



Identification of Common and Distinct Pathways in Inflammatory Bowel Disease and Colorectal Cancer: A Hypothesis Based on Weighted Gene Co-Expression Network Analysis

OPEN ACCESS

Edited by:

Damjan Glavač,
University of Ljubljana, Slovenia

Reviewed by:

Mario Gorenjak,
University of Maribor, Slovenia
Seyed Reza Mohebbi,
Shahid Beheshti University of Medical
Sciences, Iran

*Correspondence:

Nicola Silvestris
n.silvestris@oncologico.bari.it
Behzad Baradaran
baradaranb@tbzmed.ac.ir

†These authors have contributed
equally to this work

‡These authors share last authorship

Specialty section:

This article was submitted to
Computational Genomics,
a section of the journal
Frontiers in Genetics

Received: 04 January 2022

Accepted: 14 March 2022

Published: 31 March 2022

Citation:

Derakhshani A, Javadrahi D,
Hemmat N, Dufour A, Solimando AG,
Abdoli Shadbad M, Duijf PHG,
Brunetti O, Silvestris N and
Baradaran B (2022) Identification of
Common and Distinct Pathways in
Inflammatory Bowel Disease and
Colorectal Cancer: A Hypothesis
Based on Weighted Gene Co-
Expression Network Analysis.
Front. Genet. 13:848646.
doi: 10.3389/fgene.2022.848646

Afshin Derakhshani¹, Darya Javadrahi^{2†}, Nima Hemmat^{2†}, Antoine Dufour^{3,4}, Antonio Giovanni Solimando⁵, Mahdi Abdoli Shadbad⁶, Pascal H. G. Duijf^{7,8,9}, Oronzo Brunetti¹⁰, Nicola Silvestris^{11*‡} and Behzad Baradaran^{2,12*‡}

¹Department of Biochemistry and Molecular Biology, University of Calgary, Calgary, AB, Canada, ²Immunology Research Center, Tabriz University of Medical Sciences, Tabriz, Iran, ³Departments of Physiology and Pharmacology, University of Calgary, Calgary, AB, Canada, ⁴McCaig Institute, Hotchkiss Brain Institute and Snyder Institute for Chronic Diseases, University of Calgary, Calgary, AB, Canada, ⁵Department of Biomedical Sciences and Human Oncology, School of Medicine, Aldo Moro University of Bari, Bari, Italy, ⁶Student Research Committee, Tabriz University of Medical Sciences, Tabriz, Iran, ⁷School of Biomedical Sciences, Faculty of Health, Queensland University of Technology, Brisbane, QLD, Australia, ⁸Centre for Data Science, Queensland University of Technology, Brisbane, QLD, Australia, ⁹Translational Research Institute, University of Queensland Diamantina Institute, The University of Queensland, Brisbane, QLD, Australia, ¹⁰Medical Oncology Unit, IRCCS Istituto Tumori Giovanni Paolo II, Bari, Italy, ¹¹Medical Oncology Unit, Department of Human Pathology "G. Barresi", University of Messina, Messina, Italy, ¹²Department of Immunology, Faculty of Medicine, Tabriz University of Medical Sciences, Tabriz, Iran

Patients with inflammatory bowel disease (IBD), including ulcerative colitis and Crohn's disease, are at higher risk to develop colorectal cancer (CRC). However, the underlying mechanisms of this predisposition remain elusive. We performed in-depth comparative computational analyses to gain new insights, including weighted gene co-expression network analysis (WGCNA) and gene ontology and pathway enrichment analyses, using gene expression datasets from IBD and CRC patients. When individually comparing IBD and CRC to normal control samples, we identified clusters of highly correlated genes, differentially expressed genes, and module-trait associations specific for each disease. When comparing IBD to CRC, we identified common hub genes and commonly enriched pathways. Most notably, IBD and CRC share significantly increased expression of five genes (*MMP10*, *LCN2*, *REG1A*, *REG3A*, and *DUOX2*), enriched inflammatory and neutrophil activation pathways and, most notably, highly significant enrichment of IL-4 and IL-13 signaling. Thus, our work expands our knowledge about the intricate relationship between IBD and CRC development and provides new rationales for developing novel therapeutic strategies.

