

AFP computational secreted network construction and analysis between human hepatocellular carcinoma (HCC) and no-tumor hepatitis/cirrhotic liver tissues

Lin Wang · Juxiang Huang · Minghu Jiang · Xiguang Zheng

Received: 29 March 2010 / Accepted: 30 April 2010 / Published online: 8 June 2010
© International Society of Oncology and BioMarkers (ISOBM) 2010

Abstract Alpha-fetoprotein (*AFP*) computational secreted network construction and analysis of human hepatocellular carcinoma (HCC) is very useful to identify novel markers and potential targets for prognosis and therapy. By integration of gene regulatory network infer and the database for annotation, visualization, and integrated discovery, we identified and constructed significant molecule *AFP* secreted network from 25 no-tumor hepatitis/cirrhotic liver tissues and 25 HCC patients in the same GEO Dataset GSE10140-10141. Our result verified *AFP* secreted module in the upstream of no-tumor hepatitis/cirrhotic liver tissues (*AMELY*, *LCN2*, and *REG3A* activation; *DKK1*, *SFRP4*, and *SPINK1* inhibition) and its downstream (*PRSSI*, *REG3A*, and *TSHB* activation; *AMELY* and *DKK1* inhibition), and also in the upstream of HCC (*LCN2*, *REG3A*, and *SFRP4* activation; *AMELY* and *DKK1* inhibition) and its downstream (*AMELY* activation; *DKK1*, *LCN2*, *PRSSI*, *SEMA3B*, and *SPINK1* inhibition). Importantly, we data-mined that *AFP* secreted cluster of HCC is involved in disease mutation (only in HCC terms) without

cell surface receptor linked signal transduction, neuroactive ligand–receptor interaction, cell–cell signaling, and pancreas (only in no-tumor hepatitis/cirrhotic liver tissues terms), the condition which is vital to invasion of HCC. Our result demonstrated that common terms in both no-tumor hepatitis/cirrhotic liver tissues and HCC include secreted extracellular region, extracellular region part, extracellular space, signal peptide, signal, disulfide bond, glycosylation site N-linked (GlcNAc...), and glycoprotein, and these terms are less relative to invasion; therefore, we deduced the weaker *AFP* secreted network in HCC consistent with our number computation. We predicted *AFP* high expression localization within cells of HCC and without secretion to extracellular matrix. It would be necessary of *AFP* secreted function to decrease invasion of HCC.

Keywords *AFP* · Human hepatocellular carcinoma (HCC) · Network construction and analysis

Lin Wang, Juxiang Huang, and Minghu Jiang contributed equally to this work.

L. Wang (✉) · J. Huang
Biomedical Center, School of Electronics Engineering,
Beijing University of Posts and Telecommunications,
Beijing 100876, China
e-mail: wanglin98@tsinghua.org.cn

M. Jiang
Lab of Computational Linguistics, School of Humanities
and Social Sciences, Tsinghua Univ.,
Beijing 100084, China

X. Zheng
School of Electrical Computer and Telecommunications
Engineering, University of Wollongong,
Wollongong, NSW 2522, Australia

Introduction

Hepatocellular carcinoma (HCC) is one of the most common causes of cancer-related death. So, development of novel drugs in HCC has become a challenge for biologists. Alpha-fetoprotein (*AFP*) computational secreted network construction and analysis of HCC is very useful to identify novel markers and potential targets for prognosis and therapy.

AFP (fold change=6.065) is one out of 50 genes identified as significantly different in HCC vs no-tumor hepatitis/cirrhotic liver tissues. *AFP* is relevant to molecular function of transfer/carrier protein, and it is relevant to biological process of transport, developmental processes, mesoderm development, and oncogenesis (database for annotation,

visualization, and integrated discovery (DAVID)). *AFP*'s relational study can also be seen in these papers [1–10]. However, the molecular mechanism concerning *AFP* secreted network construction in HCC has received little attention.

In this paper, we first identified the significant molecule *AFP* by significant analysis of microarrays (SAM), then constructed *AFP* up- and downstream network by gene regulatory network infer (GRNInfer), and further data-mined the *AFP* secreted cluster from 25 no-tumor hepatitis/cirrhotic liver tissues from HCC patients and 25 HCC in the same GEO Dataset GSE10140-10141 by using the DAVID. We gained the negative result of *AFP* secreted module through the net numbers of activation minus inhibition compared with no-tumor hepatitis/cirrhotic liver tissues and predicted possibly the decrease of *AFP* secreted module in HCC. Our integrative result showed that *AFP* secreted cluster of HCC contained both in human no-tumor hepatitis/cirrhotic liver tissues and HCC terms of secreted, extracellular region, extracellular region part, extracellular space, signal peptide, signal, disulfide bond, glycosylation site: N-linked (GlcNAc...) and glycoprotein (with *AFP* gene), only in HCC term of disease mutation (without *AFP* gene), and none in HCC terms of cell surface receptor linked signal transduction, neuroactive ligand–receptor interaction, cell–cell signaling, and pancreas (without *AFP* gene) compared with human no-tumor hepatitis/cirrhotic liver tissues; all the conditions are vital to the invasion of HCC. Therefore, we deduced the weaker *AFP* secreted function in HCC consistent with our above computation. We predicted *AFP* high expression localization within cells of HCC and without secretion to extracellular matrix. It would be necessary of *AFP* secreted function to decrease the invasion of HCC. *AFP* secreted interaction module construction in HCC can be a new route for studying the pathogenesis of HCC. Our construction of *AFP* secreted network may be useful to identify novel markers and potential targets for prognosis and therapy of HCC.

