

# Layered Signaling Regulatory Networks Analysis of Gene Expression Involved in Malignant Tumorigenesis of Non-Resolving Ulcerative Colitis via Integration of Cross-Study Microarray Profiles

Shengjun Fan<sup>1</sup>, Zhenyu Pan<sup>2</sup>, Qiang Geng<sup>3</sup>, Xin Li<sup>1</sup>, Yefan Wang<sup>1</sup>, Yu An<sup>1</sup>, Yan Xu<sup>4</sup>, Lu Tie<sup>1</sup>, Yan Pan<sup>1</sup>, Xuejun Li<sup>1\*</sup>

**1** State Key Laboratory of Natural and Biomimetic Drugs, Department of Pharmacology, School of Basic Medical Sciences, Peking University and Beijing Key Laboratory of Tumor Systems Biology, Peking University, Beijing, China, **2** Department of Pharmacy, Xi'an Children's Hospital, Xi'an, China, **3** Department of Cardiology, Peking University People's Hospital, Beijing, China, **4** Pediatric Research Institute, Qilu Children's Hospital of Shandong University, Jinan, China

## Abstract

**Background:** Ulcerative colitis (UC) was the most frequently diagnosed inflammatory bowel disease (IBD) and closely linked to colorectal carcinogenesis. By far, the underlying mechanisms associated with the disease are still unclear. With the increasing accumulation of microarray gene expression profiles, it is profitable to gain a systematic perspective based on gene regulatory networks to better elucidate the roles of genes associated with disorders. However, a major challenge for microarray data analysis is the integration of multiple-studies generated by different groups.

**Methodology/Principal Findings:** In this study, firstly, we modeled a signaling regulatory network associated with colorectal cancer (CRC) initiation via integration of cross-study microarray expression data sets using Empirical Bayes (EB) algorithm. Secondly, a manually curated human cancer signaling map was established via comprehensive retrieval of the publicly available repositories. Finally, the co-differently-expressed genes were manually curated to portray the layered signaling regulatory networks.

**Results:** Overall, the remodeled signaling regulatory networks were separated into four major layers including extracellular, membrane, cytoplasm and nucleus, which led to the identification of five core biological processes and four signaling pathways associated with colorectal carcinogenesis. As a result, our biological interpretation highlighted the importance of EGF/EGFR signaling pathway, EPO signaling pathway, T cell signal transduction and members of the BCR signaling pathway, which were responsible for the malignant transition of CRC from the benign UC to the aggressive one.

**Conclusions:** The present study illustrated a standardized normalization approach for cross-study microarray expression data sets. Our model for signaling networks construction was based on the experimentally-supported interaction and microarray co-expression modeling. Pathway-based signaling regulatory networks analysis sketched a directive insight into colorectal carcinogenesis, which was of significant importance to monitor disease progression and improve therapeutic interventions.

**Citation:** Fan S, Pan Z, Geng Q, Li X, Wang Y, et al. (2013) Layered Signaling Regulatory Networks Analysis of Gene Expression Involved in Malignant Tumorigenesis of Non-Resolving Ulcerative Colitis via Integration of Cross-Study Microarray Profiles. PLoS ONE 8(6): e67142. doi:10.1371/journal.pone.0067142

**Editor:** Frank Emmert-Streib, Queen's University Belfast, United Kingdom

**Received:** January 22, 2013; **Accepted:** May 15, 2013; **Published:** June 25, 2013

**Copyright:** © 2013 Fan et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by the National Natural Science Foundation of China (No. 91129727, 81020108031, 30973558, 81270049) and Research Fund from Ministry of Education of China (111 Projects No.B07001). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that they have no competing interests.

\* E-mail: xjli@bjmu.edu.cn

## Introduction

As the fourth commonest carcinoma, colorectal cancer (CRC) associated with significant cause of mortality worldwide owing to its prevailing distant metastasis [1,2]. Unfortunately, there are still more than 783,000 new cases diagnosed and roughly 394,000 deaths yearly [3]. Further, it is conservatively estimated the lethality will continue to rise for the increased life-expectancy and aging population [4,5]. Epidemiological studies uncovered individuals consistently exposed to inadequate physical practices or

high-fat dietary closely interrelated with high risk of colorectal neoplasia [6]. Besides, environmental and heritable factors also made significant contributions to CRC susceptibility [7]. Traditional pathological examination have identified several causative modifiers, including *TP53*, *K-ras*, *APC*, *Wnt5*, *beta-catenin*, *DCC* or microRNAs [8,9]. Among them, *APC* and *DCC* were the most frequently detectable prognostic signatures for CRC; however, the results were perplexed and cardinal hurdles for clinical therapeutic interventions were still insurmountable. Since most of variant

genes have ineffectual profits for diagnosis and underlying mechanisms associated with CRC are ill-defined.

Considerable documents certified countless malignancies occurred in association with chronic inflammation. Infection with hepatitis B or C viruses had been found to be the main cause of hepatocellular carcinoma [10,11]. Besides, inflammation also directly related to DNA methylation and epithelial cell malignant transformation [12,13]. As an inflammatory response to infection, inflammation-mediated CRC had been reviewed [14]. Numerable evidences pointed out chronic ulcerative colitis (UC) was intimately connected with colorectal carcinogenesis [15,16,17,18]. However, our current knowledge regarding signaling regulatory networks between UC and CRC has not been unraveled yet.

Cells in multi-cellular organisms switch into diverse fates, such as division, proliferation, apoptosis or differentiation into specialized phenotypes. Genome-wide association studies (GWAS) of gene regulatory networks (GRNs) govern this process and the inference of GRNs is crucial for understanding underlying molecular mechanisms between genes and gene regulation [19,20]. Typically, GRNs are modeled as a structure of genes, *cis*-elements, and regulators. The regulators, also called transcription factors, are often described as proteins which bound to specific regions of target gene, and thereby regulated the transcription of genetic information from DNA to mRNA [21]. They functioned as an activator or repressor alone or with other proteins to regulate the recruitment of RNA polymerase to targets [22,23]. Moreover, apart from proteins, microRNAs also participate in the transcriptional and post-transcriptional regulatory manner of gene expression in plants and animals [24]. In addition, interactions among regulators are defined as *cis*-elements in GRNs and thereby control the level of gene expression during transcription.

As yet, a wide variety of approaches have been proposed for modeling GRNs, such as discrete models of Boolean networks [25,26], Bayesian networks [27,28], system of equations [29], continuous models of neural networks [30], association network [31], and co-expression models [20], etc. Among them, the most successful approaches for GRNs inference were ascribed to the reconstruction of the association network from target predictions [32]. However, the so-called 'guilt-by-association' methods [33,34] need long time consumption, making them unsuitable for GRNs inference in term of large-scale genomes [35]. Meanwhile, owing to its high risk of false-positive noise, cardinal drawback for the association network modeling was its weak correlations, which was insufficient to elaborate the real connections within the network. Needless to say, with the rapid increase of experimental approaches, the advent of microarray facilitated large-scale monitoring of gene expression under different conditions at one time and enlightened roles of genes associated with infirmities in a systematic visualization [36,37]. Thus, inference of gene regulations in GRNs based on microarray gene co-expression models provided an opportunity for understanding the underlying regulatory mechanism [38]. Meanwhile, knowledge associated with subcellular localization in a network is indispensable for grasping molecular function and intricate biological pathway at subcellular level [39]. Thus, modeling of layered GRNs by integrating large-scale microarray expression data sets contributed to elucidate illnesses in a systems biology perspective.

However, mining these data to better comprehend gene expression and regulation proposes a major challenge for bioinformatics. Since practical considerations restrain the size of samples and the overlap genes in multiple studies are limited with poor predictability [40]. If small sizes of individual studies from

different experiments are combined to increase sample size, the integrative manner is therefore a promising approach and a more accurate reconstruction of GRNs can be expected [41]. Nevertheless, real biological variation from the data set still exists [42] when different samples are added to an existing one or in a meta-analysis of multiple studies that pools microarray data across different laboratories or platforms [43]. To make up the scarcity, we perform a simulation by randomly generated 8 virtual data sets, also called random microarray data sets (RDSs), from the integrative microarray cohort.

The purpose of functional genomics in the post-genome era is to better elucidate molecular mechanisms involved in gene regulation [44]. Genes assigned to certain gene sets by clustering analysis belong to different regulatory modules or signaling pathways. GRNs clarify the interaction among transcription factors and inference of transcriptional regulatory networks assists in understanding underlying mechanism of complex cellular processes or responses [45]. Given that gene-gene interactions contribute to complex diseases, the combination of multiple variants based on biological pathways tends to uncover the synergistic effects of large-scale genes and highlight the specific signaling pathways involved in diseases [46].

Traditionally, cell signaling is described as linear diagrams. As more cross-talk between signaling has been reviewed, a systematic network view of cell signaling is proposed [47]. Among them, one of the pioneering attempts to model signaling network was assigned to the map of human cancer signaling. In Cui *et al.* network, a comprehensive analysis of human cancer signaling architectural organization assembled from cancer-associated genetically and epigenetically altered genes was established [48]. In addition, Fan and colleagues also performed a network-based pathway analysis using gene co-expression models to specify the off-target effects for torcetrapib. They highlighted that IL-2 Receptor Beta Chain in T cell Activation, Platelet-Derived Growth Factor Receptor (PDGFR) beta signaling pathway, IL-2-mediated signaling events, ErbB signaling pathway and signaling events mediated by Hepatocyte Growth Factor Receptor (HGFR, c-Met) were answered for the adverse cardiovascular effects associated with torcetrapib [49]. Thus, pathway-based signaling regulatory networks had been widely applied to complex diseases, which led to identifying diseases-susceptibility pathways for therapeutic interventions [50,51,52].

