Identification of Significant Genes with poor prognosis in Liver Metastasis of Colorectal Cancer via Bioinformatical Analysis

Junsheng Chen

Changzhi Medical College Affiliated Heping Hospital

Hongzhou Liu (✉ cjs5388@163.com)

Changzhi Medical College Affiliated Heping Hospital

Research

Keywords: Bioinformatical analysis, Colorectal Cancer, Liver Metastasis, Different expressed gene

Posted Date: September 22nd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-78881/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Colorectal cancer (CRC) is a common malignant tumor in the world wide, and more than 50% patients have liver metastases.

Purpose: The purpose of this study is to identify significant genes with poor outcome and the underlying mechanisms of CRC liver metastases.

Methods: Gene expression profiles of GSE50760, GSE41568 and GSE14297 are available on website of GEO database. Differentially expressed genes (DEGs) between CRC liver metastases and primary tissues were picked out by GEO2R tool and Venn diagram software. The Database for Annotation, Visualization and Integrated Discovery (DAVID) was used to analyze Kyoto Encyclopedia of Gene and Genome (KEGG) pathway and gene ontology (GO). Then protein-protein interaction (PPI) of these DEGs was visualized by Cytoscape with Search Tool for the Retrieval of Interacting Genes (STRING).

Results: There were total of 147 consistently expressed genes in the three datasets, including 123 up-regulated genes and 24 down-regulated genes enriched in complement and coagulation cascades, drug metabolism-cytochrome P450, metabolism of xenobiotics by cytochrome P450, prion diseases, chemical carcinogenesis, staphylococcus aureus infection and linoleic acid metabolism. Of PPI network analyzed by Molecular Complex Detection (MCODE) plug-in, all 39 genes were selected. Moreover, for the analysis of CRC survival among those genes, Kaplan–Meier analysis was implemented and 4 (SERPING1 ITIH2 CDH2 APOE) of 39 genes had a significantly worse prognosis.

Conclusion: we have identified four significant DEGs with poor prognosis in CRC liver metastases on the basis of integrated bioinformatical methods, which could be potential therapeutic targets for CRC patients with liver metastases.

Introduction

Colorectal cancer (CRC) is the third most common cancer and one of the leading causes of cancer-related death worldwide (Siegel, Miller et al. 2018). Liver metastasis is a major cause of death in patients with cancer including CRC. According to the statistics, nearly 20% of patients already have liver metastasis when they are diagnosed with colorectal cancer and more than 50% develop metastases during their disease course (Fakih 2015, Zarour, Anand et al. 2017). Although some potential genes and mechanisms related to liver metastasis of colorectal cancer have been found (Huang, Tan et al. 2018, Pretzsch, Bosch et al. 2019), the treatment of liver metastasis of colorectal cancer has not been well developed. Therefore, we should find more reliable targets to improve the treatment effect and to better understand the mechanism of CRC metastasis.

With the development of gene chip and RNA sequencing technology, we can get many gene expression data from public databases more accurately and quickly (Vogelstein, Papadopoulos et al. 2013). In

addition, the application of bioinformatics in the field of colorectal cancer in recent years has provided us with more effective tools to find new mechanisms and targets.

In this study, we selected GSE50760, GSE41568 and GSE14297 from Gene Expression Omnibus (GEO) and obtained the commonly differentially expressed genes (DEGs) in these three datasets. The Database for Annotation, Visualization and Integrated Discovery (DAVID) was used to analyze these DEGs including molecular function (MF), cellular component (CC), biological process (BP) and Kyoto Encyclopedia of Gene and Genome (KEGG) pathways. Then, we established the protein-protein interaction (PPI) network, and used the MCODE app of Cytoscape to further analyze the core genes. and verify the relationship between these major genes and prognosis ($P < 0.05$) via UCSC Xena online database. Taken above, We found that four commonly differentially expressed genes (SERPING1 ITIH2 CDH2 APOE) were associated with poor prognosis, and all these four genes were associated with extracellular exosomes. In conclusion, our bioinformatics analysis results provides additional useful biomarkers which could be an effective target for patients with CRC liver metastasis

Materials And Methods

GEO data information

We obtained the CRC liver metastasis datasets of GSE50760, GSE41568 and GSE14297 from the free public database NCBI-GEO. GSE50760, GSE41568 and GSE14297 were account of GPL11154, GPL570, GPL6370 Platforms and contained 18 liver metastases and 18 primary foci, 94 liver metastases and 39 primary foci and 18 liver metastases and 18 primary foci, respectively.

