I.A. Novikova, N.N. Timoshkina, D.S.Kutilin

DIFFERENTIAL EXPRESSION OF MICRO-RNA IN TUMOR AND NORMAL COLON TISSUES

DOI 10.25789/YMJ.2020.72.19

Non-coding RNAs (miRNAs and long non-coding RNAs (lncRNA)) play an important role in many biological processes, and dysregulation can lead to various diseases, including colorectal cancer (CRC). The aim of the study was to analyze the differential expression of miRNAs in the tumor and normal tissues of patients with CRC, as well as the identification of potential lncRNA targets and target genes using methods of machine learning. Analysis of miRNA expression was performed by multiple parallel sequencing on a MiSeq instrument. For bioinformation analysis, DESeq2, TarPmiR, ORA (Over-Representation Analysis) and FMD (Functional module detection) algorithms were used. Sequencing revealed 6 differentially expressed microRNAs (hsa-miR-143-3p, hsa-miR-26a-5p, hsa-miR-25-3p, hsa-miR-92a-3p, hsa-miR-21-5p, hsa -let-7i-5p) in the tumor tissue of the colon is relatively normal. For these microRNAs, 97 target genes and 23 potentially interacting long non-coding RNAs were identified. Together, they form a network of competitively expressed RNA characteristic of CRC, which is involved in the implementation of signaling cascades such as regulation of cell adhesion, activation of the immune response, regulation of the cellular response to hormones and stress, Wnt signaling pathway and cell migration regulation, regulation of proliferation and cell cycle, regulation interaction with viral agents, regulation of apoptosis and response to hypoxia. The data obtained expand the understanding of the mechanisms of gene expression regulation in CRC and can become the basis for a panel of tumor markers.

Keywords: colorectal cancer, microRNA, non-coding RNA, expression, metastases, multiple parallel sequencing.

Introduction. Currently, there is a stable increase in the incidence of colorectal cancer (CRC): worldwide, the number of annually registered cases is approaching 1.4 million [6]. CRC is a heterogeneous multifactorial disease, which in about 35% of cases is caused by genetic factors, including changes in the expression of noncoding RNAs (ncRNAs) [25].

The transformation of normal colon mucosa into adenocarcinoma occurs in several stages that take decades [25]. Although numerous therapies such as surgery, chemotherapy, radiation therapy, targeted therapy, and immunotherapy have been shown to reduce relapse rates and improve patient survival, 5-year survival in CRC patients is still low [24]. This emphasizes the need to search for new biomarkers that can be used for early diagnosis and prognosis of the course of this disease. A large number of studies have revealed that ncRNAs (miRNAs and long noncoding RNAs (lncRNAs)) play an important role in many biological processes, and their dysregulation can lead to various diseases, including colon cancer [29].

NOVIKOVA Inna Arnoldovna – Ph.D, deputy general director for science, National Medical Research Center of Oncology, Rostov-on-Don, Russia. e-mail: novikovainna@yahoo.com, **TIMOSHKINA Natalia N**. – Ph.D., head of the laboratory of Molecular Oncology, National Medical Research Center of Oncology, Rostov-on-Don, Russia. email: timoshkinann@rnioi.ru, **KUTILIN Denis Sergeevich** – Ph.D., Senior Researcher Laboratory of Molecular Oncology, National Medical Research Center of Oncology, Rostov-on-Don, Russia, e-mail: k.denees@yandex.ru, ORCID 0000-0002-8942-3733

MicroRNAs are short non-coding RNAs that regulate gene expression by catalyzing the destruction of mRNA or by inhibiting the translation of mRNA into protein [4]. MicroRNAs make a significant contribution to the initiation and development of various molecular events, including the initiation of oncogenesis, progression and metastasis of tumors, which makes microRNAs potential biomarkers for assessing the progression and prognosis of CRC. Although miRNAs regulate the expression of genes encoding proteins, mainly through degradation or silencing of mRNA, there is growing evidence that miRNAs can interact with lncRNA, which, in turn, also regulates the expression of target genes [1].

The study of the regulatory network of microRNA-long non-coding RNA-mRNA is of great importance both for elucidating the molecular mechanisms underlying carcinogenesis and for creating a panel of new biomarkers. Therefore, the aim of the study was to analyze the differential expression of microRNA in tumor and normal tissues of CRC patients, as well as to identify potential lncRNA targets and target genes using machine learning methods.

Materials and methods.

Patients. For multiple parallel microRNA sequencing, 80 patients (adenocarcinoma of the colon G2) aged 45-69 (median age 65) were selected, from which paired tissue biopsies (tumor and normal, 160 samples in total) were obtained.

Multiple parallel microRNA sequencing. The mirVana miRNA Isolation Kit (Ambion, Life Science Technologies, USA) was used to isolate the miRNA fraction. The creation and subsequent purification of the miRNA transcriptome library was performed using the TruSeq Small RNASample Preparation Kit (Illumina, United States). The 3'- and 5'-adapters were sequentially sewn to the miRNA fraction, reverse transcription and amplification of the created miRNA construct were carried out according to the attached protocol. The cDNA copies were purified by electrophoresis in 6% PAGE. cDNA from the gel was extracted with water and precipitated with 95% ethanol (glycogen was used as a co-precipitant). The amount of cDNA was determined on a Qubit 2.0® fluorometer using the Qubit dsDNA HS Assay Kit (Invitrogen, USA). Multiple parallel sequencing of the nucleotide sequences of the cDNA libraries was performed on a MiSeq instrument (Illumina, United States). The method of parallel multiple sequencing identified the nucleotide sequence in a random sample of generated cDNA copies. The expression of miRNA was determined by comparing the nucleotide sequence of the sequenced molecules in each sample with the known nucleotide sequences of miRNA presented in the miRBase and mirGeneDB databases. The approach is based on the algorithm implemented in miRanalyzer [14].

Statistical / bioinformatic analysis.

