European Review for Medical and Pharmacological Sciences 2015; 19: 2423-2428

Integrated regulatory network involving differently expressed genes and protein-protein interaction on pancreatic cancer

J. LONG, X.-D. WU, Z. LIU, Y.-H. XU, C.-L. GE

Department of General Surgery, the First Hospital of China Medical University, Shenyang, China

Abstract. – OBJECTIVE: Pancreatic cancer is a deadly disease with poor prognosis. However, comprehensive understanding about its pathogenesis remains insufficient. In this study, we aimed to find potential novel approaches for the treatment of pancreatic cancer and explore the regulatory mechanisms underlying pancreatic cancer progression.

MATERIALS AND METHODS: The gene expression profile data GSE32688 were downloaded from Gene Expression Omnibus database followed by background correction and normalization through GCRMA (GC Robust Multi-array Average) method. Then DEGs (differentially expressed genes) were identified using t-test method and DEGs-related PPIs (protein-protein interaction) were extracted from STRING database. The PPI networks were constructed by calculating the pearson correlation coefficient under different conditions. Moreover, the network was divided into a number of unit modules, and KEGG pathway and GO analysis were performed for genes in module networks using clusterProfiler.

RESULTS: In total, 199 DEGs (165 up-regulated genes and 34 down-regulated genes) were screened between tumor and normal samples. The integrated DEG. PPI network was established by comparing two different networks under tumor and normal conditions respectively. The top ten genes with high degrees such as *ANLN*, *PSRC1* and *ECT2* were identified in the integrated network, and they were mainly enriched in cell cycle pathway.

CONCLUSIONS: *ECT2* and *PSRC1* might be used as two novel biomarkers for diagnosis and management of pancreatic cancer.

Key Words:
Pancreatic cancer, Regulatory network, ANLN, ECT2, PSRC1, Bioinformatics methods.

Introduction

Pancreatic cancer is known as the fourth most common cause of cancer death among both men and women in the United States[1]. In 2014, 46420 new cases and 39590 deaths are estimated to occur in America[2]. Pancreatic ductal adenocarcinoma (PDAC), which is the most common and deadly form of pancreatic cancer, usually evolves from noninvasive precursor lesions, intraductal papillary mucinous neoplasms and mucinous cystic neoplasms[3]. The predominant risk factors for pancreatic cancer are smoke, age and genetic disorders. It has been found that smoke accounts for about 20% formation of pancreatic tumors[3]. In addition, pancreatic cancer is more prevalent among the elderly than the younger[4].

Pancreatic cancer is an aggressive malignancy with extremely high mortality. For this reason, surgical resection is usually not adequate for treatment of this disease[5]. The use of endoscopic ultrasound (EUS) is a great improvement in the diagnosis of pancreatic cancer[6], but it is often diagnosed in the advanced stage and only a small fraction of the tumors are localized and potentially curable[4], which make the outcome of surgical resection unsatisfactory. Therefore, numerous studies were conducted to investigate effective adjuvant chemotherapies and neoadjuvant treatments for pancreatic cancer[7,8]. Among them, gemcitabine has been demonstrated to improve median disease-free survival, and was suggested as effective adjuvant chemotherapy in resectable carcinoma of the pancreas[9]. Despite the advance in detection and management of pancreatic cancer, the overall 5-year survival rate maintained less than 5% for almost 50 years[10].

Better understanding of the pathogenesis of pancreatic cancer contributes to more effective approaches to prevent and control this disease. Previous researches indicated that the development of pancreatic cancer was associated with accumulation of genetic alterations including mutations of oncogenes such as *KRAS*, *BRAF*, *MYB*, *AKT2* and *EGFR*, as well as tumor-suppressor genes such as *MAP2K4*, *TGFBR1* and

Corresponding Author: Jin Long, MD; e-mail: jinlongdoc@163.com

FBXW7[11]. Moreover, epigenetic alterations were also demonstrated to be involved in pancreatic cancer progression[12]. Although the expression profiles analysis of pancreatic cancer has been studied to screen DEGs (differentially expressed genes) related to the pancreatic cancer[13], our analysis method was modified: the integrated PPI (protein-protein interaction) network was more conducive to finding the crucial genes involved in the process of pancreatic cancer. In this study, we downloaded gene expression profile data GSE32688 and analyzed this microarray data using bioinformatics methods, aiming to provide novel potential biomarkers for detection and prevention of pancreatic cancer.