Data processing of differentially expressed genes (DEGs)

DEGs between Primary specimen and Liver Metastasis specimen were identified with $|\log FC| > 1$ and adjust P value < 0.05 . Then, DEGs list were imputed in Venn software online to detect the commonly DEGs among these three datasets. The DEGs with $\log FC < -1$ were considered as down-regulated genes and that with $\log FC > 1$ were up-regulated genes.

Gene ontology (GO) and pathway enrichment analysis

Gene ontology analysis (GO) is an international standard classification system of gene function, which is widely used for defining genes to identify unique biological properties of high-throughput transcriptome or genome data (Ashburner, Ball et al. 2000). KEGG is a large comprehensive database including genomes, diseases, biological pathways, drugs and chemicals (Kanehisa 2000). We used DAVID, which is an online bioinformatic tool (Huang da, Sherman et al. 2009), to visualize the DEGs enrichment of BP, MF, CC and pathways ($P < 0.05$)

Protein-protein interaction (PPI) network and module analysis

PPI information can be evaluated using the online tool string (Szklarczyk, Franceschini et al. 2015). Then, PPI results was imported into the Cytoscape and the MCODE app (Shannon, Markiel et al. 2003) was

used to detect the modules in PPI network (degree cutoff = 2, max. Depth = 100, k-core = 2, node score cutoff = 0.2).

Survival Analysis Of Core Genes

UCSC Xena (<http://xena.ucsc.edu/>) was utilized to identify 39 core genes survival data. It was found that 4 genes (SERPING1 ITIH2 CDH2 APOE) had a significantly worse survival ($P < 0.05$, Table 4, Fig. 3).

Table 4
The prognostic information of the 39 key candidate genes

Category	Genes
Genes with significant worse survival ($P < 0.05$)	SERPING1 ITIH2 CDH2 APOE
Genes without significant worse or better survival ($p > 0.05$)	APOA1 AMBP APOA5 ORM2 FGA PLG SERPINF2 APOH VWF ALB RBP4 F2 APOC3 GC HP KNG1 F5 ORM1 FGG SPP1 SERPINA7 C8A HPX IGFBP1 VTN APOB CP APOA2 C3 FGB SERPINA10 SERPINC1
Genes with significant better survival ($p > 0.05$)	SERPIND1 SERPINA1

Results

Identification of DEGs between primary and liver metastasis of CRC

After analysis, we got 1449, 442 and 238 DEGs from GSE50760, GSE41568 and GSE14297, respectively. Then, Venn diagram software were used to identify the commonly DEGs in these three datasets. Results showed that a total of 147 commonly DEGs were detected, including 24 down-regulated genes ($\log_{2}FC < -1$) and 123 up-regulated genes ($\log_{2}FC > 1$) in the metastases (Table 1 & Fig. 1).

Table 1

Commonly 24 down-regulated genes and 123 up-regulated genes in the liver metastases compared to primary CRC tissues

DEGs	Genes Name
Up-regulated genes	<p>HPN ASGR2 CGNL1 REEP6 C8A RGN PON3 AQP9 MBL2 SLC10A1 AMBP RBP4 CDO1 SERPINA7</p> <p>AADAC CFI FMO3 APOC3 ANGPTL3 APOE C3 APOH ORM1 FGA IGFBP1 UPB1 FGG CYP27A1 C8G</p> <p>MASP2 SERPING1 AZGP1 LBP TFR2 F9 APOB ITIH2 FGB CDH2 AGT APCS ABCB4 CRP ADH4 ARG1</p> <p>SERPINA1 UGT2B4 COLEC11 APOC2 MTTP SLC17A2 INHBE SAA1 KNG1 CPB2 KLKB1 HP MAT1A</p> <p>LECT2 ITIH1 PLG SLC13A5 CYP2C8 SLC22A1 SLC25A18 HPR GC FGL1 SAA4 C5 APOA5 FST ORM2</p> <p>ADH1B AKR1C4 HFE2 F5 CPS1 SLC2A2 FCN3 SPP1 OTC F2 APOA2 CFHR1 HPX SULT2A1 CBS DPYS</p> <p>G6PC VTN CYP2E1 AGXT HSD11B1 C4BPA ALB SAA2 HAMP ALDOB SERPINC1 APOA1 C6 C9 APOF</p> <p>CYP4A11 SERPINA3 CPN2 HGFAC SERPINF2 GATM AOX1 SERPINA5 CP SERPIND1 HSD17B6 NNMT</p> <p>SERPINA10 CYP3A4 APOC1 BAAT ASGR1 HABP2 CYP1B1</p>
Down-regulated genes	<p>CXCL14 C15orf48 MMP3 ITLN1 REG4 SPINK4 PIGR FOXF1 FCGBP CEACAM7 LEFTY1 TSPAN1</p> <p>CLCA1 SFRP2 FBLN1 VWF CTSK MAB21L2 PLAC8 PI3 MMP1 COL12A1 PARM1 ST6GALNAC1</p>

