

Gene screening of colorectal cancers via network analysis

Vahid Mansouri¹, Mostafa Rezaei Tavirani¹, Sina Rezaei Tavirani²

¹*Proteomics Research Center, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran*

²*Proteomics Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran*

ABSTRACT

Aim: Identifying crucial genes related to colorectal cancers via protein-protein interaction (PPI) network analysis is the aim of this study.

Background: colorectal cancer as major reason of mortality is evaluated by genetic and proteomic approaches to find suitable biomarkers. Chromosomal instability plays crucial role in CRC. Expression change of large numbers of genes is reported.

Methods: Differentially expressed genes related to CRCs which obtained from different proteomic methods were extracted from a review article of Paula Álvarez-Chaver *et al.* The genes interacted by Cytoscape software via STRING database. The central nodes determined and were enriched for biological terms by ClueGO. Action map for central genes was illustrated by CluePedia. The critical genes in CRC were introduced.

Results: Among 123 query genes, 114 one recognized by software and were included in the network. SRC, EGFR, PCNA, IL8, CTNBN1, TIMP1, CDH1, and HSPD1 were determined as central genes. After gene ontology analysis SRC, EGFR, and CDH1 were identified as critical genes related to CRC.

Conclusion: It seems that SRC, EGFR, and CDH1 and the related pathways are possible biomarkers for CRC.

Keywords: Colorectal cancer, Biomarker, Gene.

(Please cite as: **Mansouri V, Rezaei Tavirani M, Rezaei Tavirani S. Gene screening of colorectal cancers via network analysis. Gastroenterol Hepatol Bed Bench 2019;12(2):149-154.**)

Introduction

Colorectal cancer (CRC) is known as a major reason of morbidity and mortality by cancers (1). There are evidence that early detection and treatment of CRCs are correlated to decrement of mortality rate (2). Serious efforts by researchers are opened new fields in personalized cancer medicine about CRCs to find precise diagnostic and therapeutic methods based on genetic milieu (3). It is reported that clinical behavior of CRCs is caused by different molecular level especially chromosomal instability (4). Dereglulation of many genes such as EGFR and cyclooxygenase 2 in CRC patients is confirmed (5, 6). It is reported that carcinoembryonic antigen (CEA) and carbohydrate antigen 19-9 (CA 19-9) levels are elevated in CRCs (7). Since expression change of large numbers of genes is

associated with diseases, introducing limited numbers of them as diagnostic or therapeutic biomarkers is aim of numerous researchers. Using screening methods and validation tests can provide useful tools to find effective biomarkers (8, 9).

PPI network analysis is an efficient method to screen large numbers of genes based on interaction between them. In this approach the query genes are interacted in and interactome (10) in two style; scale free network and non-scale free network. Usually gene networks are scale free and the elements of network can be differentiated based on centrality parameters such as degree, betweenness centrality, closeness centrality, and stress (11, 12). The nodes of network which have high numbers of connections are known as hubs and play crucial role

Received: 14 September 2018 Accepted: 18 December 2018
Reprint or Correspondence: Mostafa Rezaei Tavirani,
PhD. *Proteomics Research Center, Faculty of*

Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
E-mail: tavirany@yahoo.com
ORCID ID: 0000-0003-1767-7475

a main connected component (105 nodes and 409 edges). Top 10% of nodes based on degree value including SRC, EGFR, CTNNB1, CDH1, IL8, PCNA, TIMP1, HSPD1, PTPRC, and SERPINA1 were selected as hubs. Top 10 nodes regarding betweenness centrality contain SRC, EGFR, PCNA, IL8, CTNNB1, TIMP1, HSPB1, CDH1, SPTAN1, and HSPD1 were determined as bottlenecks. Common hubs and bottlenecks including SRC, EGFR, PCNA, IL8, CTNNB1, TIMP1, CDH1, and HSPD1 were identified as hub-bottlenecks. All hub-bottlenecks were included in top nodes based on closeness centrality and stress (see table 1). Action map including expression, activation, and inhibition relative to the hub-bottlenecks is illustrated in the figure 2.

