Int J Hum Genet, 18(2): 152-161 (2018)
DOI: 10.31901/24566330.2018/18.2.698

Significant Roles of Aberrant DNA Methylations in Colorectal Cancer

Xiao-Ling Niu[1], Yun-Lei Bai[2] and Min Pang[3]

[1]Department of Traditional Chinese Medicine, Shanghai Pudong New District Zhoupu Hospital, Shanghai, 201318, China
[2]Department of Gastroenterology, The First Hospital of Yulin, Yulin, 718000, Shaanxi Province, China
[3]Department of Ultrasonic Medicine, Xuzhou Central Hospital, Xuzhou, 221009, Jiangsu Province, China

KEYWORDS Colorectal Cancer. DNA Methylation. Differential Enrichment Analysis. Hub. Network

ABSTRACT The aim of this study was to investigate aberrant DNA methylations in colorectal cancer (CRC). Based on DNA methylation data collected from the ArrayExpress, differentially methylated genes (DMGs) between CRC patients and normal controls were identified using *t*-test. Subsequently, the feasibility of these DMGs were determined using a hierarchical clustering analysis and significant biological functions and processes involved by these DMGs were assessed by functional enrichment analyses. Ultimately, a protein-protein interaction (PPI) network was constructed and the topological centrality analysis was applied to screen the hub DMGs for CRC. A total of 1,123 DMGs between CRC samples and normal samples were obtained. Moreover, the DMGs had a good classification performance with accuracy = 0.9935, which suggested the feasibility and confidence of these DMGs. 37 significant gene ontology (GO) terms and 31 significant pathways were gained for CRC patients. 7 hub DMGs were extracted from the sub-network of PPI network

INTRODUCTION

Colorectal cancer (CRC), a cancer originates from the epithelial cells lining the colon or rectum, is one of most common malignancies and the second highest cause of cancer-related deaths in males and third largest in females (Jemal et al. 2011; Yan and Guo 2015). It is characterized by blood in the stool, a change in bowel movements, weight loss and feeling tired all the time, due to the body aging, lifestyle factors, and accumulation of epigenetic and genetic events (Giovannucci 2003; Migliore and Migheli 2011). It has been reported that more than 1.4 million new cases are diagnosed with CRC every year, and more than 690,000 die from the disease (Brenner 2014). Since CRC is often found at an advanced stage, early detection and diagnosis have become particularly important (Walsh and Terdiman 2003). Furthermore, a comprehensive understanding of pathological mechanism underlying CRC is essential and urgent for early prevention and treatment of this tumor.

For the diagnosis of CRC, biomarkers that have high sensitivity and specificity are paramount. To the best of our knowledge, multiple studies have paid attention to genetic changes in CRC (Mais et al. 2015; Qiu et al. 2016; Wood et al. 2007). Several mutated genes have been frequently detected in CRC patients, including *APC*, *BRAF*, *KRAS*, *PIK3CA* and *TP53*, and are regarded as drivers of colorectal tumorigenesis (Wood et al. 2007). However, there is a few researchers focus on epigenetic difference existed in CRC patients. Briefly, DNA methylation, is one of major types of epigenetic modifications closely linked to CRC (Baylin and Jones 2016). Besides, it is associated with numerous key processes, such as microRNA expression regulation (He et al. 2011), gene silencing (Geiman and Robertson 2002) and alternative gene splicing in cancer (Flores et al. 2012). Houlihan et al described that aberrant DNA methylation of CpG islands was deposited in the earliest detectable lesions in the colonic mucosa, aberrant crypt foci (Houlihan 2002). DNA methylation epigenotypes could be suitable for classification markers between CRC patients and normal controls

Address for correspondence:
Min Pang
Department of Ultrasonic Medicine, Xuzhou Central Hospital,
No.199 on Jiefang South Road, Xuzhou, 221009, Jiangsu Province, China
Telephone and Fax: 86-0516-83863918
E-mail: minpangxz@163.com

(Yagi et al. 2010). Importantly, epigenetic switching of H3K27me3 and DNA methylation mainly occurred at genes that were expressed in CRC (Jiang et al. 2008). Furthermore, colorectal tumors with *KRAS* mutations may also be associated with a unique DNA methylation profile (Goel et al. 2010). Despite the fact that several diagnostic panels have been developed, the pathology of CRC still remains poor.

Objective

Therefore, in this paper, to better understanding potential molecular mechanism of CRC, differentially methylated genes (DMGs) were identified based on the DNA methylation data. Subsequently, a hierarchical clustering analysis was conducted on the DMGs to verify the confidence of them. Functional enrichment analyses were employed to explore significant gene sets of the DMGs. Finally, a protein-protein interaction (PPI) network was constructed on DMGs, and further topological centrality analysis was carried out to extract hub DMGs. These results might provide evidence of the cumulative roles of epigenetic mechanisms in CRC, and shed new lights on CRC early diagnosis and treatment.