DESeq2 algorithm. When analyzing differential expression of miRNAs, the DESeq2 algorithm was used, implemented in the R. DESeq2 environment. DESeq2 provides a method for checking differential expression using negative binomial generalized linear models; variance estimates and changes in logarithmic fold. The analysis included microRNAs represented in the samples by

at least 10 copies. To take into account the effect of integral expression of miRNAs on the expression of compared miRNAs, the compared libraries were normalized using the RLE (Relative Log Expression) method. The Wald test was used to assess the statistical significance of the differences. To take into account the multiplicity of comparisons, we used the method of assessing the probability of false-positive results (false discovery rate (FDR, Benjamini-Hochberg)) [14].

Search for target genes. The search for target genes was carried out using the TarPmiR algorithm using the TargetScan, mirDB, and miRTarBase databases. TarPmiR uses a random forest approach to predict the target microRNA site. Random forest (from English - "random forest") is a machine learning algorithm that uses an ensemble of decision trees and combines the Breiman bagging method and the random subspace method. The result of the random forest model is the predicted probability that the candidate target site is the true target site. TarPmiR integrates 13 traditional functions for predicting microRNA target sites, including the availability of binding sites and thermodynamic properties such as free energy [5].

Over-Representation Analysis (ORA). This method measures the percentage of genes (or microRNAs) in the signaling pathway that are differentially expressed. The goal of ORA is to obtain a list of the most important signaling pathways, ordered according to p-value [2].

Clustering of target genes by function (Algorithm FMD (Functional module detection)). The algorithm is based on the k-nearest neighbors (KNN) algorithm and Louvain's community search algorithm to cluster closely related genes into separate functional modules. Representative processes and pathways enriched in each cluster are presented along with the cluster and their resulting Q value. The Q value of each member of the functional module is calculated using Fisher's one-sided exact test and Benjamin-Hochberg correction to correct multiple comparisons [13].

Analysis of microRNA-lncRNA interaction. HITS-CLIP, PAR-CLIP and CLASH data were obtained from the Gene Expression Omnibus database and pre-processed using FASTX-Toolkit v0.0.13 and analyzed using PARalyzer v1.1. All binding site coordinates were converted to assemblies hg19, mm9 / mm10 and ce6 / ce10, respectively, using the UCSC LiftOver Tool. The genomic coordinates of the conserved miRNA target sites predicted by TargetScan, miRanda / mirSVR, PITA, Pictar, and RNA22 were also assembled and transformed into hg19, mm9 / mm10, and ce6 / ce10 assemblies using LiftOver. The obtained coordinates were compared with the previously described CLIP clusters using BEDTools [21].

Results and discussion. As a result of multiple parallel sequencing of 160 samples, the following categories of RNA were found (average percentage is shown): microRNA (0.5%), mRNA (0.03%), antisense mRNA (0.01%), tRNA (86.7%), transcripts of unclear significance (0.03%)), other RNAs (12.7%).

The DEseq2 algorithm was used to analyze the differential expression of the detected miRNAs in tumor and normal tissue of the colon of 80 patients (Fig. 1). Six differentially expressed miRNAs were found: the expression of two, hsa-miR-143-3p and hsa-miR-26a-5p, was reduced, and four, hsa-miR-25-3p, hsa-miR-92a-3p, hsa-miR -21-5p and hsa-let-7i-5p are elevated in tumor tissue compared to normal.

The highest expression level in normal tissue was found for hsa-miR-143-3p, the lowest for hsa-miR-92a-3p; in tumor tissue, these were, respectively, hsa-miR-10b-5p and hsa-miR-25-3p (Table . one). The largest difference in miRNA expression in tumor tissue relative to normal was recorded for hsa-miR-92a-3p (7.2 times, p = 0.02), the smallest one for hsa-miR-26a-5p (3.6 times, p = 0.01) (Table 1). 1, fig. 1).

target genes presented in Table 2 made it possible to identify 3 functional clusters, including 65 genes (Fig. 2, Table 3)

From the data presented in Table 3 and Figure 2, it can be seen that the main signaling pathways in which target genes are involved are (Q <0.05): regulation of intracellular signal transmission, regulation of cell adhesion, protein dephosphorylation, activation of the immune response, regulation of the organization of the plasma membrane, regulation of cell growth, regulation of cellular response to hormones, cellular response to stress, and regulation of the Wnt signaling pathway.

An alternative FMD method for assessing the significance of differentially expressed microRNAs is overrepresentation in signaling pathways (ORA). The data obtained during the implementation of the ORA algorithm are presented in Table 4.

The main signaling cascades involving microRNAs hsa-miR-143-3p, hsa-miR-26a-5p, hsa-miR-25-3p, hsa-miR-92a-3p, hsa-miR-21-5p and hsa- let-7i-5p include (p <0.05): regulation of proliferation and the cell cycle, immune response and response to inflammation, interaction with viral agents, regulation of apoptosis, response to hypoxia, and regulation of cell migration.

Analysis of the possibility of interaction of 6 differentially expressed miRNAs with lncRNA revealed the following miRNA-lncRNA pairs (Table 5).

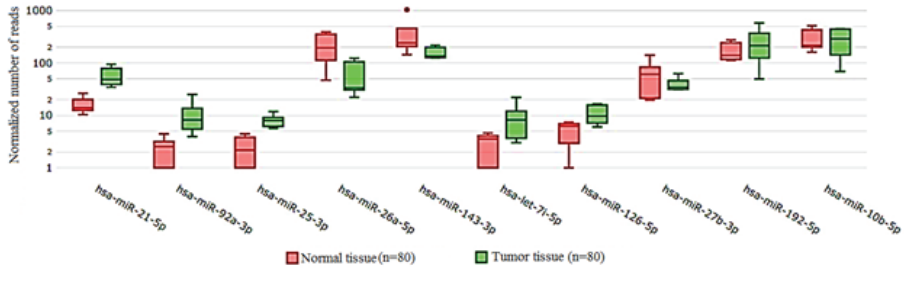

Figure 1. Expression of 10 miRNAs with the lowest FDR (6 significant miRNAs).

For microRNAs differentially expressed in tumors and in normal tissue, a search for target genes was carried out using the TarPmiR algorithm.

