

Original Article

Prospective target genes and pathways of miR-30a-5p in colorectal cancer: an investigation using TCGA and bioinformatics analysis

Li-Hua Yang^{1*}, Shu-Ya Yin^{1*}, Rong-Quan He¹, Wei-Jia Mo², Yu-Yan Pang², Yu-Zhuang Wu², Zhi-Gang Peng^{1*}, Ting-Qing Gan^{1*}

Departments of ¹Medical Oncology, ²Pathology, First Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi Zhuang Autonomous Region, People's Republic of China. *Equal contributors.

Received December 14, 2016; Accepted January 19, 2017; Epub March 15, 2017; Published March 30, 2017

Abstract: Objective: Our study aimed to explore the potential mechanism of microRNA-30a-5p (miR-30a-5p) in colorectal cancer (CRC) development using The Cancer Genome Atlas (TCGA) and a bioinformatics analysis. Methods: The miR-30a-5p expression profiles and related clinical data were retrieved from publicly available TCGA data. Kaplan-Meier survival curves were created using the SPSS software. The original microarray data were then obtained from the NCBI Gene Expression Omnibus (GEO) database. Eleven types of online software were applied to predict the genes of miR-30a-5p, and literature screening was performed for the validated collected genes. Additionally, functional annotation was then conducted with DAVID and BiNGO. A protein-protein interaction (PPI) network was also constructed with STRING10.0. Results: Among the CRC patients, there was a trend toward the high expression group surviving longer than low expression group, whereas the Kaplan-Meier curves exhibited no statistically significant difference ($P=0.0802$, $P=0.409$, and $P>0.05$). Regarding gene collection, we eventually gathered 92 potential target genes of miR-30a-5p. The bioinformatics analysis revealed the primarily enriched gene ontology (GO) terms and the top two pathways, which included amyotrophic lateral sclerosis (ALS) and basal cell carcinoma. In the PPI network, TP53, BCL2L1, H6PD and LDHA were identified as the "hub" genes. Conclusion: In conclusion, this study suggests that miR-30a-5p might exert its critical function in colorectal cancer development through prospective targets and relevant pathways. However, additional large samples are still needed to explore the prognostic role of miR-30a-5p in colorectal cancer.

Keywords: miR-30a-5p, colorectal cancer, targets, gene prediction, gene ontology, pathway analysis

Introduction

According to the latest cancer statistics in the US, colorectal cancer (CRC) has become the third deadliest disease and the most common cause of cancer death among Americans [1]. In contrast to America, the estimated incidence rate for CRC is relatively low in China, but there is a trend toward an increase in this rate each year [2]. Currently, the therapeutic methods for CRC prioritize surgery, and the others are adjuvant therapies including chemotherapy, radiotherapy and immunotherapy. Among the 70% of CRC patients with stages II and III disease, the recurrence and metastatic rates remain high despite mainstay surgery [3]. At present, the mechanisms of the occurrence and development of colon cancer remain unclear, which is

the main purpose of conducting the present study.

MicroRNAs (miRNAs) are small (20~25 nucleotides), highly conserved, non-coding RNAs that play a key role in gene regulation. MiRNAs are thought to inhibit protein expression by binding to complementary sequences of a mRNA and preventing its translation or targeting it for degradation [4]. Since miRNAs were initially described in 1933 [5], the roles of miRNAs have been revealed and include cellular differentiation, growth, apoptosis and cell cycle modulation [6]. Recently, profiling studies have demonstrated alterations of miRNAs in various types of tumors [7, 8] and their close correlation with oncogenesis and tumor suppression [9, 10]. Specifically, in CRC, some miRNAs have

been demonstrated to be silenced or overexpressed [11-13]. Among these miRNAs, miR-30a-5p, a member of the microRNA-30 family, has been clarified to be significantly underexpressed in CRC and closely linked to colon cancer cell growth [14, 15].

Accumulating evidence has revealed that miR-30a-5p is implicated in multiple tumors [16-18]. However, there are relatively few studies clarifying the regulatory mechanisms of miR-30a-5p in CRC progression; for example, Baraniskin A et al confirmed that restoring miR-30a-5p modulates denticleless protein homolog (DTL) expression, which inhibits cellular growth in CRC [19]. The results from another study indicated that miR-30a-5p targets insulin receptor substrate 2 (IRS2) to suppress CRC cell migration [15]. Additionally, miR-30a-5p has been found to negatively regulate PIK3CD, which suppresses CRC cell invasion and migration [20]. Collectively, these strongly evidenced findings suggest that miR-30a-5p may serve as a tumor suppressor in CRC by negatively regulating target genes. Nonetheless, the prognostic value of miR-30a-5p in CRC requires for further elucidation, and the detailed regulatory mechanism is also far from clear; which creates a strong need for studies revealing the prognostic role and mechanism of miR-30a-5p in CRC.

In this study, we conducted an investigation of miR-30a-5p expression patterns and survival outcomes in a cohort of CRC patients from the Cancer Genome Atlas (TCGA) to explore whether miR-30a-5p expression profiles could serve as predictors of survival outcomes in patients with CRC. Additionally, to ascertain the potential regulation mechanism of miR-30a-5p in CRC, we performed target prediction and bioinformatics analyses.

Materials and methods

Patient datasets

The miR-30a-5p expression profiles and related clinical data were publicly available from the TCGA. The inclusion criteria were the following: patients with intact overall survival (OS) times and well-characterized tumors in combination with no pretreatment.

Statistical analysis

Among these datasets, the patients were grouped into low expression and high expression groups using the median miR-30a-5p expression as the cut-off point. The overall survival was considered to be the period from surgical operation to death from any causes, and survival data between two groups were mean \pm standard deviation for statistical description. Kaplan-Meier survival curves were constructed to assess the overall survival of the patients, which were then further analyzed with the log-rank test to assess the survival difference between the two groups. The above statistical analyses were performed by SPSS software 22.0.

Acquisition of microarray datasets

GSE29921 was available from the NCBI Gene Expression Omnibus (GEO) database (<http://www.ncbi.nlm.nih.gov/geo/>). The total mRNA expression profiles derived from six HCT116 colon cancer cells were equally divided into two groups: the experimental and the control group. In the experimental group, miR-30a-5p was transfected with HCT116 cells in vitro using a sh30a-5p vector overexpressing human miR-30a-5p and a sponge vector contributed to the exclusion of miRNAs from the anti-guide strand of the sh30a-5p vector. Regarding the control group, the HCT116 cells expressed the control vector and the sponge vector referred to above. The mRNA expression profiles were measured and collected using the Agilent human whole genome array platform (G4845A AMADID 026652, cRNA 4 \times 44 k V2) [19].

