

Research Paper

Exploration of small RNA-seq data for small non-coding RNAs in Human Colorectal Cancer

Srinivas V Koduru^{1✉}, Amit K Tiwari², Sprague W Hazard³, Milind Mahajan⁴, Dino J Ravnic¹

1. Division of Plastic Surgery, Department of Surgery, College of Medicine, Pennsylvania State University, Hershey, PA, USA;
2. Department of Pharmacology & Experimental Therapeutics, College of Pharmacy & Pharmaceutical Sciences, University of Toledo - Health Sciences Campus, Toledo, OH, USA;
3. Department of Anesthesia, College of Medicine, Pennsylvania State University, Hershey, PA, USA;
4. Genomics Facility, Department of Genetics and Genomics Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA.

✉ Corresponding author: Srinivas V Koduru, PhD., Division of Plastic Surgery, Department of Surgery, College of Medicine, Pennsylvania State University, 500 University Drive, Hershey, PA 17033 Email: skoduru@hmc.psu.edu Phone: 717-531-4332

© Ivyspring International Publisher. This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY-NC) license (<https://creativecommons.org/licenses/by-nc/4.0/>). See <http://ivyspring.com/terms> for full terms and conditions.

Published: 2017.02.28

Abstract

Background: Improved healthcare and recent breakthroughs in technology have substantially reduced cancer mortality rates worldwide. Recent advancements in next-generation sequencing (NGS) have allowed genomic analysis of the human transcriptome. Now, using NGS we can further look into small non-coding regions of RNAs (sncRNAs) such as microRNAs (miRNAs), Piwi-interacting-RNAs (piRNAs), long non-coding RNAs (lncRNAs), and small nuclear/nucleolar RNAs (sn/snoRNAs) among others. Recent studies looking at sncRNAs indicate their role in important biological processes such as cancer progression and predict their role as biomarkers for disease diagnosis, prognosis, and therapy. **Results:** In the present study, we data mined publicly available small RNA sequencing data from colorectal tissue samples of eight matched patients (benign, tumor, and metastasis) and remapped the data for various small RNA annotations. We identified aberrant expression of 13 miRNAs in tumor and metastasis specimens [tumor vs benign group (19 miRNAs) and metastasis vs benign group (38 miRNAs)] of which five were upregulated, and eight were downregulated, during disease progression. Pathway analysis of aberrantly expressed miRNAs showed that the majority of miRNAs involved in colon cancer were also involved in other cancers. Analysis of piRNAs revealed six to be over-expressed in the tumor vs benign cohort and 24 in the metastasis vs benign group. Only two piRNAs were shared between the two cohorts. Examining other types of small RNAs [sn/snoRNAs, mt_rRNA, miscRNA, nonsense mediated decay (NMD), and rRNAs] identified 15 sncRNAs in the tumor vs benign group and 104 in the metastasis vs benign group, with only four others being commonly expressed. **Conclusion:** In summary, our comprehensive analysis on publicly available small RNA-seq data identified multiple differentially expressed sncRNAs during colorectal cancer progression at different stages compared to normal colon tissue. We speculate that deciphering and validating the roles of sncRNAs may prove useful in colorectal cancer prognosis, diagnosis, and therapy.

Key words: Bioinformatics, colorectal cancer, next-generation genomic sequencing, small non-coding RNA, sncRNA, miRNA, snoRNA.

Introduction

Colorectal cancer (CRC) is the third-most common cancer diagnosed in the United States in both men and woman, and the American Cancer Society estimates 134,490 new cases and 49,190 deaths

in 2016 [1]. Despite recent improvements in cancer care and awareness, incidences are still steadily growing. Although colonoscopy has allowed for earlier detection, a better understanding of tumor

biology and prognosis are still needed. Recently, advances in RNA sequencing have increased the understanding of the molecular mechanisms behind CRC. Changes in whole transcriptome profiles seem to underlie specific phenotypes. Small non-coding RNAs consist of 17-120 nucleotides, in their mature form [2] and play crucial roles in the development and progression of cancer. Innovative technologies have allowed researchers to focus on the role of small RNAs in disease progression which may contribute to the development of novel therapeutic targets. An increasing number of published studies have focused on miRNAs (17-22 nucleotides) [3], piRNAs (26-33 nucleotides) [4], small nucleolar RNAs (70-120 nucleotides) [5], and lncRNAs (>200 nucleotides) [6]. Using whole genome sequencing, present studies aim to understand the functions and transcriptional regulation of miRNA genes. Piwi-interacting RNAs (piRNAs) are one of the largest classes of the small non-coding RNAs family, and have been implicated in DNA methylation and post transcriptional-mediated repression of transposable elements [7]. However, other hypothesized functions of piRNAs remain largely unexplored [8]. Similarly, limited functional information is available for long non-coding RNAs (lncRNAs) [9]. Likewise, small nuclear RNAs (snRNAs) in eukaryotes consist of small nucleolar RNAs (snoRNAs), which carry out the fundamental role of ribosomal (rRNA) and transfer RNA (tRNA) modification and processing [10]. Two classes of snoRNAs, C/D and box H/ACA snoRNAs, have been shown to differ in sequence and structure [10].

In the present study, we carried out in-depth data analysis to identify the transcriptomic expression “signatures” of small RNAs from benign, tumor, and metastatic tissue samples from colorectal cancer patients, using existing datasets. We thoroughly analyzed the most predominant ncRNA expression in each sample type along with signaling pathways for possible biomarker development and identification of therapeutic targets in colorectal cancer.

Materials and Methods

Samples and data assembly

Colorectal cancer (CRC) small RNA sequencing (eight matched patients; 8 benign, 8 tumors, and 8 metastasis) datasets (project no: PRJNA201245; GEO: GSE46622) were downloaded from the NIH bioproject [11]. Detailed clinical information of patients was previously mentioned by Rohr et al [11]. Raw files were downloaded as *sequence raw archive* (SRA) files that were converted to FASTQ using SRA tool Kit version 2.5.7. Data was assembled using

PartekFlow® software, version 5.0 (Partek, Inc., St. Louis, MO, USA). Converted FASTQ files were uploaded to the PartekFlow® server and remapped to human genome hg19 (UCSC Genome Browser). Transcript abundances were determined and expression values were represented using reads per kilobases of exon model per million mapped reads (RPKM). All small RNA expression RPKM values >1 and representing at least 10% of the samples were considered robustly expressed and used for further analysis. Expression matrices were aligned to clinical pathologic features in order to compare miRNA, piRNA, lncRNA, and snRNA levels for association with specific CRC phenotypes. Statistical analysis between benign, tumor, and metastasis groups were carried out using the non-parametric Mann-Whitney U test followed by false discovery rate (FDR) correction through the Benjamin-Hochberg method, with the default *p* value <0.05 considered statistically significant [12].

Assembly of miRNA, piRNA, lncRNA and snoRNA annotations

Small RNA sequencing data was trimmed and aligned to the whole human genome (hg19), and BWA-0.7.12 aligner (BWA-MEM) with a few modifications (mismatch penalty 2, gap open penalty 6, clipping penalty 4 and alignment score cutoff 15) for short read mapping [13]. miRNAs were annotated from miRBase version 21 (<http://www.mirbase.org/>), which contains more than 1900 high confidence miRNAs [14]. piRNA data was generated and annotated from piRBase (regulatoryrna.org/database/piRNA), which is manually curated with a focus on piRNA functional analysis [15]. lncRNAs were quantified using reference annotation LNCipedia (<http://www.lncipedia.org>) version 3.1, downloaded from all coordinates relative to the hg19 reference genome [16]. Total small RNA (including miRNA, piRNA, snRNA, mtRNA, snoRNA, piRNA, tRF3, tRF5, tRNA, rRNA) was annotated following download of GenCode version 21 (www.Gencodegenes.org), which provides comprehensive, up-to-date information on human small non-coding RNAs, with specific regard to small nuclear and nucleolar RNAs.

Biological processes and gene network visualization: MetaCore

To investigate biological pathway interactions of small RNA expression, we used MetaCore pathway analysis of differentially expressed genes (Thomson Reuters, New York, NY) [13, 17], to delineate functional gene networks. This was based on differentially regulated gene lists with *p*<0.05 from

each group (tumor vs benign, metastasis vs benign, metastasis vs tumor) as input (Data analyzed by Gene Arrays, New York, USA) to generate disease biomarkers and GO terms.

Statistical Analysis

All data sets were subjected to paired student's *t* test with "*p*" value <0.05 being considered statistically significant. Furthermore, the Benjamini and Hochberg multiple testing adjustment method was applied to all small RNA sequencing studies and pathway analysis.

Results and Discussion

Analysis of miRNA sequencing of colorectal cancer tissue samples

To identify miRNA differential expression in colorectal cancer patient-matched tissue samples, data was downloaded from the GEO series accession number GSE46622. Using PartekFlow software miRNA analysis, we identified 19 miRNAs in the tumor vs benign cohort, and 38 miRNAs in the metastasis vs benign cohort, that were statistically significant. Top upregulated/downregulated miRNAs were determined in both tumor, metastasis stages (Figures 1A and 1B, Table 1), and hierarchical clustering/Venn diagram was compared for each group (Figure 1C). Our findings were very similar to

those originally published by Rohr et al [11], in which one of the top downregulated microRNAs was miR-1. It is downregulated during disease progression from tumor (4-fold) to metastasis (8-fold, Figure 1B), along with miR-145, miR-378a, miR-143 and miR-497 commonly regulated in both tumor and metastasis samples. Additionally, miR-31, miR-183, miR-135b, miR224 and miR-584 were upregulated with disease progression (Figure 1A, 1B and 1C). All of the miRNAs with fold change were listed in Table 1. miR-1 downregulation is consistent with other studies that showed its downregulation in colorectal cancers [18]. Several other miRNAs were expressed at low levels in normal tissue, strongly upregulated in tumors, and continued to be further upregulated in late stage (metastatic) CRC. miR-31 (9.3-fold), miR-183 (7.5-fold), miR-135b (6-fold) and miR-224 (3.6-fold), were all significantly expressed in metastasis in the sequencing data, in agreement with other studies demonstrating the same expression pattern in CRC and other disease states [19-21]. Specifically, miR-31 expression in benign samples was negligible and increased significantly with disease progression, this data is similar to previous reports showing increased presence of miR-31 in several cancer types with metastasis [22].