For 6 miRNAs that statistically significantly alter the expression in colon tumors, 5,360 target genes were predicted, of which were validated (confirmed) in the TargetScan, mirDB and miRTarBase databases of miRNA-mRNA interactions for 116 target genes, including 97 targets with minimal free energy of microRNA-mRNA interaction (Table 2).

Application of the FMD (Functional module detection) algorithm to the list of

As can be seen from the data presented in Table 5, five lncRNAs have binding sites common to several microRNAs. For example, lncRNA MALAT1 can interact with 5 out of 6 microRNAs (hsa-miR-143-3p, hsa-miR-26a-5p, hsa-miR-25-3p, hsa-miR-92a-3p, hsa-miR- 21-5p).

Significant progress in molecular biology in the last 10 years has revealed a link between microRNA and CRC. Stable dysregulation of miRNA expression in CRC leads to the stimulation of oncogenes or tumor suppressors [6].

Different studies of microRNA sequencing in CRC have identified dif-

Table1

Expression of microRNA in normal and tumor tissues *

microRNA	Normal tissue	Tumor tissue	FC	$\log_2$FC	*p*-value	P_{adj}
hsa-let-7i-5p*	1.83	8.20	4.481	2.164	0.02	0.04
hsa-miR-10a-5p	36.91	37.00	1.002	0.004	0.99	0.99
hsa-miR-10b-5p	301.14	283.17	0.940	– 0.089	0.87	0.96
hsa-miR-126-5p	4.02	10.09	2.510	1.328	0.03	0.05
hsa-miR-143-3p*	382.87	159.01	0.415	– 1.268	0.02	0.04
hsa-miR-192-5p	174.70	258.4	1.479	0.565	0.34	0.42
hsa-miR-21-5p*	15.41	57.40	3.725	1.897	0.0	0.0
hsa-miR-25-3p*	1.44	7.05	4.896	2.292	0.01	0.02
hsa-miR-26a-5p*	220.06	61.45	0.279	– 1.840	0.01	0.02
hsa-miR-27b-3p	60.77	39.16	0.644	– 0.634	0.3	0.41
hsa-miR-92a-3p*	1.35	9.65	7.148	2.838	0.0	0.02

*Note. Expression Minimum Threshold = 1; FC (fold change) - change rate; Padj is the p-value adjusted for multiple comparisons using the Benjamini-Hochberg method.

ferent specific expression profiles that may be associated with clinical or prognostic features. So Hamfjord et al. [9] identified 19 hypo-expressed and 18 hyper-expressed miRNAs in CRC. Later Schee et al. [23] identified the 5 most expressed miRNAs (hsa-miR-10a-5p, hsa-miR-21-5p, hsa-miR-22-3p, hsa-miR-143-3p, and hsa-miR-192-5p) in 88 CRC samples that the number of identified microRNAs is close to our study.

In our study, the expression of 2 microRNAs (miR-143-3p and miR-26a-5p) was decreased, and 4 (miR-25-3p, miR-92a-3p, miR-21-5p, and let-7i-5p) in tumor tissue compared to normal. The results obtained are confirmed in the works of other authors. Thus, miR-143-3p is part of a cluster associated with oncogenes, DNA repair genes, and genes that regulate the WNT and MAPK signaling pathways. Tokarz P and Blasiak J. [26] showed that hsa-miR-143-3p in CRC cells reduces the expression of DNA methyltransferases 3A (DNMT3A) and changes the phenotype of malignant transformation, and hsa-miR-143-3p is negative a prognostic factor for survival in patients with CRC. [20].

Hernández R. et al. showed that the overexpression of miR-21 leads to tumor formation, increasing invasion and metastasis [10]. Previously, Liu et al. [16] after analyzing the concentrations of miR-21 and miR-92a in the serum of CRC patients, they found higher levels of these miRNAs compared to healthy subjects. In CRC, the expression of miR-92a is significantly increased in tissues, which leads to a decrease in E-cadherin and increases the level of β-catenin and vimentin, which are involved in the regulation of the epithelial-mesenchymal transition [31]. In addition, Nishida et al. reported an association between high levels of miR-92a and lymph node and liver metastases [19]. The above results indicate that high expression of miR-92a in CRC patients is associated with poor survival, and miR-92a may be a potential biomarker for CRC.

In colon tumors, let-7 suppresses KRAS and its expression is associated with better patient survival [22].

Table2

Target genes of 6 differentially expressed microRNAs

Overexpressed microRNA						Hypoexpressed microRNA	
hsa-92a-3p		hsa-let-7i-5p		hsa-miR-25-3p	hsa-miR-21-5p	hsa-miR-26a-5p	hsa-miR-143-3p
Гены-мишени							
SPOCK2	*KLF2*	*ZNF644*	*PDP2*	*GALNT7*	*ST6GAL1*	*CHAC1*	*ZBTB44*
MEF2D	*OTUD3*	*TGFBR3*	*NR6A1*	*PER2*	*KLHL42*	*CDK8*	*MAPK7*
SNN	*CPEB4*	*SMC1A*	*PLAGL2*	*SRPRA*	*PAN3*	*TET2*	*SECISBP2L*
ATXN1	*GNAQ*	*POTEM*	*ARID3B*	*SLX4*	*GID4*	*PDE4B*	
FAR1	*PHTF2*	*AGO1*	*SMCR8*	*TMEM184B*		*ZNF608*	
GLYR1	*TSC1*	*ONECUT2*	*MBD2*	*LHFPL2*		*AGPAT5*	
PTPRJ	*G3BP2*	*BACH1*		*CPEB4*		*ESR1*	
DAB2IP	*YIPF4*	*CCNT2*		*EIF4G2*		*PIM1*	
RRBP1	*ELOVL6*	*CBX5*		*PKDCC*			
EIF4G2	*GOLGA8J*	*GRPEL2*		*C5orf24*			
PLEKHA1	*UBE2Z*	*ADIPOR2*		*FNIP1*			
SPRYD4	*PPP1R37*	*RBFOX2*		*PHLPP2*			
EVI5	*USP28*	*PDE12*		*SERTAD3*			
PER2	*KCNC4*	*KREMEN1*					
NOL4L	*FAM20C*	*CRY2*					
DNAJB12	*SOCS5*	*SURF4*					
TRAM2	*SLX4*	*IGDCC4*					
ACTC1	*MAN2A1*	*CEP135*					
CREB3L2	*ATXN7*	*SYT1*					
SOX4	*GRAMD1B*	*USP47*					
TECPR2	*PCMTD1*	*SEMA4C*					