Microarray profiling of mRNAs

The mRNA extraction from the HCT116 cell lines was conducted through a standard acid guanidinium thiocyanate-phenol-chloroform extraction procedure. As stated by the manufacturer's instructions, the QuickAmp Labeling Kit (Agilent Technologies, Palo Alto, CA, USA) was used to label the total mRNA. T7 priming and MMLV-RT were used to generate the cDNA, and the cDNA was then prepared with T7 RNA polymerase. The hybridization was performed using Agilent 10 \times blocking agent, 25 \times fragmentation buffer and 2 \times GEx hybridization buffer. Moreover, the slides were scanned following washing on the Agilent DNA Microarray Scan-

Investigation on the mechanism of miR-30a-5p in colorectal cancer

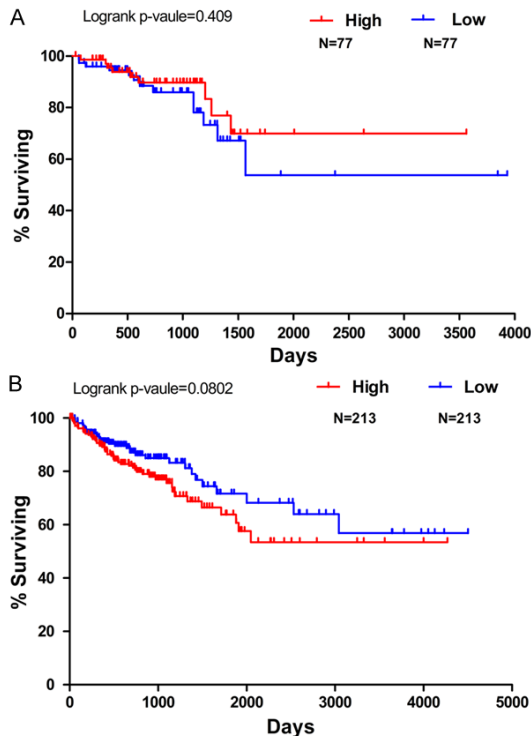


Figure 1. Influence of miR-30a-5p expression profiling on overall survival of CRC patients by Kaplan-Meier analyses. Note: A: COAD: P=0.0802; B: READ: P=0.409. Abbreviation: CRC, colorectal cancer; COAD, colon adenocarcinoma; READ, rectum adenocarcinoma.

ner (G2505B) with standard settings. Finally, the scanned images were evaluated employing the Feature Extraction Software 10.7.3.1 (Agilent) with default parameters.

Target prediction and literature screening

The predicted target genes of miR-30a-5p were retrieved with the following 11 bioinformatics tools: TargetScan, miRDB, MirTarBase, PicTar, PITA, miRanda, RNA22, TargetMiner, PolymiRTS, TarBase, and mirRNAMap. The genes that appeared in at least three types of software were chosen as the targets of miR-30a-5p for further functional annotation. Additionally, to avoid the omission of miR30a-target, the genes of miR-30a-5p that were validated by experiment were gathered from a literature screening in PubMed and two additional databases (mirTarBase and TarBase).

Functional analysis and the KEGG pathway of miR-30a-5p-target

The target genes of miR-30a-5p were functionally annotated with the Database of for Anno-

tation, Visualization, and Integrated Discovery (DAVID) v6.7 (<https://david.ncifcrf.gov/>), which provides online functional annotation and other tools, including gene conversion, functional gene classification and a gene batch viewer. After uploading the list of target gene symbols and specifying the species as homo sapiens, we chose the gene ontology and KEGG pathway as the functional analysis categories. In the functional analysis, $P < 0.05$ was defined as the selection criterion for statistical significance. Additionally, to produce a visual and intuitive representation of the results, the Biological Networks Gene Ontology tool (BiNGO, <http://www.psb.ugent.be/cbd/papers/BiNGO/>) was installed for further biological network analysis. As a cytoscape application, BiNGO is an open-source Java package that is employed to analyze GO terms that are overrepresented in biological networks and visualized in Cytoscape, and the threshold of $P < 0.05$, indicating statistically significant differences, was based on hypergeometric distribution in BiNGO.

Construction and visualization of protein-protein interaction (PPI) network

The Search Tool for the Retrieval of Interacting Genes (STRING10.0; <http://string-db.org/>) database is a tool designed to evaluate and represent protein-protein association information by scoring and integrating validated and predicted associations to produce comprehensive protein networks [21]. To determine the interactive associations, 92 targets of miR-30a-5p were input into STRING, and PPI networks were then constructed for the proteins encoded by the targets. Additionally, the Cytoscape (version 3.3; <http://cytoscape.org/>) software was applied to visualize the association between protein products based on a confidence score of > 0.4 .

Results

Prognostic value of miR-30a-5p in the TCGA cohort

After the dataset search in TCGA, a cohort of 570 patients who met the inclusion criteria was achieved. These patients were composed of 426 patients with colon adenocarcinomas (COADs) and 144 patients with rectum adenocarcinomas (READs). The two sets of patients ($n=426$ and $n=144$) were both equally stratified into low-expression and high-expression groups

Investigation on the mechanism of miR-30a-5p in colorectal cancer

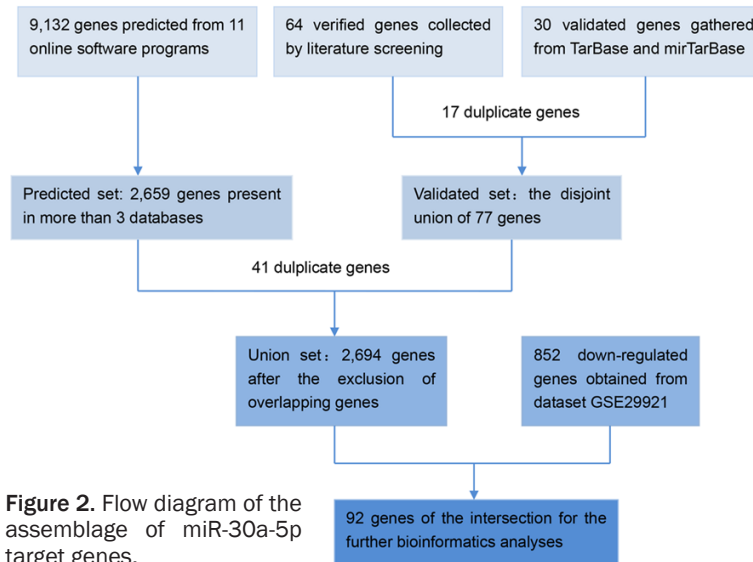


Figure 2. Flow diagram of the assemblage of miR-30a-5p target genes.

($n=213$ and $n=77$) according to the cut-off value. In the COAD set, the mean survival time of the patients with low and high miR-30a-5p levels was 2.5 ± 2.4 years and 2.4 ± 1.9 years, respectively. For the patients with READ, the survival time of the low expression group was 2.3 ± 1.9 years, and that of the high expression group was 2.2 ± 1.7 years. Based on the above, there is a downward trend in the survival time of the high expression group of CRC patients. Regarding the Kaplan-Meier curves for the COAD and READ groups (**Figure 1**), the analysis revealed no noteworthy difference in the overall survival between the low expression and high expression groups ($P=0.0802$, $P=0.409$, and $P>0.05$).

