

Research Article

Signal Propagation in Protein Interaction Network during Colorectal Cancer Progression

Yang Jiang,¹ Tao Huang,² Lei Chen,³ Yu-Fei Gao,¹ Yudong Cai,⁴ and Kuo-Chen Chou^{5,6}

¹ Department of Surgery, China-Japan Union Hospital of Jilin University, Changchun 130033, China

² Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, NY 10029, USA

³ College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China

⁴ Institute of Systems Biology, Shanghai University, Shanghai 200444, China

⁵ King Abdulaziz University, Jeddah, Saudi Arabia

⁶ Gordon Life Science Institute, 53 South Cottage Road, Belmont, MA 02478, USA

Correspondence should be addressed to Yu-Fei Gao; gaoyufei1975@yahoo.com.cn and Yudong Cai; cai_yud@yahoo.com.cn

Received 16 January 2013; Accepted 18 February 2013

Academic Editor: Bin Niu

Copyright © 2013 Yang Jiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Colorectal cancer is generally categorized into the following four stages according to its development or serious degree: Dukes A, B, C, and D. Since different stage of colorectal cancer actually corresponds to different activated region of the network, the transition of different network states may reflect its pathological changes. In view of this, we compared the gene expressions among the colorectal cancer patients in the aforementioned four stages and obtained the early and late stage biomarkers, respectively. Subsequently, the two kinds of biomarkers were both mapped onto the protein interaction network. If an early biomarker and a late biomarker were close in the network and also if their expression levels were correlated in the Dukes B and C patients, then a signal propagation path from the early stage biomarker to the late one was identified. Many transition genes in the signal propagation paths were involved with the signal transduction, cell communication, and cellular process regulation. Some transition hubs were known as colorectal cancer genes. The findings reported here may provide useful insights for revealing the mechanism of colorectal cancer progression at the cellular systems biology level.

1. Background

Cancer is a complex system disease [1]. The complexity reflects in many ways. First, it is a network disease that involves the changes of many genes and these genes are connected in a certain way. Second, the disease network is evolving all the time during the progression. Some efforts have been made to understand such dynamic network [2–6].

As the third most common cancer worldwide [7], colorectal cancer develops via a progressive accumulation of genetic mutations and pathway dysfunctions [6]. It has the following four stages from early to late [8]: Dukes A, B, C, and D. In the stage of Dukes A, the cancer is only limited to the innermost layer. In Dukes B stage, the cancer has grown through the muscle layer. In Dukes C stage, the cancer has spread to the lymph nodes nearby. In Dukes D stage, the cancer is widely spread. The stage of Dukes D is the most advanced stage of

colorectal cancer. Understanding the underlying molecular mechanisms of the pathological changes in colorectal cancer progression will facilitate the development of therapeutic treatments.

In the study of prion disease, it was found that during different stages of the disease, different regions of the network were activated and they formed a clear disease aggravation pattern on the network [2]. However, it is still not clear how one activated region is connected with another and how they can transit into one another.

To investigate the transition processes of different network states, we analyzed the gene expression profiles of 290 colorectal cancer patients, who were at different stages of Dukes A, B, C, and D. Using the Maximum Relevance and Minimum Redundancy (mRMR) [9] and Incremental Feature Selection (IFS) methods [10, 11] to compare the gene expressions among the patients of Dukes A, B, C, and D

stages, we obtained 158 early stage biomarkers and 284 late stage biomarkers, respectively. Subsequently, the early stage biomarkers and the late stage biomarkers were mapped onto the protein interaction network. If the early stage biomarker and the late stage biomarker were close to each other in the network, and also their expression levels were correlated with the patients of the Dukes B and C stages, then we assume that a signal propagation path may exist from the early stage biomarker to the late stage biomarker. Thus, by screening all the possible signal propagation paths from the early stage biomarkers to the late stage biomarkers, we have identified 632 signal propagation paths that contained 473 transition genes.

According to the Gene Ontology (GO) [12] enrichment analysis, many of the transition genes that transmitted the disease signal from the early stage biomarkers to the late stage biomarkers were involved into the signal transduction, cell communication, and cellular process regulation. Some transition hub genes were known colorectal cancer genes. They helped the transduction of the disease signal and the aggravation of colorectal cancer.

One signal propagation path from early stage biomarker MAVS to late stage biomarker GFPT1 was shown as an example. MAVS is an important immune protein and signaling protein in mitochondria [13–15] and GFPT1 is a rate-limiting enzyme of metabolism [16, 17]. It was suggested through our signal propagation analysis that MAVS responded to colorectal cancer in the early stage and then transmitted the disease signal to GFPT1 whose dysfunction further accelerated the colorectal cancer patients into late stage. This kind of in-depth analysis on the signal propagation path may provide useful insights into, or enrich, the understanding of the mechanism of colorectal cancer at the cellular or system biology level.

2. Methods

2.1. Benchmark Dataset. We downloaded the expression profiles of 19,621 genes in 290 colorectal cancer patients [18] from Gene Expression Omnibus (GEO) under accession number GSE14333. Of the 290 colorectal cancer patients, 44 were Dukes stage A, 94 Dukes stage B, 91 Dukes stage C, and 61 Dukes stage D. From Dukes A stage to Dukes D, the colorectal cancer gets more and more severe.