Table3

Functional clusters and signaling pathways target-genes of microRNA differentiall

Cluster	Function / signal pathway	Q value*	Number of genes	Гены
M1	downregulation of the protein kinase B signaling pathway	0.00056	3	*PLEKHA1,PTPRJ,OTUD3*
	protein dephosphorylation	0.00096	4	*TSC1,PDP2,PTPRJ,CRY2*
	negative regulation of intracellular signaling	0.00110	5	*PTPRJ,PLEKHA1,TSC1, ESR1, OTUD3*
	positive regulation of focal adhesion	0.00253	2	*TSC1,PTPRJ*
	positive regulation of the assembly of cellular compounds	0.00325	2	*TSC1,PTPRJ*
	positive regulation of biogenesis of cellular com-ponents	0.00422	4	*TSC1,ESR1,PTPR-J,CEP135*
	regulation of phosphoprotein phosphatase activity	0.00658	2	*CRY2,TSC1*
	cellular response to dsRNA	0.00914	2	*PDE12,ESR1*
	cell surface receptor signaling pathway that acti-vates the immune response	0.01281	2	*PLEKHA1,PTPRJ*
	regulation of the organization of the plasma mem-brane	0.02375	2	*SYT1,CEP135*
	protein stabilization	0.02442	2	*TSC1,OTUD3*
	transmission of a signal that activates the immune response	0.02984	2	*PLEKHA1,PTPRJ*
	regulation of cell growth	0.03313	2	*SYT1,PTPRJ*
	activation of the immune response	0.03807	2	*PLEKHA1,PTPRJ*
	cellular response to hormones	0.05441	2	*TSC1,ESR1*
M2	negative regulation of the cell cycle	0.02375	2	*SOX4,SMC1A*
	organization of the endomembrane system	0.03258	2	*TRAM2,SURF4*
	growth regulation	0.03429	2	*SERTAD3,PIM1*
	cellular response to organic cyclic compound	0.04654	2	*PIM1,RBFOX2*
	drug metabolism	0.04814	2	*ST6GAL1,KLF2*
M3	positive regulation of cell growth	0.00914	2	*EIF4G2,USP47*
	regulation of the canonical Wnt signaling pathway	0.02083	2	*GNAQ,USP47*
	positive regulation of GTPase activity	0.02233	2	*GNAQ,EVI5*
	lipid biosynthesis process	0.03125	2	*FAR1,ELOVL6*

*Q value - P-value adjusted using Benjamini-Hochberg correction.

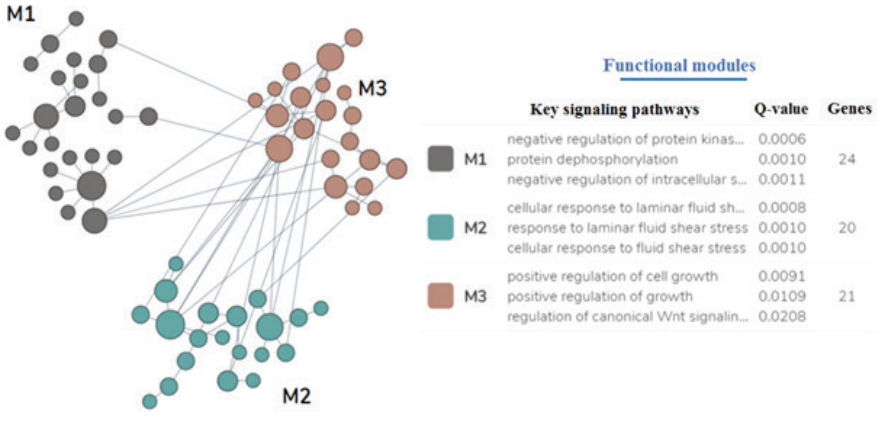

Figure 2. Functional clusters of target genes of 5 differentially expressed microRNAs in KPP

Thus, the investigated microRNAs can be divided into oncogenic and oncosuppressive ones. Those. oncogenic microRNAs mainly target and inhibit the expression of suppressor genes. The activation of these miRNAs has a significant effect on the progression of CRC (for example, hsa-miR-21 and hsa-miR-92a). However, a number of miRNAs (oncosuppressive miRNAs) play an important role in slowing tumor progression by suppressing oncogenes associated with proliferation, apoptosis, invasion, and migration (for example, hsa-miR-143-3p and hsa-let-7i-5p).

Let us consider in more detail one of the functional clusters (M3) identified by us and the effect of 5 microRNAs on it (Figure 3).

Figure 3 shows that the overexpression of miR-92a-3p can lead to suppression of the expression of the FAR1, EVI5, MAN2A1, PCMTD1, GNAQ, PHTF2, and ELOVL6 genes. The ELOVL6 gene (Fatty Acid Elongase 6) encodes an enzyme that catalyzes the first and rate-limiting reaction of the fatty acid elongation cycle and is involved in the synthesis of fatty acids that are essential in numerous biological processes as precursors of membrane lipids and lipid mediators. The FAR1 gene encodes a protein necessary for the reduction of fatty acids to

Table4

Results of analysis of overrepresentation in signaling pathways (ORA)

