

Transcriptome analysis of human colorectal cancer biopsies reveals extensive expression correlations among genes related to cell proliferation, lipid metabolism, immune response and collagen catabolism

Lai Xu^{1,*}, Rong Wang^{1,*}, Joseph Ziegelbauer², Wells W. Wu³, Rong-Fong Shen³, Hartmut Juhl⁴, Yaqin Zhang¹, Lorraine Pelosof⁵ and Amy S. Rosenberg¹

¹Office of Biotechnology Products, CDER, FDA, Silver Spring, MD 20993, USA

²HIV/AIDS Malignancy Branch, NCI, Bethesda, MD 20892, USA

³Facility for Biotechnology Resources, CBER, FDA, Silver Spring, MD 20993, USA

⁴Indivumed GMBH, Hamburg 20251, Germany

⁵Office of Hematology and Oncology Products, CDER, FDA, Silver Spring, MD 20993, USA

*These authors have contributed equally to this work

Correspondence to: Amy S. Rosenberg, email: amy.rosenberg@FDA.hhs.gov

Keywords: cell cycle, lipid metabolism, inflammation, collagen catabolism, gene co-expression network

Received: April 28, 2016 **Accepted:** July 12, 2017 **Published:** August 18, 2017

Copyright: Xu et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Precise characterization of biological processes critical to proliferation and metastasis of colorectal cancer should facilitate the development of diagnostic and prognostic biomarkers as well as novel treatments. Using mRNA-Seq, we examined the protein coding messenger RNA (mRNA) expression profiles across different histologically defined stages of primary colon cancers and compared them to their patient matched normal tissue controls. In comparing 79 colorectal cancers to their matched normal mucosa, tumors were distinguished from normal non-malignant tissues not only in the upregulation of biological processes pertaining to cell proliferation, inflammation, and tissue remodeling, but even more strikingly, in downregulated biological processes including fatty acid beta oxidization for ATP production and epithelial cell differentiation and function. A network analysis of deregulated genes revealed newly described cancer networks and putative hub genes.

Taken together, our findings suggest that, within an inflammatory microenvironment, invasive, dedifferentiated and rapidly dividing tumor cells divert the oxidation of fatty acids and lipids from energy production into lipid components of cell membranes and organelles to support tumor proliferation. A gene co-expression network analysis provides a clear and broad picture of biological pathways in tumors that may significantly enhance or supplant current histopathologic studies.

INTRODUCTION

Colorectal cancer is the third most common cancer and the third most common cause of cancer-related deaths in men and women in the United States [1]. An emerging theme in cancer biology is that metabolic regulation is intricately linked to cancer progression and therefore,

factors that promote proliferation may also directly or indirectly promote metabolic changes to support rapid proliferation and metastasis [2]. Immunohistochemical staining [3] and microarray analyses [4] have shown that cancer cells, unlike their healthy tissue counterparts, switch from aerobic mitochondrial oxidative phosphorylation to glycolysis (The Warburg Effect) as the

primary energy source even in the context of an aerobic microenvironment. Glycolysis provides cancer cells not only with energy, but also with metabolites including ribose sugars, glycerol, citrate, and nonessential amino acids which are essential for cellular proliferation [5–7].

We identified cancer specific gene expression patterns and cancer associated signaling and regulatory pathways in colorectal cancer samples as compared to their patient matched healthy colonic tissues using mRNA-Seq. Our study sought to explore expression correlations among dysregulated genes pertaining to biological pathways critical in tumor proliferation and metastasis by identifying and quantifying critical mRNA gene co-expression networks and hub genes.

RESULTS

Differential gene expression in colorectal cancer vs. normal intestinal mucosa

To identify genes that were dysregulated in tumors compared with normal tissues, 79 tumor (T) and normal (N) paired samples were evaluated. From 10,255 genes which showed abundant expression levels (FPKM >1) by mRNA-Seq, 2,358 genes were differentially expressed between tumor and normal tissues (Supplementary Table 2) using two criteria: a greater than 2 fold expression level change and p-value (FDR) ≤ 0.05 from ANOVA test. These 2,358 genes included 1,223 upregulated genes ($T/N > 2$) and 1,135 downregulated genes ($N/T > 2$) as shown in the volcano plot (Figure 1A). Unsupervised hierarchical clustering analysis (HCA) of these 2,358 genes showed that 77 normal samples and one low stage tumor sample clustered into a distinct group while 78 tumor samples and 2 normal samples associated with high stage tumors clustered into another group (Figure 1B). Similar results were revealed by a principal component analysis (PCA) (Figure 1C). These data demonstrate that colorectal cancers can generally, but not in every case, be distinguished from their adjacent normal samples using large-scale gene expression analyses.

Upregulated genes and their corresponding biological processes

Gene Ontology (GO) analysis of the 1,223 upregulated genes revealed a focus on six key biological processes as defined by DAVID bioinformatics: collagen catabolic process, chemokine-mediated signaling pathway, inflammatory response, cell proliferation, response to IFN γ , and immune response (Figure 2). Since these biological processes principally pertain to tumor environmental alterations (inflammation, proliferation, and invasion), we investigated the differences among tumors of different histologic stages and grades as well as between tumors and normal samples. The clustering results (HCA)

revealed that tumors clustered at the high expression zones (red) while normal samples clustered at the low expression zones (blue) (Figure 3). PCA plots mirrored the HCA plots showing moderate heterogeneity among tumors.

The most prominently upregulated genes consisted of families of peptidases, proteinases, and chemokines, including kallikrein-related peptidases (KLKs), matrix metalloproteinase (MMPs), and chemokines. Strikingly, the KLK peptidase genes showed the greatest upregulation: *KLK6* (231 fold increase); *KLK7* (223 fold increase); *KLK10* (43 fold increase); *KLK11* (13 fold increase); and *KLK12* (10 fold increase). In fact, *KLK6* and *KLK7* showed almost no expression in normal samples but have previously been found to be early stage biomarkers for ovarian cancer [11]. Matrix metalloproteinases (MMPs) sharply upregulated in tumors included *MMP7* (126 fold increase), *MMP3* (62 fold increase), *MMP1* (31 fold increase), *MMP10* (13 fold increase), *MMP11* (12 fold increase), and *MMP12* (10 fold increase). KLKs and MMPs are known to be involved in extracellular matrix (ECM) remodeling via degradation of extracellular proteins and collagen matrix. Such space remodeling is necessary for tumor expansion, metastasis and epithelial to mesenchymal transition [12–15]. MMPs also recruit neutrophils by generating bioactive fragments (such as N-acetyl Pro-Gly-Pro) from the ECM [16]. Tumor upregulated chemokines included the followings: *CXCL5* (138 fold increase); *CCL25* (70 fold increase); *CXCL11* (26 fold increase); *CXCL6* (20 fold increase); *CXCL1* and *CXCL10* (11 fold increase); *CXCL3* (8 fold increase); and *CXCL9*, *CXCL2*, and *CCL3L1* (6 fold increase). Other chemokines with lesser upregulation are included in Supplementary Table 2. Upregulated cytokine genes in tumors included *IL11* (26 fold increase), *IL24* (24 fold increase), *IL6* (17 fold increase), *IL1A* (14 fold increase), *IL1B* (10 fold increase), *IL23A* (8 fold increase), and *IL33* (3 fold increase). Intriguingly, although these cytokines are upregulated, the respective receptor genes for *IL1*, *IL6*, *IL11*, and *IL23* are downregulated ($N/T > 1.5$ fold). Only the receptor gene for *IL24* is upregulated ($T/N > 2$ fold) within our dataset.

The induction of genes pertaining to multiple chemokines and cytokines suggests a complex tumor microenvironment composed of different cell types that cohabit and communicate with each other via soluble mediators to promote tumor survival, proliferation and metastasis [17, 18]. The high level expression of genes related to macrophages (Supplementary Figure 1A) and neutrophils (Supplementary Figure 1B) further supports this hypothesis. Despite the elevated expression level of chemokine and cytokine genes, there was diminished expression of T and B cell related genes in tumor tissues compared to their normal control tissues (Supplementary Figure 1C). Indeed, the absence of T and B cells in colorectal cancers has been correlated with unfavorable overall survival [19]. The downregulation of T and B

cell related genes could be explained by the decreased expression of some chemokines related to the recruitment of T and B cells [20–23], such as *CCL23* (7 fold decrease), *CXCL12* (4 fold decrease), *CCL8* (3 fold decrease), and *CCL13* (2 fold decrease). Additional upregulated genes in tumors that pertain to cell proliferation include those of the regenerating family (REG) genes, such as *REG3A* (23 fold increase), *REG4* (9 fold increase), and *REG1A* (4 fold increase), which are known to be involved in cell proliferation and have a key role in carcinogenesis by activating the AKT and ERK1/2 pathways [24]. As well, upregulated mRNA expression of *REG1A* has been shown as an unfavorable prognostic marker in colorectal cancer and has been associated with peritoneal carcinomatosis [25]. Intriguingly, the mRNA and protein products of these *REG* genes and *CCL25* gene were highly expressed in autoimmune/inflammatory diseases of the intestinal tract including ulcerative colitis (UC), a condition with a high probability of development of colon cancer [26–28].

Well known oncogenes showed the expected upregulation across our samples including *MYC* (4 fold increase), karyopherin alpha 2 (*KPNA2*) (3 fold increase), pituitary tumor transforming gene (*PTTG1*) (2 fold increase), high mobility group AT-hook 1 (*HMGA1*) (3 fold increase), and nucleolar and coiled-body phosphoprotein 1 (*NOLC1*) (3 fold increase). The upregulation of *PHF19* (3 fold increase) which codes for PHD (Cys4-His-Cys motif) finger protein 19, a histone binding protein with embryonic stem cell self-renewal function [29], suggests it may have oncogenic function in colorectal cancer.