In this study, we modeled a layered signaling regulatory network associated with colorectal cancer from cross-study microarray gene expression data using the experimentally-supported interaction and microarray co-expression modeling. Cytoscape [53] in association with four plugins including BisoGenet [54], NetworkAnalyzer, AllegroMCODE and Cerebral [55] was applied for inferring the layered signaling regulatory network. To our knowledge, by far, there were no published documents expounded the pathological transition focusing on a layered signaling network and our study provided infrequent insights into the potential molecular mechanisms, which might be useful for colorectal cancer therapeutic prevention or intervention.

## Materials and Methods

### Microarray Data Sets Selection

We search the public functional genomics data repositories including ArrayExpress (<http://www.ebi.ac.uk/arrayexpress/>), Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo/>) and Stanford Microarray Database (SMD, <http://smd.stanford.edu/>) for microarray data sets satisfying the following criteria: (1) deriving from *Homo sapiens*; (2) depicting genome-

**Table 1.** Microarray data sets utilized in this experiment.

Authors	Documents (PMID)	GEO number	Array type	Number of samples
Gyorffy <i>et al.</i>	20087348	GSE4183	Affymetrix Human Genome U133 Plus 2.0 Array	30(15:15; CRC:15)
Pekow <i>et al.</i>	23388545	GSE37283	Affymetrix HT HG-U133+ PM Array	15(UC:4;CRC:11)

Abbreviations: IBD: inflammatory bowel disease, CRC: colorectal carcinoma, UC: ulcerative colitis.

doi:10.1371/journal.pone.0067142.t001

wide co-expression information of ulcerative colitis and neoplastic lesions; (3) supplying raw data files. After extensively retrieval, two different microarray expression cohorts performed on Affymetrix GeneChip platform were downloaded from GEO and selected for further investigation (Table 1).

The first data set obtained from Gyorffy *et al.* (GEO ID: GSE4183) was performed on Affymetrix Human Genome U133 Plus 2.0 Array platform [56]. This study derived from different stages of pathophysiological background of colonic diseases with 15 samples for inflammatory bowel disease (IBD) and 15 samples for colorectal carcinoma (CRC) with the purpose of developing a comprehensive comparison of preprocessing algorithms on samples to afford a data warehouse which can be further mined for in-depth pathway analyses.

The other one is a subset of 15 Affymetrix HT HG-U133 GeneChips originated from a gene expression experiment by Pekow *et al.* [57]. The aim of this original investigation was to perform a genome-wide expression profiling between chronic ulcerative colitis (UC, 4 arrays) and neoplastic lesions (11 arrays) to ultimately detect underlying prognostic gene signatures involved in this pathophysiological transition (GEO accession number GSE37283). All clinical specimens utilized in this study were directly taken from patient-sufferers either surgically disconnected or during surveillance colonoscopy.

### Computational Pipeline

To standardize the microarray data sets obtained from independent studies and reduce systematic distortions produced by different laboratories, rank normalization was introduced to effectuate robust multi-array average (RMA) analysis [58]. Subsequently, platform comparison from Affymetrix probe identifier to NCBI AILUN [59] was performed and cross-study normalization was achieved using ArrayMining (<http://www.arraymining.net/>), which pooled multiple microarray data from independent investigations or platforms into a combinational cohort [60].

### Simulated Gene Expression Data Analysis

We simulated 8 virtual cohorts, called the random microarray data sets (RDSs), from the integrated data set mentioned above [42]. For each RDSs, we created at least 12 samples. Afterwards, all the RDSs were performed Significance Analysis of Microarray (SAM, <http://www-stat.stanford.edu/tibs/SAM/>) to produce a cluster of up- or down-regulated variant genes via comparison of UC with CRC [61]. Gene expression was regarded as significantly dissimilar if the threshold of false discovery rate (FDR) less than 0.05 and fold change above 1.2.

### Cancer Signaling Regulatory Map Construction

To provide a high-quality human cancer signaling atlas with great superiority, the manually curated molecules of human cancer signaling were assembled as previously described [48]. Briefly, a map of human cancer signaling was firstly obtained from

Cui *et al.*, which united manually curated signaling molecules including BioCarta (<http://www.biocarta.com/>) [62], literature-mined signaling network [63], Cancer Cell Map including cancer mutated genes obtained from COSMIC database and other high-throughput profiling, methylated genes in cancer stem cells, cancer associated gene set acquired from plasmID (<http://plasmid.hms.harvard.edu/>) and Online Mendelian Inheritance in Man (OMIM, <http://www.ncbi.nlm.nih.gov/omim>). Subsequently, Human Protein Reference Database (HPRD, <http://www.hprd.org/>) [49] and REACTOME (<http://www.reactome.org/>) databases were appended. After adding connections between signaling molecules based on SysBiomics platform (<http://biomine.cigb.edu.cn/sysbiomics/>) [54], the original human cancer signaling map was visualized by Cytoscape (<http://www.cytoscape.org/>) [53]. Dispersive nodes without interactions and small sub-networks were discarded afterwards. Only the largest component was considered as the map of human cancer signaling. Meanwhile, duplicated edges and interaction direction were abandoned from the network utilizing NetworkAnalyzer plugin before analysis. Ultimately, the reconstructed human cancer signaling network was clustered by AllegroMCODE using Molecular Complex Detection (MCODE) clustering algorithm [64].

### Layered Signaling Regulatory Network Analysis

Subcellular localization information involved in the signaling regulatory networks driven by the co-differently-expressed genes was diffusely retrieved from Human Protein Reference Database (HPRD, <http://www.hprd.org/>), Human Proteinpedia (<http://www.humanproteinpedia.org/>) and EntrezGene (<http://www.ncbi.nlm.nih.gov/gene>). Cerebral generated a view of the layered map afterwards and each of the signaling regulatory networks was divided into four layers, including extracellular, membrane, cytoplasm and nucleus [55]. Nodes without localization information were positioned into the same layer of their adjacent neighbors.

### Gene Ontology (GO) and Pathway Analysis

Functional enrichment analysis was performed to evaluate underlying biological functions involved in the layered signaling regulatory networks. In this study, the DAVID functional annotation clustering tool (<http://david.abcc.ncifcrf.gov/>) was freely employed to enrich the potential biological functions based on "GOTERM\_BP\_FAT" option [65,66]. For pathway enrichment analysis, ToppCluster (<http://toppcluster.cchmc.org/>) was selected as gene sets category [67].

### Hierarchical Signaling Regulatory Networks Organization

To obtain an intuitionistic two-dimensional graphical representation of large-scale cluster members across an activity profile and define the main clusters by visual inspection of the resulting tree, R (<http://www.r-project.org/>) cooperated with stats package was introduced to hierarchically trace heatmap for the over-represent-

ed GO biological processes and signaling pathways in the layered signaling regulatory network.

## Results

Two independent microarray data sets depicting UC and CRC from clinical specimens were utilized in this experiment. RMA analysis was applied to implement rank normalization based on WebArray [58]. Platform comparison from Affymetrix probe identifier to NCBI was performed using AILUN [59] before cross-study normalization and integration [60]. Totally, there were 20,083 common genes, 455 unique genes in GPL570, and 433 diverse genes in GPL13158. We then performed a simulation by randomly generating 8 RDSs from the combinational cohort and carried out SAM to produce a cluster of variances.

To acquire high-quality human cancer signaling, we manually curated cancer signaling molecules thoroughly [48,49]. All the over-represented genes obtained from the virtual RDSs were mapped to portray the context-specific sub-communities and the remodeled signaling regulatory networks were redistributed according to their subcellular localization. To better expound the underlying molecular mechanisms associated with CRC initiation in the layered signaling networks, the DAVID functional annotation tool and ToppCluster were freely employed to implement Gene Ontology (GO) and pathway enrichment analysis, respectively. In general, the remodeled signaling regulatory networks mainly responded to negative regulation of DNA recombination, nucleotide-excision repair, modification-dependent protein catabolic process, protein catabolic process and modification-dependent macromolecule catabolic process. Pathway enrichment analysis indicated that EGF/EGFR signaling pathway, EPO signaling pathway, T cell signal transduction and members of the BCR signaling pathway were primarily responsible for the malignant transition of CRC.

### Identification of Co-differently-represented Genes Associated with Colorectal Cancer

To identify variant regulators between UC and CRC, two separate microarray data sets with primary intensity data files of Affymetrix CEL format were submitted to WebArray to carry out RMA analysis (Appendix S1 and S2). The output graphic plots including histogram, M-A plot and M-B plot for each study before and after within-array normalization were presented in Figure 1. We then effectuated cross-study normalization and integration using ArrayMining, which averaged across individual probe expression values and the integrative data set were attached in Appendix S3 and S4. Figure 2 pictured the density plot and Q-Q plot before and after integration based on Empirical Bayes (EB) algorithm, which focused on pooling independent microarray data sets utilizing parametric and non-parametric EB approach to filter out batch effects [68]. Finally, 8 RDSs originated from the same combinational cohort were performed SAM to produce a list of differently expressed genes for false discovery rate (FDR)<0.05 and the fold change>1.2 (Appendix S5).

### Human Cancer Signaling Network Construction

To infer a map of human cancer signaling, varieties of repositories were exhaustively retrieved. Functional relations between nodes were added using BisoGenet in SysBiomics. In total, the reconstructed human cancer signaling network consisted of 11,728 nodes and 94,471 connections, and the raw data was appended in Appendix S6.

## Signaling Regulatory Networks Organization

We next sought to pursue the decisive signaling regulatory cohort associated with the aggressive transition from UC to CRC in the map of human cancer signaling using MCODE clustering algorithm [64]. As presented in Figure 3, 8 network modules with a cluster score above 2.0 were detected. Subsequently, all the differently expressed regulators were mapped to portray the context-specific signaling sub-communities. As shown in Table 2, out of the 8 signaling regulatory modules, 3 gene sets originating from cluster 3, 6 and 7 were principally driven by most of the variance and answered for colorectal carcinoma initiation.