The protein interaction network we used was STRING v9.0 (<http://string-db.org/>) [19]. Each protein interaction in STRING has a confidence score, varying from 0.150 to 1. The confidence score is calculated by integrating the functional associations from genomic context, experiments, conserved coexpression, and previous knowledge with Bayesian method [19]. Suppose the interaction confidence score is denoted by I_{score} , it follows according to the original definition

$$I_{rank} = \begin{cases} \text{low confidence,} & \text{if } I_{score} > 0.150 \\ \text{medium confidence,} & \text{if } I_{score} > 0.400 \\ \text{high confidence,} & \text{if } I_{score} > 0.700 \\ \text{highest confidence,} & \text{if } I_{score} > 0.900, \end{cases} \quad (1)$$

where I_{rank} represents the rank of protein interaction.

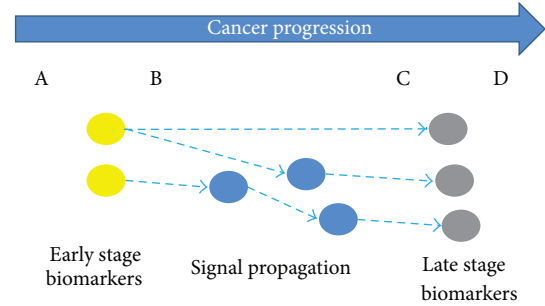


FIGURE 1: The diagram of signal propagation analysis during cancer progression. The blue arrow represents cancer progression. The colorectal cancer has four stages: Dukes A, B, C, and D. From A to D, the cancer gets worse and worse. Yellow nodes and grey nodes represent the biomarkers in the early and late stage, respectively. The goal of signal propagation analysis is to understand the transition from early stage biomarkers to late stage biomarkers by analyzing the signal propagation in the protein interaction network.

2.2. The Diagram of Signal Propagation Analysis during Cancer Progression. In studying or analyzing complex biological systems, it is quite helpful to introduce graphs or diagrams since they can provide an overall view or intuitive insights for the systems investigated, as demonstrated by a series of studies on various important biological topics (see, e.g., [20–29]). In this study, we first constructed a graph G with the PPI data from STRING. In the graph, an edge was assigned for each pair of proteins if they were in interaction with each other. There were 1375295 interaction edges among 15240 proteins. The “intimate degree” between two interacting proteins was defined by

$$I_{intimate} = 1000 \times (1 - I_{score}), \quad (2)$$

where I_{score} is the confidence score between two proteins concerned [19]. Thus, the higher the interaction confidence score between two proteins is, the closer their “interactive distance,” and hence more intimate between them.

Shown in Figure 1 is an illustration for analyzing the signal propagation during the cancer progression. The colorectal cancer has four stages: Dukes A, B, C, and D. From Dukes A to Dukes D, the cancer gets worse and worse. The blue arrow represents the cancer progression. Below, we are to identify the biomarkers in the early stage (yellow nodes) and biomarkers in the late stage (grey nodes). Subsequently, we try to understand the transition from early stage biomarkers to late stage biomarkers by analyzing the signal propagation in the protein interaction network. This kind of analysis may provide useful insights for us to in-depth understand how the signal is propagated through the network.

2.3. Identification of Biomarkers in the Early and Late Stage. The following methods were used to identify the genes between different Dukes stages. First, the Maximum Relevance and Minimum Redundancy (mRMR) [9] method was applied to select the genes that has both maximum relevance with the cancer stages and minimum redundancy to each other. The mRMR program was downloaded from

<http://penglab.janelia.org/proj/mRMR/>. Second, the mRMR ranked genes were optimized with the Incremental Feature Selection (IFS) method [8, 30–35]. During the IFS operation, the accuracies of all possible top gene sets were calculated and the gene set that had the highest prediction accuracy was chosen as the optimal gene set, that is, the biomarkers. The accuracy was examined by the jackknife test, also known as Leave-One-Out Cross Validation (LOOCV) [36–39] and the prediction model was Nearest Neighbor Algorithm (NNA) [40]. The prediction accuracy was defined as the number of correctly predicted samples divided by the number of total samples.

The early stage biomarkers were selected from the Dukes A patients and Dukes B patients with mRMR and IFS methods. The late stage biomarkers were selected from the Dukes C patients and Dukes D patients.

2.4. The Transition from the Early Stage Biomarkers to the Late Stage Biomarkers. The early stage biomarkers and late stage biomarkers were mapped onto weighted protein interaction network graph \mathbb{G} . We identified the shortest paths between them using Dijkstra's algorithm [41–43]. The path length was the sum of edge weights through which the path passed. If the path length was smaller than $1000 \times (1 - 0.700) = 300$, it had high confidence to happen.

Meanwhile, we also tested the correlation between early stage biomarkers and late stage biomarkers in Dukes B patients and Dukes C patients. The Pearson correlation test P values were adjusted with false discovery rate (FDR) [44]. The cutoff of Pearson correlation test FDR was set to 0.001.

Included were those transitions that had the length shorter than 300 and the correlation test FDR smaller than 0.001. The shortest paths from the early stage biomarkers to the late stage biomarkers in the protein interaction network were deemed as the signal propagation paths for the transition.

2.5. Statistical Significance of Signal Propagation Path Identification. To evaluate the statistical significance of the identified signal propagation paths, we estimated the FDR of the signal propagation path based on the permutation [45]. We permuted the gene symbols in protein interaction network and gene expression profiles by 20,000 times. For each of the permutations, we calculated the length of the shortest path based on the weighted protein interaction network and the Pearson correlation test P value adjusted with the FDR method based on the gene expression profiles. The FDR of the signal propagation path was defined as

$$\text{FDR}_{\text{signal-path}} = \frac{N_1}{N_2}, \quad (3)$$

where N_1 was the number of permutations in which the permuted shortest path length is shorter than the actual shortest path length and the permuted Pearson correlation test FDR is smaller than the actual Pearson correlation test FDR, while N_2 the total number of permutations which was 20,000 in this study.