Genes from microarray data with miR-30a-5p transfection

Finally, a total of 852 down-regulated genes were obtained from the experimental group relative to the control group according to the following threshold: ($|\log FC| > 1.5$ and P value < 0.05).

Target gene assemblage of miR-30a-5p

It revealed that 64 genes were verified as the target genes of miR-30a-5p in literature screening. Moreover, 30 genes were also found to be experimentally validated in the TarBase and mirTarBase databases with the inclusion criteria of Western blotting, qPCR or luciferase assays. After the exclusion of 17 genes (AVEN,

FOXD1, BDNF, BECN1, TNRC6A, SEPT7, MTDH, Snai1, PRDM1, CDH1, BCL11A, HSPA5, EYA2, SOX4, RUNX2, VIM, and DTL) that were both in the literature and the TarBase/mirTarBase database, there was ultimately a non-overlapping group of 77 genes that served as the “validated set” of miR-30a-5p target genes. Additionally, a target prediction was performed with 11 online software programs, yielding a collection of 9132 predicted genes of miR-30a-5p. Lastly, a set of 2659 genes present in more than three software programs was eventually considered as the “predicted set”

of miR-30a-5p targets. Based on the above, the disjoint union of these 77 genes (validated set) and 2659 genes (predicted set) was acquired and summed to 2694. Moreover, 41 overlapping genes (VIM, MTDH, SMAD1, Snai1, PIK3CD, PRDM1, SEPT7, AVEN, FOXD1, NEUROD1, ABL1, BDNF, NOTCH1, BECN1, TNRC6A, RUNX2, ERG, ESR2, BCL11A, BCL9, HSPA5, EYA2, SOX4, TP53, CD99, SH2B3, ITGB3, TAB3, NFATc3, Irs2, IGF1R, ATG5, Sox11, Smarcd2, Six1, NAP1L1, TMED3, CAT, DII4, TFDP1, and IDH1) between these two sets were screened out. For convenience, the union of 2694 genes was defined as the “union set” of miR-30a-5p targets. Additionally, to identify the aberrantly expressed genes of miR-30a-5p in the CRC cells, dataset GSE29921 was downloaded to obtain 852 down-regulated genes after miR-30a-5p was up-regulated as mentioned above. These genes gathered from the genechip were further designated as the “down-regulated set” of miR-30a-5p targets. The intersection of the “union set” and the “down-regulated set” was what we needed, and ultimately, the intersection of 92 target genes of miR-30a-5p in the CRC cells was ready for the following bioinformatics analyses (**Figure 2**).

Function and pathway annotation of the miR-30a-5p targets

To identify the 92 targets of miR-30a-5p at the gene level, GO term description, GO enrichment analysis and visualized network of functional annotation were performed. These GO terms of

Table 1. The GO biological process terms enriched with potential miR-30a-5p targets by DAVID

GO ID	GO Term	No.*	P value	Gene Symbol
0021954	Central nervous system neuron development	4	0.000590989	NDEL1, GNAQ, LHX8, GLI2
0060052	Neurofilament cytoskeleton organization	3	0.000724319	NDEL1, CLN8, NEFM
0021953	Central nervous system neuron differentiation	4	0.001142888	NDEL1, GNAQ, LHX8, GLI2
0043009	Chordate embryonic development	8	0.001576069	TP53, PTCH1, BCL2L1, GLI2, etc.
0009792	Embryonic development ending in birth or egg hatching	8	0.001659397	TP53, PTCH1, BCL2L1, GLI2, etc.
0007611	Learning or memory	5	0.002596621	ATXN1, ITGA5, ESR2, LHX8, etc.
0006928	Cell motion	9	0.003047345	NDEL1, UNC5C, ESR2, GLI2, etc.
0045104	Cytoskeleton Organization	3	0.004721755	NDEL1, CLN8, NEFM
0048598	Embryonic morphogenesis	7	0.005048477	GNAQ, TP53, PTCH1, GLI2, etc.
0007507	Heart development	6	0.005071914	GNAQ, PTCH1, GLI2, etc.

*: No. means the total number of targets involved in each term.

Table 2. The GO cellular component terms enriched with potential miR-30a-5p targets by DAVID

GO ID	GO Term	No.*	P value	Gene Symbol
0005856	Cytoskeleton	13	0.043826471	PKNOX2, DNMT3, FGD5, etc.
0060053	Neurofilament cytoskeleton	2	0.04972181	NDEL1, NEFM

*: No. means the total number of targets involved in each term.

Table 3. The GO molecular function terms enriched with potential miR-30a-5p targets by DAVID

GO ID	GO Term	No.*	P value	Gene Symbol
0003727	Single-stranded RNA binding	3	0.008719148	ATXN1, A1CF, MSI2
0000166	Nucleotide binding	20	0.01930989	TP53, RAB27B, RBM26, etc.
0030528	Transcription regulator activity	15	0.023218175	TP53, ARID3A, GLI2, etc.
0008266	Poly(U) RNA binding	2	0.031022979	ATXN1, MSI2, etc.
0003723	RNA binding	9	0.03329995	ATXN1, MOV10, A1CF, etc.
0008187	Poly-pyrimidine tract binding	2	0.041151719	ATXN1, MSI2

*: No. means the total number of targets involved in each term.

the 92 genes were projected to the three types of gene ontologies in DAVID, and we finally observed that the target genes of miR-30a-5p were primarily accumulated in the following biological processes: cytoskeleton organization, chordate embryonic development, learning or memory, cell motion, embryonic morphogenesis, heart development, central nervous system neuron development and differentiation ($P < 0.05$, **Table 1**). Additionally, the overrepresented pathways related to cellular components were the cytoskeleton and neurofilament cytoskeleton ($P < 0.05$, **Table 2**). Regarding the molecular functions, the genes that were prominently enriched in the pathways displayed as follows: single-stranded RNA binding, nucleotide binding, transcription regulator activity, poly(U) RNA binding, RNA binding, and poly-pyrimidine tract binding ($P < 0.05$, **Table 3**). Through performing the BinGO analysis, we assessed the statistical overrepresentation of GO categories in the set of 92 genes as visual-

ized in Cytoscape (**Figures 3-5**). The KEGG pathway analysis revealed that there were two main KEGG pathways according to the adjusted $P < 0.05$, i.e., the amyotrophic lateral sclerosis (ALS) and basal cell carcinoma pathways (**Table 4**). Additionally, according to the above results, a column chart was created to illustrate the main GO terms and pathways (**Figure 6**). Based on the above, we finally selected these genes that were significantly involved in the top two KEGG pathways, which were expected to be the promising genes of miR-30a-5p that are linked with CRC. These genes were tumor protein p53 (TP53), Bcl-2-like protein 1 (BCL2L1), neurofilament (NEFM), protein patched homolog 1 (PTCH1) and GLI family zinc finger 2 (GLI2).