### Layered Signaling Regulatory Networks Construction

Subcellular localization information involved in the context-specific sub-communities encoded by the virtual RDSs was retrieved from HPRD, Human Proteinpedia or EntrezGene, and imported as node attributes. Afterwards, Cerebral automatically generated a view of the layered signaling regulatory network which integrated 3 signaling regulatory sub-networks according to the subcellular localization, and the remodeled signaling regulatory networks were divided into four layers, including extracellular, membrane, cytoplasm and nucleus (Figure 4). Besides, nodes in the layered network were proportionate to degree and empowered polychrome schemes to distinguish these modules. Red and blue represented nodes located in cluster 3 and 6, respectively. Whereas, green stood for hubs scattered in cluster 7.

### GO Analysis

The layered signaling network associated with CRC in the context of GO was assessed by DAVID to identify significantly over-represented biological functions (FDR<0.01). Meanwhile, a heatmap with hierarchically clustering was produced to visualize these processes based on stats package in R project. As shown in Figure 5 A, most of the biological processes related to molecular metabolic, modification, biosynthetic, transcription and catabolic processes. Particularly, we underlined the importance of regulators in nucleus layer, which indicated that negative regulation of DNA recombination, nucleotide-excision repair, modification-dependent protein catabolic process, protein catabolic process and modification-dependent macromolecule catabolic process were responsible for the pathological transition from UC to colorectal carcinogenesis.

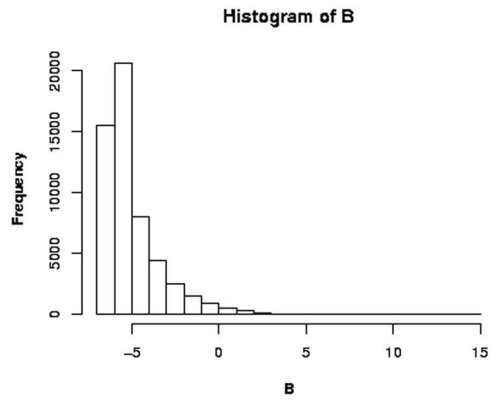
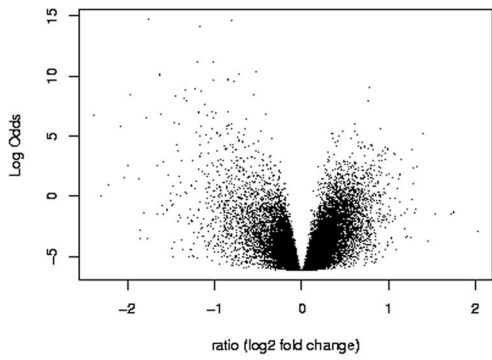
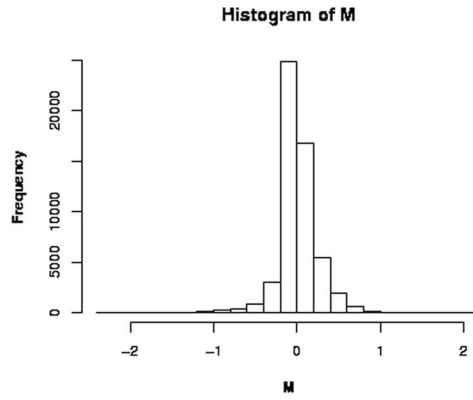
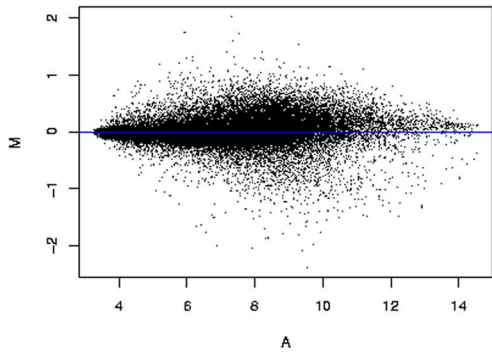
### Pathway Analysis

To gain a detailed insight into the functions of the whole regulators in the layered signaling regulatory network, we additionally performed pathway enrichment analysis using ToppCluster. As appended in Figure 5 B, genes distributed in extracellular, membrane and nucleus regions of signaling network 7 were highly associated with CRC initiation especially for EGF/EGFR signaling pathway, EPO signaling pathway, T cell signal transduction and members of the BCR signaling pathway (FDR<0.01).

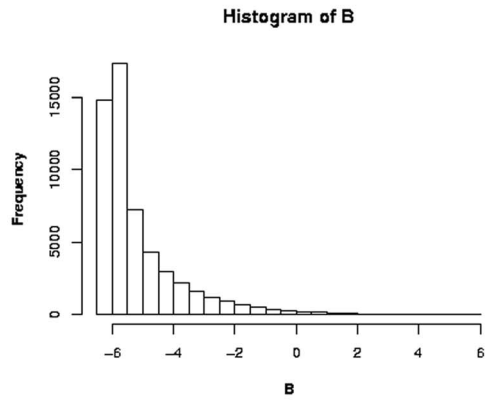
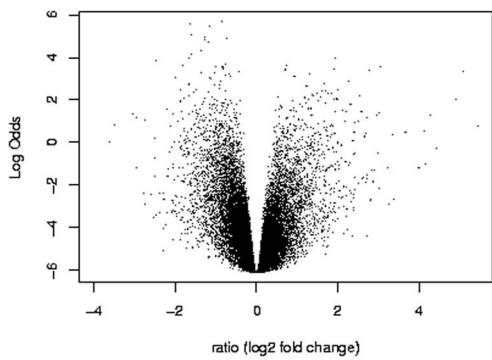
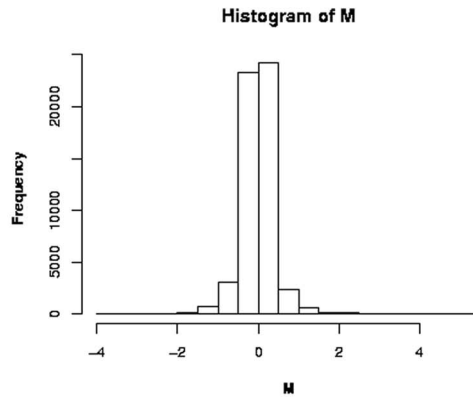
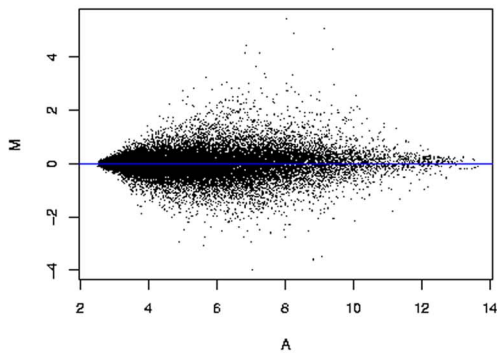
## Discussion

As microarray data on gene expression programs become available, it is profitable to create a systematic view of biological systems to improve our understanding of underlying mechanisms associated with disorders [69,70]. Given that extensive investigations of signaling had been studied over the past few decades, knowledge concerning signaling regulation had been assembled and deposited in publicly available resources. Since the abnormal expression of genes involved in diseases frequently resulted in

A



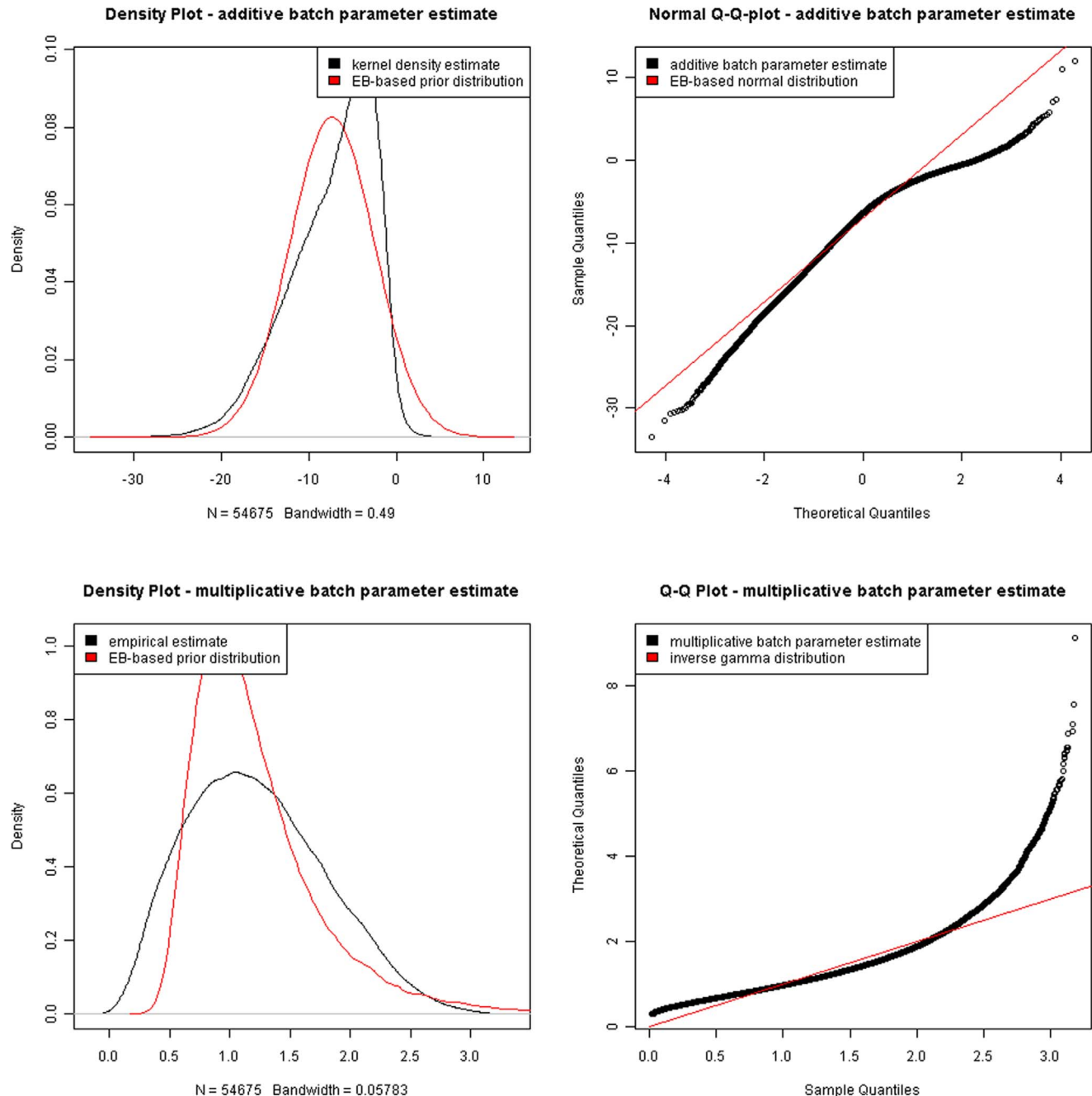
B



**Figure 1. Robust multi-array average (RMA) analysis results of microarray data based on WebArray.** (A) Statistical analysis result plot for GSE4183 included M-A plot, M-B plot, M histogram and B statistics histogram. (B) Statistical analysis result plot for GSE37283 included M-A plot, M-B plot, M histogram and B statistics histogram. M: the log-differential expression ratio; A: the log-intensity of spot, a measure of overall brightness of spot; B: B statistics, the log-odds of differential expression.  
doi:10.1371/journal.pone.0067142.g001

