protein expression, correlates with poor overall survival [279]. An experimentally established interaction between miR-4253 and the *EPHB2* gene has been confirmed, with MirTarget indicating the presence of a fully complementary interaction in the CDS. Additionally, two alternative miRNAs were identified by the program capable of binding to the CDS of *EPHB2* gene mRNA. The quantitative characteristics of the interaction between the gene's mRNA and ID00714.3p-miR, ID01403.5p-miR, and miR-4253 are outlined in Table 10.

The protein encoded by MAPK1 is involved in crucial cellular functions such as proliferation, differentiation, transcription regulation, and development. Within the CDS mRNA of the *MAPK1* gene, a compact 28 nt long cluster of miRNA binding sites exists, significantly shorter (2.9 times) than the sum of all miRNA binding site lengths (as specified in Table 10). Among the four miRNAs, three exhibit a ΔG value of more than -130 kJ/mol upon binding. These characteristics of the associations involving these three miRNAs with the *MAPK1* gene support their recommendation for gastric cancer diagnosis.

The cluster of binding sites within the mRNA of the human *MAPK1* gene (depicted in Figure 13) encodes an oligopeptide AAAAAAGAG, which is identical across P. *anubis, P. paniscus, P. troglodytes, M. mulatta, M. nemestrina, N. leucogenys, P. abelii, and G. gorilla.* Notably, in *P. abelii* and *G. gorilla*, the clusters encode slightly longer oligopeptides - AAAAAAAAAGAG and AAAAAAAAAGAG, respectively. In the mRNA of mouse and rat *MAPK1* genes, the 22 nt clusters contained a sequence of GCG triplets capable of binding ID03332.3p-miR and encoding AAAAAAG. This demonstration underscores the variability in cluster length of binding sites compared to highly conserved flanking nucleotide sequences.

```
CCAACAUGGCGGCGGCGGCGGCGGCGGCGGCGGCGGGCGCGGGCCCGGAGA ggo
CCAACAUGGCGGCGGCGGCGGCGGCGGCGGCGGCGG---GCGCGGGCCCGGAGA pab
CCAACAUGGCGGCGGCGGCGGCGGCGGCGGGCGCGGGC----------CCGGAGA hsa
CCAACAUGGCGGCGGCGGCGGCGGCGGCGGGCGCGGGC----------CCGGAGA ptr
CCAACAUGGCGGCGGCGGCGGCGGCGGCGGGCGCGGGC----------CCGGAGA ppa
CCAACAUGGCGGCGGCGGCGGCGGCGGCGGGCGCGGGC----------CCGGAGA pan
CCAACAUGGCGGCGGCGGCGGCGGCGGCGGGCGCGGGC----------CCGGAGA mml
CCAACAUGGCGGCGGCGGCGGCGGCGGCGGGCGCGGGC----------CCGGAGA mne
CCAACAUGGCGGCGGCGGCGGCGGCGGCGGGCGCGGGC----------CCGGAGA mfa
CCAACAUGGCGGCGGCGGCGGCGGCGGCGGGC----------------CCGGAGA mmu
CCAACAUGGCGGCGGCGGCGGCGGCGGCGGGC----------------CCGGAGA rno
```

Figure 13 - The nucleotide sequences of mRNA regions of orthologous *MAPK1* genes containing clusters of binding sites of four miRNAs [277].

The *PPP2CA* gene encodes PP2Acα, a vital catalytic subunit of PP2A, a prominent cellular serine-threonine phosphatase and a tumor suppressor comprised of diverse subunits that target numerous substrates [281]. An experiment conducted by Wang and colleagues [280] validated the interaction between miR-3661 and PPP2CA. Furthermore, the MirTarget program confirmed the presence of a fully complementary interaction in the CDS of the *PPP2CA* gene mRNA, as specified in Table 10.

Regarding the *TERT* gene, it encodes telomerase, which plays a critical role in cellular aging, and its dysregulation in somatic cells may contribute to carcinogenesis. Within the CDS mRNA of the *TERT* gene, a cluster of binding sites exists for three miRNAs that exhibit high free energy upon binding (as mentioned in Table 10). The nucleotide sequence of this binding site cluster remains conserved in the CDS mRNA of the *TERT* gene (as illustrated in Figure 14) across all studied objects' TERT orthologous proteins, encoding the LGAKGAAGPLPS oligopeptide.



Figure 14 - The nucleotide sequences of mRNA regions of orthologous *TERT* genes containing clusters of binding sites of three miRNAs [277].

The CDS section of the human *VEGFC* gene harbors a cluster of seven miRNA binding sites, where the collective length of these sites spans 31 nucleotides, significantly shorter by a factor of 5.1 compared to the combined lengths of individual binding sites (Table 10). While the VEGFC gene binding sites are longest in *M. nemestrina* and encode the REAPAAAAAAALE oligopeptide, in other species, the cluster's length varies due to the presence of the GCC codon. However, this length variation still allows effective binding of specific miRNAs such as ID02052.5p-miR, ID02187.5p-miR, ID01041.5p-miR, and ID01873.3p-miR. Furthermore, the surrounding nucleotides flanking the binding sites remain conserved across different species in the mRNA of the *VEGFC* gene (as illustrated in Figure 15).

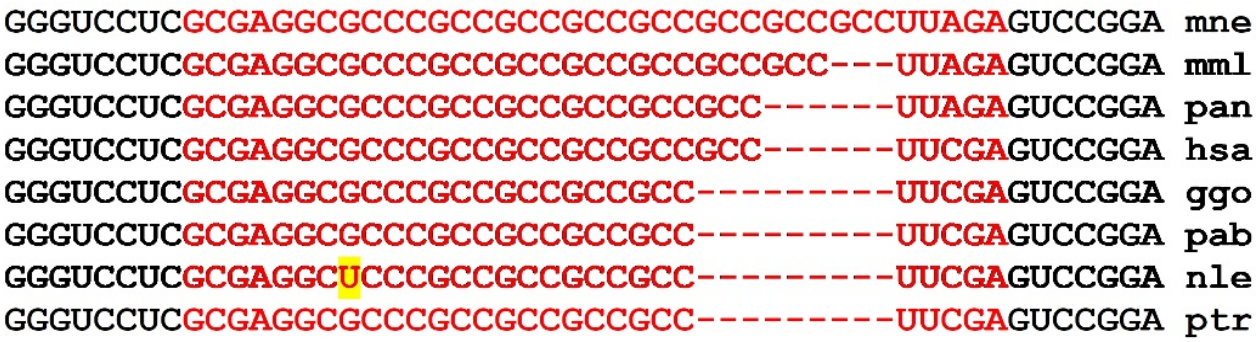


Figure 15 - The nucleotide sequences of CDS mRNA regions of orthologous *VEGFC* genes containing clusters of seven miRNA binding sites [277].

The 3'UTR of mRNAs from various gastric cancer candidate genes exhibits numerous binding sites for multiple miRNAs, as detailed in Table 11 and Table C10. Notably, among 14 candidate gastric cancer genes, a distinctive characteristic is the existence of overlapping nucleotide sequences within their 3'UTR mRNA, which accommodate binding sites for both miR-5095 and miR-619-5p (as illustrated in Figure 12). These binding schemes highlight a significant degree of complementarity in nucleotide sequences between miR-5095 and miR-619-5p and the respective mRNA of these genes, encompassing even non-standard A-C and G-U bonds.

Table 11 - The characteristics of miRNA interactions with 3'UTR mRNAs of gastric cancer candidate genes [277].

Gene; NX	miRNA	Start of site, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
1	2	3	4	5	6
ATM; 8.9	ID03006.5p-miR	9778	-121	89	24
	miR-5095	9787	-108	93	21
	miR-619-5p	9793	-119	98	22
	miR-5096	9882	-104	92	21
	miR-5585-3p	9950	-110	95	22
	miR-1273a	11054	-119	90	25
	miR-1273g-3p	11076	-113	96	21
	miR-1273e	11119	-108	93	22
	miR-5585-5p	11156	-106	91	22
FLT1; 4.6	ID01030.3p-miR	6909÷6923 (7)	-108÷-110	89÷91	23
	miR-466	6911÷6937 (9)	-106÷-108	89÷93	23
	ID00436.3p-miR	6913÷6925 (7)	-104	89	23
IGF1; 5.9	ID00470.5p-miR	4042÷4058 (9)	-108	89	23
	miR-574-5p	4042÷4062 (11)	-108÷-113	89÷93	23
	miR-1273g-3p	6009	-113	96	21
	miR-1273f	6042	-102	98	19
	miR-1273d	6043	-119	87	25
	miR-1273e	6052	-108	93	22
IGF2; 3.2	ID00470.5p-miR	2286÷2351 (6)	-108÷-113	89÷93	23
	miR-574-5p	2288, 2290	-108	89	23
	miR-574-5p	2397÷2408 (4)	-108÷-113	89÷93	23
	ID00470.5p-miR	2404, 2412	-110	91	23
	ID00470.5p-miR	2442÷2463 (3)	-108÷-113	89÷93	23
	miR-574-5p	2465, 2484	-108	89	23
	ID00470.5p-miR	2520÷2539 (3)	-108	89	23
	miR-574-5p	2522	-108	89	23
	ID00470.5p-miR	2655÷2672 (3)	-108÷-110	89÷91	23
	ID00470.5p-miR	2704, 2725	-108	89	23
	miR-574-5p	2727, 2731	-108, -110	89, 91	23
JAK2; 14.1	miR-466	5182÷5200 (10)	-104÷-106	89÷91	23
	ID01030.3p-miR	5184÷5200 (9)	-108	89	23
	ID00436.3p-miR	5184÷5202 (10)	-104÷-106	89÷91	23
SP1; 18.8	miR-466	4145÷4161 (9)	-104÷-106	89÷91	23
	ID01030.3p-miR	4147÷4159 (7)	-108	89	23
	ID00436.3p-miR	4147÷4161 (8)	-104÷-106	89÷91	23
XRCC1; 17.2	miR-574-5p	2033÷2055 (11)	-108÷-113	89÷93	23
	ID00470.5p-miR	2035÷2055 (9)	-108÷-113	89÷93	23
ZEB1; 14.8	miR-574-5p	3587÷3605 (10)	-113	93	23
	ID00470.5p-miR	3587÷3605 (10)	-108	89	23

The schemes below depicting the interaction of miR-5095 and miR-619-5p within the 3'UTR of mRNA genes demonstrate the robust effectiveness of the MirTarget program in delineating the quantitative aspects of miRNA-mRNA interactions. Across all instances shown in Figure 12, both miRNAs consistently exhibit complementary binding to the mRNA, encompassing non-standard nucleotide pairs throughout their entire length. Considering this consistency, all 14 genes involved in conjunction with miR-5095 and miR-619-5p are suggested as potential associations for the diagnosis of gastric cancer.

The sequences within the cluster of miR-5095 and miR-619-5p binding sites exhibit significant conservation (as observed in Table C10), with minor purine-to-purine (G ↔ A) and pyrimidine-to-pyrimidine (C ↔ U) substitutions that minimally impact the free energy involved in miRNA interaction with mRNA. These findings underscore the stability of these miRNAs' interactions with the mRNA of candidate genes related to gastric cancer.

AURKA,miR-5095,420*,-108,93 5'-CGCGGUGG**C**UCA**U**GCC**C**GUAA-3' 3'-GCGCCACC**A**AGU**G**CGG**A**CAUU-5'	*AURKA*,miR-619-5p,426*,-119,98 5'-GGCUCAUGCC**C**GUAAUCCCAGC-3' 3'-CCGAGUACGG**A**CAUUAGGGUCG-5'
ATM, miR-5095, 9787, -108,93 5'-CGC**A**G**C**GG**C**UCACGCCUGUAA-3' 3'-GCG**C**C**A**CC**A**AGUGCGGACAUU-5'	*ATM*,miR-619-5p,9793,-119,98 5'-GGCUCA**C**GCCUGUAAUCCCAGC-3' 3'-CCGAGU**A**CGGACAUUAGGGUCG-5'
BRCA1,miR-5095,6406,-106,91 5'-C**AU**GGUGG**C**UCACGCCUGUAA-3' 3'-G**CG**CCACC**A**AGUGCGGACAUU-5'	*BRCA1*,miR-619-5p,6412,-119,98 5'-GGCUCA**C**GCCUGUAAUCCCAGC-3' 3'-CCGAGU**A**CGGACAUUAGGGUCG-5'
CEACAM5,miR-5095,3229,-115,98 5'-CGCGGUGG**C**UCACGCCUGUAA-3' 3'-GCGCCACC**A**AGUGCGGACAUU-5'	*CEACAM5*,miR-619-5p,3235,-119,98 5'-GGCUCA**C**GCCUGUAAUCCCAGC-3' 3'-CCGAGU**A**CGGACAUUAGGGUCG-5'
CLDN18,miR-5095,1852,-106,91 5'-C**A**C**A**GUGG**C**UCACGCCUGUAA-3' 3'-G**C**G**C**CACC**A**AGUGCGGACAUU-5'	*CLDN18*,miR-619-5p,1858,-115,95 5'-GGCUCA**C**GCCUGUAAUCC**U**AGC-3' 3'-CCGAGU**A**CGGACAUUAGG**G**UCG-5'
CREB1,miR-5095,2791,-110,95 5'-CGC**A**GUGG**C**UCACGCCUGUAA-3' 3'-GCG**C**CACC**A**AGUGCGGACAUU-5'	*CREB1*,miR-619-5p,2797,-119,98 5'-GGCUCA**C**GCCUGUAAUCCCAGC-3' 3'-CCGAGU**A**CGGACAUUAGGGUCG-5'
F2R,miR-5095,3239,-115,98 5'-CGCGGUGG**C**UCACGCCUGUAA-3' 3'-GCGCCACC**A**AGUGCGGACAUU-5'	*F2R*,miR-619-5p,3245,-119,98 5'-GGCUCA**C**GCCUGUAAUCCCAGC-3' 3'-CCGAGU**A**CGGACAUUAGGGUCG-5'
HPSE,miR-5095,2759,-106,91 5'-CGC**A**GUGG**C**UCAC**A**CCUGUAA-3' 3'-GCG**C**CACC**A**AGUG**C**GGACAUU-5'	*HPSE*,miR-619-5p,2765,-110,91 5'-GGCUCA**CA**CCUGUAAUCC**U**AGC-3' 3'-CCGAGU**AC**GGACAUUAGG**G**UCG-5'
IL10,miR-5095,1210,-115,98 5'-CGCGGUGG**C**UCACGCCUGUAA-3' 3'-GCGCCACC**A**AGUGCGGACAUU-5'	*IL10*,miR-619-5p,1216,-119,98 5'-GGCUCA**C**GCCUGUAAUCCCAGC-3' 3'-CCGAGU**A**CGGACAUUAGGGUCG-5'
IRF1,miR-5095,2229,-110,95 5'-CGCGGUGG**C**UCAC**A**CCUGUAA-3' 3'-GCGCCACC**A**AGUG**C**GGACAUU-5	*IRF1*,miR-619-5p,2235,-115,95 5'-GGCUCA**CA**CCUGUAAUCCCAGC-3' 3'-CCGAGU**AC**GGACAUUAGGGUCG-5'
IRF1,miR-5095,2653,-110,95 5'-C**A**CGGUGG**C**UCACGCCUGUAA-3' 3'-G**C**GCCACC**A**AGUGCGGACAUU-5'	*IRF1*,miR-619-5p,2659,-119,98 5'-GGCUCA**C**GCCUGUAAUCCCAGC-3' 3'-CCGAGU**A**CGGACAUUAGGGUCG-5'
LEP,miR-5095,3094,-110,95 5'-**U**GCGGUGG**C**UCACGCCUGUAA-3' 3'-**G**CGCCACC**A**AGUGCGGACAUU-5'	*LEP*,miR-619-5p,3100,-119,98 5'-GGCUCA**C**GCCUGUAAUCCCAGC-3' 3'-CCGAGU**A**CGGACAUUAGGGUCG-5'
MTHFR,miR-5095,6855,-110,95 5'-CGCGGUGG**C**UCAC**A**CCUGUAA-3' 3'-GCGCCACC**A**AGUG**C**GGACAUU-5'	*MTHFR*,miR-619-5p,6861,-115,95 5'-GGCUCA**CA**CCUGUAAUCCCAGC-3' 3'-CCGAGU**AC**GGACAUUAGGGUCG-5'
PDCD4,miR-5095,3215,-106,91 5'-C**A**CGGUGG**C**UCA**U**GCCUGUAA-3' 3'-G**C**GCCACC**A**AGU**G**CGGACAUU-5'	*PDCD4*,miR-619-5p,3221,-121,100 5'-GGCUCAUGCCUGUAAUCCCAGC-3' 3'-CCGAGUACGGACAUUAGGGUCG-5'
STAT3,miR-5095,3125,-106,90 5'-C**AU**GGUGG**C**UCACGCCUGUAA-3' 3'-G**CG**CCACC**A**AGUGCGGACAUU-5'	*STAT3*,miR-619-5p,3131,-119,98 5'-GGCUCA**C**GCCUGUAAUCCCAGC-3' 3'-CCGAGU**A**CGGACAUUAGGGUCG-5'
Note. The top line contains sequentially: gene; miRNA; the beginning of the BS, nt; ΔG, kJ/mole; ΔG/ΔGm,%. * - the miRNA BS is located in the 5'UTR. The top line of the diagram is the nucleotide sequence of the miRNA binding site; the bottom line is the nucleotide sequence of the miRNA. Non-canonical pairs are in bold.	

Figure 16 - Schemes of interaction of miR-5095 and miR-619-5p in the 3'UTR of mRNA of candidate gastric cancer genes [277].

Furthermore, the *PRKAA1* gene is targeted by miR-1273d, miR-1273e, miR-1273g-5p, miR-1273f, and miR-1273h-5p, as evidenced by Figure 17 and Table C10. Except for miR-1273g-3p, the binding sites of these miRNAs form a conserved cluster within the mRNA of orthologous *PRKAA1* genes across various species like P. *troglodytes, G. gorilla, P. abelii, P. anubis, P. paniscus, N. leucogenys, M. mulatta,* and *M. nemestrina.* This coherent association between the miR-1273 family and the *PRKAA1* gene is suggested as a promising marker for the diagnosis of gastric cancer.

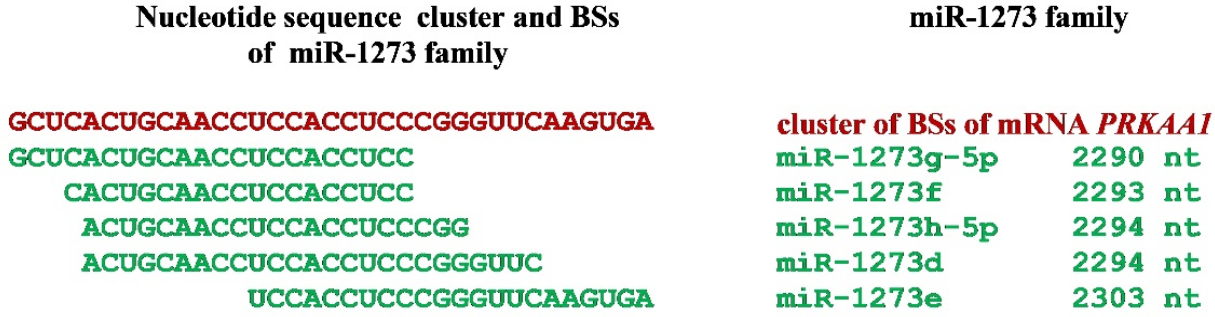
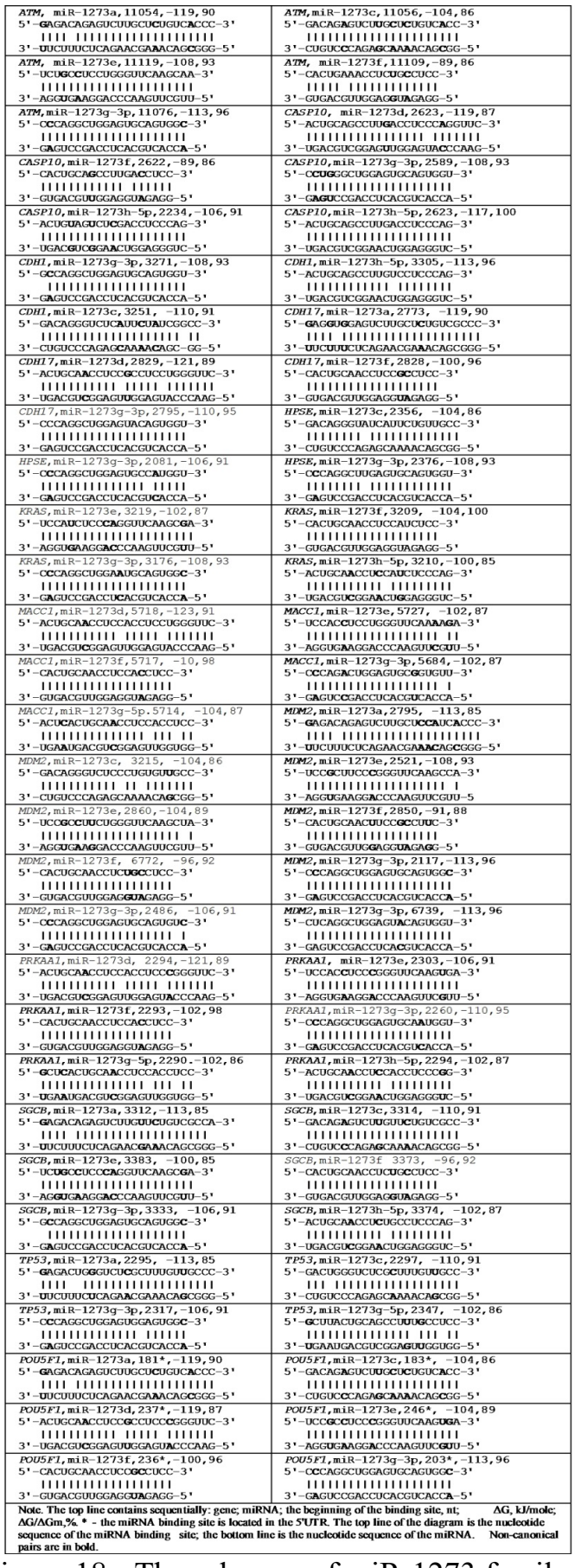


Figure 17 - The scheme of miR-1273 family binding sites in the 3'UTR of the mRNA of the candidate *PRKAA1* gene [277].

The miR-1273 family of miRNAs exhibited binding sites within the mRNA sequences of 12 gastric cancer candidate genes (detailed in Table 11 and Table C10). Through the formation of non-canonical nucleotide pairs, almost every nucleotide from each miRNA established hydrogen bonds with specific nucleotides within the mRNA structure (as depicted in Figure 18). Remarkably, certain miRNA families displayed binding sites within the mRNA that shared overlapping nucleotide sequences, thus forming clusters of binding sites. For instance, both miR-1273a and miR-1273c competed for binding to the mRNA of genes ATM, CDH1, POU5F1, SGCB, and TP53. Similarly, miR-1273e and miR-1273f exhibited overlapping mRNA binding sites for genes *ATM, KRAS, MACC1, MDM2, POU5F1, PRKAA1,* and *SGCB*. Furthermore, the mRNAs of *CDH17, MACC1, POU5F1,* and *PRKAA1* genes contained clusters of binding sites for miR-1273d and miR-1273f. These associations between the miR-1273 family of miRNAs and the candidate target genes linked to gastric cancer serve as promising markers for the diagnosis of this disease.



Note. The top line contains sequentially: gene; miRNA; the beginning of the binding site, nt; ΔG, kJ/mole; ΔG/ΔGm,%. * - the miRNA binding site is located in the 5'UTR. The top line of the diagram is the nucleotide sequence of the miRNA binding site; the bottom line is the nucleotide sequence of the miRNA. Non-canonical pairs are in bold.

Figure 18 - The schemes of miR-1273 family interaction in the 3'UTR of mRNA of candidate gastric cancer genes [277].

In the 3'UTR mRNA of five genes, clusters of binding sites were observed for miR-466, ID01030.3p-miR, and ID00436.3p-miR, featuring multiple repeats of GU dinucleotides (detailed in Table 11 and Table C10). The interaction schemes between these miRNAs and the mRNA of their respective target genes display nearly complete complementarity, involving both canonical and non-canonical nucleotide pairs. These associations of miR-466, ID01030.3p-miR, and ID00436.3p-miR with the five candidate target genes associated with gastric cancer are suggested to be utilized as diagnostic markers for this disease.

CCND1,miR-574-5p,2595,-113,93 5'-ACACACACACACACACACACA-3' |||||| |||||||||||| || 3'-UGUGUGAGUGUGUGUGUGUGAGU-5'	*CCND1*,ID00470.5p-miR,2595,-108,89 5'-ACACACAC**A**CA**C**ACACAC**A**CACA-3' ||||||||||||||||||||||| 3'-UGUGUGUG**C**GU**A**UGUGUG**C**AUGU-5'
IGF1,miR-574-5p,4042,-113,93 5'-ACACACACACACACACACACA-3' |||||| |||||||||||| || 3'-UGUGUGAGUGUGUGUGUGUGAGU-5'	*IGF1*,ID00470.5p-miR,4042,-108,89 5'-ACACACAC**A**CACACACACAC**A**CACA-3' ||||||||||||||||||||||| 3'-UGUGUGUG**C**GUAUGUGUG**C**AUGU-5'
IGF2,miR-574-5p;2299;-106;88 5'-ACACACGCACACACA**UG**CACACA-3' |||||| |||||||||||| || 3'-UGUGUGAGUGUGUGU**GU**GUGAGU-5'	*IGF2*,ID00470.5p-miR;2297;-113;93 5'-ACACACACGCA**C**ACACACA**U**G**C**ACA-3' ||||||||||||||||||||||| 3'-UGUGUGUGCGU**A**UGUGU**G**C**A**UGU-5'
IGF2,miR-574-5p;2343;-106;88 5'-ACACACGCACACACA**U**GCACACA-3' |||||| |||||||||||| || 3'-UGUGUGAGUGUGUGU**G**UGUGAGU-5'	*IGF2*,ID00470.5p-miR;2341;-113;93 5'-ACACACACGCA**C**ACACACA**U**G**C**ACA-3' ||||||||||||||||||||||| 3'-UGUGUGUGCGU**A**UGUGU**G**C**A**UGU-5'
IGF2,miR-574-5p;2522;-108;89 5'-ACACA**U**GCACACACAGCACACACA - 3' |||||| |||||||| ||||| || 3'-UGUGU**G**AGUGUGUGU-GUGUGAGU - 5'	*IGF2*,ID00470.5p-miR;2520;-108;89 5'-ACACACA**U**GCA**C**ACACAGCA**C**ACA-3' |||||||||||||||| |||||| 3'-UGUGUGU**G**CGU**A**UGUGU-GCAUGU-5'
IGF2,miR-574-5p;2835;-104;86 5'-ACACACA**UG**C**G**CACACACACGCA-3' |||||| |||||||||||| || 3'-UGUGUGA**GU**G**U**GUGUGUGUGAGU-5'	*IGF2*,ID00470.5p-miR;2837;-110;91 5'-ACACA**UG**CGCA**C**ACACACG**C**ACA-3' ||||||||||||||||||||||| 3'-UGUGU**GU**GCGU**A**UGUGUGC**A**UGU-5'
SMAD4,miR-574-5p,7744,-113,93 5'-ACACACACACACACACACACA-3' |||||| |||||||||||| || 3'-UGUGUGAGUGUGUGUGUGUGAGU-5'	*SMAD4*,ID00470.5p-miR,7744,-108,89 5'-ACACACAC**A**CA**C**ACACACAC**A**CACA- 3' ||||||||||||||||||||||| 3'-UGUGUGUG**C**GUAUGUGUG**C**AUGU-5'
XRCC1,miR-574-5p,2035,-113,93 5'-ACACACACACACACACACACA-3' |||||| |||||||||||| || 3'-UGUGUGAGUGUGUGUGUGUGAGU-5'	*XRCC1*,ID00470.5p-miR,2035,-108,89 5'-ACACACAC**A**CA**C**ACACACAC**A**CACA-3' ||||||||||||||||||||||| 3'-UGUGUGUG**C**GU**A**UGUGUGCAUGU-5'
ZEB1,miR-574-5p,3587,-113,93 5'-ACACACACACACACACACACA-3' |||||| |||||||||||| || 3'-UGUGUGAGUGUGUGUGUGUGAGU-5'	*ZEB1*,ID00470.5p-miR,3587, -108,89 5'-ACACACAC**A**CA**C**ACACACAC**A**CACA-3' ||||||||||||||||||||||| 3'-UGUGUGUG**C**GUAUGUGUG**C**AUGU-5'
Note. The top line contains sequentially: gene; miRNA; the beginning of the binding site, nt; ΔG, kJ/mole; ΔG/ΔGm,%. * - the miRNA binding site is located in the 5'UTR. The top line of the diagram is the nucleotide sequence of the miRNA binding site; the bottom line is the nucleotide sequence of the miRNA. Non-canonical pairs are in bold.	

Figure 19 - The schemes of the interaction of miR-574-5p and ID00470.5p-miR in the 3'UTR of mRNA of candidate gastric cancer genes [277].

Moreover, among the identified candidate genes for gastric cancer, six genes were identified as target genes for miR-574-5p and ID00470.5p-miR (as indicated in Table 11 and Table C10). Illustrated in Figure 19 are the interaction schemes of these miRNAs with the target gene mRNA. Notably, the binding sites of these miRNAs consist of multiple repeats of the CA dinucleotide sequence within the mRNA of the *IGF2* gene, ranging from two to ten times in the mRNA of the gene. Consequently, the associations formed by miR-574-5p and ID00470.5p-miR, in conjunction with the target genes, are recommended for use in the diagnosis of gastric cancer.

Additionally to the miRNA-gene associations present within the clusters, several other associations of miRNAs and genes outside these clusters were identified with free energy interactions of -130 kJ/mol and above (as detailed in Tables A8 and A10). Notable associations include ID02052.5p-miR with *DNMT1*, ID02457.3p-miR with *EGFR*, ID02761.3p-miR with *EZH2*, ID03047.3p-miR with *FGFR2*, miR-6789-5p with HIF1A, ID03332.3p-miR with *KRAS*, ID02064.5p-miR with *NFKB1*, ID02761.3p-miR with *PTEN*, ID02064.5p-miR with *TGFB1*, ID00252.5p-miR, ID00961, ID00049.5p-miR with *TIMP2*, ID02294.5p-miR with *ACE*, ID01895.5p-miR, ID02052.5p-miR with *CDX2*, ID00920.5p-miR with *OGG1*, and miR-4767, ID03332.3p-miR with *SIRT1*.

Moreover, in the 3'UTR mRNA of candidate gastric cancer genes, miRNA-gene associations exhibit free energy interactions of -115 kJ/mol and above, including interactions like miR-619-5p with *BRCA1*; miR-1273d, miR-1273h-5p (100%) with *CASP10*; miR-1273a, miR-1273d with CDH17; miR-5095, miR-619-5p with *CEACAM5*; miR-619-5p with ERBB3; ID03224.5p-miR with *KRAS*; miR-619-5p with *MACC1*; ID01404.5p-miR, ID02732.3p-miR with *SMAD4*; miR-6089, ID03306.3p-miR, ID03208.5p-miR, ID00296.3p-miR, miR-1183 with *TGFB1*; ID01941.5p-miR with *TIMP2*; and ID00548.3p-miR, ID02379.3p-miR with *TP53*.

Furthermore, Table 12 offers insight into miRNAs targeting two or more candidate genes, highlighting cases where certain miRNAs have multiple alternative target genes with varying free energy interactions.

Table 12 - The association of miRNA and alternative candidate genes for gastric cancer [277].

miRNA	Candidate genes
ID02064.5p-miR	*ARID1A, E2F1, NFKB1, ODC1, TDFB1, VEGFC*
miR-3960	*ARID1A, CDX2, E2F1, ODC1, VEGFC*
ID02052.5p-miR	*CDX2,DNMT1, E2F1, VEGFC*
ID01041.5p-miR	*CDX2, TBC1D9, VEGFC*
ID02761.3p-miR	*ARID1A, EZH2, PTEN*
ID01702.3p-miR	*ARID1A, PIC3CA, TBC1D9*
ID01895.5p-miR	*PIC3CA, CDX2, TBC1D9*
ID03332.3p-miR	*KRAS, MAPK1,SIRT1*
ID00296.3p-miR	*ARID1A, TGFB1*
ID02084.5p-miR	*ARID1A, ODC1*

For instance, ID02064.5p-miR has six alternative target genes within a range of -129 kJ/mol to -140 kJ/mol, and similar scenarios are observed with other miRNAs targeting multiple genes. The complexity of these associations emphasizes the necessity for a validation system to discern the most effective associations that reflect the disease state accurately. Considering the diversity of miRNA-gene associations organized within clusters and involving multiple binding sites, the creation of a chip system becomes essential to monitor and represent the status of these associations in the context of the disease.

Roughly half of the miRNAs stem from intronic origins, often co-expressed with the host genes. However, their potential involvement in disease development versus merely reflecting changes in gene expression needs highlighting. Establishing correlations between altered miRNA concentrations and disease indicators remains largely hypothetical and necessitates validation, making therapeutic approaches based on these correlations challenging. The numerous associations identified as potential disease markers complicate the selection of highly specific markers for disease and its progression. To pinpoint the most influential associations driving the disease becomes challenging due to the disease's diverse forms and characteristics. Simultaneous monitoring of changes in miRNA and mRNA concentrations via chips is crucial to identify pivotal components within these associations, given that a single component might belong to multiple associations, as illustrated in Table 12.

For instance, ID02064.5p-miR could target six genes, while six miRNAs might target the ARID1A gene. This complexity emphasizes the necessity of analyzing multiple miRNA-gene associations to comprehend their role in the disease. Chips capable of determining both miRNA and mRNA concentrations simultaneously will help delineate the key association members responsible for disease development and their respective contributions. For instance, a rise in miRNA concentration could reduce mRNA translation, and if protein synthesis alters with constant miRNA levels, gene expression modulation may be necessary to correct protein encoding. Such changes in miRNA concentration might act as the primary cause of the disease.