Signaling pathways	p-value	microRNAs
positive regulation of the activity of the tran-scription factor nf kappab	0.017	miR-143; miR-26a; miR-25; miR-92a; miR-21; let-7i
reaction to ethanol	0.017	miR-143; miR-26a; miR-25; miR-92a; miR-21; let-7i
inflammatory response	0.019	miR-143; miR-26a; miR-25; miR-92a; miR-21; let-7i
cell migration	0.020	miR-143; miR-26a; miR-25; miR-92a; miR-21; let-7i
regulation of cell proliferation	0.021	miR-143; miR-26a; miR-25; miR-92a; miR-21; let-7i
phosphatidylinositol-mediated signaling	0.023	miR-143; miR-26a; miR-25; miR-92a; miR-21; let-7i
cellular response to hypoxia	0.026	miR-143; miR-26a; miR-25; miR-92a; miR-21; let-7i
cytokine-mediated signaling pathway	0.027	miR-143; miR-26a; miR-25; miR-92a; miR-21; let-7i
ubiquitination of proteins	0.028	miR-143; miR-26a; miR-25; miR-92a; miR-21; let-7i
epidermal growth factor receptor signaling pathway	0.030	miR-143; miR-26a; miR-25; miR-92a; miR-21; let-7i
regulation of the apoptotic process	0.030	miR-143; miR-26a; miR-25; miR-92a; miR-21; let-7i
immune response	0.030	miR-143; miR-26a; miR-25; miR-92a; miR-21; let-7i
cell cycle control	0.031	miR-143; miR-26a; miR-25; miR-92a; miR-21; let-7i
interaction with the virus	0.040	miR-143; miR-26a; miR-25; miR-92a; miR-21; let-7i

Table5

MicroRNA-lncRNA interactions

microRNA	lncRNA	Genomic coordinates	Alignment
1	2	3	4
hsa-let-7i-5p	XIST	chrX:73046461-73046481	LncRNA: 5' agaACCCUAAAC-CCUACCUCu 3' | | :|||| |||||| miRNA : 3' uugUCGUGUUUGAUGAUGGAGu 5'
	NEAT1	chr11:65205185-65205207	LncRNA: 5' cugguCUUGGUCUGUCUACCUCg 3' | ::: ||: |||||| miRNA : 3' uugucGUGUUUGAU-GAUGGAGu 5'
	HELLPAR	chr12:102625721-102625741	LncRNA: 5' uuCAGAUCUCAUU-CUACCUCc 3' ||| | |:| |||||| miRNA : 3' uuGUCGUGUUUGAUGAUGGAGu 5'
	AP000766.1	chr11:107185038-107185058	LncRNA: 5' uuCACUCUUAAAU-CUACCUCa 3' || : : || | |||||| miRNA : 3' uuGUCGUGUUUGAUGAUGGAGu 5'
	LINC00265	chr7:39831428-39831448	LncRNA: 5' uuaGGCCCAUAC-CCUACCUCa 3' :|| || || |||||| miRNA : 3' uugUCGUGUUUGAUGAUGGAGu 5'
	AC124045.1	chr3:44710371-44710393	LncRNA: 5' auUAGCCUGUAAGAU-CUACCUCa 3' :||| ::||: | |||||| miRNA : 3' uuGUCG--UGUUUGAUGAUGGAGu 5'
	MIRLET7BHG	chr22:46509673-46509694	LncRNA: 5' gggAGGGCCGCCCCCUACCUCa 3' || :| : | |||||| miRNA : 3' uugUCGUGUUUGAUGAUGGAGu 5'
	OLMALINC	chr10:102133372-102133384	LncRNA: 5' ---------cuCUCCUACCUCa 3' || |||||| miRNA : 3' uugucguguuuGAUGAUGGAGu 5'
hsa-miR-143-3p	HELLPAR	chr12:102714596-102714616	LncRNA: 5' agggcaaaccuaUUCAUCUCa 3' ||||||| miRNA : 3' cucgaugucacgAAGUAGAGu 5'
	MALAT1	chr11:65269208-65269229	LncRNA: 5' ugGUUCAUAUUCAGUCAUCUCa 3' |:| |:| | |||||| miRNA : 3' cuCGA-UGUCACGAAGUAGAGu 5'