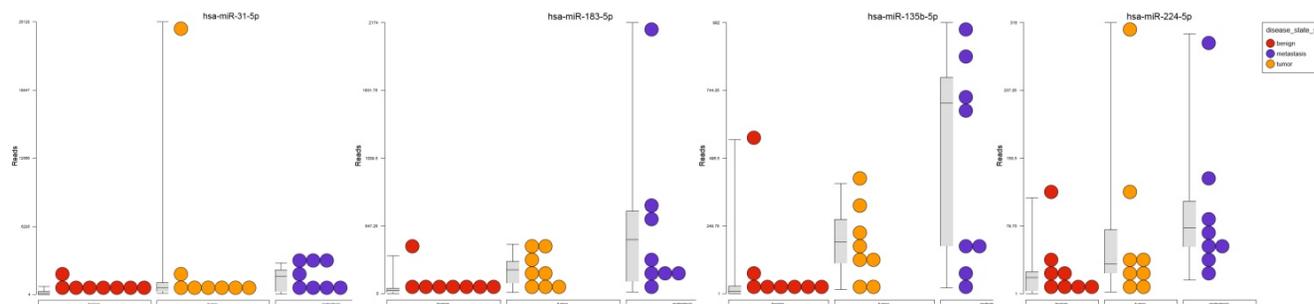
Table 1. Differential expression of miRNAs in three groups.

Metastasis vs benign		
microRNA ID	Fold Change	P-value
hsa-miR-122-5p	368.52	1.61E-02
hsa-miR-31-5p	9.28	5.99E-03
hsa-miR-183-5p	7.48	7.54E-03
hsa-miR-182-5p	6.76	1.51E-02
hsa-miR-135b-5p	5.80	6.93E-05
hsa-miR-7-5p	5.76	2.94E-02
hsa-miR-183-3p	5.58	1.48E-02
hsa-miR-96-5p	3.95	6.57E-03
hsa-miR-184	3.64	4.51E-02
hsa-miR-224-5p	3.61	1.51E-03
hsa-miR-629-5p	3.05	4.54E-02
hsa-miR-584-5p	2.80	1.69E-02
hsa-miR-21-3p	2.71	6.40E-03
hsa-miR-23a-5p	1.94	4.72E-02
hsa-miR-21-5p	1.91	3.47E-02
hsa-miR-3065-5p	-2.23	1.82E-02
hsa-miR-143-3p	-2.42	2.50E-02
hsa-miR-451a	-2.45	2.61E-02
hsa-miR-215-3p	-2.47	3.74E-02
hsa-miR-378a-5p	-2.58	1.38E-02
hsa-miR-192-5p	-2.59	1.13E-03
hsa-miR-192-3p	-2.80	2.13E-02
hsa-miR-29c-3p	-2.86	3.06E-02
hsa-miR-145-5p	-3.22	9.24E-03
hsa-miR-196b-5p	-3.33	2.05E-02
hsa-miR-144-5p	-3.34	5.73E-03
hsa-miR-378g	-3.37	2.55E-02
hsa-miR-10b-5p	-3.52	1.84E-02
hsa-miR-497-5p	-3.66	3.29E-03
hsa-miR-215-5p	-3.84	1.74E-02
hsa-miR-378c	-4.32	8.12E-03
hsa-miR-143-5p	-4.40	3.20E-02
hsa-miR-378a-3p	-4.43	9.38E-04
hsa-miR-145-3p	-4.47	4.76E-03
hsa-miR-195-5p	-5.33	6.18E-03
hsa-miR-378d	-7.32	1.43E-03
hsa-miR-144-3p	-7.76	2.76E-03
hsa-miR-1	-7.90	4.46E-04

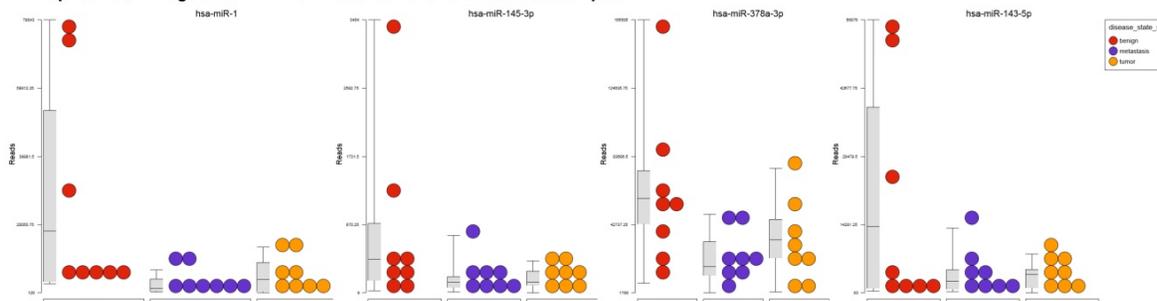
Tumor vs benign		
microRNA ID	Fold Change	P-value
hsa-miR-31-5p	21.41	9.82E-04
hsa-miR-584-5p	5.13	4.89E-02
hsa-miR-183-5p	3.45	4.91E-02
hsa-miR-135b-5p	3.07	4.92E-04
hsa-miR-224-5p	2.76	1.99E-02
hsa-miR-3117-3p	2.29	4.86E-02
hsa-miR-30a-5p	-1.47	3.77E-02
hsa-miR-29c-5p	-1.49	4.59E-02
hsa-miR-192-5p	-1.76	2.03E-02
hsa-miR-143-3p	-1.93	3.85E-02
hsa-miR-485-5p	-2.16	2.59E-02
hsa-miR-3065-5p	-2.24	1.17E-02
hsa-miR-375	-2.33	3.98E-02
hsa-miR-378a-3p	-2.54	3.64E-02
hsa-miR-147b	-2.89	2.14E-02
hsa-miR-497-5p	-2.90	1.96E-02
hsa-miR-145-3p	-3.31	2.16E-02
hsa-miR-143-5p	-3.63	3.92E-02
hsa-miR-1	-3.93	2.80E-03

Metastasis vs tumor		
microRNA ID	Fold Change	P-value
hsa-miR-204-5p	29.53	1.06E-02
hsa-miR-99b-3p	2.68	3.92E-02
hsa-miR-27b-5p	-1.67	4.43E-02
hsa-miR-23c	-6.02	4.34E-02

A. Top four upregulated miRNAs in both tumor and metastasis samples



B. Top four down regulated miRNAs in both tumor and metastasis samples



C. Differentially expressed miRNAs in colorectal cancer samples (p<0.05)

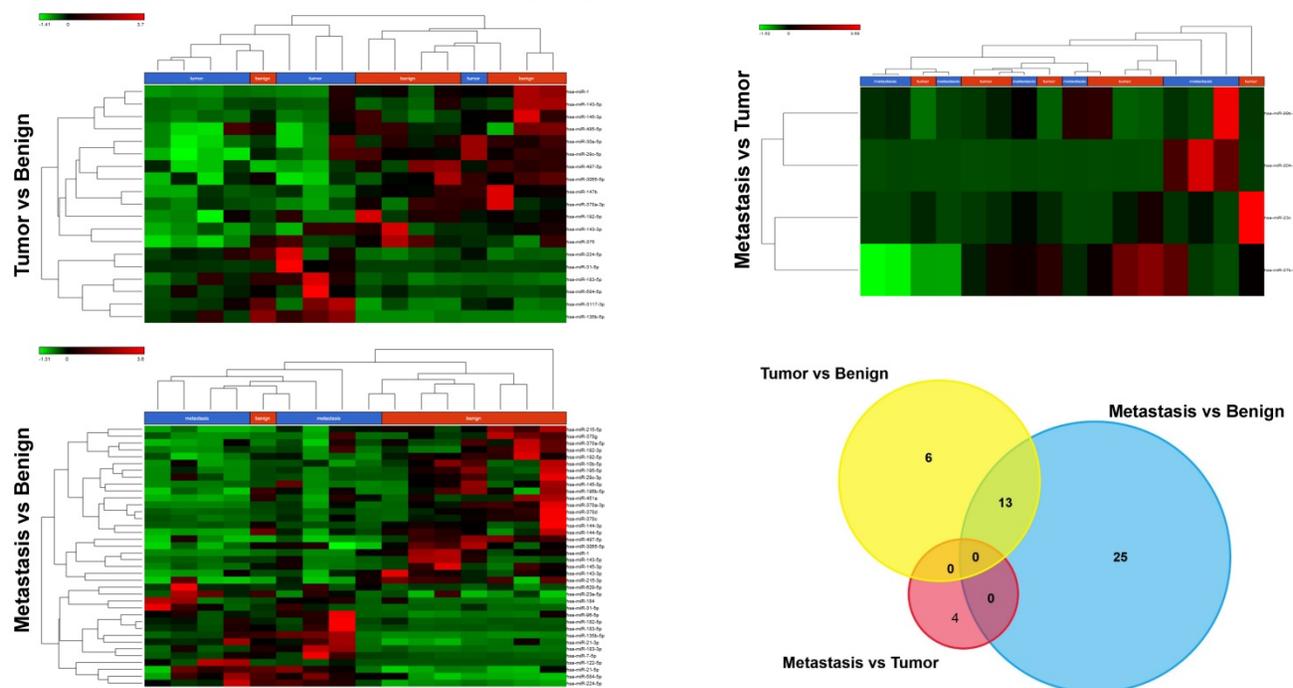


Figure 1. Differential expression of miRNAs in tumor and metastasis of colorectal cancer specimens: Top four statistically upregulated (Figure 1A) and downregulated (Figure 1B) miRNAs and dendrogram view of all miRNAs that are statistically significant ($p \leq 0.05$) in the tumor and metastatic tissue samples and associated Venn diagram (Figure 1C).

1.	tumor vs benign
2.	Metastasis vs benign



Figure 2. Comparative enrichment analysis of microRNA for Pathway Maps, Gene Ontology, Disease by Biomarker, and Network Processes. MetaCore pathway analysis of small RNA-sequencing data from the tumor and metastatic cancer samples were uploaded to MetaCore servers and most significantly affected pathways analyzed. A: Pathway Maps: Canonical pathway maps were generated, where three out of the top four pathways are involved in colorectal cancer pathogenesis (regulation of miRNAs, role of miRNAs in cell proliferation; and role of miRNAs in migration, survival and angiogenesis). B: Gene Ontology: GO processes were analyzed illustrating that gene silencing of miRNA and post transcriptional gene silencing most impacted. C: Disease status: Significantly involved miRNAs in disease were listed; most of the miRNAs were involved in adenocarcinomas. D: Comparative enrichment analysis: Differential expression of miRNAs in disease groups.