Comparing the NX values of target genes with those of their host genes (co-expressed with miRNAs) indicates comparable relationships among association components. Highly expressed genes typically lack clusters of miRNA binding sites in their mRNAs (as observed in Table 9-11 and Table C8 - Table C10), often possessing only one to three binding sites, which are usually dispersed within their mRNA sequences. This trend is evident in genes as *CDKN1A, SMAD2, LGALS3, JUN, MUC6, PGC, ERBB3, MUC1, PDCD4, STAT3, THBS1, MUC6,* and *CLDN18*, each containing a limited number of miRNA binding sites.

Understanding the expression patterns of host genes housing intronic miRNAs aids in predicting the preferential binding of competing miRNAs based not only on their free interaction energy but also on their concentrations. For example, while the free binding energy of miR-619-5p is higher (-121 kJ/mol) than that of miR-5095 (-117 kJ/mol), the latter is expressed 4.8 times more than the former, granting it an

advantage in binding within miR-619-5p and miR-5095 binding sites clusters. Significant alterations in the concentration of miRNA or mRNA within these associations might be critical in altering their interactions, potentially leading to the onset of pathology.

3.2.3 The characteristics of the miRNAs interactions with mRNAs of genes associated with small intestinal cancer

The analysis indicates that genes lacking miRNAs binding sites with a ΔG/ΔGm value higher than 86% exhibit an expression level independent of miRNAs. Among the 40 candidate genes studied, it was found that 33% are not under the regulation of miRNAs, suggesting that their expression remains unaffected or un-suppressed.

In the analysis of miRNA binding schemes with the mRNA of *ASXL1, GNAS,* and *KRAS* genes, instances of complete complementarity binding sites were discovered (as shown in Figure 20).

Specifically, the binding site of ID00061.3p-miR within the mRNA of the *ASXL1* gene is situated in the 5'UTR, displaying a ΔG/ΔGm value of 100%, with a ΔG value of -138 kJ/mol. On the other hand, the binding site of ID02093.5p-miR, consisting of 22 nucleotides, is located within the CDS of the *GNAS* gene, exhibiting a free binding energy of -127 kJ/mol.

Gene, miRNA, site, region of mRNA, characteristics of binding	Gene, miRNA, site, region of mRNA, characteristics of binding

ASXL1, ID00061.3p-miR, 340 nt,
5'UTR -138 kJ/mole, 100%, 22 nt
5' - CCACCGCCGCCGCCGCCCCAGCC - 3'
  ||||||||||||||||||| |||
3' - GGUGGCGGCGGCGGCGGGG-CGG - 5'
	GNAS, ID02093.5p-miR, 1626 nt,
	CDS, -127 kJ/mole, 100%, 22 nt
	5' - UGACGCCCCAGCCGAUCCCGAC - 3'
	  ||||||||||||||||||||||
	3' - ACUGCGGGGUCGGCUAGGGCUG - 5'

KRAS, miR-1273f, 3209 nt, 3'UTR -104 kJ/mole, 100%, 19 nt 5' - CACUGCAACCUCCAUCUCC - 3' |||||||||||||||||| 3' - GUGACGUUGGAGGUAGAGG - 5'	

Figure 20 - The scheme and characteristics of the interaction of ID00061.3p-miR, ID03064.3p-miR, ID02093.5p-miR and ID03151.3p-miR with mRNA of *ASXL1, GNAS* and *LRP1B* genes [282].

Figures 20 and 21 illustrate the interaction schemes between miRNAs and the mRNAs of genes within the 5'UTR, CDS, and 3'UTR. Numerous publications discuss various commonly used programs designed for identifying miRNA binding sites within the mRNA of target genes, typically focusing on binding sites ranging from 6 to 8 nucleotides in length, known as seed regions.

Gene, miRNA, site, region of mRNA, characteristics of binding	Gene, miRNA, site, region of mRNA, characteristics of binding
ARID1A, TJU_CMC_MD2.ID00457.3p-miR, 416 nt,CDS, -125 kJ/mole, 92%, 22 nt 5' - **A**ACCCGCCGCCGCCGCCGCCCU - 3' || |||||||||||||||||| *3' - **C**UCGGCGGCGGCGGCGGCGGGA - 5'*	*ASXL1*, TJU_CMC_MD2.ID01895.5p-miR, 400 nt, 5'UTR, -132 kJ/mole, 89%, 24 nt 5' - CGCG**U**GGAG**CU**GCCGCCGCCGCCG - 3' ||| |||||||||||||||||| 3' - GCGG**G**CCUC**AG**CGGCGGCGGCGGC - 5'
ATM, TJU_CMC_MD2.ID03006.5p-miR, 9778 nt,3'UTR,-121 kJ/mole, 89%, 24 nt 5' - **C**GGGCUGGGC**G**CAGCGGCUCACGC - 3' ||||||||||||| |||||||| 3' - **A**CCCGACCCG**U**GUCCCCGAGUGUA - 5'	*CDKN2B*, TJU_CMC_MD2.ID00470.5p-miR, 1750 nt, 3'UTR,-108 kJ/mole, 89%, 23 nt 5' - ACACACAC**A**CA**C**ACACAC**AC**ACA - 3' |||||||||||||||||||||| 3' - UGUGUGUG**C**GU**A**UGUGUG**CA**UGU - 5'
LRP1B, TJU_CMC_MD2.ID02429.3p-miR, 896 nt,5'UTR, -127 kJ/mole, 94%, 23 nt 5' - CU**G**CCGC**C**GCC**G**CCGC**C**GCCGGG - 3' ||||||||||||||||||||||| 3' - GA**U**GGCG**A**CGG**U**GGCG**A**CGGCCC - 5'	*MUC6*, TJU_CMC_MD2.ID01401.3p-miR, 7883 nt,3'UTR, -108 kJ/mole, 91%, 21 nt 5' - C**G**CCCC**C**GCUCCCA**G**CCAGGC - 3' ||||||| |||||||||||| 3' - G**U**GGGG**A**AGAGGGU**U**GGUCCG - 5'
TP53, TJU_CMC_MD2.ID01838.5p-miR, 2460 nt, 3'UTR,-110 kJ/mole,88%, 24 nt 5' - **G**AUGG**G**GUCUCACAGUGUUGC**C**CA - 3' || |||||||||| |||||||||| 3' - **U**UUCC**U**CAGAGUGACACAACG**A**GU - 5'	*VEGFA*, TJU_CMC_MD2.ID03097.3p-miR, 887 nt, CDS, -123 kJ/mole, 88%, 24 nt 5' - **C**UCC**A**GCC**G**CGCGCGCUCC**C**CAGG - 3' |||||| |||||||||||||||| 3' - **A**AGG**C**CCG**U**GCGCGCGAGG**A**GUCC - 5'
In bold type highlighted the non-canonical nucleotide (A-C, G-U) pairs	

Figure 21 - The scheme and characteristics of the interaction of miRNAs with mRNAs of genes participating in the development of small intestinal cancer [282, 283].

However, it's important to note that this perspective, prevalent in many articles, often leads to inaccurate conclusions and contributes to a high number of false-positive binding sites. Figures 20 and 21 specifically highlight the advantages offered by the MirTarget program. The free energy (ΔG) associated with miRNA binding sites across different mRNAs displayed a range from -108 to -132 kJ/mol. These variations occurred due to several nucleotide substitutions, particularly from G-C to G-U pairs, influencing the overall energy of binding.

The research examines the interaction between miRNAs and the mRNA of 17 genes in the 5'UTR, each binding from one to several miRNAs (as detailed in Table 14). Notably, some of these mRNAs bind six or more miRNAs.

Within the 5'UTR of the *ARID1A* gene, three miRNA clusters were identified (refer to Table 14). The first spans from 136 nt to 192 nt with an average free energy of interaction of -124 kJ/mol. The second cluster ranges from 206 nt to 228 nt. The third resides in position from 295 to 331 nucleotides, totaling 37 nucleotides in length, and shows an average ΔG value of -111 kJ/mol. As previously mentioned, the *ARID1A* gene functions as a tumor suppressor and plays a pivotal role in the carcinogenesis of numerous organs. ID01257.3p-miR, ID00414.3p-miR, and ID02428.3p-miR form a cluster from 295 nt to 332 nt, spanning 38 nt, with an average free energy of interaction of -111 kJ/mol.

The mRNA of the *ASXL1* gene contains a cluster of 17 miRNA binding sites positioned from 302 nt to 391 nt. The cumulative length of these miRNA binding

sites is 630 nt, condensed within a 90-nt cluster of BS—an approximate 7-time compaction. This compression fosters competition among miRNAs vying to bind to the cluster. Preferential binding is exhibited by miRNAs interacting with an mRNA displaying higher free energy and present in greater concentrations compared to competing miRNAs. Another cluster of miRNA binding sites in the mRNA of the *ASXL1* gene is situated from 172 nt to 197 nt, forming a cluster of 26 nt in total length, with an average ΔG = -129 kJ/mol. ID00021.5p-miR and ID01895.5p-miR possess binding sites from 400 nt, yet ID01895.5p-miR is likely to occupy this site due to its higher energy of interaction, facilitated by the increased concentration of GC-pairs.

In the 5'UTR mRNA of the *LRP1B* gene, 22 miRNAs form a cluster of binding sites from 887 nt to 921 nt, totaling 35 nt in length, with an average free energy of interaction at -128 kJ/mol. The cumulative length of these 22 miRNAs is 584 nt. The formation of such a binding site cluster indicates the gene's significant ability for compaction, thereby fostering competition for this specific binding site.

The mRNA of the *GNAS* gene contains two binding sites for ID02111.3p-miR and ID00724.5p-miR. Meanwhile, two miRNAs form a binding site cluster in the mRNA of the *PDXP* gene from 17 nt to 52 nt, exhibiting an average free energy of interaction equal to -124 kJ/mol.

Table 14 - The characteristics of miRNA interactions with 5'UTR mRNAs of SIC candidate genes [282].

Gene*;* NX	miRNA	Start of site, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
1	2	3	4	5	6
ARID1A; 14.4	ID02106.3p-miR	136	-123	89	23
	ID01778.3p-miR	140	-134	90	24
	ID00296.3p-miR	141, 166	-138	88	25
	miR-6081	143	-125	89	24
	ID01702.3p-miR	147	-134	89	24
	ID00465.5p-miR	148	-113	93	20
	ID01377.3p-miR	152	-117	92	20
	ID02592.5p-miR	164, 167	-123	89	23
	miR-3960	167	-115	92	20
	ID03065.3p-miR	172	-115	92	21
	ID02751.3p-miR	206	-125	92	23
	miR-483-3p	208	-104	91	21
	ID01257.3p-miR	295	-113	93	20
	ID00414.3p-miR	303	-108	93	20
	ID02428.3p-miR	310	-113	91	22
ASXL1; 15.1	ID00522.5p-miR	172	-125	89	23
	ID01804.3p-miR	173	-136	93	23
	ID02187.5p-miR	174	-123	89	23

Table 14 (continued)

1	2	3	4	5	6
	ID03064.3p-miR	220	-136	89	24
	miR-3135b	245	-113	90	22
	ID01755.3p-miR	302	-117	92	21
	ID01403.5p-miR	312	-121	89	23
	ID03367.5p-miR	313	-121	97	20
	ID00296.3p-miR	328÷338 (4)	-138÷-142	88÷91	25
	ID00089.3p-miR	329	-127	92	22
	ID01323.3p-miR	331, 333	-125, -127	92, 94	22
	ID01641.3p-miR	332, 334	-132, -140	89, 94	24
	ID02260.5p-miR	343, 361	-129, -132	92, 94	22
	ID01702.3p-miR	334	-138	92	24
	miR-3960	334, 336	-115, -117	92, 93	20
	ID02522.3p-miR	335	-125	89	23
	ID00061.3p-miR	340	-138	100	22
	ID01106.5p-miR	341	-132	89	24
	ID02229.3p-miR	343	-123	94	21
	ID02499.3p-miR	343	-121	93	21
	miR-6799-5p	357	-108	93	20
	ID02781.3p-miR	372	-121	97	20
	ID00021.5p-miR	400	-117	92	20
	ID01895.5p-miR	400	-132	89	24
GNAS; 56.4	ID02111.3p-miR	30	-115	92	22
	ID00724.5p-miR	49	-100	92	21
KRAS; 7.3	ID01310.3p-miR	17, 29	-121, -123	92, 94	22
	ID03332.3p-miR	37	-132	89	24
	ID00805.3p-miR	96	-113	91	21
LRP1B; 0.0	ID02052.5p-miR	887	-134	90	24
	ID00089.3p-miR	889	-125	91	22
	miR-3196	889	-106	96	18
	ID02294.5p-miR	889, 892	-129, -140	88, 96	24
	ID01041.5p-miR	890	-136	93	24
	ID01804.3p-miR	890	-134	91	23
	ID01873.3p-miR	890	-129	98	21
	ID00061.3p-miR	893	-129	94	22
	ID01641.3p-miR	893	-132	89	24
	ID01702.3p-miR	893	-140	93	24
	miR-3960	893, 895	-117	93	20
	ID00296.3p-miR	894	-138	88	25
	ID02429.3p-miR	896	-127	94	23
	ID01106.5p-miR	897	-132	89	24
	ID03151.3p-miR	897	-115	93	20
	ID01879.5p-miR	898	-125	92	22
	ID02064.5p-miR	898	-134	93	23
	ID02084.3p-miR	898	-138	92	24
	ID02229.3p-miR	899	-123	94	21

Table 14 (continued).

1	2	3	4	5	6
	ID02260.5p-miR	899	-127	91	22
	ID02499.3p-miR	899	-119	92	21
	ID02538.3p-miR	900	-121	90	22
NF1; 3.0	ID02352.5p-miR	157	-123	91	24
	miR-1207-5p	169	-113	91	21
	ID01787.3p-miR	297	-115	89	23
	ID01491.3p-miR	349	-125	91	23
PDXP; 2.3	ID01018.3p-miR	17	-125	89	24
	ID00714.3p-miR	29	-123	88	24
PIK3CA; 3.5	ID00071.3p-miR	4	-115	92	20
	ID01190.5p-miR	4	-140	92	24
	ID01702.3p-miR	4	-140	93	24
	ID01895.5p-miR	4	-132	89	24
	ID02595.5p-miR	4	-115	92	20
	ID00966.5p-miR	6	-136	90	24
	ID00030.3p-miR	7	-121	90	22
	ID01804.3p-miR	8	-134	91	23
	ID01873.3p-miR	8	-123	94	21
	ID03367.5p-miR	10	-115	92	20
PTEN; 10.0	ID02079.5p-miR	75	-115	92	20
	ID02611.3p-miR	486	-125	91	22
	ID03465.3p-miR	525	-119	92	22
	ID01310.3p-miR	531	-119	90	22
	ID02761.3p-miR	533	-132	89	24
	ID03037.3p-miR	536	-121	90	22
	ID01315.3p-miR	705	-115	92	20
	ID02429.3p-miR	790	-121	89	23

Additionally, the *KRAS* gene presents four binding sites, with three sites forming a cluster with overlapped nucleotide sequences. Other genes such as *APC, EGFR, ERBB2, PROM1, SOAT1*, and *SOX9* feature binding sites for single miRNAs (refer to Table C11). The *PTEN* gene serves as an onco-suppressor. Byun et al. [284] observed that *PTEN* inactivation through deletion alone initiates the development of intestinal tumors, particularly adenocarcinomas. The *PTEN* gene's mRNA contains a cluster of binding sites spanning from 525 nt to 557 nt, totaling 33 nt in length with an average ΔG of -123 kJ/mol.

In the 5'UTR region, the average free energy binding of all miRNAs with mRNAs equaled -124 kJ/mol. There were 20 miRNAs binding with mRNAs of their corresponding target genes, with 39 miRNA associations demonstrating a free energy of interaction greater than -130 kJ/mol, suggested as potential markers for early diagnosis in small intestinal cancer.

Some binding sites reside in overlapping mRNA nucleotide sequences within protein-coding regions (referenced in Table 15 and Table C12). The existence of

multiple binding sites for one or several miRNAs in one mRNA amplifies the likelihood of their interaction, subsequently reducing the translation of such mRNAs. A decrease in the binding energy of one miRNA site can be compensated by other sites. Multiple binding sites within an mRNA can intensify the inhibitory effect.

The mRNA of genes *ATM* (ΔG = -104 kJ/mol), *LRP1B* (ΔG = -96 kJ/mol), *MRC1* (ΔG = -115 kJ/mol), and *PIK3CA* (ΔG = -100 kJ/mol) are bound by single miRNAs.

The mRNA of the *MLL2* gene contains binding sites for 20 miRNAs. Kantidakis and coauthors [285] noted that *MLL2* knockdown affects adhesion-related processes and suppresses cell growth. ID02328.5p-miR and ID02562.3p-miR form a cluster of binding sites spanning from 5945 nt to 5988 nt, featuring an average ΔG of -117 kJ/mol. miR-4675 and ID01562.3p-miR form a cluster from 12225 nt to 12258 nt. Additional binding sites are located non-overlappingly within the protein-coding region of the *MLL2* gene's mRNA.

Table 15 - The characteristics of miRNA interactions with CDS mRNAs of SIC candidate genes [282].

Gene; NX	miRNA	Start of site, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
1	2	3	4	5	6
ARID1A; 14.4	miR-3960	390	-115	92	20
	ID00522.5p-miR	410	-127	91	23
	ID02052.5p-miR	410	-132	89	24
	ID02187.5p-miR	411	-127	92	23
	ID02692.5p-miR	413	-127	90	23
	ID00457.3p-miR	415	-125	92	22
	ID02064.5p-miR	417	-134	93	23
	ID02084.3p-miR	418	-138	92	24
	ID02064.5p-miR	418	-132	91	23
	ID02538.3p-miR	420	-125	94	22
	ID01704.5p-miR	471	-123	89	23
	ID02761.3p-miR	487	-140	94	24
	ID01810.3p-miR	490	-115	89	23
	ID02761.3p-miR	493	-134	90	24
	ID01753.3p-miR	749	-104	91	21
	ID00756.3p-miR	843	-125	91	23
	ID02294.5p-miR	851	-129	88	24
	ID00061.3p-miR	852	-125	91	22
	ID01473.3p-miR	1093	-125	89	23
	ID03167.3p-miR	1399	-123	91	23
	ID01610.5p-miR	1404	-110	91	21
	ID01508.5p-miR	1459	-129	90	23
	ID02945.5p-miR	4274	-108	91	21
	ID02687.5p-miR	4587	-119	92	23

Table 15 (continued).

1	2	3	4	5	6
	ID01388.5p-miR	4587	-119	89	23
	ID01565.5p-miR	4916	-115	93	21
	miR-6768-5p	4928	-115	89	23
	miR-6799-5p	5130	-108	93	20
	ID01819.5p-miR	7062	-119	89	23
ERBB2; 28.0	ID00692.3p-miR	3815	-113	90	22
	miR-4734	3817	-119	90	22
GNAS; 56.4	miR-92a-1-5p	1554	-110	90	23
	ID02093.5p-miR	1590	-123	97	22
	ID00377.3p-miR	1617	-119	90	22
	ID02093.5p-miR	1626	-127	100	22
	ID00377.3p-miR	1653	-119	90	22
	ID02093.5p-miR	1662	-117	92	22
	ID03100.3p-miR	1830	-110	91	22
MLL2; 14.9	ID02169.3p-miR	2690	-119	89	23
	ID01491.3p-miR	3300	-125	91	23
	ID01753.3p-miR	5258	-104	91	21
	miR-4446-3p	5831	-110	90	22
	ID02409.3p-miR	5878	-102	91	21
	ID02328.5p-miR	5945	-119	89	23
	ID02562.3p-miR	5967	-115	90	22
	miR-6765-3p	6653	-113	91	21
	ID02664.5p-miR	7226	-104	92	20
	ID01013.3p-miR	7483	-123	88	24
	miR-1296-5p	9261	-110	90	22
	miR-4763-3p	9991	-125	88	24
	miR-1976	10077	-106	94	20
	ID03483.3p-miR	11514	-115	90	22
	miR-4675	12225	-110	90	22
	ID01562.3p-miR	12237	-117	90	22
	ID02288.5p-miR	12756	-115	89	23
	miR-298	13065	-117	89	24
	ID02045.3p-miR	13623	-123	89	23
	miR-1260b	14398	-98	94	19
MUC5AC; 0.0	ID00481.5p-miR	668	-113	90	22
	ID02341.3p-miR	2037	-113	90	23
	ID02620.5p-miR	3277	-110	90	22
	miR-6808-5p	3342	-115	92	22
	miR-517-5p	3888	-106	91	22
	ID00967.5p-miR	4672	-115	90	22
	ID00967.5p-miR	5791	-115	90	22
	miR-214-3p	15455	-110	91	22
MUC6; 0.0	miR-6815-3p	1206	-104	91	21
	ID02841.5p-miR	1226	-108	93	20

Table 15 (continued).

1	2	3	4	5	6
	miR-939-5p	1318	-121	88	24
	ID00611.5p-miR	1574	-115	89	23
	ID02888.5p-miR	2918	-121	89	23
	ID00564.5p-miR	6488	-110	90	22
	miR-192-3p	7352	-106	91	22
	ID02507.3p-miR	7353	-115	89	24

Within the CDS, the mRNA of the *ARID1A* gene harbors three clusters of binding sites (as detailed in Table 15). The initial cluster ranges from 410 nt to 442 nt and encompasses binding sites for eight miRNAs. This cluster's length of 32 nt is significantly smaller (5.8 times less) than the sum of all binding site lengths (184 nt). Consequently, only one miRNA can bind within this cluster. Three out of eight miRNAs bind with a ΔG value equal to or greater than -130 kJ/mol, indicating a robust interaction within the associations of these miRNAs with the *ARID1A* gene. The expression level of *ARID1A* could serve as a prognostic marker in small intestinal cancer due to its correlation with high-grade tumors when exhibiting low or lost expression. Suppression of this gene's expression by miRNA encourages the development of this disease.

The *GNAS* gene's mRNA possesses a set of binding sites ranging from position 1617 nt to 1647 nt, displaying an average ΔG value of -123 kJ/mol. Another cluster of binding sites spans from 1653 nt to 1683 nt, encompassing a total length of 31 nt, and exhibits an average free energy of interaction at -118 kJ/mol.

Kumagai and coauthors [286] have detected the expression of mucin core proteins (*MUC5AC, MUC6,* and *MUC2*) in the cytoplasm of tumor cells. Seven miRNAs are identified to have binding sites within the mRNA of the *MUC5AC* gene. The mRNA of the *MUC6* gene exhibits two clusters of binding sites, one positioned from 1206 nt to 1245 nt with an average ΔG of -106 kJ/mol. miR-192-3p and ID02507.3p-miR form a cluster between 7352 nt to 7376 nt.

The overall free energy of binding for all miRNAs with mRNAs in the CDS was measured at -116 kJ/mol. Eight miRNA associations with mRNA exhibit a free energy of interaction greater than -130 kJ/mol. These findings suggest their potential as markers for early diagnosis methods in small intestinal cancer development.

Tables 16 and Table C13 illustrate the binding sites of select miRNAs with mRNA genes involved in the development of small intestinal cancer in the 3'UTR region.

Blaker and colleagues propose a frequent and complex role of impaired SMAD4 function in the pathogenesis of small intestinal adenocarcinomas [287]. The interaction between miR-574-5p and the *SMAD4* gene, experimentally established, was verified [288]. The MirTarget program revealed multiple binding sites of miR-574-5p and *SMAD4*, and identified ten alternative miRNAs capable of binding to the

3’UTR of these gene mRNA. Quantitative characteristics of mRNA interaction with ID00714.3p-miR, ID01403.5p-miR, and miR-4253 are presented in Table 16.

Notably, ID00470.5p-miR and miR-574-5p exhibit numerous binding sites within the mRNA of the *CDKN2B* gene in the same region. The miRNA with a higher ΔG value predominates in this binding site and consequently inhibits the expression of this gene.

TP53, a transcription factor regulating the cell cycle and acting as a tumor suppressor, has identified binding sites within its mRNA. ID01838.5p-miR demonstrates multiple consistently located binding sites with an average ΔG value of -113 kJ/mol.

The results obtained indicate that the interactions between the scrutinized miRNA and mRNA could potentially serve as the foundation for selecting miRNA and mRNA associations for diagnosing small intestinal cancer.

Table 16 - The characteristics of miRNA interactions with 3'UTR mRNAs of SIC candidate genes [283].

Gene; NX	miRNA	Start of site, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
1	2	3	4	5	6
ATM; 7.6	ID03006.5p-miR	9778	-121	89	24
	miR-5095	9787	-108	93	21
	miR-619-5p	9793	-119	98	22
	miR-5096	9882	-104	92	21
	miR-5585-3p	9950	-110	95	22
	miR-1273a	11054	-119	90	25
	ID00367.5p-miR	11069	-110	90	22
	miR-1273g-3p	11076	-113	96	21
	miR-1273e	11119	-108	93	22
	miR-5585-5p	11156	-106	91	22
	ID01334.3p-miR	11311	-117	93	22
	miR-1972	11316	-110	90	22
CDKN2B; 19.2	miR-574-5p	1744÷1752 (5)	-108÷-113	89÷93	23
	ID00470.5p-miR	1746÷1752 (4)	-108	89	23
KRAS; 7.3	ID03224.5p-miR	3163	-121	92	23
	miR-1273g-3p	3176	-108	93	21
	miR-1273f	3209	-104	100	19
	miR-7110-5p	4965	-110	93	21
MDM2; 5.2	miR-1273g-3p	2117	-113	96	21
	ID01404.5p-miR	2154	-110	90	23
	ID01815.5p-miR	2257	-106	89	23
	ID01360.3p-miR	2476	-104	91	21
	miR-1273g-3p	2486	-106	91	21
	miR-1273e	2521	-108	93	22
	miR-3929	3013	-115	93	23

Table 16 (continued).

1	2	3	4	5	6
	ID01838.5p-miR	3214	-110	88	24
	miR-1285-3p	3218	-106	91	22
	miR-1273g-3p	6739	-113	96	21
MLL2; 14.9	ID02459.5p-miR	16748	-110	91	21
	ID03342.3p-miR	16931	-119	92	23
	ID02547.5p-miR	17065	-113	91	21
	miR-5196-5p	17100	-113	90	22
	ID02262.3p-miR	17703	-119	89	23
	ID01254.5p-miR	17780	-110	91	21
	ID02690.5p-miR	17796	-125	89	24
	miR-7110-5p	17893	-108	91	21
	ID00564.5p-miR	17894	-113	91	22
	ID02328.5p-miR	19010	-119	89	23
SMAD4; 8.6	ID01838.5p-miR	4291	-113	90	24
	miR-1273g-3p	4312	-110	95	21
	ID01656.3p-miR	4342	-115	89	23
	miR-1273d	4346	-119	87	25
	ID01404.5p-miR	4349	-115	93	23
	miR-1972	4552	-110	90	22
	ID02732.3p-miR	7721	-123	91	23
	miR-574-5p	7740÷7756 (8)	-108÷-113	89÷93	23
	ID00470.5p-miR	7744÷7752 (5)	-108	89	23
	ID00106.5p-miR	7825	-106	91	22
	miR-574-5p	7830	-108	89	23
	ID01592.3p-miR	8227	-117	89	23
SOAT1; 4.9	miR-1273a	3804	-117	89	25
	miR-1273a	4734	-121	92	25
	miR-1273c	4736	-110	91	22
	ID01815.5p-miR	4953	-106	89	23
	miR-1285-3p	4963	-113	96	22
	miR-5684	4974	-98	92	20
	miR-1273g-3p	5485	-106	91	21
	miR-1273d	5520	-121	89	25
	ID01404.5p-miR	5523	-110	90	23
	miR-1273e	5529	-106	91	22
	miR-619-5p	5868	-110	91	22
	miR-5095	5997	-106	91	21
	miR-5585-3p	6010	-108	93	22
	miR-1285-5p	6151	-102	91	21
	miR-1303	6162	-106	91	22
TP53; 10.7	ID00548.3p-miR	1393	-115	89	23
	ID02379.3p-miR	1397	-119	89	24
	miR-1273c	2297	-110	91	22
	miR-1285-3p	2301	-110	95	22

Table 16 (continued).

1	2	3	4	5	6
	miR-1273g-3p	2317	-106	91	21
	miR-1273h-5p	2351	-106	91	21
	ID01838.5p-miR	2459	-115	92	24
	ID00785.5p-miR	2520	-113	90	23

The identification of ten miRNAs with binding sites distributed across the entire 3'UTR of *MLL2* gene mRNA provides insights into the diverse regulation of this gene. The clustering of binding sites formed by ID01254.5p-miR and ID02690.5p-miR, along with the presence of additional non-overlapping sites, highlights the complex interaction landscape within the 3'UTR of this gene.

Moreover, the presence of single miRNA binding sites in mRNAs of *APC*, *ASXL1*, and *MUC6* genes suggests a pattern where each miRNA can target one or multiple genes, and reciprocally, a single gene can be targeted by one or multiple miRNAs. This interplay between miRNAs and their target genes could potentially serve as diagnostic markers for early disease detection.

The average free energy of binding of miRNAs with mRNAs in the 3'UTR being -110 kJ/mol is indicative of the strength of these interactions. Notably, the presence of multiple binding sites for ID00470.5p-miR and miR-574-5p in the *CDKN2B* and *SMAD4* genes respectively suggests these associations as potential markers for diagnosing small intestinal cancer.

Furthermore, the identification of miRNAs forming fully complementary binding sites highlights the possibility of mRNA degradation rather than mere inhibition. The presence of slightly different binding sites for the same miRNA in different mRNA target genes showcases resistance to point mutations (single nucleotide polymorphisms) in the miRNA binding site. Additionally, the discovery of closely related miRNAs with similar sequences capable of binding to the same site with varying hybridization energy further illustrates the intricate nature of miRNA-mRNA interactions.

The identification of 79 key associations between miRNAs and mRNAs, which exhibit a free energy of interaction of -125 kJ/mol or higher, holds significant promise for diagnostic markers and potential therapeutic approaches in addressing small intestinal cancer.

Observing the preferential binding of miRNAs to different regions of mRNA, such as the 5'UTR, CDS, and 3'UTR, in genes like *ARID1A, ASXL1, EGFR, KRAS, LRP1B, PIK3CA, PTEN, CDKN2B, MSH6,* and *SMAD4,* sheds light on potential regulatory mechanisms. This preferential binding in specific mRNA regions suggests the possibility of significant mRNA translation inhibition if the concentrations of corresponding miRNAs match or exceed the concentrations of mRNAs.

Considering the role of miRNAs as endogenous regulators of gene expression, these findings open avenues for potential diagnostic markers and therapeutic targets in small intestinal cancer. Understanding these interactions at a molecular level can

offer crucial insights into the mechanisms underlying the pathogenesis of small intestinal cancer, thereby paving the way for the development of novel and effective diagnostic methods and targeted therapies for patients affected by this disease.

3.2.4 The characteristics of the miRNAs interactions with mRNAs of genes associated with colorectal cancer

The eight candidate genes have miRNA binding sites in the 5′UTR (Tables 17 and Table C14). The analysis of miRNA binding sites in the 5′UTR of candidate genes, such as the *EZH2* gene, reveals intriguing patterns of miRNA-mRNA interactions. The cluster of miRNA binding sites in the *EZH2* mRNA ranged from 112 nt to 150 nt, with the total length of five miRNA binding sites reaching 108 nt.

Table 17 - The characteristics of miRNA interactions with 5'UTR mRNAs of CRC candidate genes [289, 290].

Gene; NX	miRNA	Start of site, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
1	2	3	4	5	6
EZH2; 2.7	ID01972.5p-miR	112	-113	95	20
	ID02761.3p-miR	120	-138	93	24
	ID01643.3p-miR	122	-119	90	22
	ID03409.5p-miR	125	-115	93	20
	ID00811.3p-miR	127	-125	88	24
GSK3B; 5.8	ID02187.5p-miR	3	-123	89	23
	ID03229.5p-miR	4	-123	92	22
	ID01804.3p-miR	5, 12	-134	91	23
	ID00756.3p-miR	8	-123	89	23
	ID01041.5p-miR	8	-132	90	24
	ID00457.3p-miR	8, 11	-123, -129	91, 95	22
	ID03367.5p-miR	8, 11	-117, -119	93, 95	20
	ID00061.3p-miR	8, 17	-125, -136	91, 98	22
	ID00296.3p-miR	9	-138	88	25
	ID03151.3p-miR	9	-115	93	20
	ID01702.3p-miR	9, 12	-138, -140	92, 93	24
	ID01879.5p-miR	10,13	-123	91	22
	ID02064.5p-miR	10,13	-129, -136	90, 94	23
	miR-3960	10÷14 (4)	-115	92	20
	ID01873.3p-miR	11	-123	94	21
	ID02522.3p-miR	12	-127	91	23
	ID02429.3p-miR	14	-125	92	23
	ID02770.5p-miR	14	-117	93	20
	ID01652.3p-miR	15	-125	89	23
	ID02538.3p-miR	15	-121	90	22
	ID00061.3p-miR	17	-129	94	22
	ID01041.5p-miR	354	-129	88	24

Table 17 (continued).