Construction of the miR-30a-5p target network

A total of 92 potential genes of miR-30a-5p were mapped into the STRING database, and a network containing 92 proteins (nodes) and 35 interactions (edges) was produced (**Figure 7**).

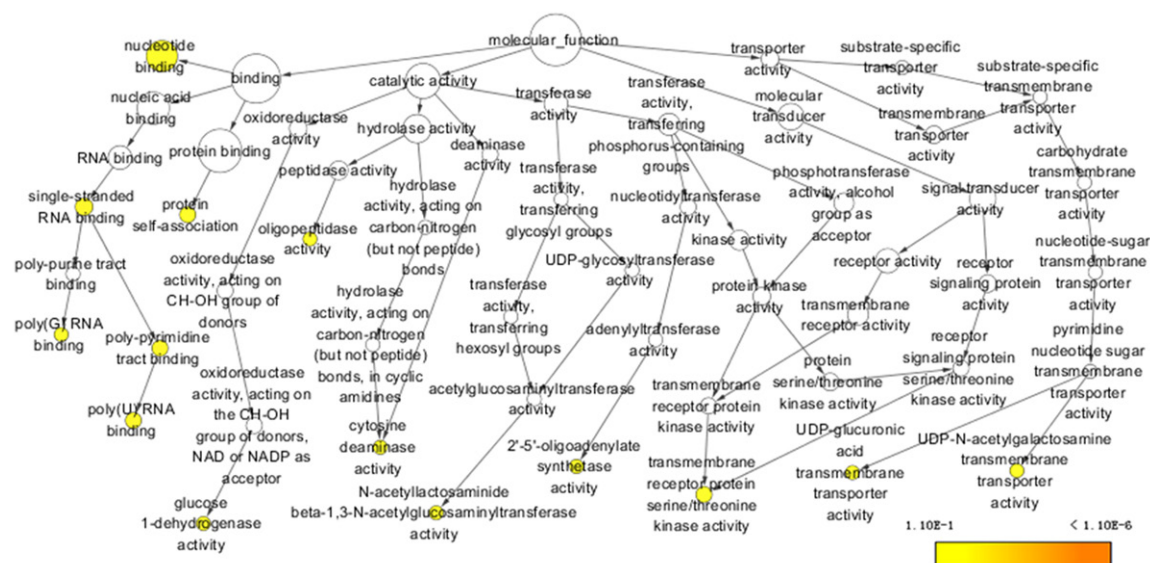


Figure 5. The hierarchical gene ontology analysis network with underlying miR-30a-5p targets of MF. Note: MF represents the networks for molecular functions. The size and color of each node correlate positively with the number of genes involved and the corresponding *P* values, respectively.

Table 4. The KEGG pathway terms enriched with potential miR-30a-5p targets by DAVID

KEGG Pathway	No.*	<i>P</i> value	Gene Symbol
Amyotrophic lateral sclerosis (ALS)	3	0.040883758	TP53, BCL2L1, NEFM
Basal cell carcinoma	3	0.043729773	TP53, PTCH1, GLI2

*: No. means the total number of targets involved in each term.

ing biomarker, miR-30a-5p has been found to be implicated in the development of multiple cancers, including CRC. In the present study, we explored the prognostic role and underlying regulation mechanisms of miR-30a-5p in CRC. In the analysis of datasets in TCGA, a relative decrease in the survival time of the patients with high expression was present. However, the Kaplan-Meier curves revealed no statistical significance by virtue of the small sample size. Therefore, to identify the prognostic value of miR-30a-5p in CRC, additional prospective research with larger capacities of samples is required. Furthermore, to gain a comprehensive understanding of the underlying regulation mechanisms of miR-30a-5p in CRC, a function enrichment analysis was performed for 92 potential genes of miR-30a-5p and followed by PPI network construction. The KEGG pathway revealed the top two pathways, i.e., the amyotrophic lateral sclerosis (ALS) pathway and the basal cell carcinoma pathway, and their respectively involved genes, i.e., TP53, BCL2L1, and

NEFM and TP53, PTCH1, and GLI2. In the PPI network, TP53, BCL2L1, and H6PD together with LDHA were identified as the key genes of miR-30a-5p. Based on the above results, TP53, BCL2L1, NEFM, PT-

CH1, GLI2, and H6PD together with LDHA are the specific genes that we intend to discuss and are closely related to the development of colon cancer.

According to the KEGG pathway analysis, there were two significant pathways, i.e., the ALS pathway and the basal cell carcinoma pathway. Thus, we speculate that miR-30a-5p likely modulates the carcinogenesis and progression of CRC via those associated pathways. However, our speculation has not been fully supported by the presently available research in which it is merely mentioned that ALS and colon cancer share the same MAPK signaling in their respective occurrence and development [22, 23]. That is, this issue still needs further investigation of the possibility that miR-30a-5p regulates colon cancer development by the ALS pathway and the basal cell carcinoma pathway.

Among the identified targets (TP53, BCL2L1, NEFM, PTCH1, GLI2, H6PD, and LDHA), TP53, a

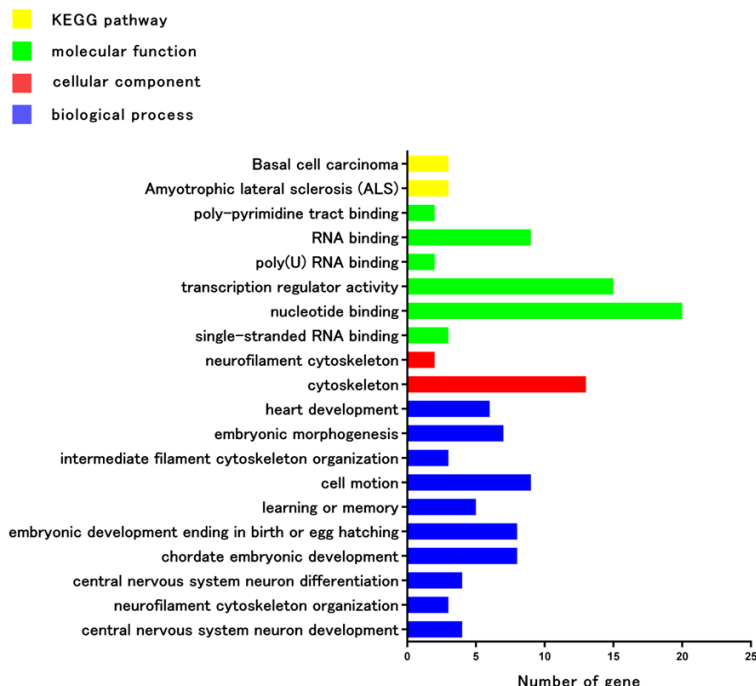


Figure 6. The column chart of significant GO terms and pathways for underlying miR-30a-5p targets related to CRC. Note: The y-axis corresponds to the main GO terms and pathways terms. The x-axis indicates the number of genes. Length of each bar refers to the specific number of genes involved in each term/pathway. Abbreviations: GO, gene ontology; CRC, colorectal cancer.