METHODOLOGY

DNA Methylation Data Collection

In this paper, the raw DNA methylation data for CRC with accessing number E-GEOD-25062 (Hinoue et al. 2012) were recruited from the public free ArrayExpress database (http://www.ebi.ac.uk/arrayexpress/). In brief, E-GEOD-25062 was consisted of 125 colorectal tumor samples and 29 normal-adjacent colonic tissue samples. Besides, it presented on A-GEOD-8490 - Illumina HumanMethylation27 BeadChip (Illumina Inc., California, USA), which enabled the simultaneous quantitative measurements of 275,578 CpG sites.

Data Quality Control and Normalization

During this step, the raw DNA methylation data were obeyed to a set of rigorous quality controls and normalizations. If probes were satisfied with one of the following criteria, it would be removed from the data. The criteria were: 1) the distance from CpG to single-nucleotide polymorphism (SNP) <2; 2) minimum allelic frequency < 0.05; 3) probes of cross-hybridising; and 4) probes on X and Y chromosomes. The filtrated methylation data were normalized according to the beta-mixture quantile normalization method, which improves the robustness of the normalization procedure and reduces the technical variation and bias (Teschendorff et al. 2013). Ultimately, a total of 25,628 CpGs were determined, and defined as DNA methylation data for in-depth exploitation.

Differential Methylation Analysis

Generally, methylation at individual CpGs is reported as a methylation β-value, which is a quantitative measure of methylation for each CpG site with range between 0 (no methylation) to 1 (completely methylated) (Wu et al. 2016). Herein, the β-values (percent methylation changes) were calculated to CRC group and normal group, respectively. In order to reduce the number of non-variable CpGs and improve the statistical power of subsequent analyses, the sits with β-values > 0.8 and < 0.2 were eliminated in all samples. And then the absolute difference for mean â-values in CRC samples and controls was counted, named as score for the CpG. Most important, we implemented the *t*-test to identify differential methylated CpGs between two groups, and the cut-off was set as $P < 0.01$ and Score > 0.05. The genes covered by differential methylated CpGs were considered to DMGs for CRC patients.

Hierarchical Clustering Analysis

For purpose of evaluating the classification performance of DMGs and validating the feasibility of the present approach, a hierarchical clustering analysis was applied across 125 CRC samples and 29 normal controls using the Cluster 3.0 (University of Tokyo, Tokyo, Japan) (Hoon 2002). The algorithm was set to complete linkage clustering using an uncentered correlation and to determine clusters of similar data in multidimensonal spaces. Ideally, the samples should be classified into 2 major clusters: CRC cases and normal controls. The present study tested the method by measuring the percentage of test samples that could be correctly classified. Supposing that CRC belonged to positive samples, and normal controls were attributed to negative

samples, accuracy was computed as following (Mohammadi et al. 2011).

$$\text{Accuracy} = \frac{TN+TP}{TN+TP+FN+FP}$$

Where TP (true positive) represented the number of positive samples correctly predicted as positive; TN (true negative) stood for the number of negative samples correctly predicted as negative; FP (false positive) was the number of negative samples incorrectly predicted as positive and FN (false negative) referred to the number of positive samples incorrectly predicted as negative. Of note, a high accuracy indicated a good classification performance, and further validated the confidence of DMGs and the feasibility of the presented method.

Functional Enrichment Analyses

To further investigate significant functional gene sets enriched by DMGs of CRC, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were conducted using the Database for Annotation, Visualization and Integrated Discovery (DAVID, https://david.ncifcrf.gov/) tool. Here, the DAVID provides a comprehensive set of functional annotations for investigators to understand biological meaning behind large list of genes (Huang da et al. 2009). Specifically, GO produces a dynamic, controlled vocabulary that can be applied to all eukaryotes even as knowledge of gene and protein roles in cells is accumulating and changing (Ashburner et al. 2000; Wang et al. 2018; Zhou et al. 2018). While KEGG is a knowledge base for systematic analysis of gene functions, linking genomic information with higher order functional information (Kanehisa and Goto 2000; Pita-Juarez and Altschuler 2018). In addition, the expression analysis systematic explored (EASE) test (Wang and Simon 2011) implemented in DAVID was utilized to filtrate significant GO terms and pathways with $P < 0.01$ and Count > 20.

PPI Network Construction and Topological Analysis

With an attempt to investigate interactions among DMGs, a PPI network was constructed based on the Search Tool for the Retrieval of Interacting Genes/proteins (STRING, http://string.embl.de/) database, which offers a critical assessment and integration of PPIs, including direct as well as indirect associations (Szklarczyk et al. 2014). Furthermore, the network was visualized by Cytoscape (http://www.cytoscape.org/) software. Here, Cytoscape is a free software package for visualizing, modeling and analyzing the integration of bimolecular interaction networks with high-throughput expression data and other molecular states (Smoot et al. 2011). In order to evaluate the importance for individual genes in the PPI network of CRC, topological centrality analysis was conducted using the degree index. Degree quantifies the local topology of each node, by summing up the number of its adjacent nodes (Wasserman 1994).

