[24412] Metastasis genes relate to colon cancer

By Haotang Wei

WORD COUNT	5688	TIME SUBMITTED	11-JUN-2018 12:38PM
		PAPER ID	38009165

Comprehensive analysis of metastasis-related genes reveals a gene signature predicting the survival of colon cancer

Running title: Metastasis genes related to colon cancer

Hao-tang Wei[1#], Ji-lin Li[2#], Ming-zhi Xie[2], Rong-e Lei[3], Bang-li Hu[2]

1 Department of Gastrointestinal Surgery, the Third Affiliated Hospital of Guangxi Medical University, Nanning 530031, China

2 Department of Research, Affiliated Tumor Hospital of Guangxi Medical University, Nanning 530021, China.

3 Department of Gastroenterology, the First Affiliated Hospital of Guangxi Medical University, Nanning 530021, China

#These authors contributed equally to this work.

Corresponding author: Bang-li Hu

Department of Research,

Affiliated Tumor Hospital of Guangxi Medical University,

71 Hedi Road, Nanning 530021, China

Tel: +86-0771- 5321865; Fax +86-0771- 5321865

E-mail: gxhubl@163.com

Abstract

Objective The mechanism underlying colon cancer metastasis remain unclear. This study aimed to elucidate the genes that are altered during the metastasis of colon cancer and identify genes that are crucial to the metastasis and survival of colon cancer patients. **Methods** The dataset of primary and metastasis tissue of colon cancer, and dataset of high and low metastasis capability of colon cancer cells were selected as training cohort, and the overlapped differentially expressed genes (DEGs) were screened from the training cohort. The functional enrichment analysis for the overlapped DEGs was performed. The prognostic value of overlapped DEGs were analyzed in TCGA dataset, and a gene signature was developed using genes that related to the overall survival (OS). The prognostic value of the gene signature was further confirmed in a validation cohort. **Results** 184 overlapped DEGs were screened from the training cohort. Functional enrichment analysis revealed the significant gene functions and pathways of the overlapped DEGs. Four hub genes (OXCT1, ACTN4, IL-8, ITGA3) were identified using protein-protein network analysis. Six genes (ALDH2, NEDD9, FLNA, LBR, TWF1, SRSF1) were closely related to the OS of colon cancer patients. A gene signature was developed using these six genes based on their risk score, and the validation cohort indicated that the prognostic value of this gene signature was high in colon cancer patients. **Conclusions** Our study has demonstrate a gene profiles involved in the metastasis of colon cancer, and identify a six-gene signature that acts as an independent biomarker on the prognosis of colon cancer.

Introduction

Colorectal cancer (CRC) is one of the leading malignant cancers in the world, and the colon cancer accounts for a large part of CRC (Siegel et al. 2017a; Siegel et al. 2017b). During the last three decades, with the development of new therapies, such as antiepidermal growth factor receptor antibody therapy and bevacizumab (Castro et al. 2013; Knijn et al. 2010), great improvement in survival has been for colon cancer patients with localized- and regional disease. However, metastasis remains the main event that leads to death of a cancer patient. Compared with the early stage of colon cancer, the prognosis of patients with distant metastasis remained poor (Siegel et al. 2017b). Progressive accumulation of genetic mutations is recognized as an important factor results in the development and progression of colon cancer (Sameer 2013). Hence, exploring the mechanisms of cancer metastasis and searching suitable predictors are crucial to the diagnosis and treatment of colon cancer.

Previously, the TNM stage and pathological characteristics of colon cancer are commonly used to predict the prognosis and facilitate treatment for colon cancer patients. But there are some limitations of these methods(Marzouk & Schofield 2011). Recently, several novel biomarkers have been tested with the aim to improve the prediction of therapeutic response and prognosis of colon cancer patients (Demirkol et al. 2017; Hu et al. 2014; Xu et al. 2017), which provide a lot of help in the diagnosis and treatment of colon cancer, but the results were inconsistent and need to further study.

Thus far, metastasis is a major factor for the low survival rate of colon cancer, and the liver

is the most common site of colon cancer metastasis, but the molecular mechanism underlying distant metastasis remains unclear. Therefore, a comprehensive analysis for the molecular alteration and identify prognostic factors is pivotal for the management of metastasis colon cancer patients. In this study, by using the colon cancer data from gene expression omnibus (GEO) and The Cancer Genome Atlas (TCGA), we analyzed the data of colon cancer in primary tumor samples and liver metastasis samples to unveil the potential bridging genes that key to the development colon cancer metastasis and the potential prognostic indicators.

Materials and methods

Patient datasets

The colon cancer tissue and cells microarray data (GSE40367(Roessler et al. 2015) and GSE2509 (Provenzani et al. 2006)) was retrieval and download from the GEO (http://www.ncbi.nlm.nih.gov/geo/) database in the National Center for Biotechnology Information (NCBI) as the training cohort. The GSE40367 dataset includes seven colon adenocarcinoma (COAD) with liver metastasis species and eight COAD primary tumor species. The GSE2905 dataset includes two colon cancer cell lines (SW480: low metastasis capability and SW620: high metastasis capability). The prognostic value of genes was analyzed using the data of COAD from TCGA. To validate the results from training cohort, we used the GSE41258(Sheffer et al. 2009) dataset that includes 390 species as validation cohort. Since the data were provided by GEO and TCGA database, additional approval by an ethics committee was not needed.

Identification of overlapped DEGs

Statistical software R (version 3.4.2) and packages of Bioconductor were applied to screen differentially expressed genes (DEGs) between primary tumor tissue and liver metastasis tissue in GSE40367, the DEGs between SW480 and SW620 in GSE2905 were also screened. Genes that fulfill the criteria of p value < 0.05 and |logFC| ≥1 were defined as the DEGs. Then the intersected DEGs of GSE40367 and GSE2905 were defined as overlapped DEGs. Before screening of DEGs, the probe level data was converted into gene expression values. If one gene is corresponding to multiple probe sets, the average data were used as gene expression values(Qin et al. 2012). We also eliminated genes with over 20% missing values as previous study did(Liew et al. 2011). After pre-processing the data, t-test methods were used to screen the DEGs using limma package.