2.6. The Transition Hubs in the Signal Propagation Paths. For each of the transition genes, we calculated the number of shortest paths that crossed it. Those genes that were crossed by more signal propagation paths were deemed more important transition hubs.

3. Results

3.1. Early and Late Stage Biomarkers. By selecting discriminative genes between the Dukes A patients and the Dukes B patients with mRMR and IFS methods, we identified the early stage biomarkers. Similarly, we obtained the late stage biomarkers from the Dukes C patients and the Dukes D patients. The IFS curves of early and late stage biomarker selection were shown in Figures 2(a) and 2(b), respectively. In Figure 2(a), the highest accuracy was 0.891 with 158 genes of the early stage biomarkers. In Figure 2(b), the highest accuracy was 0.855 with 284 genes of the early stage biomarkers. The 158 early stage biomarkers and 284 late stage biomarkers can be found in Supplemental Tables S1 and S2, available online at <http://dx.doi.org/10.1155/2013/287019> respectively.

3.2. Comparison of Early and Late Stage Biomarkers. Now let us compare the early stage biomarkers with the late stage ones. It was observed between the two kinds of biomarkers there was only one gene, RNF4, in common. The expected number of overlap genes should be 2.29 and the odds ratio was 0.432. In other words, there was less overlap than expected. It was reported that in different stages of disease, different regions of the biological network are activated [2] and the dynamics of the biological network reflects the histopathology and clinical changes [6, 46]. The shifting from the activated region of early stage biomarkers to the activated region of late stage biomarkers in the biological network explains the under overlap between the early and late stage biomarkers, which may also help understand the colorectal cancer progression. In the following section, we are to study the transition processes in which the early stage biomarkers propagate the disease-aggravating signal to the late stage biomarkers, triggering the patients to develop into the most severe condition.

3.3. From Early Stage Biomarkers to Late Stage Biomarkers: The Transition. There were 136 early stage biomarkers and 230 late stage biomarkers that could be mapped onto the STRING network. The number of all possible combination pairs between the early and late stage biomarkers was $136 \times 230 = 31,280$, for each of which we calculated their shortest path length that was the sum of the edge weights in the shortest path. Furthermore, we calculated the Pearson correlation test FDR between them in Dukes B patients and Dukes C patients. Two criteria were applied to get the signal propagation path from early stage biomarkers to late stage biomarkers: the path length should be shorter than 300 and the correlation test FDR should be smaller than 0.001. There were 632 such signal propagation paths, as given in Table S3. Such 632 signal propagation paths linked 76 early stage biomarkers and 109 late stage biomarkers. Shown in Figure 3

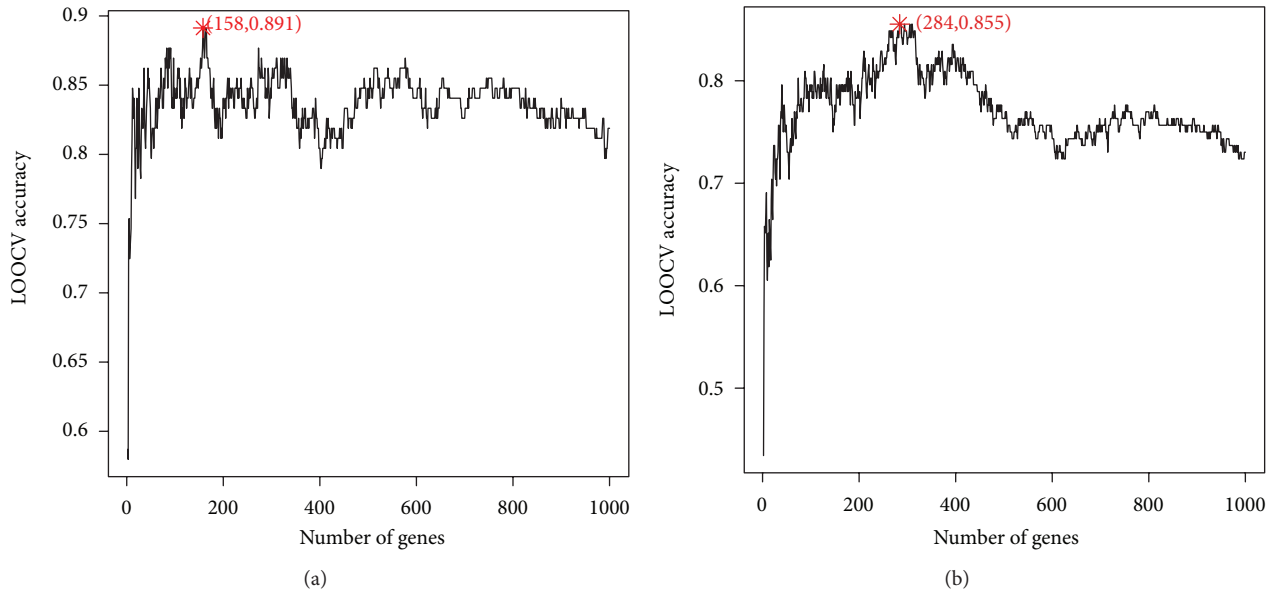


FIGURE 2: The IFS curves of early stage biomarkers and late stage biomarker. (a) The IFS curves of early stage biomarker selection. The highest accuracy was 0.891 with 158 genes which were the early stage biomarkers. (b) The IFS curves of late stage biomarker selection. The highest accuracy was 0.855 with 284 genes which were the late stage biomarkers.

are the transition networks from early stage biomarkers to late stage biomarkers.