Keywords: colorectal cancer, inflammatory bowel disease, pathogenesis, systems biology, WGCNA

Abbreviations: CRC, colorectal cancer; IBD, inflammatory bowel disease; WGCNA, weighted network correlation analysis; KEGG, kyoto encyclopedia of genes and genomes; DEGs, differentially expressed genes; MM, module membership; GS, gene significance; GEO, gene expression omnibus; GO, gene ontology; ME, module eigengene; scFv-Reg3a, single-chain variable fragment targeting REG3A; TOM, topological overlap matrix; dissTOM, dissimilarity TOM.

1 INTRODUCTION

Colorectal cancer (CRC) is one of the leading causes of cancer-related morbidity and mortality globally (Rawla et al., 2019). It is the third most common cancer and the fourth most common cause of cancer-related deaths worldwide (Hagggar and Boushey, 2009). Despite considerable advances in the treatment of CRC patients over the past decades, treatment response rates remain variable. Thus, there is a need to develop new therapies with higher efficacies and minimal side effects (Derakhshani et al., 2021a).

Inflammatory bowel disease (IBD) is a chronic multifactorial relapsing-remitting disease, and two main subtypes are Crohn's disease (CD) and ulcerative colitis (UC). In genetically predisposed individuals, incompletely identified environmental stimuli result in an aberrant immune-driven inflammatory response directed towards altered gut microbiota (Yalchin et al., 2021).

There is a close association between chronic inflammation and CRC development (Balkwill and Mantovani, 2001; Korniluk et al., 2017). The dysregulated immune responses in IBD are considered risk factors for CRC development (Kim and Chang, 2014; Zhang et al., 2017). CRC risk in patients with IBD was assessed in the first comprehensive meta-analysis in 2001 by Eaden et al., and their results showed a 2% risk of CRC at 10 years following UC diagnosis, an 8 percent risk at 20 years, and an 18 percent risk at 30 years, with an overall CRC prevalence of 3.7 percent. Patients with UC have a 1.7-fold increased risk of developing CRC, with an overall adjusted hazard ratio of 1.66 (95 percent confidence interval: 1.57–1.76), and a 1.6-fold increased risk of dying from CRC, when compared with the general population, according to the most extensive population study conducted to date ($n = 96,447$) (Eaden et al., 2001).

Additional risk factors include microbes, immunological responses, genetic mutations, and lifestyle/environmental triggers, yet the driving mechanism(s) of CRC remains uncharacterized mainly (Kaistha and Levine, 2014). Indeed, a better understanding of the molecular mechanisms of chronic inflammation could help identify novel drug targets for more personalized therapy for each patient (Ogunwobi et al., 2020).

Systems biology approaches, including transcriptomics, proteomics, and metabolomics, offer an unbiased survey of individual CRC patients (Derakhshani et al., 2020; Hemmat et al., 2020). Network-based approaches comprise reliable methods to identify specific pathways that could be regulating CRC. Since CRC is rarely caused by a single gene mutation but rather by disruptions in intricate gene networks, network-based approaches have provided a plethora of new opportunities to study the pathogenesis of cancer and to develop new cancer therapeutic approaches (Barabási et al., 2011).

Weighted gene co-expression network analysis (WGCNA) is an effective network-based approach that identifies clusters—or modules—of highly correlated genes. WGCNA is designed to analyze gene expression datasets by quantifying the associations between independent gene pairs and the degree to which these genes share the same neighbors (Hemmat et al., 2020). This tool is particularly suitable for identifying new

potential biomarkers or therapeutic target genes, a strategy that has been successfully applied to a range of cancers (Derakhshani et al., 2020; Liao et al., 2020). However, WGCNA has not yet been used to study the link between IBD and CRC. Therefore, we used WGCNA to identify differentially expressed genes and enriched pathways in IBD and CRC in our current study. The significantly changing genes and pathways that we identified will further our knowledge of the relationship between IBD and CRC development.