genome instability, analysis of gene-gene interactions in the context of gene regulatory networks could reveal the specific GRNs that led to the dysfunction of biological systems [71]. Thus, an integrative analysis of human cancer signaling network based on cross-study microarray expression profiles was of great availability in CRC initiation and progression [48].

Despite massive amounts of experimentally-supported microarray expression data bestowed in public repositories, it was still strenuous to avail legitimately. Since experiments were run on different laboratories, the combined usage of multiple platforms was necessary to overcome technical variation of individual study that resulted from sample preparation, labeling, hybridization,



**Figure 2. Cross-study normalization and integration results of two separate microarrays based upon ArrayMining.**  
doi:10.1371/journal.pone.0067142.g002

**Table 2.** Simulated results for the random microarray data sets (RDSs).

Rank of RDSs	Virtual microarray data sets	Driver genes	Signaling networks
1	IBD(GSM95520,95513,95522,95518,95521,95517); CRC(GSM915466,915461,95499,95510,915465,95501)	CDK6,CEBPD,SOC51,CD19,BCL6	3,6,7
2	IBD(GSM915451,915452,95512,95523,95515,95524,95514);CRC(GSM91546,95502,915463,915470,95506,95498)	PTGS2	6
3	IBD(GSM95515,95511,95517,95521,95513,95518); CRC(GSM95496,95497,915470,95508,95500,95509)	NEDD4L	3
4	IBD(GSM95520,95523,915452,95514,915454,95525); CRC(GSM915469,915465,95505,95506,95502,95510)	SUMO1,SYT1,SOX2,PTGS2	1,3,6,7
5	IBD(GSM95512,95520,915453,95519,95511,95524); CRC(GSM95509,95507,95510,915464,915468,915467)	STK4,MORF4L2	6,7
6	IBD(GSM95518,915452,95516,95525,95522,915115); CRC(GSM915469,95500,95496,915461,95499,95497)	STAT1	6
7	IBD(GSM915451,95523,95522,95515,915453,95521); CRC(GSM915469,915464,95504,95503,915467,915460)	PTGS2	6
8	IBD(GSM915451, 915452, 915453, 915454); CRC(GSM915460, 915461, 915462, 915463, 915464, 915465, 915466, 915467, 915468, 915469,915470)	CXCR4,SMAD3,PTGS2,PTPRC,BCL6,TLR4	3,6,7

Abbreviations: IBD: inflammatory bowel disease, CRC: colorectal carcinoma.  
doi:10.1371/journal.pone.0067142.t002

handling and other processing steps [72]. For the purpose of the study, a cross-study multiple microarray data integration and normalization method were applied to infer signaling pathways involved in colorectal carcinogenesis. Recently, Li *et al.* indicated the 'one-step-clustering' of gene expression profiles and network-based gene signatures used in the past decade was far from sufficient to produce robust gene signatures [42]. To overcome the deficiency and offer robust and accurate modulated signatures, we carried out a simulation by randomly generating a cohort of virtual microarray data set gained from the combinational one. Even though each of the RDSs was subsets of the identical cohort, the variance varied diversely. Of note, we discovered 3 signaling networks including cluster 3, 6 and 7 turned out to be robust, suggesting the stimulated re-sampling from different combinations of microarray cohort could effectively cut down the false-positive signatures in the noise.

In this study, we modeled a layered signaling regulatory network via integration cross-laboratories microarray expression data based on the experimentally-supported interaction and microarray co-expression modeling. Pathway-based signaling regulatory networks analysis revealed EGF/EGFR signaling pathway, EPO signaling pathway, T cell signal transduction and members of the BCR signaling pathway were involved in colorectal carcinogenesis. Our biological interpretations are as follows:

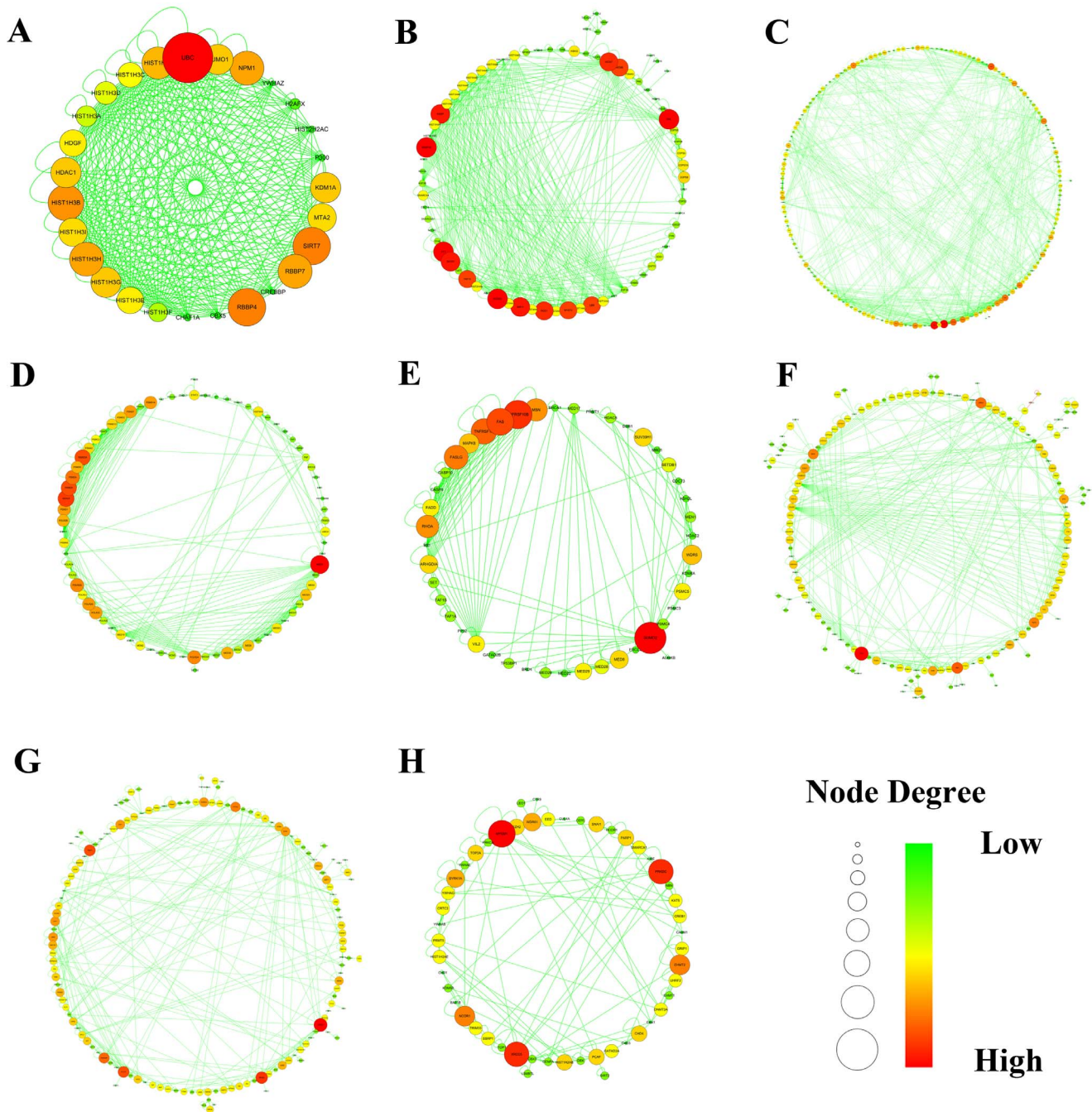
### EGF/EGFR Signaling Pathway

Numerous experimental and epidemiological investigations supplied strong evidences that EGF/EGFR signaling pathway played a pivotal role in ulcer repair. As a single-chain polypeptide of 53 amino acid residues, epidermal growth factor (EGF) could stimulate cell growth, proliferation and differentiation through autocrine, paracrine and endocrine mechanisms. It had been indicated that EGF could accelerate ulcer repair in the experimental colitis animal model [73], resulting in the alteration of downstream signaling cascades and subsequently catalyzing preferential substrates mediated by DAG, IP3 and phosphorylated RAS [74]. Recent studies demonstrated EGF could promote goblet cell mucus secretion and protect rat jejunal mucosa from injury induced by mechanical trauma [75]. Meanwhile, Luck *et al.* also discovered that EGF could significantly reduce colon ulcer and inflammation *in vivo* after intracavitary application, suggesting EGF was a protective cytokine in ulcerative colitis [76].