Downregulated genes and their corresponding biological processes

Gene Ontology analysis of 1,135 downregulated genes revealed dramatic differences between tumors and healthy colonic tissue pertaining to four biological processes: fatty acid beta oxidation, colonic detoxification

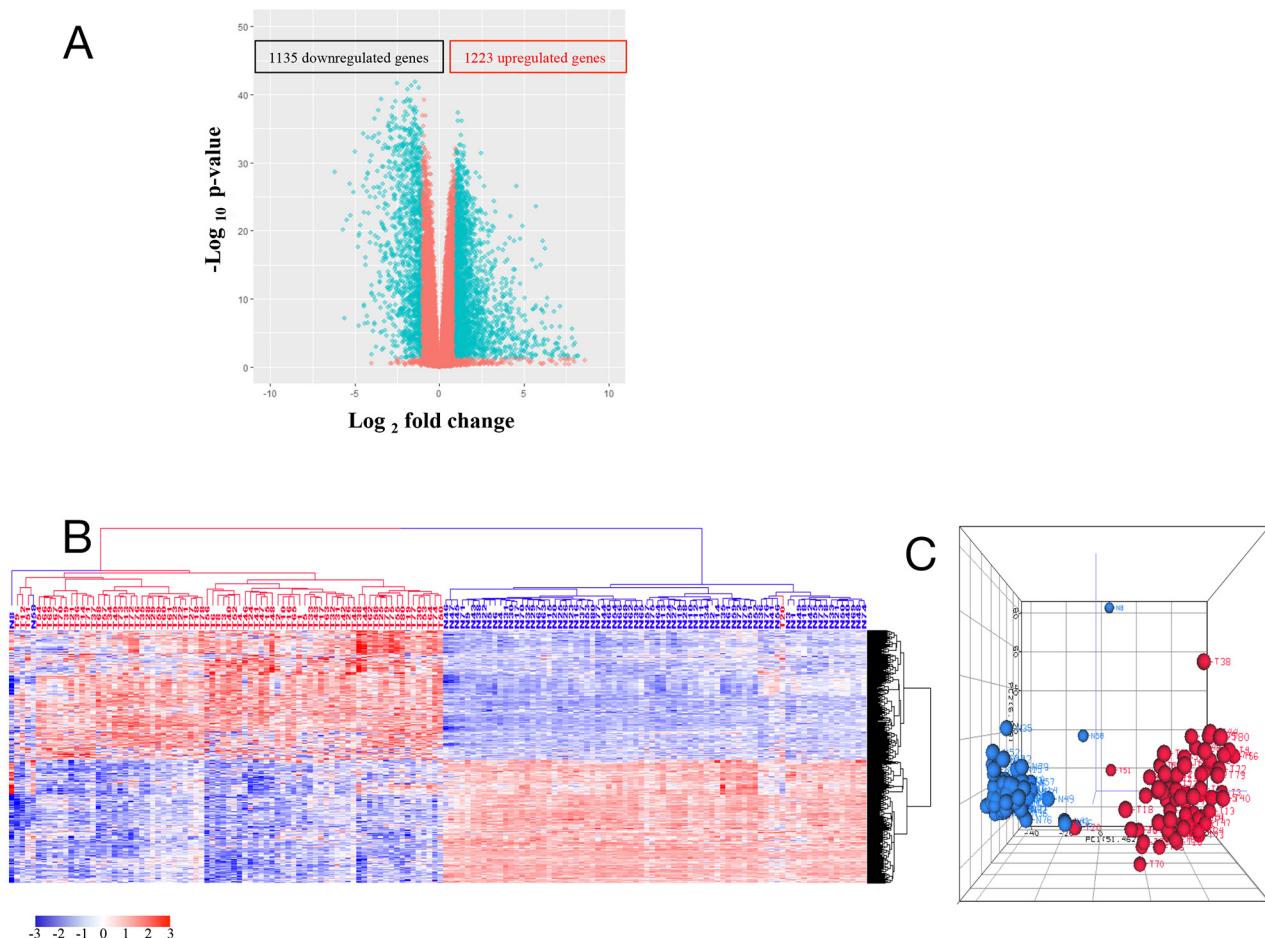


Figure 1: (A) Volcano plot of 10,255 genes with average FPKM > 1 and FDR ≤ 0.05 . Each circle corresponds to one gene. The x-axis represents gene $\log_2(T/N)$ value, and the y-axis represents \log_{10} p-value calculated from ANOVA test between tumor and normal controls. T represents gene FPKM value in tumor sample, and N represents the gene FPKM value in normal control. **(B)** and **(C)** HCA (b) and PCA (c) of 1,223 upregulated genes and 1,135 downregulated genes. Tumor is marked with red dendrogram and normal sample is marked with blue dendrogram.

of inorganic heavy metals (such as cadmium), colonic detoxification of organic toxins (such as antibiotics), and colonic epithelium differentiation. Tumors clustered at the low expression zones (blue) while normal samples

clustered at the high expression zones (red) of HCAs (Figure 4). The PCA plots, which mirrored the HCA plots, show greater heterogeneity of normal samples compared with tumors in downregulated genes (Figure

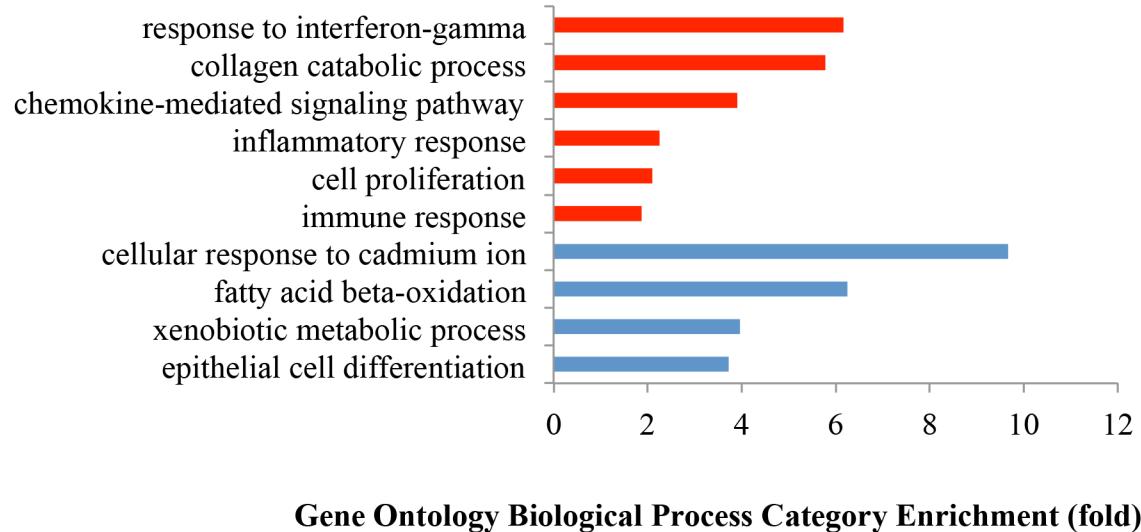


Figure 2: Gene Ontology enrichment analysis of biological processes of the differentially expressed genes between tumors and normal tissues. The biological processes were selected using a False Discovery Rate (FDR) < 0.05 in DAVID. Red bars represent the significantly upregulated biological processes with FDR < 0.05; while the blue bars represent the significantly downregulated biological processes with FDR < 0.05.

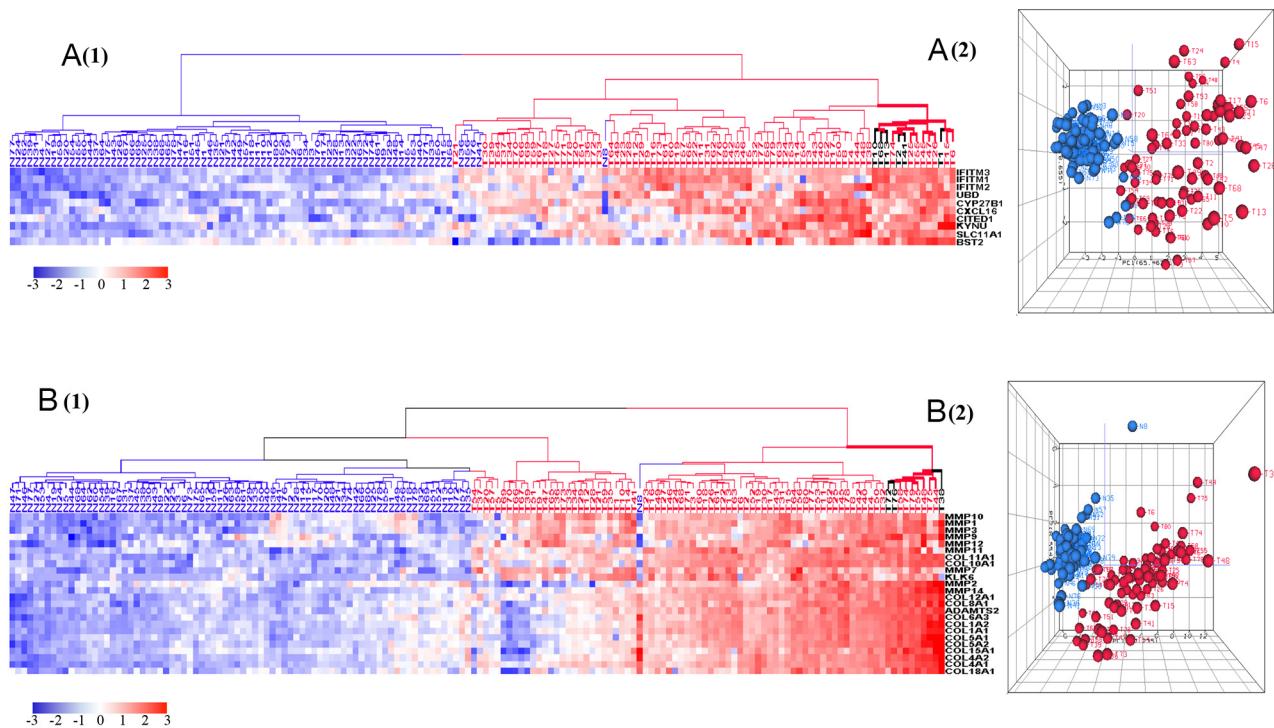


Figure 3: HCA and PCA of genes pertaining to the six most significantly upregulated biological processes in tumors vs. normal tissues. These biological processes include response to interferon-gamma (A), collagen catabolic process (B), chemokines mediated signaling pathways (*Continued*)