Functional enrichment analysis of overlapped DEGs

Gene Ontology (GO) includes three categories, namely, Biological Process (BP), Molecular Function (MF), and Cellular Component (CC). To investigate the functional level of DEGs, DEGs were underwent GO analysis by using DAVID (https://david.ncifcrf.gov/). The significant GO categories was defined with the $p < 0.05$. Then the KEGG pathway analysis was conducted to identify the significant pathways that genes enrichment using DAVID online tool. We defined the significant pathways with $p < 0.05$.

Integration of PPI network and subnetwork analysis

Protein–protein interaction (PPI) network can identity key genes and important gene modules which are involved in cancer development from interaction level. PPI information of DEGs was acquired from Search Tool for the Retrieval of Interacting Genes (STRING) database

(http://www.string- db.org/). Then, Cytoscape software was used for the establishment of PPI network for all the overlapped DEGs. After the establishment PPI network, the module analysis was carried out by Molecular Complex Detection (MCODE), a plug-ins of Cytoscape, to detect the gene modules of the PPI network, and the hub genes of each module was identified by the score of each gene in the module.

Acquisition of a gene signature from the training cohort

The association of overlapped DEGs with the overall survival (OS) of colon cancer patients was analyzed in COAD dataset from TCGA database using Cox regression analysis in survival package of R. The gene with p value <0.05 was considered to be independent prognostic factor. To evaluate the relative contribution of multiple genes for survival prediction when considering interrelationship among them, we selected the genes with p value <0.05 to develop a prognostic model by the risk scoring method as previously described. In brief, a prognosis risk score for predicting OS of colon cancer was established on the basis of a linear combination of the expression level multiplied by the regression coefficient derived from the multivariate cox regression model (β) with the following formula as previously reported: Risk score = expression of $Gene_1 \times \beta_1 Gene_1$ + expression of $Gene_2 \times \beta_2 Gene_2$ + …expression of $Gene_n \times \beta_n Gene_n$ (Bao et al. 2014). Using the median risk score as the cutoff, patients were divided into high risk group and low risk group. Kaplan-Meier curves were used to estimate the survival for patients with high risk group or low risk group. p <0.05 was defined as significantly different. The time-dependent receiver operating characteristic (ROC) curve analysis for the gene signature was preformed using the R package "survivalROC" (Heagerty et al. 2000). All statistical analyses were performed using R software and Bioconductor.

Results

Overlapped genes of CRC cells and tissues

The DEGs of GES2905 and GSE40367 were screened based on the selection criteria after preprocessing raw data. 341 DEGs between colon cancer cells lines SW420 and SW680 cells were identified in GES2905 dataset, and 7339 DEGs between primary and metastasis tumor specimens were identified in GSE40367 dataset. By overlapped the DEGs from the two datasets, we obtained 184 overlapped genes that differentially expressed in both CRC cells and tissues. The result was displayed in Fig. 1.

GO and KEGG enrichment analysis

The functions of the 184 overlapped DEGs were then analyzed by GO and KEGG enrichment analysis. Using the DAVID online tools, we found that the most enriched GO terms of DEGs that related to BP was Signal transduction, and the MF was Protein binding, and the CC was Cytoplasm. The KEGG pathway analysis based on the GO results revealed that Thyroid hormone synthesis was the most significant pathway of the overlapped DEGs. The results were shown in Fig. 2.

PPI network and Module screening analysis

Using the data from STRING database, a PPI network for the 184 DEGs consisting of 133 nodes and 138 edges was constructed by Cytoscape software. The overall PPI network was shown in Fig. 3A. Then the plug-ins MCODE was used to detect the modules in the network, and four modules were identified, with the OXCT1 (sore: 4), ACTN4 (sore: 3), IL-8 (sore: 3),

ITGA3 (sore: 3) as the hub genes of each module. The top three modules were shown in Fig. 3B-D.

Prognostic value of overlapped DEGs

The prognostic value of 184 overlapped DEGs was analyzed using the COAD dataset of TCGA by the multiple Cox regression analysis after adjusted the data of age, gender and TNM stage, the results showed that only ALDH2, NEDD9, FLNA, LBR, TWF1, SRSF1 were independent genes that associated with the OS of colon cancer patients, with the beta value as -1.343, -0.051, 0.492, -0.020, -0.181 and -1.938, respectively. We developed a six-gene signature by calculating the risk score of each gene, and divided the patients into high risk group and low risk group based on the median of risk score (Fig. 4A), the survival status and genes expression level in shown in Fig. 4B-C. The survival analysis revealed that this six-gene signature with high risk group predicted poor OS of colon cancer patients compared with low risk group, as the p value <0.001 (Fig. 5A). Using survival ROC analysis, we found that the risk score of this six-gene signature could largely predict the 1-, 3-, 5-year OS of colon cancer patients, as the value of area under ROC curve (AUC) was 0.686, 634, 618, respectively (Fig. 5B).

Validation cohort confirm the prognostic value of the six-gene signature

The prognostic value of the six-gene signature for the OS of colon patients was further determined in the validation cohort (GSE41258 datasets, 390 colon patients, mean follow-up 65.3 months). By using the same risk score model and cutoff value deriving from the training cohort, 390 patients of the validation cohort were classified into either high-risk group (n=195) or low-risk group (n=195). In consistence with results of the training cohort, the Kaplan-Meier

analysis of this six-gene signature indicated a significant difference between high risk group and low risk group with regard to the OS of colon cancer patients (p=0.005, log-rank test)

Discussion

Metastasis is the most lethal characteristic of CRC, accounting for 90% of the mortalities of patients with colon cancer (Li et al. 2017). The 5-year survival rates of colon cancer patients with localized- and regional -disease are up to 91.1% and 71.7%, respectively, but the 5-year survival rates of patients with distant metastasis drop to 13.3% (Siegel et al. 2017a). In addition, the metastatic dissemination of primary tumors is a pivotal cause for the failure of treatment (Deliu et al. 2014; Stein & Schlag 2007). At a molecular level, the distinct metastasis colon cancer is a heterogeneous group of diseases with molecularly and clinically distinct from the primary site of origin (Zarour et al. 2017). Thus, analyzing the molecular alteration of colon cancer with distinct metastasis is benefit for the identification of candidate targets for early diagnosis and treatment of advanced stage of colon cancer patients.