1	2	3	4	5	6
	ID00457.3p-miR	356	-125	92	22
	ID03367.5p-miR	356	-115	92	20
	ID02770.5p-miR	359	-115	92	20
	ID02036.3p-miR	361	-115	92	20
MTA1; 20.6	ID03332.3p-miR	37÷43 (3)	-134	90	24
	ID01310.3p-miR	38, 41	-121	92	22
	ID01810.3p-miR	52	-115	89	23
	ID00102.3p-miR	133	-117	95	20
	ID01003.3p-miR	150	-132	89	24
	ID01458.5p-miR	156	-136	93	23
	ID00020.5p-miR	159	-138	93	24
	ID02260.5p-miR	162	-127	91	22
ODC1; 24.3	ID00756.3p-miR	9	-129	94	23
	ID00722.5p-miR	11	-115	95	20
	ID03305.5p-miR	11	-113	93	20
	ID01804.3p-miR	13	-132	90	23
	ID02187.5p-miR	14	-123	89	23
	ID00457.3p-miR	15, 21	-123, -125	91, 92	22
	ID02036.3p-miR	17	-117	93	20
	ID02084.3p-miR	17	-140	93	24
	ID02064.5p-miR	17÷23 (3)	-129 ÷ -140	90÷97	23
	miR-3960	18, 21	-115	92	20
	ID01652.3p-miR	19	-125	89	23
	ID02538.3p-miR	19	-123	92	22
	ID01702.3p-miR	19, 22	-134, -142	89, 94	24
	ID02229.3p-miR	21	-121	92	21
	ID02499.3p-miR	21	-119	92	21
	ID01157.5p-miR	22	-117	93	20
	ID01377.3p-miR	23	-121	95	20
	ID01569.3p-miR	23	-117	93	20
	ID00061.3p-miR	24	-125	91	22
ONECUT2; 0.1	ID01099.5p-miR	290	-134	91	23
	ID01702.3p-miR	292	-136	90	24
	ID01184.3p-miR	301	-117	93	20
	ID03222.5p-miR	302	-113	91	21
	ID01653.5p-miR	303	-132	89	24
	miR-3960	304, 306	-115, -117	92, 93	20
	ID01453.3p-miR	306	-136	90	24
	ID01184.3p-miR	310	-123	98	20
PIK3CA; 4.2	ID00071.3p-miR	4	-115	92	20
	ID01190.5p-miR	4	-140	92	24
	ID01702.3p-miR	4	-140	93	24
	ID01895.5p-miR	4	-132	89	24
	ID02595.5p-miR	4	-115	92	20
	ID00966.5p-miR	6	-136	90	24

Table 17 (continued).

1	2	3	4	5	6
	ID00030.3p-miR	7	-121	90	22
	ID01804.3p-miR	8	-134	91	23
	ID01873.3p-miR	8	-123	94	21
	ID03367.5p-miR	10	-115	92	20
SRM; 24.0	ID02101.3p-miR	31	-119	89	23
	ID00517.5p-miR	59	-115	92	20
	ID01464.5p-miR	60	-123	92	22
	ID01816.3p-miR	65	-134	89	24
	ID01901.5p-miR	66	-121	92	21
	ID02770.5p-miR	68	-115	92	20
	ID01293.5p-miR	70	-127	94	22
	ID00009.3p-miR	72	-115	92	20
TCF4; 5.4	ID02761.3p-miR	451	-132	89	24
	ID00278.3p-miR	452, 455	-123, -132	89, 95	23
	ID03396.5p-miR	477	-125	89	24
	ID01839.3p-miR	498	-123	89	23
	ID03445.3p-miR	505	-123	91	23
	ID03332.3p-miR	508	-132	89	24

This compact clustering reduces the proportion of binding sites relative to the entire 5′UTR, creating competition among miRNAs for binding within the cluster. Competition for binding sites within the cluster is influenced by various factors, notably the concentration of miRNAs and the free energy of their interaction with the mRNA. For instance, ID02761.3p-miR exhibits a higher free energy of -138 kJ/mol when binding to the *EZH2* mRNA compared to other miRNAs. This higher energy level suggests a stronger interaction with the mRNA, providing an advantage for ID02761.3p-miR to bind and regulate *EZH2* mRNA, particularly in conditions where the concentrations of different miRNAs are equivalent. Evaluating the free energy of miRNA-mRNA interactions becomes crucial in understanding the dynamics of miRNA binding. It helps determine the preferential binding of specific miRNAs to target mRNAs, considering their concentrations and interaction energies. This knowledge is vital in deciphering the regulatory role of miRNAs in gene expression.

Absolutely, the outcome of miRNA-mRNA interactions greatly relies on the balance between miRNA and mRNA concentrations. When the concentration of miRNA is substantially lower than that of mRNA, even with a strong binding affinity between miRNA and mRNA, the translation inhibition may be insignificant. Conversely, higher concentrations of miRNA compared to mRNA can lead to a more pronounced suppression of mRNA translation, even with an average interaction between the miRNA and mRNA. Assessing the result of interactions between multiple miRNAs and one mRNA or several mRNAs of candidate genes becomes notably challenging due to these concentration dynamics. The difficulty arises from establishing the impact of various miRNAs on mRNA when considering their

differing concentrations. Moreover, numerous published studies lack crucial information concerning the concentrations of miRNA and mRNA, as well as the detailed interaction energies of all miRNA nucleotides and mRNA binding sites. This absence of information compromises the adequacy and reliability of such studies when attempting to comprehend the outcomes of miRNA-mRNA interactions. Therefore, conducting analyses of miRNA-mRNA interactions under comparable concentrations of all participants in the interaction becomes essential for drawing accurate conclusions and understanding the regulatory mechanisms involved.

Certainly, the study of miRNA interaction with the 5′UTR mRNAs of *GSK3B*, *ONECUT2, PIK3CA,* and *TCF4* genes revealed a clustered organization in the mRNA sequences. For the *GSK3B* gene mRNA with a 5'UTR length of 983 nt, the analysis indicated the presence of clustered binding sites. Specifically, the first cluster of binding sites ranged from 3 nt to 38 nt, totaling 36 nt in length, and exhibited an average ΔG of -126 kJ/mol. Although this cluster's length was 36 nt, the total length of the 32 binding sites within this cluster amounted to 705 nt, which is approximately 19.6 times larger than the cluster length. This clustering phenomenon in the *GSK3B* gene's 5'UTR, despite its length of 983 nt, suggests the gene's propensity for compaction, leading to potential competition among the involved miRNAs for the available binding sites. The second cluster spanned from 354 nt to 380 nt, containing binding sites totaling 27 nt in length. Consequently, instead of occupying the entire 983 nt length of the 5'UTR (which accounts for 97% of the total length), the binding sites within these clusters only occupied 3% of the total length. This clustering of binding sites within the 5'UTR of the GSK3B gene mRNA highlights the potential for competition among miRNAs for binding sites and suggests a compact organization within the gene's untranslated region, despite its considerable length.

The interaction analysis of miRNAs with the *MTA1* and *ODC1* gene mRNAs reveals intriguing patterns of clustered binding sites and considerations about miRNA-mRNA complex formation and stability. ID03332.3p-miR, ID01310.3p-miR, and ID01810.3p-miR collectively form a cluster of binding sites within the *MTA1* gene mRNA from 37 nt to 74 nt, exhibiting an average free energy of interaction at -126 kJ/mol. Another cluster of binding sites is situated from positions 133 nt to 183 nt, totaling 51 nt in length, and demonstrating a stronger average interaction energy of -130 kJ/mol. These characteristics indicate robust interactions between these miRNAs and the *MTA1* gene mRNA.

Nineteen miRNAs with overlapping nucleotide sequences were identified with binding sites in the ODC1 gene mRNA. A cluster composed of twenty-four binding sites starts from 9 nt to 43 nt, with a total length of 35 nt and an average ΔG of -125 kJ/mol. Despite the cumulative length of all 24 binding sites being 523 nt, longer than the length of the 5′UTR (334 nt), this cluster's limited space (35 nt) implies a competitive scenario where only a few miRNAs can occupy these positions.

The occupancy within this cluster will likely be dominated by miRNAs exhibiting interaction energies above the average.

Considering varying miRNA concentrations, the characteristics of their interactions with mRNAs, coupled with their concentrations, play a pivotal role in determining the stability and lifetime of these miRNA-mRNA complexes. Hence, monitoring and controlling the concentrations of both miRNAs and mRNAs are crucial to identify primary translation inhibitory miRNAs. This intricate interplay between miRNA binding sites, their concentrations, and the mRNA sequences emphasizes the necessity of comprehensively understanding these interactions to ascertain the regulatory effects on translation inhibition and miRNA-mRNA complex stability.

Eight miRNAs target the *ONECUT2* gene mRNA, forming a cluster of binding sites from 290 nt to 329 nt, exhibiting an average ΔG value of -125 kJ/mol. This cluster suggests a significant association between these miRNAs and the mRNA of the *ONECUT2* gene. Ten miRNAs establish a cluster of binding sites within the *PIK3CA* gene mRNA's 5′UTR, spanning from positions 4 nt to 29 nt, and displaying an average ΔG value of -128 kJ/mol. This cluster indicates a strong interaction between these miRNAs and the *PIK3CA* gene mRNA. Eight miRNAs form a cluster of binding sites in the *SRM* gene mRNA, ranging from 31 nt to 91 nt. The average ΔG value for this cluster is -121 kJ/mol, reflecting a notable affinity and interaction between these miRNAs and the *SRM* gene mRNA. The *TCF4* gene mRNA exhibits clusters of binding sites positioned from 451 nt to 531 nt, spanning a length of 81 nt. The average ΔG value for this cluster is -127 kJ/mol, indicating a robust interaction between the miRNAs and the TCF4 gene mRNA at these locations.

Based on the data acquired on miRNA-mRNA interactions with specific clusters and a ΔG value of -130 kJ/mol or higher, several genes and miRNA associations can be recommended for further consideration in colorectal cancer diagnosis: mRNA of *EZH2* gene with ID02761.3p-miR; mRNA of *GSK3B* gene with ID01804.3p-miR, ID01041.5p-miR, ID00061.3p-miR, ID00296.3p-miR, ID01702.3p-miR, ID02064.5p-miR; mRNA of *MTA1* gene with ID03332.3p-miR, ID01003.3p-miR, ID01458.5p-miR, ID00020.5p-miR; mRNA of *ODC1* gene with ID01804.3p-miR, ID02084.3p-miR, ID02064.5p-miR, ID01702.3p-miR; mRNA of *ONECUT2* gene with ID01099.5p-miR, ID01702.3p-miR, ID01653.5p-miR, ID01453.3p-miR; mRNA of *PIK3CA* gene with ID01190.5p-miR, ID01702.3p-miR, ID01895.5p-miR, ID00966.5p-miR, ID01804.3p-miR; mRNA of SRM gene with ID01816.3p-miR; mRNA of *TCF4* Gene with ID02761.3p-miR, ID00278.3p-miR, ID03332.3p-miR.

Notably, miRNAs such as ID03332.3p-miR, ID01804.3p-miR, and ID01702.3p-miR are associated with multiple genes, emphasizing their potential significance in colorectal cancer diagnosis. Additionally, in the CDS of mRNA sequences related to colorectal cancer candidate genes, clusters and multiple binding sites for various miRNAs were identified: in mRNA of *CDX2* gene six miRNAs showed binding site clusters distributed across two regions. Several miRNAs (miR-877-3p, ID03238.3p-miR, ID01352.3p-miR, ID00777.3p-miR) interacted with a cluster in the *CCND1* gene mRNA, indicating strong associations with an average ΔG = -116 kJ/mol. These

identified miRNA-mRNA interactions in clusters and CDS regions highlight specific genes and miRNAs crucial for potential diagnostic applications in colorectal cancer.

Table 18 - The characteristics of miRNA interactions with CDS mRNAs of CRC candidate genes [289, 290].

Gene; NX	miRNA	Start of site, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
CDX2; 18,9	ID01895.5p-miR	416	-138	93	24
	ID00522.5p-miR	419	-125	89	23
	ID02052.5p-miR	608	-136	91	24
	ID01804.3p-miR	608	-132	90	23
	ID01041.5p-miR	609	-129	88	24
	miR-3960	613	-117	93	20
CCND1; 9.7	miR-877-3p	1021	-108	93	21
	ID03238.3p-miR	1022	-119	92	23
	ID01352.3p-miR	1024	-115	90	23
	ID00777.3p-miR	1025	-113	90	23
ONECUT2; 0.1	ID00032.5p-miR	449	-113	90	22
	ID01868.3p-miR	456, 477	-129, -134	88, 91	24
	ID03445.3p-miR	459	-121	89	23
	ID01310.3p-miR	462, 483	-119, -123	90, 94	22
	ID02692.3p-miR	466	-125	89	23
	ID00798.3p-miR	467÷488 (4)	-132 ÷ -136	89÷91	24
	ID03291.3p-miR	470	-119	89	23
	ID01508.5p-miR	472, 481	-129	90	23
	ID02488.3p-miR	473	-121	89	23
	ID02624.3p-miR	476, 485	-119	92	21
	ID00879.3p-miR	485	-123	89	23
	ID02761.3p-miR	491	-132	89	24
	ID02094.3p-miR	509	-113	90	22
	ID02076.5p-miR	615	-113	91	21
	ID01310.3p-miR	623	-119	90	22
	ID01810.3p-miR	625	-115	89	23
	ID02043.3p-miR	637	-119	90	22
SIRT1; 6.6	ID01560.3p-miR	264	-123	89	23
	ID03332.3p-miR	280, 287	-132, -138	89, 93	24
	ID00278.3p-miR	294	-123	89	23
	ID00811.3p-miR	300	-125	88	24

In the mRNA of certain genes, the miRNA binding sites were organized into clusters with partially overlapping nucleotide sequences. For instance: The *ONECUT2* gene mRNA presented two distinct binding site clusters (as outlined in Table 18). The first cluster spanned from 449 nt to 530 nt, accommodating 13 miRNA binding sites totaling 459 nt in length. Comparatively, this cluster was 5.5

times longer than the subsequent 82-nt cluster. The second cluster, located between 615 nt to 658 nt, contained binding sites totaling 88 nt in length. Consequently, rather than occupying 793 nt (81% of the 5'UTR length), the binding sites encompassed only 9% of this length. In the *SIRT1* gene mRNA, five miRNA binding sites were identified (detailed in Table 18). The binding site cluster from 264 nt to 323 nt was one-third the length of the combined 118-nt miRNA binding sites.

Ten associations between miRNAs and target genes exhibit a free energy of miRNA interaction within the CDS greater than -130 kJ/mol, making them potential markers for colorectal cancer: *CDX2* gene and ID01895.5p-miR, ID02052.5p-miR, ID01804.3p-miR; *MSH6* gene and ID00156.5p-miR; *MTA1* gene and ID01702.3p-miR; *ONECUT2* gene and ID01868.3p-miR, ID00798.3p-miR, ID02761.3p-miR; *SIRT1* gene and ID03332.3p-miR; *TGFBR1* gene and ID03332.3p-miR

The abundance of binding sites located in the 3'UTR compared to those in the 5'UTR and CDS regions is attributed to GU-rich elements present in the 3'UTR of mRNA transcripts. These GU-rich elements play various roles in gene expression regulation by influencing mRNA stability and translation.

ID00436.3p-miR, ID01030.3p-miR, and miR-466 display multiple binding sites in the mRNAs of *ADAMTS4, CYP24A1, FLT1, GSK3B, IRS1*, and *MACC1* genes (as listed in Table 19). In the *ADAMTS4* gene mRNA, miR-1273a and miR-1273g-3p form a binding site cluster from 3568 nt to 3610 nt, while another cluster is formed by ID02880.3p-miR and miR-1972 from 3989 nt to 4012 nt. Additionally, miR-574-5p and ID00470.5p-miR demonstrate multiple binding sites in the mRNAs of *CCND1, IGF1, IGF2, PPP2R1B, SMAD4, XRCC1,* and *ZEB1* genes (as detailed in Table 19).

Certainly, here's a breakdown of the miRNA binding sites in specific gene mRNAs:

In the mRNA of *CDC27* gene the miR-548 family forms two distinct clusters of binding sites. The first cluster spans from 4971 nt to 4998 nt with an average ΔG value of -104 kJ/mol, while the second cluster is from 5012 nt to 5041 nt with an average ΔG of -102 kJ/mol. In the mRNA of *SMAD4* gene a binding site cluster is identified from 4342 nt to 4371 nt with an average ΔG of -116 kJ/mol. In mRNA of *TP53* gene a cluster formed by miR-1273c, miR-1285-3p, and miR-1273g-3p exists from 2297 nt to 2337 nt with an average ΔG of -108 kJ/mol. In mRNA of *CYP2W1* gene a seven miRNAs target this mRNA, with miR-4739 exhibiting higher free energy than miR-619-5p. A cluster of binding sites from 2160 nt to 2219 nt (59 nt in length) involves four miRNAs, which collectively occupy 1.5 times the length of the cluster, leading to competitive binding. In mRNA of *KIAA1456* gene this mRNA contains 21 binding sites, forming multiple clusters. The miR-5096 exhibits complete complementarity, indicating strong control over the expression of this gene. In mRNA of *SLC26A2* gene nine miRNAs target this mRNA, forming three clusters of two binding sites each. The presence of these clusters significantly reduces the proportion of binding sites in the total mRNA length. There is apparent competition

between miR-1285-5p and ID01237.3p-miR for binding to this mRNA. In mRNA of *SPATA13* gene six miRNAs bind to this mRNA, with binding sites of ID03024.5p-miR, together with miR-619-5p, forming a cluster.

These observations highlight the complex interactions and competitive binding among miRNAs targeting specific mRNA sequences, indicating potential regulatory roles in gene expression and molecular control mechanisms.

Table 19 - The characteristics of miRNA interactions with 3'UTR mRNAs of CRC candidate genes [289, 290].

Gene; NX	miRNA	Start of site, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
1	2	3	4	5	6
ADAMTS4; 7.9	ID00436.3p-miR	3380, 3387	-104	89	23
	ID01030.3p-miR	3383	-108	89	23
	miR-466	3383	-110	95	23
	miR-1273a	3568	-117	89	25
	miR-1273g-3p	3590	-108	93	21
	ID02880.3p-miR	3989	-108	91	22
	miR-1972	3991	-115	93	22
AGMAT; 2.1	miR-5095	2201	-106	91	21
	miR-619-5p	2207	-121	100	22
	ID01640.5p-miR	2541	-119	90	24
	miR-6719-3p	2545	-104	91	22
	miR-4452	2552	-102	89	23
CCND1; 9.7	miR-574-5p	2593÷2597 (3)	-108 ÷ -113	89÷93	23
	ID00470.5p-miR	2595, 2597	-108	89	23
CDC27; 6.7	ID03240.3p-miR	4971	-106	91	22
	miR-548 family	4974÷4980 (7)	-85 ÷ -119	95÷97	19÷25
	miR-548 family	5012÷5018 (17)	-96 ÷ -108	90÷96	20÷23
CYP2W1; 1.4	miR-7110-5p	1546	-108	91	21
	miR-4739	1739	-125	87	25
	ID01811.5p-miR	2160	-113	90	22
	miR-5095	2170	-106	91	21
	miR-619-5p	2176	-121	100	22
	ID00913.5p-miR	2196	-117	92	23
	miR-5096	2248	-108	96	21
CYP24A1; 0.1	ID02513.5p-miR	2622	-102	91	22
	ID01030.3p-miR	2627, 2631	-108	89	23
	ID01727.5p-miR	2632	-104	89	23
	miR-466	2647	-106	91	23
	ID00436.3p-miR	2651	-104	89	23
FLT1; 3.4	ID01030.3p-miR	6909÷6923 (7)	-108 ÷ -110	89÷91	23
	miR-466	6911÷6937 (9)	-104 ÷ -108	89÷93	23
	ID00436.3p-miR	6913÷6925 (7)	-104	89	23

Table 19 (continued).

1	2	3	4	5	6
GSK3B; 5.8	ID01030.3p-miR	4666÷4680 (4)	-108 ÷ -113	89÷93	23
	miR-466	4670÷4682 (6)	-104 ÷ -106	89÷91	23
	ID00436.3p-miR	4674÷4684 (3)	-104	89	23
	ID01727.5p-miR	4683	-106	91	23
IGF1; 2.3	ID00470.5p-miR	4042÷4058 (9)	-108	89	23
	miR-574-5p	4042÷4062 (11)	-108 ÷ -113	89÷93	23
	miR-1273f	6042	-102	98	19
	miR-1273d	6043	-119	87	25
	miR-1273e	6052	-108	93	22
IGF2; 7.4	ID00470.5p-miR	2286÷2351 (6)	-108 ÷ -113	89÷93	23
	miR-574-5p	2288, 2290	-108	89	23
	ID02299.5p-miR	2302	-100	94	21
	miR-574-5p	2332, 2353	-108	89	23
	ID02299.5p-miR	2346	-100	94	21
	miR-574-5p	2397÷2408 (4)	-108 ÷ -113	89÷93	23
	ID00470.5p-miR	2404, 2412	-110	91	23
	ID00470.5p-miR	2442÷2463 (3)	-110 ÷ -113	89÷93	23
	ID02299.5p-miR	2458	-100	94	21
	miR-574-5p	2465, 2484	-108	89	23
	ID00470.5p-miR	2520÷2539 (3)	-108	89	23
	miR-574-5p	2522	-108	89	23
	ID00470.5p-miR	2571	-108	89	23
	miR-297	2604	-100	92	21
	miR-574-5p	2614	-110	91	23
	ID00470.5p-miR	2620÷2672 (4)	-108 ÷ -110	89÷91	23
	ID00470.5p-miR	2704, 2725	-108	89	23
	miR-574-5p	2727, 2731	-108, -110	89, 91	23
IRS1; 5.7	miR-466	7317÷7337 (6)	-104 ÷ -108	89÷93	23
	ID00436.3p-miR	7321÷7365 (5)	-104÷ -106	89÷91	23
	ID01030.3p-miR	7321, 7333	-108	89	23
	ID01727.5p-miR	7334	-106	91	23
	ID03181.3p-miR	7351	-98	90	22
KIAA1456; 1.8	ID01334.5p-miR	2522	-110	90	22
	miR-619-5p	2536	-121	100	22
	ID02199.5p-miR	2611	-113	90	23
	miR-1303	2621	-106	91	22
	miR-1285-5p	2768	-102	91	21
	miR-1273a	3875	-121	92	25
	miR-1273c	3877	-113	93	22
	miR-1273g-3p	3897	-106	91	21
	ID01815.5p-miR	4037	-106	89	23
	ID01334.3p-miR	4132	-115	92	22
	ID02017.3p-miR	4133	-115	90	22
	miR-1972	4137	-113	91	22

Table 19 (continued).

1	2	3	4	5	6
	miR-5096	5137	-113	100	21
	miR-619-5p	5196	-117	96	22
	ID01836.5p-miR	5286	-113	90	23
	miR-1285-5p	5302	-102	91	21
	ID01360.3p-miR	8419	-104	91	21
	ID00367.5p-miR	8422	-115	93	22
	miR-1273g-3p	8429	-106	91	21
	ID02991.3p-miR	8854	-89	91	21
KISS1; 0.0	ID00415.5p-miR	622	-121	92	22
	ID02601.5p-miR	623	-121	95	21
	ID03321.3p-miR	625	-117	92	21
	miR-4433-5p	627	-110	91	21
MACC1; 0.3	ID02513.5p-miR	2908	-104	92	22
	miR-466	2914	-104	89	23
	ID01030.3p-miR	2962, 2976	-108	89	23
	ID01727.5p-miR	2963÷2981 (3)	-104	89	23
	miR-466	2966÷2990 (4)	-104 ÷ -106	89÷91	23
	ID02868.3p-miR	2982	-113	90	23
	ID00436.3p-miR	2990	-106	91	23
	ID01030.3p-miR	3949÷3965 (8)	-108 ÷ -110	89÷91	23
	miR-466	3951÷3967 (9)	-106 ÷ -108	91÷93	23
	ID00436.3p-miR	3953÷3967 (8)	-104 ÷ -106	89÷91	23
	miR-1273f	5717	-102	98	19
	miR-1273d	5718	-123	91	25
PPP2R1B; 3.7	miR-5096	3054	-113	100	21
	miR-574-5p	4561÷4581 (11)	-113	93	23
	ID00470.5p-miR	4561÷4581 (11)	-108	89	23
SLC26A2; 81.9	ID01838.5p-miR	4438	-119	95	24
	miR-1285-3p	4442	-110	95	22
	miR-7851-3p	4495	-108	91	22
	miR-619-5p	5066	-121	100	22
	miR-619-5p	5202	-110	91	22
	miR-5585-3p	5209	-113	96	22
	ID02175.3p-miR	5257	-110	91	22
	miR-1285-5p	5308	-102	91	21
	ID01237.3p-miR	5308	-113	88	24
SMAD4; 9.2	ID01656.3p-miR	4342	-115	89	23
	miR-1273d	4346	-119	87	25
	ID01404.5p-miR	4349	-115	93	23
	miR-574-5p	7740÷7756 (8)	-108 ÷ -113	89÷93	23
	ID00470.5p-miR	7744÷7752 (5)	-108	89	23
	ID00106.5p-miR	7825	-106	91	22
	miR-574-5p	7830	-108	89	23
SPATA13; 8.8	ID01412.5p-miR	3846	-113	91	22

Table 19 (continued).

1	2	3	4	5	6
	ID02484.3p-miR	4057	-104	91	21
	ID03024.5p-miR	5010	-110	91	22
	miR-619-5p	5020	-121	100	22
	miR-619-5p	5155	-117	96	22
	miR-6510-5p	5931	-110	90	22
TGFB1; 19.2	miR-6089	2060, 2065	-132, -136	89, 91	24
	ID03306.3p-miR	2060	-123	94	21
	ID01382.3p-miR	2062	-113	93	20
	ID03208.5p-miR	2066	-125	88	24
	miR-4651	2087	-113	95	20
	ID00978.5p-miR	2089	-119	90	22
	ID00296.3p-miR	2093	-140	89	25
TGFBR1; 10.9	ID02984.3p-miR	112	-123	92	22
	ID03332.3p-miR	127÷133 (3)	-134 ÷ -136	90÷91	24
	ID01310.3p-miR	128, 131	-121	92	22
TP53; 9.3	ID00548.3p-miR	1393	-115	89	23
	ID02379.3p-miR	1397	-119	89	24
	miR-1273c	2297	-110	91	22
	miR-1285-3p	2301	-110	95	22
	miR-1273g-3p	2317	-106	91	21
XRCC1; 10.1	miR-574-5p	2033÷2055 (11)	-108÷ -113	89÷93	23
	ID00470.5p-miR	2035÷2055 (9)	-108 ÷ -113	89÷93	23
ZEB1; 20.3	miR-574-5p	3587÷3605 (10)	-113	93	23
	ID00470.5p-miR	3587÷3605 (10)	-108	89	23
UQCRB; 37.7	miR-619-5p	1269	-121	100	22
	ID00913.5p-miR	1289	-115	90	23
	miR-5096	1343	-110	98	21
	miR-5585-3p	1412	-108	93	22
	ID00695.3p-miR	2233	-127	88	24

The *IGF2* gene displays seven distinct clusters of binding sites within its mRNA, specifically for ID00470.5p-miR and miR-574-5p. These multiple binding sites contribute to the robust expression of the *IGF2* gene.

In the *AGMAT* gene mRNA, binding sites for miR-5095 and miR-619-5p align in a cluster but differ by six nucleotides in their binding sites. Similar binding sites are observed in other genes, such as MTHFR. Within the *AGMAT* gene mRNA, both miR-5096 and miR-619-5p binding sites share a length difference of 74 nucleotides.

The *TGFB1* gene mRNA contains a cluster of binding sites for seven miRNAs spanning from 2060 nt to 2117 nt, with an average free energy of interaction at -125 kJ/mol.

Multiple binding site clusters involving miR-574-5p and ID00470.5p-miR are found in the *PPP2R1B, XRCC1,* and *ZEB1* gene mRNAs. Associations of miR-619-5p with *ERBB3* and miR-5096 with *PPP2R1B* exhibit complete complementarity

between the miRNA and mRNA, justifying their potential use in disease diagnosis. Furthermore, clusters of multiple binding sites involving various miRNAs are recommended for the diagnosis of colorectal cancer.

The *UQCRB* gene mRNA contains only one cluster that includes the miR-619-5p binding site (as indicated in Table 19). Additionally, single binding sites for three other miRNAs are characterized by ΔG/ΔGm values ranging from 88% to 98%.

The presence of miR-619-5p binding sites within clusters for all five genes (*CYP2W1, KIAA1456, SLC26A2, SPATA13, and UQCRB*) indicates a deliberate association of miR-619-5p with these specific genes. Furthermore, miR-619-5p strongly suppresses the expression of these genes while seemingly excluding other miRNAs from regulating their expression. Notably, in tumor cells, the synthesis of mir-619-5p is significantly higher compared to normal cells [291-295].

Existing programs, which primarily rely on the seed region of a miRNA sequence, fail to accurately detect miRNA binding sites in mRNA while determining the interaction characteristics of all nucleotides. Figure 22 illustrates the interaction characteristics across the entire length of miRNA with its binding site in mRNA. This program's unique features have enabled the discovery of novel characteristics concerning miRNA-mRNA interactions, as detailed in this paper. An essential benefit of this investigation is the utilization of endogenous biological molecules for the creation of diagnostic nanotechnologies, establishing crucial foundations for CRC diagnostics via nanotechnology.

Gene, miRNA, characteristics of binding	Gene, miRNA, characteristics of binding
5'UTR	

CDC42,miR-1273g-3p, 136 nt,
-113 kJ/mole, 96%, 21 nt
5' - C**C**CAGGCUGGAGUGCAGUGG**C** - 3'
  ||||||||||||||||||||
3' - G**A**GUCCGACCUCACGUCACC**A** - 5'
	GSK3B, ID01702.3p-miR, 12 nt,
	-140 kJ/mole, 93%, 24 nt
	5' - CGCCGCCGCCGCCGCCGCCGCCCG - 3'
	  | |||||||||||||||||||
	3' - GGGGCGGCGGCGGCGGCGGCGGGA - 5'

GSK3B, ID02064.5p-miR, 13 nt,
-136 kJ/mole, 94%, 23
5' - GCCGCCGCCGCCGCCGCCG**C**CCG - 3'
  |||||||||||||||||| ||||
3' - CGGCGGCGGCGGCGGCGGGG**A**GGC - 5
	GSK3B, ID00296.3p-miR, 9 nt,
	-138 kJ/mole, 88%, 25 nt
	5' - **U**GCC**G**CCGCCGCCGCCGCCGCCGCCGCC - 3'
	  | ||||| |||||||||||||||
	3' - **G**GGG**U**GGGGGCGGCGGCGGCGGCGG - 5'

ODC1,ID02064.5p-miR,17 nt,
-140 kJ/mole, 97%, 23 nt
5' - GCCGCCGCCGCCGCCGCCCGCCG - 3'
  |||||||||||||||||| |||
3' - CGGCGGCGGCGGCGGCGGGAGGC - 5'
	ODC1,ID02084.3p-miR,17 nt,
	-140 kJ/mole, 93%, 24 nt
	5' - GCCGCCGCCGCCGCCGCCCGCCGC - 3'
	  ||||||||||||||||||||
	3' - GGGCGGCGGCGGCGGCGGGCGGCU - 5'

CDS	

SIRT1, ID03332.3p-miR, 287 nt,
-138 kJ/mole, 93%, 24 nt
5' - GGCGGAGGC**A**GAGGCGGCGGCGGC - 3'
  |||||||||||||||||||| |
3' - CCGCCUCCG**C**CUCCGCCGCCGCGG - 5'
	SMAD7, ID00529.5p-miR, 369 nt,
	-113 kJ/mole, 95%, 21 nt
	5' - GCAGG**G**GGAGGUGGAGGAGGA - 3'
	  ||||||||||||| ||||||
	3' - CGUCC**U**CCUCCACCACCUCCU - 5'

3'UTR	

AGMAT, miR-619-5p, 2207 nt,
-121 kJ/mole, 100%, 22 nt
5' - GGCUCAUGCCUGUAAUCCCAGC - 3'
  ||||||||||||||||||||
3' - CCGAGUACGGACAUUAGGGUCG - 5'
	AGMAT, miR-5096, 2580 nt,
	-110 kJ/mole, 98%, 21 nt
	5' - GCCUG**G**CCAACAUGGUGAAAC - 3'
	  ||||||||||||||||||||
	3' - CGGAC**U**GGUUGUACCACUUUG - 5'

AGMAT, miR-1273g-3p, 2731 nt,
-113 kJ/mole, 96%, 21 nt
5' - C**C**CAGGCUGGAGUGCAGUGG**C** - 3'
  ||||||||||||||||||||
3' - G**A**GUCCGACCUCACGUCACC**A** - 5'
	AGMAT, miR-1273e, 2774 nt,
	-113 kJ/mole, 96%, 22 nt
	5' - UCC**G**C**C**UCCUGGGUUCAAGCAA - 3'
	  ||||||||||||||||||||
	3' - AGG**U**G**A**AGGACCCAAGUUCGUU - 5'

CDC27, miR-548aa, 4974 nt,
-119 kJ/mole, 97%, 25 nt
5' - UGGUGCAUAAGUAAUUGUGGUUUUU - 3'
  ||||||| ||||||||||||
3' - ACCACGUUUUCAUUAACACCAAAAA - 5'
	CDC27, miR-548t-3p, 4974 nt,
	-119 kJ/mole, 97%, 25 nt
	5' - UGGUGCAUAAGUAAUUGUGGUUUUU - 3'
	  ||||||| ||||||||||||
	3' - ACCACGUUUUCAUUAACACCAAAAA - 5'

CDC27, miR-548as-5p, 5013 nt,
-108 kJ/mole, 96%, 22 nt
5' - GGCAAAACC**U**GCAAUUACUUUU - 3'
  ||||||||||||||||||||
3' - CCGUUUUGG**G**CGUUAAUGAAAA - 5
	CDC27, miR-548x-5p, 5015 nt,
	-108 kJ/mole, 96%, 23 nt
	5' - CAAAACCUGCAAUUACUUUUGCA - 3'
	  ||||| ||||||||||||||
	3' - GUUUUUGACGUUAAUGAAAACGU - 5'

IGF1, miR-574-5p, 4042 nt,
-113 kJ/mole, 93%, 23 nt
5' - ACACACACACACACACACACACA - 3'
  |||||| ||||||||||| ||
3' - UGUGUGAGUGUGUGUGUGAGU - 5'
	IGF1, miR-1273g-3p, 6009 nt,
	-113 kJ/mole, 96%, 21 nt
	5' - C**C**CAGGCUGGAGUGCAGUGG**C** - 3'
	  ||||||||||||||||||||
	3' - G**A**GUCCGACCUCACGUCACC**A** - 5'

KRAS, miR-1273g-3p, 3176 nt,
-108 kJ/mole, 93%, 21 nt
5' - C**C**CAGGCUGGA**A**UGCAGUGG**C** - 3'
  ||||||||||||||||||||
3' - G**A**GUCCGACCU**C**ACGUCACC**A** - 5'
	KRAS, miR-1273f, 3209 nt,
	-104 kJ/mole, 100%, 19 nt
	5' - CACUGCAACCUCCAUCUCC - 3'
	  ||||||||||||||||||
	3' - GUGACGUUGGAGGUAGAGG - 5'

GSK3B,ID00436.3p-miR,4682 nt,
-104 kJ/mole, 89%, 23 nt
5'-GUGUGUG**U**GU**G**U**G**U**G**UGUGUGUG-3'
  ||||||||||||||||||||
3'-CACACAC**G**CA**U**A**U**A**U**ACACACAU-5'
	PPP2R1B, miR-4452, 3026 nt,
	-102 kJ/mole, 89%, 23 nt
	5' - AUCA**U**UU**G**AGG**U**CAAGAAUUC**G**A - 3'
	  ||||||||||||||||||||
	3' - UAGU**G**AA**U**UCC**G**GUUCUUAAG**U**U - 5'

PPP2R1B, miR-5096, 3054 nt,
-113 kJ/mole, 100%, 21 nt
5' - GCCUGACCAACAUGGUGAAAC - 3'
  ||||||||||||||||||||
3' - CGGACUGGUUGUACCACUUUG - 5'
	PPP2R1B, miR-5585-3p, 3124 nt,
	-115 kJ/mole, 98%, 22 nt
	5' - ACCUGUAGUCCCAGCUA**C**UCAG - 3'
	  ||||||||||||||||||||
	3' - UGGACAUCAGGGUCGAU**A**AGUC - 5'

STAT3,miR-619-5p,3131 nt,
-119 kJ/mole, 98%, 22 nt
5' - GGCUCA**C**GCCUGUAAUCCCAGC - 3'
  ||||||||||||||||||||
3' - CCGAGU**A**CGGACAUUAGGGUCG - 5'
	STAT3, miR-5585-3p, 3268 nt,
	-110 kJ/mole, 95%, 22 nt
	5' - **G**CCUGUAGUCCCAGCUA**C**UC**G**G - 3'
	  ||||||||||||||||||||
	3' - **U**GGACAUCAGGGUCGAU**A**AG**U**C - 5'

Note: In bold type highlighted the non-canonical A-C and G-U nucleotide pairs	

Figure 22 - The schemes of miRNAs interaction in 5'UTR, CDS and 3'UTR mRNAs of candidate genes [290].