Meanwhile, the values of FDR for the identified signal propagation paths were also calculated by first permuting the gene symbols in the protein interaction network and gene expression profiles and then comparing the permuted shortest path length and Pearson correlation FDR with the actual ones. Based on the results of the 20,000 permutations, the statistical significance of each identified signal propagation path was evaluated. It was found that all the 632 identified signal propagation paths were with FDR less than 0.05 and 81.3% of them had FDR less than 0.01.

3.4. The Transition Hubs in Signal Propagation. The 632 signal propagation paths crossed 473 genes. We ranked each of the 473 transition genes based on the number of signal propagation paths that had crossed it. The genes crossed by more signal propagation paths were regarded as more important transition hubs. The detailed results of the 473 transition genes as well as the numbers of signal propagation paths that had crossed them can be found in Table S4. The top three transition hubs were TP53 (tumor protein 53), CTNNB1 (cadherin-associated protein, beta 1), and EP300 (E1A binding protein p300). Interestingly, two of them, TP53 and EP300, were colorectal cancer genes, fully consistent with the reports in the Online Mendelian Inheritance in Man [47] (OMIM, <http://omim.org/entry/114500>).

4. Discussion

4.1. The Biological Functions of Early Stage Biomarkers, Late Stage Biomarkers, and Transition Genes. We used GATHER

[48] (<http://gather.genome.duke.edu/>) to investigate the biological functions of the 158 early stage biomarkers, the 284 late stage biomarkers, and the 473 transition genes. The Gene Ontology (GO) enrichment results thus obtained are shown in Tables 1, 2, and 3, respectively. Since the 473 transition genes were enriched into too many GO terms, only the five enriched GO terms with the highest Bayes factor [49] were shown in Table 3. It is instructive to point out that the late stage biomarkers had more enriched GO terms than the early stage biomarkers. Also, the late stage biomarkers were more enriched in the common GO terms than the early stage biomarkers, such as “GO:0009607: response to biotic stimulus,” “GO:0006952: defense response,” and “GO:0006955: immune response.” The roles of defense response and immune response in colorectal cancer [50, 51] have been widely studied. Many of the transition genes were involved in the signal transduction, cell communication, and cellular process regulation. These kinds of functions played important roles in transducing the disease signal and aggravating the colorectal cancer.

4.2. The Overlapped Gene between Early Stage Biomarkers and Late Stage Biomarkers. One overlapped gene, RNF4 (RING finger protein 4), was observed between the early stage biomarkers and the late stage biomarkers. As reported in [52], RNF4 was a patented biomarker gene of colorectal cancer. Also, as reported in [53], downregulation of RNF4 was related to the colorectal cancer risk (<http://www.wipo.int/patentscope/search/en/WO2010033371>).

Since RNF4 plays a unique role in ubiquitylation [54], DNA demethylation [35], and DNA repair [35], the colorectal cancer progression may involve the abnormal ubiquitylation and demethylation.