Materials and Methods

Microarray Data Analysis

The gene expression profile data GSE32688 were downloaded from GEO (Gene Expression Omnibus) database, which were deposited by Donahue et al[14]. The GPL570 (HG-U133_Plus_2) Affymetrix Human Genome U133 Plus 2.0 Array was used as microarray platform. The gene expression profile chips were comprised of 25 tumor samples and 7 normal samples. Additionally, the microarray data sets were preprocessed using GCRMA (GC Robust Multi-array Average) for background correction and normalization[14]. The probes with high loss frequency (more than 30%) were deleted.

DEGs Screening

Since one gene may correspond to multiple probes, the mean value of the corresponding probes was calculated as the gene expression value. Moreover, the gene expression matrix consisting of 20539 genes was established.

The differential expression analysis between tumor and normal samples was performed using *t*-test method. The FDR (false discovery rate) less than 0.01 was set as the cut-off criterion for DEGs screening.

Construction of the Integrated PPI Network

The STRING database (STRING 9.1)[15] was used to obtain information on human PPI, which comprised of 2159978 interaction relationships among 18105 genes. Then the DEGs-related PPIs were extracted, and the pearson correlation coefficient under tumor and normal conditions was separately calculated. FDR less than 0.01 was used as the cut-off criterion. Thus, two types of PPI networks were constructed, named disease network and control network respectively. By comparing the two networks, the integrated DEG. PPI network was obtained by retaining differential interaction under different conditions. In the DEG. PPI network, an unit module is defined as a network consisting of a DEG and its interacted proteins.

Function Analysis of the Genes in Network Modules

The clusterProfiler[16] package was recruited to perform KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis and GO (Gene Ontology) functional enrichment analysis for genes in module networks. The FDR less than 0.05 was chosen as the cut-off criterion for both KEGG and GO enrichment analysis.

Results

DEGs Between Tumor and Normal Samples

Based on the differential expression analysis, a total of 199 DEGs were screened between 25 tumor samples and 7 normal samples, including 165 (83%) up-regulated DEGs and 34 (17%) down-regulated DEGs.

Integrated DEG PPI Network

By comparing disease network and control network, the integrated DEG. PPI network was established (Figure 1). The top ten genes with high node degrees in the network was presented in Table I, including up-regulated genes such as *ANLN*, *PSRC1*, *ECT2*, *SHFM1*, *DLAPH3*, *RACH1*, *F12*, *IQGAP3* and *TSG101*, as well as a down-regulated gene *RPL15*.

The nodes in the DEG. PPI networks were divided into two categories: the DEN (differentially expressed node) and non-DEN (non-differentially expressed node). Moreover, the edges in the network could also be divided into two categories: TSE (tumor sample-specific edges) and NSE (normal sample-specific edges). As a result, a total of 155 (77.9%) DENs interacted with other proteins in the DEG. PPI networks. Besides, approximately 89.3% of edges were TSE (Table II).