DEGs gene ontology and KEGG pathway analysis in CRC liver metastases

All 147 DEGs were analyzed by DAVID and the results of GO analysis showed that 1) for biological processes (BP), DEGs were particularly enriched in acute-phase response, negative regulation of endopeptidase activity, platelet degranulation, fibrinolysis, regulation of complement activation, triglyceride metabolic process, receptor-mediated endocytosis, lipoprotein metabolic process, phospholipid efflux, positive regulation of cholesterol esterification, complement activation-classical pathway, etc.; 2) for molecular function (MF), DEGs were enriched in serine-type endopeptidase inhibitor activity, serine-type endopeptidase activity, heparin binding, phosphatidylcholine-sterol O-acyltransferase activator activity, lipase inhibitor activity, receptor binding, etc.; 3) for GO cell component (CC), DEGs were significantly enriched in extracellular region, blood microparticle, extracellular space, extracellular exosome, platelet alpha granule lumen, chylomicron, etc. (Table 2 and Table S1). KEGG analysis results demonstrated that DEGs were particularly enriched in complement and coagulation cascades, drug

metabolism-cytochrome P450, metabolism of xenobiotics by cytochrome P450, prion diseases, chemical carcinogenesis, staphylococcus aureus infection and linoleic acid metabolism ($P < 0.05$) (Table 3, Table S2).

Table 2
GO analysis of DEGs in liver metastases

Category	Term	Count	%	PValue	FDR
GOTERM_BP_DIRECT	GO:0006953 ~ acute-phase response	16	10.884	7.10E-22	1.12E-18
	GO:0010951 ~ negative regulation of endopeptidase activity	19	12.925	8.35E-18	1.32E-14
	GO:0002576 ~ platelet degranulation	16	10.884	7.44E-15	1.18E-11
	GO:0042730 ~ fibrinolysis	9	6.1224	3.75E-12	5.94E-09
	GO:0030449 ~ regulation of complement activation	9	6.1224	1.01E-10	1.60E-07
	GO:0006641 ~ triglyceride metabolic process	9	6.1224	3.93E-10	6.22E-07
	GO:0006898 ~ receptor-mediated endocytosis	15	10.204	5.43E-10	8.58E-07
	GO:0042157 ~ lipoprotein metabolic process	9	6.1224	8.00E-10	1.27E-06
	GO:0033700 ~ phospholipid efflux	7	4.7619	8.95E-10	1.42E-06
	GO:0010873 ~ positive regulation of cholesterol esterification	6	4.0816	4.77E-09	7.54E-06
GOTERM_CC_DIRECT	GO:0007596 ~ blood coagulation	14	9.5238	5.00E-09	7.91E-06
	GO:0006958 ~ complement activation, classical pathway	11	7.483	1.04E-08	1.65E-05
	GO:0005576 ~ extracellular region	79	53.741	5.87E-44	6.92E-41
	GO:0072562 ~ blood microparticle	36	24.49	6.19E-42	7.30E-39
	GO:0005615 ~ extracellular space	71	48.299	1.34E-40	1.58E-37
	GO:0070062 ~ extracellular exosome	79	53.741	6.53E-27	7.69E-24
	GO:0031093 ~ platelet alpha granule lumen	14	9.5238	3.16E-16	3.89E-13
	GO:0042627 ~ chylomicron	9	6.1224	3.82E-14	4.50E-11

Category	Term	Count	%	PValue	FDR
	GO:0034364 ~ high-density lipoprotein particle	10	6.8027	4.53E-14	5.34E-11
GOTERM_MF_DIRECT	GO:0004867 ~ serine-type endopeptidase inhibitor activity	15	10.204	5.85E-14	8.07E-11
	GO:0004252 ~ serine-type endopeptidase activity	19	12.925	3.79E-12	5.24E-09
	GO:0008201 ~ heparin binding	13	8.8435	8.90E-09	1.23E-05
	GO:0060228 ~ phosphatidylcholine-sterol O-acyltransferase activator activity	5	3.4014	6.71E-08	9.26E-05
	GO:0055102 ~ lipase inhibitor activity	5	3.4014	6.71E-08	9.26E-05
	GO:0005102 ~ receptor binding	16	10.884	2.44E-07	3.37E-04