Continuation of table. 5

1	2	3	4
	AC245884.8	chr19:54950318-54950338	LncRNA: 5' ucauguCUGUAAGUCAUCUCc 3' | || |||||| miRNA : 3' cucgauGUCACGAAGUAGAGu 5'
	MEOX2-AS1	chr7:15734898-15734919	LncRNA: 5' gcGUUGUGAAUCAGUCAUCUCa 3' |:|:: | | |||||| miRNA : 3' cuCGAUG-UCACGAAGUAGAGu 5'
	MEG3	chr14:101315825-101315845	LncRNA: 5' augggAUUGUCAUUCAUCUCa 3' |: || ||||||| miRNA : 3' cucgaUGUCACGAAGUAGAGu 5'
hsa-miR-26a-5p	TUG1	chr22:31371700-31371719	LncRNA: 5' augCUCUACUGG--UACUUGAa 3' || | |||| |||||| miRNA : 3' ucgGAUAGGACCUAAUGAACUu 5
	NORAD	chr20:34637129-34637150	LncRNA: 5' acaaauuaacuccUUACUUGAa 3' |||||||| miRNA : 3' ucggauaggaccuAAUGAACUu 5'
	HCG11	chr6:26523678-26523694	LncRNA: 5' -----uuaacuccUUACUUGAa 3' |||||||| miRNA : 3' ucggauaggaccuAAUGAACUu 5'
	MALAT1	chr11:65268969-65268990	LncRNA: 5' aaCUUGUUAUUUUUUACUUGAa 3' |:|:|: | ||||||| miRNA : 3' ucGGAUAGGACCUAAUGAACUu 5
	WASIR2	chr16:75102-75123	LncRNA: 5' uuauaccauaaaAUUACUUGAa 3' |||||||| miRNA : 3' ucggauaggaccUAAUGAACUu 5
	AC005332.6	chr17:66132006-66132023	LncRNA: 5' acaUUAUCU----UUACUUGAg 3' :||||: ||||||| miRNA : 3' ucgGAUAGGACCUAAUGAACUu 5'
	AC005332.7	chr17:66125163-66125184	LncRNA: 5' uucuggauaaaacAUUACUUGAg 3' |||||||| miRNA : 3' ucggauaggaccUAAUGAACUu 5'
	EBLN3P	chr9:37087347-37087368	LncRNA: 5' uccagggaaagagcUACUUGAc 3' |||||| miRNA : 3' ucggauaggaccuaAUGAACUu 5'
hsa-miR-25-3p	NORAD	chr20:34636526-34636547	LncRNA: 5' ccauuguuAUGUGUGUGCAAUu 3' | ::: |||||| miRNA : 3' agucuggcUCUGUUCACGUUAc 5'
	XIST	chrX:73063643-73063665	LncRNA: 5' uuGGACUGUUAAUAUGUGCAAUu 3' :|||:| |:| |||||| miRNA : 3' agUCUGGC-UCUGUUCACGUUAc 5'
	MALAT1	chr11:65268090-65268110	LncRNA: 5' caGGAAGGAG-CGAGUGCAAUu 3' :|| ||| |:||||||| miRNA : 3' agUCUGGCUCUGUUCACGUUAc 5'
	PURPL	chr5:27494380-27494405	LncRNA: 5' gaAGACCUGAUUUAUCAUGUGCAAUa 3' ||||| || | || |||||| miRNA : 3' agUCUGG-CU--CU-GUUCACGUUAc 5'
hsa-miR-92a-3p	OIP5-AS1	chr15:41592688-41592709	LncRNA: 5' uaaaaCCGGGAUAUGUGCAAUa 3' |||||:| |||||| miRNA : 3' uguccGGCCCUGUUCACGUUAu 5'
	NORAD	chr20:34636526-34636547	LncRNA: 5' ccauuguuauguguGUGCAAUu 3' |||||| miRNA : 3' uguccggcccuguuCACGUUAu 5'
	XIST	chrX:73063643-73063664	LncRNA: 5' uggacuguuaAUAUGUGCAAUu 3' |:| |||||| miRNA : 3' uguccggcccUGUUCACGUUAu 5'
	MALAT1	chr11:65268089-65268110	LncRNA: 5' ccAGGAAGGAGCGAGUGCAAUu 3' ||| || :|:||||||| miRNA : 3' ugUCCGGCCCUGUUCACGUUAu 5'

The end of the table. 5

1	2	3	4
	SNHG14	chr15:25364949-25364970	LncRNA: 5' agcaccaaauuucuGUGCAAUg 3' |||||| miRNA : 3' uguccggcccuguuCACGUUAu 5'
hsa-miR-21-5p	TUG1	chr22:31372028-31372048	LncRNA: 5' gauAAAUGAG-CUAAUAAGCUu 3' | || | | |||||| miRNA : 3' aguUGUAGUCAGACUAUUCGAu 5'
	XIST	chrX:73043927-73043947	LncRNA: 5' cugcCACCCAUAU-AUAAGCUa 3' || | | |||||| miRNA : 3' aguuGUAGUCAGACUAUUCGAu 5'
	MALAT1	chr11:65266342-65266365	LncRNA: 5' uuuGCAUUCAAGUUCCAUAAGCUg 3' :|||: |||: |||||| miRNA : 3' aguUGUAG--UCAGACUAUUCGAu 5'
	FAM66E	chr8:7841811-7841833	LncRNA: 5' acuugAUUGUGGUGGAUAAGCUu 3' ||:: | : ||||||| miRNA : 3' aguugUAGU-CAGACUAUUCGAu 5'

fatty alcohols, which is necessary for lipid synthesis. Suppression of expression of the FAR1 and ELOVL6 genes can lead to changes in the lipid composition of cell membranes and disruption of the normal presentation of cellular antigens, which is important for masking tumor cells from the immune system [8]. The PHTF2 (Putative Homeodomain Transcription Factor 2) and EVI5 (Ecotropic Viral Integration Site 5) genes function as regulators of the cell cycle and transcription of a wide range of genes associated with tumor development (according to https://www.genecards.org). GNAQ (G Protein Subunit Alpha Q) is involved in the modulation of various transmembrane signaling cascades, regulates the chemotaxis of dendritic cells, and is also involved in the regulation (agonist) of the canonical Wnt signaling pathway.

At the same time, miR-26a-5p hypoexpression can lead to an increase in the expression of the CDK8, TET2, and AGPAT5 genes (Figure 3). TET2 encodes a protein that catalyzes the conversion of the modified DNA base methylcytosine to 5-hydroxymethylcytosine, i.e. plays a key role in active demethylation of DNA, which leads to changes in the epigenetic regulation of gene expression during cell differentiation [11]. AGPAT5 encodes an integral membrane protein that converts lysophosphatidic acid to phosphatidic acid, which is a step 2 of de novo phospholipid biosynthesis. This can also contribute to a change in the representation of tumor antigens on the cell surface. The protein encoded by the CDK8 gene is a member of the cyclin-dependent protein kinase family. CDK8 can have an activating or inhibiting effect on the function of transcription factors [18]. CDK8 can act as an oncogene in colorectal cancer (transcription of the CDK8 gene in tumors of this type is significantly increased), caused by over-activation of the Wnt / β-catenin signaling pathway. [7].

Thus, changes in the expression of these microRNAs in CRC can promote epithelial-mesenchymal transition and masking of tumor cells from the immune system, which is also confirmed in other studies [10].

However, it is obvious that data on microRNA-mRNA interactions are insufficient for a complete understanding of the mechanisms of carcinogenesis. In recent years, data have been obtained on complex regulatory interactions (by competitive binding) between lncRNA, microRNA, and the target genome in CRC patients [29, 3, 16, 25].