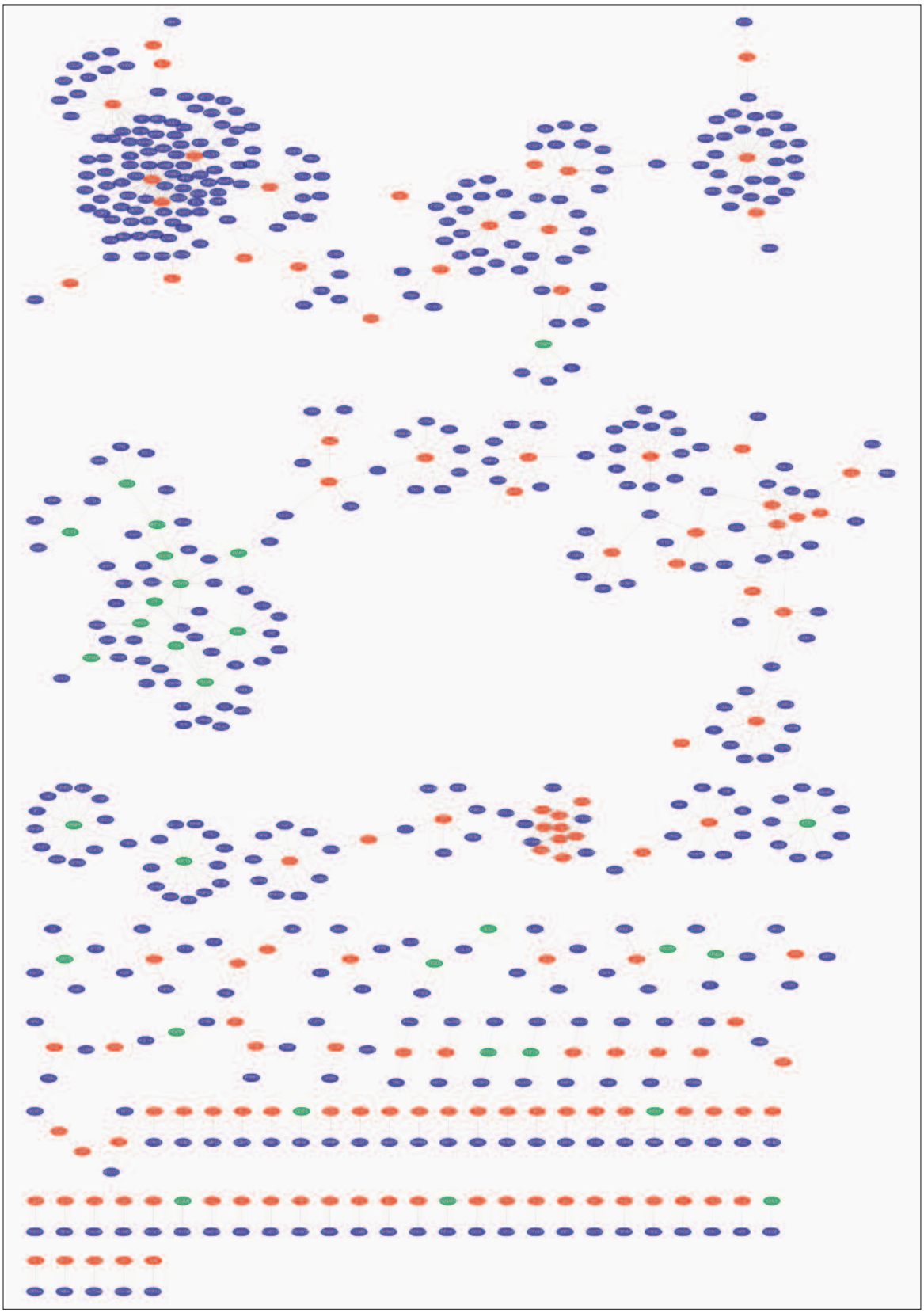

Figure 1. Integrated protein-protein interaction network of DEGs (differentially expressed genes). Red nodes represent up-regulated DEGs; green nodes represent down-regulated DEGs; blue nodes represent non-differentially expressed genes.

Table I. Top ten DEGs (differently expressed genes) with high degrees in the integrated network.

Gene	de_state	node_degree
ANLN	1	64
PSRC1	1	53
ECT2	1	37
SHFM1	1	27
DIAPH3	1	21
RAC1	1	21
F12	1	16
IQGAP3	1	16
RPL15	-1	13
TSG101	1	13

de-state: 1 represents up-regulated DEGs and -1 represents down-regulated DEGs in the pancreatic cancer samples,

Table II. Characteristics of integrated protein-protein interaction network.

Subtypes	Numbers
DEN	155
Non-DEN	475
	630
TSE	616
NSE	80
	696

DEN: differentially expressed node; Non-DEN: non-differentially expressed node; TSE: tumor sample-specific edges; NSE: normal sample-specific edges.