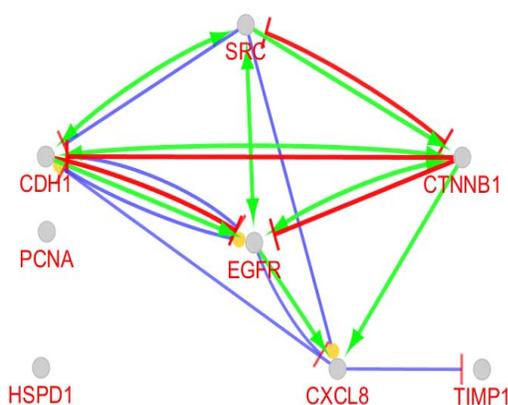


Figure 2. Action map including activation, inhibition, and expression relative to the 8 hub-bottlenecks. Blue, green, and red colors refer to expression, activation, and inhibition actions. Kapa score is considered as default value in CluePedia.

Further investigation about biological terms related to the hub-bottlenecks is shown in the figure 3. Numbers of 60 terms are identified which are attributed to the 8 central nodes. The determined terms are clustered in 6 classes including SRC auto phosphorylation is positively regulated by In1A-bound CDH1, Phosphorylation of EGFR by SRC kinase, proMMP9 binds TIMP1, interleukin-8 receptor binding, purine-specific mismatch base pair DNA N-glycosylase activity, and regulation of T cell mediated immune response to tumor cell. Graphical presentation of 6 groups is shown in the figure 4.

Discussion

Introducing 123 cancer markers is an opportunity to find the best ones. It is possible to screen and select the

crucial individuals that are prominent roles in onset and progress of CRC. Numbers of 114 genes among 123 ones (about 93%) were recognized by STRING database. The network was included 9 isolated nodes (about 8%) and a main connected component contains 105 nodes and 409 edges. Numbers of 8 central nodes were introduced. In many studies hubs or bottlenecks are introduced as central node, however it is reported that hub-bottlenecks are crucial genes or proteins.

In this investigation the hub-bottlenecks which were top nodes based on both closeness centrality and stress were introduced as central nodes. This procedure indicates that the selected central nodes are potent central nodes and with high rate of confidence are the critical nodes of the studied network. Action map analysis revealed that except PCNA and HSPD1 the other central nodes are participated in the action map. As it is shown in the table 1, these two genes are in the row positions 6 and 8, therefore the weaker hubs are among the identified central nodes. Also, it is clear that row 7 is occupied by TIMP1. This gene is an individual among interacted genes in action map which only is regulated by CXCL8 (IL8) and has no regulatory effect on the other genes. Ignoring Timp1, the 5 remained genes including SRC, EGFR, IL8, CTNNB1, and CDH1 are highly connected to each other. As it is depicted in the figures 3 and 4, numbers of 60 biological terms related to the 8 central nodes are identified which are groups in the 6 clusters. Two clusters including “SRC autophosphorylation is positively regulated by In1A-bound CDH1” and “Phosphorylation of EGFR by SRC kinase” contain 51 terms or 85% all biological terms. It seems that these two classes of biological terms are the core of terms related to the central nodes. As it is appeared both groups are concerned by SRC, CDH1, and EGFR genes. As it is shown in the action map IL8 (CXCL8) has no inhibition link with the other elements of network and it is activated by CTNNB1 and EGFR. IL8 also has no activator activity on the other nodes of the network. CTNNB1 is the third elements in the table 1. Comparison between CTNNB1 and the top 2 central nodes shows considerable differences between them based on degree vale and the other centrality parameters. Finally based on these findings, SCR, EGFR, and CDH1 can be considering as critical genes which are involved in CRC. In following part, functional roles of these three critical genes in body and cancers are discussed:

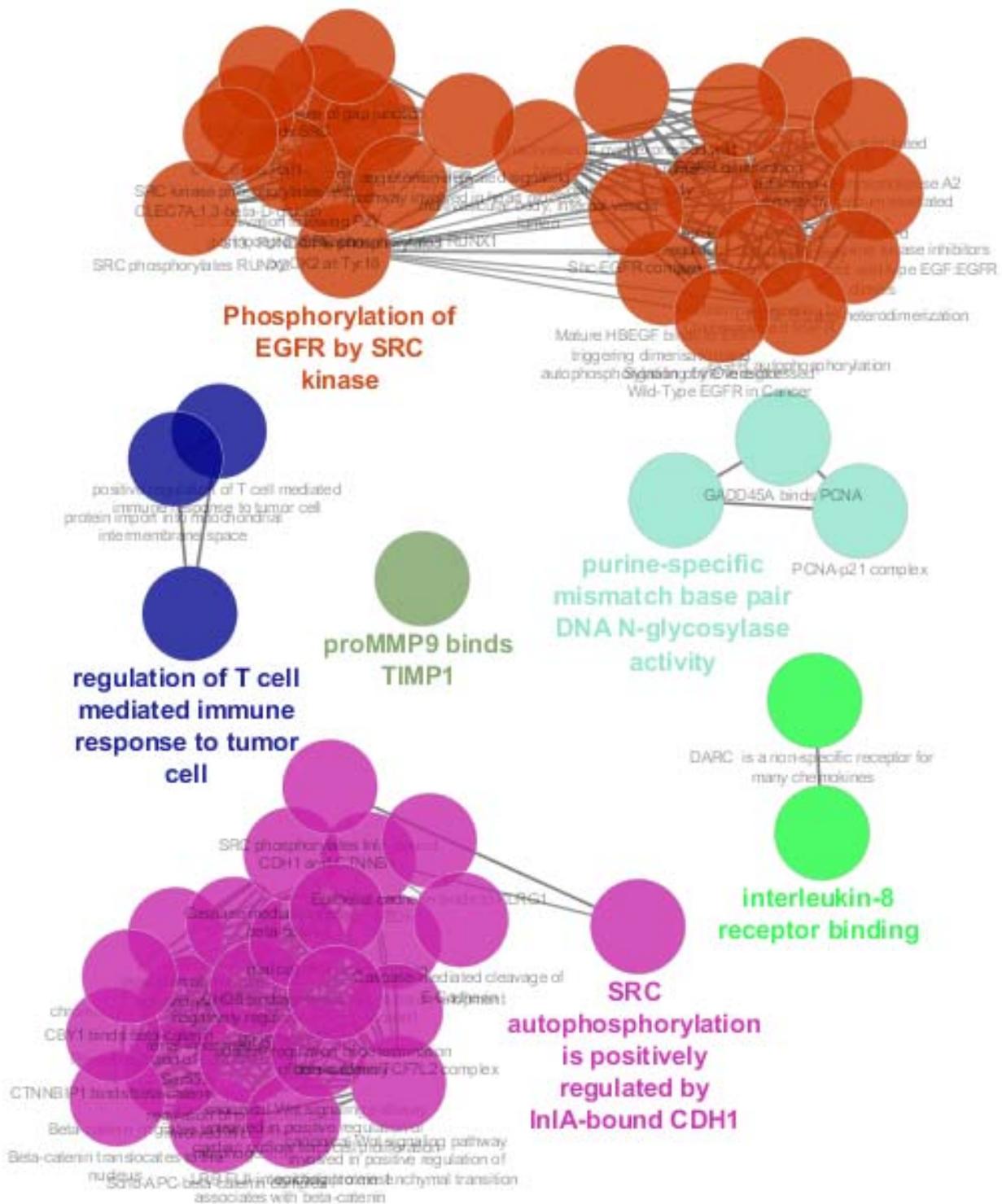


Figure 3. Biological terms relative to the 8 hub-bottlenecks. The colored terms are names of groups. Grouping p value and p value ≤ 0.05 were considered.

SRC is an important element in signal transduction pathways in normal and cancerous cells. Different

pathways are identified that are related to SRC such as cell survival, adhesion, division and motility (20).