Materials and methods

Microarray data

We used microarrays containing 6,144 genes from 25 no-tumor hepatitis/cirrhotic liver tissues from HCC patients and 25 HCC in the same GEO Dataset GSE10140-10141 [11].

Gene selection algorithms

Fifty potential HCC molecular markers were identified using SAM. SAM is a statistical technique for finding significant genes in a set of microarray experiments. The input to SAM is gene expression measurements from a set

of microarray experiments, as well as a response variable from each experiment. The response variable may be a grouping like untreated, treated (either unpaired or paired), and so on. SAM computes a statistic d_i for each gene i , measuring the strength of the relationship between gene expression and the response variable. It uses repeated permutations of the data to determine if the expression of any genes is significantly related to the response. The cutoff for significance is determined by a tuning parameter δ , chosen by the user based on the false-positive rate. We normalized data by \log_2 and selected minimum fold change = 2.78. Here, we chose the 50 top-fold significant (a big difference compared with no-tumor hepatitis/cirrhotic liver tissues) genes under the false-discovery rate, and q values are 0%. The q value (invented by John Storey [12]) for each gene is the lowest false-discovery rate at which that gene is called significant. It is like the well-known p value, but adapted to multiple-testing situations.

Network establishment of candidate genes

The entire network was constructed using GRNInfer [13] and GVedit tools. GRNInfer is a novel mathematic method called gene network reconstruction tool based on linear programming and a decomposition procedure for inferring gene networks. The method theoretically ensures the derivation of the most consistent network structure with respect to all of the datasets, thereby not only significantly alleviating the problem of data scarcity but also remarkably improving the reconstruction reliability. Equation 1 represents all of the possible networks for the same dataset.

$$J = (X' - A)U\Lambda^{-1}V^T + YV^T = \hat{J} + YV^T \quad (1)$$

We established a network based on the 50 top-fold distinguished genes and selected parameters as λ 0.0 because we used one dataset, threshold 0.000001. λ is a positive parameter, which balances the matching and sparsity terms in the objective function. Using different thresholds, we can predict various networks with different edge density.

Functional annotation clustering

The DAVID Gene Functional Clustering Tool provides typical batch annotation and gene-GO term enrichment analysis for highly throughput genes by classifying them into gene groups based on their annotation term co-occurrence [14, 15]. DAVID uses a novel algorithm to measure relationships among the annotation terms based on the degrees of their co-association genes to group similar annotation contents from the same or different resources into annotation groups. The grouping algorithm

is based on the hypothesis that similar annotations should have similar gene members. The functional annotation clustering integrates the same techniques of kappa statistics to measure the degree of the common genes between two annotations, and fuzzy heuristic clustering to classify the groups of similar annotations according to kappa values. The tool also allows observation of the internal relationships of the clustered terms by comparing to the typical linear, redundant term report, over which similar annotation terms may be distributed among many other terms.

Results

Identification of HCC molecular markers

AFP is one out of 50 genes identified as high expression in HCC vs no-tumor hepatitis/cirrhotic liver tissues. The 50

significant genes included serine peptidase inhibitor Kazal type 1 (*SPINK1*); alpha-fetoprotein (*AFP*); P antigen family member 4 (*PAGE4*); v-myb myeloblastosis viral oncogene homolog-like 2 (*MYBL2*); cyclin-dependent kinase inhibitor 3 (*CDKN3*); thyroid stimulating hormone beta (*TSHB*); dickkopf homolog 1 (*DKK1*); ribonucleotide reductase M2 polypeptide (*RRM2*); S100 calcium binding protein P (*S100P*); actinin alpha 2 (*ACTN2*); trophinin associated protein (*TROAP*); ELAV-like 3 (*ELAVL3*); discs large homolog 7 (*DLG7*); NIMA-related kinase 2 (*NEK2*); lipocalin 2 (*LCN2*); histone cluster 1 H3 (*HIST1H3H*); ninjurin 2 (*NINJ2*); protease serine 1 (*PRSS1*); nucleolar and spindle associated protein 1 (*NUSAP1*); insulin-like growth factor 2 mRNA binding protein 3 (*IGF2BP3*); sema domain immunoglobulin domain short basic domain secreted 3B (*SEMA3B*); forkhead box M1 (*FOXM1*); regenerating islet-derived 3 alpha (*REG3A*); FLJ33790; secreted frizzled-related protein 4 (*SFRP4*); baculoviral IAP repeat-containing 5 (*BIRC5*); solute carrier family 17 member 7 (*SLC17A7*); breast cancer 1