RESULTS

DMGs

In the present study, DNA methylation data with 275,578 CpGs for CRC patients were recruited from the ArrayExpress database. When conducting quality control and normalization, the CpGs which were not satisfied with the criteria were removed. As a result, 25,628 methylated CpGs remained in the final dataset of 154 samples for CRC. Here, the β-values were utilized in graphical representations of the data and demonstrated the percentage of methylation counted by methylated. And then a *t*-test was utilized to evaluate the differentially methylated CpGs under the thresholds of $P < 0.01$. As shown in Figure 1, a volcano plot exhibiting the distribution of the 4,792 differentially methylated CpGs (representing 3,503 genes). Furthermore, a score was computed for everyone of 4,792 differentially CpGs based on its methylation β-value, and the Score > 0.05 was regarded as the other threshold for selection of DMGs. Consequently, a total of 1,494 differentially CpGs (covering 1,123 genes) were reserved, and the 1,123 genes were defined as DMGs between CRC patients and normal controls for further exploitation in the subsequent analysis.

Evaluation for the DMGs

To assess the classification performance for 1,123 DMGs of CRC and validate the feasibility of the present method, a hierarchical clustering analysis was conducted, and the cluster heatmap was illustrated in Figure 2. Globally, there were distinguished methylation patterns for CRC samples and normal controls, which segregated

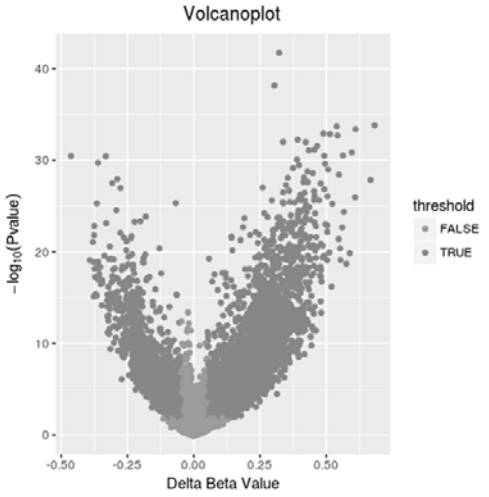

Fig. 1. Volcano plot exhibiting methylation data of 125 colorectal cancer (CRC) samples and 6 normal samples. X axis stood for the mean methylation differences between CRC and normal. Y axis represented the log transformed P values. A total of 4,792 CpG sites were considered significantly differently methylated, shown in blue
Source: Author

samples into two distinct groups approximately. Specifically, 124 CRC samples and 29 normal controls were correctly distinct, whereas 1 CRC samples were wrongly attributed to normal samples. Thus, the accuracy for these DMGs was 0.9935, which suggested that the DMGs had a good classification performance across CRC patients and normal controls and even enhanced the stability and confidence of the present study.

Enriched GO Terms by DMGs

For purpose of revealing significant biological functions in CRC compared with normal controls, GO functional enrichment analyses on DMGs were performed. When setting the thresholds as P < 0.01 and Count > 20, a total of 264 significant GO terms were detected. Step further, to reinforce the relationship between these terms and CRC progression, we refined the 264 significant GO terms by the condition of Count > 50, and the results were displayed in Table 1. Total 37 significant GO terms with P < 0.01 and Count > 50 were gained, of which calcium ion homeostasis (P = 3.68E-14, Count = 82), cell-cell

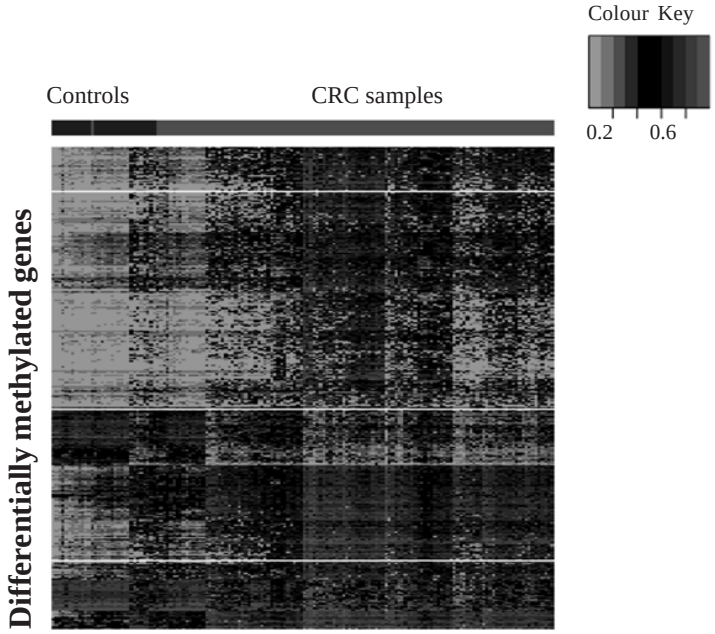

Fig. 2. Hierarchical clustering analysis of significantly differentially methylated genes (DMGs) between colorectal cancer (CRC) patients and normal samples. Each row was an individual DMG and each column stood for a different sample. Color gradation from green to red denoted low to high DNA methylation respectively, with â-values ranging from 0 (no methylation; green) to 1 (complete methylation; red)
Source: Author