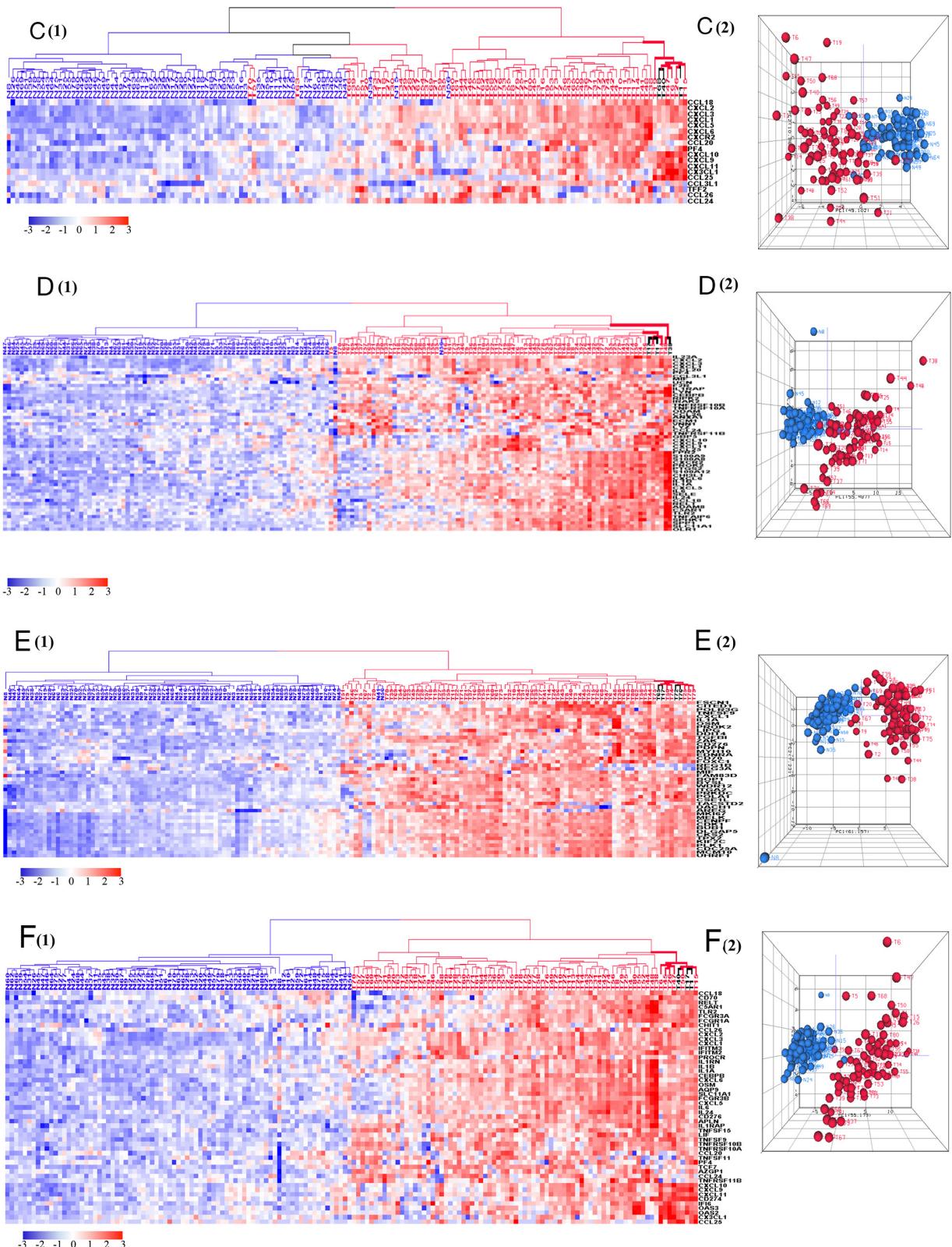


Figure 3 (Continued): HCA and PCA of genes pertaining to the six most significantly upregulated biological processes in tumors vs. normal tissues. (C), inflammatory response (D), cell proliferation (E), immune response (F). Majority of tumors were clustered together within the upregulated side (red) of HCAs. Tumor is marked with red dendrogram and normal sample is marked with blue dendrogram. The bold dendograms in HCAs mark subgroups of tumors with the greatest upregulated genes. Both high stage tumors (labeled in red) and low stage tumors (labeled in black) are presented in bold dendograms. Tumors were more heterogeneous than normal controls in PCA plots.

4). Five isoforms of carbonic anhydrase (CA), involved in the conversion of CO_2 and water into bicarbonate and protons, are among the top downregulated genes (Supplementary Table 2): *CA1* (41 fold decrease), *CA7* (25 fold decrease), *CA2* (15 fold decrease), *CA4* (18 fold decrease) and *CA12* (4 fold decrease). These data suggest

that the acidic environment generated by tumor cells plays an important role in cell proliferation and metastasis [30] but the signaling pathways and genes that lead to such are not known and require investigation. In contrast to the above carbonic anhydrase genes which were sharply downregulated in tumors, carbonic anhydrase 9 (*CA9*),

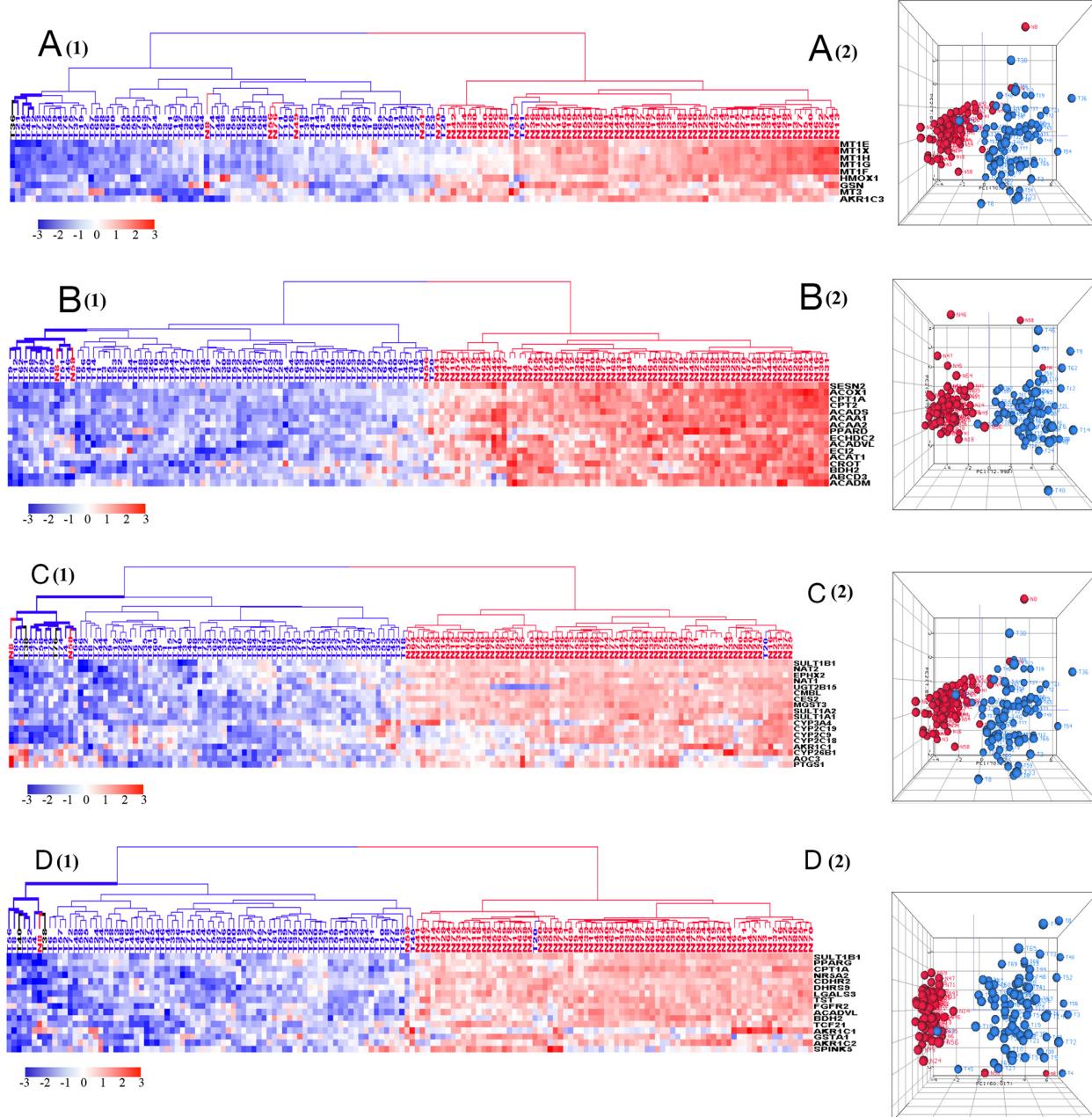


Figure 4: HCA and PCA of genes pertaining to the four significantly downregulated biological processes in tumors vs. normal tissues. These biological processes include cellular response to cadmium ion (A), fatty acid beta-oxidation (B), xenobiotic metabolic process (C), and epithelial cell differentiation (D). Majority of tumors were clustered together within the downregulated side (blue) of HCAs. Tumor is marked with blue dendrogram and normal sample is marked with red dendrogram. The bold dendograms in HCA mark subgroups of tumors with the greatest downregulated genes. Both high stage tumors (labeled in blue) and low stage tumors (labeled in black) are presented in bold dendograms. Normal controls were more heterogeneous than tumors in PCA plots.