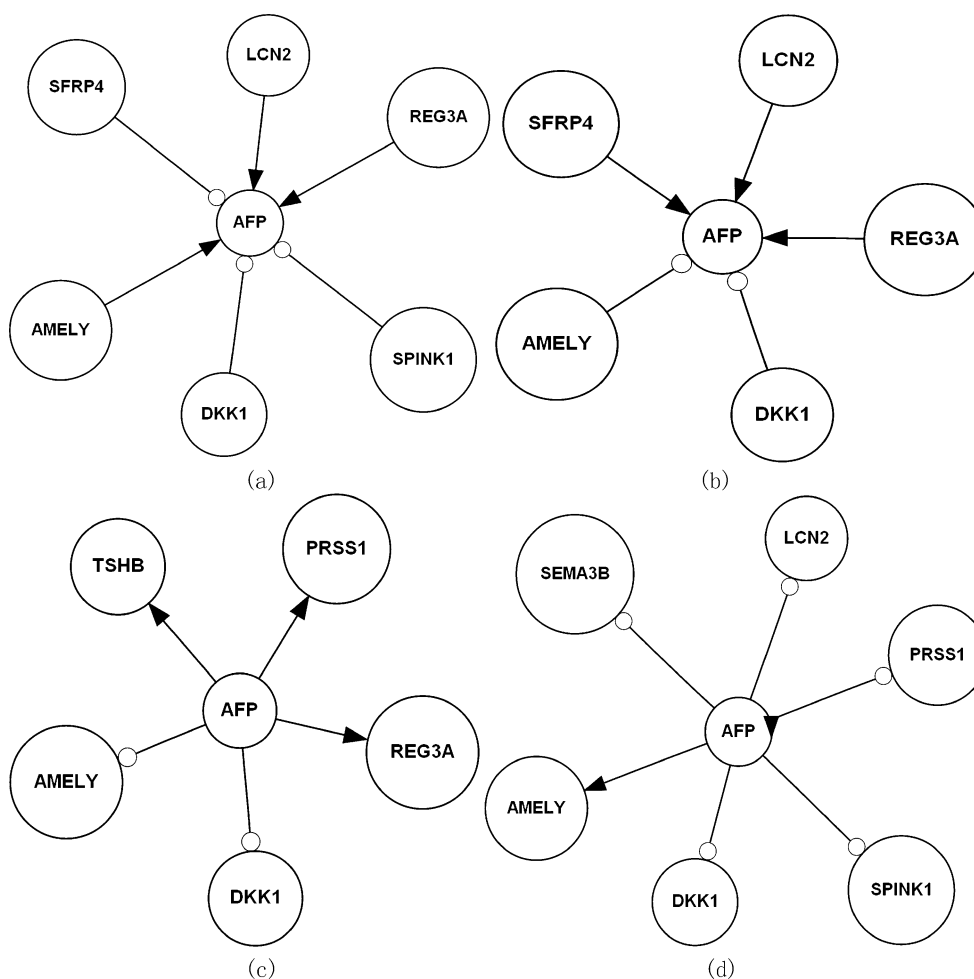


Fig. 1 *AFP* up- and downstream secreted network construction in human no-tumor hepatitis/cirrhotic liver tissues by infer (a, c). *AFP* up- and downstream secreted network construction in HCC by infer (b, d). Arrowheads represent activation; cycle represents inhibition

(*BRCAl*); topoisomerase II alpha (*TOP2A*); nuclear receptor subfamily 5 group A member 1 (*NRS1A1*); X-ray repair complementing defective repair in Chinese hamster cells 2 (*XRCC2*); E2F transcription factor 1 (*E2F1*); desmuslin (*DMN*); roundabout axon guidance receptor homolog 1 (*ROBO1*); aldehyde dehydrogenase 3 family member A1 (*ALDH3A1*); non-SMC condensin I complex subunit H (*NCAPH*); budding uninhibited by benzimidazoles 1 homolog beta (*BUB1B*); homeobox A5 (*HOXA5*); pituitary tumor-

transforming 1 (*PTTG1*); glycerophosphodiester phosphodiesterase domain containing 5 (*GDPD5*); thymidine kinase 1 (*TK1*); amelogenin Y-linked (*AMELY*); protein kinase cGMP-dependent type II (*PRKG2*); ubiquitin-conjugating enzyme E2C (*UBE2C*); sulfotransferase family cytosolic 1C member 2 (*SULT1C2*); glutamate receptor metabotropic 1 (*GRM1*); sex comb on midleg-like 2 (*SCML2*); TTK protein kinase (*TTK*); cyclin E2 (*CCNE2*); cell division cycle 20 homolog (*CDC20*).