To date, some biomarkers has been identified to be the candidate targets for early diagnosis and treatment of colon cancer and rectum cancer, including genes, miRNA, lncRNA and the related signatures. The search for molecular signatures present in a primary tumor that can identify the metastasis potential of each tumor has been described with promising results. Vellinga et al(Vellinga et al. 2017) designed a lymphangiogenic gene set and applied it to large datasets of CRC, and found that this lymphangiogenic gene set was correlated with worse prognosis and consensus molecular subtype-4 in both primary and liver metastatic CRC. Rokavec et al (Rokavec et al. 2017) reported a single gene, RBM47, was down-regulation during CRC progression may promote epithelial-mesenchymal transition and metastasis. Further,

proteomic studies of exosomes from cancer cell lines have identified four candidates genes (MET, S100A8, S100A9, TNC) involved in CRC metastasis (Ji et al. 2013). Other biomarkers, such as a four -miRNA signature (let-7i, miR-10b, miR-221 and miR-320a) (Hur et al. 2015), and a six- lncRNA signature were reported to be promising biomarkers for the metastasis and prognosis of CRC (Hu et al. 2014).

In this study, we identified 4 hub genes from the subnetwork of the PPI network, that is OXCT1, ACTN4, IL-8 and ITGA3. OXCT1 is a key enzyme in ketone body metabolism that catalyzes the first and rate-determining step of ketolysis. The product of OXCT1 converts to acetyl-CoA and finally fed into the tricarboxylic acid cycle for oxidation and ATP production(Zhang & Xie 2017). The role of OXCT1 has been defined in several cancers, included colorectal cancer, and the OXCT1 expressed at higher levels in the metastatic CRC cell line CC-M3 (Lee et al. 2016). ACTN4 is a non-muscle-type alpha-actinin, it play an important role in regulating cytoskeleton organization and involving transcriptional regulation of gene expression. ACTN4 encodes a nonmuscle, alpha actinin isoform which is concentrated in the cytoplasm, and involves in metastatic processes, and was reported to enhance cancer cell motility, invasion, and metastasis. In CRC, ACTN4 was reported to promote CRC cell line invasion by suppressing focal adhesion maturation (Fukumoto et al. 2015). ITGA3, a member of the integrin family, joins a beta 1 subunit to form an intact integrin and interacts with extracellular matrix proteins, including members of the laminin family(Nagata et al. 2013). ITGA3 levels represent biological traits associated with lymphatic dissemination and local invasiveness. One study showed that ITGA3 was over-expressed in stages III versus I of CRC patients, and related to the OS and disease-free survival (Linhares et al. 2015). IL-8 is a pro-inflammatory chemokine produced by various cell types to recruit leukocytes to sites of infection or tissue injury.

Emerging research now indicates that paracrine signaling by tumor-derived IL-8 promotes the trafficking of neutrophils and myeloid-derived suppressor cells into the tumor microenvironment, which have the ability to dampen anti-tumor immune responses (David et al. 2016). Lambrechts et al (Lambrechts et al. 2015) observed that IL-8 plasma levels at baseline and subsequent increases in IL-8 were associated with worse progression-free survival of metastatic CRC patients. These studies confirmed our results that the hub genes of the PPI network were crucial to the metastasis of colon cancer, but further studies remains warrant to determine the underlying mechanism.

Similar to the previous studies, in this study, a signature constructed by six genes was showed to be good predictor for the OS of colon cancer, among these six genes, the role of ALDH2, NEDD9, SRSF1 and FLNA were reported to be associated with the CRC in several studies. ALDH2 is essential for the metabolism and detoxification of a wide range of endogenous and exogenous aldehyde substrates. it is the rate-limiting enzyme in the ethanol metabolism, oxidizing acetaldehyde to acetic acid both in the liver and other tissues (Chen et al. 2016). As a novel biological marker, ALDH2 displays an attractive prospect in the screening, diagnosis and evaluation of the prognosis of many diseases, and the genetic polymorphism of ALDH2 was significantly correlated with the susceptibility to CRC (Li et al. 2016). NEDD9 is a non-catalytic scaffolding protein, assembles complexes involving oncogenic kinases, and regulates the magnitude and duration of cell signaling cascades that controls multiple processes, which are crucial for tumorigenesis and metastases(Shagisultanova et al. 2015). Study has showed that downregulation of NEDD9 by apigenin can suppresses migration, invasion, and metastasis of CRC cells (Dai et al. 2016). With regard to the SRSF1, there was a study reported that phosphorylation of SRSF1 regulated alternative splicing of tumor-related Rac1b in CRC

cells(Goncalves et al. 2014). SRSF1 is a prototypical splicing factor mostly recruited to SREs classified as exonic splicing enhancers. SRSF1 recognizes degenerate purine-rich sequence motifs and its binding promotes recognition of both constitutive and alternative exons during spliceosomal assembly (Sanford et al. 2009). FLNA is an actin-binding protein expressed ubiquitously within the body with multiple roles both in cell signaling and maintenance of cell shape and motility. A well-known association already exists between this mutation and disorders of neuronal migration, vascular function, connective tissue integrity, and skeletal development(Shelmerdine et al. 2017); FLNA showed low expression in CRC, and was high correlated with the incidence and development of CRC (Tian et al. 2015). LBR a transmembrane protein of the inner nuclear membrane. LBR interacts through its nucleoplasmic amino-terminal domain with both heterochromatin and B-type lamins, and is phosphorylated throughout the cell cycle, but on different sites in interphase and mitosis (Duband-Goulet et al. 1998). TWF1 is a conserved actin-binding protein with two actin depolymerizing factor homology domains16 and belongs to the ADF-H family. It regulates diverse morphological and motile processes by both sequestering ADP-actin monomers and capping filament barbed ends(Paavilainen et al. 2007). However, no study has yet reported the role of LBR and TWF1 in CRC, so their role need be further studied in colon cancers.