Table 3
KEGG pathway analysis of different expressed genes in CRC liver metastases

Term	Count	%	PValue	FDR	Genes
hsa04610:Complement and coagulation cascades	27	18.36	9.64E-33	1.09E-29	MBL2 C9 C3 MASP2 C6 C5 FGG FGA FGB SERPINA5 CFI KLKB1 SERPINC1 SERPINA1 KNG1 F9 SERPING1 C4BPA PLG C8G C8A VWF F5 SERPINF2 F2 SERPIND1 CPB2
hsa00980:Metabolism of xenobiotics by cytochrome P450	9	6.12	4.67E-06	0.005273	CYP3A4 AKR1C4 CYP1B1 SULT2A1 HSD11B1 ADH1B UGT2B4 CYP2E1 ADH4
hsa05204:Chemical carcinogenesis	9	6.12	8.42E-06	0.009513	CYP3A4 CYP1B1 SULT2A1 CYP2C8 ADH4 HSD11B1 ADH1B, UGT2B4, CYP2E1
hsa00982:Drug metabolism - cytochrome P450	8	5.44	2.63E-05	0.02971	CYP3A4 CYP2C8 ADH4 AOX1 FMO3 ADH1B UGT2B4 CYP2E1
hsa05150:Staphylococcus aureus infection	7	4.76	6.61E-05	0.074694	MBL2 FGG C3 MASP2 C5 CFI PLG
hsa00140:Steroid hormone biosynthesis	7	4.76	9.94E-05	0.112235	CYP3A4 AKR1C4 CYP1B1 HSD11B1 HSD17B6 UGT2B4 CYP2E1
hsa00830:Retinol metabolism	7	4.76	1.73E-04	0.195168	CYP3A4 CYP2C8 ADH4 AOX1 HSD17B6 ADH1B UGT2B4
hsa05020:Prion diseases	5	3.40	9.41E-04	1.058441	C8A C9 C6 C5 C8G

Term	Count	%	PValue	FDR	Genes
hsa03320:PPAR signaling pathway	6	4.08	0.001749	1.958076	APOA2 APOA1 CYP27A1 APOC3 APOA5 MMP1
hsa04976:Bile secretion	6	4.08	0.001994	2.229643	BAAT AQP9 SULT2A1 SLC22A1 ABCB4 SLC10A1
hsa01230:Biosynthesis of amino acids	6	4.08	0.002407	2.686273	ARG1 MAT1A OTC ALDOB CPS1 CBS
hsa01100:Metabolic pathways	27	18.37	0.005903	6.470185	CYP3A4 ALDOB ADH1B DPYS AGXT ARG1 AKR1C4 MAT1A ADH4 HSD17B6 RGN GATM CYP2C8 OTC UPB1 CYP2E1 CDO1 CPS1 ST6GALNAC1 G6PC BAAT CYP27A1 AOX1 HSD11B1 UGT2B4 CBS NNMT

Protein–protein interaction network (PPI) and modular analysis

The results of STRING and Cytoscape analysis showed that total of 138 DEGs (121 up-regulated and 17 down-regulated genes) of the 147 DEGs were filtered into the DEGs PPI network complex, containing 1391 edges (Fig. 2A). After MCODE analysis, there were 33 central nodes were identified among the 138 nodes (Fig. 2B).

Discussion

To identify more potential prognostic biomarker in CRC liver metastases, three GEO datasets (GSE50760, GSE41568 and GSE14297) associated with CRC liver metastasis were analyzed by bioinformatical methods. One hundred and thirty-liver metastases and seventy-five primary CRC specimens were enrolled in the present research. Via GEO2R and Venn software, we revealed a total of 147 commonly changed DEGs ($|\log FC| > 1$ and adjust P value < 0.05) including 123 up-regulated ($\log FC > 1$) and 24 down-regulated ($\log FC < -1$) DEGs. The DAVID methods were used to analyze Gene Ontology and Pathway Enrichment and the results showed that 1) for biological processes (BP), DEGs were particularly enriched in acute-phase

response, negative regulation of endopeptidase activity, platelet degranulation, fibrinolysis, regulation of complement activation, triglyceride metabolic process, receptor-mediated endocytosis, lipoprotein metabolic process, phospholipid efflux, positive regulation of cholesterol esterification, complement activation-classical pathway, etc.; 2) for molecular function (MF), DEGs were enriched in serine-type endopeptidase inhibitor activity, serine-type endopeptidase activity, heparin binding, phosphatidylcholine-sterol O-acyltransferase activator activity, lipase inhibitor activity, receptor binding, etc.; 3) for GO cell component (CC), DEGs were significantly enriched in extracellular region, blood microparticle, extracellular space, extracellular exosome, platelet alpha granule lumen, chylomicron, etc.