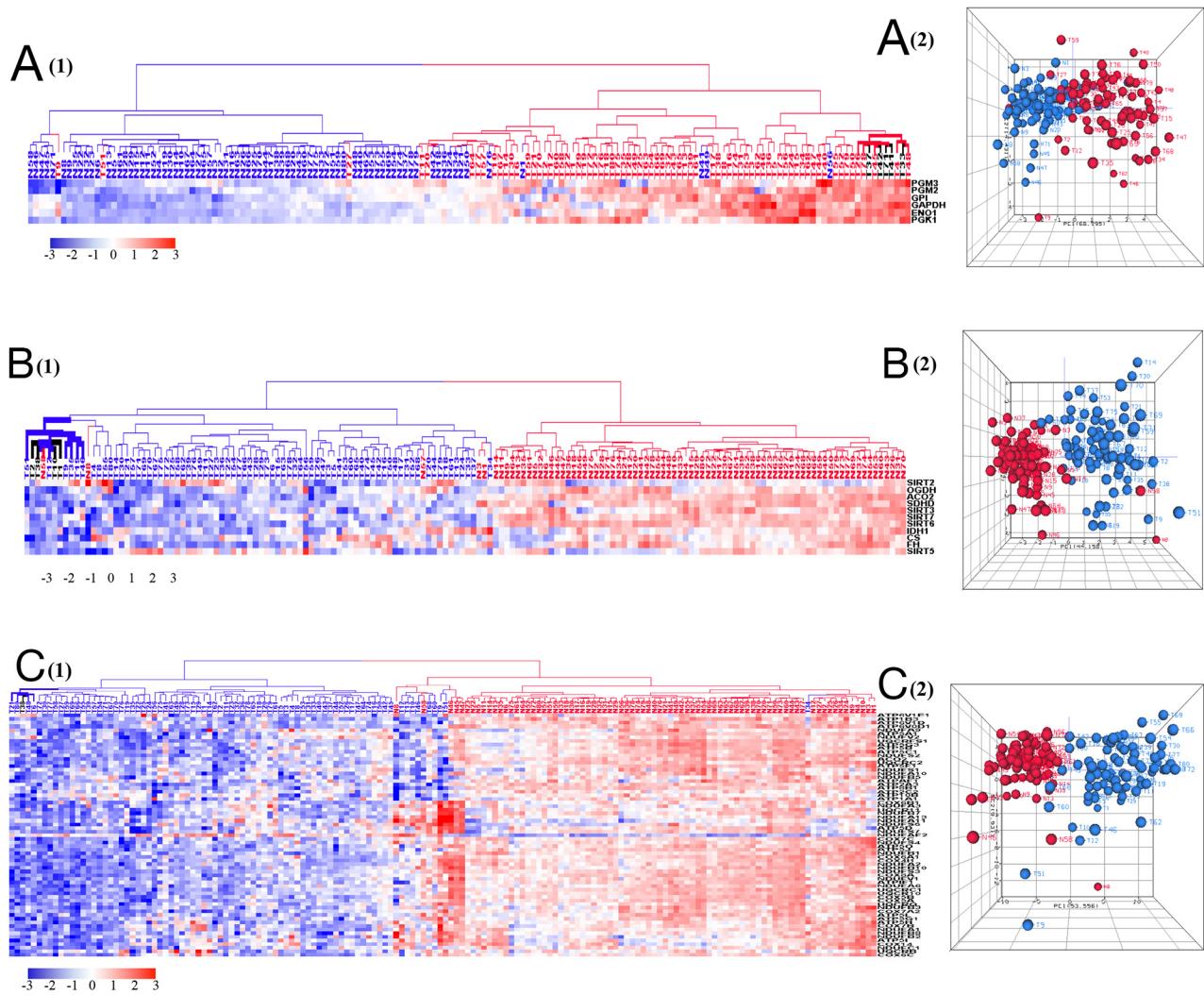
was overexpressed with a 47 fold expression increase in tumor samples. Upregulation of this *C4* gene is associated with poor prognosis [31].

Downregulation of genes pertaining to membrane transporters included bestrophin 4 (*BEST4*, chloride channel) (42 fold decrease) and solute carrier family genes such as *SLC30A10* (28 fold decrease), *SLC6A19* (23 fold decrease), *SLC26A2* (20 fold decrease), and *SLC26A3* (20 fold decrease). Downregulation of previously reported membrane transporters correlated with the loss of normal intestinal epithelial function in tumor cells [32]. Moderately downregulated genes included Acyl-CoA

dehydrogenase (*ACADS*) (5 fold decrease) and Acyl-CoA Dehydrogenase Very Long Chain (*ACADVL*) (2 fold decrease) which are involved in fatty acid beta oxidation.

Switch of ATP production from mitochondria to cytosol in colorectal cancer

The down-regulation of mitochondrial fatty acid beta-oxidation for ATP production prompted us to examine other energy generating pathways in tumor tissues including glycolysis, the tricarboxylic acid (TCA) cycle, glycolysis suppressive Sirtuin (*SIRT*) pathways,



and the mitochondrial respiratory chain. Genes related to glycolysis, TCA/SIRT pathways, and the mitochondrial respiratory chain were identified from the 10,255 gene data set with an average FPKM >1 and FDR ≤ 0.05 in ANOVA. For glycolysis related genes, tumors clustered at the high expression zone (red) while normal samples clustered at the low expression zone (blue) (Figure 5A). The genes involved in the SIRT pathway, TCA cycle, and mitochondrial respiratory chain related functions were sharply downregulated with tumors clustered at the low expression zone in dark blue and normal samples at the high expression zone in red (Figure 5B, 5C). The PCA plots showed similar results with tumors separated from normal samples (Figure 5). The downregulated fatty acid oxidation, TCA/SIRT pathway, and mitochondrial respiratory chain pathway, as well as upregulated glycolysis support the classical Warburg effect [5–7] in colorectal cancer as in other tumor tissues [33–36].

Determination of gene co-expression network and hub genes in colorectal cancer

To explore possible interactions of deregulated biological processes, we generated a comprehensive gene co-expression network from the identified 2,358 deregulated genes. This co-expression network is composed of 1,355 genes with a total of 54,018 strong connections as defined by the threshold value of > 0.8 for the absolute value of the Pearson's correlation coefficient (cc) between the expression levels of two genes (Figure 6). Two upregulated modules included 344 cell proliferation related genes as well as 184 collagen catabolism and inflammation related genes; while 3 downregulated modules included 610 lipid catabolic process related genes, and 164 cell differentiation related genes, as well as 53 genes pertaining to T and B cell related immune

system process including *LY9*, *CR2*, *CXCR5*, *CD19*, *CD79A*, and *CD79B*. The following correlations among modules were observed: first, lipid catabolic process positively correlated ($cc > 0.8$) with cell differentiation and T cell / B cell related immune system process; second, lipid catabolic process negatively correlated ($cc < -0.8$) with cell proliferation; third, cell proliferation strongly positively correlated ($cc > 0.8$) with collagen catabolism and inflammation (Figure 6). Although the increased de novo synthesis of fatty acids has been reported in various cancer cells [37, 38], the association of decreased lipid catabolism, such as fatty acid beta-oxidation, with cell proliferation, dedifferentiation and diminished lymphocyte markers in tumor is a novel finding. Similarly, cell proliferation, collagen catabolism and inflammation have been recognized as hallmarks of cancers [12–16, 39], but the correlated dysregulation with respect to metabolic and differentiation parameters has not previously been reported.

To probe for candidate hub genes we identified 5 up and downregulated genes with the greatest number of co-expressed genes (from 106 to 383 genes) (Table 1). Five upregulated candidate hub genes included methylenetetrahydrofolate dehydrogenase (NADP⁺ dependent) 1-like (*MTHFD1L*), mitotic checkpoint serine/threonine kinase (*BUB1*), thyroid hormone receptor interactor 13 (*TRIP13*), mitotic spindle assembly checkpoint protein (*MAD2L1*), and phosphoribosyl pyrophosphate amidotransferase (*PPAT*). Five downregulated putative hub genes included amyloid beta precursor like protein 2 (*APPL2*), cyclin M4 (*CNNM4*), monoacylglycerol O-acyltransferase 2 (*MOGAT2*), cyclin-dependent kinase inhibitor 2B (*CDKN2B*), and leucine rich repeat containing 19 (*LRRK19*). In the co-expression network, the upregulated hub genes correlated with

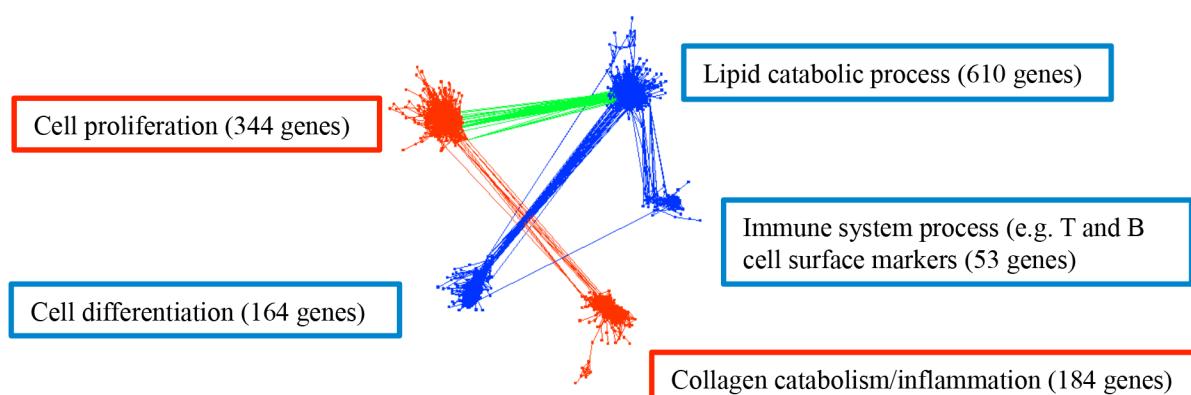


Figure 6: Gene co-expression network in colorectal cancer with 1,355 genes and 54,018 connections. Upregulated biological processes for the corresponding genes are labeled in red boxes and downregulated biological processes are in blue boxes. Red edges are the connections between upregulated processes, blue edges are the connections between downregulated processes, and green edges are the connections between up- and down-regulated processes. Gene number in each module is presented in each corresponding box. The representative biological processes were determined in DAVID with FDR < 0.05.

225 genes related to cell cycle, and the downregulated hub genes correlated with 382 genes related to lipid metabolism (Figure 7). Interestingly, *MTHFD1*, an enzyme involved in the synthesis of tetrahydrofolate

(THF) within the mitochondrion [40], correlated positively with upregulated genes and negatively with downregulated genes, thus acting as a potential link between cell cycle and lipid metabolism.