Enrichment Analysis of the Genes in Module Networks

GO and KEGG pathway enrichment analyses were performed for the unit modules, only four modules were significantly enriched in KEGG pathways. As shown in Table III, the genes with high degrees such as *ANLN*, *PSRC1* and *ECT2* were mainly enriched in the pathways correlated with cell cycle and oocyte meiosis.

Discussion

Pancreatic cancer is one of the most deadly cancer diseases because of the poor prognosis. The improvement of the diagnosis and therapeutic methods of this disease rely on the potent molecular markers[17]. In the present study, we identified 199 DEGs between normal and tumor samples, consisting of 165 up-regulated genes and 34 down-regulated genes. Besides, the integrated DEG. PPI network was constructed, and ten genes with high degrees including *ANLN*, *PSRC1* and *ECT2* were identified. The enrichment analysis indicated that these genes were mainly associated with cell cycle process.

ANLN (anillin) is a 124 kDa protein with highly conserved multidomain, which interacts with cytoskeletal components[18]. The gene *ANLN* was found to be over-expressed in pancreatic cancer, and it had tumor-related expression pattern[19]. Due to the significantly different expression between cancer and normal samples (cancer/normal 19.8-fold), *ANLN* was supposed to be a molecular marker of pancreatic cancer[17].

Cytokinesis is a dynamic and plastic process involving the co-ordinated regulation of many components. It proceeds through the formation of actomyosin-based contractile ring (CR)[20,21]. ANLN could interact with F-actin and regulate the contractile activity of myosin II. Besides, it is essential for the formation of cleavage furrow and is suggested to be a pivotal organizer of the cytokinetic machinery. In addition, ANLN contains a crucial C-terminal domain, which shares homology with the RhoA binding protein Rhotekin and directly interacts with RhoA[22]. RhoA is a small GTPase which concentrates at

Table III. Pathway analyses for genes in unit module networks.

Symbol	ID	Description	Count	BgRatio	*p*-value
ECT2	hsa04110	Cell cycle	6	128/5894	6.25E-07
PSRC1	hsa04110	Cell cycle	8	128/5894	1.07E-10
PSRC1	hsa04114	Oocyte meiosis	7	114/5894	2.58E-09
ANLN	hsa04110	Cell cycle	9	128/5894	6.23E-11
ANLN	hsa04114	Oocyte meiosis	8	114/5894	9.69E-10
RPL15	hsa03010	Ribosome	9	92/5894	7.82E-15

BgRatio: Background ratio.

the site of nascent cleavage furrow formation and acts as a crucial regulator of cytokinesis[23]. Further structure investigation demonstrated that ANLN acted as a bifunctional linker coordinating the transition between CR and MR (midbody ring, which forms to stabilize the intercellular bridge until abscission) during cytokinesis [20]. The present study indicated that *ANLN* was identified as a key gene related to pancreatic cancer and the pathway of cell cycle. All these results illustrated that *ANLN* may play critical roles in pancreatic cancer progression via the regulation of CR during cytokinesis.

The protein encoded by *ECT2* (epithelial cell transforming sequence 2 gene) is a guanine nucleotide exchange factor (GEF) and transforming protein[24]. As a RhoGEF, ECT2 is required for RhoA localization and furrowing, and plays critical roles in Rho activation during cytokinesis[25]. Only after activated by *ECT2* can RhoA exert its roles in the process of cytokinesis[22]. The study of Zhang et al[26] indicated that ECT2 may participate in pancreatic tumorgenesis because that it was highly expressed in pancreatic tumor tissues compared with normal pancreatic tissues. Besides, the expression of *ECT2* is elevated with the onset of DNA synthesis and remains elevated during G2 and M phases in mice cells, suggesting that ECT2 has a cell cycle-dependent expression pattern and participates in the regulation of cytokinesis[27]. In our study, *ECT2* was another pivotal gene correlated with pancreatic cancer based on the DEG. PPI network. These evidences implied that *ECT2* had great potential to be used as a novel biomarker for the detection and prevention of pancreatic cancer. In addition, ANLN and ECT2 interacted with each other in the PPI network, and they were both related to the regulation of RohA. Thus, we speculated that *ANLN* and *ETC2* might participate in pancreatic cancer development through the regulation of RhoA in the process of cytokinesis.