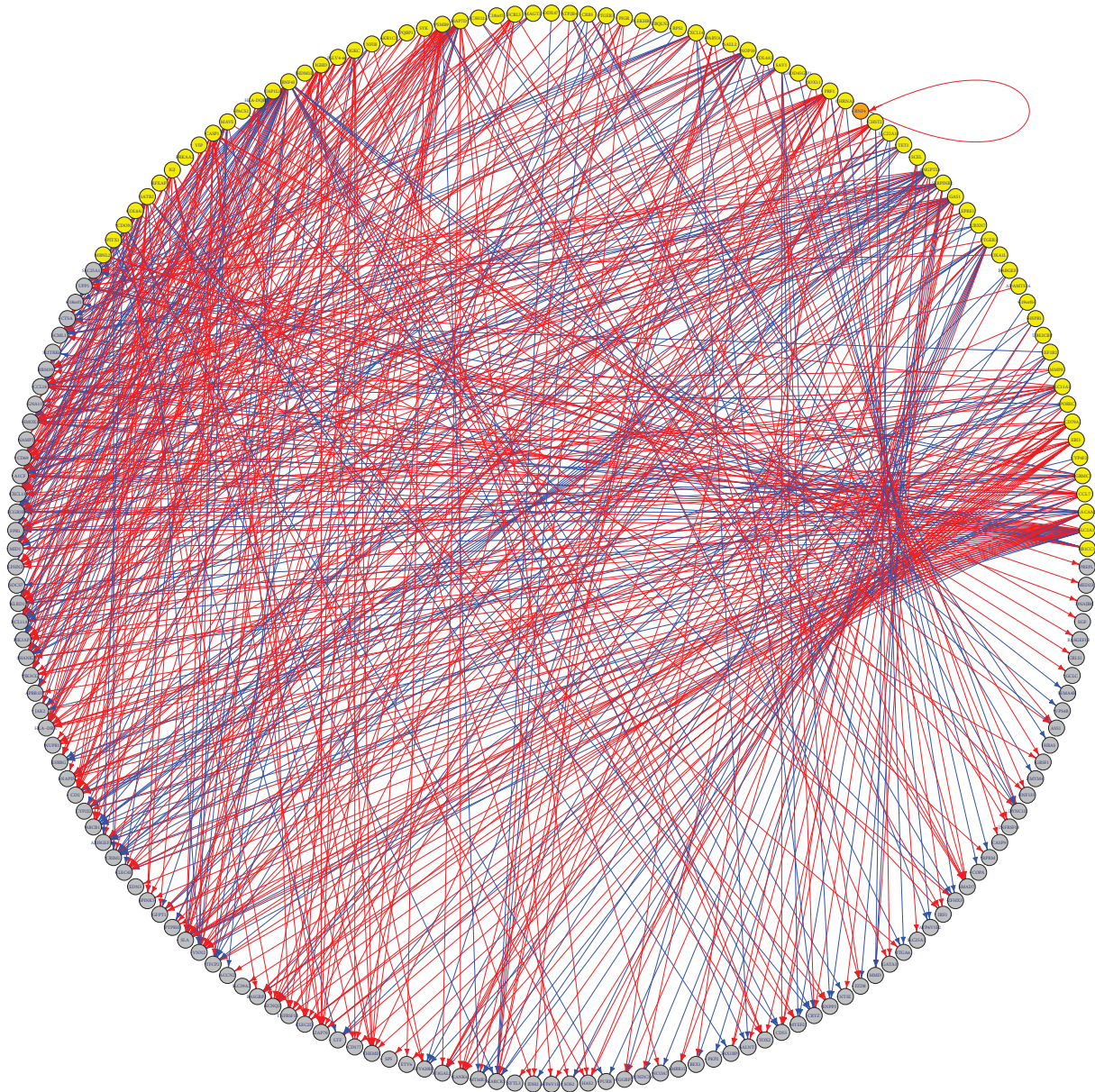


FIGURE 3: The transition network from early stage biomarkers to late stage biomarkers. The yellow and grey nodes were early and late stage biomarkers, respectively. The orange node, RNF4, was both early and late stage biomarker. The red and blue edges indicated that the early and late stage biomarkers were positively and negatively correlated.

4.3. *The Signal Propagation Path from the Early Stage Biomarker MAVS to the Late Stage Biomarker GFPT1.* It is interesting to see that GFPT1 was ranked no. 1 among the late stage biomarkers although it was even not a biomarker in the early stage. We traced back in the signal propagation paths and found GFPT1 was the downstream of the following seven early stage biomarkers: MAVS, TET3, GAS1, ANGPTL4, MAP7D1, CEACAM1, and PGRMCL. Among the 158 early stage biomarkers, MAVS was ranked no. 4, but MAVS was not a late stage biomarker. The Pearson correlation test P value and Pearson correlation coefficient between the expression levels of MAVS and GFPT1 in the Dukes B patients and the Dukes C patients were $1.09e - 05$ and 0.317, respectively. Shown in Figure 4 is the signal propagation path from MAVS

to GFPT1 in the STRING network: MAVS \rightarrow IRF3 \rightarrow CREBBP \rightarrow TP53 \rightarrow ATF3 \rightarrow ATF4 \rightarrow ASNS \rightarrow GLUL \rightarrow GFPT1.

Mitochondrial antiviral signaling (MAVS) protein is important in innate immunity [13–15]. The antibody able to induce immune responses can be used to treat cancer [55]. Immune responses usually occur early in the cancer progression stage but later the cancer cells may develop an ability to escape the immune-mediated lysis [56]. This might explain why MAVS was an early stage biomarker, but not a late stage biomarker.

GFPT1 is the key enzyme in hexosamine synthesis pathway whose products have been implicated in O-linked N-acetylglucosamine (O-GlcNAc) protein modification, insulin

TABLE 1: The enriched GO terms of the 158 early stage biomarkers with adjusted P value less than 0.01.

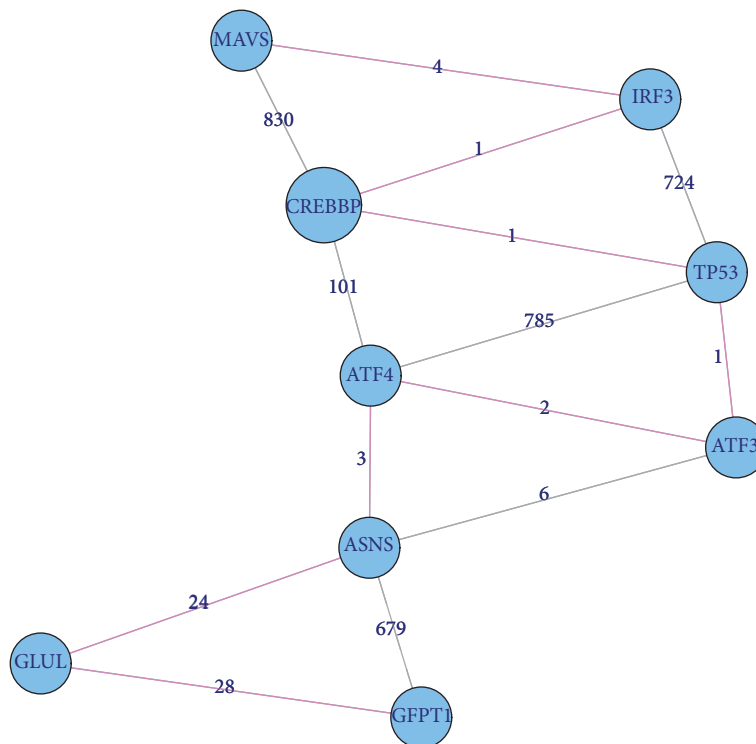
Gene ontology	Number of input genes with the annotation	Adjusted P value
GO:0009607: response to biotic stimulus	16	0.001
GO:0006952: defense response	14	0.004
GO:0006955: immune response	13	0.004

TABLE 2: The enriched GO terms of the 284 late stage biomarkers with adjusted P value less than 0.01.