However, the over-expression of EGF was an aggressive biological behavior, as increased levels of EGF/EGFR were detected in innumerable types of carcinomas [77,78]. In other words, high levels of EGF or the activation of EGF/EGFR pathway had been found to significantly associate with tumor initiation and proliferation. Numerous signaling cascades including *KRAS/BRAF* and *PI3K/AKT* were certified to entangle in colorectal cancer [79]. In the presence of EGFR, the accumulation of EGF in solid tumor led to the activation of the *PI3K/AKT* pathway which subsequently gave rise to solid tumor proliferation. Therefore, tactics targeting EGF/EGFR cascade system would have been an effective and timely approach for clinical prevention and therapeutic intervention for the transition from chronic non-resolving UC to pernicious colorectal carcinoma [80].

### EPO Signaling Pathway

Pathway-based approach for analysis of GWAS also proposed the assumption that the dysfunction of EPO signaling pathway was assigned to be a highly aggressive biological process associated with colorectal carcinogenesis. Erythropoietin (EPO) was a primary growth factor regulating erythroid progenitor proliferation and maturation [81]. Despite insufficient evidences conclud-

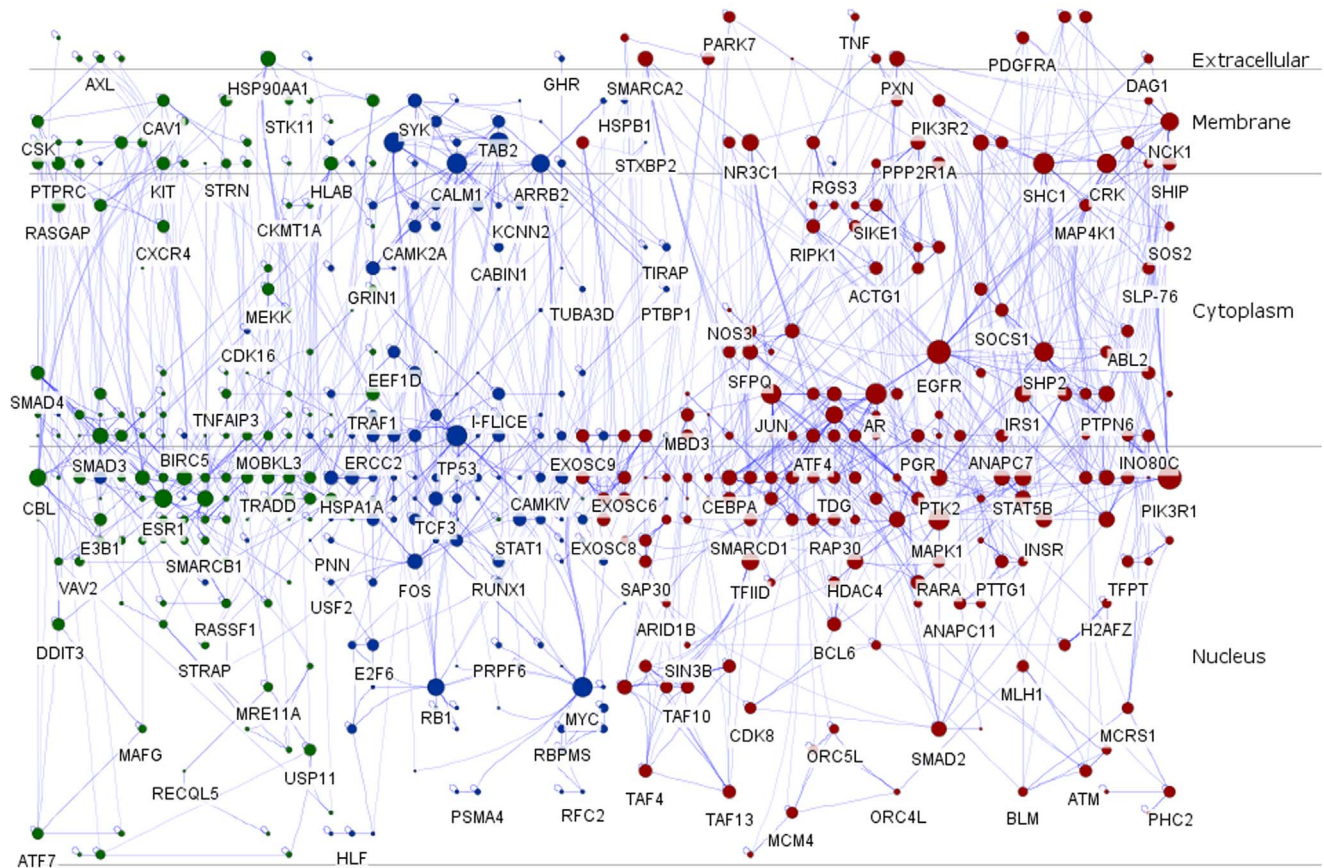


**Figure 3. Signaling regulatory modules of human cancer signaling network generated by AllegroMCODE based on molecular complex detection (MCODE) algorithm.** A-H represented signaling regulatory networks 1–8. Circle dots in the networks corresponded to genes. Red represented high degree connectivity, whereas green stood for low degree connectivity. doi:10.1371/journal.pone.0067142.g003

ed, several recent investigations still cling to provide infrequent clues that EPO signaling pathway was closely linked to IBDs. Pro-inflammatory cytokines, such as interleukin-1beta (IL-1beta), tumor necrosis factor alpha (TNF-alpha), and interferon gamma (INF-gamma), are produced in increased amounts by peripheral-blood monocytes and mononuclear cells in intestinal lamina propria in patients with IBD [82,83]. *In vitro* and *in vivo* models indicating administration erythropoietin could drastically reverse the suppression related to pro-inflammatory cytokines of cellular maturation of the erythroid lineage [84,85]. What's more, Lee

*et al.* discovered recombinant human erythropoietin (rhEPO) treatment remarkably attenuated hyperoxia-induced lung injury by down-regulating inflammatory responses in neonatal rat model [86]. It has been indicated that hypoxia and necrosis were common features of ulcerative colitis [87,88]. The elevated hypoxia-inducible factors (HIFs) level induced by hypoxic conditions in ulcerative colitis was a direct catalyst which accelerated the synthesis and release of EPO [89]. According to the documented literatures, erythropoietin deficiency contributed to the development of chronic anemia in patients with IBD [90].





**Figure 4. Layered signaling regulatory networks driven by co-differently-expressed microarray genes involved in malignant transition of colorectal cancer from the benign chronic non-solving ulcerative colitis to the more aggressive one.** In the layered signaling regulatory networks, the size of each node was proportional to the degree. In addition, red, blue and green represented nodes stemmed from signaling networks 3, 6 and 7, respectively.  
doi:10.1371/journal.pone.0067142.g004

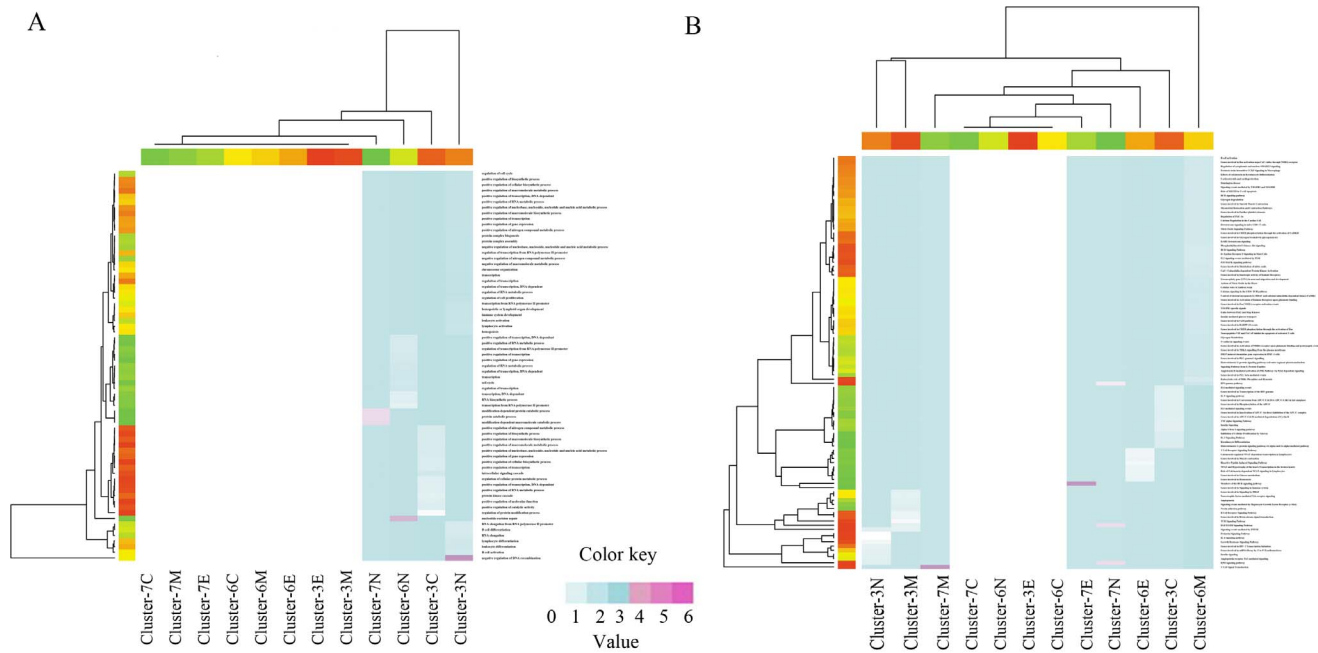
Further, in patients combined with refractory anemia and IBD, treatment with oral iron and rhEPO could raise hemoglobin and improve hematocrit [91]. Thus, high levels of EPO in ulcerative colitis led to the activation of EPO signaling pathway and therefore served as an anti-inflammatory factor against inflammation progression.