key tumor suppressor encoding tumor protein, has been reported to be implicated in the pathogenesis and development of multiple cancers, including colon cancer [24]. As an experimentally verified target of miR-30a-5p, TP53 has been suggested to link with miR-30a and CAGE to confer resistance to anti-cancer drugs in CRC. Park et al evidenced that TP53 forms the feedback loop by binding to the promoter of miR-30a and CAGE [25]. Furthermore, in colon cancer cells, it was found that the p53-CDKN1A signaling pathway played a key role in modulating cell cycle via targeting DTL [19]. These findings indicate that TP53 is an important target gene of miR-30a-5p that contributes to tumor development and may be regarded as a potential target for colon carcinoma therapy. Additionally, abundant evidence has confirmed that TP53 exerts its functions in cell-cycle arrest, DNA repair and apoptosis through the product p53-regulating target genes [26]. For example, Ohashi et al reported that p53 participates in apoptosis and the inhibition of tumor cell proliferation in colorectal cancer by transactivating the gene AKR1B10 and binding

directly to its promoter [27]. In another study, activated p53 was verified to downregulate HMGB2 transcription and thereby make the cells more sensitive to radiation in CRC [28]. Additionally, TP53 mutations are the most common mutations in human cancers [29], and research has demonstrated that CRC patients with mutant TP53 exhibit more chemo-resistance and shorter survival compared with those with wide-type TP53 [30]. In contrast, Lee et al observed that the activation of the TP53 pathway induces apoptosis and autophagy in CRC cells processed with the specific activators ilimaquinone and ethylsmenoquinone [31]. Taken together, these findings indicate that TP53 suppresses tumorigenesis and has a critical role in the pathological process of CRC.

BCL2L1, also known as Bcl-xL, has been characterized as a specific gene that is implicated in apoptosis regulation across various cancers, and the encoded protein belongs to the Bcl-2 anti-apoptosis protein family [32]. Increasing data suggests an over-expression of Bcl-xL in colon cancer as well as a close association with the invasion and development of cancer cells [33]. Moreover, it has been clearly indicated that Bcl-xL, to a lower extent, is a key anti-apoptotic factor in CRC [34], which actually provides an insight into therapeutic agents for CRC. As expected, one study observed that phosphorylating Bcl-xL at serine 62 reduces the anti-apoptotic activity of colon cancer cells, which is induced by mapatumumab and oxaliplatin in combination with hyperthermia [35]. Additionally, the phosphorylation of Bcl-xL has been observed to function in the G2 checkpoint in another study [36]. Recently, Eichhorn et al found that Bcl-xL is capable of compensating for the loss of Mcl-1 in colon cancer cell lines to overcome the anti-cancer drug ABT-263 [37]. Together, the close relationship of BCL2L1 with colorectal cancer was strongly evidenced by

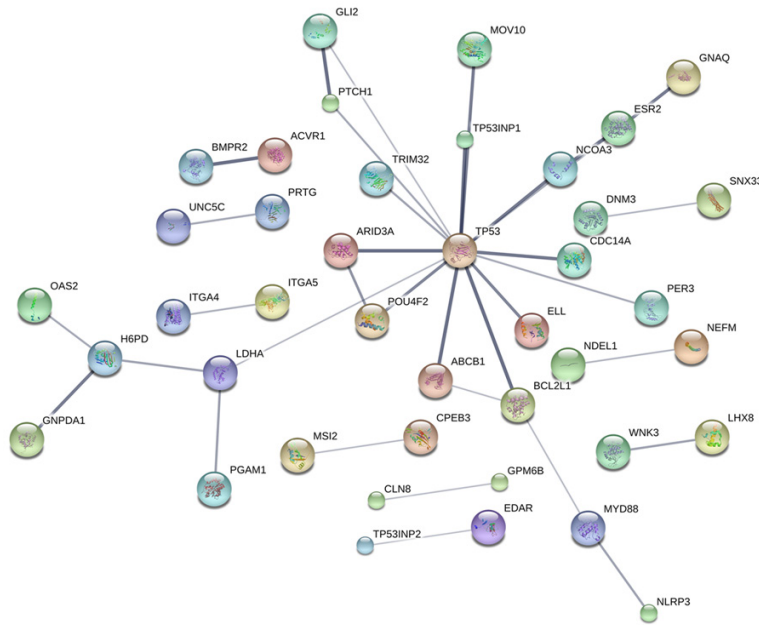


Figure 7. The PPI network of the underlying targets of miR-30a-5p. Note: The network shows hub genes (nodes) interacted with other genes, and the interactions (edges) between any two genes interactions. The line thicknesses indicate the strength of data support. Abbreviation: PPI, protein-protein interaction.

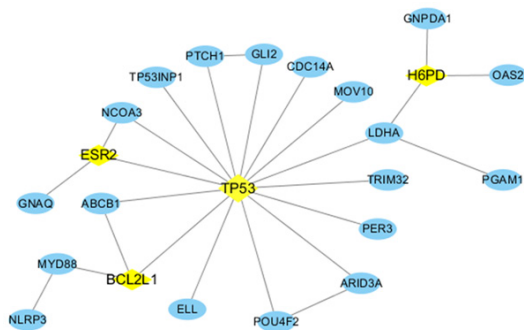


Figure 8. The PPI network visualized in Cytoscape. Note: Each blue circle stands for a gene, and yellow diamond represents a gene with a high degree (hub genes). Edges correspond to interactions between genes. Abbreviation: PPI, protein-protein interaction.

these studies, which indicates that BCL2L1 is likely a target gene upregulated by miR-30a-5p during ectopic expression in CRC to promote the carcinogenesis and development of CRC.

NEFM is known to be a marker of early neuronal differentiation. NEFM plays an essential role in nerve fiber growth [38]. In some cancers, NEFM has been revealed to be linked with risk variants and prognosis [39, 40]. However, the relationship between NEFM and colorectal cancer

has not yet been explored. There is only one study that speculates that DNA methylation-mediated inactivation of NEFM cells could possibly enhance cell proliferation in CRC [41]. Therefore, the role of NEFM in CRC still requires investigation.