Gene ontology	Number of input genes with the annotation	Adjusted P value
GO:0006952: defense response	25	0.0002
GO:0006955: immune response	23	0.0002
GO:0016064: immunoglobulin mediated immune response	8	0.0006
GO:0006959: humoral immune response	9	0.001
GO:0009607: response to biotic stimulus	25	0.002
GO:0019730: antimicrobial humoral response	6	0.005

TABLE 3: The most enriched five GO terms of the 473 transition genes.

Gene ontology	Number of input genes with the annotation	Adjusted P value	Bayes factor
GO:0008283: cell proliferation	107	<0.0001	47
GO:0007154: cell communication	219	<0.0001	43
GO:0007165: signal transduction	191	<0.0001	43
GO:0051244: regulation of cellular physiological process	71	<0.0001	40
GO:0050794: regulation of cellular process	84	<0.0001	38

FIGURE 4: The signal propagation path from MAVS to GFPT1. The signal propagation path from MAVS to GFPT1 was MAVS \rightarrow IRF3 \rightarrow CREBBP \rightarrow TP53 \rightarrow ATF3 \rightarrow ATF4 \rightarrow ASNS \rightarrow GLUL \rightarrow GFPT1. The genes in the signal propagation path were mapped onto STRING network. The number on the edge was the edge weight. The edges on the signal propagation path were highlighted with pink color.

resistance, and glucose toxicity [16, 17]. It is a molecular therapeutic target for type-2 diabetes [57, 58]. As a metabolic disease, cancer is always accompanied with impaired mitochondrial function and dysfunctional energy metabolism [59].

Accordingly, it is rational to deduce the signal propagation from MAVS to GFPT1 as follows: in mitochondria, as an important innate immunity protein, MAVS may response to colorectal cancer in a very early stage. Then as a signaling protein, it transmits its signal to GFPT1 that has close relationship with mitochondria. The perturbation of GFPT1 may cause the dysfunction of mitochondria in the energy metabolism. The fates of the cells may be doomed by the collapse of their energy systems.

5. Conclusions

Our results indicated that the strong signals of early stage biomarkers would not necessarily disappear during the colorectal cancer progression, but might be transferred to other late stage biomarkers. This finding may provide useful insights for in-depth analyzing the signal propagation paths and helping to reveal the cellular mechanism of colorectal cancer aggravation.

Authors' Contribution

Y. Jiang and T. Huang contributed equally to this work.

Acknowledgments

This work was supported by Grants from National Basic Research Program of China (2011CB510102, 2011CB510101), Innovation Program of Shanghai Municipal Education Commission (no. 12YZ120, no. 12ZZ087), Natural Science Fund Projects of Jilin province (201215059), Development of Science and Technology Plan Projects of Jilin province (20100733, 201101074), and SRF for ROCS, SEM (2009-36), Scientific Research Foundation (Jilin Department of Science & Technology, 200705314, 20090175, 20100733), Scientific Research Foundation (Jilin Department of Health, 2010Z068), SRF for ROCS (Jilin Department of Human Resource & Social Security, 2012-2014).