The schemes illustrate nucleotide interactions spanning the entire length of miRNA with binding sites on mRNA. Incorporating non-canonical pairs of nucleotides such as A-C and G-U enhances the free energy of interaction between miRNA and mRNA through additional hydrogen bonding. Moreover, the stacking interaction maintained in the double-stranded helix structure contributes to a more stable interaction of the miRNA-RISC with mRNA, enabling longer interactions.

Out of the 50 identified binding sites, 45 were localized in the 3'UTR, four in the CDS, and one in the 5'UTR. While clusters with only 2-3 binding sites were observed, other genes exhibited clusters consisting of more than ten miRNA binding sites within their mRNA.

The effectiveness of the MirTarget program is evidenced by the interaction patterns between miRNA and mRNA nucleotides. Figure 23 demonstrates the interaction patterns of various miRNAs with mRNA nucleotides from the five examined genes. In the interaction scheme of miR-619-5p with the *CYP2W1* gene mRNA, all nucleotides bind complementarily, reflecting absolute affinity between the miRNA and mRNA. For other miRNAs, interactions with mRNA often involve the formation of non-canonical A-C and G-U pairs. While these interactions are weaker than canonical pairs, the stacking interaction between nucleotides in each antiparallel RNA strand is maintained.

CYP2W1; miR-619-5p; 2176; -121; 100; 22 5'-GGCUCAUGCCUGUAAUCCCAGC-3' 3'-CCGAGUACGGACAUUAGGGUCG-5'	*CYP2W1*; ID00913.5p-miR; 2196; -117; 92; 23 5'-**G**CACUUUGGGAG**G**CCGAGGCAGG-3' 3'-**U**GUGAAACCCUC**U**CGCUCCGUCC-5'
KIAA1456; ID00367.5p-miR; 8422; -115; 93; 22 5'-UCUGUCACCCAGG**C**UGGAGGGC-3' 3'-AGACAGUGGGUCC**A**ACCUCCCC-5'	*KIAA1456*; ID01334.3p-miR; 4132; -115; 92; 22 5'-AGG**C**GUGAGCCACCGCGCCCGG-3' 3'-UCC**A**CACUCGGUGGCGCGUUCC-5'
SLC26A2; ID01237.3p-miR; 5308; -113; 88; 24 5'-**C**CUGGGUGACAGAGC**G**AGACUCC**G**-3' 3'-**A**GACCAACUGUCUCG**U**UCUGUGG**U**-5'	*SLC26A2*; ID01838.5p-miR; 4439; -113; 90; 24 5'-**G**A**GA**G**G**GUCUCACUGUGUUGC**C**CA-3' 3'-**UU**U**C**C**U**CAGAGUGACACAACG**A**GU-5'
SPATA13; ID00913.5p-miR; 5040; -113; 88; 23 5'-**G**CACUUUGGGAG**G**CC**A**AGGCAGG-3' 3'-**U**GUGAAACCCUC**U**CG**C**UCCGUCC-5'	*UQCRB*; ID00913.5p-miR; 1289; -115; 90; 23 5'-**G**CACUUUGGGAG**G**CCGAGGC**G**GG-3' 3'-**U**GUGAAACCCUC**U**CGCUCCG**U**CC-5'
KIAA1456; miR-1273a; 3875; -121; 92; 25 5'-**G**AGACAGAGUCU**C**GCU**C**UGUCGCCC-3' 3'-**U**UCUUUCUCAGA**A**CGA**A**ACAGCGGG-5'	*KIAA1456*; miR-1273c; 3877; -113; 93; 22 5'-GACAG**A**GUCUCG**C**U**C**UGUCGCC-3' 3'-CUGUC**C**CAGAGC**AA**ACAGCGG-5'
KIAA1456; miR-5096; 5137; -113; 100; 21 5'-GCCUGACCAACAUGGUGAAAC-3' 3'-CGGACUGGUUGUACCACUUUG-5'	*SLC26A2*; miR-5585-3p; 5209; -113; 96; 22 5'-**G**CCUGUAGUCCCAGCUA**C**UCAG-3' 3'-**U**GGACAUCAGGGUCGAU**A**AGUC-5'
UQCRB; miR-5096; 1343; -110; 98; 21 5'-GCCUG**G**CCAACAUGGUGAAAC-3' 3'-CGGAC**U**GGUUGUACCACUUUG-5'	*UQCRB*; miR-5585-3p; 1412; -108; 93; 22 5'-ACCUGUA**A**UCCCAGCUA**C**UC**G**G-3' 3'-UGGACAU**C**AGGGUCGAU**A**AG**U**C-5'
Note: Gene; miRNA; start of binding site (nt); the free energy, ΔG (kJ/mole); the ΔG/ΔGm (%); length of miRNA (nt). In bold type highlighted the non-canonical A-C and G-U pairs of nucleotide	

Figure 23 - The schemes of miRNAs interaction in 3'UTR mRNAs of candidate genes [290].

The impact of miRNA on gene expression suppression is influenced by the relative concentrations of miRNA and mRNA. Strong binding between miRNAs and mRNAs might not effectively suppress gene expression if the miRNA concentration is significantly lower than that of the mRNA. Conversely, when there's an excess of miRNA compared to mRNA, average miRNA-mRNA interactions can lead to substantial inhibition of translation. Moreover, given the competition among miRNAs for binding to mRNAs, selecting an effective miRNA for a specific gene becomes more complex, especially when managing the expression of several or numerous miRNAs and genes.

The approaches we've proposed enable the selection from thousands of potential miRNAs likely to interact with candidate genes, significantly reducing material costs and time spent searching for miRNA-gene associations. Our studies have identified miRNAs capable of effectively binding to specific candidate gene mRNAs based on their interaction characteristics.

Validating the stability of miRNA-gene associations can be achieved through identification in orthologous genes. For instance, we observed that in orthologous genes of *CYP2W1*, the bond between ID03126.5p-miR and the orthologous genes in various primates remains conservative. This conservation is reflected in the identity of the LGLLGLWG oligopeptide in the orthologous proteins of primates, as depicted in Figure 24.

```
MALLLLLFLGLLGLWGLLCACAQD Hsa
MALLLLLFLGLLGLWGLLRACARD Ptr
MALLLLLLLGLLGLWGLLRACARD Pab
MALLLLLLLGLLGLWGLLRACTRD Nle
MALLLLLLLGLLGLWGLLRAYARD Mml
MALLLLLLLGLLGLWGLLRAYARD Csa
```

Figure 24 - Regions of orthologous *CYP2W1* proteins containing the LGLLGLWG oligopeptide encoded by ID03126.5p-miR binding sites [290].

The oligopeptides situated before the binding site in the *Hsa* and *Ptr* sequences exhibit dissimilarities compared to other primate species. Conversely, the oligopeptides located after the binding site display higher variability among the studied subjects.

Notably, the binding sites of miR-619-5p in the mRNA of the five examined genes are highly conservative. However, the surrounding nucleotide sequences exhibit variability, as illustrated in Figure 25.

The region preceding the conserved binding site of mir-619-5p exhibits significant variability in oligonucleotides. Additionally, after a conserved oligonucleotide at the 3’-end, there is also high nucleotide variability [296].

CTGACCCGGTGCGGT**GGCTCATGCCTGTAATCCCAGC**ACTTTGGGAGGCCGAGGCAGGCG *CYP2W1*
TAGGCTGAGCATGGT**GGCTCATGCCTGTAATCCCAGC**ACTTTGGGAGGCCAAGGCAGAAG *KIAA1456*
TCGGCCAGGTGCAGT**GGCTCATGCCTGTAATCCCAGC**ACGTTGGGAGGCCGAGGCGGGTG *SLC26A2*
AAGGCTGGGTGCTGT**GGCTCATGCCTGTAATCCCAGC**ACTTTGGGAGGCCAAGGCAGGTG *SPATA13*
TTGGCAAGGCATGGT**GGCTCATGCCTGTAATCCCAGC**ACTTTGGGAGGCCGAGGCGGGTA *UQCRB*

Figure 25 - The regions of mRNA nucleotide sequences of orthologous genes containing miR-619-5p binding sites [290, 296]

From the obtained results, it is deduced that 115 candidate genes were targeted by 448 miRNAs, with binding sites detected in the 5'UTR, CDS, and 3'UTR of the respective gene mRNAs. Most miRNA binding sites spanned the mRNA length without overlapping nucleotide sequences, although clusters with overlapping sequences were identified in certain cases. The average free energy of miRNA-mRNA interactions appeared higher in the 5'UTR and CDS, indicating a preference for miRNA binding in these regions among the studied genes.

Notably, a significant majority of miRNA binding sites, including those of miR-619-5p, were located in the 3'UTR [296]. Specifically, in the mRNAs of target genes like *CYP2W1, KIAA1456, SLC26A2, SPATA13,* and *UQCRB*, multiple binding sites were observed for seven, twenty, nine, five, and four miRNAs respectively. These genes showed clusters of miRNA binding sites, albeit with only two or three binding sites in some cases. Furthermore, established miRNA binding sites were conserved in the mRNA of orthologous primate genes. Diagrams depicting miRNA-mRNA nucleotide interactions revealed the significance of non-canonical A-C and G-U pairs, previously overlooked by other researchers, in elevating the free energy of miRNA-mRNA interaction.

The identified associations between miRNA and *CYP2W1, KIAA1456, SLC26A2, SPATA13,* and *UQCRB* target genes suggest their potential as markers in developing diagnostic methods for gastrointestinal tract cancer.

Based on quantitative interaction characteristics between miRNAs and mRNAs, associations with high free energy were identified between miRNAs and candidate genes, recommended as markers for colorectal cancer diagnosis. Notably, there were 29 associations in the 5'UTR with an interaction free energy higher than -130 kJ/mol. Instances of multiple binding sites for ID03332.3p-miR in the mRNA CDS of *SIRT1* and *TGFBR1* genes highlight the importance of regulating this miRNA's expression. The average free energy of miRNA interaction with candidate gene mRNA 3'UTRs was -109 kJ/mol, with four associations having an interaction free energy greater than -125 kJ/mol. Notably, associations involving miR-6089, ID03208.5p-miR, ID00296.3p-miR with *TGFB1*, and ID03332.3p-miR with *TGFBR1* showcased higher free energies of interaction, potentially serving as early diagnostic markers for colorectal cancer.

Conclusion

The focus of this dissertation is on identifying crucial genes and miRNA molecules involved in gastrointestinal (GI) tract cancer development. By establishing associations between miRNA nanomolecules and mRNAs of candidate genes related to GI tract cancer, the aim is to devise methods for early pathology diagnosis. Consequently, this study presents potential associations of select miRNA nanomolecules with mRNA target genes, recommended as markers for GI tract cancer.

To achieve this, two databases were created: one cataloging genes exhibiting mutations in GI tract cancer and another compiling miRNA nanomolecules, noting expression alterations in esophageal, stomach, small intestinal, and colorectal cancers. From this data, 372 genes, crucial in function and extensively studied, were selected. Using the MirTarget program, which regulates these genes, a search for miRNA targets was conducted, employing a technique based on the score parameter ($\Delta G/\Delta Gm$) that we introduced. Consequently, 6274 miRNAs were identified as potential key regulators of these 372 target genes. These discovered miRNA binding sites display substantial complementarity, mostly referring to non-canonical sites with G:U pairs and unpaired nucleotides at the 5' or 3’ ends of the miRNA, often displaying high regulatory efficiency.

Predicting miRNA nanomolecule binding sites computationally is a crucial step in understanding the biological function of miRNAs. These findings aid in understanding the regulation of gene features at post-transcriptional levels through miRNAs, particularly those involved in various cellular processes related to GI tract cancer.

The majority of the 372 studied genes encode transcription factors implicated in GI tract cancer manifestation. mRNA of these genes, highly associated with miRNAs, potentially undergo concentration changes, affecting GI tract cancer development. However, the actual impact of predicted miRNA on the expression of investigated genes remains poorly studied. Nevertheless, the obtained data indicate a high probability of interaction between miRNA and mRNA of these genes, presenting promising avenues for further research.

The observed clustered organization of binding sites in the 5'UTR, CDS, and 3'UTR suggests an ordered arrangement of binding sites among several miRNA groups within mRNA fragments, albeit partially combined.

The characteristics of the interaction of miRNA with mRNA genes are specific for each localization of GI tract cancer, and can be used in the selection of associations for early diagnosis. Each group of candidate genes contains associations of miRNAs and their target genes, which have an increased degree of interaction and can serve as markers for the development of early diagnostic methods.

The human miRNA nanomolecules database and the database on 372 candidate genes associated with the development of cancer of the esophagus, stomach, small

intestine, and large intestine, created in the format for the MirTarget program, can be used to develop test systems for early diagnosis of these diseases.

The human genome contains candidate gastrointestinal cancer genes, having miRNA binding sites with high nucleotide complementarity in the 5'UTR, CDS and 3'UTR of their mRNA. The miRNA binding sites in the CDS of candidate genes have high phylogenetic conservatism.

Most candidate genes are targeted by one or more miRNAs, the binding sites of which are separate from each other in the mRNA. In the mRNAs of other candidate genes, miRNA binding sites are located overlapping nucleotide sequences, forming clusters of binding sites for two or more miRNAs. The clusters of miRNA binding sites are characterized by their compact arrangement and competition between miRNAs, since their binding sites are located with overlapping nucleotides.

The quantitative characteristics of the interaction of miRNA with mRNA of candidate target genes indicate a potentially high dependence of the expression of these genes on the concentration of miRNA and mRNA. They serve as the basis for the selection of miRNA and mRNA associations for the diagnosis of gastrointestinal tract cancer.

Associations of miRNA with target genes for the following localizations of gastrointestinal tract cancer have been proposed.

For esophageal adenocarcinoma - ID00792.3p-miR and *ALDH1A2*; ID01778.3p-miR, ID02084.3p-miR and *ARID1A;* ID02052.5p-miR, ID02064.5p-miR and *ARID1A, BARX1;* ID00296.3p-miR, ID02294.5p-miR, ID01702.3p-miR and *ARID1A, IGFBP2;* ID03332.3p-miR and *AR, FOXF1, FOXP1, GDF7;* ID00961.3p-miR and *BARX1*. For esophageal squamous cell carcinoma - ID01041.5p-miR and *ELN, SIX4, EGR1;* ID02761.3p-miR and *COL27A1;* ID01804.3p-miR and *CCNE1, EGR1, ELN;* ID00049.5p-miR and *C2orf40;* ID00966.5p-miR and *EGR1;* ID00602.3p-miR and *BMP7;* ID02668.5p-miR and *LAMC2;* ID01641.3p-miR and *ADAMTS2*.

For gastric cancer - miR-3120-3p and *DNM3OS;* miR-3661 and *PPP2CA;* miR-4253 and *EPHB2;* miR-4273 and *ZNF717;* miR-1273 family and *ATM, CASP10, CDH1, CDH17, HPSE, IGF1, KRAS, MACC1, MDM2, PRKAA1, TP53;* miR-619-5p and *CASP10, CEACAM5, CLDN18, CREB1, ERBB3, F2R, HPSE, IL10, IRF1, LEP, MACC1, MTHFR, PDCD4, STAT3;* miR-1183 and *THBS1;* miR-4767 and *SIRT1;* ID02051.3p-miR and *E2F1;* ID01190.5p-miR and *PIK3CA;* ID01816.3p-miR and *TERT;* ID00798.3p-miR and *MAPK1*; ID00920.5p-miR and *OGG1;* ID03491.3p-miR and *TWIST1;* ID01338.5p-miR and *TERT;* ID01662.3p-miR and *MYC;* ID01546.5p-miR, ID01293.5p-miR and *MAPK1;* ID01458.5p-miR and *VEGFC*.

For small intestinal cancer - ID02093.5p-miR and *GNAS;* miR-1273f and *KRAS;* ID00061.3p-miR, ID02260.5p-miR, ID03064.3p-miR, ID00296.3p-miR, ID01641.3p-miR and *ASXL1;* ID02457.3p-miR and *EGFR;* ID02052.5p-miR, ID02294.5p-miR, ID01041.5p-miR, ID01804.3p-miR ID01702.3p-miR, ID02064.5p-miR, ID02084.3p-miR and *LRP1B;* ID01190.5p-miR, ID01702.3p-miR, ID01895.5p-miR, ID00966.5p-miR, ID01804.3p-miR and *PIK3CA*.

For colorectal cancer - miR-548 family and *CDC27, ONECUT2;* miR-619-5p and *AGMAT, ERBB3;* miR-5096 and *PPP2R1B;* miR-574-5p, ID00470.5p-miR and *CCND1, IGF1, PPP2R1B, SMAD4, XRCC1, ZEB1;* miR-466, ID00436.3p-miR, ID01030.3p-miR and *FLT1, GSK3B, IRS1, MACC1;* miR-6089 and *TGFB1;* ID00296.3p-miR and *FMNL3, GSK3B, TGFB1;* ID03332.3p-miR and *MTA1, SIRT1, TCF4, TGFBR1;* miR-4665-3p and *STAT3;* miR-6789-5p and *HIF1A;* ID01453.3p-miR, ID01099.5p-miR, ID01868.3p-miR, and *ONECUT2;* ID01003.3p-miR and *MTA1;* ID01492.3p-miR and *SMAD2,* ID00156.5p-miR and *MSH6;* ID03064.3p-miR and *SMAD7.*

Key research findings and conclusions:

1. In the format for the MirTarget program, a database of genes associated with the development of gastrointestinal tract cancer was created, consisting of genes: esophageal adenocarcinoma - 38 genes, esophageal squamous cell carcinoma - 66 genes, gastric cancer - 215 genes, small intestinal cancer - 40 genes, colorectal cancer - 135 genes. A database of 6274 nucleotide miRNA sequences was created in the format for the MirTarget program.

2. In genes associated with the development of esophageal cancer, 279 binding sites have been found for 184 miRNAs, which can regulate 32 out of 38 candidate esophageal adenocarcinoma genes. The 246 miRNAs form 321 mRNA binding sites in 53 of the 66 candidate genes for esophageal squamous cell carcinoma. The associations of miRNA and mRNA, having a free energy of interaction equal to -130 kJ/mol and more, have been identified for their recommendation as markers for the early diagnosis of esophageal cancer.

3. Based on the study of the characteristics of the interaction of 6274 miRNAs with mRNAs of 215 candidate gastric cancer genes, associations with an average free energy of interaction equal to -130 kJ/mol and higher in 5'UTR, CDS, 3'UTR were identified. 153 out of 215 mRNA candidate genes linked to gastric cancer exhibit a strong affinity for 309 miRNA molecules at the nanomolecular level. Identification of miRNA binding sites within mRNA clusters of certain genes has been accomplished. Moreover, there exists a notable preservation of miRNA binding sites among the mRNA of orthologous genes such as *ARID1A, E2F1, TBC1D9, MAPK1, TERT*, and *VEGFC*.

4. In genes associated with the development of small intestinal cancer, 211 miRNAs have binding sites that can efficiently regulate the expression of 32 out of 40 candidate genes. The mRNAs of the *ASXL1, GNAS,* and *KRAS* genes have completely complementary miRNA binding sites. The mRNA of the *SMAD4* gene contained multiple binding sites for miR-574-5p and ID00470.5p-miR forming clusters in the same region.

5. Within the mRNA of genes associated with colorectal cancer development, extensive miRNA binding has been observed: 233 binding sites for 165 miRNAs in the 5'UTR, 193 binding sites for 167 miRNAs in the protein-coding region, and 450 binding sites for 151 miRNAs in the 3'UTR have been identified. Notably, the miR-548 family, along with miR-619-5p, miR-6089, ID03208.5p-miR, ID00296.3p-miR,

and ID03332.3p-miR, have been proposed as potential candidates for colorectal cancer diagnosis.

6. Associations of miRNA nanomolecules with target genes for the following localizations of gastrointestinal tract cancer were revealed: for esophageal adenocarcinoma - 18 associations, esophageal squamous cell carcinoma - 12 associations, gastric cancer - 42 associations, small intestinal cancer - 12 associations, colorectal cancer - 35 associations of miRNA and mRNA.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. // CA Cancer J Clin. – 2021. – V. 71, № 3. – 209-249.
2. Syngal S., Brand R.E., Church J.M., Giardiello F.M., Hampel H.L., Burt R.W. ACG clinical guideline: Genetic testing and management of hereditary gastrointestinal cancer syndromes // Am J Gastroenterol. – 2015. – Vol. 110, № 2. – P. 223–62.
3. Zeng H., Zheng R., Zhang S., Zuo T., Xia C., Zou X., Chen W. Esophageal cancer statistics in China, 2011: Estimates based on 177 cancer registries // Thorac Cancer. – 2016. – Vol. 7. – P. 232–237.
4. Zaninotto G., Bennett C., Boeckxstaens G., et al. The 2018 ISDE achalasia guidelines // Dis Esophagus. – 2018. – Vol. 31.
5. Park GW, Kim SK, Lee CH, Kim CR, Jeong HJ, Kim DK. Effect of chronic obstructive pulmonary disease on swallowing function in stroke patients // Ann Rehabil Med. – 2015. – Vol. 39, № 2. – P. 218–25.
6. Gillies CL., Farrukh A., Abrams KR., Mayberry JF. Risk of esophageal cancer in achalasia cardia: A meta–analysis // JGH Open. – 2019. – Vol. 3, №3. – P. 196–200.
7. Belsey R. Functional disease of the esophagus // J Thorac Cardiovasc Surg. – 1966. – Vol. 52. – P. 164–88.
8. Lortat–Jacob JL., Richard CA., Fekete F., et al. Cardiospasm and esophageal carcinoma: report of 24 cases // Surgery. – 1969. –Vol. 66. – P. 969–75.
9. Pierce WS., MacVaugh H. 3rd, Johnson J. Carcinoma of the esophagus arising in patients with achalasia of the cardia // J Thorac Cardiovasc Surg. – 1970. – Vol. 59. – P. 335–9.
10. Hankins JR, McLaughlin JS. The association of carcinoma of the esophagus with achalasia // J Thorac Cardiovasc Surg. – 1975. – Vol. 69. – P. 355–60.
11. 11–1. Boeckxstaens GE., Zaninotto G., Richter JE. Achalasia // Lancet. – 2014. – Vol. 383. – P. 83–93.
12. Lopes AB, Fagundes RB. Esophageal squamous cell carcinoma – precursor lesions and early diagnosis // World J Gastrointest Endosc. – 2012. – Vol. 4. – P. 9–16.
13. Tustumi F, Bernardo WM, da Rocha JRM, et al. Esophageal achalasia: a risk factor for carcinoma. A systematic review and meta–analysis // Dis Esophagus. – 2017. – Vol. 30. – P. 1–8.
14. Leeuwenburgh I, Haringsma J, Van Dekken H, et al. Long–term risk of oesophagitis, Barrett's oesophagus and esophageal cancer in achalasia patients // Scand J Gastroenterol Suppl. – 2006. – P. 7-10.

15. Wroblewski LE, Peek RM Jr, Wilson KT. Helicobacter pylori and gastric cancer: factors that modulate disease risk // Clin Microbiol Rev. – 2010. – Vol. 23, № 4. – P. 713-739.
16. Murata-Kamiya N., K. Kikuchi T. Hayashi H. Higashi, M. Hatakeyama. Helicobacter pylori exploits host membrane phosphatidylserine for delivery, localization, and pathophysiological action of the CagA oncoprotein // Cell Host Microbe. – 2010. – V. 7. – P. 399-411.
17. Bartchewsky W.Jr., Martini MR., Masiero M., Squassoni AC., Alvarez MC., Ladeira MS., Salvatore D., Trevisan M., Pedrazzoli J Jr., Ribeiro ML. Effect of Helicobacter pylori infection on IL-8, IL-1beta and COX-2 expression in patients with chronic gastritis and gastric cancer // Scand J Gastroenterol. – 2009. – Vol. 44, № 2. – P. 153-61.
18. Gonzalez CA, Figueiredo C, Lic CB, Ferreira RM, Pardo ML, Ruiz Liso JM, et al. Helicobacter pylori cagA and vacA genotypes as predictors of progression of gastric preneoplastic lesions: A long-term follow-up in a high-risk area in Spain // Am J Gastroenterol. – 2011. – Vol. 106. – P. 867–74.
19. Chen EY., Vaccaro GM. Small Bowel Adenocarcinoma // Clin Colon Rectal Surg. – 2018. – Vol. 31, № 5. – P. 267-277
20. Weedon DD, Shorter RG, Ilstrup DM, Huizenga KA, Taylor WF. Crohn's disease and cancer // N Engl J Med. –1973. – Vol. 289. – P. 1099–1103.
21. Freeman HJ. Colorectal cancer risk in Crohn's disease // World J Gastroenterol. - 2008. – Vol. 14, № 12. – P. 1810-1811.
22. Freeman HJ. Colorectal cancer complicating Crohn's disease // Can J Gastroenterol. – 2001. – Vol. 15. – 231–236.
23. Jamali L., Tofigh R., Tutunchi S., Panahi G., Borhani F., et al. Circulating microRNAs as diagnostic and therapeutic biomarkers in gastric and esophageal cancers // J Cell Physiol. – 2018. – Vol. 233, №11. – P. 8538–8550.
24. Wang Z., Zhao Z., Yang Y., Luo M., Zhang M., et al. MiR–99b–5p and miR–203a–3p Function as Tumor Suppressors by Targeting IGF–1R in Gastric Cancer // Sci Rep. – 2018. – Vol. 8, №1. – P. 10119.
25. Feng C., Xian Q., Liu S. MicroRNA–518 inhibits gastric cancer cell growth by inducing apoptosis via targeting MDM2 // Biomed Pharmacother. – 2018. – Vol. 97. – P. 1595–1602.
26. Lopez-Lazaro M. What Is The Main Cause Of Cancer? // CancerStudTher J. – 2016. – Vol. 1, № 1 . – P.1-2.
27. Vogelstein B, Kinzler KW (2004) Cancer genes and the pathways they control // Nat Med. – Vol. 10. – P. 789-799.
28. Vaux DL. In defense of the somatic mutation theory of cancer // Bioessays. – 2011. – Vol. 33. – P. 341–343;
29. Vogelstein B., Kinzler KW. The Path to Cancer --Three Strikes and You're Out // N Engl J Med. – 2015. – Vol. 373. – P. 1895–1898

30. Martinac, B. The ion channels to cytoskeleton connection as potential mechanism of mechanosensitivity // Biochim. Biophys Acta. – 2014. – Vol.1838. – P. 682-691.
31. Bruce JIE., James AD. Targeting the Calcium Signalling Machinery in Cancer // Cancers. – 2020. – Vol. 12, № 9 . – P. 2351.
32. Krakhmal NV., Zavyalova MV., Denisov EV., Vtorushin SV., Perelmuter VM. Cancer Invasion: Patterns and Mechanisms // Acta Naturae. – 2015. – Vol. 7, № 2. – P.17-28.
33. Spano D., Heck C., De Antonellis P., Christofori G., Zollo M. Molecular networks that regulate cancer metastasis // Semin Cancer Biol. – 2012. – Vol. 22, № 3. – P. 234-249.
34. Lambert AW., Pattabiraman DR., Weinberg RA. Emerging Biological Principles of Metastasis // Cell. – 2017. – Vol. 168, № 4 . – P. 670-691.
35. Romano M. F., Targeting TGFbeta-mediated processes in cancer // Curr Opin Drug Discov Devel. – 2009. – Vol. 12. – P. 253-263.
36. Zhang Y., Xia M., Jin K. et al. Function of the c-Met receptor tyrosine kinase in carcinogenesis and associated therapeutic opportunities // Mol Cancer. – 2018. – Vol. 17. – P. 45.
37. Winkler J., Abisoye-Ogunniyan A., Metcalf K.J. et al. Concepts of extracellular matrix remodelling in tumour progression and metastasis // Nat Commun. – 2020. – V. 11. – P. 5120.
38. Banin Hirata BK., Oda JM., Losi Guembarovski R, Ariza CB, de Oliveira CE, Watanabe MA. Molecular markers for breast cancer: prediction on tumor behavior // Dis Markers. – 2014. – Vol. 2014. – P. 513158.
39. Sharma S. Tumor markers in clinical practice: General principles and guidelines // Indian J Med Paediatr Oncol. – 2009. – Vol. 30(1) . – P. 1-8.
40. Buonaguro FM., Holdenrieder S., Pagliaro L., Morgenstern D., Dayyani F. Clinically Meaningful Use of Blood Tumor Markers in Oncology // Biomed Research International. – 2016. – Vol. 2016. – P. 1-9.
41. Jelski W, Mroczko B. Biochemical Markers of Colorectal Cancer – Present and Future // Cancer Manag Res. – 2020. – Vol. 12. – P. 4789-4797.
42. Wu N., Sun H., Sun Q., Zhang F., Ma L., Hu Y., Cong X. Circulating microRNAs as diagnostic biomarkers for melanoma: a systematic review and meta-analysis // BMC cancer. – 2023. – Vol. 23, № 1. – P. 414.
43. To KK., Tong CW., Wu M., Cho WC. MicroRNAs in the prognosis and therapy of colorectal cancer: From bench to bedside // World J Gastroenterol. – 2018. – Vol. 24, № 27. – P. 2949-2973.
44. Bottani M., Banfi G., Lombardi G. Circulating miRNAs as Diagnostic and Prognostic Biomarkers in Common Solid Tumors: Focus on Lung, Breast, Prostate Cancers, and Osteosarcoma // J Clin Med. – 2019. – Vol. 8, № 10. – P. 1661.

45. Condrat CE., Thompson DC., Barbu MG. miRNAs as Biomarkers in Disease: Latest Findings Regarding Their Role in Diagnosis and Prognosis // Cells. – 2020. – Vol. 9, № 2. – P. 276.
46. Salas-Vega S., Iliopoulos O., Mossialos E. Assessment of Overall Survival, Quality of Life, and Safety Benefits Associated With New Cancer Medicines // JAMA Oncol. – 2017. – Vol. 3, № 3. – P. 382–390.
47. Yumo Xie, Lishuo Shi, Xiaosheng He, Yanxin Luo, Gastrointestinal cancers in China, the USA, and Europe // Gastroenterology Report. – 2021. – goab010.
48. Bray F., Ferlay J., Soerjomataram I., Siegel R.L., Torre L.A., Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries // CA: a cancer journal for clinicians. – 2018. – Vol. 68, № 6. – P. 394–424.
49. Snider E. J., Freedberg D.E., Abrams J.A. Potential Role of the Microbiome in Barrett's Esophagus and Esophageal Adenocarcinoma // Digestive diseases and sciences. – 2016. – Vol. 61, № 8. – P. 2217–2225.
50. Maghsudlu M., Farashahi Yazd E. Heat-induced inflammation and its role in esophageal cancer // Journal of Digestive Diseases. – 2017. – Vol.18, № 8. – P. 431–444.
51. Wang, D. H., Souza, R. F. Biology of Barrett's esophagus and esophageal adenocarcinoma // Gastrointestinal endoscopy clinics of North America. – 2011. – Vol. 21, № 1. – P. 25–38.
52. Batra R., Malhotra G. K., Singh S., Are, C. Managing Squamous Cell Esophageal Cancer // The Surgical clinics of North America. – 2019. – Vol. 99, № 3. – P. 529–541.
53. Wang QL., Xie SH., Wahlin K., Lagergren J. Global time trends in the incidence of esophageal squamous cell carcinoma // Clin Epidemiol . – 2018. – Vol. 10. – P. 717–728.
54. Schlottmann F., Molena D., Patti MG. Gastroesophageal reflux and Barrett's esophagus: a pathway to esophageal adenocarcinoma // Updates Surg. – 2018. – Vol. 70, № 3. – P. 339-342.
55. Harris C., Croce B., Munkholm-Larsen S. Esophageal cancer // Ann Cardiothorac Surg. – 2017. – Vol. 6, № 2. – P. 190.
56. Mönig S., Chevallay M., Niclauss N., Zilli T., Fang W., Bansal A., Hoeppner J. Early esophageal cancer: the significance of surgery, endoscopy, and chemoradiation // Ann N Y Acad Sci. – 2018 . – Vol. 34, № 1. – P. 115-123
57. Wei Q., Li X., Yu W., Zhao K., Qin G., Chen H., Gu Y., Ding F., Zhu Z., Fu X., Sun M. microRNA-messenger RNA regulatory network of esophageal squamous cell carcinoma and the identification of miR-1 as a biomarker of patient survival. – 2019. – J Cell Biochem. – Vol. 120. – P. 12259-12272;
58. Hemmatzadeh M., Mohammadi H., Karimi M., Musavishenas MH., Baradaran B. Differential role of microRNAs in the pathogenesis and treatment of Esophageal cancer // Biomed Pharmacother. – 2016. – Vol. 82. – P. 509-19.