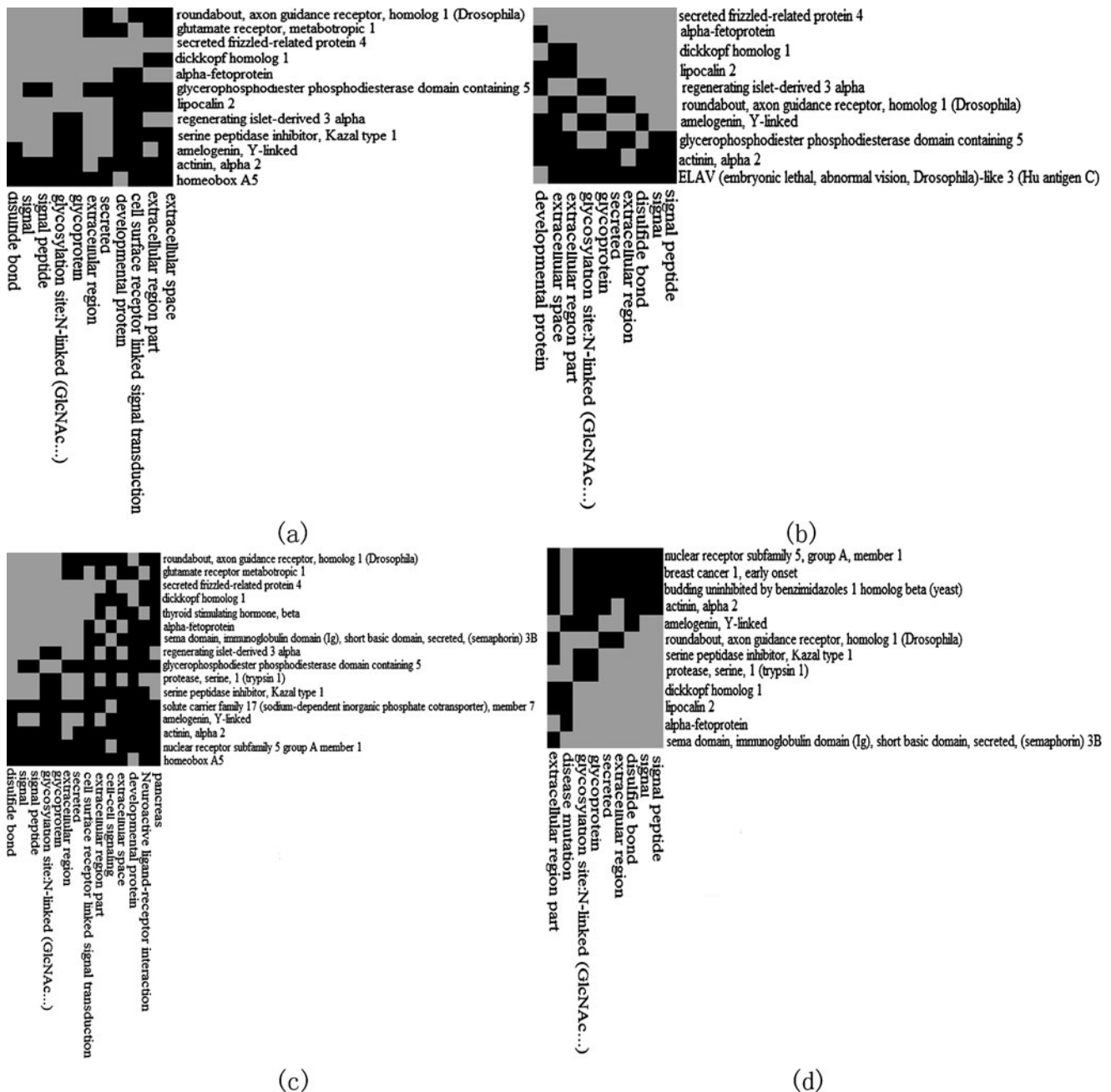


Fig. 2 *AFP* up- and downstream secreted cluster in human no-tumor hepatitis/cirrhotic liver tissues by DAVID (**a**, **c**). *AFP* up- and downstream secreted cluster by DAVID in HCC (**b**, **d**). Gray color represents gene-term association positively reported; black color represents gene-term association not reported yet

Table 1 Up- and downstream gene numbers of activation and inhibition of each module with *AFP* gene in *AFP* secreted cluster between human no-tumor hepatitis/cirrhotic liver tissues and HCC

Term	<i>AFP</i> upstream				<i>AFP</i> downstream			
	con (act)	con (inh)	exp (act)	exp (inh)	con (act)	con (inh)	exp (act)	exp (inh)
Extracellular space	1	1	2	0	2	0		
Extracellular region part 2	1	2	1	2	1	2	1	
Glycosylation site: N-linked (GlcNAc...)	3	3	3	2	2	3	1	4
Glycoprotein 3	3	3	2	2	3	1	4	
Disulfide bond	4	2	4	0	4	1	1	5
Secreted	3	3	3	2	3	2	2	5
Extracellular region	4	3	4	2	3	3	2	6
Signal	5	3	4	2	4	3	2	6
Signal Peptide	5	3	4	2	4	3	2	6

con control (human no-tumor hepatitis/cirrhotic liver tissues), *exp* experiment (HCC), *act* activation, *inh* inhibition

AFP up- and downstream secreted network construction in human no-tumor hepatitis/cirrhotic liver tissues and HCC

In the no-tumor hepatitis/cirrhotic liver tissues, *AFP* upstream secreted network appeared that *AMELY*, *LCN2*, *REG3A* activate *AFP*, and *DKK1*, *SFRP4*, *SPINK1* inhibit *AFP*, as shown in Fig. 1a, whereas in HCC, *AFP* upstream secreted network showed that *LCN2*, *REG3A*, *SFRP4* activate *AFP*, and *AMELY*, *DKK1* inhibit *AFP*, as shown in Fig. 1b.

In the no-tumor hepatitis/cirrhotic liver tissues, *AFP* downstream secreted network reflected that *AFP* activates *PRSS1*, *REG3A*, *TSHB* and inhibits *AMELY*, *DKK1*, as shown in Fig. 1c, whereas in HCC, *AFP* downstream secreted network appeared that *AFP* activates *AFP*, *AMELY* and inhibits *DKK1*, *LCN2*, *PRSS1*, *SEMA3B*, *SPINK1*, as shown in Fig. 1d.

Identification of *AFP* up- and downstream secreted cluster in human no-tumor hepatitis/cirrhotic liver tissues and HCC by DAVID

AFP secreted cluster terms both in human no-tumor hepatitis/cirrhotic liver tissues and HCC include extracellular region, secreted, extracellular region part, extracellular space, signal peptide, signal, disulfide bond, glycosylation

site: N-linked (GlcNAc...) and glycoprotein (with *AFP* gene). Other *AFP* secreted cluster terms only in human no-tumor hepatitis/cirrhotic liver tissues containing cell surface receptor linked signal transduction, neuroactive ligand–receptor interaction (without *AFP* gene), as shown in Fig. 2a, c. However, *AFP* secreted cluster terms only in HCC cover disease mutation (without *AFP* gene), as shown in Fig. 2b, d.