Compared with previous studies, this study screened the metastasis related genes by overlapping the DEGs from cancer tissues and cell lines. These tissues included primary colon cancer tissues and liver metastasis tissues, and the cell lines included primary (SW480) and metastatic (SW620) human isogenic colorectal cancer cell lines, thus, the overlapped DEGs could be more reliable in reflecting the genes alteration of metastatic colon cancer. We also performed a comprehensive analysis for these overlapped DEGs, using GO and KEGG analysis,

277 and identified the function of the genes and the pathways they involved, then the PPI network
278 and subnetwork analysis revealed the gene-gene interaction and identified four hub genes that
279 crucial to the network, which provided an insight into the mechanism of colon cancer metastasis.
280 Furthermore, we analyzed the prognostic value of the overlapped DEGs, and identified six genes
281 that related to the OS of colon cancer patients. We finally developed a six-gene signature to test
282 its prognostic value and validated by an independent dataset. Therefore, these results were more
283 informative and provided a reliable prognostic factors pinpointing a subset of patients with poor
284 prognosis.

285 However, some limitations need to be noted in this study. First, although metastasis-related
286 genes of colon cancer were identified and the prognostic value of them were validated in our
287 study, the results were calculated from microarray or RNA-sequencing technique datasets, thus,
288 lack of functional validation of the target genes is one of the major limitations of this study.
289 Therefore, a thorough functional experiments for these genes and corresponding downstream
290 events to reveal novel diagnostic and therapeutic targets for colon cancer is necessary. Second,
291 the development of colon cancer metastasis can be caused by many factors, such as KRAS,
292 BRAF mutation, microsatellite instability, which has been proven to be closely related to the
293 colon cancer, but due to the limited of datasets, we did not perform stratified analysis based on
294 these factors, future studies should conduct this analysis to explore the difference under different
295 conditions. Third, the mean time of follow-up in validation cohort was 65.3 months, thus, a study
296 including a longer follow-up time is warranted to validate our results in the future.

297 In conclusion, this study has screened a gene profile involving in the metastasis of colon
298 cancer, and identify four hub genes from the gene profile. We also identify and validate a six-
299 gene signature that can be served as an indicator of prognosis of colon cancer. Some genes that

are not yet proved to be associated with colon cancer metastasis may represent new therapeutic targets.

Abbreviation:

CRC: colorectal cancer; DEGs: differentially expressed genes; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; OS: overall survival; ROC: receiver operating characteristic curve; OXCT1: 3-oxoacid CoA-transferase 1; ACTN4: actinin alpha 4; IL-8: interleukin 8; ITGA3: integrin subunit alpha 3; ALDH2: aldehyde dehydrogenase 2; NEDD9: neural precursor cell expressed, developmentally down-regulated 9; FLNA: filamin A; LBR: lamin B receptor; TWF1: twinfilin actin binding protein 1; SRSF1: serine and arginine rich splicing factor 1; TNC: tenascin C; RBM47: RNA binding motif protein 47; MET: MET proto-oncogene, receptor tyrosine kinase

Reference

Bao ZS, Li MY, Wang JY, Zhang CB, Wang HJ, Yan W, Liu YW, Zhang W, Chen L, and Jiang T. 2014. Prognostic value of a nine-gene signature in glioma patients based on mRNA expression profiling. *CNS Neurosci Ther* 20:112-118. 10.1111/cns.12171

Castro E, Goh C, Olmos D, Saunders E, Leongamornlert D, Tymrakiewicz M, Mahmud N, Dadaev T, Govindasami K, Guy M, Sawyer E, Wilkinson R, Ardern-Jones A, Ellis S, Frost D, Peock S, Evans DG, Tischkowitz M, Cole T, Davidson R, Eccles D, Brewer C, Douglas F, Porteous ME, Donaldson A, Dorkins H, Izatt L, Cook J, Hodgson S, Kennedy MJ, Side LE, Eason J, Murray A, Antoniou AC, Easton DF, Kote-Jarai Z, and Eeles R. 2013. Germline BRCA mutations are associated with higher risk of nodal involvement,

distant metastasis, and poor survival outcomes in prostate cancer. *J Clin Oncol* 31:1748-1757. 10.1200/JCO.2012.43.1882

Chen CH, Joshi AU, and Mochly-Rosen D. 2016. The Role of Mitochondrial Aldehyde Dehydrogenase 2 (ALDH2) in Neuropathology and Neurodegeneration. *Acta Neurol Taiwan* 25(4):111-123. 10196099/254111

Dai J, Van Wie PG, Fai LY, Kim D, Wang L, Poyil P, Luo J, and Zhang Z. 2016. Downregulation of NEDD9 by apigenin suppresses migration, invasion, and metastasis of colorectal cancer cells. *Toxicol Appl Pharmacol* 311:106-112. 10.1016/j.taap.2016.09.016

David JM, Dominguez C, Hamilton DH, and Palena C. 2016. The IL-8/IL-8R Axis: A Double Agent in Tumor Immune Resistance. *Vaccines (Basel)* 4. 10.3390/vaccines4030022

Deliu IC, Georgescu EF, and Bezna MC. 2014. Analysis of prognostic factors in colorectal carcinoma. *Rev Med Chir Soc Med Nat Iasi* 118:808-816.

Demirkol S, Gomceli I, Isbilen M, Dayanc BE, Tez M, Bostanci EB, Turhan N, Akoglu M, Ozyerli E, Durdu S, Konu O, Nissan A, Gonen M, and Gure AO. 2017. A Combined ULBP2 and SEMA5A Expression Signature as a Prognostic and Predictive Biomarker for Colon Cancer. *J Cancer* 8:1113-1122. 10.7150/jca.17872

Duband-Goulet I, Courvalin JC, and Buendia B. 1998. LBR, a chromatin and lamin binding protein from the inner nuclear membrane, is proteolyzed at late stages of apoptosis. *J Cell Sci* 111 (Pt 10):1441-1451.