2 MATERIALS AND METHODS

2.1 Microarray Datasets and Processing

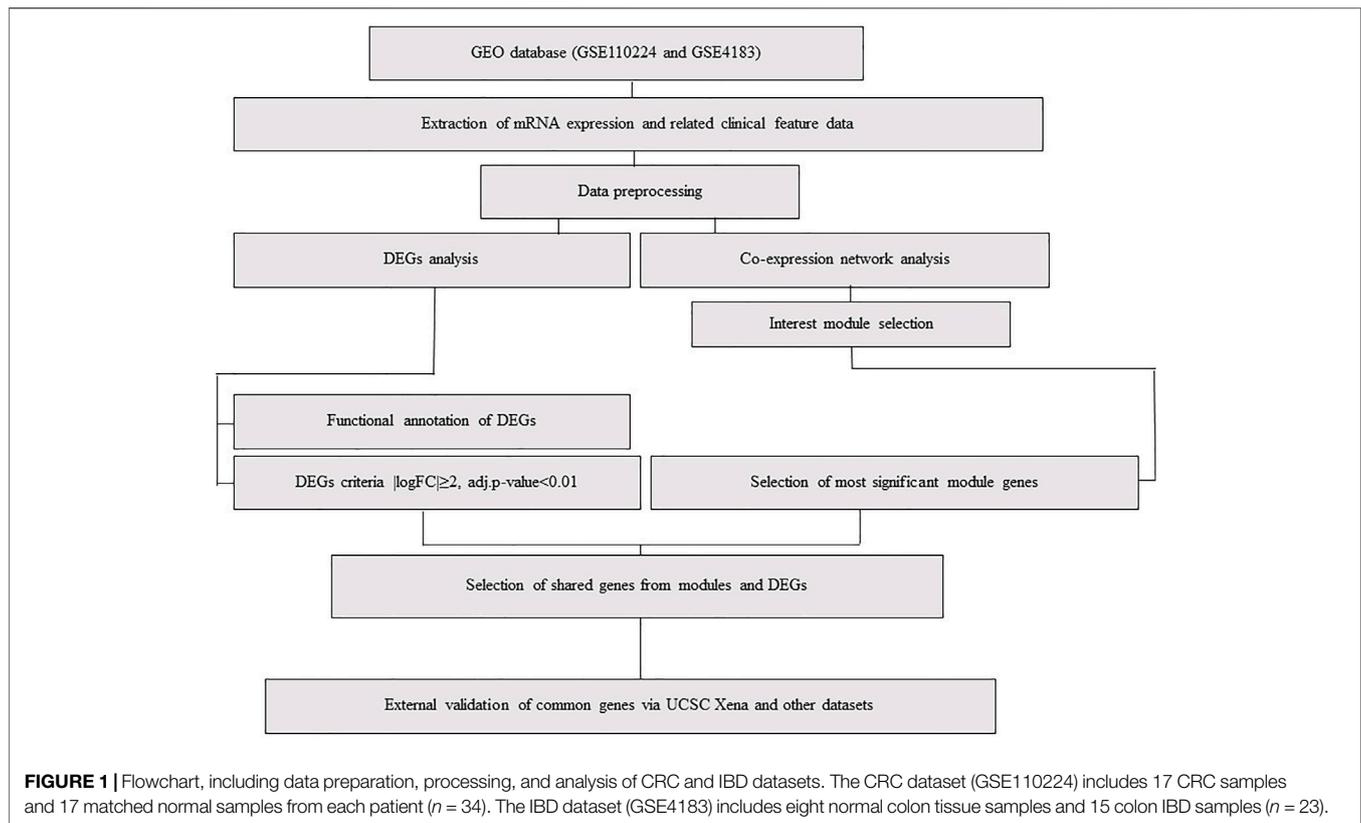
The GSE110224 and GSE4183 microarray datasets were downloaded from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>). The GSE110224 dataset was based on GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array with 34 samples (17 patients with colorectal cancer and 17 matched adjacent normal tissue samples). All samples of this dataset were used for downstream analyses ($n = 34$). The GSE4183 dataset was created on the GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array and contained 53 samples, including 15 patients with CRC, 15 with adenoma, 15 IBD tissue samples, and eight healthy normal control tissue samples. Eight control samples and the 15 IBD samples of this dataset were used for analyses ($n = 23$; **Supplementary Table S1**). The raw data were corrected, quantile-normalized, and probe IDs were converted to gene symbols. If there were multiple probes for genes, probe signals for the same gene were collapsed by averaging to yield a single expression level per gene per sample. Gene symbols were filtered across all samples through their variance. Only genes with variances ranked in the top 5,000 were selected for subsequent analyses.