In the no-tumor hepatitis/cirrhotic liver tissues, *AFP* upstream modules mainly include secreted *AMELY*, *LCN2*, *REG3A*, *DKK1*, *SFRP4*, *SPINK1*, *AFP*, etc., as shown in Fig. 2a. In HCC, *AFP* upstream modules mainly cover secreted *LCN2*, *REG3A*, *SFRP4*, *AMELY*, *DKK1*, *AFP*, etc., as shown in Fig. 2b. In the no-tumor hepatitis/cirrhotic liver tissues, *AFP* downstream modules mainly consist of secreted *PRSS1*, *REG3A*, *TSHB*, *AMELY*, *DKK1*, *AFP*, etc., as shown in Fig. 2c. In HCC, *AFP* downstream modules mainly contain secreted *AFP*, *AMELY*, *DKK1*, *LCN2*, *PRSS1*, *SEMA3B*, *SPINK1*, *AFP*, etc., as shown in Fig. 2d.

Discussion

We have already done some work in this relative field about gene network construction and analysis seen in our

Table 2 Activation and inhibition gene names of *AFP* up- and downstream secreted cluster in human no-tumor hepatitis/cirrhotic liver tissues and HCC

Term	con (act)	con (inh)	exp (act)	exp (inh)
<i>AFP</i> upstream				
Secreted	<i>AMELY</i> , <i>LCN2</i> , <i>REG3A</i>	<i>DKK1</i> , <i>SFRP4</i> , <i>SPINK1</i>	<i>LCN2</i> , <i>REG3A</i> , <i>SFRP4</i>	<i>AMELY</i> , <i>DKK1</i>
<i>AFP</i> downstream				
Secreted	<i>PRSS1</i> , <i>REG3A</i> , <i>TSHB</i>	<i>AMELY</i> , <i>DKK1</i>	<i>AFP</i> , <i>AMELY</i>	<i>DKK1</i> , <i>LCN2</i> , <i>PRSS1</i> , <i>SEMA3B</i> , <i>SPINK1</i>

published papers [16–19]. Here, we first indentified the 50 top-fold significant genes in HCC by SAM. Then, we compared *AFP* up- and downstream gene numbers of activation and inhibition between no-tumor hepatitis/cirrhrotic liver tissues and HCC (Table 1). We identified *AFP* secreted module based on the integration of infer and DAVID as follows.

In *AFP* secreted module of upstream network of human no-tumor hepatitis/cirrhrotic liver tissues, our integrative result showed that *AMELY*, *LCN2*, and *REG3A* activate *AFP*, and *DKK1*, *SFRP4*, *SPINK1* inhibit *AFP*, whereas in that of HCC, *LCN2*, *REG3A*, and *SFRP4* activate *AFP*, and *AMELY* and *DKK1* inhibit *AFP*. In *AFP* secreted module of downstream network of human no-tumor hepatitis/cirrhrotic liver tissues, our integrative result reflected that *AFP* activates *PRSSI*, *REG3A*, and *TSHB* and inhibits *AMELY* and *DKK1*, whereas in that of HCC, *AFP* activates *AFP* and *AMELY* and inhibits *DKK1*, *LCN2*, *PRSSI*, *SEMA3B*, and *SPINK1* (Figs. 1, 2; Table 2). *AMELY* has been reported to have molecular function of extracellular matrix, extracellular matrix structural protein, miscellaneous function, and structural protein, and it is relevant to biological process of developmental processes, skeletal development, and mesoderm development (DAVID). *AMELY*'s relational study also can be seen in these papers [20–24]. *SEMA3B*'s molecular function contains signaling molecule and membrane-bound signaling molecule, and it is relevant to biological process of signal transduction, cell surface receptor mediated signal transduction, receptor protein tyrosine kinase signaling pathway, developmental processes, neurogenesis, ectoderm development, tumor suppressor, and oncogenesis (DAVID). *SEMA3B*'s relational study also can be seen in these papers [25–29]. *SPINK1* is relevant to molecular function of select regulatory molecule, protease inhibitor and serine protease inhibitor, and biological process of protein metabolism and modification, proteolysis (DAVID). *SPINK1*'s relational study also can be seen in these papers [30–33]. *PRSSI* has been proved to be concerned with molecular function of protease and serine protease, and biological process of protein metabolism and modification, and proteolysis (DAVID). *PRSSI*'s relational study also can be seen in these papers [34–37]. We gained the negative

result of *AFP* secreted module through the net numbers of activation minus inhibition compared with no-tumor hepatitis/cirrhrotic liver tissues and predicted possibly the decrease of *AFP* secreted module in HCC.