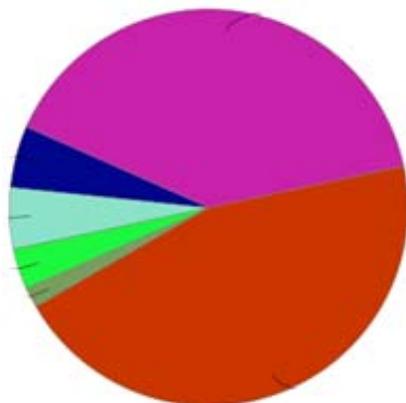


Figure 4. Six clusters of biological term relative to the 8 central nodes are shown. Names of groups correspond to the group colors are written in the foot of figure.

- SRC auto phosphorylation is positively regulated by InLA-bound CDH.
- Phosphorylation of EGFR by SRC kinase.
- proMMP9 binds TIMP1.
- interleukin-8 receptor binding.
- purine-specific mismatch base pair DNA N-glycosylase activity.
- Regulation of T cell mediated immune response to tumor cell.

Investigations indicate that SRC is up-regulated in various cancers such as brain, colon, pancreas, and breast cancers. High level of SRC protein is found in cancerous tissue relative to adjacent normal tissue. It is reported that there is correlation between SRC level and malignancy process progression (21). Bosutinib, dasatinib, saracatinib, and KX01 are introduced as SRC inhibitors (22). EGFR the other crucial genes codes the transmembrane tyrosine kinase which activated by EGF family ligand. EGFR activation leads to activation of MAPK signaling pathway that is responsible for cell proliferation promotion (23). Deregulation of EGFR in wide varieties of tumors is reported. Since inhibition of EGFR in chemotherapy is an important aim, two classes of EGFR inhibitors are identified including monoclonal antibodies such as cetuximab (Erbix) which target extracellular domain of EGFR and the other class of small molecule tyrosine kinase inhibitors as like erlotinib and gefitinib (24).

CDH1 the third crucial protein in this analysis is a critical agent in cell-cell adhesion. CDH1 is known as tumor suppressor and its down regulation in several cancers such as gastric cancer is reported (25). There is evidence that CDH1 gene hyper methylation is associated with CRC (26). As it is reported CDH1 and CDH13 methylation in serum can be considered as cervical cancer markers (27).

It seems that SRC, EGFR, and CDH1 are suitable biomarker candidates for CRC. However, quantity profiling of the elements of this panel may be useful tool in diagnosis and follow of CRC patients. Besides diagnostic features, these genes may be a appreciate drug targets in CRC disease.

Acknowledgment

This project is supported by Shahid Beheshti University of Medical Sciences

Conflict of interests

The authors declare that they have no conflict of interest.

References

1. Hagggar FA, Boushey RP. Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors. *Clin Colon Rectal Surg* 2009;22:191-7
2. Winawer SJ, Fletcher RH, Miller L, Godlee F, Stolar M, Mulrow C, et al. Colorectal cancer screening: clinical guidelines and rationale. *Gastroenterol* 1997;112:594-642.
3. Bardelli A, Siena S. Molecular mechanisms of resistance to cetuximab and panitumumab in colorectal cancer. *J Clin Oncol* 2010;28:1254-61.
4. Markowitz SD, Bertagnolli MM. Molecular basis of colorectal cancer. *New England J Med* 2009;361:2449-60.
5. Khambata-Ford S, Garrett CR, Meropol NJ, Basik M, Harbison CT, Wu S, et al. Expression of epiregulin and amphiregulin and K-ras mutation status predict disease control in metastatic colorectal cancer patients treated with cetuximab. *J Clin Oncol* 2007;25:3230-7.
6. Kashfi SM, Behboudi Farahbakhsh F, Nazemalhosseini Mojarad E, Mashayekhi K, Azimzadeh P, Romani S, et al. Interleukin-16 polymorphisms as new promising biomarkers for risk of gastric cancer. *Tumour Biol*. 2016;37:2119-26.
7. Kim NH, Lee MY, Park JH, Park DI, Sohn CI, Choi K, et al. Serum CEA and CA 19-9 Levels are Associated with the Presence and Severity of Colorectal Neoplasia. *Yonsei Med J* 2017;58:918-24.