Fukumoto M, Kurisu S, Yamada T, and Takenawa T. 2015. alpha-Actinin-4 enhances colorectal cancer cell invasion by suppressing focal adhesion maturation. *PLoS One* 10:e0120616. 10.1371/journal.pone.0120616

346 Goncalves V, Henriques AF, Pereira JF, Neves Costa A, Moyer MP, Moita LF, Gama-Carvalho
347 M, Matos P, and Jordan P. 2014. Phosphorylation of SRSF1 by SRPK1 regulates
348 alternative splicing of tumor-related Rac1b in colorectal cells. *RNA* 20:474-482.
349 10.1261/rna.041376.113
350 Heagerty PJ, Lumley T, and Pepe MS. 2000. Time-dependent ROC curves for censored survival
351 data and a diagnostic marker. *Biometrics* 56:337-344.
352 Hu Y, Chen HY, Yu CY, Xu J, Wang JL, Qian J, Zhang X, and Fang JY. 2014. A long non-
353 coding RNA signature to improve prognosis prediction of colorectal cancer. *Oncotarget*
354 5:2230-2242. 1895 10.18632/oncotarget.1895
355 Hur K, Toiyama Y, Schetter AJ, Okugawa Y, Harris CC, Boland CR, and Goel A. 2015.
356 Identification of a metastasis-specific MicroRNA signature in human colorectal cancer. *J*
357 *Natl Cancer Inst* 107. 10.1093/jnci/dju492
358 Ji H, Greening DW, Barnes TW, Lim JW, Tauro BJ, Rai A, Xu R, Adda C, Mathivanan S, Zhao
359 W, Xue Y, Xu T, Zhu HJ, and Simpson RJ. 2013. Proteome profiling of exosomes
360 derived from human primary and metastatic colorectal cancer cells reveal differential
361 expression of key metastatic factors and signal transduction components. *Proteomics*
362 13:1672-1686. 10.1002/pmic.201200562
363 Knijn N, Tol J, and Punt CJ. 2010. Current issues in the targeted therapy of advanced colorectal
364 cancer. *Discov Med* 9:328-336.
365 Lambrechts D, Thienpont B, Thuillier V, Sagaert X, Moisse M, Peuteman G, Pericay C,
366 Folprecht G, Zalcberg J, Zilocchi C, Margherini E, Chiron M, and Van Cutsem E. 2015.
367 Evaluation of efficacy and safety markers in a phase II study of metastatic colorectal
368 cancer treated with aflibercept in the first-line setting. *Br J Cancer* 113:1027-1034.

10.1038/bjc.2015.329

Lee CL, Huang CJ, Yang SH, Chang CC, Huang CC, Chien CC, and Yang RN. 2016. Discovery of genes from feces correlated with colorectal cancer progression. *Oncol Lett* 12:3378-3384. 10.3892/ol.2016.5069

Li H, Chen YX, Wen JG, and Zhou HH. 2017. Metastasis-associated in colon cancer 1: A promising biomarker for the metastasis and prognosis of colorectal cancer. *Oncol Lett* 14:3899-3908. 10.3892/ol.2017.6670

Li R, Zhao Z, Sun M, Luo J, and Xiao Y. 2016. ALDH2 gene polymorphism in different types of cancers and its clinical significance. *Life Sci* 147:59-66. 10.1016/j.lfs.2016.01.028

Liew AW, Law NF, and Yan H. 2011. Missing value imputation for gene expression data: computational techniques to recover missing data from available information. *Brief Bioinform* 12:498-513. 10.1093/bib/bbq080

Linhares MM, Affonso RJ, Jr., Viana Lde S, Silva SR, Denadai MV, de Toledo SR, and Matos D. 2015. Genetic and Immunohistochemical Expression of Integrins ITGAV, ITGA6, and ITGA3 As Prognostic Factor for Colorectal Cancer: Models for Global and Disease-Free Survival. *PLoS One* 10:e0144333. 10.1371/journal.pone.0144333

Marzouk O, and Schofield J. 2011. Review of histopathological and molecular prognostic features in colorectal cancer. *Cancers (Basel)* 3:2767-2810. 10.3390/cancers3022767

Nagata M, Noman AA, Suzuki K, Kurita H, Ohnishi M, Ohyama T, Kitamura N, Kobayashi T, Uematsu K, Takahashi K, Kodama N, Kawase T, Hoshina H, Ikeda N, Shingaki S, and Takagi R. 2013. ITGA3 and ITGB4 expression biomarkers estimate the risks of locoregional and hematogenous dissemination of oral squamous cell carcinoma. *BMC Cancer* 13:410. 10.1186/1471-2407-13-410

392 Paavilainen VO, Hellman M, Helfer E, Bovellan M, Annila A, Carlier MF, Permi P, and
393 Lappalainen P. 2007. Structural basis and evolutionary origin of actin filament capping
394 by twinfilin. *Proc Natl Acad Sci U S A* 104:3113-3118. 10.1073/pnas.0608725104
395 Provenzani A, Fronza R, Loreni F, Pascale A, Amadio M, and Quattrone A. 2006. Global
396 alterations in mRNA polysomal recruitment in a cell model of colorectal cancer
397 progression to metastasis. *Carcinogenesis* 27:1323-1333. 10.1093/carcin/bgi377
398 Qin S, Kim J, Arafat D, and Gibson G. 2012. Effect of normalization on statistical and biological
399 interpretation of gene expression profiles. *Front Genet* 3:160. 10.3389/fgene.2012.00160
400 Roessler S, Lin G, Forgues M, Budhu A, Hoover S, Simpson RM, Wu X, He P, Qin LX, Tang
401 ZY, Ye QH, and Wang XW. 2015. Integrative genomic and transcriptomic
402 characterization of matched primary and metastatic liver and colorectal carcinoma. *Int J*
403 *Biol Sci* 11:88-98. 10.7150/ijbs.10583
404 Rokavec M, Kaller M, Horst D, and Hermeking H. 2017. Pan-cancer EMT-signature identifies
405 RBM47 down-regulation during colorectal cancer progression. *Sci Rep* 7:4687.
406 10.1038/s41598-017-04234-2
407 Sameer AS. 2013. Colorectal cancer: molecular mutations and polymorphisms. *Front Oncol*
408 3:114. 10.3389/fonc.2013.00114
409 Sanford JR, Wang X, Mort M, Vanduyn N, Cooper DN, Mooney SD, Edenberg HJ, and Liu Y.
410 2009. Splicing factor SFRS1 recognizes a functionally diverse landscape of RNA
411 transcripts. *Genome Res* 19:381-394. 10.1101/gr.082503.108
412 Shagisultanova E, Gaponova AV, Gabbasov R, Nicolas E, and Golemis EA. 2015. Preclinical
413 and clinical studies of the NEDD9 scaffold protein in cancer and other diseases. *Gene*
414 567:1-11. 10.1016/j.gene.2015.04.086