Otherwise, our integrative result showed that *AFP* secreted cluster of HCC contained both in human no-tumor hepatitis/cirrhrotic liver tissues and HCC terms of secreted, extracellular region, extracellular region part, extracellular space, signal peptide, signal, disulfide bond, and glycosylation site: N-linked (GlcNAc...) and glycoprotein (with *AFP* gene), only in HCC term of disease mutation (without *AFP* gene), and none in HCC terms of cell surface receptor linked signal transduction, neuroactive ligand–receptor interaction, cell–cell signaling, and pancreas (without *AFP* gene) compared with human no-tumor hepatitis/cirrhrotic liver tissues; all the conditions are vital to the invasion of HCC (Fig. 2). As we all know, relationships of annotation terms in one cluster have similar annotation contents and correlative functions. Koelink et al. indicated that a slow *AFP* decrease is an important predictor of liver cancer development in further life [38]. Ishigami et al. observed that *AFP*-positive gastric cancer has an aggressive behavior with hematogenous metastasis [39]. Therefore, we deduced the weaker *AFP* secreted function in HCC consistent with our above computation. We predicted *AFP* high expression localization within cells of HCC and without secretion to extracellular matrix. It would be necessary of *AFP* secreted function to decrease the invasion of HCC.

In conclusion, we predicted *AFP* high expression localization within cells of HCC and without secretion to extracellular matrix and possibly the decrease of *AFP* secreted module in HCC compared with no-tumor hepatitis/cirrhrotic liver tissues. It would be necessary of *AFP* secreted module to decrease the invasion of HCC. *AFP* secreted interaction module construction in HCC can be a new route to study the pathogenesis of HCC.

Acknowledgments This work was supported by the National Natural Science Foundation in China (no. 60871100) and the Teaching and Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry. State Key Lab of Pattern Recognition Open Foundation, key project of philosophical and social science of MOE (07JZD0005).

Appendix Table

ID	Gene name	Fold change	<i>q</i> values (%)
<i>SPINK1</i>	Serine peptidase inhibitor Kazal type 1	7.1868541	0
<i>AFP</i>	Alpha-fetoprotein	6.0653532	0
<i>PAGE4</i>	P antigen family, member 4	5.1761949	0
<i>MYBL2</i>	v-myb myeloblastosis viral oncogene homolog-like 2	4.8323113	0

<i>CDKN3</i>	Cyclin-dependent kinase inhibitor 3	4.7876185	0
<i>TSHB</i>	Thyroid-stimulating hormone beta	4.509559	0
<i>DKK1</i>	Dickkopf homolog 1	4.5082799	0
<i>RRM2</i>	Ribonucleotide reductase M2 polypeptide	4.3636217	0
<i>S100P</i>	S100 calcium-binding protein P	4.2705617	0
<i>ACTN2</i>	Actinin alpha 2	4.2504895	0
<i>TROAP</i>	Trophinin-associated protein	4.1758078	0
<i>ELAVL3</i>	ELAV-like 3	4.0291492	0
<i>DLG7</i>	Disks large (<i>Drosophila</i>) homolog-associated protein 5	4.0100004	0
<i>NEK2</i>	NIMA-related kinase 2	3.829819	0
<i>LCN2</i>	Lipocalin 2	3.8033255	0
<i>HIST1H3H</i>	Histone cluster 1, H3	3.7414943	0
<i>NINJ2</i>	Ninjurin 2	3.7372496	0
<i>PRSS1</i>	Protease serine 1	3.7284463	0
<i>NUSAP1</i>	Nucleolar and spindle associated protein 1	3.57542	0
<i>IGF2BP3</i>	Insulin-like growth factor 2 mRNA binding protein 3	3.5582517	0
<i>SEMA3B</i>	Sema domain immunoglobulin domain short basic domain secrete3d.5	34B0229	0
<i>FOXM1</i>	Forkhead box M1	3.4067646	0
<i>REG3A</i>	Regenerating islet-derived 3 alpha	3.3982422	0
<i>FLJ33790</i>	Kelch-like 35	3.3802468	0
<i>SFRP4</i>	Secreted frizzled-related protein 4	3.3698296	0
<i>BIRC5</i>	Baculoviral IAP repeat-containing 5	3.3151171	0
<i>SLC17A7</i>	Solute carrier family 17 member 7	3.2780314	0
<i>BRCA1</i>	Breast cancer 1	3.1450336	0
<i>TOP2A</i>	Topoisomerase (DNA) II alpha	3.1409126	0
<i>NR5A1</i>	nuclear receptor subfamily 5 group A member 1	3.1296242	0
<i>XRCC2</i>	X-ray repair complementing defective repair in Chinese hamster3.c1e2H0s4	2712	0
<i>E2F1</i>	E2F transcription factor 1	3.1123178	0
<i>DMN</i>	Synemin, intermediate filament protein	3.0962649	0
<i>ROBO1</i>	Roundabout, axon guidance receptor homolog 1	3.0915801	0
<i>ALDH3A1</i>	Aldehyde dehydrogenase 3 family memberA1	3.0364491	0
<i>NCAPH</i>	Non-SMC condensin I complex subunit H	3.0068688	0
<i>BUB1B</i>	Budding uninhibited by benzimidazoles 1 homolog beta	2.9853921	0
<i>HOXA5</i>	Homeobox A5	2.9526933	0
<i>PTTG1</i>	Pituitary tumor-transforming 1	2.9346163	0
<i>GDPD5</i>	Glycerophosphodiester phosphodiesterase domain containing 5	2.917215	0
<i>TK1</i>	Thymidine kinase 1	2.9003364	0
<i>AMELY</i>	Amelogenin Y-linked	2.8942756	0
<i>PRKG2</i>	Protein kinase cGMP-dependent type II	2.8765877	0
<i>UBE2C</i>	Ubiquitin-conjugating enzyme E2C	2.8695949	0
<i>SULT1C2</i>	Sulfotransferase family cytosolic 1C member 2	2.849562	0
<i>GRM1</i>	Glutamate receptor metabotropic 1	2.833317	0
<i>SCML2</i>	Sex comb on midleg-like 2	2.817374	0
<i>TTK</i>	TTK protein kinase	2.8139951	0
<i>CCNE2</i>	Cyclin E2	2.7810875	0
<i>CDC20</i>	Cell division cycle 20 homolog	2.778426	0