For pathway analysis, DEGs were particularly enriched in complement and coagulation cascades, drug metabolism-cytochrome P450, metabolism of xenobiotics by cytochrome P450, prion diseases, chemical carcinogenesis, staphylococcus aureus infection and linoleic acid metabolism ($P < 0.05$). Next, DEGs PPI network complex of 138 nodes and 1391 edges was constructed via the STRING online database and Cytoscape software. Then, 39 vital genes were screened from the PPI network complex by Cytotype MCODE analysis. In addition, through Kaplan Meier plotter analysis, we found that 4 (SERPING1 ITIH2 CDH2 APOE) of the 39 genes had a significantly worse survival. Finally, we did literature research on these four genes.

Complement and coagulation cascades is the most significance pathway in our results. Complement is considered as the first line of defense against non self or unwanted host elements. Therefore, the researchers designed different strategies to treat tumors by increasing complement activity (Kolev, Towner et al. 2011). However, as early as 1975, shearer et al. reported that complement is able to stimulate tumor growth when treating tumor with low concentration antibody (Shearer 1975). And recent studies have confirmed that complement can promote tumor growth in mouse models (Markiewski and Lambris 2009). Moreover, in vivo and in vitro data show that tumor cells can activate complement (Bjorge, Hakulinen et al. 2005, Corrales, Ajona et al. 2012) and Shi et al. (Shi, Fang et al. 2017) reported that complement component 1, q subcomponent binding protein (C1QBP), in lipid rafts mediates liver metastasis of pancreatic cancer by regulating IGF-1 / IGF1R signaling. They reported that the expression of C1QBP in many human cancers is higher than that in normal tissues, including thyroid, lung, esophageal, gastric and colon cancer, and refer to previous reports that C1QBP mediates EGF induced chemotaxis and distant metastasis by activating receptor tyrosine kinase. It has been accepted that complement and coagulation cascades is key pathway promoting the growth and metastasis of different cancers. However, the mechanisms and effects of complement specific deregulation on tumor microenvironment are not clear. More works are needed to elucidate the complex process of cancer cells controlling complement activation and how complement affects tumor progression

SERPING1 is also known as C1IH or CINH. SERPING1 belongs to the serpins family of serine protease inhibitors, which regulate the proteases involved in fibrinolysis, coagulation, inflammation, cell migration, cell differentiation and apoptosis (Hayashi, Ushizawa et al. 2011). Many studies have reported that the SERPING1 expression is up-regulated in tumors, including glioblastoma (Fornvik, Maddahi et al. 2017), breast cancer and gastric cancer (Wojtukiewicz 1998). Fornvik, K. et al. (Fornvik, Maddahi et al. 2017)

found that, in glioblastoma, the high SERPING1 expression is correlated with dismal prognosis and the survival period can be significantly prolonged after the SERPING1 being blocked by antibody. The same is true for colorectal cancer, and SERPING1 can be used as an independent predictor of cancer mortality (Kocsis, Meszaros et al. 2011)¶

ITIH2, is also known as SHAP. The complex of ITIH2 and hyaluronan plays a significant role in the inflammatory response and the high expression of ITIH2 is associated with poor outcome of NSCLC (Wu, Wang et al. 2013). In addition, a strong link between the expression levels of ITIH2 and estrogen levels has been demonstrated, since ITIH2 contains an estrogen-binding domain which might be critical for metastasis and tumor growth because estrogen has a profound effect on extracellular matrix integrity (Kopylov, Stepanov et al. 2020). At present, there is not much research on the role of ITIH2 in tumor growth and metastasis, which is also a potential direction for the study of tumor metastasis mechanism in the future

CDH2 is a member of the cadherin family, which regulates many cellular processes including apoptosis, angiogenesis, and chemoresistance (Miao, Wang et al. 2018). CDH2 is up-regulated in many kinds of cancers, including colorectal cancer (Wang 2015), prostate cancer (Zhang, Shen et al. 2014), gastric cancer (Zhang, Wang et al. 2014), bladder cancer (van der Horst, Bos et al. 2014), lung cancer (Zhu 2015) and glioma (Chen, Cai et al. 2018). Moreover, it is reported that CDH2 play an important role in the process of epithelia-mesenchymal transition (Yang, Wang et al. 2015). In the lung cancer research, it was found that CDH2 expression is associated the gefitinib-resistance (Yamauchi 2011) and to the brain metastases of lung cancer cells (Grinberg-Rashi, Ofek et al. 2009). In glioma, CDH2 is recognized as a independent prognostic factor and its high expression is correlated with high-grade glioma and worse outcome (Chen, Cai et al. 2018).