References

- [1] I. G. Khalil and C. Hill, "Systems biology for cancer," *Current Opinion in Oncology*, vol. 17, no. 1, pp. 44–48, 2005.
- [2] D. Hwang, I. Y. Lee, H. Yoo et al., "A systems approach to prion disease," *Molecular Systems Biology*, vol. 5, p. 252, 2009.
- [3] T. Huang, Y.-D. Cai, L. Chen et al., "Selection of reprogramming factors of induced pluripotent stem cells based on the protein interaction network and functional profiles," *Protein and Peptide Letters*, vol. 19, no. 1, pp. 113–119, 2012.
- [4] T. Huang, J. Zhang, L. Xie et al., "Crosstissue coexpression network of aging," *OMICS A Journal of Integrative Biology*, vol. 15, no. 10, pp. 665–671, 2011.
- [5] T. Huang, L. Liu, Q. Liu et al., "The role of Hepatitis C Virus in the dynamic protein interaction networks of Hepatocellular cirrhosis and Carcinoma," *International Journal of Computational Biology and Drug Design*, vol. 4, no. 1, pp. 5–18, 2011.
- [6] E. R. Fearon and B. Vogelstein, "A genetic model for colorectal tumorigenesis," *Cell*, vol. 61, no. 5, pp. 759–767, 1990.
- [7] N. Ismaili, "Treatment of colorectal liver metastases," *World Journal of Surgical Oncology*, vol. 9, article 154, 2011.
- [8] Drug-and-Therapeutics-Bulletin, "Population screening for colorectal cancer," *Drug and Therapeutics Bulletin*, vol. 44, no. 9, pp. 65–68, 2006.
- [9] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1226–1238, 2005.
- [10] T. Huang, J. Zhang, Z.-P. Xu et al., "Deciphering the effects of gene deletion on yeast longevity using network and machine learning approaches," *Biochimie*, vol. 94, no. 4, pp. 1017–1025, 2012.
- [11] T. Huang, Z. S. He, W. R. Cui et al., "A sequence-based approach for predicting protein disordered regions," *Protein and Peptide Letters*, vol. 20, no. 3, pp. 243–248, 2013.
- [12] M. Ashburner, C. A. Ball, J. A. Blake et al., "Gene ontology: tool for the unification of biology," *Nature Genetics*, vol. 25, no. 1, pp. 25–29, 2000.
- [13] M. Papatrifiantafyllou, "Innate immunity: MAVS build-ups for defence," *Nature Reviews Immunology*, vol. 11, no. 9, pp. 570–571, 2011.
- [14] D. Arnoult, F. Soares, I. Tattoli, and S. E. Girardin, "Mitochondria in innate immunity," *EMBO Reports*, vol. 12, no. 9, pp. 901–910, 2011.
- [15] F. Hou, L. Sun, H. Zheng, B. Skaug, Q.-X. Jiang, and Z. J. Chen, "MAVS forms functional prion-like aggregates to activate and propagate antiviral innate immune response," *Cell*, vol. 146, no. 3, pp. 448–461, 2011.
- [16] K. Liu, G. Wang, S. H. Zhao et al., "Molecular characterization, chromosomal location, alternative splicing and polymorphism of porcine GFAT1 gene," *Molecular Biology Reports*, vol. 37, no. 6, pp. 2711–2717, 2010.
- [17] T.-J. Hsieh, T. Lin, P.-C. Hsieh, M.-C. Liao, and S.-J. Shin, "Suppression of Glutamine:Fructose-6-phosphate amidotransferase-1 inhibits adipogenesis in 3T3-L1 adipocytes," *Journal of Cellular Physiology*, vol. 227, no. 1, pp. 108–115, 2012.
- [18] R. N. Jorissen, P. Gibbs, M. Christie et al., "Metastasis-associated gene expression changes predict poor outcomes in patients with Dukes stage B and C colorectal cancer," *Clinical Cancer Research*, vol. 15, no. 24, pp. 7642–7651, 2009.
- [19] L. J. Jensen, M. Kuhn, M. Stark et al., "STRING 8—a global view on proteins and their functional interactions in 630 organisms," *Nucleic Acids Research*, vol. 37, no. 1, pp. D412–D416, 2009.
- [20] G. P. Zhou and M. H. Deng, "An extension of Chou's graphic rules for deriving enzyme kinetic equations to systems involving parallel reaction pathways," *Biochemical Journal*, vol. 222, no. 1, pp. 169–176, 1984.
- [21] K. C. Chou, "Graphic rules in steady and non-steady state enzyme kinetics," *Journal of Biological Chemistry*, vol. 264, no. 20, pp. 12074–12079, 1989.