The roles of EPO had been well documented in erythrocytopoiesis, but its clinical relevance in carcinoma remained controversial and deserved for further investigations. Activation of EPO signal transduction pathway was significantly dysregulated during disorders progression to a more aggressive phenotype. As early as 2003, EPO was suggested as a rewarding guideline for cancer patients. However, owing to an observed higher mortality, concerns were raised that EPO was unprofitable for survival in cancer patients [92]. In 2003, Henke and his colleagues pointed out EPO was prosperous to correct anemia in head-and-neck cancer patients; however, it failed to improve, and even impair, cancer patient-sufferers [93]. In addition, high levels of EPO and EPO receptor (EPOR) were found in various cancer cell types and the EPO/EPOR system was known to induce proliferation, angiogenesis and even inhibit apoptosis [94]. Yasuda *et al.* examined the expression of the EPO/EPOR in 24 kinds of malignant cancerous cell lines and confirmed EPO signal transduction system was engaged in tumorigenesis of almost all malignancies [95]. Meanwhile, Mohyeldin and Lai also attested EPO was a crucial clinical signatures for head and neck squamous

cell carcinoma diagnosis [96,97]. Analogous conclusions could also be obtained in lung cancer [98], prostate cancer [99] and ovarian cancer [100], which indicated EPO/EPOR signaling system was tightly connected with tumor cell apoptosis, hypoxia resistance and metastasis. Furthermore, it had been confirmed that cancerous cell lines with EPO pretreatment rendered them less sensitive to the cytotoxicity of cisplatin [95]. In general, EPO signaling pathway appeared to be considerably altered in the malignant transition from ulcerative colitis to colorectal carcinogenesis.

### T Cell Signal Transduction

Imbalance of anti-inflammatory mediators (e.g., IL-10 and TGF) and excessive pro-inflammatory responses appeared to be a risk factor for chronic IBD [101]. Mice with IL-10-KO or lymphocyte-deficient Rag-KO were reported to develop spontaneous IBD [102]. On the contrary, several of the pro-inflammatory factors had been identified to associate with T cell production, such as IL-12, IL-23, IL-6, as well as IL-17. Of note, elevated production of memory CD4<sup>+</sup> regulatory T cells (Treg) specifically stimulated by IL-23 is especially relevant to tissue inflammation [102,103]. In the case of IBD models mentioned above, memory CD4<sup>+</sup> Treg constituted 60–80% of the whole T cells in mice with UC [104]. In addition, Fiona *et al.* also discovered that CD4/CD45RB<sup>high</sup> Treg could be confirmed as an initiator of IBD, suggesting T cell signal transduction was highly associated with chronic IBD [105].



**Figure 5. Heatmap of the over-represented biological processes and enrichment pathways using stats package in R environment.** For each figure, columns correspond to biological processes (A) or signaling pathways (B), and rows correspond to gene cluster category and subcellular localization. Expression values are logarithm of ratio value utilizing log transform data. Red and blue in each grid represented positive, while white represented null. (A) Significantly over-represented GO biological processes in differential cluster and layers. (B) Significantly enriched signaling pathways in differential cluster and layers. E: Extracellular. M: Membrane. C: Cytoplasm. N: Nucleus. doi:10.1371/journal.pone.0067142.g005

Several studies had shown a huge accumulation of Treg distributed in solid tumors and increased with the transition from benign to a malignant state [106]. Large numbers of cells with the phenotype of Treg in patients with early stage lung cancer have implications in the pathogenesis [107]. A recent study in mice revealed that the efficacy of therapeutic cancer vaccination in mice could be enhanced by removing CD4<sup>+</sup>CD25<sup>+</sup> Treg, suggesting T cell signal transduction played a role in immunosuppressive effect. Meanwhile, activated suppression by Treg had been confirmed to play an important role in T cells down-regulation [106]. Thus, interventions including selective depletion Treg, inhibition of the proliferation and immune regulation of Treg or suppression Treg from tumor microenvironment aggregation, might be an available choice [108,109]. In brief, T cell signal transduction was found to play an essential role during the transition from a benign UC to a malignant CRC.

### Members of the BCR Signaling Pathway

B cell receptor (BCR), a multi-protein complex with an antigen binding subunit and a signaling subunit, was crucial for inflammatory signaling initiation and propagation. With the assistance of costimulatory signals such as natural killer (NK) or M cytokines, B cell superantigen (BSAg) could interact with B cells and Ig, and participate in immune system or inflammatory disorders via cross-linking with BCR [110,111]. BSAg could enlarge immune systems via activation of autoreactive B cells which subsequently led to the amplification of local or systemic inflammatory reaction through interacting with BCR [112]. Silverman *et al.* uncovered, owing to the damage involved in intestinal digestive tract, polyclonal activators could mediate local immune response through penetrating lymphoid tissues [113].

Aberrant BCR signaling pathway plays significant roles in the pathogenesis of tumor immunity. Recent studies demonstrated that over-expression of regulatory B cell (Breg) could enhance tumor immunity [114]. As a member of the FOX protein family involved in immune system response, FOXP3 (forkhead box P3) was a master regulator in the development of Treg [115]. Elevated levels of Breg disturbed in tumor could transform CD4<sup>+</sup> Treg into FoxP3<sup>+</sup>Treg via direct interaction with transforming growth factor-beta (TGF-beta) pathways, resulting in tumor proliferation. However, in the absence of Breg, the transformation of FoxP3<sup>+</sup>Treg was impeded and inhibited breast cancer metastasis [116]. Thus, in view of the negative regulation of Breg, depletion of B cells might be propitious to enhance the anti-tumor capacity [117] and the application of rituximab could significantly reduce the number of mature B cells in body, which resulted in an inhibitory effect on colorectal carcinoma initiation and progression [118].

### Conclusions

A layered signaling regulatory network involved in the transition from primary UC to aggressive CRC was successfully modeled via linking cross-study microarray co-expression models with experimentally-supported interactions from the map of human cancer signaling, which led to identifying several chief cellular processes and signaling pathways. These pathological processes had been documented to be characteristics for colorectal carcinogenesis, and could be summarized into four main pathways, including EGF/EGFR signaling pathway, EPO signaling pathway, T cell signal transduction and members of the BCR signaling pathway. Therefore, our biological interpretation was focused on these pathways and their potential contributions to the transition from benign UC to aggressive or malignant CRC.

Our model for inferring the layered signaling regulatory network has two key points. For one thing, the connections between nodes or hubs in the signaling regulatory networks were integrated from cross-laboratories microarray expression profiling. For another, the layered signaling regulatory network, as opposed to the traditional ones, provided infrequent insights into molecular function and the intricate signaling pathways at the subcellular level. However, due to its positive connections and some noises that could not be avoided, the layered signaling regulatory network presented here is still not comprehensive. Nevertheless, we still confirmed that our inference would provide incentive illustration for the malignant transition of CRC, and supplied directive significance for future therapeutic intervention.