415 Sheffer M, Bacolod MD, Zuk O, Giardina SF, Pincas H, Barany F, Paty PB, Gerald WL,
416 Notterman DA, and Domany E. 2009. Association of survival and disease progression
417 with chromosomal instability: a genomic exploration of colorectal cancer. *Proc Natl*
418 *Acad Sci U S A* 106:7131-7136. 10.1073/pnas.0902232106
419 Shelmerdine SC, Semple T, Wallis C, Aurora P, Moledina S, Ashworth MT, and Owens CM.
420 2017. Filamin A (FLNA) mutation-A newcomer to the childhood interstitial lung disease
421 (ChILD) classification. *Pediatr Pulmonol* 52:1306-1315. 10.1002/ppul.23695
422 Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RGS, Barzi A, and Jemal A. 2017a.
423 Colorectal cancer statistics, 2017. *CA Cancer J Clin* 67:177-193. 10.3322/caac.21395
424 Siegel RL, Miller KD, and Jemal A. 2017b. Cancer Statistics, 2017. *CA Cancer J Clin* 67:7-30.
425 10.3322/caac.21387
426 Stein U, and Schlag PM. 2007. Clinical, biological, and molecular aspects of metastasis in
427 colorectal cancer. *Recent Results Cancer Res* 176:61-80.
428 Tian ZQ, Shi JW, Wang XR, Li Z, and Wang GY. 2015. New cancer suppressor gene for
429 colorectal adenocarcinoma: filamin A. *World J Gastroenterol* 21:2199-2205.
430 10.3748/wjg.v21.i7.2199
431 Vellinga TT, Kranenburg O, Frenkel N, Ubink I, Marvin D, Govaert K, van Schelven S,
432 Hagendoorn J, and Borel Rinkes IH. 2017. Lymphangiogenic Gene Expression Is
433 Associated With Lymph Node Recurrence and Poor Prognosis After Partial Hepatectomy
434 for Colorectal Liver Metastasis. *Ann Surg*. 10.1097/SLA.0000000000002430
435 Xu G, Zhang M, Zhu H, and Xu J. 2017. A 15-gene signature for prediction of colon cancer
436 recurrence and prognosis based on SVM. *Gene* 604:33-40. 10.1016/j.gene.2016.12.016
437 Zarour LR, Anand S, Billingsley KG, Bisson WH, Cercek A, Clarke MF, Coussens LM, Gast CE,

438 Geltzeiler CB, Hansen L, Kelley KA, Lopez CD, Rana SR, Ruhl R, Tsikitis VL, Vaccaro
439 GM, Wong MH, and Mayo SC. 2017. Colorectal Cancer Liver Metastasis: Evolving
440 Paradigms and Future Directions. *Cell Mol Gastroenterol Hepatol* 3:163-173.
441 10.1016/j.jcmgh.2017.01.006
442 Zhang S, and Xie C. 2017. The role of OXCT1 in the pathogenesis of cancer as a rate-limiting
443 enzyme of ketone body metabolism. *Life Sci* 183:110-115. 10.1016/j.lfs.2017.07.003
444
445
446

[24412] Metastasis genes relate to colon cancer

ORIGINALITY REPORT

37%

SIMILARITY INDEX

PRIMARY SOURCES

1	www.spandidos-publications.com Internet	122 words — 3%
2	Jiang-Hui Zeng, Liang Liang, Rong-Quan He, Rui-Xue Tang, Xiao-Yong Cai, Jun-Qiang Chen, Dian-Zhong Luo, Gang Chen. "Comprehensive investigation of a novel differentially expressed lncRNA expression profile signature to assess the survival of patients with colorectal adenocarcinoma", Oncotarget, 2017 Crossref	111 words — 3%
3	Meng Zhou, Yanying Sun, Yifan Sun, Wanying Xu, Zhaoyue Zhang, Hengqiang Zhao, Zhaohua Zhong, Jie Sun. "Comprehensive analysis of lncRNA expression profiles reveals a novel lncRNA signature to discriminate nonequivalent outcomes in patients with ovarian cancer", Oncotarget, 2016 Crossref	93 words — 2%
4	Chaoyang Zhang, Li Peng, Yaqin Zhang, Zhaoyang Liu, Wenling Li, Shilian Chen, Guancheng Li. "The identification of key genes and pathways in hepatocellular carcinoma by bioinformatics analysis of high-throughput data", Medical Oncology, 2017 Crossref	69 words — 2%
5	mdpi.com Internet	58 words — 1%
6	www.biorxiv.org Internet	48 words — 1%
7	www.dovepress.com Internet	47 words — 1%

8	Susan C. Shelmerdine, Thomas Semple, Colin Wallis, Paul Aurora, Shahin Moledina, Michael T. Ashworth, Catherine M. Owens. " Filamin A ( ) mutation-A newcomer to the childhood interstitial lung disease (ChILD) classification ", Pediatric Pulmonology, 2017 Crossref	46 words — 1%
9	Song Zhang, Caifeng Xie. "The role of OXCT1 in the pathogenesis of cancer as a rate-limiting enzyme of ketone body metabolism", Life Sciences, 2017 Crossref	40 words — 1%
10	jcs.biologists.org Internet	39 words — 1%
11	preview-wjso.biomedcentral.com Internet	37 words — 1%
12	repositorio.insa.pt Internet	36 words — 1%
13	Bo Gao, Tian Yu, Dongbo Xue, Boshi Sun, Qin Shao, Hani Choudhry, Victoria Marcus, Jiannis Ragoussis, Yuguo Zhang, Weihui Zhang, Zu-hua Gao. "A multidimensional integration analysis reveals potential bridging targets in the process of colorectal cancer liver metastasis", PLOS ONE, 2017 Crossref	36 words — 1%
14	www.oncotarget.com Internet	32 words — 1%
15	wjso.biomedcentral.com Internet	30 words — 1%
16	Jiang, Qian, You-Cheng Yu, Xiao-Jun Ding, Yin Luo, and Hong Ruan. "Bioinformatics Analysis Reveals Significant Genes and Pathways to Targetfor Oral Squamous Cell Carcinoma", Asian Pacific Journal of Cancer Prevention, 2014. Crossref	27 words — 1%