Once miRNA gene expression was determined, we performed biological pathway analysis, using MetaCore software (previous known as GeneGo; v6.25), to identify candidate “molecular signatures.” Pathway analysis on differentially expressed benign, tumor, and metastasis samples were obtained. As expected, most of the affected miRNAs were involved in cell regulation, cell proliferation, migration, and angiogenesis of colorectal cancer (Figure 2A). Gene ontology (Figure 2B) and disease by biomarkers (Figure 2C) clearly showed that most of the miRNAs were known to be involved in colorectal and/or other cancer types. Regulation and proliferation pathway analysis data of miRNAs in tumor vs benign and metastasis vs benign cancer samples, showed most of the miRNAs were involved in colon cancer and cellular migration (Figures 3A-3E). We further viewed the network analysis of the miRNAs that were involved in signaling pathways. We observed three major pathways were commonly affected in the tumor and metastasis groups (Figures 3F-3H) and only specific to the metastasis group (Figures 3I-3K).

Differentially expressed piRNAs in colorectal cancer tissue samples

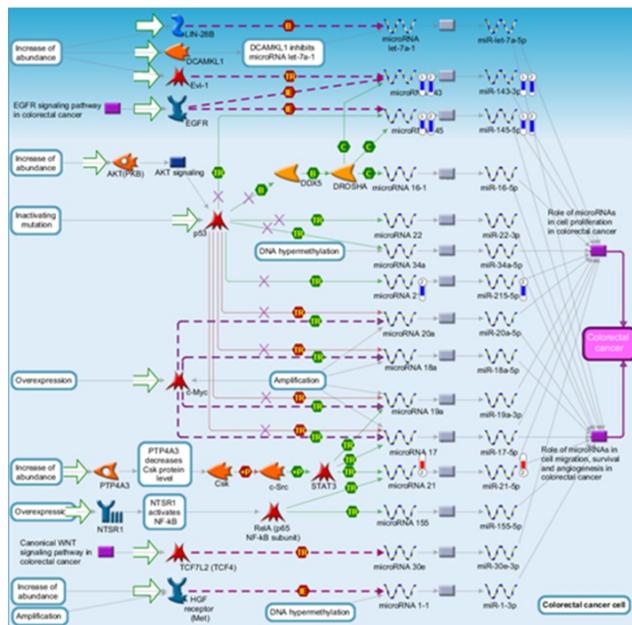
Piwi-interacting RNAs are the largest class of non-coding RNAs [23] recently shown to play

important roles in cancer progression. They also represent a third class of small RNA silencers (other two classes are miRNAs and siRNAs) that form RNA-protein complexes by interacting with piwi-proteins. These are required for both epigenetic and post-transcriptional gene silencing of retrotransposons and other genetic elements in germ line cells, particularly during spermatogenesis [24]. We remapped aligned small RNA sequenced data with piRBase annotation and identified six piRNAs (including two piRNAs with alternative transcripts) whose expression levels were statistically significant ($p < 0.05$; Figure 4A, and group wise hierarchical clustering with Venn diagram in Figure 4B). Two piRNAs were commonly upregulated in both tumor and metastasis groups. Tumor vs benign group enriched only four piRNAs which piR-hsa-25447 (3-fold), piR-hsa-23992 (3-fold), and piR-hsa-1043 (2.4-fold) were upregulated and the remaining one downregulated [piR-hsa-28876 (2.3-fold)]. In the metastasis vs benign group, 27 piRNAs were enriched (with multiple transcripts) of which 22 were upregulated [Top five: piR-hsa-22842 (22-fold), piR-hsa-23317 (9.5-fold), piR-hsa-26131 (6-fold) piR-hsa-23210.1 (6-fold) and piR-hsa-25447 (5.7-fold) and the remaining five downregulated [piR-hsa-27729 (3.4-fold), piR-hsa-7193.1 (3.4-fold), piR-hsa-7193.2

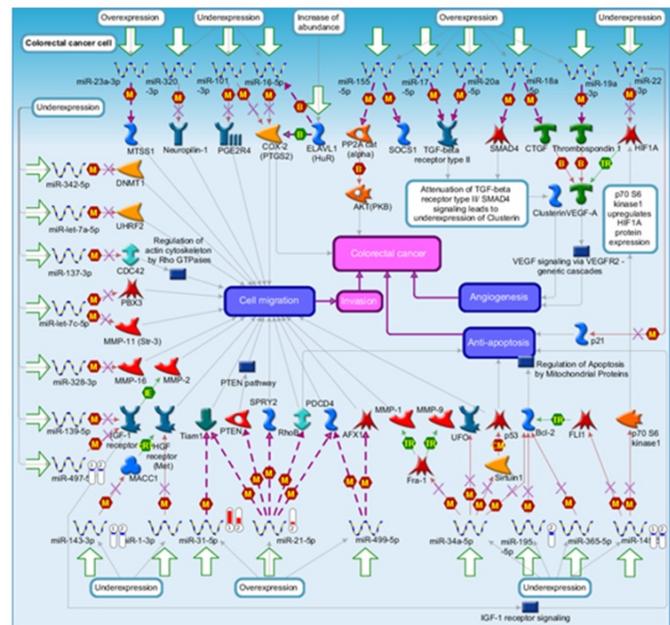
(3.2-fold) piR-hsa-27729.1 (3-fold) and piR-hsa-27730.1 (2.4-fold)] (All piRNAs were listed in the Table 2). Recently, more attention is being focused on piRNAs in relation to biologic functions and epigenetic involvement of the biomarker for diagnosis, prognosis, and therapeutic use, which still remains largely unknown. Although piRNA detection in cancer correlates with poorer clinical outcomes suggesting a functional role [25], data available so far is not sufficient to entirely discriminate between a ‘passenger’ role for the ectopic expression versus a

‘driver’ role in the pathogenesis of these diseases [25]. Cancer associated piRNAs are especially upregulated in breast cancer (piR-4987, piR-20365, piR-20485, piR-20582 and piR-932) [26, 27], and gastric cancer (piR-823 and piR651) [28, 29]. In kidney cancer both upregulated piRNAs (piR-32051, piR-39894, piR-43607) and down-regulated piRNAs (piR-38756, piR-57125, piR-30924) [30] have been identified. piRNA down-regulation has also been observed in pancreatic cancer (piR-017061) [31], multiple myeloma (piR-823), and lung cancer (piR-L-163) [32].

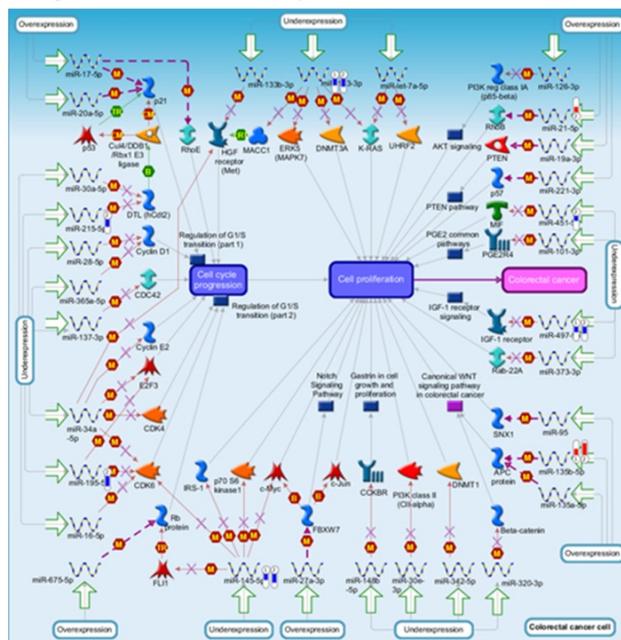
A. Regulation of microRNAs in colorectal cancer



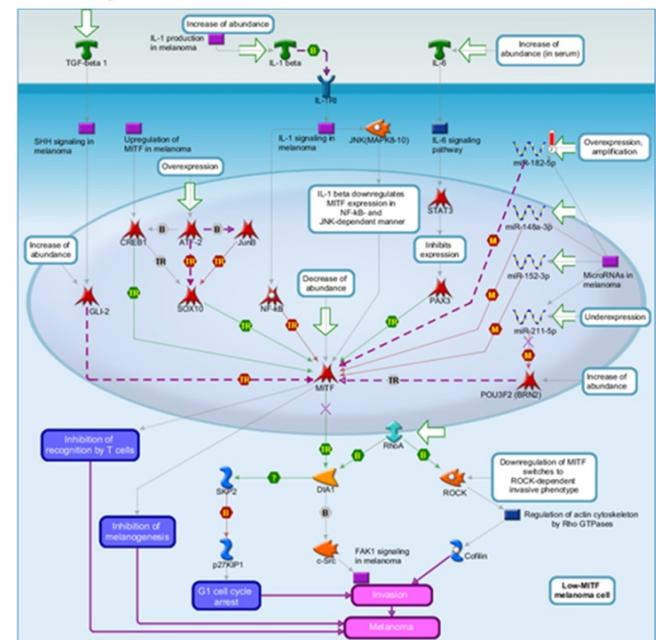
B. Role of microRNAs in cell migration, survival and angiogenesis in colorectal cancer



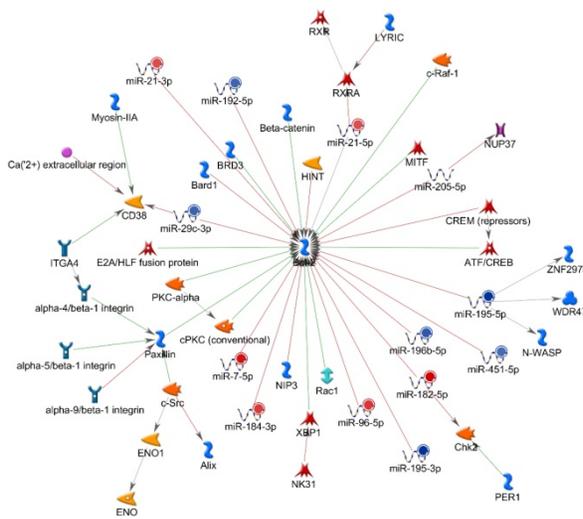
C. Regulation of microRNAs in cell proliferation in colorectal cancer