59. Harada K., Baba Y., Ishimoto T., Shigaki H., Kosumi K., Yoshida N., Watanabe M., Baba H. The role of microRNA in esophageal squamous cell carcinoma // J Gastroenterol. – 2016. – Vol. 51. – P. 520–30.
60. Morgan E., Arnold M., Camargo MC., et al. The current and future incidence and mortality of gastric cancer in 185 countries, 2020-40: A population-based modelling study // EClinicalMedicine. – 2022. – Vol. 47. – P. 101404.
61. Richman D.M., Tirumani SH., Hornick JL., et al. Beyond gastric adenocarcinoma: Multimodality assessment of common and uncommon gastric neoplasms // Abdom Radiol (NY) . – 2017. – Vol. 42, № 1. – P. 124-140.
62. Fischbach W. Gastric mucosal-associated lymphoid tissue lymphoma // Gastroenterol Clin North Am. – 2013. – Vol. 42, № 2. – P. 371-380.
63. Buas MF., Vaughan TL. Epidemiology and risk factors for gastroesophageal junction tumors: understanding the rising incidence of this disease // Semin Radiat Oncol. – 2013. – Vol. 23, № 1. – P. 3-9.
64. Berlth F., Bollschweiler E., Drebber U., Hoelscher AH., Moenig S. Pathohistological classification systems in gastric cancer: diagnostic relevance and prognostic value // World J Gastroenterol. – 2014. – Vol. 20, № 19. – P. 5679-5684.
65. Sun RJ., Tang L., Chen Y., et al. Feasibility of differentiating T3 from T4a gastric cancer in different Lauren classification by determining serosa invasion: Diagnostic performance of high enhanced serosa sign // Chin J Cancer Res. – 2018. – Vol. 30, № 2. – P.263-271.
66. Correa P., Piazuelo MB. Helicobacter pylori Infection and Gastric Adenocarcinoma // US Gastroenterol Hepatol Rev. – 2011. – Vol. 7, № 1. – P. 59-64.
67. Assumpção PP., Barra WF., Ishak G., et al. The diffuse-type gastric cancer epidemiology enigma // BMC Gastroenterol. – 2020. – Vol. 20, № 1. – P. 223.
68. Khatoon J., Rai RP., Prasad KN. Role of Helicobacter pylori in gastric cancer: Updates // World J Gastrointest Oncol. – 2016. – Vol. 8, № 2. – P. 147-158.
69. Jorge MM., William OR., Martín GZ. How can Helicobacter pylori eradication therapies be improved? // Revista Colombiana de Gastroenterologia. – 2018. – Vol. 33, № 4. – P. 437-447.
70. Sheh A., Fox JG. The role of the gastrointestinal microbiome in Helicobacter pylori pathogenesis // Gut Microbes. – 2013. – 4, № 6. – 505-531.
71. Kakiuchi, M. et al. Recurrent gain of function mutations of RHOA in diffuse-type gastric carcinoma. – Nat. Genet. . – 2014. – Vol. 46. – P. 583–587.
72. Wang K., Kan J., Yuen ST., et al. Exome sequencing identifies frequent mutation of ARID1A in molecular subtypes of gastric cancer // Nat. Genet. . – 2011. – Vol. 43. – P. 1219–1223.
73. Wang K., Yuen S., Xu J., et al. Whole-genome sequencing and comprehensive molecular profiling identify new driver mutations in gastric cancer // Nat. Genet. . – 2014. – Vol. 46. – P. 573–582.

74. Zang ZJ., Cutcutache I., Poon SL., et al. Exome sequencing of gastric adenocarcinoma identifies recurrent somatic mutations in cell adhesion and chromatin remodeling genes // Nat. Genet. – 2012. – Vol. 44. – P. 570–574.
75. Deng N., Goh LK., Wang H., et al. A comprehensive survey of genomic alterations in gastric cancer reveals systematic patterns of molecular exclusivity and co occurrence among distinct therapeutic targets. – 2012. – Gut. – Vol. 61. – P. 673–684.
76. Pearson A, Smyth E, Babina IS, et al. High-level clonal FGFR amplification and response to FGFR inhibition in a translational clinical trial // Cancer Discov. – 2016. – Vol. 6. – P. 838–851.
77. Van Cutsem E., Bang YJ., Mansoor W., et al. A randomized, open-label study of the efficacy and safety of AZD4547 monotherapy versus paclitaxel for the treatment of advanced gastric adenocarcinoma with FGFR2 polysomy or gene amplification // Ann Oncol. – 2017. – Vol. 28, № 6. – P. 1316-1324.
78. Volk N., Lacy B. Anatomy and Physiology of the Small Bowel // Gastrointest Endosc Clin N Am. – 2017. – Vol. 27, № 1. – P. 1-13.
79. Vagholkar K., Mathew T. Adenocarcinoma of the small bowel: a surgical dilemma // Saudi J Gastroenterol. – 2009. – Vol. 15, № 4. – P. 264-267.
80. Barsouk A, Rawla P, Barsouk A, Thandra KC. Epidemiology of Cancers of the Small Intestine: Trends, Risk Factors, and Prevention // Med Sci (Basel). – 2019. – Vol. 7, № 3. – P. 46.
81. Shenoy S. Genetic risks and familial associations of small bowel carcinoma // World J Gastrointest Oncol. – 2016. – Vol. 8, № 06. – P. 509–519.
82. Bilimoria K. Y., Bentrem D. J., Wayne J. D., Ko C. Y., Bennett C. L., Talamonti, M. S. Small Bowel Cancer in the United States // Ann. Surg. – 2009. – Vol. 249(1). – P. 63–71.
83. Liu Y., Zhang S., Zhou W., Hu D., Xu H., Ji G. Secondary Bile Acids and Tumorigenesis in Colorectal Cancer // Front Oncol. – 2022. – Vol. 12. – P. 813745.
84. The Bethesda Handbook of Clinical Oncology. 2022. Chapter 11. p.176.
85. Pan SY., Morrison H. Epidemiology of cancer of the small intestine // World J Gastrointest Oncol. – 2011. – Vol. 3(3) . – P. 33-42.
86. Grossmann AH., Layfield LJ., Randall RL. Classification, molecular characterization, and the significance of pten alteration in leiomyosarcoma // Sarcoma. – 2012. – Vol. 2012. – P. 380896.
87. Beham AW., Schaefer IM., Schüler P., Cameron S., Ghadimi BM. Gastrointestinal stromal tumors // Int J Colorectal Dis. – 2012. – Vol. 27. – P. 689-700.
88. Rammohan A., Sathyanesan J., Rajendran K., et al. A gist of gastrointestinal stromal tumors: A review // World J Gastrointest Oncol. – 2013. – Vol. 5, № 6. – P.102-112.

89. Isozaki K., Hirota S. Gain-of-Function Mutations of Receptor Tyrosine Kinases in Gastrointestinal Stromal Tumors // Curr Genomics. – 2006. – Vol. 7, № 8. – P. 469-475.
90. Zhou L., Liao Y., Wu J., et al. Small bowel gastrointestinal stromal tumor: a retrospective study of 32 cases at a single center and review of the literature // Ther Clin Risk Manag. – 2018. – Vol. 14. – P. 1467-1481.
91. Vernuccio F., Taibbi A., Picone D., LA Grutta L., Midiri M., Lagalla R., Lo Re G., Bartolotta T. V. Imaging of Gastrointestinal Stromal Tumors: From Diagnosis to Evaluation of Therapeutic Response // Anticancer research. – 2016. – Vol. 36, № 6. – P. 2639–2648.
92. Bennett C.M., Coleman H.G., Veal P.G., Cantwell M.M., Lau C.C., Murray L.J. Lifestyle factors and small intestine adenocarcinoma risk: A systematic review and meta-analysis // Cancer Epidemiol. – 2015. – Vol. 39, № 3. – P. 265-73.
93. Morgan E., Arnold M., Gini A, et al. Global burden of colorectal cancer in 2020 and 2040: incidence and mortality estimates from GLOBOCAN // Gut. – 2023. – Vol. 72, № 2. – P.338-344.
94. Haggar F.A. Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors/ F.A. Haggar, R.P. Boushey// Clin. Colon Rectal Surg. – 2009.– Vol. 22.– P. 191-197.
95. Birgisson H. Survival endpoints in colorectal cancer and the effect of second primary other cancer on disease free survival/ H. Birgisson, U. Wallin, L. Holmberg, B. Glimelius// BMC Cancer –2011. – Vol.11. – P. 438
96. J. Ferlay, H.R. Shin, F. Bray et al.Cancer Incidence and Mortality Worldwide // Lyon, France: International Agency for Research on Cancer. – 2010. – Vol.15.
97. Choi A.R., Park J.R., Kim R.J. et al. Inhibition of Wnt1 expression reduces the enrichment of cancer stem cells in a mouse model of breast cancer // Biochem. Biophys. Res. Commun. – 2012. –Vol. 425. –P.436-442.
98. Simon K. Colorectal cancer development and advances in screening // Clin Interv Aging. – 2016. – Vol. 11. – P. 967-976.
99. Testa U., Pelosi E., Castelli G. Colorectal cancer: genetic abnormalities, tumor progression, tumor heterogeneity, clonal evolution and tumor-initiating cells // Med Sci (Basel) . – 2018. – Vol. 6, № 2. – P. 31.
100. Nguyen, H. T., Duong, H. The molecular characteristics of colorectal cancer: Implications for diagnosis and therapy (Review) // Oncology Letters. – 2018. . – Vol. 16. – P. 9-18.
101. Steinke V., Engel C., Büttner R., Schackert HK., Schmiegel WH., Propping P. Hereditary nonpolyposis colorectal cancer (HNPCC)/Lynch syndrome // Dtsch Arztebl Int. – 2013. – Vol. 110, № 3. – P. 32-38.
102. Foersch S., Neurath MF. Colitis-associated neoplasia: molecular basis and clinical translation // Cell Mol Life Sci. – 2014. – Vol. 71, № 18. – P. 3523-35.
103. Barzi A. Comparative effectiveness of screening strategies for Lynch syndrome/ A. Barzi, S. Sadeghi, M.W. Kattan, N.J. Meropol // J.Natl.Cancer.Inst. – 2015. – Vol. 107, № 4. – P.1-9.

104. Mangifesta M. Mucosal microbiota of intestinal polyps reveals putative biomarkers of colorectal cancer/ M. Mangifesta, L. Mancabelli, C. Milani, F. Gaiani, N. de'Angelis et al. // Sci. Rep. – 2018. –Vol.8, № 1. – P. 13974.
105. Cyrany J., Bureš J., Rejchrt S., Kopáčová M. Familial adenomatous polyposis: complex patient management // Vnitr Lek. – 2018. – Vol. 64, № 6. –P. 635-641.
106. Matthew D. Familial Colorectal Cancer: Understanding the Alphabet Soup // Alabama Clin. Colon. Rect.al Surg. – 2016. – Vol.29. –P.185-195.
107. Sehej P., Easwaran H. Genetic and epigenetic dependencies in colorectal cancer development // Gastroenterology report. – 2022. – Vol. 10. – goac035.
108. Xicola R.M., Manojlovic Z., Augustus G.J., Kupfer S.S., Emmadi R., Alagiozian-Angelova V., Triche T.Jr., Salhia B., Carpten J., Llor X., Ellis N.A. Lack of APC somatic mutation is associated with earlyonset colorectal cancer in African Americans // Carcinogenesis. – 2018. – Vol. 39, № 11. – P. 1331-1341.
109. Gagné-Sansfaçon J., Coulombe G., Langlois M.J., Langlois A., Paquet M., Carrier J., Feng G.S., Qu C.K., Rivard N. SHP-2 phosphatase contributes to KRAS-driven intestinal oncogenesis but prevents colitis-associated cancer development // Oncotarget. – 2016. – Vol.13, № 3. – P.206-215
110. Pino MS., Chung DC. The chromosomal instability pathway in colon cancer // Gastroenterology. – 2010. – Vol. 138, № 6. – P. 2059-2072.
111. Rapozo DC., Grinmann AB., Carvalho AT., de Souza HS., Soares-Lima SC., de Almeida Simão T., de Paiva D., Abby F., Albano RM., Pinto LF. Analysis of mutations in TP53, APC, K-ras, and DCC genes in the non-dysplastic mucosa of patients with inflammatory bowel disease // Int J Colorectal Dis. – 2009. – Vol. 24, № 10. – P. 1141-8.
112. Grover S., Kastrinos F., Steyerberg EW., et al. Prevalence and phenotypes of APC and MUTYH mutations in patients with multiple colorectal adenomas // JAMA . – 2012. – Vol. 308, № 5. – P. 485-92.
113. Nojadeh J. N., Behrouz Sharif S., Sakhinia E. Microsatellite instability in colorectal cancer // EXCLI journal. – 2018. – Vol. 17. – P. 159–168.
114. Conesa-Zamora P., Solano J. G., García F. G., Turpin M.C., Santos J. T., Moreno D. T., et al. Expression profiling shows differential molecular pathways and provides potential new diagnostic biomarkers for colorectal serrated adenocarcinoma // Int. J. Cancer. – 2013. – Vol.132. – P.297-307.
115. Stefanius K., Ylitalo L., Tuomisto A., Kuivila R., Kantola T., Sirnio P. et al. Frequent mutations of KRAS in addition to BRAF in colorectal serrated adenocarcinoma // Histopathology. – 2011. – Vol. 58. – P.679- 692.
116. Steelman LS., Chappell WH., Abrams SL., et al. Roles of the Raf/MEK/ERK and PI3K/PTEN/Akt/mTOR pathways in controlling growth and sensitivity to therapy-implications for cancer and aging // Aging (Albany NY) . – 2011. – Vol. 3, № 3. – P. 192-222.
117. Braicu C., Buse M., Busuioc C., et al. A Comprehensive Review on MAPK: A Promising Therapeutic Target in Cancer. Cancers (Basel) . – 2019. – Vol. 11, № 10. – P. 1618.

118. Munshi A., Ramesh R. Mitogen-activated protein kinases and their role in radiation response // Genes Cancer. – 2013. – Vol. 4. – P. 401-408.
119. Cargnello M., Roux PP. Activation and function of the MAPKs and their substrates, the MAPK-activated protein kinases // Microbiol Mol Biol Rev. – 2011. – Vol. 75, № 1. – 50-83.
120. Lee S., Rauch J., Kolch W. Targeting MAPK Signaling in Cancer: Mechanisms of Drug Resistance and Sensitivity. – Int J Mol Sci. – 2020. – Vol. 21(3) . – P. 1102.
121. Krasinskas AM. EGFR Signaling in Colorectal Carcinoma // Patholog Res Int. – 2011. – Vol. 2011. – P. 932932.
122. Ben Brahim E.I., Ayari I., Jouini R., Atafi S., Koubaa W., Elloumi H., Chadli A. Expression of epidermal growth factor receptor (EGFR) in colorectal cancer: An immunohistochemical study/ // Arab. J. Gastroenterol. – 2018. – Vol. 19, № 3. – P. 121-124.
123. Tyrrell HEJ., Kerr D. Prognostic markers for colorectal cancer // Oncotarget. – 201. – Vol. 9, № 69 – P. 33060-33061.
124. Saeed O., Lopez-Beltran A., Fisher KW., et al. RAS genes in colorectal carcinoma: pathogenesis, testing guidelines and treatment implications // J Clin Pathol. – 2019. – Vol. 72, № 2. – P. 135-139.
125. Guo YJ., Pan WW., Liu SB., Shen ZF., Xu Y., Hu LL. ERK/MAPK signalling pathway and tumorigenesis // Exp Ther Med. – 2020. – Vol. 19, № 3. – P. 1997-2007.
126. Walaszczyk A., Gabryś D. Molecular markers used in breast cancer diagnosis — current practice and future perspectives. Nowotwory // Journal of Oncology. – 2018. – Vol. 68, № 5. – P. 259-267.
127. Nagpal M., Singh S., Singh P., Chauhan P., Zaidi MA. Tumor markers: A diagnostic tool // Natl J Maxillofac Surg. – 2016. – Vol. 7, № 1. – P. 17-20.
128. Gold P., Freedman SO. Demonstration of tumor-specific antigens in human colonic carcinomata by immunological tolerance and absorption techniques // J Exp Med. – 1965. – Vol. 121. – P. 439-462.
129. Tong G., Xu W., Zhang G., et al. The role of tissue and serum carcinoembryonic antigen in stages I to III of colorectal cancer-A retrospective cohort study // Cancer Med. – 2018. – Vol. 7, № 11. – P. 5327-5338.
130. Ghareib AF., Mohamed RH., Abdel-Fatah AR., Saadawy SF. Association between serum microrna-21 gene expression, carcinoemberyonic antigen and clinicopathological character for colorectal cancer Zagazig university medical journal . – 2020 . – Vol. 26 . – P. 919-927.
131. Sisik A., Kaya M., Bas G., Basak F., Alimoglu O. CEA and CA 19-9 are Still Valuable Markers for the Prognosis of Colorectal and Gastric Cancer Patients // Asian Pac. J. Cancer Prev. – 2013. – Vol. 14. – P. 4289–4294.
132. Lakemeyer L., Sander S., Wittau M., Henne-Bruns D., Kornmann M., Lemke J. Diagnostic and Prognostic Value of CEA and CA19-9 in Colorectal Cancer // Diseases. – 2021. – Vol. 9, № 1. – P. 21.

133. Chen Q., Kong H., Qi X., et al. Carcinoembryonic Antigen: A Potential Biomarker to Evaluate the Severity and Prognosis of COVID-19 // Front Med (Lausanne). – 2020. – Vol. 7. – P. 579543.
134. Huang CS., Chen CY., Huang LK., Wang WS., Yang SH. Prognostic value of postoperative serum carcinoembryonic antigen levels in colorectal cancer patients who smoke // PLoS One. – 2020. – Vol. 15, № 6. – P. e0233687.
135. Bergamaschi S., Morato E., Bazzo M., et al. Tumor markers are elevated in patients with rheumatoid arthritis and do not indicate presence of cancer // Int J Rheum Dis. – 2012. – Vol. 15, № 2. – P. 179-182.
136. Lingen MW., Kalmar JR., Karrison T., Speight PM. Critical evaluation of diagnostic aids for the detection of oral cancer // Oral Oncol. – 2008. – Vol. 44, № 1. – P. 10-22.
137. https://training.seer.cancer.gov/diagnostic/markers.html
138. Koprowski H., Steplewski Z., Mitchell K., Herlyn M., Herlyn D., Fuhrer P. Colorectal carcinoma antigens detected by hybridoma antibodies // Somatic Cell Genet. – 1979. – Vol. 5, № 6. – 957-71.
139. Lee T., Teng TZJ., Shelat VG. Carbohydrate antigen 19-9 - tumor marker: Past, present, and future // World J Gastrointest Surg. – 2020. – Vol. 12, № 12. – P. 468-490.
140. Lee SP., Sung IK., Kim JH., Lee SY., Park HS., Shim CS. Usefulness of Carbohydrate Antigen 19-9 Test in Healthy People and Necessity of Medical Follow-up in Individuals with Elevated Carbohydrate Antigen 19-9 Level // Korean J Fam Med. – 2019. – Vol. 40, № 5. – P. 314-322.
141. Zhang Y, Yang J, Li H, Wu Y, Zhang H, Chen W. Tumor markers CA19-9, CA242 and CEA in the diagnosis of pancreatic cancer: a meta-analysis // Int J Clin Exp Med. – 2015. – Vol. 8. – P. 11683–91.
142. Ning S., Wei W., Li J., et al. Clinical significance and diagnostic capacity of serum TK1, CEA, CA 19-9 and CA 72-4 levels in gastric and colorectal cancer patients // J Cancer. – 2018. – Vol. 9, № 3. – P. 494-501.
143. Yanoma T., Fukuchi M., Sakurai S., Shoji H., Naitoh H., Kuwano H. Granular cell tumor of the esophagus with elevated preoperative serum carbohydrate antigen 19-9: a case report // Int Surg. – 2015. – Vol. 100, № 2. – P. 365-369.
144. Kim NH., Lee MY., Park JH., Park DI., Sohn CI., Choi K., et al. Serum CEA and CA 19-9 levels are associated with the presence and severity of colorectal neoplasia // Yonsei Med J. – 2017. – Vol. 58. – P. 918–24.
145. Lee SP., Sung IK., Kim JH., Lee SY., Park HS., Shim CS. Usefulness of Carbohydrate Antigen 19-9 Test in Healthy People and Necessity of Medical Follow-up in Individuals with Elevated Carbohydrate Antigen 19-9 Level // Korean J Fam Med. – 2019. – Vol. 40, № 5. – P. 314-322.
146. Holdenrieder S., Dharuman Y., Standop J., Trimpop N., Herzog M., Hettwer K., Simon K., Uhlig S., Micallef J. Novel serum nucleosomics biomarkers for the detection of colorectal cancer // Anticancer Res. – 2014. – Vol. 34. – P. 2357–2362.

147. Wang J., Wang X., Yu F., Chen J., Zhao S., Zhang D., Yu Y., Liu X., Tang H., Peng Z. Combined detection of preoperative serum CEA, CA19-9 and CA242 improve prognostic prediction of surgically treated colorectal cancer patients // Int J Clin Exp Pathol. – 2015. – Vol. 8. – P. 14853–14863.
148. Ringel J., Lohr M. The MUC gene family: their role in diagnosis and early detection of pancreatic cancer // Mol. Cancer. – 2003. – Vol. 2. – P. 9.
149. You W., Sheng N., Yan L., et al. The difference in prognosis of stage II and III colorectal cancer based on preoperative serum tumor markers // J Cancer. – 2019. – Vol. 10, № 16. – P. 3757-3766.
150. ATTALLAH AM., AL E. Clinical Value of Serum CEA, CA19-9, CA 242 and AFP in Diagnosis of Gastrointestinal Tract Cancers // Int J Cancer Res. – 2006. – Vol. 2, № 1. – P. 50-56.
151. Hu PJ., Chen MY., Wu MS., et al. Clinical Evaluation of CA72-4 for Screening Gastric Cancer in A Healthy Population: A Multicenter Retrospective Study // Cancers (Basel). – 2019. – Vol. 11, № 5. – P. 733.
152. Mariampillai AI., Cruz JPD., Suh J., Sivapiragasam A., Nevins K., Hindenburg AA. Cancer Antigen 72-4 for the Monitoring of Advanced Tumors of the Gastrointestinal Tract, Lung, Breast and Ovaries // Anticancer Res. – 2017. – Vol. 37, № 7. – P. 3649-3656.
153. Moldogazieva N.T., Ostroverkhova D.S., Kuzmich N.N., Kadochnikov V.V., Terentiev A.A., Porozov Y.B. Elucidating Binding Sites and Affinities of ERα Agonists and Antagonists to Human Alpha-Fetoprotein by In Silico Modeling and Point Mutagenesis // International Journal of Molecular Sciences. – 2020. – Vol. 21, № 3. – P. 893.
154. Sauzay C., Petit A., Bourgeois A.M., Barbare J.C., Chauffert B., Galmiche A., Houessimon A. Alpha-foetoprotein (AFP): A multi-purpose marker in hepatocellular carcinoma // Clin. Chim. Acta. – 2016. – Vol. 463. – P. 39–44.
155. Zhang J., Liang R., Wei J., Ye J., He Q., ChunlingYuan Y.J., Li Y., Liu Z., Kin Y. Identification of candidate biomarkers in malignant ascites from patients with hepatocellular carcinoma by iTRAQ-based quantitative proteomic analysis // Biomed. Res. Int. – 2018. – Vol. 2018. – p. 5484976.
156. Mizejewski G.J. Biological role of alpha-fetoprotein in cancer: Prospects for anticancer therapy // Expert Rev. Anticancer Ther. – 2002. – Vol. 2. – P. 709–735.
157. Căinap C., Nagy V., Gherman A., et al. Classic tumor markers in gastric cancer. Current standards and limitations // Clujul Med. – 2015. – Vol. 88, № 2. – P. 111-115.
158. Cidón EU., Bustamante, R. Gastric cancer: tumor markers as predictive factors for preoperative staging // Journal of gastrointestinal cancer. – 2011. – Vol. 42, № 3. – P. 127–130.
159. Lee R.C., Feinbaum R.L., Ambros V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. // Cell. – 1993. – Vol. 75, № 5. – P. 843–854

160. Lagos-Quintana M., Rauhut R., Lendeckel W., et al. Identification of novel genes coding for small expressed RNAs // Science. – 2001. – Vol.294. – P. 853-8.
161. Reid G, Kirschner MB, van Zandwijk N. Circulating microRNAs: Association with disease and potential use as biomarkers // Crit Rev OncolHematol. – 2011. – Vol. 80. – P. 193-208.
162. Wang B. Base Composition Characteristics of Mammalian miRNAs // J Nucleic Acids. – 2013. – Vol. 2013. – P. 951570.
163. Vorozheykin, P.S., Titov, I.I. Erratum to: How Animal miRNAs Structure Influences Their Biogenesis // Russ J Genet. – 2020. – Vol. 56. – P. 1012–1024.
164. Gu LQ., Wanunu M., Wang MX., McReynolds L., Wang Y. Detection of miRNAs with a nanopore single-molecule counter // Expert Rev Mol Diagn. – 2012. – Vol. 6. – P. 573-84.
165. Axtell M.J., Westholm J.O., Lai E.C. Vive la différence: biogenesis and evolution of microRNAs in plants and animals. // Genome Biol. BioMed Central, 2011. Vol. 12, № 4. – P.221.
166. Griffiths-Jones S. et al. miRBase: tools for microRNA genomics. // Nucleic Acids Res. 2008. – Vol. 36. – P. 154-8.
167. Kishikawa T., Otsuka M., Ohno M., Yoshikawa T., Takata A., Koike K. Circulating RNAs as new biomarkers for detecting pancreatic cancer // World J. Gastroenterol. – 2015. – Vol. 21. – P. 8527–8540.
168. Han BW., Wang W., Li C., Weng Z., Zamore PD. Noncoding RNA. piRNA-guided transposon cleavage initiates Zucchinidependent, phased piRNA production // Science. NIH Public Access. – 2015. – Vol. 348, № 6236. – P. 817–821.
169. Hyun S. Small RNA Pathways That Protect the Somatic Genome. // Int. J. Mol. Sci. Multidisciplinary Digital Publishing Institute (MDPI) . – 2017. – Vol. 18, № 5.
170. Herrmann A. et al. CTLA4 aptamer delivers STAT3 siRNA to tumor-associated and malignant T cells. // J. Clin. Invest. American Society for Clinical Investigation. – 2014. – Vol. 124, № 7. – P. 2977–2987.
171. Kang S., Kang D., Bakhtiar U., Islam SM., Je S., Kim JH., Song JJ. Silencing Daxx increases the anti-tumor activity of a TRAIL/shRNA Bcl-xL-expressing oncolytic adenovirus through enhanced viral replication and cellular arrest // Cell Signal. – 2015. – Vol. 27(6) . – P. 1214-1224.
172. Taft R.J. et al. The relationship between transcription initiation RNAs and CCCTC-binding factor (CTCF) localization. // Epigenetics Chromatin. BioMed Central. – 2011. – Vol. 4. – P. 13.
173. Ambros V. et al. A uniform system for microRNA annotation. // RNA. – 2003. – Vol. 9, № 3. – P. 277–279.
174. Saini H.K., Enright A.J., Griffiths-Jones S. Annotation of mammalian primary microRNAs. // BMC genomics. – 2008. – Vol. 9. – P. 564

175. Steri M., Idda ML., Whalen MB., Orrù V. Genetic variants in mRNA untranslated regions. // Wiley Interdiscip Rev RNA. – 2018. – Vol. 9, № 4. – P. e1474.
176. Ono M. et al. Identification of human miRNA precursors that resemble box C/D snoRNAs. // Nucleic Acids Res. Oxford University Press. – 2011. – Vol. 39, № 9. – P. 3879–3891.
177. Velayudha Vimala Kumar K., Srikakulam N., Padbhanabhan P., Pandi G. Deciphering microRNAs and Their Associated Hairpin Precursors in a Non-Model Plant, Abelmoschus esculentus // Noncoding RNA. – 2017. – Vol. 3, № 2. – P. 19.
178. Devor E.J. Primate MicroRNAs miR-220 and miR-492 Lie within Processed Pseudogenes. // J. Hered. Oxford University Press, 2006. – Vol. 97, № 2. – P. 186–190.
179. Xiong P., Schneider R.F., Hulsey C.D. et al. Conservation and novelty in the microRNA genomic landscape of hyperdiverse cichlid fishes // Sci Rep. – 2019. – Vol. 9. – P. 13848.
180. Kozomara A., Griffiths-Jones S.. miRBase: annotating high confidence microRNAs using deep sequencing data. // Nucleic Acids Res. – 2014. – Vol. 42. – P. 68–73.
181. Lewis BP., Shih IH., Jones-Rhoades MW., Bartel DP., Burge CB. Prediction of mammalian microRNA targets. // Cell. – 2003. – Vol. 115, № 7. – P. 787-98.
182. Shin C. et al. Expanding the microRNA targeting code: functional sites with centered pairing. // Mol. Cell. Elsevier Ltd. – 2010. – Vol. 38, № 6. – P. 789–802.
183. Bazzini AA., Johnstone TG., Christiano R., Mackowiak SD., Obermayer B., Fleming ES, et al. Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation // EMBO J. – 2014. – Vol. 33, № 9. – P. 981–93.
184. Johnstone TG., Bazzini AA., Giraldez AJ. Upstream ORFs are prevalent translational repressors in vertebrates. // EMBO J. – 2016. – Vol. 35, № 7. – P. 706–23.
185. Moran Y., Agron M., Praher D., Technau U. The evolutionary origin of plant and animal microRNAs. // Nat Ecol Evol. – 2017. – Vol. 1, № 3. – P. 27.
186. Lee H., Han S., Kwon CS., Lee D. Biogenesis and regulation of the let-7 miRNAs and their functional implications. // Protein Cell. – 2016. – Vol. 7, № 2. – P. 100-113.
187. O'Brien J., Hayder H., Zayed Y., Peng C. Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. // Front Endocrinol (Lausanne) . – 2018. – Vol. 9. – P. 402.
188. Macfarlane LA., Murphy PR. MicroRNA: Biogenesis, Function and Role in Cancer. // Curr Genomics. – 2010. – Vol. 11, № 7. – P. 537-561.
189. Ha M., Kim V. Regulation of microRNA biogenesis. // Nat Rev Mol Cell Biol. – 2014. – Vol. 15. – P. 509–524.

190. Koturbash I., Tolleson WH., Guo L., et al. microRNAs as pharmacogenomic biomarkers for drug efficacy and drug safety assessment. // Biomark Med. – 2015. – Vol. 9, № 11. – 1153-1176.
191. Gu S., Jin L., Zhang F. et al. Biological basis for restriction of microRNA targets to the 39 untranslated region in mammalian mRNAs // NSMB. - 2009. - № 16. - P. 144-150.
192. Filipowicz W., Bhattacharyya S.N., Sonenberg N. Mechanisms of post-transcriptional regulation by microRNAs: Are the answers in sight? // Nat. Rev. Genet. - 2008. - Vol. 9. - P. 102-114.
193. Chendrimada TP., Finn KJ., Ji X. et al. MicroRNA silencing through RISC recruitment of eIF6 // Nature. – 2007. – V. 447. – P. 823-828.
194. Hendrickson D.G., Hogan D.J., McCullough H.L. et al. Concordant regulation of translation and mRNA abundance for hundreds of targets of a human microRNA // PLoS Biol. – 2009. – Vol. 7. – P. 1000238.
195. Luo Y., Na Z., Slavoff SA. P-Bodies: Composition, Properties, and Functions // Biochemistry. – 2018. – Vol. 57, № 17. – P.2424–2431.
196. Eulalio A., Behm-Ansmant I., Izaurralde E. P bodies: at the crossroads of post-transcriptional pathways // Nat Rev Mol Cell Biol. – 2007. – V. 8, № 1. – P. 9-22.
197. Eulalio A., Behm-Ansmant I., Schweizer D. et al. P-body formation is a consequence, not the cause, of RNA-mediated gene silencing // Mol. Cell Biol. - 2007. – V. 27. – P. 3970-3981.
198. Condrat CE., Thompson DC., Barbu MG, et al. miRNAs as Biomarkers in Disease: Latest Findings Regarding Their Role in Diagnosis and Prognosis. // Cells. – 2020. – V.9, № 2. – P. 276.
199. McDonald JS., Milosevic D., Reddi HV., Grebe SK., Algeciras-Schimnich A. Analysis of circulating microRNA: preanalytical and analytical challenges. // Clin Chem. – 2011. – V. 57. – P. 833–840.
200. Gidlöf O., Andersson P., van der Pals J., Götberg M., Erlinge D. Cardiospecific microRNA plasma levels correlate with troponin and cardiac function in patients with ST elevation myocardial infarction, are selectively dependent on renal elimination, and can be detected in urine samples. // Cardiology. – 2011. – V. 118. – P. 217–226.
201. Hanke M., Hoefig K., Merz H., Feller AC., Kausch I., Jocham D, et al. A robust methodology to study urine microRNA as tumor marker: microRNA-126 and microRNA-182 are related to urinary bladder cancer. // Urol Oncol. – 2010. – V. 28. – P. 655–661.
202. 212. Weber JA., Baxter DH., Zhang S., Huang DY., Huang KH., Lee MJ., et al. The microRNA spectrum in 12 body fluids. // Clin Chem. – 2010. – V. 56. – P. 1733–1741.
203. Wang J., Zhang KY., Liu SM., Sen S. Tumor-associated circulating microRNAs as biomarkers of cancer. // Molecules. – 2014. – V. 19. – P. 1912-1938.