Table 1: Significant GO terms with P < 0.01 and Count > 50

ID	*GO term*	*P value*	*Count*
GO:0055074	calcium ion homeostasis	3.68E-14	82
GO:0198738	cell-cell signaling by wnt	3.68E-14	76
GO:0006874	cellular calcium ion homeostasis	3.50E-13	70
GO:0051962	positive regulation of nervous system development	2.45E-11	69
GO:0007423	sensory organ development	6.03E-09	69
GO:0044708	single-organism behavior	1.08E-12	67
GO:0048514	blood vessel morphogenesis	7.12E-08	65
GO:0007389	pattern specification process	2.85E-09	64
GO:0006875	cellular metal ion homeostasis	1.23E-07	63
GO:0044057	regulation of system process	4.71E-07	63
GO:0072507	divalent inorganic cation homeostasis	3.30E-09	62
GO:0008015	blood circulation	5.67E-07	62
GO:0003013	circulatory system process	7.21E-07	62
GO:0007507	heart development	1.03E-06	62
GO:0048568	embryonic organ development	1.18E-08	60
GO:0010720	positive regulation of cell development	9.34E-08	60
GO:0042391	regulation of membrane potential	4.41E-11	59
GO:0030900	forebrain development	9.86E-11	59
GO:0055074	calcium ion homeostasis	3.07E-09	59
GO:0072503	cellular divalent inorganic cation homeostasis	8.69E-09	59
GO:0007409	axonogenesis	5.91E-09	57
GO:0050769	positive regulation of neurogenesis	8.91E-09	57
GO:0010975	regulation of neuron projection development	6.22E-08	57
GO:0001501	skeletal system development	1.94E-05	57
GO:0015672	monovalent inorganic cation transport	3.41E-05	57
GO:0034762	regulation of transmembrane transport	2.38E-07	55
GO:0023061	signal release	5.67E-07	55
GO:0034765	regulation of ion transmembrane transport	1.87E-07	54
GO:0001525	angiogenesis	1.72E-06	54
GO:0045165	cell fate commitment	3.50E-13	52
GO:0003012	muscle system process	1.37E-06	52
GO:0010817	regulation of hormone levels	2.69E-04	51
GO:0022604	regulation of cell morphogenesis	2.69E-04	51
GO:0061564	axon development	4.42E-04	51
GO:0043410	positive regulation of MAPK cascade	4.64E-04	51
GO:0050878	regulation of body fluid levels	1.05E-03	51

signaling by wnt (P = 3.68E-14, Count = 76) and cellular calcium ion homeostasis (P = 3.50E-13, Count = 70) were the three most significant ones in the progression of CRC patients.

Enriched KEGG Pathways by DMGs

In this section, KEGG pathway enrichment analysis was performed on DMGs of CRC, and the results were listed in Table 2. There were 31 significant pathways met to the criterion P < 0.01 and Count > 20 in total. Particularly, the top three significant pathways were Calcium signaling pathway (P = 4.19E-20, Count = 67), Wnt signaling pathway (P = 5.18E-10, Count = 55) and Cell adhesion molecules (CAMs) (P = 2.22E-07, Count = 63), which might play more important roles than the others of the 31 significant pathways for CRC process.

Hub DMGs

As mentioned above, the feasibility and confidence of DMGs had been validated. Furthermore, to explore the interactions or co-operations among the 1,123 DMGs, a PPI network was constructed dependent on the STRING database. Besides, the inherent value in the STRING database for an interaction was defined as its combined score. Note that only the interactions of required confidence (combined score) greater than 0.8 were permitted to be retained in the PPI network, naming sub-network. The sub-network containing 236 nodes representing DMGs and 1,361 edges representing interactions with DMGs was illustrated in Figure 3. Subsequently, topological centrality analysis was carried out on nodes in the sub-network using the degree index, and the top 3 percent in descending order

Table 2: Significant KEGG pathways with P < 0.01 and Count > 20

ID	*Pathway*	*P value*	*Count*
PATH:hsa04020	Calcium signaling pathway	4.19E-20	67
PATH:hsa04310	Wnt signaling pathway	5.18E-10	55
PATH:hsa04514	Cell adhesion molecules (CAMs)	2.22E-07	63
PATH:hsa04024	cAMP signaling pathway	1.14E-06	31
PATH:hsa04970	Salivary secretion	1.42E-06	29
PATH:hsa04725	Cholinergic synapse	5.32E-06	24
PATH:hsa04724	Glutamatergic synapse	5.48E-06	25
PATH:hsa05032	Morphine addiction	9.84E-06	47
PATH:hsa04713	Circadian entrainment	1.81E-05	52
PATH:hsa04727	GABAergic synapse	2.43E-05	59
PATH:hsa04723	Retrograde endocannabinoid signaling	4.21E-05	60
PATH:hsa04916	Melanogenesis	1.11E-04	38
PATH:hsa04971	Gastric acid secretion	1.44E-04	44
PATH:hsa04015	4015 Rap1 signaling pathway	2.03E-04	55
PATH:hsa04730	Long-term depression	2.18E-04	21
PATH:hsa04740	Olfactory transduction	2.46E-04	37
PATH:hsa05200	Pathways in cancer	4.79E-04	38
PATH:hsa05033	Nicotine addiction	5.38E-04	48
PATH:hsa04540	Gap junction	8.97E-04	49
PATH:hsa04915	Estrogen signaling pathway	9.89E-04	49
PATH:hsa04062	Chemokine signaling pathway	1.63E-03	53
PATH:hsa04080	Neuroactive ligand-receptor interaction	2.15E-03	62
PATH:hsa04972	Pancreatic secretion	2.32E-03	58
PATH:hsa05414	Dilated cardiomyopathy (DCM)	2.98E-03	39
PATH:hsa03010	Ribosome	3.24E-03	46
PATH:hsa05016	Huntington's disease	3.45E-03	50
PATH:hsa04976	Bile secretion	3.94E-03	31
PATH:hsa04022	cGMP - PKG signaling pathway	4.51E-03	57
PATH:hsa04911	Insulin secretion	5.65E-03	42
PATH:hsa05217	Basal cell carcinoma	5.75E-03	47
PATH:hsa04060	Cytokine-cytokine receptor interaction	5.98E-03	38