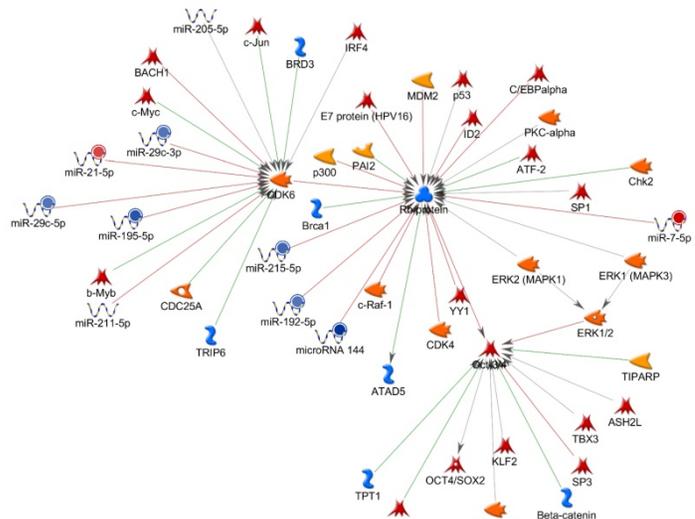
D. Down regulation of MITF in melanoma



I. Biological network in metastasis vs benign



J. Biological network in metastasis vs benign



K. Biological network in metastasis vs benign

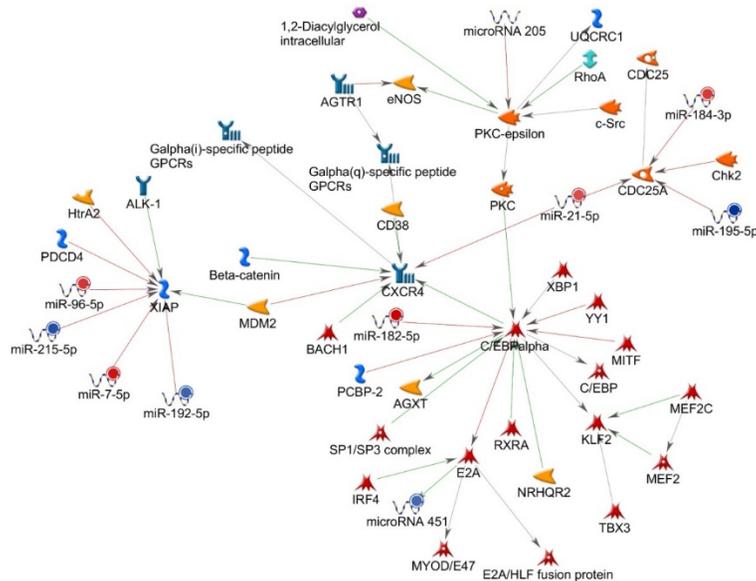


Figure 3. Top regulated Pathway Maps and Network processes. Small RNA-seq data from the tumor and metastasis cancer samples were entered into MetaCore and top regulated pathways and network processes were presented. **A to E: Regulation of miRNAs in Colorectal Cancer:** Top five highest scored pathway maps linked to colorectal cancers with involved in migration and angiogenesis. Upward thermometers have red color and indicate up-regulated signals and downward (blue) ones indicate downregulated expression levels of the genes. **F-H: Generation of biological networks using comparison Analyze Networks (cAN) algorithm common in both Colorectal Cancer groups:** Comparison of the disease group enrichment analysis revealed three commonly affected pathways in both groups. **I-K: Generation of biological networks using Analyze Networks (AN) algorithm in metastasis Colorectal Cancer group only:** This is a variant of the shortest paths algorithm with main parameters of enrichment with enriched miRNAs prioritized based on the number of fragments of canonical pathways in the networks. Up-regulated genes are marked with red circles; down-regulated with blue circles. The 'checkerboard' color indicates mixed expression of the gene between files or between multiple tags for the same gene.

A. Two commonly upregulated piRNAs in tumor and metastasis samples

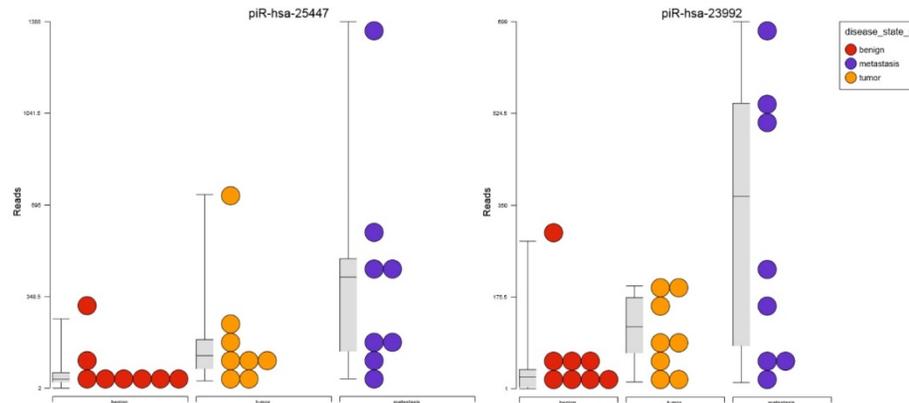
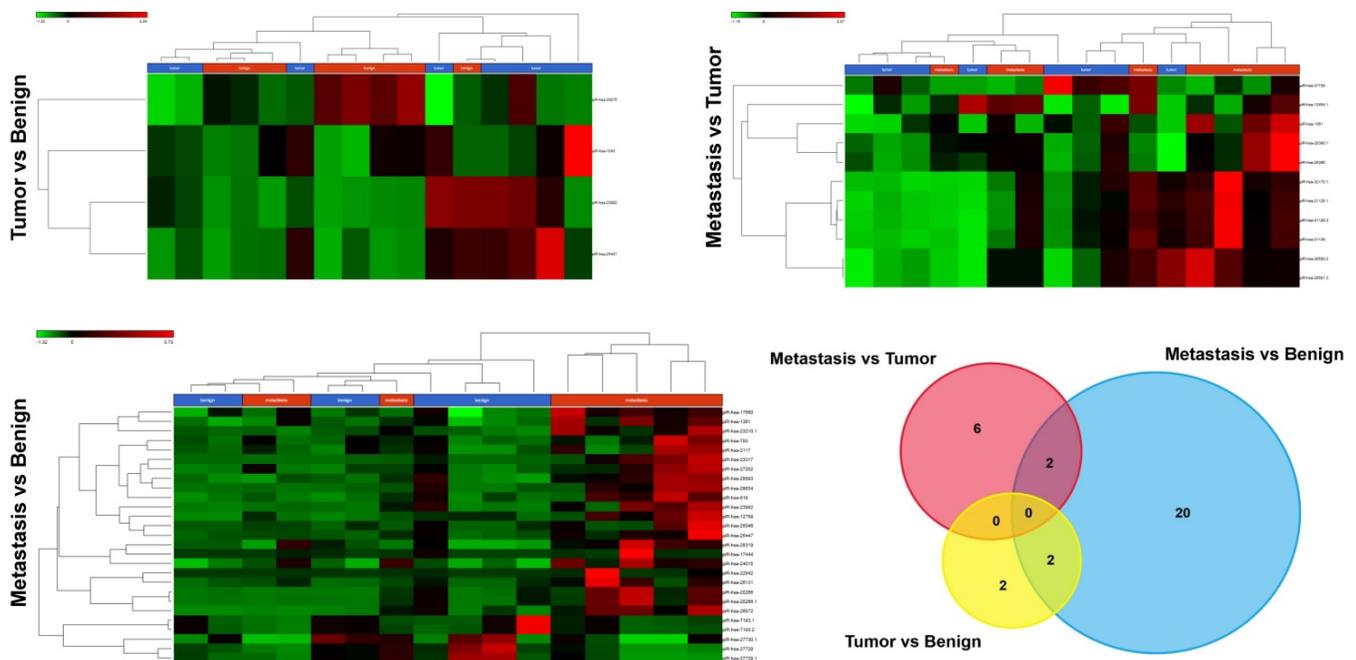
B. Differentially expressed piRNAs in tumor and metastasis cancer samples ($p < 0.05$)

Figure 4. Differential expression of piRNAs in colorectal cancer: piRNA analysis revealed a total of two piRNAs statistically regulated commonly between two groups. They are (Figure 4A) listed as dot plots. A dendrogram view of all statistically significant piRNAs ($p \leq 0.05$) are shown as hierarchical clustering and along with a Venn diagram (Figure 4B).

Differentially expressed long non-coding RNAs in colorectal cancer tissue samples

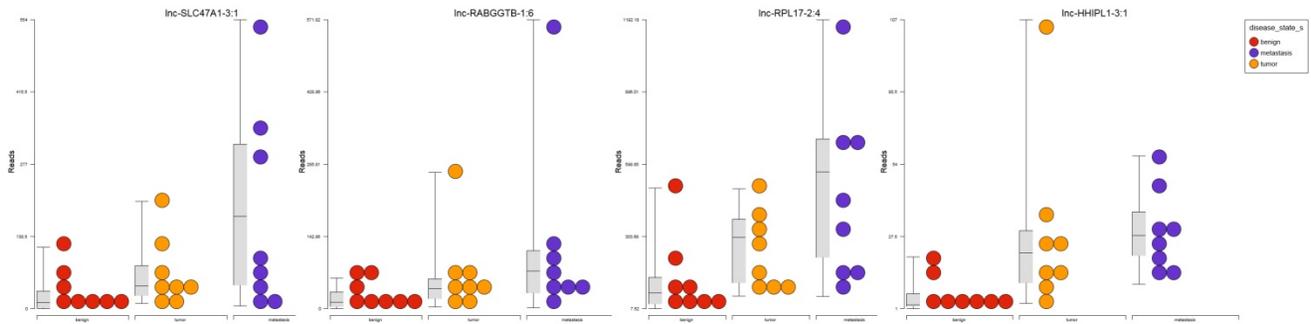
Long non-coding RNAs, larger than 200 nucleotides [24], are the most recent and least characterized ncRNAs. lncRNAs have tissue-specific expression and are expressed in a regulated manner correlating with distinct gene sets that influence cell function [33]. They are also expressed as both tumor suppressors and promoters [34-36]. Our small RNA sequencing data was remapped to identify differentially expressed lncRNAs in colorectal cancer samples. It revealed 18 lncRNAs in tumor vs benign,

89 in metastasis vs benign, and 15 in metastasis vs tumor groups being significantly expressed ($p < 0.05$). The top upregulated ones being lnc-SLC47A1-3:1, lnc-RABGGTB-1:6, lnc-RPL17-2:4 and lnc-HHIPL1-3:1 (Figure 5A); and the top downregulated lncRNAs being lnc-ABHD3-1:4, lnc-ABHD3-1:2, lnc-ZNF860-1:1 and lnc-PCYOX1L-2:15 (Figure 5B). Stage wise hierarchical clustering with Venn diagram for both tumor and metastasis groups is illustrated in Figure 5C along with differentially expressed lncRNAs listed in the Table 3. lncRNAs are a heterogeneous group of transcripts, with diverse mechanisms and differential expression in many diseases including tumors [35].