204. To KK., Tong CW., Wu M., Cho WC. MicroRNAs in the prognosis and therapy of colorectal cancer: From bench to bedside. // World J Gastroenterol. . – 2018. – V. 24. – P. 2949-2973.
205. Bottani M., Banfi G., Lombardi G. Circulating miRNAs as Diagnostic and Prognostic Biomarkers in Common Solid Tumors: Focus on Lung, Breast, Prostate Cancers, and Osteosarcoma. // J Clin Med. – 2019. – V. 8. – P. 1661.
206. Calin G. A., Sevignani C., Dumitru C.D., Hyslop T., Noch E., Yendamuri S., Shimizu M., Rattan S., Bullrich F., Negrini M., Croce C.M. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. // Proceedings of the National Academy of Sciences of the United States of America. – 2014. – Vol. 101, № 9. – P. 2999–3004.
207. Zhang L., Huang J., Yang N., Greshock J., Megraw M. S., Giannakakis A., Liang S., Naylor TL., Barchetti A., Ward MR., Yao G., Medina A., O'brien-Jenkins A., Katsaros D., Hatzigeorgiou A., Gimotty PA., Weber B.L., Coukos G. microRNAs exhibit high frequency genomic alterations in human cancer. – 2006. – Proceedings of the National Academy of Sciences of the United States of America. – Vol. 103(24). – P. 9136–9141.
208. Czubak K., Lewandowska MA., Klonowska K., et al. High copy number variation of cancer-related microRNA genes and frequent amplification of DICER1 and DROSHA in lung cancer. Oncotarget. – 2015. – Vol. 6, № 27. – P. 23399-23416.
209. Pekarsky Y, Croce CM. Role of miR-15/16 in CLL // Cell Death Differ. – 2015. – V. 22, № 1. – P. 6-11.
210. He L., Thomson JM., Hemann MT., Hernando-Monge E., Mu D., Goodson S., Powers S., Cordon-Cardo C., Lowe SW., Hannon GJ., Hammond SM. A microRNA polycistron as a potential human oncogene // Nature. – 2005. – V. 435, № 7043. – P.828-33.
211. Lu Y., Thomson JM., Wong HY., Hammond SM., Hogan BL. Transgenic over-expression of the microRNA miR-17-92 cluster promotes proliferation and inhibits differentiation of lung epithelial progenitor cells. // Dev. Biol. – 2007. – V. 310, № 2. – P. 442–453.
212. Ventura A., Young AG., Winslow MM., Lintault L., Meissner A., Erkeland SJ., Newman J., Bronson RT., Crowley D., Stone JR., Jaenisch R., Sharp PA., Jacks T. Targeted Deletion Reveals Essential and Overlapping Functions of the miR17w92 Family of miRNA Clusters // Cell. – 2008. V.132. – P. 875–886
213. Fuziwara CS., Kimura ET. Insights into Regulation of the miR-17-92 Cluster of miRNAs in Cancer. Front Med (Lausanne). – 2015. – V. 2. – P. 64.
214. van Haaften G., Agami R. Tumorigenicity of the miR-17-92 cluster distilled // Genes Dev. – 2010. – V. 24, № 1. – P. 1-4.
215. Dal Bo M., Bomben R., Hernández L., Gattei V. The MYC/miR-17-92 axis in lymphoproliferative disorders: A common pathway with therapeutic potential // Oncotarget. – 2015. – V. 6, № 23. – P. 19381-19392.

216. Lu Y., Thomson JM., Wong HY., Hammond SM., Hogan BL. Transgenic over-expression of the microrna mir-17-92 cluster promotes proliferation and inhibits differentiation of lung epithelial progenitor cells // Dev Biol. – 2007. – V. 310. – P. 442–453.
217. Farazi TA., Horlings HM., Ten Hoeve JJ., Mihailovic A., Halfwerk H., Morozov P., et al. Microrna sequence and expression analysis in breast tumors by deep sequencing // Cancer Res. – 2011. – V. 71. – P. 4443–4453.
218. Tsuchida A., Ohno S., Wu W., Borjigin N., Fujita K., Aoki T., Ueda S., Takanashi M., Kuroda M. miR-92 is a key oncogenic component of the miR-17-92 cluster in colon cancer // Cancer Sci. Wiley. – 2011. – V. 102, № 12. – P. 2264–2271.
219. Olive V., Bennett MJ., Walker JC., Ma C., Jiang I., Cordon-Cardo C., Li QJ., Lowe SW., Hannon GJ., He L. miR-19 is a key oncogenic component of mir-17-92.miR-19 is a key oncogenic component of mir-17-92 // Genes Dev. – 2009. – V. 23, № 24. – P. 2839–2849.
220. Mavrakis KJ., Wolfe AL., Oricchio E., Palomero T., de Keersmaecker K., McJunkin K., Zuber J., James T., Khan AA., Leslie CS., Parker JS., Paddison PJ., Tam W., Ferrando A., Wendel HG. Genome-wide RNA-mediated interference screen identifies miR-19 targets in Notch-induced T-cell acute lymphoblastic leukaemia // Nat. Cell Biol. – 2010. – V. 12, № 4. – P. 372–379.
221. Gantier MP., Stunden HJ., McCoy CE., Behlke MA., Wang D., Kaparakis-Liaskos M., Sarvestani ST., Yang YH., Xu D., Corr SC., Morand EF., Williams BR. A miR-19 regulon that controls NF-κB signaling. // Nucleic Acids Res. – 2012. – Vol. 40, № 16. – P. 8048–8058.
222. Xiang J., Wu J. Feud or Friend? The Role of the miR-17-92 Cluster in Tumorigenesis // Curr. Genomics. – 2010. – Vol. 11, № 2. – P. 129–135.
223. Liu CZ, Liu W, Zheng Y, et al. PTEN and PDCD4 are bona fide targets of microRNA-21 in human cholangiocarcinoma. – Chin Med Sci J. – 2012. – V. 27, № 2. – P. 65-72.
224. Wang Y, Li Z, Zhao X, Zuo X, Peng Z. miR-10b promotes invasion by targeting HOXD10 in colorectal cancer. Oncol Lett. – 2016. – V. 12, № 1. – P. 488-494.
225. Costinean S. et al. Pre-B cell proliferation and lymphoblastic leukemia/high-grade lymphoma in E -miR155 transgenic mice // Proc. Natl. Acad. Sci. – 2006. – Vol. 103, № 18. – P. 7024–7029.
226. Mu P. et al. Genetic dissection of the miR-17 92 cluster of microRNAs in Myc-induced B-cell lymphomas // Genes Dev. – 2009. – Vol. 23, № 24. – P. 2806–2811.
227. Garzon R. et al. MicroRNA signatures associated with cytogenetics and prognosis in acute myeloid leukemia // Blood. American Society of Hematology. –2008. – Vol. 111, № 6. – P. 3183–3189.
228. Pidíkova P, Reis R, Herichova I. miRNA Clusters with Down-Regulated Expression in Human Colorectal Cancer and Their Regulation. // Int J Mol Sci. –2020 – V. 21, № 13 – P. 4633.

229. Croce C.M. Causes and consequences of microRNA dysregulation in cancer // Nat. Rev. Genet. – 2009. – Vol. 10, № 10. – P. 704–714.
230. Croce C.M. MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review // EMBO Mol. Med. – 2012. – Vol. 4, № 3. –P. 143–159.
231. Liu H, Liu Y, Sun P, et al. Colorectal cancer-derived exosomal miR-106b-3p promotes metastasis by down-regulating DLC-1 expression. // Clin Sci (Lond). – 2020 – V. 134, № 4. – P. 419-434.
232. Yoshii S, Hayashi Y, Iijima H, et al. Exosomal microRNAs derived from colon cancer cells promote tumor progression by suppressing fibroblast TP53 expression. // Cancer Sci. – 2019. – V. 110, № 8. – P. 2396-2407.
233. Wang SE, Lin RJ. MicroRNA and HER2-overexpressing cancer // Microrna. 2013. – V. 2, № 2. – P. 137-147.
234. Gorbatenko A., Sokilde R., Sorensen EE., et al. HER2 and p95HER2 differentially regulate miRNA expression in MCF-7 breast cancer cells and downregulate MYB proteins through miR-221/222 and miR-503. // Sci Rep. – 2019. –V. 9, № 1. – 3352.
235. Hu M., Zhu S., Xiong S., Xue X., Zhou X. MicroRNAs and the PTEN/PI3K/Akt pathway in gastric cancer (Review). // Oncol Rep. – 2019. – V. 41, № 3. –P. 1439-1454.
236. Liu J, Ke F, Chen T, et al. MicroRNAs that regulate PTEN as potential biomarkers in colorectal cancer: a systematic review. // J Cancer Res Clin Oncol. –2020. – V. 146, № 4. – P. 809-820.
237. Zhang Y. et al. Emerging Role of MicroRNAs in mTOR Signaling // Cell. Mol. Life Sci. Springer International Publishing. – 2017. – V.74. – P. 2613-2625.
238. Safa A., Abak A., Shoorei H., Taheri M., Ghafouri-Fard S. MicroRNAs as regulators of ERK/MAPK pathway: A comprehensive review. // Biomed Pharmacother. – V. 132. – P. 110853.
239. Olejniczak M., Kotowska-Zimmer A., Krzyzosiak W. Stress-induced changes in miRNA biogenesis and functioning. // Cell Mol Life Sci. – 2018. – V. 75, № 2. – P.177-191.
240. 250. Kabekkodu SP., Shukla V., Varghese VK., D' Souza J., Chakrabarty S., Satyamoorthy K. Clustered miRNAs and their role in biological functions and diseases // Biol Rev Camb Philos Soc. – 2018. – V.93, № 4. – P.1955-1986.
241. 251. Bavelloni A., Ramazzotti G., Poli A, et al. MiRNA-210: A Current Overview. // Anticancer Res. – 2017. – V. 37, № 12. – P. 6511-6521.
242. 252. Bertero T., Rezzonico R., Pottier N., Mari B. Impact of MicroRNAs in the Cellular Response to Hypoxia // Int Rev Cell Mol Biol. – 2017. – V. 333. – P. 91-158.
243. 253. Berindan-Neagoe I., Calin G.A. Molecular Pathways: microRNAs, Cancer Cells, and Microenvironment // Clin. Cancer Res. – 2014. – Vol. 20, № 24. – P. 6247–6253.

244. Rosenfeld N. et al. MicroRNAs accurately identify cancer tissue origin // Nat. Biotechnol. – 2008. – Vol. 26, № 4. – P. 462–469
245. Cheerla N., Gevaert O. MicroRNA based Pan-Cancer Diagnosis and Treatment Recommendation. // BMC Bioinformatics. BioMed Central. – 2017. – Vol. 18, № 1. – P. 32.
246. Birkman, EM., Mansuri, N., Kurki, S. et al. Gastric cancer: immunohistochemical classification of molecular subtypes and their association with clinicopathological characteristics. // Virchows Arch. – 2018. – V. 472. – P. 369–382.
247. Li C., Zheng H., Jia H., et al. Prognosis of three histological subtypes of colorectal adenocarcinoma: A retrospective analysis of 8005 Chinese patients. // Cancer Med. – 2019. – V. 8, № 7. – P. 3411-3419.
248. Lin S., Gregory RI. MicroRNA biogenesis pathways in cancer. // Nat Rev Cancer. 2015. – V. 15, № 6. – P. 321-333.
249. Iorio MV., Croce CM. Causes and consequences of microRNA dysregulation // Cancer J. – 2012. – V. 18, № 3 . – P. 215-222.
250. Catalanotto C., Cogoni C., Zardo G. MicroRNA in Control of Gene Expression: An Overview of Nuclear Functions. // Int J Mol Sci. – 2016. – V. 17, № 10. – P. 1712.
251. Londin E., Loher P., Telonis AG., et al. Analysis of 13 cell types reveals evidence for the expression of numerous novel primate- and tissue-specific microRNAs. // Proc Natl Acad Sci USA. – 2015. – V. 112, № 10. – P. 1106-1115.
252. Catalanotto C., Cogoni C., Zardo G. MicroRNA in Control of Gene Expression: An Overview of Nuclear Functions // Int J Mol Sci. – 2012. – Vol. 17, №10. – P. 1-17.
253. Peltomki P. DNA mismatch repair and cancer // Mutat. Res. – 2001. – Vol. 488. – P. 77-85.
254. Oliveto S., Mancino M., Manfrini N., Biffo S. Role of microRNAs in translation regulation and cancer. // World J Biol Chem. – 2017. – V. 8, № 1. – P. 45-56.
255. Chouchane L., Boussen H., Sastry K.S. Breast cancer in Arab populations: molecular characteristics and disease management implications // Lancet. Oncol. - 2013. - Vol. 14, № 10. - P. 417-424.
256. Yepes S., López R., Andrade R.E., Rodriguez-Urrego P.A., López-Kleine L., Mercedes Torres M. Co-expressed miRNAs in gastric adenocarcinoma // Genomics.- 2016. - Vol. 2016.- P. 1-9.
257. Redova M., Sana J., Slaby O., Circulating miRNAs as new blood-based biomarkers for solid cancers // Future Oncol. - 2013. - Vol. 9, № 3. - P. 387-402.
258. Witkos T.M., Koscianska E., Krzyzosiak W.J. Practical Aspects of microRNA Target Prediction // Curr Mol Med. - 2011. - Vol. 11, №2. - P. 93-109.
259. Akhtar M.M., Micolucci L., Islam M.S., Olivieri F., Procopio A.D. A Practical Guide to miRNA Target Prediction. Methods // Mol Biol. - 2019. - Vol. 1970. - P. 1-13.

260. Saçar Demirci M.D., Yousef M., Allmer J. Computational Prediction of Functional MicroRNA–mRNA Interactions // Methods in Molecular Biology. - 2019. -Vol. 1912. - P. 175–196.
261. Yurikova, O.Y., Aisina, D.E., Niyazova, R.E. et al. The Interaction of miRNA-5p and miRNA-3p with the mRNAs of Orthologous Genes. // Mol Biol. – 2019. – V. 53. – P. 612–623.
262. Ivashchenko A., Berillo O., Pyrkova A., Niyazova R. (2014) Binding Sites of miR-1273 Family on the mRNA of Target Genes. // Biomed Res Int. – 2014. – V. 2014. – P.11.
263. Ivashchenko A., Berillo O., Pyrkova A., Niyazova R., Atambayeva S. The properties of binding sites of miR-619-5p, miR-5095, miR-5096 and miR- 5585-3p in the mRNAs of human genes. // Biomed Res Int. – 2014. – V. 2014. – P. 1-8.
264. Ivashchenko A.T., Pyrkova A.Y., Niyazova R.Y., Alybayeva A., Baskakov K. Prediction of miRNA binding sites in mRNA. //Bioinformation. – 2016. – V. 12, №. 4. – P. 237-240.
265. Kondybayeva A.M., Akimniyazova A.N., Kamenova S.U., Ivashchenko A.T. The characteristics of miRNA binding sites in mRNA of ZFHX3 gene and its orthologs. // Vavilov journal of genetics and breeding. – 2018. – V.22, № 4. – P.438-444.
266. Wu R.C., Wang T.L., Shih IeM. The emerging roles of arid1a in tumor suppression. // Cancer Biol Ther. – 2014. – V. 15. – P. 655–664.
267. Streppel M.M., Lata S., DelaBastide M., Montgomery E.A., Wang J.S., Canto M.I., Macgregor-Das A.M., Pai S., Morsink FHM, Offerhaus G.J., et al. Next-generation sequencing of endoscopic biopsies identifies ARID1A as a tumor-suppressor gene in Barrett's esophagus. // Oncogene. – 2014. – V. 33. – P. 347–57.
268. Akimniyazova A. N., Ivashchenko A. T. Characteristics of miRNA interaction with mRNA candidate genes of esophageal adenocarcinoma // Experimental Biology. – 2019. – V. 80. – P. 106-116.
269. Mikhail S., Albanese C., Pishvaian MJ. Cyclin-dependent kinase inhibitors and the treatment of gastrointestinal cancers. // Am J Pathol. – 2015. – V. 185. – P. 1185-97.
270. Smith E., Palethorpe H.M., Ruszkiewicz A.R., Edwards S., Leach DA., Underwood T.J., Need E.F., Drew P.A. Androgen Receptor and Androgen-Responsive Gene FKBP5 Are Independent Prognostic Indicators for Esophageal Adenocarcinoma. // Dig Dis Sci. – 2016. – V. 61. – P. 433-43.
271. Du Y., Liu P., Zang W., Wang Y., Chen X., Li M., Zhao G. BTG3 upregulation induces cell apoptosis and suppresses invasion in esophageal adenocarcinoma. // Mol Cell Biochem. – 2015. – V. 404. – P. 31-8.

272. Ou Y.H., Chung P.H., Hsu F.F., Sun T.P., Chang W.Y., Shieh S.Y. The candidate tumor suppressor BTG3 is a transcriptional target of p53 that inhibits E2F1. // EMBO J. – 2007. – V. 26. – P. 3968-80.
273. Ivashchenko A. T., Niyazova R. Ye., Atambayeva Sh. A., Pyrkova A. Yu., Aisina D. E., Yurikova O. Yu., Kondybayeva A., Akimniyazova A., Bayzhigitova D., Bolshoy A. A. miRNA: achievements, misconceptions, perspectives. // NEWS of NAS RK. – 2018. – V. 4. – P. 36-46.
274. Heng L., Jia Z., Sun J., Zhao Y., Zhang K., Zhu Y., Lu S. Integrated Analysis of Competing Endogenous RNAs Network Reveals Potential Signatures in Osteosarcoma Development // Technol Cancer Res Treat. – 2020. – V. 19. – P.396-406.
275. Акимниязова А.Н., Ким А.С., Кдырбаева Ж.С., Габдулхаева Б.Б. Взаимодействие miRNA с mRNA генов, участвующих в развитии плоскоклеточного рака пищевода. // XX межд. научно-практическая конференция "Химия, физика, биология, математика: теоретические и прикладные исследования", Россия, 04.02.2019
276. Akimniyazova A.N., Kim A.S. The interaction of miRNAs with mRNAs of genes in esophageal squamous cell carcinoma. // MCCMB 2019, Moscow, Russia, P. 1-4.
277. Akimniyazova A., Pyrkova A., Uversky V., Ivashchenko A. Predicting Associations of miRNAs and Candidate Gastric Cancer Genes for Nanomedicine. // Nanomaterials. – 2021. – V. 11, № 3. – P. 691.
278. Wang S., Ni B., Zhang Z., Wang C., Wo L., Zhou C., Zhao Q., Zhao E. Long non-coding RNA DNM3OS promotes tumor progression and EMT in gastric cancer by associating with Snail. // Biochemical and Biophysical Research Communications. – 2019. – V. 511, № 1. – P. 57-62.
279. Yin, J., Li, Z., Ye, L., Birkin, E., Li, L., Xu, R., Chen, G., Ji, J., Zhang, Z., Jiang, W. G., Cui, Y. EphB2 represents an independent prognostic marker in patients with gastric cancer and promotes tumour cell aggressiveness. // Journal of Cancer. –2020. – V. 11, № 10. – P. 2778–2787.
280. Wang J., Li Z., Liu B., Chen G., Shao N., Ying X., Wang, Y. Systematic study of cis-antisense miRNAs in animal species reveals miR-3661 to target PPP2CA in human cells. // RNA. – 2016. – V. 22, № 1. – P.87–95.
281. Seshacharyulu P., Pandey P., Datta K., Batra SK. Phosphatase: PP2A structural importance, regulation and its aberrant expression in cancer. // Cancer Lett. – 2013. – V. 335, № 1. – P. 9-18.
282. Akimniyazova A. N., Régnier M., Atambayeva Sh. A., Niyazova R. E., Ivashchenko A. T. Characteristics of MiRNA Interaction With MRNA of Candidate Genes of Small Intestinal Cancer. // International Journal of Biology and Chemistry. – 2018. – V. 11, № 2. – P. 11-20.

283. Akimniyazova A.N. Characteristics of miRNAs interaction with mRNAs of candidate genes in small intestinal cancer. // Biotechnology: state of the art and perspectives. 25 - 27 february 2019. P. 385-386
284. Byun D., Ahmed N., Nasser S., Shin J., Al-Obaidi S., Goel S., Corner G., Wilson A., Flanagan D., Williams D., Augenlicht L., Vincan E, Mariadason J. Intestinal epithelial-specific PTEN inactivation results in tumor formation. // Am J Physiol Gastrointest Liver Physiol. – 2011. – V.301. – P. 856-64.
285. Kantidakis T., Saponaro M., Mitter R., Horswell S., Kranz A., Boeing S., Aygün O., Kelly G.P., Matthews N., Stewart A., Stewart A.F., Svejstrup J.Q. Mutation of cancer driver MLL2 results in transcription stress and genome instability. // Genes Dev. – 2016. – V.30. – P. 408-420.
286. Kumagai R., Kohashi K., Takahashi S., Ya-mamoto H., Hirahashi M., Taguchi K., Nishiyama K., Oda Y. Mucinous phenotype and CD10 expression of primary adenocarcinoma of the small intestine. // World J Gastroenterol. – 2015. – V. 21. – P. 2700-2710.
287. Bläker H., Aulmann S., Helmchen B., Otto H. F., Rieker R. J., Penzel, R. Loss of SMAD4 function in small intestinal adenocarcinomas: comparison of genetic and immunohistochemical findings. // Pathology - Research and Practice. – 2004. – V. 200, № 1. – P. 1–7.
288. Sun FD., Wang PC., Luan RL., Zou SH., Du X. MicroRNA-574 enhances doxorubicin resistance through down-regulating SMAD4 in breast cancer cells. // Eur Rev Med Pharmacol Sci. – 2018. – V. 22, № 5. – P. 1342-1350.
289. Akimniyazova A.N. Computational prediction of miRNA binding sites in mRNA of colorectal cancer candidate genes. Proceedings of BGRS/SB-2020 Bioinformatics of Genome Regulation and Structure/Systems Biology, Novosibirsk, 2020. – P. 402-403.
290. Akimniyazova, A.N., Pyrkova, A.Y., Aisina, D.E., Ivashchenko A.T. miRNA Binding Site Clusters in mRNAs of Colorectal Cancer Candidate Genes. // Nanotechnol Russia. – 2020. – V. 15. – P. 807–818.
291. Kim DH., Park S., Kim H., Choi YJ., Kim SY., Sung KJ., Sung YH., Choi CM., Yun M., Yi YS., Lee CW., Kim SY., Lee JC., Rho JK. Tumor-derived exosomal miR-619-5p promotes tumor angiogenesis and metastasis through the inhibition of RCAN1.4. // Cancer Lett. – 2020. – V.475. – P. 2-13.
292. Chen G., Gu Y., Han P., Li Z., Zhao JL., Gao MZ. Long noncoding RNA SBF2-AS1 promotes colorectal cancer proliferation and invasion by inhibiting miR-619-5p activity and facilitating HDAC3 expression. // J Cell Physiol. – 2019. – V. 234. – P. 18688-18696.
293. Shkurnikov M.Y., Makarova Y.A., Knyazev E.N., Fomicheva K.A., Galatenko A.V., Nyushko K.M., Galatenko V.V., Vechorko V.I., Alekseev B.Y. Plasma Level of hsa-miR-619-5p microRNA Is Associated with Prostatic Cancer Dissemination beyond the Capsule. // Bull Exp Biol Med. – 2017. – V. 163. – P. 475-477.

294. Qiu G., Zhang XB., Zhang SQ., Liu PL., Wu W., Zhang JY., Dai SR. Dysregulation of MALAT1 and miR-619-5p as a prognostic indicator in advanced colorectal carcinoma. // OncolLett. – 2016. – V. 12. – P. 5036-5042.
295. Knyazev E.N., Fomicheva K.A., Mikhailenko D.S., Nyushko K.M., Samatov T.R., Alekseev B.Y., Shkurnikov M.Y. Plasma Levels of hsa-miR-619-5p and hsa-miR-1184 Differ in Prostatic Benign Hyperplasia and Cancer. // Bull Exp Biol Med. – 2016. – V. 161. – P. 108-11.
296. Atambayeva S., Niyazova R., Ivashchenko A., Pyrkova A., Pinsky I., Akimniyazova A., Labeit S. The Binding Sites of miR-619-5p in the mRNAs of Human and Orthologous Genes. // BMC Genomics. – 2017. – V. 18. – P. 428-438.

Supplementary materials A



ҚАЗАҚСТАН РЕСПУБЛИКАСЫ

РЕСПУБЛИКА КАЗАХСТАН

СВИДЕТЕЛЬСТВО

О ВНЕСЕНИИ СВЕДЕНИЙ В ГОСУДАРСТВЕННЫЙ РЕЕСТР ПРАВ НА ОБЪЕКТЫ, ОХРАНЯЕМЫЕ АВТОРСКИМ ПРАВОМ

№ 15600 от «2» марта 2021 года

Фамилия, имя, отчество, (если оно указано в документе, удостоверяющем личность) автора (ов):
ПЫРКОВА АННА ЮРЬЕВНА, АКИМНИЯЗОВА АЙГУЛЬ НУРЛАНКЫЗЫ, ИВАЩЕНКО АНАТОЛИЙ ТИМОФЕЕВИЧ, РАХМЕТУЛЛИНА АЙЖАН КАЗИЕВНА

Вид объекта авторского права: **программа для ЭВМ**

Название объекта: **MirTarSeq (MiRNA target sequence - олигопептиды и олигонуклеотиды)**

Дата создания объекта: **22.01.2021**

Құжат тұлнұсқалығын http://www.kazpatent.kz/ru сайтының "Авторлық құқық" бөлімінде тексеруге болады https://copyright.kazpatent.kz

Подлинность документа возможно проверить на сайте kazpatent.kz в разделе «Авторское право» https://copyright.kazpatent.kz

Подписано ЭЦП

Оспанов Е.К.

Supplementary materials B

The certificate of implementation of the research results in the course «"Omics" technologies in medicine and pharmacy» of the master's educational program 7M101 "Public Health".

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РЕСПУБЛИКИ КАЗАХСТАН
КАЗАХСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ имени АЛЬ-ФАРАБИ

СОГЛАСОВАНО
И.о. проректора по научно-инновационной деятельности
_____ Рамазанов Т.С.
«_____»_____2021 г.

УТВЕРЖДАЮ
Член правления - проректор по академическим вопросам, председатель комиссии
_____ Жакыпова Ф.Н.
«25» 06 2021 г.

АКТ
о внедрении завершенной научно-исследовательской работы (этапа) в учебный процесс

Комиссия Казахского национального университета имени аль-Фараби в составе: председатель: Жакыпова Фатима Надыровна - член правления - проректор по академическим вопросам, члены: директор департамента по академическим вопросам Мухитдинова Тансулу Мирасбековна, директор департамента по науке и инновационной деятельности Мухамбетжанов Серик Копжасарович, руководитель службы методической работы Жакупова Гульназия Толгаевна, декан факультета медицины и здравоохранения Калматаева Жанна Амантаевна, председатель методического совета Высшей школы медицины факультета медицины и здравоохранения Джумашева Рахима Тажибаевна, заведующая кафедрой фундаментальной медицины Сарсенова Лаззат Кадиргалиевна составили настоящий акт о том, что в 2020/2021 учебном году на кафедре фундаментальной медицины внедрены результаты научно-исследовательской работы PhD-докторанта 3-го курса Акимниязовой А.Н. на тему: «Взаимодействие miRNA с mRNA кандидатных генов рака органов желудочно-кишечного тракта» (по госбюджету, №AP05132460).

№ п/п	Форма внедрения (наименование нового курса, спецкурса, раздела лекций, лаб. работы, установки, учебные пособия и т. п.); курс, специальность	Объем внедрения (количество работ, лекционных часов)	Краткое содержание внедренной работы
1	В курс «"Omics" technologies in medicine and pharmacy» (5 кредитов) магистратуры 1-го курса, образовательная программа 7М101 «Health care»	**Лекции –** 3 часа. 1. Bioinformatics (1 час) 2.Personalized medicine (the future of medicine) (1 час) 3. The development of new drugs by using the omics technologies (1 час)	Используются результаты, полученные при изучении характеристик взаимодействия miRNA с mRNA кандидатных генов, участвующих в развитии рака органов желудочно-кишечного тракта.
		Семинары – 6 часов 1. Bioinformatics (2 часа) 2. Personalized medicine (the future of medicine) (2 часа)	Используются методы анализа нуклеотидных последовательностей генов и аминокислотных последовательностей белков, разработанных при

		3. The development of new drugs by using the omics technologies (2 часа)	выполнении темы «Взаимодействие miRNA с mRNA кандидатных генов рака органов желудочно-кишечного тракта», а также приводятся примеры диагностических маркеров, найденных с помощью этих методов.
2	В курс «Mechanisms of Defense and Disease» (10 кредитов) бакалавриата 2-го курса, образовательная программа 6B10103 «General medicine»	**Лекции -** 2 часа 1. Cancer Genetics and Genomics (1 час) 2. Polygenic disorders: developmental malformation (1 час)	Используются результаты, полученные при выяснении молекулярно-генетических основ онкогенеза с использованием геномных технологий и учетом полигенной природы онкогенеза.
		Семинары - 4 часа 1. Cancer Genetics and Genomics (2 часа) 2. Polygenic disorders: developmental malformation (2 часа)	Используются информационные технологии анализа нуклеотидных последовательностей кандидатных генов онкологических заболеваний.

Материалы к настоящему акту рассмотрены на заседании методического совета Высшей школы медицины факультета медицины и здравоохранения (протокол № 09 от 27.05. 2021 г.)

Члены комиссии:

Директор департамента по академическим вопросам — Мухитдинова Т.М.

Директор департамента по науке и инновационной деятельности — Мухамбетжанов С.К.

Руководитель службы методической работы — Жакупова Г.Т.

Декан факультета медицины и здравоохранения — Калматаева Ж.А.

Председатель методического совета Высшей школы медицины факультета медицины и здравоохранения — Джумашева Р.Т.

Заведующая кафедрой фундаментальной медицины — Сарсенова Л.К.

Supplementary materials C

Table C1 - The database of candidate GI tract cancer genes.

Gene	PMID	Type	Gene	PMID	Type	Gene	PMID	Type
1	2	3	1	2	3	1	2	3
ABCB1	24958102	GC	*FBXW7*	25944166	GC	*NOTCH1*	26612211	GC
ABCG2	28633632	ESCC		28617917	SIC		28984154	CRC
	24151392	GC	*FGFR4*	23901234	GC	*NOX5*	26901778	EAC
	27449042	CRC	*FHIT*	22957075	GC	*ODC1*	25079701	GC
ABO	27191536	GC	*FKBP5*	26467701	EAC		27433286	CRC
ACE	18059164	GC	*FLT1*	24398900	GC	*OGG1*	21822670	GC
ADAMTS12	25823933	ESCC		26041938	CRC		27022219	CRC
ADAMTS2	29358879	ESCC	*FMNL3*	25758200	CRC	*OLFM4*	26303970	GC
ADAMTS4	25786261	CRC	*FN1*	30214289	CRC	*ONECUT2*	24402783	CRC
ADAMTS5	25786261	CRC	*FOS*	22002117	GC	*OXT*	26406593	EAC
ADH1A	29248712	ESCC	*FOXF1*	26383589	EAC	*OXTR*	26406593	EAC
ADH1B	27038040	ESCC	*FOXM1*	25889361	EAC	*PARP1*	23757351	EAC
	25524923	GC		25907137	GC		26540566	GC
AFP	23382965	GC	*FOXP1*	26383589	EAC	*PCNA*	27207660	CRC
AGMAT	27433286	CRC	*FOXP3*	24657498	GC	*PDCD4*	27021515	GC
AKAP4	26590805	CRC	*FSCN1*	25823933	ESCC	*PDGFRA*	29358879	ESCC
AKT1	26818920	GC	*GCOM1*	25823933	ESCC		25264210	SIC
ALDH1A2	25447851	EAC	*GDF15*	18264720	GC	*PDXP*	21297586	SIC
ALDH2	29423536	ESCC		26497212	CRC	*PGC*	25549793	GC
	25524923	GC	*GDF7*	26783083	EAC	*PIK3CA*	26543351	GC
	26804999	CRC	*GKN1*	25752269	GC		28617917	SIC
ALOX15	27320813	CRC	*GKN2*	24408014	GC		27511117	CRC
AMD1	27433286	CRC	*GLI1*	24081672	GC	*PIK3CB*	26543351	GC
APC	26073472	GC	*GNAS*	28617917	SIC	*PIK3CD*	26543351	GC
	29575536	SIC	*GPBAR1*	28293080	EAC	*PIK3CG*	26543351	GC
	27325016	CRC	*GREM1*	29804199	CRC	*PIK3CG*	26404261	CRC
APEX1	25733810	GC	*GSK3B*	24982245	CRC	*PLAU*	29358879	ESCC
APOBEC1	25085003	EAC	*GSTK1*	27006283	GC		18466392	GC
AQP3	26768614	GC	*GSTM1*	24244742	GC		30175151	CRC
AR	26467701	EAC		26219826	CRC	*PLCE1*	26554163	GC
ARHGAP24	25929336	CRC	*GSTP1*	26467157	GC	*PMS2*	25029614	SIC
ARID1A	28440661	EAC		24970398	CRC	*POLB*	25561897	GC
	24744582	GC	*GSTT1*	27006283	GC	*POLD1*	30086056	CRC
	28617917	SIC		25249451	CRC	*POLE*	30175151	CRC
ARID2	28617917	SIC	*HGF*	27683052	GC	*POSTN*	25823933	ESCC
ASXL1	28617917	SIC	*HIF1A*	26931435	GC	*POU5F1*	22429998	GC
ATF3	27043582	CRC		27421843	CRC	*PPARD*	26004416	CRC
ATM	26587992	GC	*HLA-A*	24398900	GC	*PPARG*	25516376	GC
	28617917	SIC	*HLA-G*	29067119	ESCC		24631504	CRC
AURKA	25823933	ESCC	*HMGB1*	25821182	GC	*PPP1R1B*	24291029	GC
	26346770	GC		25778491	CRC	*PPP1R3C*	25823933	ESCC
AXIN2	26297437	EAC	*HMOX1*	26087182	CRC	*PPP2R1B*	26247730	CRC

Table C1 (continued).