of degree distribution were regarded as hub DMGs of this sub-network. A total of 7 hub DMGs were identified, CASR (Degree = 77), PLCB1 (Degree = 68), AGTR1 (Degree = 61), MCHR2 (Degree = 60), GRM4 (Degree = 59), ADCY1 (Degree = 53), and ADCY8 (Degree = 49).

DISCUSSION

DNA methylation, is a result of the covalent addition of a methyl group at the 5' position of the pyrimidine ring of cytosines within the context of CpG dinucleotides, and has a variety of important functions, including control of gene expression, cellular differentiation, genomic imprinting and X-chromosome inactivation (Hermann et al. 2004; Sibbons et al. 2018). Particularly, DNA methylation is stable and easily detected qualitatively or quantitatively. Hence it has been taken as the most promising diagnostic marker for early detection of cancer (Zhao et al. 2014). Meanwhile, If CRC was detected early before metastasis established, it would be easily cured by surgical procedures (Guo et al. 2018; Toiyama et al. 2014). Consequently, early detection of CRC is a decisive step in the successful cure of this tumor.

Therefore, in the present study, the researchers aimed to identify aberrant DNA methylations as significant signatures of CRC, and reveal epigenetic mechanism underlying CRC by multiple bioinformatics analyses. Particularly, using the DNA methylation data, DMGs between CRC samples and normal controls were identified using *t*-test. And the DMGs were confirmed with a good classification performance between CRC and controls, which enhanced the feasibility of this study. Results of the functional enrichment analyses showed that 37 significant GO terms and 31 significant pathways were detected. Moreover, a total of 7 hub DMGs were investigated by topological centrality analysis for subnetwork of PPI network. On the basis of these results, the potential mechanisms of CRC were