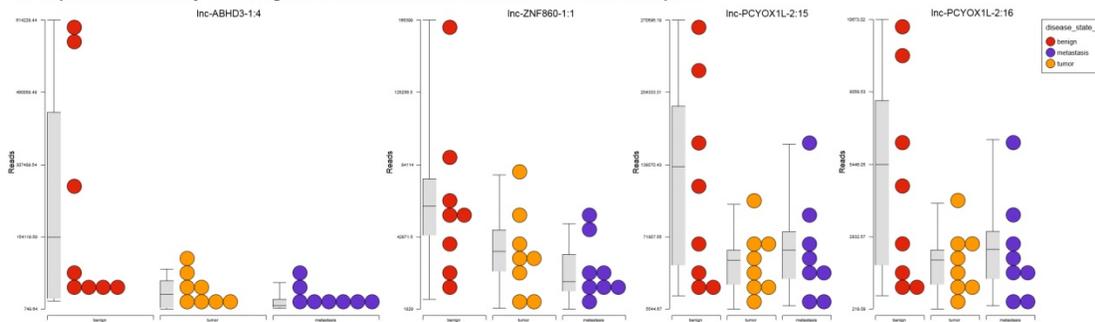
ABHD3 is an alpha/beta hydrolase protein found in a wide range of enzyme domains whose function remains unknown, and multiple transcripts of this

gene were downregulated in tumor and metastasis samples in this study. Further exploration is needed to assess their function in cancer and other diseases.

A. Top four commonly upregulated lncRNAs in both tumor and metastasis samples



B. Top four commonly downregulated lncRNAs in tumor and metastasis samples



C. Differentially expressed lncRNAs in both tumor and metastasis samples (p<0.05)

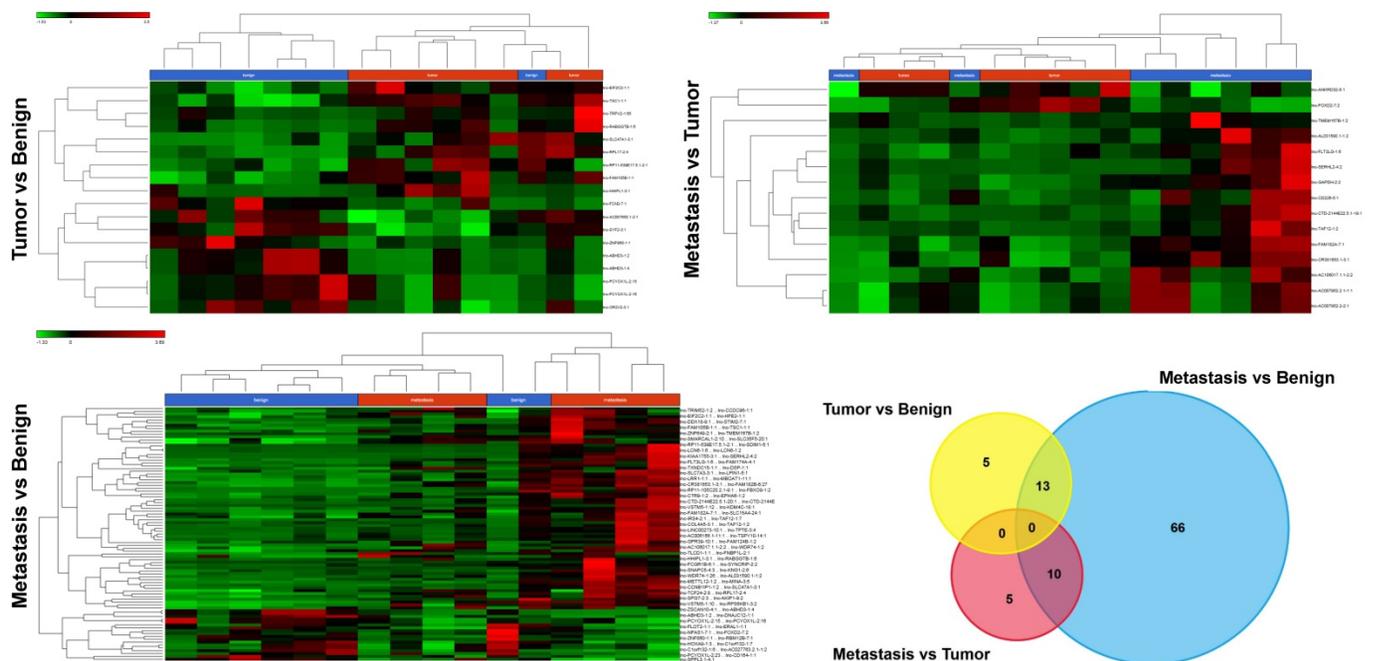


Figure 5. Differential expression of long non-coding RNAs (lncRNAs) in colorectal cancer: Top four upregulated lncRNA were shown in Figure 5A and top four downregulated lncRNAs shown in Figure 5B. Dendrogram view of lncRNAs that are statistically significant ($p \leq 0.05$) as a hierarchical tree with stage wise comparison along with a Venn diagram (Figure 5C).

Table 2. Differential expression of piRNAs in three groups.

Metastasis vs benign		
piRNA ID	Fold Change	P-value
piR-hsa-22842	22.24	2.17E-02
piR-hsa-23317	9.52	2.60E-02
piR-hsa-26131	5.96	1.46E-02
piR-hsa-23210	5.87	4.40E-02
piR-hsa-25447	5.73	2.65E-03
piR-hsa-20266	5.57	3.02E-02
piR-hsa-17444	5.46	4.24E-02
piR-hsa-20266	5.45	3.63E-02
piR-hsa-23992	5.06	3.62E-03
piR-hsa-28634	4.95	1.66E-02
piR-hsa-25046	4.95	2.38E-02
piR-hsa-26872	4.76	2.81E-02
piR-hsa-27282	4.58	3.17E-03
piR-hsa-619	4.02	3.12E-02
piR-hsa-793	3.85	4.45E-02
piR-hsa-28319	3.60	4.19E-02
piR-hsa-12789	3.54	1.53E-02
piR-hsa-28593	3.47	2.38E-02
piR-hsa-1361	3.41	6.20E-03
piR-hsa-2117	3.10	3.89E-02
piR-hsa-24015	2.99	5.88E-03
piR-hsa-17560	2.02	4.39E-02
piR-hsa-27730	-2.43	3.51E-02
piR-hsa-27729	-3.07	1.89E-02
piR-hsa-7193	-3.25	4.67E-02
piR-hsa-7193	-3.39	4.23E-02
piR-hsa-27729	-3.40	1.26E-02

Tumor vs benign		
piRNA ID	Fold Change	P-value
piR-hsa-25447	3.17	1.72E-02
piR-hsa-23992	3.02	9.82E-03
piR-hsa-1043	2.43	4.19E-02
piR-hsa-28876	-2.30	3.47E-02

Metastasis vs tumor		
piRNA ID	Fold Change	P-value
piR-hsa-21126	3.08	2.68E-02
piR-hsa-21126	3.00	4.91E-02
piR-hsa-32170	2.79	4.87E-02
piR-hsa-21126	2.79	4.71E-02
piR-hsa-28390	2.47	4.23E-02
piR-hsa-28390	2.42	4.77E-02
piR-hsa-1361	2.27	3.96E-02
piR-hsa-12454	2.26	4.88E-02
piR-hsa-26589	2.18	4.80E-02
piR-hsa-26591	2.18	4.80E-02
piR-hsa-27729	-2.03	4.81E-02

Differentially expressed sn/snoRNAs, mt_rRNA, miscRNA, nonsense mediated decay and rRNAs in colorectal cancer tissue samples

Small nuclear RNAs, form a class of RNA molecules that localize within the nucleus of eukaryotic cells [24]. Their primary function is pre-mRNA processing, for which they are always associated with a set of specific proteins and the complexes are referred to as small nuclear ribonucleoproteins (snRNPs). A subclass of snRNAs called small nucleolar RNAs that localize to the nucleolus are implicated in the maturation of ribosomal and transfer RNA molecules by specific chemical modifications [24]. We remapped the aligned reads to the GenCode database, which contains most of the curated small RNAs. We specifically looked at sn/snoRNAs, mt_rRNA, misc_rRNA, nonsense mediated decay (NMD) and rRNAs. All the mentioned snRNAs were significantly upregulated during CRC disease progression (Figure 6A, and stage wise hierarchical clustering with Venn diagram in Figure 6C). In our investigation, nine differential expression genes were noted in the tumor vs benign group, 104 in metastasis vs benign (with multiple transcripts), and 15 in metastasis vs tumor group. This snRNA cohort included all sn/snoRNAs, mt_rRNA, misc_rRNA, NMD and rRNAs, with most belonging to sn/snoRNAs (listed in the Table 4). Increased

expression was noted during metastasis, suggesting potential biomarker use in late stage cancers. Previous studies have shown the RNU family of snRNAs to be expressed in pancreatic, colorectal and lung cancers [37] and decreasing after tumor removal by surgery. Recent studies have also shown that the SNHG (small nucleolar RNA host genes) family of snoRNAs plays a crucial role in cancer. For example, Zhang *et al.*, showed upregulation of SNHG1 in hepatocellular carcinoma, affecting tumor suppressor genes such as p53 and others [38], while Zhao *et al.*, showed downregulation of SNHG5 in gastric cancer [39]. Conversely, in our study, SNHG12 (two transcripts, 4.2-fold) and SNHG1 (3.4-fold) increased. Other snoRNAs, i.e., snoRD14 (5-fold), snoRD15B (5-fold), snoRD32A (4.7-fold) were only upregulated in the metastasis stage (Table 4). Further validation of identified snRNAs are needed to confirm their biological functions and possible use as biomarkers.