## Supporting Information

**Appendix S1 Robust multi-array average (RMA) analysis result of GSE37283 based upon WebArray platform.**  
(ZIP)

**Appendix S2 Robust multi-array average (RMA) analysis result of GSE4183 based upon WebArray platform.**  
(ZIP)

## References

- Boyle P, Langman JS (2000) ABC of colorectal cancer: Epidemiology. *BMJ* 321: 805–808.
- Boyle P (1995) Progress in preventing death from colorectal cancer. *Br J Cancer* 72: 528–530.
- Hendon SE, DiPalma JA (2005) U.S. practices for colon cancer screening. *Keio J Med* 54: 179–183.
- Devroede GJ, Taylor WF, Sauer WG, Jackman RJ, Stickler GB (1971) Cancer risk and life expectancy of children with ulcerative colitis. *N Engl J Med* 285: 17–21.
- Yancik R (2005) Population aging and cancer: a cross-national concern. *Cancer J* 11: 437–441.
- Giovannucci E, Colditz GA, Stampfer MJ, Willett WC (1996) Physical activity, obesity, and risk of colorectal adenoma in women (United States). *Cancer Causes Control* 7: 253–263.
- Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, et al. (2000) Environmental and heritable factors in the causation of cancer—analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med* 343: 78–85.
- Rossi S, Kopetz S, Davuluri R, Hamilton SR, Calin GA (2010) MicroRNAs, ultraconserved genes and colorectal cancers. *Int J Biochem Cell Biol* 42: 1291–1297.
- Kanthan R, Senger JL, Kanthan SC (2012) Molecular events in primary and metastatic colorectal carcinoma: a review. *Patholog Res Int* 2012: 597497.
- Nunnari G, Berretta M, Pinzone MR, Di Rosa M, Berretta S, et al. (2012) Hepatocellular carcinoma in HIV positive patients. *Eur Rev Med Pharmacol Sci* 16: 1257–1270.
- Smolle E, Zohrer E, Bettermann K, Haybaeck J (2012) Viral hepatitis induces hepatocellular cancer: what can we learn from epidemiology comparing iran and worldwide findings? *Hepat Mon* 12: e7879.
- Hartnett L, Egan LJ (2012) Inflammation, DNA methylation and colitis-associated cancer. *Carcinogenesis* 33: 723–731.
- Fuseya C, Horiuchi A, Hayashi A, Suzuki A, Miyamoto T, et al. (2012) Involvement of pelvic inflammation-related mismatch repair abnormalities and microsatellite instability in the malignant transformation of ovarian endometriosis. *Hum Pathol* 43: 1964–1972.
- Izkowitz SH, Greenwald B, Meltzer SJ (1995) Colon carcinogenesis in inflammatory bowel disease. *Inflamm Bowel Dis* 1: 142–158.
- Lakatos PL, Lakatos L (2012) Challenges in calculating the risk for colorectal cancer in patients with ulcerative colitis. *Clin Gastroenterol Hepatol* 10: 1179–1180.
- Rubin DC, Shaker A, Levin MS (2012) Chronic intestinal inflammation: inflammatory bowel disease and colitis-associated colon cancer. *Front Immunol* 3: 107.
- Okayasu I (2012) Development of ulcerative colitis and its associated colorectal neoplasia as a model of the organ-specific chronic inflammation-carcinoma sequence. *Pathol Int* 62: 368–380.
- Jiang W, Shadrach B, Carver P, Goldblum JR, Shen B, et al. (2012) Histomorphologic and molecular features of pouch and peripouch adenocarcinoma: a comparison with ulcerative colitis-associated adenocarcinoma. *Am J Surg Pathol* 36: 1385–1394.
- de Matos SR, Tripathi S, Emmert-Streib F (2012) Organizational structure and the periphery of the gene regulatory network in B-cell lymphoma. *BMC Syst Biol* 6: 38.
- Gu Z, Zhang C, Wang J (2012) Gene regulation is governed by a core network in hepatocellular carcinoma. *BMC Syst Biol* 6: 32.
- Latchman DS (1997) Transcription factors: an overview. *Int J Biochem Cell Biol* 29: 1305–1312.
- Lee TI, Young RA (2000) Transcription of eukaryotic protein-coding genes. *Annu Rev Genet* 34: 77–137.
- Nikolov DB, Burley SK (1997) RNA polymerase II transcription initiation: a structural view. *Proc Natl Acad Sci U S A* 94: 15–22.
- Chen K, Rajewsky N (2007) The evolution of gene regulation by transcription factors and microRNAs. *Nat Rev Genet* 8: 93–103.
- Kauffman S (1974) The large scale structure and dynamics of gene control circuits: an ensemble approach. *J Theor Biol* 44: 167–190.
- Kauffman SA (1969) Metabolic stability and epigenesis in randomly constructed genetic nets. *J Theor Biol* 22: 437–467.
- Rice JJ, Tu Y, Stolovitzky G (2005) Reconstructing biological networks using conditional correlation analysis. *Bioinformatics* 21: 765–773.
- Friedman N (2004) Inferring cellular networks using probabilistic graphical models. *Science* 303: 799–805.
- Chen T, He HL, Church GM (1999) Modeling gene expression with differential equations. *Pac Symp Biocomput*: 29–40.
- Dehmer M, Emmert-Streib F (2009) *Analysis of Complex Networks*. Wiley-Blackwell.
- Margolin AA, Nemenman I, Basso K, Wiggins C, Stolovitzky G, et al. (2006) ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics* 7 Suppl 1: S7.
- Shalgi R, Lieber D, Oren M, Pilpel Y (2007) Global and local architecture of the mammalian microRNA-transcription factor regulatory network. *PLoS Comput Biol* 3: e131.
- Wu LF, Hughes TR, Davierwala AP, Robinson MD, Stoughton R, et al. (2002) Large-scale prediction of *Saccharomyces cerevisiae* gene function using overlapping transcriptional clusters. *Nat Genet* 31: 255–265.
- Walker MG, Volkmut W, Sprinzak E, Hodgson D, Klingler T (1999) Prediction of gene function by genome-scale expression analysis: prostate cancer-associated genes. *Genome Res* 9: 1198–1203.
- Mostafavi S, Ray D, Warde-Farley D, Grouios C, Morris Q (2008) GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. *Genome Biol* 9 Suppl 1: S4.
- Macarron R, Banks MN, Bojanic D, Burns DJ, Cirovic DA, et al. (2011) Impact of high-throughput screening in biomedical research. *Nat Rev Drug Discov* 10: 188–195.
- Mullard A (2012) Oncology trials gear up for high-throughput sequencing. *Nat Rev Drug Discov* 11: 339–340.
- Zhang X, Zhao XM, He K, Lu L, Cao Y, et al. (2012) Inferring gene regulatory networks from gene expression data by path consistency algorithm based on conditional mutual information. *Bioinformatics* 28: 98–104.
- Zhu W, Yang L, Du Z (2009) Layered functional network analysis of gene expression in human heart failure. *PLoS One* 4: e6288.
- Kelly DJ, Ghosh S (2005) RNA profiling for biomarker discovery: practical considerations for limiting sample sizes. *Dis Markers* 21: 43–48.

**Appendix S3 Cross-study normalization and integration results (part 1) using ArrayMining.**

(ZIP)

**Appendix S4 Cross-study normalization and integration results (part 2) using ArrayMining.**

(ZIP)

**Appendix S5 Detailed list of the differently expressed genes by Significant Analysis of Microarray (SAM) originated from the simulated microarray data sets or virtual random microarray data sets (RDSs).**

(ZIP)

**Appendix S6 A map of human cancer signaling network.**

(ZIP)

## Author Contributions

Conceived and designed the experiments: XJL SF. Performed the experiments: XJL SF. Analyzed the data: XJL SF. Contributed reagents/materials/analysis tools: XJL SF ZP QG. Wrote the paper: XJL SF. Corrected this manuscript: XJL YW YA YX LT YP.