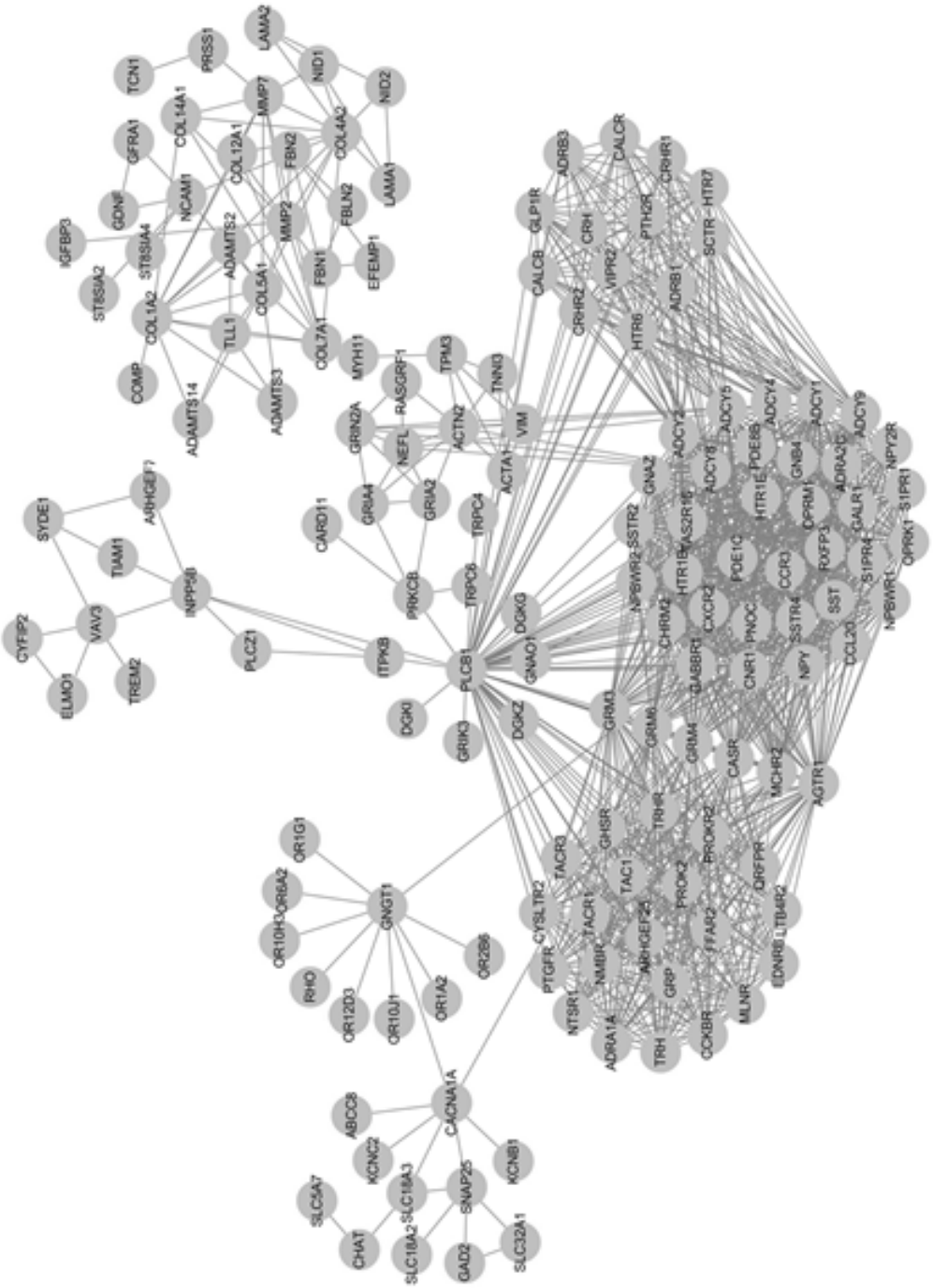

Fig. 3. Sub-network of protein-protein interaction (PPI) network constructed on differentially methylated genes (DMGs). Nodes referred to DMGs, while edges represented the interactions between any two of DMGs with combined score > 0.8
Source: Author

revealed, which shed new insights into CRC diagnosis and therapy.

A recent large-scale comparison between genes mutated and hypermethylated in CRC revealed significant overlap between these two alterations (Boot et al. 2017; Chan et al. 2008). Over-expression of immediate-early response genes significantly correlate to downstream protein synthesis and small-molecule metabolite production, and further take participated in the colorectal carcinogenesis (Qiu et al. 2016). Coincidently, in the present work, we found that two significant pathways, cAMP signaling pathway and Cell adhesion molecules (CAMs), were closely related to small-molecule metabolite processes. Particularly, the most significant pathway enriched by DMGs was calcium signaling pathway. In addition, results of GO enrichment analysis also showed that calcium ion homeostasis and cellular calcium ion homeostasis were significant between CRC patients and controls. Hence, we might infer that calcium is crucial for CRC, which has been confirmed by previous literatures (Keum et al. 2014; Z et al. 2011). Increasing intakes of total calcium, dietary calcium, calcium-containing supplement may reduce the risk of CRC (Z et al. 2011). Most important, significant GO term cell-cell signaling by wnt and significant pathway wnt signaling pathway both uncovered that the wnt signaling acted as critical roles in CRC patients, which had been confirmed by Bienz et al. (Bienz and Clevers 2000). Above all, significant GO terms and pathways enriched by DMG are related to CRC tightly, which validated the feasibility and enhanced the confidence of these DMGs of CRC.

Significantly, 7 hub DMGs were identified for CRC, which was comprised of CASR, PLCB1, AGTR1, MCHR2, GRM4, ADCY1 and ADCY8. In details, CASR was the one with the highest degree distribution. CASR (calcium sensing receptor) is a G protein-coupled receptor, expressed by tissues involved in the control of Ca^{2+} homeostasis. It could act as a disulfide-linked dimmer to couple to multiple signaling pathways (Hendy et al. 2015). It had been reported that higher blood 25-hydroxyvitamin D levels decreased the risk of developing CRC (Fedirko et al. 2012). Coincidently, CASR is one of vitamin D related genes. On the other hand, functions of CASR in the progression of CRC might be relate to the calcium content. Antiproliferative effects of Ca^{2+} are impaired or lost in CRC, and this could be due to reduced CASR expression in the tumors (Rogers et al. 2012; Sukawa et al. 2011). Hence, this hub DMG played significant roles in CRC patients.

CONCLUSION

The present study provided a comprehensive bioinformatics analysis of DMGs which might be involved in the development and progression of CRC. Further, significant gene sets enriched by DMGs and hub DMGs might be potential biomarkers for treatment of CRC and provide a clue for understanding the potential pathogenesis of CRC. However, the validations of these hub DMGs should be performed in CRC cell lines as soon as possible.

RECOMMENDATIONS

All investigations obtained from the current study will provide potential targets for CRC treatment, and shed new lights on revealing molecular mechanism underlying this tumor.

ACKNOWLEDGEMENTS

This work was supported by Shanghai Pudong New Area Zhoupu Hospital college key discipline construction projects (No.ZP-XK-2015B-5) and Shanghai Pudong New Area of traditional Chinese medicine clinical Subject construction of clinical demonstration projects (No. PDZYXK-3-2013008).

REFERENCES

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT et al. 2000. Gene Ontology: Tool for the unification of biology. *Nat Genet*, 25: 25-29.