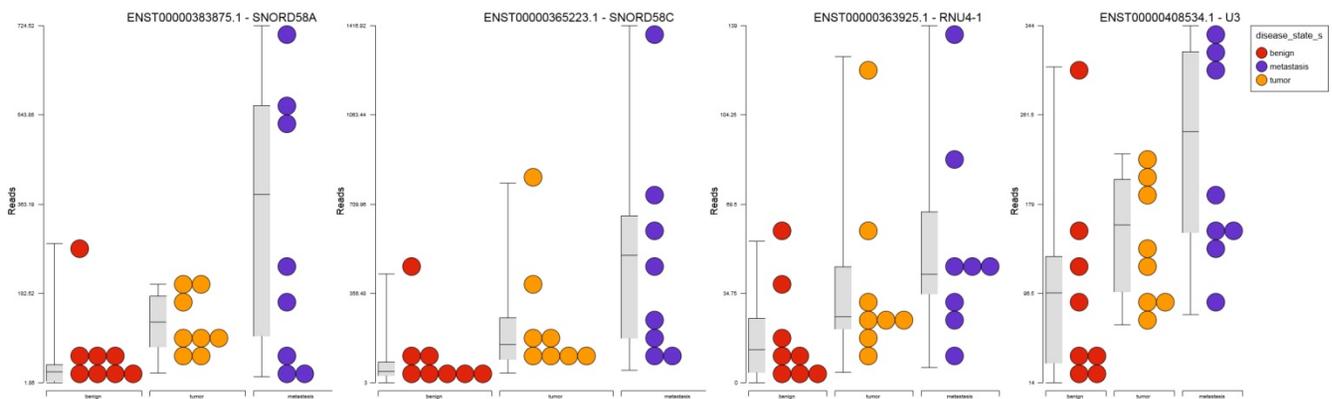
Conclusions

High-throughput sequencing technologies have effectively delineated the crucial roles of non-coding RNAs in normal cellular biology and cancer progression. We have identified several small non-coding RNAs which may be involved in the stepwise progression of normal colon tissue to cancer and subsequent metastatic disease. However, their functions are still relatively unknown. Our analysis can provide a comprehensive list of the miRNAs,

piRNAs, lncRNAs and sn/snoRNAs specific to each stage with the commonly expressed genes listed in Tables 1-4. Further understanding and validating will allow us to link these distinct small RNAs with the phenotypic traits seen in the evolution of colorectal cancer, thereby offering a potential for biomarker

development. Likewise, further understanding of their roles in cancer signaling pathways may represent a novel target for treatment. Detailed analysis of the individual transcriptome can personalize both prognosis and intervention strategies in colorectal cancer.

A. Differential expression of sn/snoRNAs in both tumor and metastasis groups



B. Differentially expressed sn/snoRNAs, mt_rRNA, miscRNA, NMD and rRNAs in both tumor and metastasis samples (p<0.05)

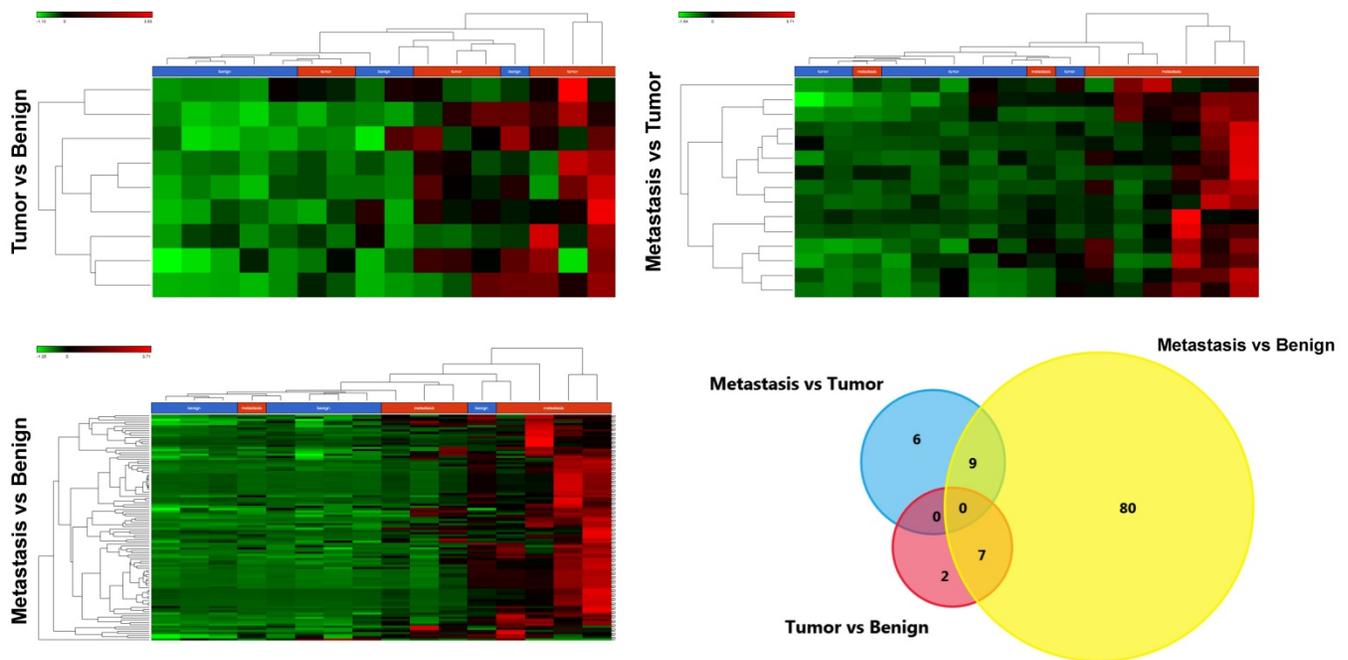


Figure 6. Differential expression of sn/snoRNAs, mt_rRNA, miscRNA, nonsense mediated decay and rRNAs in colorectal cancer tissue: small RNA-seq data was analyzed for sn/snoRNAs, mt_rRNA, miscRNA, nonsense mediated decay and rRNAs with GenCode annotation and found nine genes in tumor vs benign, 104 in metastasis vs benign and 15 in metastasis vs tumor groups. Two commonly upregulated sn/snoRNAs are shown in Figure 6A and hierarchical clustering is shown in Figure 6B with an associated Venn diagram.

Table 3. Differential expression of lncRNAs in three groups.

Metastasis vs benign		
lncRNA ID	Fold Change	P-value
lnc-FCGR1B-6:1	13.14	2.03E-02
lnc-SERHL2-4:2	10.66	5.78E-03
lnc-VSTM5-1:12	8.69	1.24E-02
lnc-KIAA1755-3:1	8.40	2.76E-02
lnc-SLC47A1-3:1	6.70	5.27E-03
lnc-TMEM167B-1:2	6.45	1.19E-02
lnc-RABGGTB-1:6	6.32	2.88E-02
lnc-AL031590.1-1:2	5.89	4.43E-04
lnc-SNAPC5-4:3	5.73	2.58E-02
lnc-LINC00273-10:1	5.54	5.82E-03
lnc-AC106017.1.1-2:2	5.12	9.48E-05
lnc-KNG1-2:6	4.87	3.50E-02
lnc-GPR39-10:1	4.84	1.21E-02
lnc-CCNB1IP1-1:2	4.82	2.74E-02
lnc-TPTE-3:4	4.74	1.03E-02
lnc-EPHA6-1:2	4.71	2.82E-02
lnc-LCN6-1:2	4.70	1.37E-02
lnc-LCN6-1:6	4.70	1.38E-02
lnc-CTD-2144E22.5.1-20:1	4.67	6.00E-03
lnc-CTD-2144E22.5.1-19:1	4.59	8.98E-03
lnc-TAF12-1:2	4.47	1.08E-02
lnc-RP11-105C20.2.1-9:1	4.45	9.90E-03
lnc-TAF12-1:7	4.37	3.31E-03
lnc-TRIM52-1:2	4.32	3.11E-02
lnc-SDIM1-5:1	4.08	2.98E-02
lnc-MINA-3:5	4.01	4.59E-02
lnc-AC006156.1-11:1	3.85	7.75E-03
lnc-DDX18-9:1	3.62	2.89E-02
lnc-FNBP1L-2:1	3.54	1.01E-02
lnc-TSPY10-14:1	3.54	6.87E-03
lnc-LPIN1-5:1	3.53	1.93E-02
lnc-RPL17-2:4	3.48	1.38E-03
lnc-SLC7A3-3:1	3.47	4.17E-02
lnc-FAM174A-4:1	3.44	3.69E-02
lnc-HHIPL1-3:1	3.43	4.91E-03
lnc-FAM182A-7:1	3.41	1.10E-03
lnc-FLT3LG-1:6	3.33	4.12E-03
lnc-SPG7-2:3	3.27	7.38E-03
lnc-LRR1-1:1	3.26	2.11E-02
lnc-SYNCRIP-2:2	3.25	3.44E-02
lnc-MBOAT1-11:1	3.25	1.50E-02
lnc-STIM2-7:1	3.24	1.98E-02
lnc-SMARCAL1-2:10	3.22	5.48E-03
lnc-CR381653.1-3:1	3.19	1.58E-03
lnc-METTL12-1:2	3.15	9.92E-03
lnc-WDR74-1:2	3.12	2.93E-03
lnc-COL4A5-3:1	3.04	4.80E-02
lnc-FAM182B-6:27	3.01	2.72E-03
lnc-TSC1-1:1	3.00	7.74E-03
lnc-FAM124B-1:2	2.91	3.30E-02
lnc-KDM4C-18:1	2.91	2.66E-02
lnc-FAM105B-1:1	2.87	1.61E-02
lnc-CTR9-1:2	2.86	3.70E-02
lnc-EIF2C2-1:1	2.84	1.81E-03
lnc-DSP-1:1	2.64	1.83E-02
lnc-TCF24-2:8	2.59	4.14E-02
lnc-ZNF649-2:1	2.54	1.39E-02
lnc-TXNDC15-1:1	2.51	2.36E-02
lnc-TLCD1-1:1	2.47	7.48E-03
lnc-WDR74-1:26	2.45	4.36E-02
lnc-SLC15A4-24:1	2.39	1.31E-02
lnc-CCDC96-1:1	2.39	8.43E-03
lnc-FBXO9-1:2	2.10	4.98E-02
lnc-ZSCAN10-4:1	2.07	1.92E-02
lnc-HFE2-1:1	2.06	2.10E-02
lnc-RPS6KB1-3:2	2.05	2.12E-02
lnc-IRS4-2:1	1.89	3.46E-02
lnc-RP11-539E17.5.1-2:1	1.87	2.06E-02
lnc-SLC35F5-20:1	1.73	4.53E-02
lnc-VSTM5-1:10	1.73	3.94E-02
lnc-AKIP1-9:2	1.44	2.89E-02
lnc-C1orf132-1:6	-1.98	4.61E-02
lnc-C1orf132-1:7	-2.00	4.28E-02
lnc-SPPL3.1-4:1	-2.04	2.10E-02
lnc-ERAL1-1:1	-2.28	2.43E-02
lnc-PCYOX1L-2:16	-2.59	1.71E-02
lnc-PCYOX1L-2:15	-2.60	1.54E-02
lnc-FOXD2-7:2	-3.09	2.55E-02
lnc-NPASI-7:1	-3.10	3.97E-02
lnc-CD164-1:1	-3.32	9.09E-03
lnc-HOXA9-1:3	-3.55	2.22E-02
lnc-FLOT2-1:1	-3.59	1.13E-02
lnc-PCYOX1L-2:23	-3.66	7.64E-03
lnc-DNAJC12-1:1	-3.69	1.39E-02
lnc-AC027763.2.1-1:2	-3.99	4.58E-03
lnc-ZNF860-1:1	-4.47	1.09E-03
lnc-RBM12B-7:1	-5.98	5.19E-03
lnc-ABHD3-1:2	-9.79	3.63E-04
lnc-ABHD3-1:4	-9.84	4.32E-04