41. Xu L, Tan AC, Winslow RL, Geman D (2008) Merging microarray data from separate breast cancer studies provides a robust prognostic test. *BMC Bioinformatics* 9: 125.
42. Li J, Lenferink AE, Deng Y, Collins C, Cui Q, et al. (2010) Identification of high-quality cancer prognostic markers and metastasis network modules. *Nat Commun* 1: 34.
43. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, et al. (2004) Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. *Proc Natl Acad Sci U S A* 101: 9309–9314.
44. Woychik RP, Klebig ML, Justice MJ, Magnuson TR, Avner ED (1998) Functional genomics in the post-genome era. *Mutat Res* 400: 3–14.
45. Xing B, van der Laan MJ (2005) A causal inference approach for constructing transcriptional regulatory networks. *Bioinformatics* 21: 4007–4013.
46. Pedroso I (2010) Gaining a pathway insight into genetic association data. *Methods Mol Biol* 628: 373–382.
47. Natarajan M, Lin KM, Hsueh RC, Sternweis PC, Ranganathan R (2006) A global analysis of cross-talk in a mammalian cellular signalling network. *Nat Cell Biol* 8: 571–580.
48. Cui Q, Ma Y, Jaramillo M, Bari H, Awan A, et al. (2007) A map of human cancer signaling. *Mol Syst Biol* 3: 152.
49. Fan S, Geng Q, Pan Z, Li X, Tie L, et al. (2012) Clarifying off-target effects for torcetrapib using network pharmacology and reverse docking approach. *BMC Syst Biol* 6: 152.
50. van Lunteren E, Moyer M (2009) Gene expression profiling in the type 1 diabetes rat diaphragm. *PLoS One* 4: e7832.
51. Andrew AS, Hu T, Gu J, Gui J, Ye Y, et al. (2012) HSD3B and gene-gene interactions in a pathway-based analysis of genetic susceptibility to bladder cancer. *PLoS One* 7: e51301.
52. Mahurkar S, Moldovan M, Suppiah V, O'Doherty C (2013) Identification of shared genes and pathways: a comparative study of multiple sclerosis susceptibility, severity and response to interferon Beta treatment. *PLoS One* 8: e57655.
53. Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, et al. (2007) Integration of biological networks and gene expression data using Cytoscape. *Nat Protoc* 2: 2366–2382.
54. Martin A, Ochagavia ME, Rabasa LC, Miranda J, Fernandez-de-Cossio J, et al. (2010) BisoGenet: a new tool for gene network building, visualization and analysis. *BMC Bioinformatics* 11: 91.
55. Barsky A, Gardy JL, Hancock RE, Muzner T (2007) Cerebral: a Cytoscape plugin for layout of and interaction with biological networks using subcellular localization annotation. *Bioinformatics* 23: 1040–1042.
56. Gyorffy B, Molnar B, Lage H, Szallasi Z, Eklund AC (2009) Evaluation of microarray preprocessing algorithms based on concordance with RT-PCR in clinical samples. *PLoS One* 4: e5645.
57. Pekow J, Dougherty U, Huang Y, Gomez E, Nathanson J, et al. (2013) Gene signature distinguishes patients with chronic ulcerative colitis harboring remote neoplastic lesions. *Inflamm Bowel Dis* 19: 461–470.
58. Xia X, McClelland M, Wang Y (2005) WebArray: an online platform for microarray data analysis. *BMC Bioinformatics* 6: 306.
59. Chen R, Li L, Butte AJ (2007) AILUN: reannotating gene expression data automatically. *Nat Methods* 4: 879.
60. Glaab E, Garibaldi JM, Krasnogor N (2009) ArrayMining: a modular web-application for microarray analysis combining ensemble and consensus methods with cross-study normalization. *BMC Bioinformatics* 10: 358.
61. Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 98: 5116–5121.
62. Awan A, Bari H, Yan F, Moksong S, Yang S, et al. (2007) Regulatory network motifs and hotspots of cancer genes in a mammalian cellular signalling network. *IET Syst Biol* 1: 292–297.
63. Ma'ayan A, Jenkins SL, Neves S, Hasseldine A, Grace E, et al. (2005) Formation of regulatory patterns during signal propagation in a Mammalian cellular network. *Science* 309: 1078–1083.
64. Bader GD, Hogue CW (2003) An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* 4: 2.
65. Huang DW, Sherman BT, Lempicki RA (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 37: 1–13.
66. Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4: 44–57.
67. Chen J, Bardes EE, Aronow BJ, Jegga AG (2009) ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res* 37: W305–W311.
68. Johnson WE, Li C, Rabinovic A (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8: 118–127.
69. Zhu H, Qian J (2012) Applications of functional protein microarrays in basic and clinical research. *Adv Genet* 79: 123–155.
70. Tiwari M (2012) Microarrays and cancer diagnosis. *J Cancer Res Ther* 8: 3–10.
71. de la Fuente A (2010) From 'differential expression' to 'differential networking' - identification of dysfunctional regulatory networks in diseases. *Trends Genet* 26: 326–333.
72. Zakharkin SO, Kim K, Mehta T, Chen L, Barnes S, et al. (2005) Sources of variation in Affymetrix microarray experiments. *BMC Bioinformatics* 6: 214.
73. Procaccino F, Reinshagen M, Hoffmann P, Zeeh JM, Lakshmanan J, et al. (1994) Protective effect of epidermal growth factor in an experimental model of colitis in rats. *Gastroenterology* 107: 12–17.
74. King AC, Cuatrecasas P (1981) Peptide hormone-induced receptor mobility, aggregation, and internalization. *N Engl J Med* 305: 77–88.
75. Ishikawa S, Cepinskas G, Specian RD, Itoh M, Kvietyts PR (1994) Epidermal growth factor attenuates jejunal mucosal injury induced by oleic acid: role of mucus. *Am J Physiol* 267: G1067–G1077.
76. Luck MS, Bass P (1993) Effect of epidermal growth factor on experimental colitis in the rat. *J Pharmacol Exp Ther* 264: 984–990.
77. Lafky JM, Wilken JA, Baron AT, Mailhe NJ (2008) Clinical implications of the ErbB/epidermal growth factor (EGF) receptor family and its ligands in ovarian cancer. *Biochim Biophys Acta* 1785: 232–265.
78. Henson ES, Gibson SB (2006) Surviving cell death through epidermal growth factor (EGF) signal transduction pathways: implications for cancer therapy. *Cell Signal* 18: 2089–2097.
79. Berg M, Soreide K (2012) EGFR and downstream genetic alterations in KRAS/BRAF and PI3K/AKT pathways in colorectal cancer: implications for targeted therapy. *Discov Med* 14: 207–214.
80. Merla A, Goel S (2012) Novel drugs targeting the epidermal growth factor receptor and its downstream pathways in the treatment of colorectal cancer: a systematic review. *Chemother Res Pract* 2012: 387172.
81. Bouscary D, Pene F, Claessens YE, Muller O, Chretien S, et al. (2003) Critical role for PI 3-kinase in the control of erythropoietin-induced erythroid progenitor proliferation. *Blood* 101: 3436–3443.
82. Mahida YR, Wu K, Jewell DP (1989) Enhanced production of interleukin 1-beta by mononuclear cells isolated from mucosa with active ulcerative colitis of Crohn's disease. *Gut* 30: 835–838.
83. Reinecker HC, Steffen M, Witthoef T, Pflueger I, Schreiber S, et al. (1993) Enhanced secretion of tumour necrosis factor-alpha, IL-6, and IL-1 beta by isolated lamina propria mononuclear cells from patients with ulcerative colitis and Crohn's disease. *Clin Exp Immunol* 94: 174–181.
84. Smith MA, Knight SM, Maddison PJ, Smith JG (1992) Anaemia of chronic disease in rheumatoid arthritis: effect of the blunted response to erythropoietin and of interleukin 1 production by marrow macrophages. *Ann Rheum Dis* 51: 753–757.
85. Faquin WC, Schneider TJ, Goldberg MA (1992) Effect of inflammatory cytokines on hypoxia-induced erythropoietin production. *Blood* 79: 1987–1994.
86. Lee JH, Sung DK, Koo SH, Shin BK, Hong YS, et al. (2007) Erythropoietin attenuates hyperoxia-induced lung injury by down-modulating inflammation in neonatal rats. *J Korean Med Sci* 22: 1042–1047.
87. Ordas I, Eckmann L, Talamini M, Baumgart DC, Sandborn WJ (2012) Ulcerative colitis. *Lancet* 380: 1606–1619.
88. Parray FQ, Wani ML, Malik AA, Wani SN, Bijli AH, et al. (2012) Ulcerative colitis: a challenge to surgeons. *Int J Prev Med* 3: 749–763.
89. Phillips TM, Kim K, Vlashi E, McBride WH, Pajonk F (2007) Effects of recombinant erythropoietin on breast cancer-initiating cells. *Neoplasia* 9: 1122–1129.
90. Erslev AJ, Caro J, Miller O, Silver R (1980) Plasma erythropoietin in health and disease. *Ann Clin Lab Sci* 10: 250–257.
91. Schreiber S, Howaldt S, Schnoor M, Nikolaus S, Bauditz J, et al. (1996) Recombinant erythropoietin for the treatment of anemia in inflammatory bowel disease. *N Engl J Med* 334: 619–623.
92. Leyland-Jones B (2003) Breast cancer trial with erythropoietin terminated unexpectedly. *Lancet Oncol* 4: 459–460.
93. Henke M, Laszig R, Rube C, Schafer U, Haase KD, et al. (2003) Erythropoietin to treat head and neck cancer patients with anaemia undergoing radiotherapy: randomised, double-blind, placebo-controlled trial. *Lancet* 362: 1255–1260.
94. Giatromanolaki A, Sivridis E, Maltezos E, Papazoglou D, Simopoulos C, et al. (2003) Hypoxia inducible factor 1alpha and 2alpha overexpression in inflammatory bowel disease. *J Clin Pathol* 56: 209–213.
95. Yasuda Y, Fujita Y, Matsuo T, Koinuma S, Hara S, et al. (2003) Erythropoietin regulates tumour growth of human malignancies. *Carcinogenesis* 24: 1021–1029.
96. Mohyeldin A, Lu H, Dalgard C, Lai SY, Cohen N, et al. (2005) Erythropoietin signaling promotes invasiveness of human head and neck squamous cell carcinoma. *Neoplasia* 7: 537–543.
97. Lai SY, Childs EE, Xi S, Coppelli FM, Gooding WE, et al. (2005) Erythropoietin-mediated activation of JAK-STAT signaling contributes to cellular invasion in head and neck squamous cell carcinoma. *Oncogene* 24: 4442–4449.
98. Dagnon K, Pacary E, Commo F, Antoine M, Bernaudin M, et al. (2005) Expression of erythropoietin and erythropoietin receptor in non-small cell lung carcinomas. *Clin Cancer Res* 11: 993–999.
99. Feldman L, Wang Y, Rhim JS, Bhattacharya N, Loda M, et al. (2006) Erythropoietin stimulates growth and STAT5 phosphorylation in human prostate epithelial and prostate cancer cells. *Prostate* 66: 135–145.
100. McBroom JW, Acs G, Rose GS, Krivak TC, Mohyeldin A, et al. (2005) Erythropoietin receptor function and expression in epithelial ovarian carcinoma. *Gynecol Oncol* 99: 571–577.

101. Braegger CP (1994) Immunopathogenesis of chronic inflammatory bowel disease. *Acta Paediatr Suppl* 83: 18–21.
102. Yen D, Cheung J, Scheerens H, Poulet F, McClanahan T, et al. (2006) IL-23 is essential for T cell-mediated colitis and promotes inflammation via IL-17 and IL-6. *J Clin Invest* 116: 1310–1316.
103. Oppmann B, Lesley R, Blom B, Timans JC, Xu Y, et al. (2000) Novel p19 protein engages IL-12p40 to form a cytokine, IL-23, with biological activities similar as well as distinct from IL-12. *Immunity* 13: 715–725.
104. Davidson NJ, Leach MW, Fort MM, Thompson-Snipes L, Kuhn R, et al. (1996) T helper cell 1-type CD4+ T cells, but not B cells, mediate colitis in interleukin 10-deficient mice. *J Exp Med* 184: 241–251.
105. Powrie F, Correa-Oliveira R, Mauze S, Coffman RL (1994) Regulatory interactions between CD45RB<sup>high</sup> and CD45RB<sup>low</sup> CD4+ T cells are important for the balance between protective and pathogenic cell-mediated immunity. *J Exp Med* 179: 589–600.
106. Woo EY, Chu CS, Goletz TJ, Schlienger K, Yeh H, et al. (2001) Regulatory CD4(+)/CD25(+) T cells in tumors from patients with early-stage non-small cell lung cancer and late-stage ovarian cancer. *Cancer Res* 61: 4766–4772.
107. von Boehmer H (2005) Mechanisms of suppression by suppressor T cells. *Nat Immunol* 6: 338–344.
108. Steitz J, Bruck J, Lenz J, Knop J, Tuting T (2001) Depletion of CD25(+)/CD4(+) T cells and treatment with tyrosinase-related protein 2-transduced dendritic cells enhance the interferon alpha-induced, CD8(+) T-cell-dependent immune defense of B16 melanoma. *Cancer Res* 61: 8643–8646.
109. Shimizu J, Yamazaki S, Sakaguchi S (1999) Induction of tumor immunity by removing CD25+CD4+ T cells: a common basis between tumor immunity and autoimmunity. *J Immunol* 163: 5211–5218.
110. Snapper CM, Mond JJ (1996) A model for induction of T cell-independent humoral immunity in response to polysaccharide antigens. *J Immunol* 157: 2229–2233.
111. Kristiansen SV, Pascual V, Lipsky PE (1994) Staphylococcal protein A induces biased production of Ig by VH3-expressing B lymphocytes. *J Immunol* 153: 2974–2982.
112. Levy JA (1993) Pathogenesis of human immunodeficiency virus infection. *Microbiol Rev* 57: 183–289.
113. Silverman GJ (1997) B-cell superantigens. *Immunol Today* 18: 379–386.
114. DiLillo DJ, Matsushita T, Tedder TF (2010) B10 cells and regulatory B cells balance immune responses during inflammation, autoimmunity, and cancer. *Ann N Y Acad Sci* 1183: 38–57.
115. Zhang L, Zhao Y (2007) The regulation of Foxp3 expression in regulatory CD4(+)/CD25(+)T cells: multiple pathways on the road. *J Cell Physiol* 211: 590–597.
116. Olkhanud PB, Damdinsuren B, Bodogai M, Gress RE, Sen R, et al. (2011) Tumor-evoked regulatory B cells promote breast cancer metastasis by converting resting CD4(+) T cells to T-regulatory cells. *Cancer Res* 71: 3505–3515.
117. Tadmor T, Zhang Y, Cho HM, Podack ER, Rosenblatt JD (2011) The absence of B lymphocytes reduces the number and function of T-regulatory cells and enhances the anti-tumor response in a murine tumor model. *Cancer Immunol Immunother* 60: 609–619.
118. Barbera-Guillem E, Nelson MB, Barr B, Nyhus JK, May KJ, et al. (2000) B lymphocyte pathology in human colorectal cancer. Experimental and clinical therapeutic effects of partial B cell depletion. *Cancer Immunol Immunother* 48: 541–549.