Baylin SB, Jones PA 2016. Epigenetic determinants of cancer. *Cold Spring Harbor Perspectives in Biology*, 6.

Bienz M, Clevers H 2000. Linking colorectal cancer to Wnt signaling. *Cell*, 103: 311.

Boot A, Oosting J, van Eendenburg JDH, Kuppen PJK, Morreau H, van Wezel T 2017. Methylation associated transcriptional repression of ELOVL5 in novel colorectal cancer cell lines. *PLoS One*, 12: e0184900.

Brenner H 2014. Colorectal cancer. *Lancet*, 383: 1490-1502.

Chan TA, Glockner S, Yi JM, Chen W, Van NL, Cope L, Herman JG, Velculescu V, Schuebel KE, Ahuja N 2008. Convergence of mutation and epigenetic alterations identifies common genes in cancer that predict for poor prognosis. *Plos Medicine*, 5: e114.

De Hoon Me, Eisen Michael 2002 Cluster 3.0 Manual. University of Tokyo. From <http://bonsai.ims.u-tokyo. ac.jp>.

Fedirko V, Riboli E, Tjønneland A, Ferrari P, Olsen A, Buenodemesquita HB, van Duijnhoven FJ, Norat T, Jansen EH, Dahm CC 2012. Pre-diagnostic 25-Hydroxyvitamin D, VDR and CASR Polymorphisms, and survival in patients with colorectal cancer in Western European populations. *Cancer Epidemiology Biomarkers & Prevention*, 21: 582.

Flores K, Wolschin F, Corneveaux JJ, Allen AN, Huentelman MJ, Amdam GV 2012. Genome-wide association between DNA methylation and alternative splicing in an invertebrate. *BMC Genomics*, 13: 1-9.

Geiman TM, Robertson KD 2002. Chromatin remodeling, histone modifications, and DNA methylation—how does it all fit together? *Journal of Cellular Biochemistry*, 87: 117-125.

Giovannucci E 2003. Diet, body weight, and colorectal cancer: A summary of the epidemiologic evidence. *Journal of Womens Health*, 12: 173-182.

Goel A, Xicola RM, Nguyen TP, Doyle BJ, Sohn VR, Bandipalliam P, Rozek LS, Reyes J, Cordero C, Balaguer F 2010. Aberrant DNA methylation in hereditary nonpolyposis colorectal cancer without mismatch repair deficiency. *Gastroenterology*, 138: 1854-1862.e1851.

Guo Y, Wu R, Gaspar JM, Sargsyan D, Su ZY, Zhang C, Gao L, Cheng D, Li W, Wang C et al. 2018. DNA methylome and transcriptome alterations and cancer prevention by curcumin in colitis-accelerated colon cancer in mice. *Carcinogenesis*, 10.1093/carcin/bgy043. [Epub ahead of print].

He Y, Cui Y, Wang W, Gu J, Guo S, Ma K, Luo X 2011. Hypomethylation of the hsa-miR-191 locus causes high expression of hsa-mir-191 and promotes the epithelial-to-mesenchymal transition in hepatocellular carcinoma. *Neoplasia*, 13: 841-853.

Hendy GN, D'Souza-Li L, Yang B, Canaff L, Cole DE 2015. Mutations of the calcium-sensing receptor CASR in familial hypocalciuric hypercalcemia, neonatal severe hyperparathyroidism, and autosomal dominant hypocalcemia. *Human Mutation*, 82: 281-296.

Hermann A, Gowher H, Jeltsch A 2004. Biochemistry and biology of mammalian DNA methyltransferases. *Cellular & Molecular Life Sciences Cmls*, 61: 2571.

Hinoue T, Weisenberger DJ, Lange CP, Shen H, Byun HM, Van DBD, Malik S, Pan F, Noushmehr H, van Dijk CM 2012 Genome-scale analysis of aberrant DNA methylation in colorectal cancer. *Genome Research*, 22: 271-282.

Houlihan PS 2002. CpG island methylation in aberrant crypt foci of the colorectum. *American Journal of Pathology*, 160: 1823-1830.

Huang da W, Sherman BT, Lempicki RA 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*, 4: 44-57.

Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D 2011. Global cancer statistics. *CA Cancer J Clin*, 61: 69-90.

Jiang X, Tan J, Li J, Kivimäe S, Yang X, Zhuang L, Lee PL, Chan MT, Stanton L, Liu ET 2008. DACT3 is an epigenetic regulator of Wnt/β-Catenin Signaling in colorectal cancer and is a therapeutic target of histone modifications. *Cancer Cell*, 13: 529-541.

Kanehisa M, Goto S 2000. KEGG: Kyoto Encyclopaedia of Genes and Genomes. *Nucleic Acids Research*, 28: 27-3024.

Keum N, Aune D, Greenwood DC, Ju W, Giovannucci EL 2014. Calcium intake and colorectal cancer risk: Dose-response meta-analysis of prospective observational studies. *International Journal of Cancer Journal International Du Cancer*, 135: 1940-1948.

Mais L, Sarrabi M, Desseigne F, Guibert P, Castillo C, Foulfoin M, Karabajakian A, Christelle DLF 2015. Retrospective study of BRAF-mutant metastatic colorectal cancer patients and influence of the mismatch repair efficiency. *Journal of Neurosurgery*, 97: 794-802.