Tumor vs benign		
lncRNA ID	Fold Change	P-value
lnc-HHIPL1-3:1	4.47	7.67E-03
lnc-RABGGTB-1:6	3.91	3.81E-02
lnc-TRPV2-1:65	3.17	3.95E-02
lnc-SLC47A1-3:1	3.07	9.59E-03
lnc-RPL17-2:4	2.63	6.59E-03
lnc-FAM105B-1:1	2.42	1.34E-02
lnc-TSC1-1:1	2.09	2.01E-02
lnc-EIF2C2-1:1	2.06	1.88E-02
lnc-RP11-539E17.5.1-2:1	1.85	4.42E-02
lnc-AC087650.1-2:1	-1.83	3.72E-02
lnc-PCYOX1L-2:15	-2.09	2.81E-02
lnc-PCYOX1L-2:16	-2.10	3.02E-02
lnc-SYF2-3:1	-2.20	4.46E-02
lnc-OR2V2-3:1	-2.24	4.60E-02
lnc-ZNF860-1:1	-2.53	4.12E-02
lnc-FCN2-7:1	-3.02	3.38E-02
lnc-ABHD3-1:4	-4.50	3.99E-03
lnc-ABHD3-1:2	-4.51	3.42E-03

Metastasis vs tumor		
lncRNA ID	Fold Change	P-value
lnc-SERHL2-4:2	10.66	3.80E-02
lnc-TAF12-1:2	5.81	1.83E-02
lnc-CTD-2144E22.5.1-19:1	4.35	3.37E-02
lnc-TMEM167B-1:2	4.20	2.65E-02
lnc-CD226-5:1	4.08	3.16E-02
lnc-GAPDH-2:2	3.75	3.16E-02
lnc-AL031590.1-1:2	3.59	2.51E-02
lnc-FLT3LG-1:6	3.08	1.05E-02
lnc-AC106017.1.1-2:2	3.06	8.88E-03
lnc-FAM182A-7:1	2.50	2.15E-02
lnc-AC007952.2-2:1	2.35	4.59E-02
lnc-CR381653.1-3:1	2.31	3.59E-02
lnc-AC007952.2.1-1:1	2.26	4.53E-02
lnc-ANKRD32-5:1	-1.81	2.94E-02
lnc-FOXD2-7:2	-3.10	2.18E-02

Table 4. Differential expression of sn/snoRNAs in three groups.

Metastasis vs benign			Tumor vs benign			Metastasis vs tumor		
Gene ID	Fold Change	P-value	Gene ID	Fold Change	P-value	Gene ID	Fold Change	P-value
RNU2-6P	18.79	2.42E-02	RNU4-2	4.08	3.27E-02	RNVU1-9	5.30	4.43E-02
RNU2-48P	15.60	3.99E-02	RNU4-1	3.60	3.11E-02	SNORD14	5.01	7.90E-03
RNU2-33P	15.55	4.00E-02	Y_RNA	3.20	3.17E-02	RNA5SP263	4.94	3.40E-02
RNU2-61P	15.28	3.68E-02	SNORD58A	3.13	4.28E-03	SNORD15B	4.88	3.62E-02
RNU2-37P	14.93	3.10E-02	SNORD58C	2.68	1.65E-02	SNORD32A	4.69	3.05E-02
RNU2-36P	12.50	3.68E-02	U3	1.85	4.87E-02	SNORD83B	4.33	1.68E-02
RNU2-5P	11.40	4.50E-02				RNA5SP502	4.30	4.41E-02
RNU2-2P	10.74	3.55E-02				SNORD53	3.78	1.74E-02
SNORD83A	9.63	9.78E-03				SCARNA10	3.75	3.00E-02
Y_RNA	9.62	2.26E-02				SNORD33	3.15	5.91E-03
RNU12	8.98	1.64E-02				SNORD14C	2.89	1.94E-02
RNA5SP486	8.84	4.44E-02				SNORD94	2.87	1.65E-02
SNORD12C	8.42	1.27E-02				SNORD127	2.58	4.48E-02
SNORA18	7.95	2.23E-02				SNORD119	2.21	3.58E-02
SNORD62	7.31	2.83E-02				SNORD43	1.73	3.20E-02
SNORD83B	6.73	4.07E-04						
SNORD14	6.63	2.71E-03						
RNA5SP214	6.16	2.23E-02						
RNU5B-1	6.02	2.53E-02						
RNA5SP74	5.94	3.73E-02						
RNA5S15	5.94	2.95E-02						
RNU5E-1	5.91	2.09E-02						
RNU2-70P	5.85	2.75E-02						
RNA5S16	5.81	3.93E-02						
RNA5S5	5.74	3.26E-02						
RNA5S7	5.67	3.39E-02						
RNA5S12	5.66	4.86E-02						
RNA5S14	5.58	4.83E-02						
RNA5S3	5.57	3.86E-02						
RNVU1-19	5.53	3.52E-02						
SNORD15B	5.51	1.26E-02						
RNA5-8SP2	5.44	6.42E-03						
RNA5S11	5.42	4.69E-02						
SCARNA4	5.40	1.15E-02						
RNU2-64P	5.38	2.96E-02						
RNA5SP481	5.24	2.53E-02						
SNORA81	5.23	2.83E-02						
RNU1-11P	5.20	3.05E-02						
RNA5SP149	5.17	3.05E-02						
SNORD58A	5.08	1.73E-03						
RNVU1-10	4.94	4.76E-02						
SNORD53	4.88	4.72E-03						
RNVU1-1	4.67	4.86E-02						
SNORD96A	4.67	1.30E-02						
RNU1-27P	4.63	2.46E-02						
Y_RNA	4.60	2.64E-02						
RNVU1-18	4.60	2.05E-02						
SNORD16	4.59	2.88E-02						
RNA5-8SP6	4.56	2.29E-03						
SNORD58C	4.49	2.38E-03						
RN7SKP76	4.45	8.97E-03						
SNORD89	4.34	3.15E-02						
SNHG12	4.20	5.38E-03						
RNU1-2	4.18	1.48E-02						
RNVU1-9	4.12	4.17E-02						
SNORA14B	4.08	1.83E-02						
RNU1-4	4.07	4.72E-02						
SNORD95	4.06	4.86E-02						
SNORD78	3.91	2.36E-02						
RNY5	3.90	2.14E-02						
SNHG5	3.89	2.23E-02						
SNHG1	3.87	2.75E-03						
RNVU1-7	3.86	4.96E-02						
SNORD94	3.79	3.60E-03						
SNORD33	3.72	2.01E-03						
MT-RNR1	3.57	1.52E-02						
RN7SL674P	3.53	1.67E-02						
RNU4-1	3.53	2.33E-02						
U3	3.47	4.20E-02						
SNHG1	3.41	3.54E-02						
RNA5-8SP4	3.35	1.47E-02						
SNORD51	3.34	7.83E-03						
SNORD100	3.31	1.34E-02						
SNORD15A	3.27	1.33E-02						
SNORD14C	3.27	7.82E-03						
RN7SL128P	3.25	1.17E-02						
SNORD68	3.24	1.08E-02						
SNORD38A	3.21	1.15E-02						
SNHG1	3.15	2.43E-02						
RN7SL2	3.06	1.65E-02						
RPL37A	3.01	6.22E-03						
SNORD97	2.98	3.23E-02						
SNORA50	2.94	4.24E-02						
RN7SL5P	2.91	2.63E-02						
GAS5	2.81	6.39E-03						
Y_RNA	2.81	2.25E-02						
SNHG1	2.80	1.67E-02						
LINC00910	2.79	4.85E-02						
SNORD82	2.74	2.02E-02						
SNORD5	2.65	1.56E-02						
RNU1-82P	2.57	3.42E-02						
SNORD55	2.54	2.24E-02						
SNORD4A	2.46	1.04E-02						
SNHG1	2.31	3.83E-02						
SNORD58B	2.29	1.21E-02						
SNORD20	2.19	1.75E-02						
SNORD6	2.19	4.05E-02						
RN7SK	2.17	3.87E-02						
SNHG16	1.89	2.71E-02						
SNORD45A	1.89	4.18E-02						
U3	1.87	2.32E-02						
SNORD43	1.86	2.15E-02						
U3	1.73	4.83E-02						

Abbreviations

ncRNA: non-coding RNA
 sncRNA: small non-coding RNA
 miRNA: micro RNA
 piRNA: piwi-interacting RNA
 lncRNA: long non-coding RNA
 snRNA: small nuclear RNA
 snoRNA: small nucleolar RNA
 mt_rRNA: mitochondrial rRNA
 SNHG: small nucleolar RNA host genes
 CRC: colorectal cancer
 SRA: sequence raw archive
 FDR: false discovery rate
 NGS: Next Generation Sequencing
 GO: Gene ontology

Acknowledgements

We would like to acknowledge Gene Arrays (An entity of Vedic Research, Inc., USA) for MetaCore data analysis free of charge and the Department of Surgery at Penn State University for startup support.