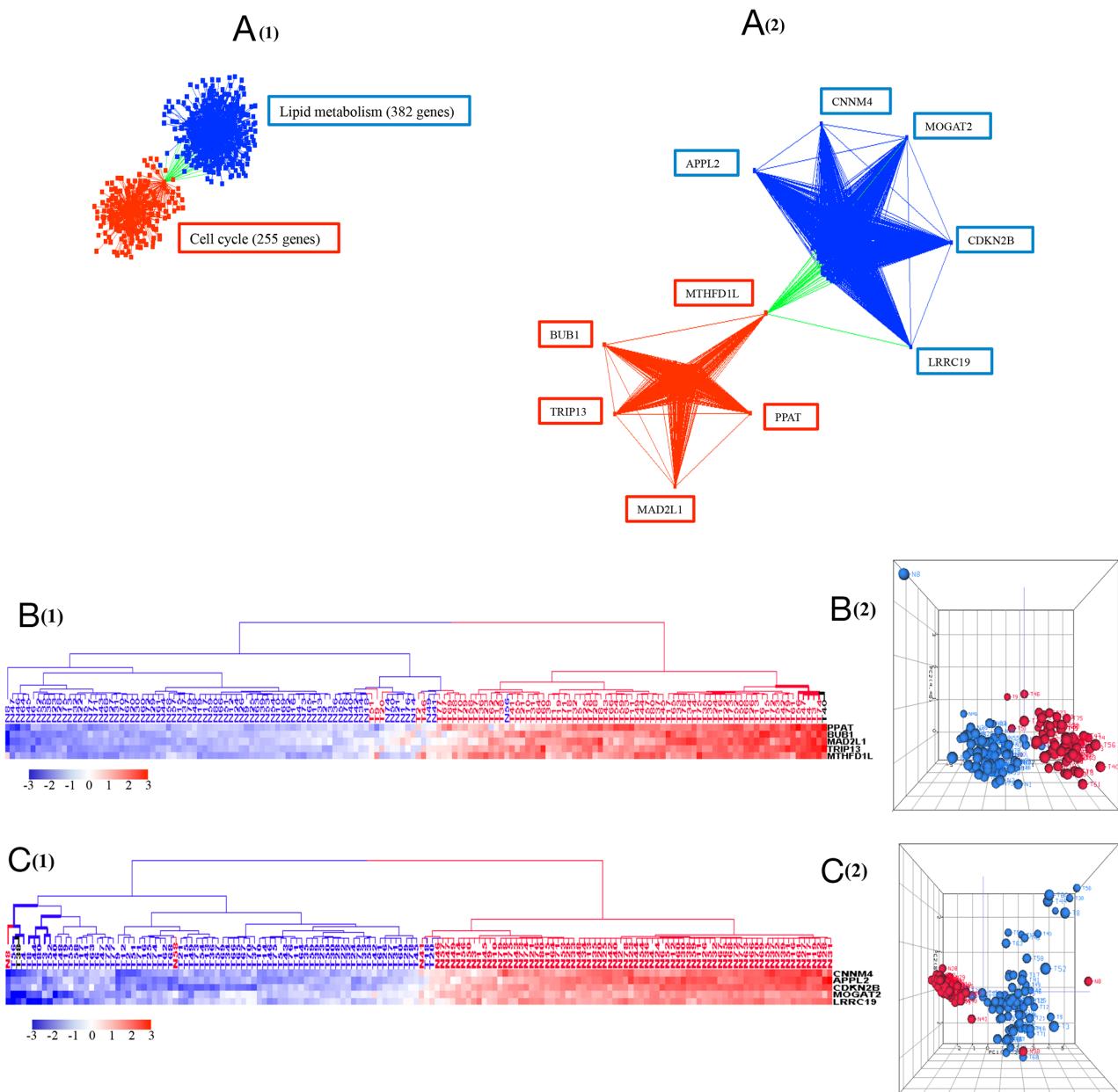


Figure 7: (A) The connections of ten potential hub genes and their correlated 607 genes with 2,080 interactions. The upregulated hub genes are in red squares and the downregulated hub genes are in blue squares. Red edges are the connections between upregulated genes, blue edges are the connections between downregulated genes, and green edges are the connections between up and down regulated genes. The represented biological processes for the corresponding genes, determined in DAVID with FDR < 0.05, are labelled in boxes. (B) and (C) HCA and PCA of the upregulated hub genes and downregulated hub genes. The five upregulated hub genes showed increased expression in 76 out of 79 tumors, and the five downregulated hub genes showed decreased expression in 77 out of 79 tumors when compared with normal samples. (B) Tumor is marked with red dendrogram and normal sample is marked with blue dendrogram. The bold dendograms in HCAs mark subgroups of tumors with the greatest upregulated hub genes. Both high stage tumors (labeled in red) and low stage tumors (labeled in black) are presented in bold dendograms. (C) Tumor is marked with blue dendrogram and normal sample is marked with red dendrogram. The bold dendograms in HCAs mark subgroups of tumors with the greatest downregulated hub genes. Both high stage tumors (labeled in blue) and low stage tumors (labeled in black) are presented in bold dendograms.

Table 1: Five upregulated hub genes and five downregulated hub genes

Upregulated hub genes	Proteins coded	Gene function	Expression fold change (T/N)	Number of correlated genes ($ CC \geq 0.8$)
MTHFD1L	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1-Like	de novo synthesis of purines and thymidylate	5.51	176
BUB1	mitotic checkpoint Serine/ Threonine kinase	activate spindle checkpoint	3.15	150
TRIP13	thyroid hormone receptor interactor 13	inactivate spindle assembly checkpoint	4.09	145
MAD2L1	mitotic arrest deficient 2 like 1	mitotic spindle assembly checkpoint protein	3.08	109
PPAT	phosphoribosyl pyrophosphate amidotransferase	de novo synthesis of purines	3.06	106
Downregulated hub genes	Proteins coded	Gene function	Expression fold change (T/N)	Number of correlated genes ($ CC \geq 0.8$)
APPL2	amyloid beta precursor like protein 2	cell proliferation regulation	0.29	383
CNNM4	cyclin M4	metal transporter	0.33	365
MOGAT2	monoacylglycerol O-Acyltransferase 2	fat digestion and absorption	0.16	361
CDKN2B	cyclin-dependent kinase inhibitor 2B	inhibit cell cycle G1 progression	0.14	357
LRRC19	leucine rich repeat containing 19	Membrane protein	0.17	357

DISCUSSION

In this study, we used high-throughput mRNA-Seq technology to examine the whole transcriptome in a total of 79 colon cancers and matched normal tissue controls. Our study provides a comprehensive gene expression analysis in colon cancer tissues. The upregulated response to IFN γ , inflammation response, immune and chemokine mediated signaling pathways (Figure 2, 3) demonstrate that inflammation is a major feature of the tumor microenvironment in colorectal cancer. Interferon induced transmembrane proteins (*IFITMs*), can be transcriptionally induced by type I and type II interferons [41]. We observed that tumor samples had upregulated *IFITMs* including *IFITM1*, *IFITM2*, and *IFITM3* (Figure 3), consistent with previous observations in colorectal cancer [42–44]. Although it has been used to treat tumors due to its immunostimulatory function, paradoxically IFN γ , has recently been found to promote tumor progression by upregulating checkpoint inhibitors [45, 46]. This tumor-promoting function for IFN γ is supported by the upregulated IFN γ (Supplementary Figure 2) observed in

our study. Such activity, however, may lead to favorable responses to checkpoint inhibition in this setting [47, 48].

Five of 9 metallothionein (*MT*) genes were downregulated in tumors (Figure 4A). The MT proteins are capable of binding heavy metal (cadmium, mercury, lead, and arsenic) and their diminished expression is associated with poorer survival in colorectal cancer [49]. We also found that 5 of 22 cytochrome P450 (*CYP*) genes, whose associated enzymes oxidize small foreign organicmolecules such as toxins or drugs [50], were downregulated in tumors (Figure 4C). These findings suggest that colorectal cancer cells lose key physiological functions of detoxification which could be the result of epithelial cell dedifferentiation (Figure 2, 4) [51]. These findings are in keeping with observations that colonic dedifferentiation may initiate intestinal tumorigenesis [52].

The identification of highly connected potential hub genes from the gene co-expression network should facilitate the identification of candidate biomarkers and therapeutic targets for colorectal cancer. Among 10 putative hub genes (Table 1, Figure 7), all 5 upregulated hub genes have been identified as oncogenes [40, 53–59],