The gene *PSRC1* encodes paraspeckle, a nucleolar protein, which localizes to punctate subnuclear structures that occur close to splicing speckles. Alternative splicing of this gene could result in multiple transcript variants[28]. *PSRC1* has been demonstrated to be regulated by p53 and involved in p53-mediated growth suppression in mouse. Moreover, *PSRC1* was reported to participate in microtubules formation and assembly of mitotic spindle, implying its potential involvement in cytokinesis[29]. Previous study showed that *PSRC1* was identified as one of three candidate susceptibility genes (*SORT1*, *CELSR2* and *PSRC1*) for cardiovascular disease[30]. However, elucidations on the corresponding relationships between this gene and other diseases were deficient. According to the results of present study, *PSRC1* was predicted as a critical gene associated with pancreatic cancer and cell cycle events, which meant that *PSRC1* might be a novel biomarker for the diagnosis and management of pancreatic cancer.

Conclusions

Three vital DEGs were identified to be correlative with pancreatic cancer and participate in cell cycle pathway. Among them, both *ANLN* and *ECT2* might involve in pancreatic cancer progression via the regulation of RhoA during cytokinesis. Furthermore, *PSRC1* and *ECT2* might be used as two novel biomarkers for the diagnosis and treatment of pancreatic cancer.

Conflict of Interest

The Authors declare that there are no conflicts of interest.

References

1) Campbell PJ, Yachida S, Mudie LJ, Stephens PJ, Pleasance ED, Stebbings LA, Morsberger LA, Latimer C, Mclaren S, Lin M-L. The patterns and dynamics of genomic instability in metastatic pancreatic cancer. Nature 2010; 467: 1109-1113.
2) Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. CA Cancer J Clin 2014; 64: 9-29.
3) Vincent A, Herman J, Schulick R, Hruban RH, Goggins M. Pancreatic cancer. Lancet 2011; 378: 607-620.
4) Hidalgo M. Pancreatic cancer. New Engl J Med 2010; 362: 1605-1617.
5) Li C, Heidt DG, Dalerba P, Burant CF, Zhang L, Adsay V, Wicha M, Clarke MF, Simeone DM. Identification of pancreatic cancer stem cells. Cancer Res 2007; 67: 1030-1037.
6) Galasso D, Carnuccio A, Larghi A. Pancreatic cancer: diagnosis and endoscopic staging. Eur Rev Med Pharmacol Sci 2010; 14: 375-385.
7) Neoptolemos JP, Stocken DD, Bassi C, Ghaneh P, Cunningham D, Goldstein D, Padbury R, Moore MJ, Gallinger S, Mariette C. Adjuvant chemotherapy with fluorouracil plus folinic acid vs gemcitabine following pancreatic cancer resection: a randomized controlled trial. JAMA 2010; 304: 1073-1081.

8) Gillen S, Schuster T, Zum Büschenfelde CM, Friess H, Kleeff J. Preoperative/neoadjuvant therapy in pancreatic cancer: a systematic review and meta-analysis of response and resection percentages. PLoS Med 2010; 7: e1000267.

9) Oettle H, Post S, Neuhaus P, Gellert K, Langrehr J, Ridwelski K, Schramm H, Fahlke J, Zuelke C, Burkart C. Adjuvant chemotherapy with gemcitabine vs observation in patients undergoing curative-intent resection of pancreatic cancer: a randomized controlled trial. JAMA 2007; 297: 267-277.

10) Jemal A, Siegel R, Xu J, Ward E. Cancer statistics, 2010. CA Cancer J Clin 2010; 60: 277-300.

11) Rozenblum E, Schutte M, Goggins M, Hahn SA, Panzer S, Zahurak M, Goodman SN, Sohn TA, Hruban RH, Yeo CJ. Tumor-suppressive pathways in pancreatic carcinoma. Cancer Res 1997; 57: 1731-1734.