17	Fleshman, . "RNA expression of the molecular signature genes for metastasis in colorectal cancer", Oncology Reports, 2011. Crossref	25 words — 1%
18	Sun, Baoyong, Fangxin Wang, Min Li, and Mingshan Yang. "Identifications of genetic differences between metastatic and non-metastatic osteosarcoma samples based on bioinformatics analysis", Medical Oncology, 2015. Crossref	25 words — 1%
19	Thomas T. Vellinga, Onno Kranenburg, Nicola Frenkel, Inge Ubink et al. "Lymphangiogenic Gene Expression Is Associated With Lymph Node Recurrence and Poor Prognosis After Partial Hepatectomy for Colorectal Liver Metastasis", Annals of Surgery, 2017 Crossref	24 words — 1%
20	www.nature.com Internet	21 words — 1%
21	dknet.org Internet	21 words — 1%
22	Hao-Yuan Wang, Ji-Ye Li, Xiu Liu, Xiao-Yan Yan et al. "A three ion channel genes-based signature predicts prognosis of primary glioblastoma patients and reveals a chemotherapy sensitive subtype", Oncotarget, 2016 Crossref	19 words — < 1%
23	Bockhorn, Jessica, Rachel Dalton, Chika Nwachukwu, Simo Huang, Aleix Prat, Kathy Yee, Ya-Fang Chang, Dezheng Huo, Yujia Wen, Kaitlin E. Swanson, Tyler Qiu, Jun Lu, Seo Young Park, M. Eileen Dolan, Charles M. Perou, Olufunmilayo I. Olopade, Michael F. Clarke, Geoffrey L. Greene, and Huiping Liu. "MicroRNA-30c inhibits human breast tumour chemotherapy resistance by regulating TWF1 and IL-11", Nature Communications, 2013. Crossref	18 words — < 1%
24	Elena Shagisultanova, Anna V. Gaponova, Rashid Gabbasov, Emmanuelle Nicolas, Erica A. Golemis.	18 words — < 1%

"Preclinical and clinical studies of the NEDD9 scaffold protein in cancer and other diseases", Gene, 2015
Crossref

25 Huang, Xin-Ping, Jin Hou, Xiao-Yun Shen, Chao-Yuan Huang, Xue-Hui Zhang, Yu-An Xie, and Xiao-Ling Luo. "Microrna-486-5p, downregulated in hepatocelluar carcinoma, suppresses tumor growth by targeting p85α", FEBS Journal, 2014.
Crossref
18 words — < 1%

26 Mingwei Chen, Baoquan Liu, Jianbing Xiao, Yingnan Yang, Yafang Zhang. "A novel seven-long non-coding RNA signature predicts survival in early stage lung adenocarcinoma", Oncotarget, 2017
Crossref
15 words — < 1%

27 pdfs.semanticscholar.org
Internet
15 words — < 1%

28 Luai R. Zarour, Sudarshan Anand, Kevin G. Billingsley, William H. Bisson et al. "Colorectal Cancer Liver Metastasis: Evolving Paradigms and Future Directions", Cellular and Molecular Gastroenterology and Hepatology, 2017
Crossref
14 words — < 1%

29 Yanli Wang, Ke Lv, Mei Zhao, Fengji Liang et al. "Expression profiles and functional annotation analysis of mRNAs in suprachiasmatic nucleus of Clock mutant mice", Gene, 2018
Crossref
14 words — < 1%

30 Workenhe, S.T.. "Infectious salmon anaemia virus (ISAV) isolates induce distinct gene expression responses in the Atlantic salmon (Salmo salar) macrophage/dendritic-like cell line TO, assessed using genomic techniques", Molecular Immunology, 200909
Crossref
14 words — < 1%

31 www.infona.pl
Internet
13 words — < 1%

32	"Encyclopedia of Signaling Molecules", Springer Nature, 2018 Crossref	13 words — < 1%
33	Guangru Xu, Minghui Zhang, Hongxing Zhu, Jinhua Xu. "A 15-gene signature for prediction of colon cancer recurrence and prognosis based on SVM", Gene, 2017 Crossref	12 words — < 1%
34	NA. "Supplement 2, Proceedings of the 14th World Conference on Lung Cancer, Book 2", Journal of Thoracic Oncology, 2011 Crossref	12 words — < 1%
35	www.wjgnet.com Internet	12 words — < 1%
36	f1000.com Internet	12 words — < 1%
37	www.jove.com Internet	11 words — < 1%
38	Upender Manne, Trafina Jadhav, Balananda-Dhurjati Kumar Putcha, Temesgen Samuel, Shivani Soni, Chandrakumar Shanmugam, Esther A. Suswam. "Molecular Biomarkers of Colorectal Cancer and Cancer Disparities: Current Status and Perspective", Current Colorectal Cancer Reports, 2016 Crossref	11 words — < 1%
39	C Criscitiello, M A Bayar, G Curigliano, F W Symmans et al. "A gene signature to predict high tumor-infiltrating lymphocytes after neoadjuvant chemotherapy and outcome in patients with triple-negative breast cancer", Annals of Oncology, 2017 Crossref	11 words — < 1%
40	Chenhua Sun, Qi Yuan, Dongdong Wu, Xiaohu Meng, Baolin Wang. "Identification of core genes and outcome in gastric cancer using bioinformatics analysis",	11 words — < 1%

Oncotarget, 2017
Crossref

41	spandidos-publications.com Internet	11 words — < 1%
42	ub-ed.ub.uni-greifswald.de Internet	10 words — < 1%
43	biblioteca.universia.net Internet	10 words — < 1%
44	www.science.gov Internet	10 words — < 1%
45	Sven H. Loosen, Frank Tacke, Marcel Binnebösel, Catherine Leyh et al. "Serum levels of soluble urokinase plasminogen activator receptor (suPAR) predict outcome after resection of colorectal liver metastases", Oncotarget, 2018 Crossref	10 words — < 1%
46	Xianglong Tian, Xiaoqiang Zhu, Tingting Yan, Chenyang Yu, Chaoqin Shen, Ye Hu, Jie Hong, Haoyan Chen, Jing-Yuan Fang. "Recurrence-associated gene signature optimizes recurrence-free survival prediction of colorectal cancer", Molecular Oncology, 2017 Crossref	10 words — < 1%
47	SÃ¸rensen, Karina Dalsgaard. "Predicting the future of advanced prostate cancer", The Lancet Oncology, 2012. Crossref	10 words — < 1%
48	open.uct.ac.za Internet	9 words — < 1%
49	rnajournal.cshlp.org Internet	9 words — < 1%
50	Lindsey E Montefiori, Debora R Sobreira, Noboru J Sakabe, Ivy Aneas et al. "A promoter interaction	9 words — < 1%

map for cardiovascular disease genetics", Cold Spring Harbor Laboratory, 2018
Crossref Posted Content