PTCH1 and GLI2 are identified to be involved in the Hedgehog (Hh) signaling pathway, which is activated and has been reported to drive cell survival in human CRC [42-44]. Indeed, activating the Hh signaling pathway requires the final activation of GLI2, which is regulated by PTCH1 binding to the Hh ligand [45]. Therefore, to elucidate the underlying roles of PTCH1 and GLI2 in CRC, Peng et al revealed that the aberrant methylation of PTCH1 is likely an early event in colon

tumorigenesis because the methylation reflects an association with the down-regulated PTCH1 expression [46]. Moreover, Chung et al demonstrated that the presence of the PTCH1 mutation in CRC activates the Hh signaling pathway, which results in the promotion of colorectal carcinogenesis [47]. Regarding GLI2, it has been reported that this gene is highly expressed in colorectal carcinoma and may act as a predictor of poor survival [48]. Recent studies [49, 50] have demonstrated the blockage of Hedgehog survival signaling-induced DNA damage and cell death by decreasing GLI2 expression in CRC cells. Collectively, these findings suggest that the genes PTCH1 and GLI2 have opposite effects on the modulation of the pathogenesis of CRC via the common Hh signaling pathway.

In our study, TP53, BCL2L1, H6PD and LDHA were identified as the “hub” genes in the PPI network. Among these promising genes, the gene H6PD encodes the enzymes that mainly control the pentose phosphate pathway (PPP) in the oxidative branch, which has been reported to be a crucial regulator of tumor progression [51-53]. However, the studies that have investigated the potential association of H6PD

with CRC are rare. *Vizan et al* reported that H6PD promotes cell cycle progression in human CRC by increasing enzyme activities at late G1 and S phases [54]. Additionally, this study revealed that over-expressed H6PD is correlated with metastases in CRC by virtue of STAT3-mediated epithelial-mesenchymal transition (EMT) induction [55]. Moreover, the gene lactate dehydrogenase A (LDHA) has been reported to exert an indispensable effect on the metabolism of tumor cells [56-59]. Previous studies have revealed high levels of LDHA in CRC cells, but the potential mechanism of this action remains largely unclear. Delightedly, the fact that LDHA participates in CRC cell growth through aerobic glycolysis was strongly supported by *Billiard et al*, who demonstrated that inhibiting LDHA with the inhibitor quinoline 3-sulfonamides reverses aerobic glycolysis and thus leads to the impairment of cell survival [60]. Subsequently, the research of *Wang et al* clarified the role of LDHA in CRC via the knock-down of LDHA; these authors found that LDHA knockdown inhibits cell proliferation by decreasing lactate production as well as glucose uptake [61]. Taken together these limited findings provide us with powerful evidence of the functioning of H6PD and LDHA in CRC; however, further investigations are still desirable to establish the roles of these genes.

In summary, the identified genes, including TP53, BCL2L1, PTCH1, GLI2, H6PD and LDHA, were found to be closely correlated with CRC tumorigenesis and progression, with the exception of the unexplored gene NEFM. These findings demonstrate that miR-30a-5p most likely modulates the pathogenesis of colorectal cancer through the above well-studied genes. However, the relationships of the two significant KEGG pathways of miR-30a-5p with CRC remain to be explored. Moreover, whether miR-30a-5p might act as a prognostic signature of survival in patients with CRC remains to be investigated by additional research with large sample sizes. Overall, the bioinformatics analysis present in our study provides a comprehensive understanding of the underlying mechanism of the function of miR-30a-5p in CRC via its relevant target genes, which, to some extent, helps to support further research exploring the detailed function of miR-30a-5p in colorectal cancer.

Acknowledgements

We would like to thank the Fund of the Future Academic Stars of Guangxi Medical University (WLXSZX16042) and the Guangxi Natural Science Foundation (2015GXNSFBA139157). The funders had no part in the study design, data collection and analysis, the manuscript preparation, or the decision to publish.

Disclosure of conflict of interest

None.

Address correspondence to: Drs. Zhi-Gang Peng and Ting-Qing Gan, Department of Medical Oncology, First Affiliated Hospital of Guangxi Medical University, 6 Shuangyong Road, Nanning, Guangxi Zhuang Autonomous Region, People's Republic of China. Tel: +86-771-5353121; E-mail: Drpzg001@163.com (ZGP); 77827825@qq.com (TQG)

References

- [1] Siegel RL, Miller KD and Jemal A. Cancer statistics, 2016. *CA Cancer J Clin* 2016; 66: 7-30.
- [2] Wang S and El-Deiry WS. p73 or p53 directly regulates human p53 transcription to maintain cell cycle checkpoints. *Cancer Res* 2006; 66: 6982-6989.
- [3] You YN, Rustin RB and Sullivan JD. Oncotype DX((R)) colon cancer assay for prediction of recurrence risk in patients with stage II and III colon cancer: a review of the evidence. *Surg Oncol* 2015; 24: 61-66.
- [4] Hudson T. PROs' new quality improvement focus: will it work in practice? *Hospitals* 1991; 65: 48-50.
- [5] Lee RC, Feinbaum RL and Ambros V. The *C. Elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 1993; 75: 843-854.
- [6] Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 2004; 116: 281-297.
- [7] Dudda JC, Salaun B, Ji Y, Palmer DC, Monnot GC, Merck E, Boudousquie C, Utzschneider DT, Escobar TM, Perret R, Muljo SA, Hebeisen M, Rufer N, Zehn D, Donda A, Restifo NP, Held W, Gattinoni L and Romero P. MicroRNA-155 is required for effector CD8+ T cell responses to virus infection and cancer. *Immunity* 2013; 38: 742-753.
- [8] Guz M, Rivero-Muller A, Okon E, Stenzel-Bembenek A, Polberg K, Slomka M and Stepulak A. MicroRNAs-role in lung cancer. *Dis Markers* 2014; 2014: 218169.