12) Omura N, Goggins M. Epigenetics and epigenetic alterations in pancreatic cancer. Int J Clin Exp Pathol 2009; 2: 310-326.

13) Yang D, Zhu Z, Wang W, Shen P, Wei Z, Wang C, Cai Q. Expression profiles analysis of pancreatic cancer. Eur Rev Med Pharmacol Sci 2013; 17: 311-317.

14) Donahue TR, Tran LM, Hill R, Li Y, Kovochich A, Calvopina JH, Patel SG, Wu N, Hindoyan A, Farrell JJ. Integrative survival-based molecular profiling of human pancreatic cancer. Clin Cancer Res 2012; 18: 1352-1363.

15) Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, Von Mering C. STRING v9. 1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res 2013; 41: D808-D815.

16) Yu G, Wang L-G, Han Y, He Q-Y. ClusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012; 16: 284-287.

17) Olakowski M, Tyszkiewicz T, Jarzab M, Król R, Oczko-Wojciechowska M, Kowalska M, Kowal M, Gala GM, Kajor M, Lange D. NBL1 and anillin (ANLN) genes over-expression in pancreatic carcinoma. Folia Histochem Cytobiol 2009; 47: 249-248.

18) Suzuki C, Daigo Y, Ishikawa N, Kato T, Hayama S, Ito T, Tsuchiya E, Nakamura Y. ANLN plays a critical role in human lung carcinogenesis through the activation of RHOA and by involvement in the phosphoinositide 3-kinase/AKT pathway. Cancer Res 2005; 65: 11314-11325.

19) Chen X, Leung SY, Yuen ST, Chu K-M, Ji J, Li R, Chan AS, Law S, Troyanskaya OG, Wong J. Variation in gene expression patterns in human gastric cancers. Mol Biol Cell 2003; 14: 3208-3215.

20) Kechad A, Jananji S, Ruella Y, Hickson GR. Anillin acts as a bifunctional linker coordinating midbody ring biogenesis during cytokinesis. Curr Biol 2012; 22: 197-203.

21) Hickson GR, O'farrell PH. Anillin: a pivotal organizer of the cytokinetic machinery. Biochem Soc Trans 2008; 36: 439-441.

22) Piekny AJ, Glotzer M. Anillin is a scaffold protein that links RhoA, actin, and myosin during cytokinesis. Curr Biol 2008; 18: 30-36.

23) Piekny A, Werner M, Glotzer M. Cytokinesis: welcome to the Rho zone. Trends Cell Biol 2005; 15: 651-658.

24) Salhia B, Tran NL, Chan A, Wolf A, Nakada M, Rutka F, Ennis M, Mcdonough WS, Berens ME, Symons M. The guanine nucleotide exchange factors trio, Ect2, and Vav3 mediate the invasive behavior of glioblastoma. Am J Pathol 2008; 173: 1828-1838.

25) Nishimura Y, Yonemura S. Centralspindlin regulates ECT2 and RhoA accumulation at the equatorial cortex during cytokinesis. J Cell Sci 2006; 119: 104-114.

26) Zhang M-L, Lu S, Zhou L, Zheng S-S. Correlation between ECT2 gene expression and methylation change of ECT2 promoter region in pancreatic cancer. Hepatobiliary Pancreat Dis Int 2008; 7: 533-538.

27) Sakata H, Rubin JS, Taylor WG, Miki T. A rho-specific exchange factor ect2 is induced from S to M phases in regenerating mouse liver. Hepatology 2000; 32: 193-199.

28) Homeister JW, Willis MS. Atherosclerosis: Pathogenesis, Genetics and Experimental Models. eLS 2010; 15 Nov, DOI:IO 10002/9780470015902. a0005998 pub2.

29) Korpetinou A, Skandalis SS, Labropoulou VT, Smirlaki G, Noulas A, Karamanos NK, Theocharis AD. Serglycin: at the crossroad of inflammation and malignancy. Front Oncol 2013; 3: 327.

30) Schadt EE. Molecular networks as sensors and drivers of common human diseases. Nature 2009; 461: 218-223.