APOE, which has been proved to be up regulated in patients with colorectal cancer (Borgquist, Butt et al. 2016). It has been reported that overexpression of APOE was related to the growth of tumor in the advanced stage and increaseds the risk of tumor metastasis and invasion (Kopylov, Stepanov et al. 2020), which may be related to the key role of APOE in cell proliferation and DNA synthesis (Niemi and Yla-Herttuala 2002). Moreover,APOE is one of the regulators of PI3K/Akt/mTOR pathway, which plays an important role in cell migration and proliferation, therefore, APOE can promote tumor progression by enhancing cell polarity (Vergadi, Ieronymaki et al. 2017). The value of APOE in CRC species has been reported as a prognostic indicator (Borgquist, Butt et al. 2016).

At present, the relationship between colorectal cancer and these four genes has been reported, but most of them are single gene research. It has not been reported that using these four genes as a panel to predict the metastasis and prognosis of colorectal cancer. Therefore, our data can provide a more comprehensive and accurate method to predict the metastasis and prognosis of colorectal cancer and direction for the future study of CRC liver metastasis.

Declarations

Funding

No funding was received.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

Junsheng Chen and Hongzhou Liu conceived and designed the idea to this paper; collected and analyzed the data, and drafted the paper. Both the authors read and approved the final manuscript.

Competing interests

The authors declare no competing financial interests.

Ethics approval and consent to participate

Not available

Patient consent for publication

Not available

References

- Ashburner, M., C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin and G. Sherlock (2000). "Gene ontology: tool for the unification of biology. The Gene Ontology Consortium." Nat Genet **25**(1): 25-29.
- Bjorge, L., J. Hakulinen, O. K. Vintermyr, H. Jarva, T. S. Jensen, O. E. Iversen and S. Meri (2005). "Ascitic complement system in ovarian cancer." Br J Cancer **92**(5): 895-905.
- Borgquist, S., T. Butt, P. Almgren, D. Shiffman, T. Stocks, M. Orho-Melander, J. Manjer and O. Melander (2016). "Apolipoproteins, lipids and risk of cancer." Int J Cancer **138**(11): 2648-2656.
- Chen, Q., J. Cai and C. Jiang (2018). "CDH2 expression is of prognostic significance in glioma and predicts the efficacy of temozolomide therapy in patients with glioblastoma." Oncol Lett **15**(5): 7415-7422.
- Corrales, L., D. Ajona, S. Rafail, J. J. Lasarte, J. I. Riezu-Boj, J. D. Lambris, A. Rouzaut, M. J. Pajares, L. M. Montuenga and R. Pio (2012). "Anaphylatoxin C5a creates a favorable microenvironment for lung cancer

progression." J Immunol **189**(9): 4674-4683.

Fakih, M. G. (2015). "Metastatic colorectal cancer: current state and future directions." J Clin Oncol **33**(16): 1809-1824.

Fornvik, K., A. Maddahi, O. Persson, K. Osther, L. G. Salford and H. Nittby Redebrandt (2017). "C1-inactivator is upregulated in glioblastoma." PLoS One **12**(9): e0183086.

Grinberg-Rashi, H., E. Ofek, M. Perelman, J. Skarda, P. Yaron, M. Hajdуч, J. Jacob-Hirsch, N. Amariglio, M. Krupsky, D. A. Simansky, Z. Ram, R. Pfeffer, I. Galanter, D. M. Steinberg, I. Ben-Dov, G. Rechavi and S. Izraeli (2009). "The expression of three genes in primary non-small cell lung cancer is associated with metastatic spread to the brain." Clin Cancer Res **15**(5): 1755-1761.

Hayashi, K. G., K. Ushizawa, M. Hosoe and T. Takahashi (2011). "Differential gene expression of serine protease inhibitors in bovine ovarian follicle: possible involvement in follicular growth and atresia." Reprod Biol Endocrinol **9**: 72.

Huang da, W., B. T. Sherman and R. A. Lempicki (2009). "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources." Nat Protoc **4**(1): 44-57.

Huang, S., X. Tan, Z. Huang, Z. Chen, P. Lin and S. W. Fu (2018). "microRNA biomarkers in colorectal cancer liver metastasis." J Cancer **9**(21): 3867-3873.