2.2 Construction of Co-Expression Modules of CRC and IBD

Co-expression networks for gene expression data of patients and the control group were reconstructed using the R (<http://www.r-project.org/index.html>) and R/Bioconductor packages (Gentleman et al., 2004), such as WGCNA (Langfelder and Horvath, 2008). Briefly, the matrix of the gene expression profiles was converted into the matrix of pairwise gene similarity according to the Pearson test, followed by conversion into a matrix of adjacency. According to the already represented scale-free gene co-expression topological algorithm, when the β value is considered 12 and 7 (GSE110224 and GSE4183, respectively), the adjacency matrix met the scale-free topology criteria. Then, a topological overlap matrix (TOM) and a “dissimilarity” matrix (dissTOM) were created using TOM similarity and dissimilarity modules, based on the correlation of the pairwise gene-co-expression and according to **Eq. I**.

$$dissTOM = 1 - TOM \quad (I)$$

Finally, the clusters of highly interconnected genes were created with a minimum module size of 30 genes and a cut height of 0.001 (Horvath, 2011).



2.3 Construction of Module-Trait Relationships and Hub Genes of CRC and IBD

To identify modules significantly related to the evaluated clinical trait, the expression profiles of each module were summarized via their module eigengene (ME) as the eigenvector correlated to the first principal component of the expression matrix. The gene significance (GS) values were used for measuring individual gene associations with CRC and IBD. Also, module membership (MM) was defined as the correlation of the ME and the gene expression profile for each module. When the GS and MM were highly correlated (determined by Pearson correlation analysis), the most significant (central) elements in the modules were also closely associated with the trait (Zhang and Horvath, 2005). The top-ranked significant module genes from CRC were compared with the most significant IBD module genes using the freely available *Venny* tool (version 2.1) (<http://bioinfogp.cnb.csic.es/tools/venny/>). Then, Co-expression networks of shared gene modules between CRC and IBD were then constructed using GeneMANIA (<https://genemania.org/>).

2.4 Identification of DEGs and Biological Pathways for CRC and IBD

Identification of DEGs among patients and normal samples was carried out using a volcano plot, indicating the upregulated and downregulated genes in CRC and IBD samples. DEGs were called if they met two criteria: (1) $|\log_2(\text{FC})| \geq 2$ and (2) $p_{\text{adjust}} < 0.01$. Herein, FC represents the fold change in gene expression and

p_{adjust} denotes the Benjamini and Hochberg-adjusted p -value to compensate for multiple hypothesis testing. To investigate the biological functions and pathways of selected DEGs and gene modules, functional gene ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed, and results were visualized using the *ClueGO* v2.5.7 and *CluePedia* v1.5.7 plugins of Cytoscape v3.8.1 [18]. Enriched ontological terms and pathways with a threshold of Benjamini-adjusted p -value < 0.000001 were evaluated. Alternatively, pathway analysis was performed using *Reactome* (Jassal et al., 2020) and Benjamini-adjusted p -values.