1	2	3	1	2	3	1	2	3
AZIN1	27433286	CRC	*HNF4A*	25965823	GC	*PP2CA*	28904398	GC
BAG3	26577854	CRC	*HOXA10*	29358879	ESCC		30296597	CRC
BARX1	26383589	EAC	*HPGD*	29358879	ESCC	*PRKAA1*	26098866	GC
BCL2	26554163	GC	*HPGDS*	26759244	GC	*PROM1*	21064103	SIC
	25977444	CRC	*HPSE*	25727320	GC		27435662	CRC
BCL2L1	26401016	GC	*HSPA4*	24473833	GC	*PSCA*	26848528	GC
	29393430	CRC	*IBSP*	29358879	ESCC	*PSG2*	24305936	GC
BIRC5	26277021	GC	*ICAM1*	24548863	GC	*PTEN*	26753960	GC
	26442524	CRC	*IGF1*	17608636	GC		28617917	SIC
BMI1	28633632	ESCC		26381944	CRC		26098881	CRC
	24330850	GC	*IGF1R*	26337161	GC	*PTGS2*	31518663	GC
BMP2	21570392	GC	*IGF2*	24451943	GC		25818893	CRC
BMP4	25647270	CRC		27337954	CRC	*PTP4A3*	26563151	CRC
BMP7	29358879	ESCC	*IGF2BP3*	25822686	CRC	*RASSF1*	26735582	GC
BRAF	23511561	GC	*IGFBP2*	26317790	EAC		25942531	CRC
	26892442	SIC	*IGFBP3*	29358879	ESCC	*REG4*	21780125	GC
	25975986	CRC		24451943	GC	*RELA*	26801246	GC
BRCA1	25266802	GC	*IL10*	26088449	GC	*REM1*	25793713	GC
BTG3	25701359	EAC	*IL17A*	26287940	GC	*ROCK2*	26901778	EAC
C2orf40	12800218	ESCC		26559812	CRC	*RUNX3*	25229459	EAC
CAB39L	25823933	ESCC	*IL17F*	25147431	GC		25884934	GC
CASP10	15288293	GC	*IL1A*	25740226	GC	*S100A4*	26497012	GC
CASP3	26662569	GC	*IL1B*	26805397	GC	*S100A8*	23704827	GC
	26844701	CRC	*IL1RN*	24615437	GC	*SASH1*	25823933	ESCC
CASR	25879211	CRC	*IL2*	19058298	GC	*SDHA*	21771581	GC
CAV1	26817615	GC	*IL4*	24072495	GC	*SDHAF2*	21771581	GC
CCKBR	28109268	GC	*IL6*	25655003	GC	*SDHB*	21771581	GC
CCNA2	28800315	ESCC	*IL6ST*	26750536	GC	*SDHC*	21771581	GC
CCND1	28800315	ESCC	*IRF1*	27866197	GC	*SDHD*	21771581	GC
	25338835	GC	*IRS1*	26577117	CRC	*SEPP1*	22715394	EAC
	25809706	CRC	*ITGA3*	29358879	ESCC	*SEPT9*	27304597	CRC
CCNE1	28800315	ESCC	*ITGB1*	24870620	GC	*SERPINE1*	27383203	GC
CCR7	24040244	GC	*ITGB4*	29358879	ESCC		30214289	CRC
CD14	24978812	GC	*JAK2*	25055044	GC	*SETD2*	26760764	GC
CD34	25264210	SIC	*JUN*	22002117	GC	*SFRP1*	23943784	GC
CD44	26416034	GC	*KDR*	22842485	GC	*SFRP2*	29361151	CRC
	26010608	CRC	*KIAA1456*	24743840	CRC	*SGCB*	26850596	GC
CD55	26202380	EAC	*KISS1*	26847533	CRC	*SH3BGRL2*	25823933	ESCC
CDC27	26821069	CRC	*KLF6*	22581522	GC	*SHH*	27039174	GC
CDC42	28460460	CRC	*KRAS*	27699948	GC	*SIRT1*	24107295	GC
CDH1	27356686	GC		26892442	SIC		26975631	CRC
	26742007	CRC		27325016	CRC	*SIX1*	29358879	ESCC
CDH17	24465527	GC	*KSR1*	26673620	CRC	*SIX4*	29358879	ESCC
CDK9	28404924	EAC	*LAMA3*	29358879	ESCC	*SKP2*	24971481	GC
CDKN1A	26309358	GC	*LAMB3*	29358879	ESCC	*SLC22A3*	26136075	GC
CDKN1B	22343731	GC	*LAMC2*	29358879	ESCC	*SLC26A2*	23840040	CRC

Table C1 (continued).

1	2	3	1	2	3	1	2	3
CDKN2A	25280564	EAC	*LEP*	24569475	EAC	*SLCO6A1*	27006283	GC
	23504555	GC		24548863	GC	*SMAD2*	27383203	GC
	28617917	SIC	*LGALS3*	24971481	GC		25980495	CRC
	23703248	CRC	*LGALS9*	28586026	EAC	*SMAD4*	24952744	EAC
CDKN2B	28617917	SIC	*LGR5*	27435662	CRC		25947258	GC
CDX1	24415870	GC	*LRP1B*	28617917	SIC		28617917	SIC
	26537799	CRC	*LTA*	26554163	GC		27325016	CRC
CDX2	22838504	GC	*MACC1*	26919111	GC	*SMAD7*	26779637	CRC
	25663765	CRC		25884643	CRC	*SMOX*	27433286	CRC
CEACAM3	26471376	GC	*MAL*	25823933	ESCC	*SMR3B*	26548949	GC
	26556959	CRC	*MAP2K7*	27699948	GC	*SOAT1*	25987131	SIC
CEACAM5	24305936	GC	*MAPK1*	25516376	GC	*SOCS2*	25025962	CRC
CEP72	27527254	EAC	*MAPK3*	26683357	GC	*SOCS6*	25025962	CRC
CHEK2	23296741	GC	*MCL1*	26276722	GC	*SOX2*	28633632	ESCC
CHFR	24748501	GC	*MDK*	25345957	GC		28692180	EAC
CLDN18	24073219	GC	*MDM2*	27006283	GC		24325912	GC
COL10A1	29358879	ESCC		28617917	SIC	*SOX9*	28617917	SIC
COL11A1	29358879	ESCC		26840028	CRC	*SP1*	24613927	GC
COL1A1	29358879	ESCC	*MEIS1*	29358879	ESCC	*SPATA13*	17599059	CRC
COL27A1	29358879	ESCC	*MET*	26543351	GC	*SPARC*	25516351	GC
COL3A1	30175151	CRC	*MFAP2*	29358879	ESCC	*SPP1*	29358879	ESCC
COL7A1	29358879	ESCC	*MGLL*	25823933	ESCC		25847246	GC
CREB1	25825983	GC	*MGMT*	27291049	GC		30175151	CRC
CTNNB1	26675260	GC		27006309	CRC	*SRM*	27433286	CRC
	27546842	SIC	*MLF1*	29358879	ESCC	*STAT3*	26494556	GC
	26490308	CRC	*MLH1*	26121593	GC		26024373	CRC
CTSE	25348778	EAC		25029614	SIC	*SULT1A1*	26455829	ESCC
CXCL8	26199092	GC		25908759	CRC	*SULT1A2*	26455829	ESCC
CXCR1	25022956	GC	*MLL2*	28617917	SIC	*TBC1D9*	22781708	GC
CXCR4	26091251	GC	*MME*	25759539	SIC	*TBX5*	26783083	EAC
CYP1A1	26455829	ESCC	*MMP1*	25823933	ESCC	*TCF3*	29358879	ESCC
	24443269	GC		24505369	GC	*TCF4*	25961950	CRC
	26514676	CRC	*MMP10*	29358879	ESCC	*TCF7L2*	26240283	CRC
CYP1B1	27111221	CRC	*MMP11*	29358879	ESCC	*TERT*	2677859	GC
CYP24A1	28811712	CRC	*MMP12*	29358879	ESCC	*TFF1*	26240479	GC
CYP2E1	25596089	GC	*MMP13*	29358879	ESCC	*TFF2*	26265233	GC
CYP2W1	26563883	CRC	*MMP14*	26610210	GC	*TFF3*	26265233	GC
DCC	22252584	GC	*MMP2*	29344220	ESCC	*TGFB1*	27997896	GC
	25249451	CRC		25136862	GC		26873488	CRC
DCTN1	29864111	CRC	*MMP3*	29358879	ESCC	*TGFBR1*	26714817	CRC
DCTN2	29864111	CRC	*MMP7*	28968424	ESCC	*TGFBR2*	27156840	CRC
DCTN5	29864111	CRC		26983663	GC	*THBS1*	23765259	GC
DCTN6	29864111	CRC	*MMP9*	29358879	ESCC	*THBS2*	30214289	CRC
DKK3	26093488	EAC		25724188	GC	*TIMP2*	24440352	GC
DNM3OS	30770102	GC	*MORC2*	26098774	GC	*TIMP3*	27383203	GC
DNMT1	25595591	GC	*MRC1*	26530135	SIC	*TLR2*	26880585	GC

Table C1 (continued).

1	2	3	1	2	3	1	2	3
	26497367	CRC	*MSH2*	22386861	GC		26167073	CRC
DNMT3B	26172537	GC		25759539	SIC	*TLR4*	23160836	GC
	26497367	CRC		25534380	CRC	*TNF*	26088449	GC
DPYD	25112781	GC	*MSH6*	25029614	SIC		26572151	CRC
	26281864	CRC		26099011	CRC	*TNFRSF4*	26439988	CRC
E2F1	26867682	GC	*MTA1*	27320813	CRC		29351612	ESCC
EGF	25111897	GC	*MTAP*	27433286	CRC		26733670	EAC
	27997901	GC		28046029	ESCC	*TP53*	26867682	GC
EGFR	26892442	SIC	*MTHFR*	26314858	GC		27546842	SIC
	27064574	CRC		25823789	CRC		26991699	CRC
EGR1	29047233	ESCC	*MTOR*	24729090	GC	*TSC1*	24770722	CRC
ELN	25823933	ESCC		28212575	EAC	*TWIST1*	25368021	GC
EMP1	25823933	ESCC	*MUC1*	25658482	GC	*TYMS*	25185651	GC
EPCAM	26028961	CRC		27207660	CRC		25677447	CRC
EPHB2	26768611	GC		25704078	GC	*UBAP2L*	27456362	CRC
	24151090	EAC	*MUC2*	25759539	SIC	*UGT1A1*	24114122	SIC
ERBB2	26970343	GC		27207660	CRC	*UQCRB*	22545919	CRC
	24797764	SIC	*MUC5AC*	25759539	SIC	*VDR*	25910066	EAC
	26991109	CRC	*MUC6*	25704078	GC		27000664	GC
	24151090	EAC		25759539	SIC	*VEGFA*	27546842	SIC
ERBB3	24973425	GC	*MUTYH*	22126480	GC		26896610	CRC
	26367378	CRC		26377631	CRC	*VEGFC*	29393383	GC
ERCC1	29188501	ESCC	*MYC*	24130168	GC	*VIM*	26264654	GC
	23884439	GC		26152922	CRC	*WIF1*	29568950	CRC
ESR1	26406593	EAC	*NAT1*	22905173	CRC	*XAF1*	20198350	GC
	25970543	CRC		25886288	ESCC	*XRCC1*	26846663	GC
ESR2	25650184	CRC	*NAT2*	25051916	GC		25582318	CRC
EZH2	25618371	GC		26683305	CRC	*ZEB1*	22466758	GC
	30214616	CRC	*NF1*	28617917	SIC		26387539	CRC
F2R	25231630	GC	*NFKB1*	26801246	GC	*ZEB2*	27761656	GC
FASN	26109148	CRC	*NME1*	22903495	GC	*ZNF717*	25422082	GC
FGFR2	26516773	GC	*NOD2*	26378020	CRC	*ZNRD1*	23893241	GC

Table C2 - The characteristics of miRNA interactions with 5'UTR mRNAs of EAC candidate genes.

Gene; NX	miRNA	Position, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
1	2	3	4	5	6
ALDH1A2; 2,3	ID00792.3p-miR	10	-132	95	22
	ID00744.3p-miR	122	-125	89	23
AR; 2,2	ID01739.3p-miR	621	-115	95	21
BARX1; 0,3	ID01675.5p-miR	61	-121	92	21
BTG3; 19,9	ID01123.3p-miR	155	-127	88	24
	ID02803.5p-miR	186	-108	91	21
	ID00278.3p-miR	230	-125	91	23
CD55; 36,1	miR-1587	148	-108	94	20

Table C2 (continued)

1	2	3	4	5	6
ERBB2; 40,9	miR-6090	25	-110	95	19
ERBB3; 29,4	ID00099.3p-miR	114	-113	93	21
FKBP5; 29,2	ID01708.5p-miR	11	-115	89	23
FOXM1; 13,5	ID01615.3p-miR	68	-115	90	22
GDF7; 0,9	ID03054.3p-miR	125	-113	90	22
	ID02682.5p-miR	339	-113	93	20
	ID01112.3p-miR	446	-113	93	20
GPBAR1; 1,3	miR-4726-5p	836	-119	90	23
ROCK2; 12,4	ID01184.3p-miR	248	-117	93	20
SMAD4; 21,7	ID00577.3p-miR	160	-106	94	20
	ID00961.3p-miR	248	-127	90	23
SOX2; 15,9	ID01749.3p-miR	344	-117	89	23

Table C3 - The characteristics of miRNA interactions with CDS mRNAs of EAC candidate genes.

Gene; NX	miRNA	Position, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
1	2	3	4	5	6
AR; 2,2	ID00372.5p-miR	1363	-127	94	24
	ID01398.3p-miR	1517	-115	90	22
	ID01261.5p-miR	1744	-113	95	20
ARID1A; 23,2	miR-4767	938	-121	85	23
	ID01473.3p-miR	1093	-125	89	23
	ID01508.5p-miR	1459	-129	90	23
	ID02945.5p-miR	4274	-108	91	21
	miR-6799-5p	5130	-108	93	20
	ID01819.5p-miR	7062	-119	89	23
AXIN2; 6,5	ID01796.3p-miR	1526	-125	89	24
	ID00648.5p-miR	1765	-125	92	22
	ID02344.3p-miR	2086	-127	88	24
BARX1; 0,3	ID01757.3p-miR	192	-117	93	21
	ID02052.5p-miR	311	-132	89	24
	ID02429.3p-miR	314	-121	89	23
	ID02079.5p-miR	500	-115	92	20
	ID00380.3p-miR	558	-117	89	23
BTG3; 19,9	ID02017.3p-miR	664	-117	92	22
CD55; 36,1	miR-1273d	1425	-119	87	25
	miR-5585-5p	1471	-106	91	22
CEP72; 4,6	miR-939-3p	1238	-113	91	21
	ID01839.3p-miR	1363	-123	89	23
DKK3; 8,8	ID03402.5p-miR	295	-117	92	22
	miR-4269	870	-113	93	21
ESR1; 1,1	ID02606.5p-miR	409	-106	91	22
	ID02556.3p-miR	1852	-115	90	23
FOXF1; 17,9	ID00267.3p-miR	55	-113	91	21
	miR-6720-5p	299	-121	90	23
	ID00407.3p-miR	718	-119	90	22

Table C3 (continued)

1	2	3	4	5	6
FOXM1; 13,5	ID01279.5p-miR	1897	-115	93	22
GDF7; 0,9	ID00102.3p-miR	784	-117	95	20
	ID01419.3p-miR	876	-115	89	23
	ID01346.3p-miR	1576	-119	92	22
IGFBP2; 26,4	ID01859.5p-miR	198	-117	90	22
ROCK2; 12,4	miR-1237-3p	535	-108	91	21
RUNX3; 3,5	ID02259.5p-miR	593	-119	89	23
	ID00024.5p-miR	1406	-119	93	21
SOX2; 15,9	ID03289.5p-miR	455	-106	93	20
	ID01259.3p-miR	1218	-121	90	23
TBX5; 6,4	miR-671-5p	1558	-119	90	23

Table C4 - The characteristics of miRNA interactions with 3'UTR mRNAs of EAC candidate genes.

Gene; NX	miRNA	Position, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
AR; 2,2	ID02403.3p-miR	4096	-108	89	23
AXIN2; 6,5	ID01257.3p-miR	2960	-113	93	20
	ID02524.5p-miR	3702	-93	90	22
ERBB3; 29,4	miR-619-5p	4950	-117	96	22
	miR-619-5p	5104	-121	100	22
ESR1; 1,1	ID03196.3p-miR	3339	-121	88	24
	miR-6879-5p	3593	-113	90	22
FKBP5; 29,2	ID02017.3p-miR	1434	-115	90	22
	ID00625.5p-miR	5727	-106	91	21
	miR-619-5p	7114	-121	100	22
	miR-5096	7186	-106	94	21
	miR-5585-3p	7257	-113	96	22
	ID02175.3p-miR	7306	-110	91	22
FOXM1; 13,5	miR-6809-5p	2864	-104	91	22
	ID00962.3p-miR	3260	-119	90	23
FOXP1; 19,6	ID03465.3p-miR	3843	-117	90	22
LEP; 0,0	miR-5096	3172	-104	92	21
	miR-5585-3p	3240	-108	93	22
ROCK2; 12,4	miR-1277-5p	5212	-96	88	24
	miR-619-5p	6272	-110	91	22
	ID01640.5p-miR	6308	-117	89	24
	miR-5096	6346	-104	92	21
	ID01836.5p-miR	6497	-115	92	23
	miR-1277-5p	6988	-96	88	24
RUNX3; 3,5	ID02589.5p-miR	2956	-113	93	21
SMAD4; 21,7	ID01838.5p-miR	4291	-113	90	24
	miR-1972	4552	-110	90	22
	ID01592.3p-miR	8227	-117	89	23
SOX2; 15,9	ID00037.3p-miR	1671	-121	90	23
	ID00125.3p-miR	1671	-113	90	22

Table C5 - The characteristics of miRNA interactions with 5'UTR mRNAs of ESCC candidate genes.

Gene; NX	miRNA	Position, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
ADH1B; 33,5	ID02831.3p-miR	57	-110	91	22
ALDH2; 26,4	ID02142.3p-miR	8	-123	92	21
AURKA; 7,7	ID01697.5p-miR	184	-108	93	20
	miR-1285-3p	352	-106	91	22
BMP7; 17,2	ID00962.3p-miR	181	-117	89	23
C2orf40; 16,5	ID00813.5p-miR	85	-125	88	24
CCNE1; 4,7	ID02692.5p-miR	1	-129	91	23
EGR1; 47,2	miR-4763-3p	52	-125	88	24
	ID00966.5p-miR	90	-138	92	24
	ID00051.3p-miR	216	-123	89	24
HPGD; 32,0	ID02556.3p-miR	168	-115	90	23
ITGB4; 24,0	miR-3656	27	-102	98	17
LAMA3; 12,4	ID03332.3p-miR	13	-138	93	24
MGLL; 63,8	miR-4687-5p	433	-117	93	22
MMP11; 0,8	ID00959.3p-miR	4	-106	93	20
SASH1; 35,6	ID02660.3p-miR	448	-119	90	22
SIX4; 0,5	ID00057.3p-miR	19	-108	91	22
SOX2; 15,9	ID01749.3p-miR	344	-117	89	23
SPP1; 1,5	miR-1913	61	-117	92	22

Table C6 - The characteristics of miRNA interactions with CDS mRNAs of ESCC candidate genes.

Gene; NX	miRNA	Position, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
1	2	3	4	5	6
ABCG2; 1,5	ID02408.3p-miR	1672	-104	91	22
ADAMTS12; 1,4	ID02607.3p-miR	2817	-117	93	22
ADH1B; 33,5	miR-4647	505	-110	90	23
BMP7; 17,2	ID03473.3p-miR	1617	-102	91	21
	ID03162.3p-miR	1686	-106	91	22
CCND1; 37,7	miR-4731-3p	508	-110	91	22
CCNE1; 4,7	ID00275.3p-miR	1356	-108	91	22
COL10A1; 0,5	ID01242.3p-miR	892	-121	88	24
COL1A1; 17,4	ID01731.5p-miR	1458	-115	89	23
	ID02001.5p-miR	1715	-110	90	22
	ID00707.5p-miR	3860	-110	91	22
COL7A1; 11,3	miR-3677-5p	879	-113	90	22
	miR-659-3p	1853	-110	90	22
	miR-671-5p	4160	-119	90	23
	ID00356.3p-miR	4491	-110	90	22

Table C6 (continued)

1	2	3	4	5	6
	miR-378a-5p	5130	-119	98	22
	ID00101.3p-miR	6143	-113	90	22
	ID02050.5p-miR	6240	-117	89	24
	ID03025.5p-miR	7016	-104	92	20
	ID03317.3p-miR	7503	-113	88	24
	ID00542.3p-miR	8089	-123	88	24
COL27A1; 3,6	ID00135.5p-miR	468	-110	95	21
	ID01745.3p-miR	693	-110	90	23
	miR-6803-5p	1294	-121	92	22
	miR-4763-3p	1308	-129	91	24
	ID00646.5p-miR	2115	-110	90	23
	miR-939-3p	3221	-113	91	21
	miR-6852-5p	3392	-106	91	21
	ID02318.5p-miR	3989	-115	89	23
	miR-3615	4064	-113	93	21
	ID02238.5p-miR	4817	-121	90	23
CYP1A1; 1,2	miR-6894-5p	567	-119	89	24
EGR1; 47,2	ID00564.5p-miR	806	-110	90	22
	ID01787.3p-miR	1682	-115	89	23
ELN; 30,1	miR-331-5p	173	-110	91	22
	ID00372.5p-miR	844	-125	92	24
	ID02634.3p-miR	1417	-117	90	22
	ID01107.3p-miR	2215	-119	90	23
FSCN1; 22,3	ID02191.5p-miR	405	-110	91	21
HLA-G; 0,1	ID01960.3p-miR	649	-106	93	20
HOXA10; 0,3	ID03242.5p-miR	171	-108	93	20
	ID01415.5p-miR	718	-106	91	21
	miR-4449	751	-121	92	22
	ID00338.3p-miR	784	-119	93	21
IBSP; 0,1	miR-562	978	-93	92	20
ITGA3; 12,8	ID03393.3p-miR	3448	-115	92	22
ITGB4; 24,0	miR-3940-5p	192	-108	93	20
	ID01388.5p-miR	200	-121	90	23
	ID01739.3p-miR	2335	-110	91	21
	ID00261.3p-miR	2339	-110	90	22
	ID02375.3p-miR	3093	-123	89	24
	ID01972.5p-miR	4431	-110	93	20
	ID00811.3p-miR	4443	-125	88	24
	ID01995.3p-miR	4452	-106	94	19
	miR-1225-3p	5167	-117	90	22
	miR-8071	5481	-108	94	20
	ID01005.5p-miR	5497	-110	90	23
LAMA3; 12,4	ID00586.3p-miR	96	-125	91	23
	ID02198.3p-miR	4532	-102	91	21
	miR-4300	6804	-93	96	18

Table C6 (continued)

1	2	3	4	5	6
LAMB3; 22,1	ID02430.3p-miR	1716	-121	89	23
LAMC2; 4,2	ID02668.5p-miR	347	-132	91	24
	miR-617	375	-106	91	22
	miR-29a-3p	2312	-102	91	22
	miR-935	2716	-115	89	23
MEIS1; 7,0	ID00610.3p-miR	761	-117	90	22
	ID01633.3p-miR	765	-106	91	21
MFAP2; 1,0	ID00455.5p-miR	394	-115	89	23
	miR-6871-3p	670	-113	90	22
MGLL; 63,8	ID00979.5p-miR	982	-117	89	23
MMP2; 52,3	ID02146.5p-miR	379	-125	91	23
	miR-1285-5p	1376	-104	92	21
	ID01636.5p-miR	1681	-108	91	21
	ID01456.3p-miR	1691	-113	90	22
MMP3; 0,2	ID00314.3p-miR	132	-119	93	23
MMP9; 0,7	ID01341.5p-miR	220	-113	90	22
	ID01104.5p-miR	698	-113	93	22
MMP11; 0,8	miR-4479	38	-121	92	22
MMP12; 0,1	miR-1245b-5p	780	-96	92	21
MMP13; 0,0	miR-30d-3p	207	-104	91	22
SASH1; 35,6	ID03170.5p-miR	1478	-102	91	21
	ID02281.5p-miR	4059	-115	90	22
SIX4; 0,5	ID00101.3p-miR	247	-113	90	22
	ID01868.3p-miR	280	-129	88	24
SOX2; 15,9	ID03289.5p-miR	455	-106	93	20
	ID01259.3p-miR	1218	-121	90	23
SULT1A1; 2,7	ID00111.3p-miR	595	-106	93	20
	ID02210.5p-miR	795	-110	90	22
SULT1A2; 0,2	ID00111.3p-miR	473	-106	93	20
	ID02210.5p-miR	673	-110	90	22
TCF3; 8,0	ID03058.3p-miR	490	-119	89	24
	ID02050.3p-miR	1217	-117	89	23

Table C7 - The characteristics of miRNA interactions with 3'UTR mRNAs of ESCC candidate genes.

Gene; NX	miRNA	Position, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
1	2	3	4	5	6
ADAMTS2; 6,0	ID01758.3p-miR	5915	-121	88	24
AURKA; 7,7	miR-506-3p	1905	-102	92	21
BMP7; 17,2	ID01153.3p-miR	2164	-110	91	21
	miR-5096	2815	-110	98	21
	miR-5585-3p	2883	-108	93	22
CCND1; 37,7	ID01220.5p-miR	1159	-117	90	24

Table C7 (continued)

1	2	3	4	5	6
	miR-6865-3p	1350	-106	91	21
	ID00108.3p-miR	3401	-125	89	24
COL1A1; 17,4	ID00076.5p-miR	4903	-106	91	21
COL11A1; 0,0	miR-624-3p	5903	-98	92	21
ELN; 30,1	ID02119.3p-miR	2395	-113	90	22
	miR-4800-5p	2773	-106	93	21
	miR-4800-5p	3134	-108	94	21
ERCC1; 10,0	ID01358.5p-miR	1544	-127	94	24
	miR-1273g-3p	2612	-108	93	21
	miR-1273e	2655	-113	96	22
FSCN1; 22,3	miR-5739	1760	-104	92	20
	miR-4763-3p	2482	-125	88	24
	miR-4670-3p	2713	-102	91	22
HPGD; 32,0	ID01950.5p-miR	1562	-96	90	22
IGFBP3; 11,7	ID01696.3p-miR	1753	-108	91	21
ITGA3; 12,8	ID02900.5p-miR	4576	-106	91	22
MGLL; 63,8	ID01543.3p-miR	1538	-115	90	22
	miR-3136-3p	2407	-104	91	22
	ID02815.3p-miR	3331	-110	91	21
MTHFR; 3,5	ID03245.5p-miR	3350	-117	89	24
	miR-8089	3460	-121	88	24
	ID00345.3p-miR	6281	-119	89	23
	miR-5585-3p	6300	-108	93	22
	ID03407.3p-miR	6342	-108	91	22
	miR-1285-5p	6399	-104	92	21
	miR-5585-3p	7003	-110	95	22
	ID02175.3p-miR	7051	-113	93	22
SASH1; 35,6	ID01334.3p-miR	5752	-113	90	22
SH3BGRL2; 29,3	miR-4666a-3p	4116	-96	94	21
	miR-3658	4602	-98	90	22
SIX4; 0,5	miR-5585-5p	4131	-106	91	22
SULT1A1; 2,7	miR-1285-5p	1529	-108	96	21
TCF3; 8,0	miR-328-5p	2787	-123	89	23
	miR-7843-5p	3007	-110	90	22

Table C8 - The characteristics of the miRNA interactions with 5′UTR mRNAs of gastric cancer candidate genes.

Gene; NX	miRNA	Start of site, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
1	2	3	4	5	6
AURKA; 4.3	miR-5095	420	-108	93	21
	miR-619-5p	426	-119	98	22
DNMT1; 9.0	ID02052.5p-miR	137	-134	90	24

Table C8 (continued)

1	2	3	4	5	6
EGFR; 7.8	ID02457.3p-miR	89	-132	95	22
EZH2; 2.7	ID02761.3p-miR	120	-138	93	24
FGFR2; 9.3	ID03047.3p-miR	48	-132	89	24
JUN; 47.0	ID02008.5p-miR	269	-123	91	22
KRAS; 23.6	ID01310.3p-miR	17, 29	-121, -123	92,94	22
	ID03332.3p-miR	37	-132	89	24
LGALS3; 36.9	ID00329.3p-miR	21	-125	91	22
MMP2; 20.5	ID00278.3p-miR	110	-123	89	23
	ID03345.5p-miR	124	-127	90	24
NFKB1; 14.6	ID02064.5p-miR	128	-132	91	23
POU5F1; 6.3	miR-1273a	181	-119	90	25
	miR-1273g-3p	203	-113	96	21
	miR-1273d	237	-119	87	25
PTEN; 17.8	ID01310.3p-miR	531	-119	90	22
	ID02761.3p-miR	533	-132	89	24
	ID03037.3p-miR	536	-121	90	22
SFRP1; 11.0	ID01641.3p-miR	111	-132	89	24
SMAD2; 26.3	ID01492.3p-miR	16	-129	100	23
TGFB1; 12.8	ID00457.3p-miR	209	-129	95	22
	ID02064.5p-miR	211	-132	91	23
THBS1; 46.5	miR-328-5p	84	-123	89	23
	miR-6786-5p	88	-119	93	21
	ID01652.3p-miR	113	-129	92	23
TIMP2; 24.*2*	ID01520.3p-miR	60	-123	92	21
	ID00252.5p-miR	61	-134	90	24
	ID02668.5p-miR	219	-127	88	24
	ID00961.3p-miR	239	-132	93	23
	ID00049.5p-miR	269	-136	90	24
	ID00417.3p-miR	278	-123	94	21
	ID01293.5p-miR	281	-127	94	22
TIMP3; 22.9	ID02903.3p-miR	1102	-121	90	22
TYMS; 4.2	ID00252.5p-miR	87	-132	89	24

Table C9 - The characteristics of miRNA interactions with CDS mRNAs of gastric cancer candidate genes.

Gene; NX	miRNA	Position, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
1	2	3	4	5	6
ACE; 1.9	ID00522.5p-miR	62	-125	89	23
	ID02294.5p-miR	64	-132	90	24
ARID1A; 19.5	ID01473.3p-miR	1093	-125	89	23
	ID03167.3p-miR	1399	-123	91	23
	ID01508.5p-miR	1459	-129	90	23
CDX2; 0.1	ID01895.5p-miR	416	-138	93	24
	ID00522.5p-miR	419	-125	89	23

Table C9 (continued)

1	2	3	4	5	6
	ID02052.5p-miR	608	-136	91	24
	ID01041.5p-miR	609	-129	88	24
	miR-3960	613	-117	93	20
EPHB2; 2.3	miR-4253	1089	-102	100	18
IGFBP3; 8.3	ID02982.3p-miR	435	-123	100	21
MUC1; 105.0	ID00645.5p-miR	507	-110	93	20
	miR-6752-5p	510	-119	90	22
MUC6; 76.6	miR-6815-3p	1218	-104	91	21
	ID02841.5p-miR	1238	-108	93	20
	miR-939-5p	1330	-121	88	24
OGG1; 10.8	ID00920.5p-miR	882	-134	94	24
PARP1; 15.9	ID01616.3p-miR	1275	-119	90	23
PGC; 963.7	ID01303.5p-miR	715	-115	92	20
SIRT1; 13.2	miR-4767	236	-134	94	23
	ID01560.3p-miR	264	-123	89	23
	ID03332.3p-miR	280, 287	-132, 138	89, 93	24
	ID00278.3p-miR	294	-123	89	23
	ID00811.3p-miR	300	-125	88	24

Table C10 - The characteristics of miRNA interactions with 3'UTR mRNAs of gastric cancer candidate genes.