- [9] Ahmad J, Hasnain SE, Siddiqui MA, Ahamed M, Musarrat J and Al-Khedhairi AA. MicroRNA in carcinogenesis & cancer diagnostics: a new paradigm. *Indian J Med Res* 2013; 137: 680-694.
- [10] Kent OA and Mendell JT. A small piece in the cancer puzzle: microRNAs as tumor suppressors and oncogenes. *Oncogene* 2006; 25: 6188-6196.
- [11] Yu Y, Nangia-Makker P, Farhana L, G Rajendra S, Levi E and Majumdar AP. miR-21 and miR-145 cooperation in regulation of colon cancer stem cells. *Mol Cancer* 2015; 14: 98.
- [12] Geng L, Chaudhuri A, Talmon G, Wisecarver JL, Are C, Brattain M and Wang J. MicroRNA-192 suppresses liver metastasis of colon cancer. *Oncogene* 2014; 33: 5332-5340.
- [13] Bovell LC, Shanmugam C, Putcha BD, Katkoori VR, Zhang B, Bae S, Singh KP, Grizzle WE and Manne U. The prognostic value of microRNAs varies with patient race/ethnicity and stage of colorectal cancer. *Clin Cancer Res* 2013; 19: 3955-3965.
- [14] Wang X, Chen L, Jin H, Wang S, Zhang Y, Tang X and Tang G. Screening miRNAs for early diagnosis of colorectal cancer by small RNA deep sequencing and evaluation in a Chinese patient population. *Onco Targets Ther* 2016; 9: 1159-1166.
- [15] Zhang Q, Tang Q, Qin D, Yu L, Huang R, Lv G, Zou Z, Jiang XC, Zou C, Liu W, Luo J, Zhao Z, Muhammad S, Wang G, Chen YG and Wang X. Role of microRNA 30a targeting insulin receptor substrate 2 in colorectal tumorigenesis. *Mol Cell Biol* 2015; 35: 988-1000.
- [16] Zhang N, Wang X, Huo Q, Sun M, Cai C, Liu Z, Hu G and Yang Q. MicroRNA-30a suppresses breast tumor growth and metastasis by targeting metadherin. *Oncogene* 2014; 33: 3119-3128.
- [17] Jiang BY, Zhang XC, Su J, Meng W, Yang XN, Yang JJ, Zhou Q, Chen ZY, Chen ZH, Xie Z, Chen SL and Wu YL. BCL11A overexpression predicts survival and relapse in non-small cell lung cancer and is modulated by microRNA-30a and gene amplification. *Mol Cancer* 2013; 12: 61.
- [18] Wang T, Li F and Tang S. MiR-30a upregulates BCL2A1, IER3 and cyclin D2 expression by targeting FOXL2. *Oncol Lett* 2015; 9: 967-971.
- [19] Baraniskin A, Birkenkamp-Demtroder K, Magh-nouj A, Zollner H, Munding J, Klein-Scory S, Reinacher-Schick A, Schwarte-Waldhoff I, Schmiegel W and Hahn SA. MiR-30a-5p suppresses tumor growth in colon carcinoma by targeting DTL. *Carcinogenesis* 2012; 33: 732-739.
- [20] Zhong M, Bian Z and Wu Z. miR-30a suppresses cell migration and invasion through down-regulation of PIK3CD in colorectal carcinoma. *Cell Physiol Biochem* 2013; 31: 209-218.
- [21] Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, von Mering C and Jensen LJ. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res* 2013; 41: D808-815.
- [22] Spencer P, Fry RC and Kisby GE. Unraveling 50-year-old clues linking neurodegeneration and cancer to cycad toxins: are microRNAs common mediators? *Front Genet* 2012; 3: 192.
- [23] Kim EK and Choi EJ. Compromised MAPK signaling in human diseases: an update. *Arch Toxicol* 2015; 89: 867-882.
- [24] Zuckerman V, Wolynec K, Sionov RV, Haupt S and Haupt Y. Tumour suppression by p53: the importance of apoptosis and cellular senescence. *J Pathol* 2009; 219: 3-15.
- [25] Park D, Kim H, Kim Y and Jeoung D. miR-30a Regulates the expression of CAGE and p53 and regulates the response to anti-cancer drugs. *Mol Cells* 2016; 39: 299-309.
- [26] Li XL, Zhou J, Chen ZR and Chng WJ. P53 mutations in colorectal cancer-molecular pathogenesis and pharmacological reactivation. *World J Gastroenterol* 2015; 21: 84-93.
- [27] Ohashi T, Idogawa M, Sasaki Y, Suzuki H and Tokino T. AKR1B10, a transcriptional target of p53, is downregulated in colorectal cancers associated with poor prognosis. *Mol Cancer Res* 2013; 11: 1554-1563.
- [28] Shin YJ, Kim MS, Kim MS, Lee J, Kang M and Jeong JH. High-mobility group box 2 (HMGB2) modulates radioresponse and is downregulated by p53 in colorectal cancer cell. *Cancer Biol Ther* 2013; 14: 213-221.
- [29] Kandath C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, Xie M, Zhang Q, McMichael JF, Wyczalkowski MA, Leiserson MD, Miller CA, Welch JS, Walter MJ, Wendl MC, Ley TJ, Wilson RK, Raphael BJ and Ding L. Mutational landscape and significance across 12 major cancer types. *Nature* 2013; 502: 333-339.
- [30] Iacopetta B. TP53 mutation in colorectal cancer. *Hum Mutat* 2003; 21: 271-276.
- [31] Lee HY, Chung KJ, Hwang IH, Gwak J, Park S, Ju BG, Yun E, Kim DE, Chung YH, Na M, Song GY and Oh S. Activation of p53 with ilimaquinone and ethylsmenoquinone, marine sponge metabolites, induces apoptosis and autophagy in colon cancer cells. *Mar Drugs* 2015; 13: 543-557.
- [32] Danial NN and Korsmeyer SJ. Cell death: critical control points. *Cell* 2004; 116: 205-219.
- [33] Zhang YL, Pang LQ, Wu Y, Wang XY, Wang CQ and Fan Y. Significance of Bcl-xL in human colon carcinoma. *World J Gastroenterol* 2008; 14: 3069-3073.