In our study, 6 lncRNAs interact with more than one differentially expressed microRNA. One of these lncRNAs, MALAT1, has binding sites for miR-143-3p, miR-26a-5p, miR-25-3p, miR-92a-3p, and miR-21-5p. MALAT1 (metastasis associated lung adenocarcinoma transcript 1) is a long spliced non-coding RNA that is involved in the epigenetic modulation of gene expression (in particular, those associated with cell migration) and is closely associated with the development of cancer. [thirty]. Long noncoding RNA XIST (X-inactive specific transcript) has binding sites for 4 microRNAs (let-7i-5p, miR-25-3p, miR-92a-3p, and miR-21-5p). Xist is a key effector in the inactivation of the X chromosome, thereby providing dose equality (ie, the number of active variants of one gene) [12]. Long non-coding RNA NORAD (Non-Coding RNA Activated By DNA Damage) has binding sites for 3 microRNAs (miR-26a-5p, miR-25-3p, and miR-92a-3p) and is associated with cancers such as pancreatic cancer and bladder. Its overexpression reduces the invasive ability of tumor cells, while the rate of apoptosis increases significantly [15]. Long noncoding RNA HELLP, potentially interacting with let-7i-5p and miR-143-3p, activates a large set of genes that are involved in the cell cycle and regulate the invasive ability of cells [27]. Long noncoding RNA TUG1 (taurine upregulated gene 1), potentially interacting with miR-26a-5p and miR-21-5p, plays an import-

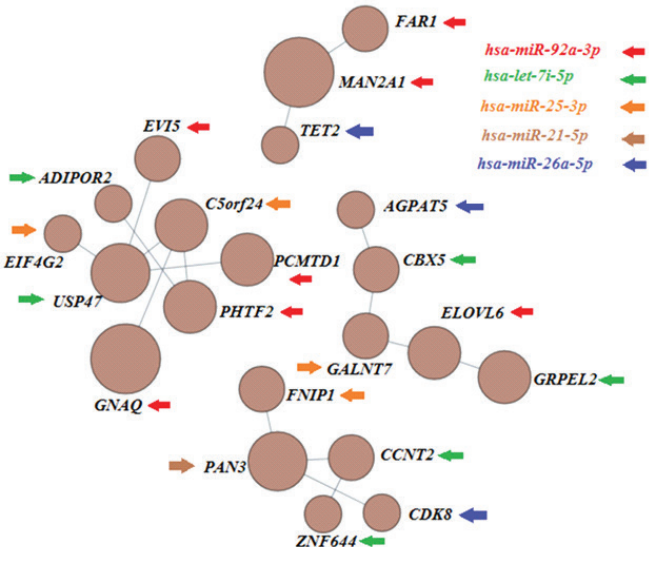

Figure 3. Functional cluster M3 consisting of 21 target genes and 5 microRNAs differentially expressed in CRC

ant role in the epigenetic regulation of transcription, promotes proliferation, and is activated in tumor cells [25].

Currently, the role of the eRNA network (a network of competitively expressed RNAs) consisting of microRNAs (miR-143-3p, miR-26a-5p, miR-25-3p, miR-92a-3p, miR-21-5p and let- 7i-5p), long non-coding RNAs (listed in Table 5), and messenger RNAs encoded by 97 target genes (listed in Table 2), the development of CRC remains completely unclear. And the data obtained in the course of bioinformatic analysis can become the basis for further studies of the complex regulatory network of mRNA-lncRNA-miRNA for a better understanding of the mechanisms of carcinogenesis and identification of new CRC biomarkers.

Conclusion. Thus, as a result of multiple parallel sequencing, 6 differentially expressed microRNAs (hsa-miR-143-3p and hsa-miR-26a-5p, hsa-miR-25-3p, hsa-miR-92a-3p, hsa- miR-21-5p, hsa-let-7i-5p) in the tumor tissue of the colon is relatively normal. For these micro-RNAs, 97 target genes were identified, belonging to 3 functional clusters, and 23 long non-coding RNAs potentially interacting with them. Together they constitute a network of competitively expressed RNAs (keRNA network) characteristic of CRC, which is involved in the implementation of signaling cascades such as the regulation of cell adhesion, activation of the immune response, regulation of the cellular response to hormones and stress, regulation of the Wnt signaling pathway and cell migration, regulation of proliferation and cell cycle, regulation of interaction with viral agents, regulation of apoptosis and response to hypoxia. The data obtained broaden the understanding of the mechanisms of gene regulation in the conditions of this oncopathology and can become the basis for a panel of highly specific tumor markers.

References

1. Abdelsattar ZM, Wong SL, Regenbogen SE, Jomaa DM, Hardiman KM, Hendren S. Colorectal cancer outcomes and treatment patterns in patients too young for average-risk screening. Cancer. 2016; 122: 929-934.

2. Backes C, Khaleeq QT, Meese E, Keller A. miEAA: microRNA enrichment analysis and annotation. Nucleic Acids Res. 2016;44(W1):W110-6.

3. Cui C., D. Zhai, L. Cai, Q. Duan, L. Xie, J. Yu. Long noncoding RNA HEIH promotes colorectal cancer tumorigenesis via counteracting mir-939mediated transcriptional repression of Bcl-Xl. Cancer Res. Treat., 50 (2018), pp. 992-1008

4. Dimitriadi TA, Burtsev DV, Dzhenkova EA, Kutilin DS. MicroRNA as markers of pre-cancer diseases progression in cervical cancer. Modern problems of science and education. 2020; 1. URL: http://science-education.ru/ru/article/view?id=29529 (in Russ)

5. Ding J, Li X, Hu H. TarPmiR: a new approach for microRNA target site prediction. Bioinformatics. 2016;32(18):2768-75.

6. Ding L, Lan Z, Xiong X, Ao H, Feng Y, Gu H, Yu M, Cui Q. The Dual Role of MicroRNAs in Colorectal Cancer Progression. Int. J. Mol. Sci. 2018; 19(9):2791

7. Firestein R, Bass AJ, Kim SY, Dunn IF, Silver SJ, Guney I, Freed E, Ligon AH, Vena N, Ogino S, Chheda MG, Tamayo P, Finn S, Shrestha Y, Boehm JS, Jain S, Bojarski E, Mermel C, Barretina J, Chan JA, Baselga J, Tabernero J, Root DE, Fuchs CS, Loda M, Shivdasani RA, Meyerson M, Hahn WC (Sep 2008). "CDK8 is a colorectal cancer oncogene that regulates beta-catenin activity". Nature. 455 (7212): 547–51.