Gene; NX	miRNA	Start of site, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
1	2	3	4	5	6
AQP3; 3.9	miR-466	1255	-106	91	23
	ID00436.3p-miR	1255	-104	89	23
	ID01030.3p-miR	1259, 1269	-108	89	23
BRCA1; 5.1	miR-5095	6406	-106	91	21
	miR-619-5p	6412	-119	98	22
	miR-5096	6486	-110	98	21
	miR-5585-3p	6554	-110	95	22
CASP10; 8.7	miR-1273h-5p	2234	-106	91	21
	miR-1285-3p	2357	-106	91	22
	miR-1273g-3p	2589	-108	93	21
	miR-1273d	2623	-119	87	25
	miR-1273h-5p	2623	-117	100	21
	miR-619-5p	3247	-113	93	22
	miR-5585-3p	3389	-108	93	22
CCND1; 19.2	miR-574-5p	2593÷2597 (3)	-108÷-113	89÷93	23
	ID00470.5p-miR	2595, 2597	-108	89	23
CDH1; 21.6	miR-1273c	3251	-110	91	22
	miR-1273g-3p	3271	-108	93	21
	miR-1273h-5p	3305	-113	96	21
CDH17; 0.3	miR-1273a	2773	-119	90	25
	miR-1273g-3p	2795	-110	95	21
	miR-1273d	2829	-121	89	25

Table C10 (continued)

1	2	3	4	5	6
CDKN1A; 25.2	miR-4433-3p	1601	-106	91	21
CEACAM5; 20.1	miR-5585-3p	2441	-108	93	22
	miR-5095	3229	-115	98	21
	miR-619-5p	3235	-119	98	22
	miR-5585-3p	3378	-113	96	22
CLDN18; 126.9	miR-5095	1852	-106	91	21
	miR-619-5p	1858	-115	95	22
	miR-1273h-3p	2053	-113	90	22
CREB1; 13.3	miR-5095	2791	-110	95	21
	miR-619-5p	2797	-119	98	22
	miR-5096	2871	-106	94	21
	miR-619-5p	2932	-110	91	22
	miR-5585-3p	2939	-108	93	22
ERBB3; 32.9	miR-619-5p	4950	-117	96	22
	miR-619-5p	5104	-121	100	22
F2R; 4.5	miR-5095	3239	-115	98	21
	miR-619-5p	3245	-119	98	22
	miR-5585-3p	3387	-113	96	22
HPSE; 2.4	miR-1273g-3p	2081	-106	91	21
	miR-1273g-3p	2376	-108	93	21
	miR-5095	2759	-106	91	21
	miR-619-5p	2765	-110	91	22
	miR-5096	2839	-110	98	21
IL10; 2.0	miR-5095	1210	-115	98	21
	miR-619-5p	1216	-119	98	22
	miR-5096	1290	-106	94	21
IRF1; 17.5	miR-5095	2229	-110	95	21
	miR-619-5p	2235	-115	95	22
	miR-619-5p	2523	-110	91	22
	miR-5096	2597	-110	98	21
	miR-5095	2653	-110	95	21
	miR-619-5p	2659	-119	98	22
	miR-5585-3p	2800	-110	95	22
KRAS; 23.6	ID03224.5p-miR	3163	-121	92	23
	miR-1273g-3p	3176	-108	93	21
	miR-1273f	3209	-104	100	19
LEP; 0.2	miR-5095	3094	-110	95	21
	miR-619-5p	3100	-119	98	22
	miR-5096	3172	-104	92	21
	miR-5585-3p	3240	-108	93	22
MACC1; 5.0	ID01030.3p-miR	2962, 2976	-108	89	23
	ID01727.5p-miR	2963÷2981 (3)	-104	89	23
	miR-466	2966÷2990 (4)	-104÷-106	89÷91	23
	ID01030.3p-miR	3949÷3965 (8)	-108÷-110	89÷91	23
	miR-466	3951÷3967 (9)	-106÷-108	91÷93	23
	ID00436.3p-miR	3953÷3967 (8)	-104÷-106	89÷91	23
	miR-619-5p	4856	-117	96	22
	miR-1273f	5717	-102	98	19

Table C10 (continued)

1	2	3	4	5	6
	miR-1273d	5718	-123	91	25
MDM2; 9.8	miR-1273g-3p	2117	-113	96	21
	miR-1273g-3p	2486	-106	91	21
	miR-1273e	2521	-108	93	22
	miR-1273g-3p	6739	-113	96	21
MTHFR; 13.6	miR-5095	6855	-110	95	21
	miR-619-5p	6861	-115	95	22
	miR-5585-3p	7003	-110	95	22
MUC6; 76.6	ID01401.3p-miR	7895	-108	91	21
PDCD4; 27.6	miR-5095	3215	-106	91	21
	miR-619-5p	3221	-121	100	22
	miR-619-5p	3355	-110	91	22
PRKAA1; 22.1	miR-1273g-3p	2260	-110	95	21
	miR-1273f	2293	-102	98	19
	miR-1273d	2294	-121	89	25
	ID01404.5p-miR	2297	-113	91	23
	miR-1273e	2303	-106	91	22
SGCB; 9.0	miR-1273c	3314	-110	91	22
	miR-1273g-3p	3332	-110	95	21
SMAD4; 19.3	ID01838.5p-miR	4291	-113	90	24
	miR-1273g-3p	4312	-110	95	21
	ID01656.3p-miR	4342	-115	89	23
	miR-1273d	4346	-119	87	25
	ID01404.5p-miR	4349	-115	93	23
	ID02732.3p-miR	7721	-123	91	23
	miR-574-5p	7740÷7756 (8)	-108÷-113	89÷93	23
	ID00470.5p-miR	7744÷7752 (5)	-108	89	23
	ID00106.5p-miR	7825	-106	91	22
	miR-574-5p	7830	-108	89	23
STAT3; 29.4	miR-5095	3125	-106	91	21
	miR-619-5p	3131	-119	98	22
	miR-5585-3p	3268	-110	95	22
TGFB1; 12.8	miR-6089	2060, 2065	-132, -136	89, 91	24
	ID03306.3p-miR	2060	-123	94	21
	ID01382.3p-miR	2062	-113	93	20
	ID03208.5p-miR	2066	-125	88	24
	miR-4651	2087	-113	95	20
	ID00978.5p-miR	2089	-119	90	22
	ID00296.3p-miR	2093	-140	89	25
	miR-1183	4361	-132	90	27
TIMP2; 24.2	ID01941.5p-miR	1427	-117	89	24
TIMP3; 22.9	miR-1224-5p	3268	-104	96	19
TP53; 12.0	ID00548.3p-miR	1393	-115	89	23
	ID02379.3p-miR	1397	-119	89	24
	miR-1273c	2297	-110	91	22
	miR-1285-3p	2301	-110	95	22
	miR-1273g-3p	2317	-106	91	21
	miR-1273h-5p	2351	-106	91	21

Table C10 (continued)

1	2	3	4	5	6
	ID01838.5p-miR	2460	-110	88	24
	ID00785.5p-miR	2520	-113	90	23

Table C11 - Characteristics of the binding sites of miRNA and mRNA in 5'UTR of genes involved in the development of SIC.

Gene; NX	miRNA	Position, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
APC; 4,7	miR-6809-3p	147	-102	91	21
CTNNB1; 56,1	ID00477.5p-miR	77	-113	95	20
	ID03206.5p-miR	152	-115	92	20
EGFR; 5,9	ID02457.3p-miR	89	-132	95	22
ERBB2; 28,0	miR-6090	25	-110	95	19
PROM1; 7,7	miR-1292-3p	288	-110	93	20
SMAD4; 8,6	ID00577.3p-miR	160	-106	94	20
	ID00961.3p-miR	248	-127	90	23
SOAT1; 4,9	ID03036.3p-miR	46	-115	89	23
SOX9; 8,2	ID01969.3p-miR	294	-113	96	20

Table C12 - Characteristics of the binding sites of miRNA and mRNA in CDS of genes involved in the development of SIC.

Gene; NX	miRNA	Position, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
1	2	3	4	5	6
ATM; 7,6	miR-3681-3p	9439	-104	91	22
ASXL1; 15,1	ID01859.5p-miR	2370	-117	90	22
	ID03444.5p-miR	4714	-106	93	22
CDKN2B; 19,2	miR-1229-3p	387	-115	89	23
	ID02899.3p-miR	412	-132	89	24
CTNNB1; 56,1	miR-6083	1773	-98	92	20
EGFR; 5,9	ID02344.3p-miR	1779	-127	88	24
	miR-525-5p	1854	-102	91	21
LRP1B; 0,0	ID02669.5p-miR	14051	-96	92	20
FBXW7; 4,3	ID02514.3p-miR	1243	-108	93	22
	miR-3162-3p	2137	-108	91	21
MRC1; 9,7	ID02169.5p-miR	1259	-115	89	23
MSH6; 3,7	ID00156.5p-miR	246	-134	90	24
	miR-4527	2816	-104	91	22
MUC2; 19,0	miR-6840-3p	381	-117	93	22
	ID00910.3p-miR	8453	-110	90	22
PDXP; 2,3	ID01242.3p-miR	173	-125	91	24
	ID01877.3p-miR	386	-123	88	24
	ID01610.5p-miR	646	-110	91	21
PIK3CA; 3,5	ID00724.5p-miR	3043	-100	92	21

Table C12 (continued)

1	2	3	4	5	6
SOX9; 8,2	ID00978.5p-miR	1065	-119	90	22
	ID02011.3p-miR	1446	-106	93	20
UGT1A1; 3,1	ID00091.5p-miR	465	-117	89	23
VEGFA; 38,3	ID00678.3p-miR	627	-113	91	21
	ID03097.3p-miR	888	-121	89	23

Table C13 - Characteristics of the binding sites of miRNA and mRNA in 3'UTR of genes involved in the development of SIC.

Gene; NX	miRNA	Position, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
APC; 4,7	miR-4693-5p	9182	-108	94	23
ASXL1; 15,1	miR-1910-3p	6549	-100	92	20
CD34; 18,1	miR-6124	1302	-102	92	20
	miR-4716-3p	2459	-106	91	22
MUC6; 0,0	ID01401.3p-miR	7883	-108	91	21
NF1; 3,0	ID02733.5p-miR	10571	-89	91	21
	miR-8067	10908	-102	92	22
SOX9; 8,2	ID00509.3p-miR	2153	-98	92	21
	miR-466	3414	-104	89	23
VEGFA; 38,3	miR-1277-5p	2085	-96	88	24
	miR-1277-5p	3275	-96	88	24
	miR-1277-5p	3314	-96	88	24

Table C14 - Characteristics of the binding sites of miRNA and mRNA in 5'UTR of genes involved in the development of CRC.

Gene; NX	miRNA	Position, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
1	2	3	4	5	6
ADAMTS4; 7,9	ID00183.3p-miR	73	-102	91	21
ADAMTS5; 1,9	ID00119.3p-miR	288	-121	92	22
ALDH2; 124,7	ID02142.3p-miR	8	-123	92	21
APC; 3,3	miR-6809-3p	147	-102	91	21
ATF3; 17,3	ID02770.5p-miR	153	-115	92	20
BAG3; 37,2	miR-6850-5p	253	-121	92	22
BCL2L1; 20,5	miR-877-5p	149	-104	92	20
BIRC5; 0,7	ID01145.3p-miR	110	-125	89	23
CD44; 19,1	ID01213.5p-miR	129	-121	90	23
	ID02860.5p-miR	376	-113	91	21
CDC42; 46,3	miR-8088	52	-104	91	21
	ID01838.5p-miR	115	-121	97	24
	ID01767.5p-miR	120	-110	91	22
	miR-1273g-3p	136	-113	96	21
CDH1; 26,4	ID00894.5p-miR	57	-123	89	23

Table C14 (continued)

1	2	3	4	5	6
	ID02781.3p-miR	60	-117	93	20
CTNNB1; 45,4	ID00477.5p-miR	77	-113	95	20
	ID03206.5p-miR	152	-115	92	20
DCC; 0,0	miR-711	83	-110	90	22
	ID01643.3p-miR	440	-121	92	22
	ID00777.3p-miR	482	-115	92	23
	miR-466	535, 543	-104	89	23
DCTN1; 29,9	ID01911.5p-miR	287	-125	91	23
DNMT1; 7,3	ID02052.5p-miR	137	-134	90	24
DNMT3B; 0,5	ID02106.3p-miR	49	-123	89	23
EGFR; 7,0	ID02457.3p-miR	89	-132	95	22
EPCAM; 140,8	ID01857.3p-miR	237	-117	92	21
	ID00316.3p-miR	244	-119	89	23
	ID01317.5p-miR	314	-115	89	23
	miR-328-5p	316	-123	89	23
ERBB2; 21,5	miR-6090	25	-110	95	19
ERBB3; 17,5	ID00099.3p-miR	114	-113	93	21
ESR2; 0,1	ID01280.3p-miR	41	-117	92	22
FASN; 19,5	miR-4430	52	-98	96	18
	ID02256.3p-miR	76	-125	91	22
	ID03120.3p-miR	81	-119	93	21
	miR-6785-5p	87	-113	91	22
FMNL3; 3,3	ID00584.5p-miR	51	-117	90	22
	miR-3960	52	-115	92	20
	ID00296.3p-miR	53	-140	89	25
	D01634.3p-miR	185	-121	89	23
FN1; 181,7	miR-6892-3p	17	-110	93	21
	ID01723.5p-miR	109	-115	90	23
	miR-1914-3p	230	-117	90	22
GSTP1; 161,0	ID00267.3p-miR	77	-115	93	21
	ID03331.3p-miR	91	-129	90	23
HIF1A; 18,9	miR-6789-5p	54	-132	90	24
	ID01675.5p-miR	55	-121	92	21
	ID02822.5p-miR	67	-127	91	23
HMOX1; 16,1	ID01152.3p-miR	75	-113	95	20
IGF2; 7,4	miR-3141	266	-104	94	19
KRAS; 7,5	ID01310.3p-miR	17, 29	-121, -123	92, 94	22
	ID03332.3p-miR	37	-132	89	24
	ID00805.3p-miR	96	-113	91	21
MYC; 29,2	ID01662.3p-miR	14	-132	93	23
	ID02429.3p-miR	16	-121	89	23
	ID01881.5p-miR	22	-113	91	21
	ID01432.3p-miR	296	-110	90	22
ODC1; 24,3	ID02624.3p-miR	114	-119	92	21
	ID03156.5p-miR	177	-129	88	24
ONECUT2; 0,0	ID01774.5p-miR	151	-129	90	23
PPP2R1B; 3,7	ID01034.5p-miR	15	-121	92	21

Table C14 (continued)

1	2	3	4	5	6
PROM1; 6,0	miR-1292-3p	288	-110	93	20
PTEN; 10,9	ID02079.5p-miR	75	-115	92	20
	ID02611.3p-miR	486	-125	91	22
	ID03465.3p-miR	525	-119	92	22
	ID01310.3p-miR	531	-119	90	22
	ID02761.3p-miR	533	-132	89	24
	ID03037.3p-miR	536	-121	90	22
	ID01315.3p-miR	705	-115	92	20
	ID02429.3p-miR	790	-121	89	23
PTGS2; 9,2	ID03397.3p-miR	108	-123	92	21
SEPT9; 25,4	ID03387.3p-miR	361	-125	88	24
	ID00149.3p-miR	370	-113	90	22
	ID00529.5p-miR	370	-110	93	21
	miR-6833-3p	371	-104	94	21
	ID01352.3p-miR	694	-119	93	23
	ID03324.3p-miR	696	-115	90	22
SERPINE1; 24,5	ID01098.3p-miR	30	-123	88	24
SFRP2; 47,6	ID03385.3p-miR	87	-117	90	23
	ID00909.3p-miR	94	-119	89	23
SMAD2; 5,1	ID01492.3p-miR	16	-129	100	23
	ID01018.3p-miR	225	-123	88	24
SMAD4; 9,2	ID00577.3p-miR	160	-106	94	20
	ID00961.3p-miR	248	-127	90	23
SMAD7; 9,1	miR-6786-5p	128	-117	92	21
	ID03064.3p-miR	134	-138	90	24
SMOX; 2,6	ID01319.5p-miR	13	-106	93	20
	ID00443.5p-miR	40	-123	89	24
	ID00570.3p-miR	98	-113	90	22
SOCS2; 6,4	ID00850.3p-miR	445	-117	90	22
SOCS6; 5,6	miR-6081	60	-123	88	24
	ID01911.5p-miR	86	-123	89	23
SPP1; 16,8	miR-1913	61	-117	92	22
STAT3; 40,7	ID00570.3p-miR	28	-115	92	22
	miR-4665-3p	39	-136	85	26
TCF4; 5,4	miR-5196-5p	364	-115	92	22
	ID01734.5p-miR	371	-113	90	22
TCF7L2; 12,0	miR-6124	308	-102	92	20
TGFB1; 19,2	miR-6877-5p	5	-113	90	22
	ID03306.3p-miR	6	-121	92	21
	ID02770.5p-miR	6	-115	92	20
	ID00795.5p-miR	186	-117	92	22
	ID00457.3p-miR	209	-129	95	22
	ID02064.5p-miR	211	-132	91	23
	ID00529.5p-miR	232	-108	91	21
	miR-877-3p	233	-108	93	21
	miR-6824-5p	708	-113	90	22
TGFBR1; 10,9	ID01590.3p-miR	16	-117	92	22
	ID02174.5p-miR	38	-123	89	23

Table C14 (continued)

1	2	3	4	5	6
TYMS; 4,2	ID00252.5p-miR	87	-132	89	24
	ID01029.5p-miR	113	-115	90	22
	ID02106.3p-miR	114	-125	91	23
UBAP2L; 15,4	ID03324.3p-miR	45	-115	90	22
WIF1; 0,1	ID01355.3p-miR	237	-115	92	20
XRCC1; 10,1	ID00850.3p-miR	37	-117	90	22
ZEB1; 20,3	ID01777.3p-miR	1	-121	89	23
	ID00849.3p-miR	1	-117	90	22
	ID01545.3p-miR	8	-110	91	21

Table C15 - Characteristics of the binding sites of miRNA and mRNA in CDS of genes involved in the development of CRC.

Gene; NX	miRNA	Position, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
1	2	3	4	5	6
ABCG2; 2,8	ID02408.3p-miR	1672	-104	91	22
ADAMTS4; 7,9	ID03489.5p-miR	573	-123	88	24
	ID00265.3p-miR	909	-115	92	22
ADAMTS5; 1,9	ID01242.3p-miR	1097	-121	88	24
	ID00306.5p-miR	1178	-119	92	22
	miR-511-5p	2729	-102	92	21
AKAP4; 0,0	miR-7162-3p	302	-98	94	19
ALDH2; 124,7	ID03270.3p-miR	153	-108	91	21
	miR-1226-5p	156	-127	86	26
	miR-4687-3p	171	-110	91	21
	ID00223.5p-miR	1559	-119	89	23
ALOX15; 0,6	ID01385.5p-miR	258	-110	91	21
ARHGAP24; 1,6	ID01944.5p-miR	1165	-102	91	21
BCL2; 3,7	miR-1343-5p	606	-121	90	22
	ID02256.3p-miR	715	-125	91	22
BIRC5; 0,7	miR-5095	352	-106	91	21
	ID00913.5p-miR	378	-115	90	23
BMP4; 6,8	ID01699.5p-miR	906	-106	91	21
CASR; 0,0	ID01787.3p-miR	2421	-115	89	23
CCND1; 9,7	miR-4731-3p	508	-110	91	22
CDC27; 6,7	miR-3127-3p	372	-113	90	22
CDH1; 26,4	miR-7160-3p	186	-106	91	21
	ID01845.5p-miR	484	-119	90	22
	ID00254.3p-miR	573	-110	90	22
	ID00982.3p-miR	2454	-100	92	20
CDX1; 41,0	ID01491.3p-miR	189	-123	89	23
	ID00663.5p-miR	190	-125	89	23
	ID01181.5p-miR	193	-110	93	20
	ID00087.3p-miR	309	-115	92	22
COL3A1; 180,1	miR-6798-5p	651	-121	90	23

Table C15 (continued)

1	2	3	4	5	6
	miR-3909	3517	-113	93	22
	ID02852.5p-miR	3593	-117	92	23
CTNNB1; 45,4	miR-6083	1773	-98	92	20
CYP1A1; 1,0	miR-6894-5p	567	-119	89	24
CYP1B1; 2,4	ID01181.5p-miR	543	-113	95	20
	ID03408.3p-miR	816	-110	90	22
DCC; 0,0	ID00432.3p-miR	2971	-102	91	22
DCTN1; 29,9	miR-4633-5p	1498	-100	92	20
DNMT1; 7,3	ID00389.5p-miR	3797	-125	94	22
	ID03063.3p-miR	4833	-110	90	22
DNMT3B; 0,5	ID02759.5p-miR	788	-108	93	21
DPYD; 4,8	ID02288.5p-miR	1897	-115	89	23
EGFR; 7,0	ID02344.3p-miR	1779	-127	88	24
	miR-525-5p	1854	-102	91	21
ERBB2; 21,5	ID00692.3p-miR	3815	-113	90	22
	miR-4734	3817	-119	90	22
ESR1; 0,2	ID02606.5p-miR	409	-106	91	22
	ID02556.3p-miR	1852	-115	90	23
FASN; 19,5	miR-7161-3p	146	-127	86	28
	miR-3656	1345	-102	98	17
	ID01548.5p-miR	3054	-119	89	23
	ID03291.3p-miR	3428	-119	89	23
	ID02020.3p-miR	5146	-115	90	22
	ID01995.3p-miR	5324	-106	94	19
	ID01336.3p-miR	5991	-138	90	24
	miR-3940-5p	7090	-113	96	20
FMNL3; 3,3	ID00564.5p-miR	1626	-110	90	22
FN1; 181,7	miR-6756-5p	1156	-123	92	23
	ID02776.5p-miR	1392	-110	90	22
	miR-3926	2596	-104	92	21
GSTM1; 10,1	ID01955.3p-miR	462	-108	89	23
HMGB1; 37,4	miR-3653	264	-89	98	18
	ID03324.3p-miR	746	-115	90	22
	ID00777.3p-miR	748	-113	90	23
IRS1; 5,7	ID02344.3p-miR	412	-127	88	24
	ID03141.5p-miR	1416	-110	91	21
	ID02266.5p-miR	2084	-102	92	20
	ID00372.5p-miR	2085	-121	89	24
	miR-4655-3p	2380	-113	91	21
	ID01753.3p-miR	3446	-106	93	21
KISS1; 0,0	ID02978.3p-miR	560	-119	90	22
KSR1; 4,1	ID03489.3p-miR	472	-121	88	24
	ID02362.5p-miR	664	-108	93	20
MSH6; 4,2	ID00156.5p-miR	246	-134	90	24
	miR-4527	2816	-104	91	22
MTA1; 20,6	D01524.3p-miR	406	-113	91	22
	ID03238.3p-miR	421	-117	90	23

Table C15 (continued)

1	2	3	4	5	6
	ID01593.5p-miR	1003	-119	89	23
	ID01702.3p-miR	2271	-134	89	24
MUC1; 54,8	ID00645.5p-miR	507	-110	93	20
	miR-6752-5p	510	-119	90	22
MUC2; 65,5	miR-6840-3p	381	-117	93	22
	ID00910.3p-miR	8453	-110	90	22
MUTYH; 5,7	miR-331-5p	1212	-110	91	22
MYC; 29,2	miR-6761-5p	989	-104	91	21
NOTCH1; 14,1	ID01953.5p-miR	46	-123	91	22
	miR-1271-5p	1242	-108	91	22
	ID03029.3p-miR	2480	-108	91	21
	ID02396.5p-miR	4096	-113	91	22
	ID02151.5p-miR	4415	-119	93	23
	ID03324.3p-miR	4972	-115	90	22
	ID02890.3p-miR	5181	-119	89	23
	miR-486-3p	6091	-106	91	21
	ID01469.5p-miR	6766	-123	92	22
OGG1; 5,1	ID02210.5p-miR	443	-110	90	22
	ID03142.5p-miR	443	-104	92	20
	ID00920.5p-miR	882	-134	94	24
PIK3CA; 4,2	ID00724.5p-miR	3043	-100	92	21
PIK3CG; 0,3	miR-6815-5p	1022	-115	90	23
	miR-3129-3p	3077	-104	92	22
POLD1; 5,4	ID03196.3p-miR	103	-121	88	24
	ID03387.3p-miR	203	-129	91	24
	ID03114.3p-miR	1069	-113	91	21
	miR-629-3p	1073	-110	91	22
	ID03320.3p-miR	1607	-113	90	22
	ID03070.3p-miR	2588	-119	90	22
POLE; 4,6	ID00782.3p-miR	4899	-110	90	22
	ID02748.3p-miR	6116	-119	89	23
	miR-3192-3p	6150	-108	93	21
PPARD; 14,7	miR-4751	517	-117	89	24
SEPT9; 25,4	ID03428.5p-miR	1039	-117	90	22
	miR-8069	1350	-119	89	23
	miR-4507	1362	-108	93	20
SERPINE1; 24,5	miR-4758-3p	277	-119	90	23
	ID01786.5p-miR	542	-125	88	24
SFRP2; 47,6	miR-3677-5p	243	-115	92	22
	miR-4465	642	-110	93	22
SIRT1; 6,6	ID03324.3p-miR	435	-115	90	22
SMAD7; 9,1	ID00529.5p-miR	369	-113	95	21
	miR-6729-5p	722	-119	90	22
SMOX; 2,6	miR-5047	593	-102	91	21
	ID00211.3p-miR	1093	-113	90	22
	ID03343.3p-miR	1101	-110	90	22
	miR-3148	1240	-102	91	22

Table C15 (continued)

1	2	3	4	5	6
SOCS2; 6,4	ID02930.3p-miR	681	-132	89	24
SRM; 24,0	ID02434.5p-miR	110	-125	89	24
	ID01652.3p-miR	112	-125	89	23
	miR-6786-5p	114	-121	95	21
STAT3; 40,7	ID00857.5p-miR	1116	-119	90	23
TGFB1; 19,2	miR-6742-5p	2047	-110	90	22
TGFBR1; 10,9	ID02984.3p-miR	112	-123	92	22
	ID03332.3p-miR	127	-134	90	24
	ID01310.3p-miR	128	-121	92	22
	ID03332.3p-miR	130	-134	90	24
	ID01310.3p-miR	131	-121	92	22
	ID03332.3p-miR	133	-136	91	24
	miR-6887-3p	733	-104	91	21
THBS2; 12,4	miR-192-3p	1597	-108	93	22
	miR-598-3p	2942	-104	91	22
TLR2; 3,4	ID00935.5p-miR	295	-104	94	20
TNF; 0,6	ID02050.3p-miR	230	-121	92	23
TNFRSF4; 1,7	ID02542.5p-miR	51	-127	88	24
	miR-4494	384	-110	90	22
TSC1; 7,9	miR-324-3p	3302	-110	96	20
VEGFA; 25,0	ID00678.3p-miR	627	-113	91	21
	ID03097.3p-miR	888	-121	89	23
XRCC1; 10,1	miR-4763-3p	964	-125	88	24
ZEB1; 20,3	ID00777.3p-miR	3246	-117	93	23
	ID03324.3p-miR	3250	-115	90	22
	ID00777.3p-miR	3252	-113	90	23

Table C16 - Characteristics of the binding sites of miRNA and mRNA in 3'UTR of genes involved in the development of CRC.

Gene; NX	miRNA	Position, nt	ΔG, kJ/mol	ΔG/ΔGm, %	Length, nt
ADAMTS4;7,9	ID01334.3p-miR	4131	-113	90	22
AGMAT; 2,1	miR-1273g-3p	1617	-108	93	21
	ID01237.3p-miR	1988	-113	88	24
	miR-4452	2253	-102	89	23
	miR-5096	2281	-110	98	21
	miR-5585-3p	2349	-108	93	22
	miR-1285-5p	2448	-104	92	21
	ID02744.3p-miR	2509	-102	91	22
	miR-619-5p	2517	-119	98	22
	miR-5096	2580	-110	98	21
	miR-1273g-3p	2731	-113	96	21
	ID01404.5p-miR	2768	-110	90	23
	miR-1273e	2774	-113	96	22
ALOX15; 0,6	miR-1273g-3p	2394	-115	98	21

Table C16 (continued)

1	2	3	4	5	6
	miR-1273d	2431	-121	89	25
APC; 3,3	miR-4693-5p	9182	-108	94	23
ATF3; 17,3	miR-6746-5p	1690	-119	93	22
BAG3; 37,2	miR-6515-3p	2353	-100	92	20
BCL2; 3,7	ID03441.3p-miR	3613	-110	90	22
BCL2L1; 20,5	miR-6842-5p	1693	-115	93	22
	ID02328.5p-miR	2459	-123	92	23
CASR; 0,0	ID01309.3p-miR	4776	-113	90	23
CCND1; 9,7	ID01220.5p-miR	1159	-117	90	24
	miR-6865-3p	1350	-106	91	21
	ID01543.3p-miR	2963	-117	92	22
	ID00529.5p-miR	2972	-113	95	21
	ID00108.3p-miR	3401	-125	89	24
CDC27; 6,7	ID02587.3p-miR	4448	-102	91	22
CDH1; 26,4	miR-1273c	3251	-110	91	22
	miR-1273g-3p	3271	-108	93	21
	miR-1273h-5p	3305	-113	96	21
	ID01838.5p-miR	3413	-110	88	24
CEACAM3; 0,2	ID00128.3p-miR	797	-115	93	22
CYP1B1; 2,4	ID02086.5p-miR	2334	-104	91	23
DCTN5; 10,1	ID00750.3p-miR	2706	-110	91	22
	ID01692.5p-miR	3069	-104	91	22
	miR-619-5p	6471	-110	91	22
	ID01836.5p-miR	6696	-115	92	23
	miR-1285-5p	6712	-106	94	21
	miR-1303	6722	-106	91	22
	miR-5585-3p	7172	-106	91	22
ERBB3; 17,5	miR-619-5p	4950	-117	96	22
	miR-619-5p	5104	-121	100	22
ESR1; 0,2	ID03196.3p-miR	3339	-121	88	24
	miR-6879-5p	3593	-113	90	22
FASN; 19,5	ID00978.5p-miR	7689	-119	90	22
	ID00611.5p-miR	7810	-117	90	23
FMNL3; 3,3	miR-4436b-5p	4403	-115	92	22
	ID01746.3p-miR	6295	-106	91	21
	ID01754.3p-miR	6572	-121	88	24
	miR-6788-5p	8531	-102	91	21
GREM1; 29,7	ID03324.3p-miR	2759	-115	90	22
	ID02426.5p-miR	3153	-91	90	22
HMOX1; 16,1	miR-3155a	1228	-106	91	21
IGF2; 7.4	ID00470.5p-miR	2837	-110	91	23
	ID01151.3p-miR	3500	-121	88	24
	miR-5002-5p	3716	-106	88	24
	ID03383.3p-miR	4292	-115	90	23
IGF2BP3; 0,0	ID03447.3p-miR	4088	-104	91	22
IRS1; 5,7	ID00436.3p-miR	7402÷7425 (3)	-104	89	23
KRAS; 7,5	ID03224.5p-miR	3163	-121	92	23

Table C16 (continued)

1	2	3	4	5	6
	miR-1273g-3p	3176	-108	93	21
	miR-1273f	3209	-104	100	19
	miR-7110-5p	4965	-110	93	21
MACC1; 0,3	miR-619-5p	4856	-117	96	22
	ID00913.5p-miR	4876	-115	90	23
	ID03437.5p-miR	4958	-96	92	22
	miR-5585-3p	4997	-113	96	22
MDM2; 4,7	miR-1273g-3p	2117	-113	96	21
	ID01404.5p-miR	2154	-110	90	23
	ID01815.5p-miR	2257	-106	89	23
	ID01360.3p-miR	2476	-104	91	21
	miR-1273g-3p	2486	-106	91	21
	miR-1273e	2521	-108	93	22
	miR-3929	3013	-115	93	23
	ID01838.5p-miR	3214	-110	88	24
	miR-1285-3p	3218	-106	91	22
	miR-1273g-3p	6739	-113	96	21
MTA1; 20,6	ID00791.5p-miR	2360	-127	88	24
	miR-6829-5p	2383	-104	92	20
	ID00899.5p-miR	2736	-115	92	21
MTAP; 3,2	miR-5708	2277	-102	92	20
	miR-5585-3p	2431	-115	98	22
	ID01836.5p-miR	2530	-117	93	23
	miR-1285-5p	2546	-106	94	21
	ID02799.5p-miR	3532	-117	89	23
MTHFR; 4,5	ID03245.5p-miR	3350	-117	89	24
	miR-8089	3460	-121	88	24
	ID00345.3p-miR	6281	-119	89	23
	miR-5585-3p	6300	-108	93	22
	ID03407.3p-miR	6342	-108	91	22
	miR-1285-5p	6399	-104	92	21
	ID01811.5p-miR	6844	-117	93	22
	miR-5095	6855	-110	95	21
	miR-619-5p	6861	-115	95	22
	miR-5585-3p	7003	-110	95	22
	ID02175.3p-miR	7051	-113	93	22
NOTCH1; 14,0	ID02808.3p-miR	8813	-125	91	23
ONECUT2; 0,0	miR-548f-5p	3190	-96	90	22
	miR-548ah-5p	3465	-93	94	20
	miR-548aj-3p	4398	-91	93	21
	ID02759.5p-miR	12801	-108	93	21
PIK3CG; 0,3	miR-1285-5p	5164	-108	96	21
	miR-5585-3p	5226	-106	91	22
	ID02566.3p-miR	5262	-113	91	22
	miR-1303	5335	-106	91	22
PPARD; 14,7	miR-4632-5p	2738	-119	89	23
PPP2R1B; 3,7	miR-7850-5p	2198	-102	91	21

Table C16 (continued)

1	2	3	4	5	6
	miR-4452	3026	-102	89	23
	miR-5585-3p	3124	-115	98	22
PTP4A3; 29,5	miR-4539	926	-113	90	22
SEPT9; 25,4	ID00222.3p-miR	3244	-121	90	22
	ID03308.3p-miR	3244	-119	92	21
	ID00414.3p-miR	4238	-108	93	20
SMAD2; 5,1	miR-1273a	6065	-117	89	25
	ID01838.5p-miR	6066	-113	90	24
	ID01404.5p-miR	6124	-110	90	23
SMAD4; 9,2	ID01838.5p-miR	4291	-113	90	24
	miR-1273g-3p	4312	-110	95	21
	ID02732.3p-miR	7721	-123	91	23
	ID00106.5p-miR	7825	-106	91	22
	miR-574-5p	7830	-108	89	23
	ID01592.3p-miR	8227	-117	89	23
SMOX; 2,6	miR-423-5p	2280	-113	90	23
SRM; 24,0	ID03281.3p-miR	1177	-100	94	20
STAT3; 40,7	miR-5095	3125	-106	91	21
	miR-619-5p	3131	-119	98	22
	ID00345.3p-miR	3249	-119	89	23
	miR-5585-3p	3268	-110	95	22
	ID01836.5p-miR	3359	-117	93	23
	ID01774.5p-miR	3782	-129	90	23
TGFBR1; 10,9	miR-938	1901	-106	91	22
TP53; 9,3	miR-1273h-5p	2351	-106	91	21
	ID01838.5p-miR	2459	-115	92	24
	ID00785.5p-miR	2520	-113	90	23
UBAP2L; 15,4	ID01054.5p-miR	3531	-119	90	22
VEGFA; 25,0	miR-1277-5p	2085	-96	88	24
	miR-1277-5p	3275	-96	88	24
	miR-1277-5p	3314	-96	88	24