References

1. Beale G, Chattopadhyay D, Gray J, Stewart S, Hudson M, Day C, et al. Afp, pivkaii, gp3, scca-1 and follisatin as surveillance biomarkers for hepatocellular cancer in non-alcoholic and alcoholic fatty liver disease. *BMC Cancer*. 2008;8:200.
2. Cheng HT, Chang YH, Chen YY, Lee TH, Tai DI, Lin DY. Afp-l3 in chronic liver diseases with persistent elevation of alpha-fetoprotein. *J Chin Med Assoc*. 2007;70:310–7.

3. Chiba N, Yoshioka T, Sakayori M, Mikami Y, Miyazaki S, Akiyama S. Afp-producing hepatoid adenocarcinoma in association with Barrett's esophagus with multiple liver metastasis responding to paclitaxel/cddp: a case report. *Anticancer Res.* 2005;25:2965–8.
4. Hatta W, Kumagai S, Imamura J, Oikawa T, Asano Y, Fujita H. A case of curatively resected afp producing gastric cancer that responded remarkably to 1 course of ts-1 and showed complete loss of multiple liver metastatic tumors. *Gan to kagaku ryoho.* 2005;32:855–8.
5. Hirashima Y, Kitajima K, Sugi S, Kagawa K, Kumamoto T, Murakami K, et al. Successful bi-weekly paclitaxel treatment of an afp-producing gastric cancer patient with peritoneal dissemination and multiple liver metastasis. *Gan to kagaku ryoho.* 2006;33:517–9.
6. Lee HY, Jung JH, Kang YS, Kim YS, Moon HS, Park KO, et al. Clinical significance of transiently elevated serum afp level in developing hepatocellular carcinoma in hbsag positive-liver cirrhosis. *The Korean journal of gastroenterology = Taehan Sohwagi Hakhoe chi.* 2004;43:252–9.
7. Okuse C, Yotsuyanagi H, Takahashi Y, Hayashi T, Suzuki M, Iino S, et al. Gastric cancer with liver metastasis producing alpha-fetoprotein (afp) and protein induced by vitamin k absence or antagonist ii (pivka-ii); a case report. *Nippon Shokakibyo Gakkai zasshi The Japanese journal of gastro-enterology.* 2003;100:28–34.
8. Oshiro Y, Takada Y, Enomoto T, Fukao K, Ishikawa S, Iijima T. A resected case of metachronous liver metastasis from lung cancer producing alpha-fetoprotein (afp) and protein induced by vitamin k absence or antagonist ii (pivka-ii). *Hepato-gastroenterology.* 2004;51:1144–7.
9. Sotiropoulos GC, Molmenti EP, Lang H. Liver transplantation for hepatocellular carcinoma in the meld era: Leading roles of meld score, afp level, and recipient age as predictors of survival. *Dig Dis Sci.* 2009;54:917.
10. Yan Z, Qiao J, Gong H, Hou Y. A novel liver-directed gene delivery system using an autonomously replicating vector specifically expressed in afp positive hepatoma cells. *Science in China.* 1998;41:80–6.
11. Hoshida Y, Villanueva A, Kobayashi M, Peix J, Chiang DY, Camargo A, et al. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. *N Engl J Med.* 2008;359:1995–2004.
12. Storey JD. A direct approach to false discovery rates. *J Roy Stat Soc, Ser B.* 2002;64:479–98.
13. Wang Y, Joshi T, Zhang XS, Xu D, Chen L. Inferring gene regulatory networks from multiple microarray datasets. *Bioinformatics (Oxford, England).* 2006;22:2413–20.
14. da Huang W, Sherman BT, Tan Q, Collins JR, Alvord WG, Roayaei J, et al. The david gene functional classification tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol.* 2007;8:R183.
15. Dennis Jr G, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. David: database for annotation, visualization, and integrated discovery. *Genome Biol.* 2003;4:P3.
16. Sun Y, Wang L, Jiang M, Huang J, Liu Z, Wolf S. Secreted phosphoprotein 1 upstream invasive network construction and analysis of lung adenocarcinoma compared with human normal adjacent tissues by integrative biocomputation. *Cell Biochem Biophys.* 2009;56:59–71. ISSN: 1085-9195 (Print) 1559-0283 (Online).
17. Sun Y, Wang L, Lui L. Integrative decomposition procedure and kappa statistics set up atf2 ion binding module in malignant pleural mesothelioma (mpm). *Frontiers of Electrical and Electronic Engineering in China.* 2008;3:381–7.
18. Wang L, Sun Y, Jiang M, Zhang S, Wolf S. Fos proliferating network construction in early colorectal cancer (crc) based on integrative significant function cluster and inferring analysis. *Cancer Invest.* 2009;27:816–24.
19. Wang L, Sun Y, Jiang M, Zheng X (2009) Integrative decomposition procedure and kappa statistics for the distinguished single molecular network construction and analysis. *J Biomed Biotechnol* 2009. doi:10.1155/2009/726728.