- [22] K. C. Chou, "Applications of graph theory to enzyme kinetics and protein folding kinetics. Steady and non-steady-state systems," *Biophysical Chemistry*, vol. 35, no. 1, pp. 1–24, 1990.
- [23] I. W. Althaus, A. J. Gonzales, J. J. Chou et al., "The quinoline U-78036 is a potent inhibitor of HIV-1 reverse transcriptase," *Journal of Biological Chemistry*, vol. 268, no. 20, pp. 14875–14880, 1993.
- [24] K. C. Chou, F. J. Kezdy, and F. Reusser, "Kinetics of processive nucleic acid polymerases and nucleases," *Analytical Biochemistry*, vol. 221, no. 2, pp. 217–230, 1994.
- [25] J. Andraos, "Kinetic plasticity and the determination of product ratios for kinetic schemes leading to multiple products without rate laws—new methods based on directed graphs," *Canadian Journal of Chemistry*, vol. 86, no. 4, pp. 342–357, 2008.
- [26] K. C. Chou, "Graphic rule for drug metabolism systems," *Current Drug Metabolism*, vol. 11, no. 4, pp. 369–378, 2010.
- [27] G. P. Zhou, "The disposition of the LZCC protein residues in wenxiang diagram provides new insights into the protein-protein interaction mechanism," *Journal of Theoretical Biology*, vol. 284, no. 1, pp. 142–148, 2011.
- [28] K. C. Chou, W. Z. Lin, and X. Xiao, "Wenxiang: a web-server for drawing wenxiang diagrams," *Natural Science*, vol. 3, pp. 862–865, 2011.
- [29] G. P. Zhou, "The structural determinations of the leucine zipper coiled-coil domains of the cGMP-dependent protein kinase I α and its interaction with the myosin binding subunit of the myosin light chains phosphase," *Protein and Peptide Letters*, vol. 18, no. 10, pp. 966–978, 2011.
- [30] B. Q. Li, T. Huang, L. Liu, Y. D. Cai, and K. C. Chou, "Identification of colorectal cancer related genes with mRMR and shortest path in protein-protein interaction network," *PLoS ONE*, vol. 7, Article ID e33393, 2012.
- [31] H. Huang, J. Wang, Y. D. Cai, H. Yu, and K. C. Chou, "Hepatitis C virus network based classification of Hepatocellular cirrhosis and carcinoma," *PLoS ONE*, vol. 7, Article ID e34460, 2012.
- [32] T. Huang, Z. Xu, L. Chen, Y. D. Cai, and X. Kong, "Computational analysis of HIV-1 resistance based on gene expression profiles and the virus-host interaction network," *PLoS ONE*, vol. 6, no. 3, Article ID e17291, 2011.
- [33] T. Huang, S. Wan, Z. Xu et al., "Analysis and prediction of translation rate based on sequence and functional features of the mRNA," *PLoS ONE*, vol. 6, no. 1, Article ID e16036, 2011.
- [34] T. Huang, W. Cui, L. Hu, K. Feng, Y. X. Li, and Y. D. Cai, "Prediction of pharmacological and xenobiotic responses to drugs based on time course gene expression profiles," *PLoS ONE*, vol. 4, no. 12, Article ID e8126, 2009.
- [35] X. V. Hu, T. M. A. Rodrigues, H. Tao et al., "Identification of RING finger protein 4 (RNF4) as a modulator of DNA demethylation through a functional genomics screen," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 34, pp. 15087–15092, 2010.
- [36] T. Huang, C. Wang, G. Zhang, L. Xie, and Y. Li, "SySAP: a system-level predictor of deleterious single amino acid polymorphisms," *Protein and Cell*, vol. 3, no. 1, pp. 38–43, 2012.
- [37] T. Huang, S. Niu, Z. Xu et al., "Predicting transcriptional activity of multiple site P53 mutants based on hybrid properties," *PLoS ONE*, vol. 6, no. 8, Article ID e22940, 2011.
- [38] K. C. Chou, "Prediction of protein cellular attributes using pseudo amino acid composition," *Proteins*, vol. 43, pp. 246–255, 2001, Erratum: vol. 44, p. 60, 2001.
- [39] K. C. Chou and C. T. Zhang, "Review: prediction of protein structural classes," *Critical Reviews in Biochemistry and Molecular Biology*, vol. 30, pp. 275–349, 1995.
- [40] K. C. Chou, "Some remarks on protein attribute prediction and pseudo amino acid composition (50th Anniversary Year Review)," *Journal of Theoretical Biology*, vol. 273, pp. 236–247, 2011.
- [41] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Mathematik*, vol. 1, no. 1, pp. 269–271, 1959.
- [42] G. Chartrand and O. R. Oellermann, *Applied and Algorithmic Graph Theory*, McGraw-Hill College, 1992.
- [43] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, MIT Press and McGraw-Hill, 2nd edition, 2001.
- [44] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *Journal of the Royal Statistical Society B*, vol. 57, pp. 289–300, 1995.
- [45] Y. Xie, W. Pan, and A. B. Khodursky, "A note on using permutation-based false discovery rate estimates to compare different analysis methods for microarray data," *Bioinformatics*, vol. 21, no. 23, pp. 4280–4288, 2005.
- [46] Y. Cai, T. Huang, L. Hu, X. Shi, L. Xie, and Y. Li, "Prediction of lysine ubiquitination with mRMR feature selection and analysis," *Amino Acids*, vol. 42, no. 4, pp. 1387–1395, 2011.
- [47] A. Hamosh, A. F. Scott, J. S. Amberger, C. A. Bocchini, and V. A. McKusick, "Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders," *Nucleic Acids Research*, vol. 33, pp. D514–D517, 2005.
- [48] J. T. Chang and J. R. Nevins, "GATHER: a systems approach to interpreting genomic signatures," *Bioinformatics*, vol. 22, no. 23, pp. 2926–2933, 2006.
- [49] S. N. Goodman, "Toward evidence-based medical statistics. 2: the Bayes factor," *Annals of Internal Medicine*, vol. 130, no. 12, pp. 1005–1013, 1999.
- [50] M. Czéh, C. Loddenkemper, S. Shalpour et al., "The immune response to sporadic colorectal cancer in a novel mouse model," *Oncogene*, vol. 29, no. 50, pp. 6591–6602, 2010.
- [51] C. S. D. Roxburgh and D. C. McMillan, "The role of the in situ local inflammatory response in predicting recurrence and survival in patients with primary operable colorectal cancer," *Cancer Treatment Reviews*, vol. 38, no. 5, pp. 451–466, 2012.
- [52] Q. Chen, Z. Ye, S. C. Lin, and B. Lin, "Recent patents and advances in genomic biomarker discovery for colorectal cancers," *Recent Patents on DNA and Gene Sequences*, vol. 4, no. 2, pp. 86–93, 2010.
- [53] H.-J. Terng, W.-J. Lee, and C.-Y. Chen, "Molecular markers for lung and colorectal carcinomas," WO2010033371, 2010.
- [54] A. Plechanovová, E. G. Jaffray, S. A. McMahon et al., "Mechanism of ubiquitylation by dimeric RING ligase RNF4," *Nature Structural and Molecular Biology*, vol. 18, no. 9, pp. 1052–1059, 2011.
- [55] J. Hess, P. Ruf, and H. Lindhofer, "Cancer therapy with trifunctional antibodies: linking innate and adaptive immunity," *Future Oncology*, vol. 8, no. 1, pp. 73–85, 2012.

- [56] I. B. . Barsoum, T. K. Hamilton, X. Li et al., "Hypoxia induces escape from innate immunity in cancer cells via increased expression of ADAM10: role of nitric oxide," *Cancer Research*, vol. 71, no. 24, pp. 7433–7441, 2011.
- [57] K. C. Chou, "Molecular therapeutic target for type-2 diabetes," *Journal of Proteome Research*, vol. 3, no. 6, pp. 1284–1288, 2004.
- [58] K. C. Chou, "Structural bioinformatics and its impact to biomedical science," *Current Medicinal Chemistry*, vol. 11, no. 16, pp. 2105–2134, 2004.
- [59] T. N. Seyfried and L. M. Shelton, "Cancer as a metabolic disease," *Nutrition and Metabolism*, vol. 7, article 7, 2010.