- [34] Schulze-Bergkamen H, Ehrenberg R, Hickmann L, Vick B, Urbanik T, Schimanski CC, Berger MR, Schad A, Weber A, Heeger S, Galle PR and Moehler M. Bcl-x(L) and Myeloid cell leukaemia-1 contribute to apoptosis resistance of colorectal cancer cells. *World J Gastroenterol* 2008; 14: 3829-3840.
- [35] Song X, Kim SY and Lee YJ. The role of Bcl-xL in synergistic induction of apoptosis by mapatumumab and oxaliplatin in combination with hyperthermia on human colon cancer. *Mol Cancer Res* 2012; 10: 1567-1579.
- [36] Wang J, Beauchemin M and Bertrand R. Phospho-Bcl-x(L)(Ser62) plays a key role at DNA damage-induced G(2) checkpoint. *Cell Cycle* 2012; 11: 2159-2169.
- [37] Eichhorn JM, Alford SE, Sakurikar N and Chambers TC. Molecular analysis of functional redundancy among anti-apoptotic Bcl-2 proteins and its role in cancer cell survival. *Exp Cell Res* 2014; 322: 415-424.
- [38] Garcia ML, Lobsiger CS, Shah SB, Deerinck TJ, Crum J, Young D, Ward CM, Crawford TO, Gotow T, Uchiyama Y, Ellisman MH, Calcutt NA and Cleveland DW. NF-M is an essential target for the myelin-directed "outside-in" signaling cascade that mediates radial axonal growth. *J Cell Biol* 2003; 163: 1011-1020.
- [39] Penney KL, Sinnott JA, Tyekucheva S, Gerke T, Shui IM, Kraft P, Sesso HD, Freedman ML, Loda M, Mucci LA and Stampfer MJ. Association of prostate cancer risk variants with gene expression in normal and tumor tissue. *Cancer Epidemiol Biomarkers Prev* 2015; 24: 255-260.
- [40] Ricketts CJ, Morris MR, Gentle D, Shuib S, Brown M, Clarke N, Wei W, Nathan P, Latif F and Maher ER. Methylation profiling and evaluation of demethylating therapy in renal cell carcinoma. *Clin Epigenetics* 2013; 5: 16.
- [41] Calmon MF, Jeschke J, Zhang W, Dhir M, Siebenkas C, Herrera A, Tsai HC, O'Hagan HM, Pappou EP, Hooker CM, Fu T, Schuebel KE, Gabrielson E, Rahal P, Herman JG, Baylin SB and Ahuja N. Epigenetic silencing of neurofilament genes promotes an aggressive phenotype in breast cancer. *Epigenetics* 2015; 10: 622-632.
- [42] Mazumdar T, DeVecchio J, Shi T, Jones J, Agyeman A and Houghton JA. Hedgehog signaling drives cellular survival in human colon carcinoma cells. *Cancer Res* 2011; 71: 1092-1102.
- [43] Shi T, Mazumdar T, DeVecchio J, Duan ZH, Agyeman A, Aziz M and Houghton JA. cDNA microarray gene expression profiling of hedgehog signaling pathway inhibition in human colon cancer cells. *PLoS One* 2010; 5.
- [44] Tremblay MR, McGovern K, Read MA and Castro AC. New developments in the discovery of small molecule Hedgehog pathway antagonists. *Curr Opin Chem Biol* 2010; 14: 428-435.
- [45] Agyeman A, Jha BK, Mazumdar T and Houghton JA. Mode and specificity of binding of the small molecule GANT61 to GLI determines inhibition of GLI-DNA binding. *Oncotarget* 2014; 5: 4492-4503.
- [46] Peng L, Hu J, Li S, Wang Z, Xia B, Jiang B, Li B, Zhang Y, Wang J and Wang X. Aberrant methylation of the PTCH1 gene promoter region in aberrant crypt foci. *Int J Cancer* 2013; 132: E18-25.
- [47] Chung JH and Bunz F. A loss-of-function mutation in PTCH1 suggests a role for autocrine hedgehog signaling in colorectal tumorigenesis. *Oncotarget* 2013; 4: 2208-2211.
- [48] Singovski G, Bernal C, Kuciak M, Siegl-Cachedenier I, Conod A and Ruiz I Altaba A. In vivo epigenetic reprogramming of primary human colon cancer cells enhances metastases. *J Mol Cell Biol* 2016; 8: 157-173.
- [49] Mazumdar T, DeVecchio J, Agyeman A, Shi T and Houghton JA. Blocking Hedgehog survival signaling at the level of the GLI genes induces DNA damage and extensive cell death in human colon carcinoma cells. *Cancer Res* 2011; 71: 5904-5914.
- [50] Mazumdar T, DeVecchio J, Agyeman A, Shi T and Houghton JA. The GLI genes as the molecular switch in disrupting hedgehog signaling in colon cancer. *Oncotarget* 2011; 2: 638-645.
- [51] Boren J, Montoya AR, de Atauri P, Comin-Anduix B, Cortes A, Centelles JJ, Frederiks WM, Van Noorden CJ and Cascante M. Metabolic control analysis aimed at the ribose synthesis pathways of tumor cells: a new strategy for antitumor drug development. *Mol Biol Rep* 2002; 29: 7-12.
- [52] Zhang J, Wang G, Mao Q, Li S, Xiong W, Lin Y and Ge J. Glutamate dehydrogenase (GDH) regulates bioenergetics and redox homeostasis in human glioma. *Oncotarget* 2016; [Epub ahead of print].
- [53] Lucarelli G, Galleggiante V, Rutigliano M, Sanguedolce F, Cagiano S, Bufo P, Lastilla G, Maiorano E, Ribatti D, Giglio A, Serino G, Vavallo A, Bettocchi C, Selvaggi FP, Battaglia M and Ditonno P. Metabolomic profile of glycolysis and the pentose phosphate pathway identifies the central role of glucose-6-phosphate dehydrogenase in clear cell-renal cell carcinoma. *Oncotarget* 2015; 6: 13371-13386.
- [54] Vizan P, Alcarraz-Vizan G, Diaz-Moralli S, Solovjeva ON, Frederiks WM and Cascante M. Modulation of pentose phosphate pathway during cell cycle progression in human colon adenocarcinoma cell line HT29. *Int J Cancer* 2009; 124: 2789-2796.

- [55] Liu G, Zhu J, Yu M, Cai C, Zhou Y, Yu M, Fu Z, Gong Y, Yang B, Li Y, Zhou Q, Lin Q, Ye H, Ye L, Zhao X, Li Z, Chen R, Han F, Tang C and Zeng B. Glutamate dehydrogenase is a novel prognostic marker and predicts metastases in colorectal cancer patients. *J Transl Med* 2015; 13: 144.
- [56] Shi M, Cui J, Du J, Wei D, Jia Z, Zhang J, Zhu Z, Gao Y and Xie K. A novel KLF4/LDHA signaling pathway regulates aerobic glycolysis in and progression of pancreatic cancer. *Clin Cancer Res* 2014; 20: 4370-4380.
- [57] Mirebeau-Prunier D, Le Pennec S, Jacques C, Fontaine JF, Gueguen N, Boutet-Bouzamondo N, Donnart A, Malthiery Y and Savagner F. Estrogen-related receptor alpha modulates lactate dehydrogenase activity in thyroid tumors. *PLoS One* 2013; 8: e58683.
- [58] Zhao D, Zou SW, Liu Y, Zhou X, Mo Y, Wang P, Xu YH, Dong B, Xiong Y, Lei QY and Guan KL. Lysine-5 acetylation negatively regulates lactate dehydrogenase A and is decreased in pancreatic cancer. *Cancer Cell* 2013; 23: 464-476.
- [59] Yang Y, Su D, Zhao L, Zhang D, Xu J, Wan J, Fan S and Chen M. Different effects of LDH-A inhibition by oxamate in non-small cell lung cancer cells. *Oncotarget* 2014; 5: 11886-11896.
- [60] Billiard J, Dennison JB, Briand J, Annan RS, Chai D, Colon M, Dodson CS, Gilbert SA, Greshock J, Jing J, Lu H, McSurdy-Freed JE, Orband-Miller LA, Mills GB, Quinn CJ, Schneck JL, Scott GF, Shaw AN, Waitt GM, Wooster RF and Duffy KJ. Quinoline 3-sulfonamides inhibit lactate dehydrogenase A and reverse aerobic glycolysis in cancer cells. *Cancer Metab* 2013; 1: 19.
- [61] Wang J, Wang H, Liu A, Fang C, Hao J and Wang Z. Lactate dehydrogenase A negatively regulated by miRNAs promotes aerobic glycolysis and is increased in colorectal cancer. *Oncotarget* 2015; 6: 19456-19468.