Table 2A: Discrepancy between histological diagnosis and molecular profiles

Dysregulated biological processes	A few tumors clustered with most normal samples		A few normal samples clustered with most tumors	
	High stage tumors	Low stage tumors	Normal samples adjacent to high stage tumors	Normal samples adjacent to low stage tumors
upregulated response to interferon-gamma (10 genes)	0	0	N8	N36, N41, N66, N76
upregulated collagen catabolic process (24 genes)	0	0	N8	0
upregulated chemokines mediated signaling pathways (18 genes)	T57, T63	T70	N15, N34, N56	0
upregulated inflammatory response (56 genes)	0	0	N8	0
upregulated cell proliferation (50 genes)	0	0	N56	N41
upregulated immune response (51 genes)	0	0	0	0
upregulated glycolysis (6 genes)	T9, T51, T27	0	N56, N49, N50, N21, N34, N2	N1, N18, N36, N41, N42, N76
upregulated hub genes (5 genes)	T9, T46, T51	T2, T20	N56	0
downregulated cellular response to cadmium (9 genes)	T51	T20, T31	N8, N49, N56	N41, N76
downregulated acid beta-oxidation (16 genes)	0	0	N8, N56, N58	0
downregulated xenobiotic metabolic process (18 genes)	0	T20	N8, N58	0
downregulated epithelial cell differentiation (16 genes)	0	T20	N8, N58	0
downregulated SIRT and TCA cycle (11 genes)	0	0	N3, N8, N57, N58	N1
downregulated respiratory chain (67 genes)	T34	0	N8, N58	0
downregulated hub genes (5 genes)	0	0	N8, N58	N41

and all 5 downregulated hub genes have been identified as tumor suppressor genes [60–64]. *MTHFD1L* is the only hub gene which has both positive and negative correlations with genes pertaining to cell cycle and lipid metabolism respectively (Figure 7A). Specifically, the *MTHFD1L* enzyme is involved in the synthesis of tetrahydrofolate (THF) within the mitochondrion [40], in contrast to cytosolic methylene tetrahydrofolate reductase (*MTHFR*), a drug target of methotrexate (MTX) [65]. Inhibition of mitochondrial *MTHFD1L* expression specifically in tumors may thus provide a new therapeutic target. Overexpression of *MAD2L1* and *BUB1*, candidate hub genes in our study, has previously been associated with high grade breast tumors and poor survival of breast cancer patients [66, 67]. *TRIP13*, critical for the inactivation of the spindle assembly checkpoint, is associated with the progression of certain cancers, and was found to be highly expressed in multiple colorectal cancer tissues as assessed by RT-PCR [58]. *PPAT* gene was identified as a prognostic biomarker in aggressive lung adenocarcinoma

by RNA-Seq analysis [68]. These overexpressed potential hub genes may have relevance as diagnostic or prognostic biomarkers for colorectal cancer. Of note, the diminished expression of putative hub genes such as *CDKN2B*, *CNNM4*, *LRRC19*, and *MOGAT2* might be the result of inactivation of genes through DNA methylation as their methylation was identified in colorectal, gastric, and breast cancers as well as in leukaemia [62, 69–71]. The DNA demethylation agent arsenic trioxide was able to restore the expression of methylation-silenced *CDKN2B* in the human T lymphoblast cell line Molt4 and human myelodysplastic syndrome-(MDS-REBT) cell line Mutz-1 [72], indicating the potential of DNA demethylation agents to contribute to cancer treatment in some instances.

Not to be overlooked, in our analyses we observed that a few normal controls clustered with high stage and low stage tumors while multiple low stage tumors clustered with high stage tumors in certain biological pathways (Figure 3, 4, 5, 7, Table 2A). These discrepancies between histological determination and molecular profiles

Table 2B: Low stage tumors and normal samples show similar gene expression profiles with high stage tumors.

Dysregulated biological processes	Tumors and normal samples clustered within high expression zones		
	High stage tumors	Low stage tumors	Normal samples adjacent to high stage tumors
upregulated response to interferon-gamma	T6, T14, T16, T25, T26, T46, T47, T55	T1, T2, T10, T13, T41, T68	0
upregulated collagen catabolic process	T44, T48, T55, T56, T74, T75, T79	T38, T72, T76	0
upregulated chemokines mediated signaling pathways	T6, T15, T26, T47	T1, T40, T68	0
upregulated inflammatory response	T14, T15, T48	T1, T17, T38	0
upregulated cell proliferation	T25, T74, T50, T78, T79	T17, T65, T72, T76	0
upregulated immune response	T6, T15, T26, T45, T47, T50	T1, T17, T40	0
upregulated glycolysis	T4, T26, T44, T45, T48, T75	T33, T37, T41, T42, T71	0
upregulated hub genes	T3, T47, T56, T74	T40	0
Tumors and normal samples clustered within low expression zones			
Dysregulated biological processes	High stage tumors	Low stage tumors	Normal samples adjacent to high stage tumors
downregulated acid beta-oxidation	T9, T12, T21, T46, T51, T62, T57, T58, T59, T77, T80	T36	N8, N58
downregulated xenobiotic metabolic process	T4, T5, T35, T8, T14, T49, T52, T56, T58, T74, T75, T79, T80	T38, T76	N8, N58
downregulated epithelial cell differentiation	T3, T4, T8, T52, T56	T38, T40	N8
downregulated SIRT and TCA cycle	T9, T12, T19, T35, T48, T51	T1, T2, T10, T38	N8, N58
downregulated respiratory chain	T6, T21, T48, T80	T38	0
downregulated hub genes	T3, T4, T8, T52, T56, T80	T38	N8

suggest the necessity of molecular subtyping of colorectal cancers pertaining to numerous attributes correlated with tumor invasiveness and metastasis [73, 74]. Moreover, the so-called normal tissues adjacent to either high or low stage tumors in the heatmaps (Table 2B) showed molecular profiles of malignancy, indicating the presence of either histologically undetected tumor budding or altered metabolism imposed by proximal tumors. As the tumor margin determines “surgical clearance” which, in turn, impacts the potential for both local recurrence and overall survival [75-77], a molecular evaluation of tumor adjacent tissues rather than sole reliance on a histological determination may better determine true negative margins.

In summary, in applying mRNA-Seq to analyze transcriptome profiles in 79 primary colorectal cancers which were stringently collected, stored and processed,

we showed that colorectal cancer cells/tissues exhibit malignant features such as cell proliferation, tissue remodeling and cytokine related inflammation and lose the physiological functions of normal colorectal tissue such as metabolism of toxins. Our comprehensive co-expression network provides a global view of interacting biological processes in colorectal cancer which should facilitate discovery of tumor specific targets for the treatment of this deadly disease.

MATERIALS AND METHODS

Cohort

Seventy-nine paired-tissues (79 tumor and 79 normal controls, Supplementary Table 1) of pretreatment

colorectal cancers were collected from 38 male and 41 female patients by Indivumed GmbH (Germany) for mRNA sequencing. To evaluate tumor content, hematoxylin and eosin stained microscopic slices were examined by pathologists to determine the tumor cell and normal cell areas, respectively. Histologically, tumor content is 50-70% in tumors and 0% in normal tissues. Normal tissues were collected from a site at a minimum of 5 cm from the tumor margin. Ischemia time to freeze was 6-11 minutes. The short cold ischemia increases the likelihood that the data do not relate to postsurgical tissue processing artifacts [8]. According to the medical pathology report, tumors were classified as well, moderately, and poorly differentiated tumors following international guideline UICC TNM-classification [9]. For the convenience of analysis, 26 tumors with stage 1 and 2 were considered as low stage tumors, while 53 tumors with stage 3 and 4 as well as lymph node (LN) and lymphatic vessel (LV) positive were considered as high stage tumors. The ratio of high stage tumors vs. low stage tumors is 2 to 1. There were 17 well (low grade) differentiated, 36 moderate (medium grade) differentiated, and 26 poorly (high grade) differentiated tumors. Clinical and histopathological characteristics of the patients as well as tumor location are summarized in Supplementary Table 1. Among these 80 tumor pairs, 79 pairs were sequenced except the T7/N7 pair.

mRNA sequencing

RNA quality was assessed using the Agilent 2100 Bioanalyzer, with cellular RNA analyzed using the RNA 6000 Nano Kit (Agilent). Samples with a RNA Integrity Number (RIN) of 7 or higher were processed to generate libraries for mRNA sequencing following the Illumina® TruSeq Stranded mRNA Sample Preparation Guide. In this method, poly (A) RNA was purified from 0.5 µg total RNA, fragmented and reverse-transcribed into cDNAs. Double strand cDNAs were adenylated at the 3' ends and ligated to indexed sequencing adaptors, followed with briefly amplification for 15 cycles. One femtomole of the sequencing libraries (median size ~260nt) were denatured and loaded onto a flow cell for cluster generation using the Illumina cBot. Every six samples were loaded onto each lane of a rapid run flow cell. Paired-end sequencing was carried out on HiSeq 2500 sequencer (Illumina, San Diego, CA, USA) for 100 x 2 cycles. For each sample, we obtained ~50 million 100-bp reads that passed preset filtering parameters.

Sequencing data analysis

For mRNA sequencing, Tophat V.2.0.11 and Cufflinks V.2.2.1 were used to align the short reads in fastq files to the RefSeq UCSC human hg19 transcript reference

genome annotation database and the quantification of relative abundance of each transcript was reported as fragments per kilobase of transcript per million mapped reads (FPKM). ANOVA test was conducted (on Partek genomics suite) to identify mRNAs with differential expression between tumors and matched normal adjacent tissues using the threshold False Discovery Rate (FDR) ≤ 0.05 . Genes with at least 2 fold change (increase or decrease) of expression levels as well as FDR ≤ 0.05 in ANOVA test were selected as significantly dysregulated genes between tumor and normal tissues. These genes were used for correlation analysis by Partek. Genes with Pearson's correlation coefficient (cc) > 0.8 or < -0.8 were used for network construction in Cytoscape (2.8.2). The genes with the greatest number of correlated genes were defined as potential hub genes. The unsupervised hierarchical clustering analysis (HCA) and principal component analysis (PCA) were used to explore the gene expression profiles in ArrayTrack (The National Center for Toxicological Research, U.S. Food and Drug Administration).

Colorectal cancer gene expression landscape

A total of 25,761 genes were detected and 17,055 genes were differentially expressed (FDR ≤ 0.05 in ANOVA). There were 9,211 genes (35% of total genes) with average FPKM (across 158 samples) < 1 and 16,550 genes (65% of total genes) with average FPKM > 1 . Since relative higher FPKM values generally confer more reliable quantitation of the genes in samples [10], our data analysis focused on 10,255 genes (40% of total genes) with average FPKM > 1 and differential expression between tumor and normal controls (FDR ≤ 0.05 in ANOVA).

Gene ontology and KEGG pathway analysis

Gene ontology (GO) analysis was performed on genes with differential expression between tumor and normal samples using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) v6.8 (<https://david.ncifcrf.gov/>), NIAID/NIH. False Discovery Rate (FDR) < 0.05 was used as the criteria for GO category enrichment.

Ethics statement

Written informed consent was obtained from all participants involved in the study via the procuring laboratory at Indivumed (Hamburg, Germany).

Author contributions

RW, WW, YQZ and RFS carried out experiments. RW and LX performed data analysis. AR, LX, RW, JZ, HJ and LP designed experiments and interpreted results. AR, RW and LX wrote the manuscript, and all authors edited it.