51	Nagata, Masaki, Arhab A Noman, Kenji Suzuki, Hiroshi Kurita, Makoto Ohnishi, Tokio Ohyama, Nobutaka Kitamura, Takanori Kobayashi, Kohya Uematsu, Katsu Takahashi, Naoki Kodama, Tomoyuki Kawase, Hideyuki Hoshina, Nobuyuki Ikeda, Susumu Shingaki, and Ritsuo Takagi. "ITGA3 and ITGB4 expression biomarkers estimate the risks of locoregional and hematogenous dissemination of oral squamous cell carcinoma", BMC Cancer, 2013. Crossref	9 words — < 1%
52	Hur, K., Y. Toiyama, A. J. Schetter, Y. Okugawa, C. C. Harris, C. R. Boland, and A. Goel. "Identification of a Metastasis-Specific MicroRNA Signature in Human Colorectal Cancer", JNCI Journal of the National Cancer Institute, 2015. Crossref	9 words — < 1%
53	ar.iiarjournals.org Internet	9 words — < 1%
54	bmccancer.biomedcentral.com Internet	8 words — < 1%
55	Xin Zhou, Zebo Huang, Lei Xu, Mingxia Zhu, Lan Zhang, Huo Zhang, Xiaping Wang, Hai Li, Wei Zhu, Yongqian Shu, Ping Liu. "A panel of 13-miRNA signature as a potential biomarker for predicting survival in pancreatic cancer", Oncotarget, 2016 Crossref	8 words — < 1%
56	www.ijcep.com Internet	8 words — < 1%
57	repositorio.uam.es Internet	8 words — < 1%
58	Rong Xu, Alin Rai, Maoshan Chen, Wittaya Suwakulsiri, David W. Greening, Richard J.	8 words — < 1%

Simpson. "Extracellular vesicles in cancer — implications for future improvements in cancer care", Nature Reviews Clinical Oncology, 2018
Crossref

59 Jun Wang, Peng Gao, Jingxu Sun, Jinxin Shi, Zhonghua Wu, Xi Zhong, Yongxi Song, Zhenning Wang. "A six-gene prognostic predictor for patients with gastric cancer", Oncotarget, 2018 — 8 words — < 1%
Crossref

60 Shun Liu, Tian-Bo Yang, Yue-Li Nan, An-Hua Li, Dong-Xiang Pan, Yang Xu, Shu Li, Ting Li, Xiao-Yun Zeng, Xiao-Qiang Qiu. "Genetic variants of cell cycle pathway genes predict disease-free survival of hepatocellular carcinoma", Cancer Medicine, 2017 — 8 words — < 1%
Crossref

61 Okugawa, Yoshinaga, Yuji Toiyama, Yasuhiko Mohri, Koji Tanaka, Mikio Kawamura, Junichiro Hiro, Toshimitsu Araki, Yasuhiro Inoue, Chikao Miki, and Masato Kusunoki. "Elevated serum concentration of monocyte chemotactic protein 4 (MCP-4) as a novel non-invasive prognostic and predictive biomarker for detection of metastasis in colorectal cancer : Serum MCP-4 in Colorectal Cancer", Journal of Surgical Oncology, 2016. — 8 words — < 1%
Crossref

62 onlinelibrary.wiley.com — 8 words — < 1%
Internet

63 Han, Cheng-Bo Sun, Li Ma, Jie-Tao Li, Ya. "The influence of mtDNA deletion on lung cancer cells under the conditions of hypoxia and irradiation", Lung, Dec 2014 Issue — 8 words — < 1%
Publications

64 Zhou, Meng, Maoni Guo, Dongfeng He, Xiaojun Wang, Yinqiu Cui, Haixiu Yang, Dapeng Hao, and Jie Sun. "A potential signature of eight long non-coding RNAs predicts survival in patients with non-small cell lung cancer", Journal of Translational Medicine, 2015. — 7 words — < 1%
Crossref

65 Dai, Jin, Peter G. Van Wie, Leonard Yenwong Fai, Donghern Kim, Lei Wang, Pratheeshkumar Poyil, Jia Luo, and Zhuo Zhang. "Downregulation of NEDD9 by apigenin suppresses migration, invasion, and metastasis of colorectal cancer cells", Toxicology and Applied Pharmacology, 2016.
Crossref
6 words — < 1%

66 Li, Rui, Zihan Zhao, Mingyang Sun, Jiachi Luo, and Yechen Xiao. "ALDH2 gene polymorphism in different types of cancers and its clinical significance", Life Sciences, 2016.
Crossref
6 words — < 1%

67 Recent Results in Cancer Research, 2016.
Crossref
6 words — < 1%

68 Bin Xiao, Lidan Chen, Yongli Ke, Jianfeng Hang et al. "Identification of methylation sites and signature genes with prognostic value for luminal breast cancer", BMC Cancer, 2018
Crossref
6 words — < 1%

69 Min, Weijie, Yanan Li, Yihui Zhang, Dongwei Dai, Yiqun Cao, Zhijian Yue, and Jianmin Liu. "Role of the anti-glioma drug AT13148 in the inhibition of Notch signaling pathway", Gene, 2015.
Crossref
6 words — < 1%

70 "Non-coding RNAs in Colorectal Cancer", Springer Nature, 2016
Crossref
6 words — < 1%

71 Huarong Chen, Xiaoqiang Sun, Weiting Ge, Yun Qian, Rui Bai, Shu Zheng. "A seven-gene signature predicts overall survival of patients with colorectal cancer", Oncotarget, 2016
Crossref
6 words — < 1%

72 "Abstract", Hepatology International, 06/2008
Crossref
6 words — < 1%

EXCLUDE QUOTES	OFF	EXCLUDE MATCHES	OFF
EXCLUDE BIBLIOGRAPHY	ON		