Migliore L, Migheli F 2011. Genetics, cytogenetics, and epigenetics of colorectal cancer. *Journal of Biomedicine & Biotechnology*, 2011: E1172-E1179

Mohammadi A, Saraee MH, Salehi M 2011. Identification of disease-causing genes using microarray data mining and gene ontology. *BMC Medical Genomics*, 4: 12.

Pita-Juarez Y, Altschuler G 2018. The Pathway Coexpression Network: Revealing pathway relationships. *PLoS Computational Biology*, 14(3):e1006042.

Qiu Y, Cai G, Su M, Chen T, Zheng X, Xu Y, Ni Y, Zhao A, Xu LX, Cai S 2016. Serum metabolite profiling of human colorectal cancer using GC"TOFMS and UPLC"QTOFMS. *Journal of Proteome Research*, 8: 4844.

Rogers AC, Hanly AM, Collins D, Baird AW, Winter DC 2012. Review article: Loss of the calcium-sensing receptor in colonic epithelium is a key event in the pathogenesis of colon cancer. *Clinical Colorectal Cancer*, 11: 24-30.

Sibbons CM, Irvine NA, Perez-Mojica JE, Calder PC, Lillycrop KA, Fielding BA, Burdge GC 2018. Polyunsaturated fatty acid biosynthesis involving Delta8 desaturation and differential DNA Methylation of FADS2 Regulates proliferation of human peripheral blood mononuclear cells. *Front Immunol*, 9: 432.

Smoot ME, Ono K, Ruscheinski J, Wang P-L, Ideker T 2011 Cytoscape 2.8: New features for data integration and network visualization. *Bioinformatics*, 27: 431-432

Sukawa Y, Yamamoto H, Nosho K, Miyamoto C, Miyamoto N, Taniguchi H, Adachi Y, Nakazawa M, Suzuki H, Sasaki S 2011. Epigenetic inactivation of calcium-sensing receptor in colorectal carcinogenesis. *Modern Pathology An Official Journal of the United States & Canadian Academy of Pathology Inc*, 24: 876-884.

Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP 2015. STRING v10: Protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Research*, 43: 447-452.

Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, Beck S 2013. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics*, 29: 189-196.

Toiyama Y, Okugawa Y, Goel A 2014. DNA methylation and microRNA biomarkers for noninvasive detection of gastric and colorectal cancer. *Biochemical & Biophysical Research Communications,* 455: 43-57.

Walsh JM, Terdiman JP 2003. Colorectal cancer screening: Scientific review. *Jama the Journal of the American Medical Association,* 289: 1288-1296.

Wang CM, Yang XL, Liu MH, Cheng BH, Chen J, Bai B 2018. High-throughput sequencing analysis of differentially expressed miRNAs and target genes in ischemia/reperfusion injury and apelin-13 neuroprotection. *Neural Regen Res,* 13: 265-271.

Wang X, Simon R 2011. Microarray-based cancer prediction using single genes. *BMC Bioinformatics,* 12: 391.

Wasserman S 1994. *Social Network Analysis: Methods And Applications.* Volume 8. New York and. Cambridge, ENG: Cambridge University Press.

Wood LD, Parsons DW, Jones S, Lin J, Sjöblom T, Leary RJ, Shen D, Boca SM, Barber T, Ptak J 2007. The genomic landscapes of human breast and colorectal cancers. *Science,* 318: 1108.

Wu DP, Farrell PW, Haworth DK, Emes PR, Kitchen DM, Glossop DJ, Hanna DF, Fryer AA 2016. Maternal genome-wide DNA methylation profiling in gestational diabetes shows distinctive disease-associated changes relative to matched healthy pregnancies. *Epigenetics,* 13: 122-128.

Yagi K, Akagi K, Hayashi H, Nagae G, Tsuji S, Isagawa T, Midorikawa Y, Nishimura Y, Sakamoto H, Seto Y 2010. Three DNA methylation epigenotypes in human colorectal cancer. *Clinical Cancer Research,* 16: 21.

Yan W, Guo M 2015. Epigenetics of colorectal cancer. *Gastroenterology,* 1238: 405-424.

Z S, PP W, B R, M C, R G, S B, J Z, J S, Y Z 2011. Calcium and vitamin D and risk of colorectal cancer: Results from a large population-based case-control study in Newfoundland and Labrador and Ontario. *Can J Public Health,* 102: 382-389.

Zhao Y, Sun J, Zhang H, Guo S, Gu J, Wang W, Tang N, Zhou X, Yu J 2014. High-frequency aberrantly methylated targets in pancreatic adenocarcinoma identified via global DNA methylation analysis using methylCap-seq. *Clinical Epigenetics,* 6: 1-14.

Zhou Q, Hong L, Wang J 2018. Identification of key genes and pathways in pelvic organ prolapse based on gene expression profiling by bioinformatics analysis. *Arch Gynecol Obstet,* 297(5):1 323-1332.

Paper received for publication on March 2018
Paper accepted for publication on May 2018