8. Groza N.V., Golovanov A.B., Nalivayko Ye.A., Myagkova G.I. The therapeutic role of polyunsaturated fatty acids and their derivatives in pathophysiological processes. Bulletin of MITHT, 2012, vol. 7, No. 5. 3-16

9. Hamfjord J, Stangeland AM, Hughes T, et al. Differential expression of miRNAs in colorectal cancer: comparison of paired tumor tissue and adjacent normal mucosa using high-throughput sequencing. PLoS One. 2012;7:e34150. doi: 10.1371/journal.pone.0034150

10. Hernández R, Sánchez-Jiménez E, Melguizo C, Prados J, Rama AR. Downregulated microRNAs in the colorectal cancer: diagnostic and therapeutic perspectives. BMB Rep. 2018 Nov;51(11):563-571. doi: 10.5483/BMBRep.2018.51.11.116.

11. Ito S., D'Alessio A.C., Taranova O.V., Hong K., Sowers L.C., Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification (англ.) // Nature : journal. — 2010. — August (vol. 466, no. 7310). — P. 1129—1133

12. Kalantry S., Purushothaman S., Bowen R. B., Starmer J., Magnuson T. Evidence of Xist RNA-independent initiation of mouse imprinted X-chromosome inactivation. (англ.) // Nature. - 2009. - Vol. 460, no. 7255. - P. 647-651.

13. Krishnan A, Zhang R, Yao V, Theesfeld CL, Wong AK, Tadych A, Volfovsky N, Packer A, Lash A, Troyanskaya OG. Genome-wide prediction and functional characterization of the genetic basis of autism spectrum disorder. Nature Neuroscience.2016;19(11):1454-1462

14. Kutilin DS. Regulation of Gene Expression of Cancer/Testis Antigens in Colorectal Cancer Patients. Molecular biology. 2020.Vol. 54.No. 4.P. 580-595.

15. Lei Y., Wang Y.H., Wang X.F., Bai J. LINC00657 promotes the development of colon cancer by activating PI3K/AKT pathway. Eur Rev Med Pharmacol Sci. 2018 Oct;22(19):6315-6323. doi: 10.26355/eurrev_201810_16042.

16. Li Y., C. Zeng, J. Hu, Y. Pan, Y. Shan, B. Liu, L. Jia Long non-coding RNA-SNHG7 acts as a target of miR-34a to increase GALNT7 level and regulate PI3K/Akt/mTOR pathway in colorectal cancer progression J. Hematol. Oncol., 11 (2018), p. 89

17. Liu GH, Zhou ZG, Chen R, et al. Serum miR-21 and miR-92a as biomarkers in the diagnosis and prognosis of colorectal cancer. Tumour Biol. 2013;34:2175–2181

18. Nemet J, Jelicic B, Rubelj I, Sopta M (Feb 2014). "The two faces of Cdk8, a positive/negative regulator of transcription". Biochimie. 97: 22–7.

19. Nishida N., Nagahara M., Sato T., Mimori K., Sudo T., Tanaka F., Shibata K., Ishii H., Sugihara K., Doki Y., et al. Microarray analysis of colorectal cancer stromal tissue reveals upregulation of two oncogenic miRNA clusters. Clin. Cancer Res. 2012;18:3054–3070

20. Qian X, Yu J, Yin Y, et al. MicroRNA-143 inhibits tumor growth and angiogenesis and sensitizes chemosensitivity to oxaliplatin in colorectal cancers. Cell Cycle. 2013;12:1385–1394

21. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features, Bioinformatics. 2010; 26: 841-842.

22. Saridaki Z., Weidhaas J.B., Lenz H.J., Laurent-Puig P., Jacobs B., De Schutter J., De Roock W., Salzman D.W., Zhang W., Yang D., et al. A let-7 microRNA-binding site polymorphism in KRAS predicts improved outcome in patients with metastatic colorectal cancer treated with salvage cetuximab/panitumumab monotherapy. Clin. Cancer Res. 2014;20:4499–4510

23. Schee K, Lorenz S, Worren MM, et al. Deep sequencing the microRNA transcriptome in colorectal cancer. PLoS One. 2013;8:e66165

24. Siegel RL, Miller KD, Jemal A. Cancer statistics. CA A Cancer J. Clin. 2019; 69: 7-34.

25. Tang XJ, Wang W, Hann SS. Interactions among lncRNAs, miRNAs and mRNA in colorectal cancer. Biochimie. 2019; 163:58-72.

26. Tokarz P, Blasiak J. The role of microRNA in metastatic colorectal cancer and its significance in cancer prognosis and treatment. Acta Biochim Pol. 2012;59:467–474

27. van Dijk M, Thulluru HK, Mulders J, Michel OJ, Poutsma A, Windhorst S, Kleiverda G, Sie D, Lachmeijer AM, Oudejans CB. HELLP babies link a novel lincRNA to the trophoblast cell cycle. J Clin Invest. 2012 Nov;122(11):4003-11. doi: 10.1172/JCI65171. Epub 2012 Oct 24. PMID: 23093777; PMCID: PMC3484460.

28. Wang C., D.Z. Hu, J.Z. Liu, Identification of critical TF-miRNA-mRNA regulation loops for colorectal cancer metastasis Genet. Mol. Res., 14 (2015), pp. 5485-5495

29. Weng M, Wu D, Yang C, Peng H, Wang G, Wang T, Li X. Noncoding RNAs in the development, diagnosis, and prognosis of colorectal cancer. Transl. Res. 2017; 181:108-120.

30. Yoshimoto R, Mayeda A, Yoshida M, Nakagawa S (January 2016). "MALAT1 long non-coding RNA in cancer". Biochimica et Biophysica Acta. 1859 (1): 192–199.

31. Zhang G., Zhou H., Xiao H., Liu Z., Tian H., Zhou T. MicroRNA-92a functions as an oncogene in colorectal cancer by targeting PTEN. Dig. Dis. Sci. 2014;59:98–107