20. Fontanesi L, Scotti E, Russo V. Differences of the porcine amelogenin x and y chromosome genes (amelx and amely) and their application for sex determination in pigs. *Mol Reprod Dev.* 2008;75:1662–8.
21. Mitchell RJ, Kreskas M, Baxter E, Buffalino L, Van Oorschot RA. An investigation of sequence deletions of amelogenin (amely), a y-chromosome locus commonly used for gender determination. *Ann Hum Biol.* 2006;33:227–40.
22. Murphy KM, Cohen JS, Goodrich A, Long PP, Griffin CA. Constitutional duplication of a region of chromosome yp encoding amely, prky, and tbl1y: implications for sex chromosome analysis and bone marrow engraftment analysis. *J Mol Diagn.* 2007;9:408–13.
23. Reddy KS. Clinical management of a rare de novo translocation 46, x, t(y;15) (p11.2 approximately 11.3;q11.2). *Ish t(y;15)(dyz3+, amely+, snrpn+;d15z+)* found prenatally. *Prenat Diagn.* 1998;18:294–7.
24. Yong RY, Gan LS, Chang YM, Yap EP. Molecular characterization of a polymorphic 3-mb deletion at chromosome yp11.2 containing the amely locus in Singapore and Malaysia populations. *Hum Genet.* 2007;122:237–49.
25. Grote HJ, Schmiemann V, Geddert H, Rohr UP, Kappes R, Gabbert HE, et al. Aberrant promoter methylation of p16(ink4a), rarb2 and sema3b in bronchial aspirates from patients with suspected lung cancer. *Int J Cancer.* 2005;116:720–5.
26. Ito M, Ito G, Kondo M, Uchiyama M, Fukui T, Mori S, et al. Frequent inactivation of rassf1a, blu, and sema3b on 3p21.3 by promoter hypermethylation and allele loss in non-small cell lung cancer. *Cancer Lett.* 2005;225:131–9.
27. Koyama N, Zhang J, Huqun, Miyazawa H, Tanaka T, Su X, et al. Identification of igfbp-6 as an effector of the tumor suppressor activity of sema3b. *Oncogene.* 2008;27:6581–9.
28. Marsit CJ, Wiencke JK, Liu M, Kelsey KT. The race associated allele of semaphorin 3b (sema3b) t415i and its role in lung cancer in African-Americans and Latino-Americans. *Carcinogenesis.* 2005;26:1446–9.
29. Nair PN, McArdle L, Cornell J, Cohn SL, Stallings RL. High-resolution analysis of 3p deletion in neuroblastoma and differential methylation of the sema3b tumor suppressor gene. *Cancer Genet Cytogenet.* 2007;174:100–10.
30. O'Reilly DA, Witt H, Rahman SH, Schulz HU, Sargen K, Kage A, et al. The spink1 n34s variant is associated with acute pancreatitis. *Eur J Gastroenterol Hepatol.* 2008;20:726–31.
31. Rerknimitr R, Shotelersuk V, Buranasupkajorn P, Plengpanich W, Snabboon T. A novel spink1 gene mutation, c.206c>t, in a Thai patient with chronic alcoholic pancreatitis. *Jop.* 2008;9:33–6.
32. Shimosegawa T, Kume K, Masamune A. Spink1, adh2, and ald2 gene variants and alcoholic chronic pancreatitis in Japan. *J Gastroenterol Hepatol.* 2008;23 Suppl 1:S82–6.
33. Tomlins SA, Rhodes DR, Yu J, Varambally S, Mehra R, Perner S, et al. The role of spink1 in ETS rearrangement-negative prostate cancers. *Cancer Cell.* 2008;13:519–28.
34. de las Heras-Castano G, Castro-Senosiain B, Fontalba A, Lopez-Hoyos M, Sanchez-Juan P. Hereditary pancreatitis: clinical features and inheritance characteristics of the r122c mutation in the cationic trypsinogen gene (prss1) in six Spanish families. *Jop.* 2009;10:249–55.
35. Felderbauer P, Stricker I, Schnekenburger J, Bulut K, Chromik AM, Belyaev O, et al. Histopathological features of patients with chronic pancreatitis due to mutations in the prss1 gene: Evaluation of braf and kras2 mutations. *Digestion.* 2008;78:60–5.

36. Liu QC, Gao F, Ou QS, Zhuang ZH, Lin SR, Yang B, et al. Novel mutation and polymorphism of prss1 gene in the Chinese patients with hereditary pancreatitis and chronic pancreatitis. *Chin Med J*. 2008;121:108–11.
37. Pho-Iam T, Thongnoppakhun W, Yenchitsomanus PT, Limwongse C. A Thai family with hereditary pancreatitis and increased cancer risk due to a mutation in prss1 gene. *World J Gastroenterol*. 2005;11:1634–8.
38. Koelink CJ, van Hasselt P, van der Ploeg A, van den Heuvel-Eibrink MM, Wijburg FA, Bijleveld CM, et al. Tyrosinemia type I treated by nbc: how does afp predict liver cancer? *Mol Genet Metab*. 2006;89:310–5.
39. Ishigami S, Natsugoe S, Nakashima H, Tokuda K, Nakajo A, Okumura H, et al. Biological aggressiveness of alpha-fetoprotein (afp)-positive gastric cancer. *Hepato-gastroenterology*. 2006;53:338–41.