154 Genes expression difference among hepatitis B infected and uninfected individuals

8. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. *Genome Biol* 2011;12:R60.
9. Rifai N, Gillette MA, Carr SA. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat Biotechnol* 2006;24:971.
10. Network CGA. Comprehensive molecular characterization of human colon and rectal cancer. *Nat* 2012;487:330.
11. Safari-Alighiarloo N, Rezaei-Tavirani M, Taghizadeh M, Tabatabaei SM, Namaki S. Network-based analysis of differentially expressed genes in cerebrospinal fluid (CSF) and blood reveals new candidate genes for multiple sclerosis. *Peer J* 2016;4:e2775.
12. Rezaei-Tavirani M, Rezaei-Tavirani S, Ahmadi N, Naderi N, Abdi S. Pancreatic adenocarcinoma protein-protein interaction network analysis. *Gastroenterol Hepatol Bed Bench* 2017;10:S85.
13. Zamanian-Azodi M, Rezaei-Tavirani M, Hasanzadeh H, Rad SR, Dalilan S. Introducing biomarker panel in esophageal, gastric, and colon cancers; a proteomic approach. *Gastroenterol Hepatol Bed Bench* 2015;8:6.
14. Zamanian-Azodi M, Rezaei-Tavirani M, Rahmati-Rad S, Hasanzadeh H, Tavirani MR, Seyyedi SS. Protein-Protein Interaction Network could reveal the relationship between the breast and colon cancer. *Gastroenterol Hepatol Bed Bench* 2015;8:215.
15. Álvarez-Chaver P, Otero-Estévez O, De la Cadena MP, Rodríguez-Berrocal FJ, Martínez-Zorzano VS. Proteomics for discovery of candidate colorectal cancer biomarkers. *World Journal of Gastroenterology: WJG* 2014;20:3804.
16. Zhi J, Sun J, Wang Z, Ding W. Support vector machine classifier for prediction of the metastasis of colorectal cancer. *Int J Mol Med* 2018;41:1419-26.
17. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res* 2017;45:D362-8.
18. Bindea G, Galon J, Mlecnik B. CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. *Bioinformatics* 2013;29:661-3.
19. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 2009;25:1091-3.
20. Sen B, Johnson FM. Regulation of SRC family kinases in human cancers. *J Signal Transduct* 2011;2011:865819.
21. Irby RB, Yeatman TJ. Role of Src expression and activation in human cancer. *Oncogene* 2000;19:5636.
22. Creedon H, Brunton VG. Src kinase inhibitors: promising cancer therapeutics? *Crit Rev Oncog* 2012;17:145-59.
23. Metzger B, Chambeau L, Begon DY, Faber C, Kayser J, Berchem G, et al. The human epidermal growth factor receptor (EGFR) gene in European patients with advanced colorectal cancer harbors infrequent mutations in its tyrosine kinase domain. *BMC Med Genetics* 2011;12:144.
24. Kohsaka S, Nagano M, Ueno T, Suehara Y, Hayashi T, Shimada N, et al. A method of high-throughput functional evaluation of EGFR gene variants of unknown significance in cancer. *Sci Trans Med* 2017;9:eaan6566.
25. Liu X, Chu KM. E-cadherin and gastric cancer: cause, consequence, and applications. *Biomed Res Int* 2014;2014:637308.
26. Wheeler J, Kim H, Efstathiou J, Ilyas M, Mortensen NM, Bodmer W. Hypermethylation of the promoter region of the E-cadherin gene (CDH1) in sporadic and ulcerative colitis associated colorectal cancer. *Gut* 2001;48:367-71.
27. Widschwendter A, Ivarsson L, Blassnig A, Müller HM, Fiegl H, Wiedemair A, et al. CDH1 and CDH13 methylation in serum is an independent prognostic marker in cervical cancer patients. *Int J cancer* 2004;109:163-6.