Kanehisa, M. G., S. (2000). "KEGG: Kyoto Encyclopedia of Gene and Genomes." Nucleic Acids Res **28**(1): 27-30.

Kocsis, J., T. Meszaros, B. Madaras, E. K. Toth, S. Kamondi, P. Gal, L. Varga, Z. Prohaszka and G. Fust (2011). "High levels of acute phase proteins and soluble 70 kDa heat shock proteins are independent and additive risk factors for mortality in colorectal cancer." Cell Stress Chaperones **16**(1): 49-55.

Kolev, M., L. Towner and R. Donev (2011). "Complement in cancer and cancer immunotherapy." Arch Immunol Ther Exp (Warsz) **59**(6): 407-419.

Kopylov, A. T., A. A. Stepanov, K. A. Malsagova, D. Soni, N. E. Kushlinsky, D. V. Enikeev, N. V. Potoldykova, A. V. Lisitsa and A. L. Kaysheva (2020). "Revelation of Proteomic Indicators for Colorectal Cancer in Initial Stages of Development." Molecules **25**(3).

Markiewski, M. M. and J. D. Lambris (2009). "Unwelcome complement." Cancer Res **69**(16): 6367-6370.

Miao, J., W. Wang, S. Wu, X. Zang, Y. Li, J. Wang, R. Zhan, M. Gao, M. Hu, J. Li and S. Chen (2018). "miR-194 Suppresses Proliferation and Migration and Promotes Apoptosis of Osteosarcoma Cells by Targeting CDH2." Cell Physiol Biochem **45**(5): 1966-1974.

- Niemi, M. H., T.; Karttunen, T.J.; Eskelinen, S.; Kervinen, K.; Savolainen, M.J.; Lehtola, J.; Makela, J.; and S. K. Yla-Herttuala, Y.A. (2002). "Apolipoprotein E and colon cancer: Expression in normal and malignant human intestine and effect on cultured human colonic adenocarcinoma cells." Eur J Intern Med **13**(1): 37-34.
- Pretzsch, E., F. Bosch, J. Neumann, P. Ganschow, A. Bazhin, M. Guba, J. Werner and M. Angele (2019). "Mechanisms of Metastasis in Colorectal Cancer and Metastatic Organotropism: Hematogenous versus Peritoneal Spread." J Oncol **2019**: 7407190.
- Shannon, P., A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski and T. Ideker (2003). "Cytoscape: a software environment for integrated models of biomolecular interaction networks." Genome Res **13**(11): 2498-2504.
- Shearer, W. T. A., J.P.; Frank, M.M.; Parker, C.W. (1975). "Humoral immunostimulation IV Role of complement." J Exp Med **141**(4): 736-752.
- Shi, H., W. Fang, M. Liu and D. Fu (2017). "Complement component 1, q subcomponent binding protein (C1QBP) in lipid rafts mediates hepatic metastasis of pancreatic cancer by regulating IGF-1/IGF-1R signaling." Int J Cancer **141**(7): 1389-1401.
- Siegel, R. L., K. D. Miller and A. Jemal (2018). "Cancer statistics, 2018." CA Cancer J Clin **68**(1): 7-30.
- Szklarczyk, D., A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, M. Simonovic, A. Roth, A. Santos, K. P. Tsafou, M. Kuhn, P. Bork, L. J. Jensen and C. von Mering (2015). "STRING v10: protein-protein interaction networks, integrated over the tree of life." Nucleic Acids Res **43**(Database issue): D447-452.
- van der Horst, G., L. Bos, M. van der Mark, H. Cheung, B. Heckmann, P. Clement-Lacroix, G. Lorenzon, R. C. Pelger, R. F. Bevers and G. van der Pluijm (2014). "Targeting of alpha-v integrins reduces malignancy of bladder carcinoma." PLoS One **9**(9): e108464.
- Vergadi, E., E. Ieronymaki, K. Lyroni, K. Vaporidi and C. Tsatsanis (2017). "Akt Signaling Pathway in Macrophage Activation and M1/M2 Polarization." J Immunol **198**(3): 1006-1014.
- Vogelstein, B., N. Papadopoulos, V. E. Velculescu, S. Zhou, L. A. Diaz, Jr. and K. W. Kinzler (2013). "Cancer genome landscapes." Science **339**(6127): 1546-1558.
- Wang, X. X., X.; Wu, J.; Wan, Y.; Hui, H.; Cao, X. (2015). "MicroRNA-206 Attenuates Tumor Proliferation and Migration Involving the Downregulation of NOTCH3 in Colorectal Cancer." Oncol Rep **33**(3): 1402-1410.
- Wojtukiewicz, W. Z. R., M.; Kloczko, J.; Dib, A.; Galar, M. (1998). "Profiles of Plasma Serpins in Patients with Advanced Malignant Melanoma, Gastric Cancer and Breast Cancer " Haemostasis **28**(1): 7-13.