ACKNOWLEDGMENTS

The authors would like to thank Drs. Daniela I. Verthelyi, Ashutosh V. Rao, Jennifer S. Dickey, and Dickran G. Kazandjian for their critical review and comments on this manuscript. The authors would like to thank Dr. Leihong Wu for the help in the construction of gene co-expression network.

CONFLICTS OF INTEREST

The authors have declared that no conflicts of interest exists.

FINANCIAL SUPPORT

This work was supported by FDA intramural program funds (to Amy Rosenberg).

REFERENCES

1. Brenner H, Kloor M, Pox CP. Colorectal cancer. *Lancet*. 2014; 383:1490–502.
2. Boroughs LK, DeBerardinis RJ. Metabolic pathways promoting cancer cell survival and growth. *Nat Cell Biol*. 2015; 17:351–59.
3. Feichtinger RG, Zimmermann F, Mayr JA, Neureiter D, Hauser-Kronberger C, Schilling FH, Jones N, Sperl W, Kofler B. Low aerobic mitochondrial energy metabolism in poorly- or undifferentiated neuroblastoma. *BMC Cancer*. 2010; 10:149.
4. Chaika NV, Yu F, Purohit V, Mehla K, Lazenby AJ, DiMaio D, Anderson JM, Yeh JJ, Johnson KR, Hollingsworth MA, Singh PK. Differential expression of metabolic genes in tumor and stromal components of primary and metastatic loci in pancreatic adenocarcinoma. *PLoS One*. 2012; 7:e32996.
5. Turner JS, Brittain EG. Oxygen as a factor in photosynthesis. *Biol Rev Camb Philos Soc*. 1962; 37:130–70.
6. Zelitch I. Photosynthesis, Photorespiration, and Plant Productivity. New York: Academic Press. 1971: 253–255.
7. Gatenby RA, Gillies RJ. Why do cancers have high aerobic glycolysis? *Nat Rev Cancer*. 2004; 4:891–99.
8. Unger FT, Lange N, Krüger J, Compton C, Moore H, Agrawal L, Juhl H, David KA. Nanoproteomic analysis of ischemia-dependent changes in signaling protein phosphorylation in colorectal normal and cancer tissue. *J Transl Med*. 2016; 14:6.
9. Puppa G, Sonzogni A, Colombari R, Pelosi G. TNM staging system of colorectal carcinoma: a critical appraisal of challenging issues. *Arch Pathol Lab Med*. 2010; 134:837–52.
10. Toung JM, Morley M, Li M, Cheung VG. RNA-sequence analysis of human B-cells. *Genome Res*. 2011; 21:991–98.
11. Dong Y, Loessner D, Irving-Rodgers H, Obermair A, Nicklin JL, Clements JA. Metastasis of ovarian cancer is mediated by kallikrein related peptidases. *Clin Exp Metastasis*. 2014; 31:135–47.
12. Griffith JW, Sokol CL, Luster AD. Chemokines and chemokine receptors: positioning cells for host defense and immunity. *Annu Rev Immunol*. 2014; 32:659–702.
13. Lu P, Takai K, Weaver VM, Werb Z. Extracellular matrix degradation and remodeling in development and disease. *Cold Spring Harb Perspect Biol*. 2011; 3:a005058.
14. Kessenbrock K, Plaks V, Werb Z. Matrix metalloproteinases: regulators of the tumor microenvironment. *Cell*. 2010; 141:52–67. <https://doi.org/10.1016/j.cell.2010.03.015>.
15. Tamir A, Jag U, Sarojini S, Schindewolf C, Tanaka T, Gharbaran R, Patel H, Sood A, Hu W, Patwa R, Blake P, Chirina P, Oh Jeong J, et al. Kallikrein family proteases KLK6 and KLK7 are potential early detection and diagnostic biomarkers for serous and papillary serous ovarian cancer subtypes. *J Ovarian Res*. 2014; 7:109.
16. Johnson-Holiday C, Singh R, Johnson E, Singh S, Stockard CR, Grizzle WE, Lillard JW Jr. CCL25 mediates migration, invasion and matrix metalloproteinase expression by breast cancer cells in a CCR9-dependent fashion. *Int J Oncol*. 2011; 38:1279–85.
17. Rowland KJ, Diaz-Miron J, Guo J, Erwin CR, Mei J, Worthen GS, Warner BW. CXCL5 is required for angiogenesis, but not structural adaptation after small bowel resection. *J Pediatr Surg*. 2014; 49:976–80.
18. Esquivel-Velázquez M, Ostoa-Saloma P, Palacios-Arreola MI, Nava-Castro KE, Castro JI, Morales-Montor J. The role of cytokines in breast cancer development and progression. *J Interferon Cytokine Res*. 2015; 35:1–16.
19. Kinoshita T, Muramatsu R, Fujita T, Nagumo H, Sakurai T, Noji S, Takahata E, Yaguchi T, Tsukamoto N, Kudo-Saito C, Hayashi Y, Kamiyama I, Ohtsuka T, et al. Prognostic value of tumor-infiltrating lymphocytes differs depending on histological type and smoking habit in completely resected non-small-cell lung cancer. *Ann Oncol*. 2016; 27:2117–23.
20. Dembic Z. Chemotactic Factors—Advances in Research and Application: 2013 Edition. Scholarly Editions. 2013; 29.
21. Choy JC, Yi T, Rao DA, Tellides G, Fox-Talbot K, Baldwin WM 3rd, Pober JS. CXCL12 induction of inducible nitric oxide synthase in human CD8 T cells. *J Heart Lung Transplant*. 2008; 27:1333–39.

22. Debes GF, Diehl MC. CCL8 and skin T cells—an allergic attraction. *Nat Immunol.* 2011; 12:111–12.
23. Huber AK, Irani DN. Targeting CXCL13 During Neuroinflammation. *Adv Neuroimmune Biol.* 2015; 6:1–8.
24. Liu X, Wang J, Wang H, Yin G, Liu Y, Lei X, Xiang M. REG3A accelerates pancreatic cancer cell growth under IL-6-associated inflammatory condition: involvement of a REG3A-JAK2/STAT3 positive feedback loop. *Cancer Lett.* 2015; 362:45–60.
25. van Beelen Granlund A, Østvik AE, Brenna Ø, Torp SH, Gustafsson BI, Sandvik AK. REG gene expression in inflamed and healthy colon mucosa explored by in situ hybridisation. *Cell Tissue Res.* 2013; 352:639–46.
26. Astrosini C, Roeefzaad C, Dai YY, Dieckgraefe BK, Jöns T, Kemmner W. REG1A expression is a prognostic marker in colorectal cancer and associated with peritoneal carcinomatosis. *Int J Cancer.* 2008; 123:409–13.
27. Trivedi PJ, Bruns T, Ward S, Mai M, Schmidt C, Hirschfield GM, Weston CJ, Adams DH. Intestinal CCL25 expression is increased in colitis and correlates with inflammatory activity. *J Autoimmun.* 2016; 68:98–104.
28. Itatani Y, Kawada K, Inamoto S, Yamamoto T, Ogawa R, Taketo MM, Sakai Y. The Role of Chemokines in Promoting Colorectal Cancer Invasion/Metastasis. *Int J Mol Sci.* 2016; 17:E643.
29. Brien GL, Gambero G, O'Connell DJ, Jerman E, Turner SA, Egan CM, Dunne EJ, Jurgens MC, Wynne K, Piao L, Lohan AJ, Ferguson N, Shi X, et al. Polycomb PHF19 binds H3K36me3 and recruits PRC2 and demethylase NO66 to embryonic stem cell genes during differentiation. *Nat Struct Mol Biol.* 2012; 19:1273–81.
30. Kato Y, Ozawa S, Miyamoto C, Maehata Y, Suzuki A, Maeda T, Baba Y. Acidic extracellular microenvironment and cancer. *Cancer Cell Int.* 2013; 13:89.
31. Kalavská K, Cierna Z, Chovanec M, Takacova M, Svetlovska D, Miskovska V, Obertova J, Palacka P, Rajec J, Sycova-Mila Z, Machalekova K, Kajo K, Spanik S, et al. Prognostic value of intratumoral carbonic anhydrase IX expression in testicular germ cell tumors. *Oncol Lett.* 2017; 13:2177–85.
32. Bhutia YD, Babu E, Ramachandran S, Yang S, Thangaraju M, Ganapathy V. SLC transporters as a novel class of tumour suppressors: identity, function and molecular mechanisms. *Biochem J.* 2016; 473:1113–24.
33. Vander Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science.* 2009; 324:1029–33.
34. Li WJ, Zhao YF. Warburg Effect and Mitochondrial Metabolism in Skin Cancer. *J Carcinogene Mutagene.* 2012; S4:002. <https://doi.org/10.4172/2157-2518.S4-002>.
35. Shuch B, Linehan WM, Srinivasan R. Aerobic glycolysis: a novel target in kidney cancer. *Expert Rev Anticancer Ther.* 2013; 13:711–19.
36. Pate KT, Stringari C, Sprowl-Tanio S, Wang K, TeSlaa T, Hoverter NP, McQuade MM, Garner C, Digman MA, Teitell MA, Edwards RA, Gratton E, Waterman ML. Wnt signaling directs a metabolic program of glycolysis and angiogenesis in colon cancer. *EMBO J.* 2014; 33:1454–73.
37. Menendez JA, Lupu R. Fatty acid synthase and the lipogenic phenotype in cancer pathogenesis. *Nat Rev Cancer.* 2007; 7:763–77.
38. Sánchez-Martínez R, Cruz-Gil S, Gómez de Cedrón M, Álvarez-Fernández M, Vargas T, Molina S, García B, Herranz J, Moreno-Rubio J, Reglero G, Pérez-Moreno M, Feliu J, Malumbres M, Ramírez de Molina A. A link between lipid metabolism and epithelial-mesenchymal transition provides a target for colon cancer therapy. *Oncotarget.* 2015; 6:38719–36. <https://doi.org/10.18632/oncotarget.5340>.
39. Coussens LM, Werb Z. Inflammation and cancer. *Nature.* 2002; 420:860–67.
40. Pike ST, Rajendra R, Artzt K, Appling DR. Mitochondrial C1-tetrahydrofolate synthase (MTHFD1L) supports the flow of mitochondrial one-carbon units into the methyl cycle in embryos. *J Biol Chem.* 2010; 285:4612–20.
41. Lewin AR, Reid LE, McMahon M, Stark GR, Kerr IM. Molecular analysis of a human interferon-inducible gene family. *Eur J Biochem.* 1991; 199:417–23.
42. Andreu P, Colnot S, Godard C, Laurent-Puig P, Lamarque D, Kahn A, Perret C, Romagnolo B. Identification of the IFITM family as a new molecular marker in human colorectal tumors. *Cancer Res.* 2006; 66:1949–55.
43. Diamond MS, Farzan M. The broad-spectrum antiviral functions of IFIT and IFITM proteins. *Nat Rev Immunol.* 2013; 13:46–57.
44. Platanias LC. Mechanisms of type-I- and type-II-interferon-mediated signalling. *Nat Rev Immunol.* 2005; 5:375–86.
45. Abiko K, Matsumura N, Hamanishi J, Horikawa N, Murakami R, Yamaguchi K, Yoshioka Y, Baba T, Konishi I, Mandai M. IFN- γ from lymphocytes induces PD-L1 expression and promotes progression of ovarian cancer. *Br J Cancer.* 2015; 112:1501–09.
46. Mandai M, Hamanishi J, Abiko K, Matsumura N, Baba T, Konishi I. Dual Faces of IFN γ in Cancer Progression: A Role of PD-L1 Induction in the Determination of Pro- and Antitumor Immunity. *Clin Cancer Res.* 2016; 22:2329–34.
47. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer.* 2012; 12:252–64.
48. Buferne M, Chasson L, Grange M, Mas A, Arnoux F, Bertuzzi M, Naquet P, Leserman L, Schmitt-Verhulst AM, Auphan-Anezin N. IFN γ producing CD8(+) T cells modified to resist major immune checkpoints induce regression of MHC class I-deficient melanomas. *OncolImmunology.* 2015; 4:e974959.
49. Arriaga JM, Levy EM, Bravo AI, Bayo SM, Amat M, Aris M, Hannois A, Bruno L, Roberti MP, Loria FS, Pairola A, Huertas E, Mordoh J, Bianchini M. Metallothionein