Funding

Funding was provided by the Penn State Department of Surgery.

Ethics

Data presented in the study was download from the NIH data sets (GSE46622) and original authors obtained the formal consents. We did not recruit any human subjects for this study (Not applicable).

Authors contributions

SVK: designed and supervised the study, experimentation, data analysis and manuscript preparation. DJR and AKT: data analysis and manuscript editing. SWH and MM: manuscript editing. All authors have read and approved the final manuscript.

Data availability and materials

Data available at NIH GEO website: project no: PRJNA201245; GEO: GSE46622.

Competing Interests

The authors have declared that no competing interest exists.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA: A Cancer Journal for Clinicians*. 2016; 66: 7-30.
2. Eddy SR. Non-coding RNA genes and the modern RNA world. *Nat Rev Genet*. 2001; 2: 919-29.

3. Lee RC, Feinbaum RL, Ambros V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*. 1993; 75: 843-54.
4. Sai Lakshmi S, Agrawal S. piRNABank: a web resource on classified and clustered Piwi-interacting RNAs. *Nucleic Acids Research*. 2008; 36: D173-D7.
5. Scott MS, Ono M, Yamada K, Endo A, Barton GJ, Lamond AI. Human box C/D snoRNA processing conservation across multiple cell types. *Nucleic Acids Research*. 2011; 40: 3676-88.
6. Cao J. The functional role of long non-coding RNAs and epigenetics. *Biological Procedures Online*. 2014; 16: 11.
7. Moyano M, Stefani G. piRNA involvement in genome stability and human cancer. *Journal of Hematology & Oncology*. 2015; 8.
8. Thomson T, Lin H. The Biogenesis and Function PIWI Proteins and piRNAs: Progress and Prospect. *Annual review of cell and developmental biology*. 2009; 25: 355-76.
9. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Res*. 2012; 22: 1775-89.
10. Scott MS, Ono M. From snoRNA to miRNA: Dual function regulatory non-coding RNAs. *Biochimie*. 2011; 93: 1987-92.
11. Röhr C, Kerick M, Fischer A, Kühn A, Kashofer K, Timmermann B, et al. High-Throughput miRNA and mRNA Sequencing of Paired Colorectal Normal, Tumor and Metastasis Tissues and Bioinformatic Modeling of miRNA-1 Therapeutic Applications. *PLoS ONE*. 2013; 8: e67461.
12. Firmino N, Martinez VD, Rowbotham DA, Enfield KS, Bennewith KL, Lam WL. HPV status is associated with altered PIWI-interacting RNA expression pattern in head and neck cancer. *Oral Oncol*. 2016; 55: 43-8.
13. Koduru S, Tiwari A, Leberfinger A, Hazard S, Kawasawa Y, Mahajan M, et al. A Comprehensive NGS Data Analysis of Differentially Regulated miRNAs, piRNAs, lncRNAs and sn/snoRNAs in Triple Negative Breast Cancer. *Journal of Cancer*. 2017; 8: 578-96.
14. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Research*. 2014; 42: D68-D73.
15. Zhang P, Si X, Skogerbø G, Wang J, Cui D, Li Y, et al. piRBase: a web resource assisting piRNA functional study. *Database*. 2014; 2014: bau110.
16. Volders P-J, Verheggen K, Menschaert G, Vandepoele K, Martens L, Vandesompele J, et al. An update on LNCipedia: a database for annotated human lncRNA sequences. *Nucleic Acids Research*. 2015; 43: D174-D80.
17. Sehgal K, Guo X, Koduru S, Shah A, Lin A, Yan X, et al. Plasmacytoid dendritic cells, interferon signaling, and FcγRIIIb contribute to pathogenesis and therapeutic response in childhood immune thrombocytopenia. *Sci Transl Med*. 2013; 5: 193ra89.
18. Taniguchi K, Sakai M, Sugito N, Kumazaki M, Shinohara H, Yamada N, et al. PTBP1-associated microRNA-1 and -133b suppress the Warburg effect in colorectal tumors. *Oncotarget*. 2016; 7: 18940-52.
19. Wu W, Wang Z, Yang P, Yang J, Liang J, Chen Y, et al. MicroRNA-135b regulates metastasis suppressor 1 expression and promotes migration and invasion in colorectal cancer. *Molecular and Cellular Biochemistry*. 2014; 388: 249-59.
20. Nagel R, le Sage C, Diosdado B, van der Waal M, Oude Vrielink JAF, Bolijn A, et al. Regulation of the Adenomatous Polyposis Coli Gene by the miR-135 Family in Colorectal Cancer. *Cancer Research*. 2008; 68: 5795-802.
21. Bandrés E, Cubedo E, Agirre X, Malumbres R, Zárata R, Ramirez N, et al. Identification by Real-time PCR of 13 mature microRNAs differentially expressed in colorectal cancer and non-tumoral tissues. *Mol Cancer*. 2006; 5: 29.
22. O'Day E, Lal A. MicroRNAs and their target gene networks in breast cancer. *Breast Cancer Res*. 2010; 12: 201.
23. Fu A, Jacobs DI, Hoffman AE, Zheng T, Zhu Y. PIWI-interacting RNA 021285 is involved in breast tumorigenesis possibly by remodeling the cancer epigenome. *Carcinogenesis*. 2015; 36: 1094-102.
24. Morceau F, Chateauvieux S, Gaigneaux A, Dicato M, Diederich M. Long and Short Non-Coding RNAs as Regulators of Hematopoietic Differentiation. *International Journal of Molecular Sciences*. 2013; 14: 14744-70.
25. Moyano M, Stefani G. piRNA involvement in genome stability and human cancer. *Journal of Hematology & Oncology*. 2015; 8: 38.
26. Huang G, Hu H, Xue X, Shen S, Gao E, Guo G, et al. Altered expression of piRNAs and their relation with clinicopathologic features of breast cancer. *Clinical and Translational Oncology*. 2013; 15: 563-8.

27. Zhang H, Ren Y, Xu H, Pang D, Duan C, Liu C. The expression of stem cell protein Piwil2 and piR-932 in breast cancer. *Surgical Oncology*. 2013; 22: 217-23.
28. Cheng J, Deng H, Xiao B, Zhou H, Zhou F, Shen Z, et al. piR-823, a novel non-coding small RNA, demonstrates in vitro and in vivo tumor suppressive activity in human gastric cancer cells. *Cancer Letters*. 2012; 315: 12-7.
29. Cui L, Lou Y, Zhang X, Zhou H, Deng H, Song H, et al. Detection of circulating tumor cells in peripheral blood from patients with gastric cancer using piRNAs as markers. *Clinical Biochemistry*. 2011; 44: 1050-7.
30. Busch J, Ralla B, Jung M, Wotschovsky Z, Trujillo-Arribas E, Schwabe P, et al. Piwi-interacting RNAs as novel prognostic markers in clear cell renal cell carcinomas. *Journal of Experimental & Clinical Cancer Research* : CR. 2015; 34: 61.
31. Müller S, Raulefs S, Bruns P, Afonso-Grunz F, Plötner A, Thermann R, et al. Next-generation sequencing reveals novel differentially regulated mRNAs, lncRNAs, miRNAs, sdRNAs and a piRNA in pancreatic cancer. *Mol Cancer*. 2015; 14: 94.
32. Mei Y, Wang Y, Kumari P, Shetty AC, Clark D, Gable T, et al. A piRNA-like small RNA interacts with and modulates p-ERM proteins in human somatic cells. *Nature Communications*. 2015; 6: 7316.
33. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature*. 2009; 458: 223-7.
34. Huarte M, Guttman M, Feldser D, Garber M, Koziol MJ, Kenzelmann-Broz D, et al. A large intergenic non-coding RNA induced by p53 mediates global gene repression in the p53 response. *Cell*. 2010; 142: 409-19.
35. Huarte M. The emerging role of lncRNAs in cancer. *Nat Med*. 2015; 21: 1253-61.
36. Zheng GXY, Do BT, Webster DE, Khavari PA, Chang HY. Dicer-microRNA-Myc circuit promotes transcription of hundreds of long noncoding RNAs. *Nat Struct Mol Biol*. 2014; 21: 585-90.
37. Köhler J, Schuler M, Gauler TC, Nöpel-Dünnebacke S, Ahrens M, Hoffmann A-C, et al. Circulating U2 small nuclear RNA fragments as a diagnostic and prognostic biomarker in lung cancer patients. *Journal of Cancer Research and Clinical Oncology*. 2016; 142: 795-805.
38. Zhang M, Wang W, Li T, Yu X, Zhu Y, Ding F, et al. Long noncoding RNA SNHG1 predicts a poor prognosis and promotes hepatocellular carcinoma tumorigenesis. *Biomedicine & Pharmacotherapy*. 2016; 80: 73-9.
39. Zhao L, Guo H, Zhou B, Feng J, Li Y, Han T, et al. Long non-coding RNA SNHG5 suppresses gastric cancer progression by trapping MTA2 in the cytosol. *Oncogene*. 2016; 35: 5770-80.