Wu, X., L. Wang, Y. Ye, J. A. Aakre, X. Pu, G. C. Chang, P. C. Yang, J. A. Roth, R. S. Marks, S. M. Lippman, J. Y. Chang, C. Lu, C. Deschamps, W. C. Su, W. C. Wang, M. S. Huang, D. W. Chang, Y. Li, V. S. Pankratz, J. D. Minna, W. K. Hong, M. A. Hildebrandt, C. A. Hsiung and P. Yang (2013). "Genome-wide association study of genetic predictors of overall survival for non-small cell lung cancer in never smokers." Cancer Res **73**(13): 4028-4038.

Yamauchi, M. Y., I.; Yamaguchi, R.; Shimamura, T.; Nagasaki, M.; Imoto, S.; Niida, A.; Koizumi, F.; Kohno, T.; Yokota, J.; Miyano, S.; Gotoh, N. (2011). "N-cadherin expression is a potential survival mechanism of gefitinib-resistant lung cancer cells." Am J Cancer Res **1**(7): 823-833.

Yang, H., L. Wang, J. Zhao, Y. Chen, Z. Lei, X. Liu, W. Xia, L. Guo and H. T. Zhang (2015). "TGF-beta-activated SMAD3/4 complex transcriptionally upregulates N-cadherin expression in non-small cell lung cancer." Lung Cancer **87**(3): 249-257.

Zarour, L. R., S. Anand, K. G. Billingsley, W. H. Bisson, A. Cercek, M. F. Clarke, L. M. Coussens, C. E. Gast, C. B. Geltzeiler, L. Hansen, K. A. Kelley, C. D. Lopez, S. R. Rana, R. Ruhl, V. L. Tsikitis, G. M. Vaccaro, M. H. Wong and S. C. Mayo (2017). "Colorectal Cancer Liver Metastasis: Evolving Paradigms and Future Directions." Cell Mol Gastroenterol Hepatol **3**(2): 163-173.

Zhang, J., C. Shen, L. Wang, Q. Ma, P. Xia, M. Qi, M. Yang and B. Han (2014). "Metformin inhibits epithelial-mesenchymal transition in prostate cancer cells: involvement of the tumor suppressor miR30a and its target gene SOX4." Biochem Biophys Res Commun **452**(3): 746-752.

Zhang, J., P. Wang, J. Zhu, W. Wang, J. Yin, C. Zhang, Z. Chen, L. Sun, Y. Wan, X. Wang, G. Chen and Y. Liu (2014). "SPARC expression is negatively correlated with clinicopathological factors of gastric cancer and inhibits malignancy of gastric cancer cells." Oncol Rep **31**(5): 2312-2320.

Zhu, J. Y., L.; You, W.; Cui, X.; Chen, Y.; Hu, J.; Liu, W.; Li, S.; Song, X.; Wei, Y.; Zhang, W.; Li, F. (2015). "Genetic variation in miR-100 rs1834306 is associated with decreased risk for esophageal squamous cell carcinoma in Kazakh patients in northwest China." Int J Clin Exp Pathol **8**(6): 7332-7340.

Figures

Fig 1

Figure 1

Identification of 147 commonly DEGs in the three datasets by online Venn diagrams software (<http://bioinformatics.psb.ugent.be/webtools/Venn/>). Different colors represent different datasets. (A) 123 DEGs were up-regulated in the three datasets ($\log_{2}FC > 1$). (B) 24 DEGs were down-regulated in three datasets ($\log_{2}FC < -1$)

Fig2

Figure 2

Commonly DEGs PPI network constructed by STRING online database and Module analysis. (A) There were 138 DEGs in the PPI network complex. The nodes meant proteins; the edges meant the interaction of proteins; yellow circles meant down-regulated DEGs and red circles meant up-regulated DEGs. (B) Module analysis by MCODE app in the Cytoscape software (degree cutoff = 2, node score cutoff = 0.2, k-core = 2, and max. Depth = 100)

Fig 3

Figure 3

The prognostic information of the 39 core genes. 4 of 39 genes had a significantly worse survival rate ($P < 0.05$)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS1GOanalysisofDEGsinlivermetastases.xlsx](#)
- [TableS2KEGGpathwayanalysisofdifferentexpressedgenes.xlsx](#)