- expression in colorectal cancer: relevance of different isoforms for tumor progression and patient survival. *Hum Pathol.* 2012; 43:197–208.
50. Crago J, Klaper RD. Influence of gender, feeding regimen, and exposure duration on gene expression associated with xenobiotic metabolism in fathead minnows (*Pimephales promelas*). *Comp Biochem Physiol C Toxicol Pharmacol.* 2011; 154:208–12.
51. Mimoun S, Andriamihaja M, Chaumontet C, Atanasiu C, Benamouzig R, Blouin JM, Tomé D, Bouillaud F, Blachier F. Detoxification of H(2)S by differentiated colonic epithelial cells: implication of the sulfide oxidizing unit and of the cell respiratory capacity. *Antioxid Redox Signal.* 2012; 17:1–10.
52. Schwitalla S, Fingerle AA, Cammareri P, Nebelsiek T, Göktuna SI, Ziegler PK, Canli O, Heijmans J, Huels DJ, Moreaux G, Rupec RA, Gerhard M, Schmid R, et al. Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties. *Cell.* 2013; 152:25–38.
53. Lehman NL. Future potential of thymidylate synthase inhibitors in cancer therapy. *Expert Opin Investig Drugs.* 2002; 11:1775–87.
54. Grady WM. Genomic instability and colon cancer. *Cancer Metastasis Rev.* 2004; 23:11–27.
55. Ricke RM, Jeganathan KB, van Deursen JM. Bub1 overexpression induces aneuploidy and tumor formation through Aurora B kinase hyperactivation. *J Cell Biol.* 2011; 193:1049–64.
56. Sotillo R, Hernando E, Diaz-Rodríguez E, Teruya-Feldstein J, Cordón-Cardo C, Lowe SW, Benezra R. Mad2 overexpression promotes aneuploidy and tumorigenesis in mice. *Cancer Cell.* 2007; 11:9–23.
57. Jo YS, Oh HR, Kim MS, Yoo NJ, Lee SH. Frameshift mutations of OGDH, PPAT and PCCA genes in gastric and colorectal cancers. *Neoplasma.* 2016; 63:681–86.
58. Kurita K, Maeda M, Mansour MA, Kokuryo T, Uehara K, Yokoyama Y, Nagino M, Hamaguchi M, Senga T. TRIP13 is expressed in colorectal cancer and promotes cancer cell invasion. *Oncol Lett.* 2016; 12:5240–46.
59. Zhong R, Chen X, Chen X, Zhu B, Lou J, Li J, Shen N, Yang Y, Gong Y, Zhu Y, Yuan J, Xia X, Miao X. MAD1L1 Arg558His and MAD2L1 Leu84Met interaction with smoking increase the risk of colorectal cancer. *Sci Rep.* 2015; 5:12202.
60. Suzuki H, Zhou X, Yin J, Lei J, Jiang HY, Suzuki Y, Chan T, Hannon GJ, Mergner WJ, Abraham JM, Meltzer SJ. Intragenic mutations of CDKN2B and CDKN2A in primary human esophageal cancers. *Hum Mol Genet.* 1995; 4:1883–87.
61. Pyrzynska B, Banach-Orlowska M, Teperek-Tkacz M, Miekus K, Drabik G, Majka M, Miaczynska M. Multifunctional protein APPL2 contributes to survival of human glioma cells. *Mol Oncol.* 2013; 7:67–84.
62. Funato Y, Yamazaki D, Mizukami S, Du L, Kikuchi K, Miki H. Membrane protein CNNM4-dependent Mg²⁺ efflux suppresses tumor progression. *J Clin Invest.* 2014; 124:5398–410.
63. Kim JH, Kang S, Kim TW, Yin L, Liu R, Kim SJ. Expression profiling after induction of demethylation in MCF-7 breast cancer cells identifies involvement of TNF-α mediated cancer pathways. *Mol Cells.* 2012; 33:127–33.
64. Narimatsu T, Matsuura K, Nakada C, Tsukamoto Y, Hijiyama N, Kai T, Inoue T, Uchida T, Nomura T, Sato F, Seto M, Takeuchi I, Mimata H, Moriyama M. Downregulation of NDUFB6 due to 9p24.1-p13.3 loss is implicated in metastatic clear cell renal cell carcinoma. *Cancer Med.* 2015; 4:112–24.
65. Stankova J, Lawrence AK, Rozen R. Methylenetetrahydrofolate reductase (MTHFR): a novel target for cancer therapy. *Curr Pharm Des.* 2008; 14:1143–50.
66. Han JY, Han YK, Park GY, Kim SD, Kim JS, Jo WS, Lee CG. Bub1 is required for maintaining cancer stem cells in breast cancer cell lines. *Sci Rep.* 2015; 5:15993.
67. Wang Z, Katsaros D, Shen Y, Fu Y, Canuto EM, Benedetto C, Lu L, Chu WM, Risch HA, Yu H. Biological and Clinical Significance of MAD2L1 and BUB1, Genes Frequently Appearing in Expression Signatures for Breast Cancer Prognosis. *PLoS One.* 2015; 10:e0136246.
68. Goswami MT, Chen G, Chakravarthi BV, Pathi SS, Anand SK, Carskadon SL, Giordano TJ, Chinnaian AM, Thomas DG, Palanisamy N, Beer DG, Varambally S. Role and regulation of coordinately expressed de novo purine biosynthetic enzymes PPAT and PAICS in lung cancer. *Oncotarget.* 2015; 6:23445–61. <https://doi.org/10.18632/oncotarget.4352>.
69. Spisák S, Kalmár A, Galamb O, Wichmann B, Sipos F, Péterfia B, Csabai I, Kovácszky I, Semsey S, Tulassay Z, Molnár B. Genome-wide screening of genes regulated by DNA methylation in colon cancer development. *PLoS One.* 2012; 7:e46215.
70. Lima PS, Molffeta GA, Araujo AG, Zago MA, Silva WA Jr. DNA methylation analysis of the tumor suppressor gene CDKN2B in Brazilian leukemia patients. *Genet Mol Biol.* 2008; 31:632–38.
71. Van der Auwera I, Yu W, Suo L, Van Neste L, van Dam P, Van Marck EA, Pauwels P, Vermeulen PB, Dirix LY, Van Laere SJ. Array-based DNA methylation profiling for breast cancer subtype discrimination. *PLoS One.* 2010; 5:e12616.
72. Fu HY, Shen JZ, Wu Y, Shen SF, Zhou HR, Fan LP. Arsenic trioxide inhibits DNA methyltransferase and restores expression of methylation-silenced CDKN2B/CDKN2A genes in human hematologic malignant cells. *Oncol Rep.* 2010; 24:335–43.
73. Vargas AJ, Harris CC. Biomarker development in the precision medicine era: lung cancer as a case study. *Nat Rev Cancer.* 2016; 16:525–37.

74. Achyut BR, Bader DA, Robles AI, Wangsa D, Harris CC, Ried T, Yang L. Inflammation-mediated genetic and epigenetic alterations drive cancer development in the neighboring epithelium upon stromal abrogation of TGF- β signaling. *PLoS Genet.* 2013; 9:e1003251.
75. Ker CG. Surgical safety margin of gastroenterological cancer surgery: A truth or a dream? *Formosan Journal of Surgery.* 2014; 47:83–89.
76. Chu CM, Yao CT, Chang YT, Chou HL, Chou YC, Chen KH, Terng HJ, Huang CS, Lee CC, Su SL, Liu YC, Lin FG, Wetter T, Chang CW. Gene expression profiling of colorectal tumors and normal mucosa by microarrays meta-analysis using prediction analysis of microarray, artificial neural network, classification, and regression trees. *Dis Markers.* 2014; 2014:634123.
77. Chan SK, Griffith OL, Tai IT, Jones SJ. Meta-analysis of colorectal cancer gene expression profiling studies identifies consistently reported candidate biomarkers. *Cancer Epidemiol Biomarkers Prev.* 2008; 17:543–52.