2.5 Identification of DEGs for CRC and IBD

The top significant module genes of each dataset and DEGs were selected for comparison with other datasets. Common genes were chosen for further analysis using the *Venny* tool (version 2.1) software available freely at (<http://bioinfogp.cnb.csic.es/tools/venny/>).

The *pheatmap* visualization R package (<https://CRAN.R-project.org/package=pheatmap>) was used to show gene expression patterns between IBD and normal tissues. Additionally, the approach was used to demonstrate the gene expression patterns between CRC and normal samples. Then the shared genes were visualized.

2.6 External Validation of Common Genes

For CRC, external validation of common IBD and CRC genes was carried out by using The Cancer Genome Atlas (TCGA) dataset in the UCSC Xena browser (<https://xenabrowser.net/>). The gene expression levels in primary tumors, normal solid tissue,

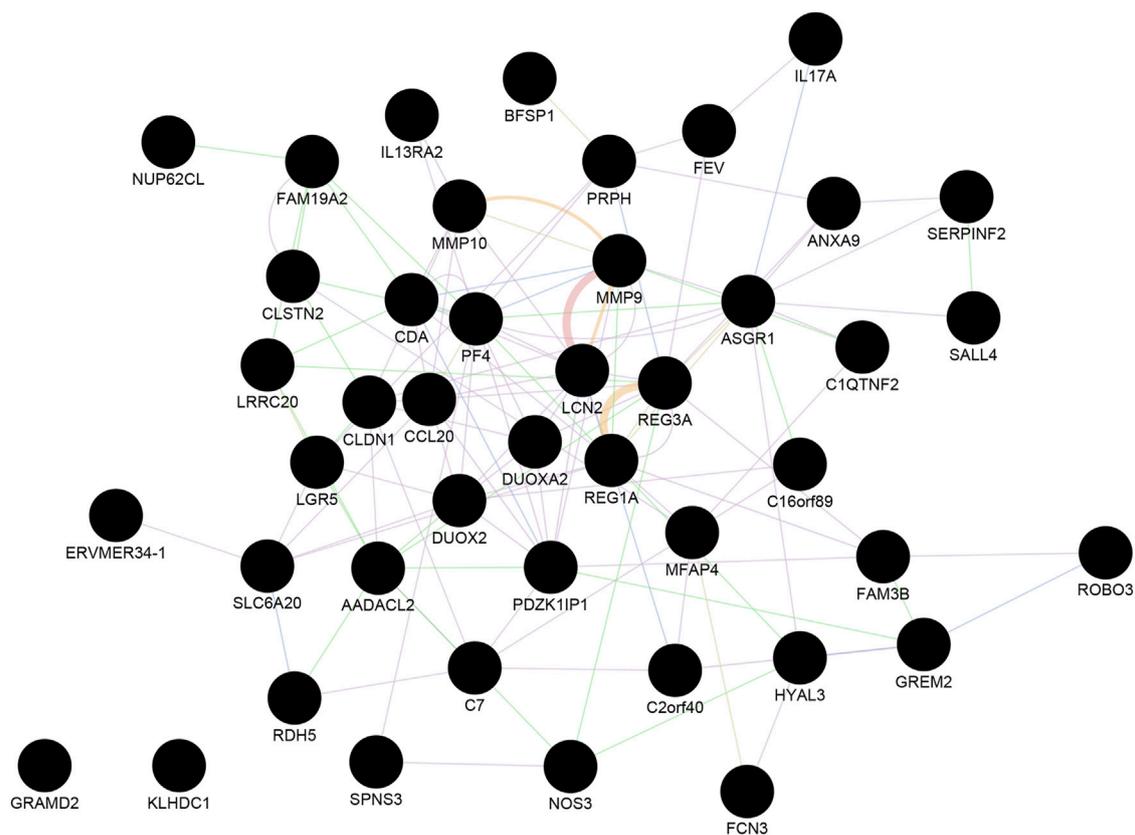


FIGURE 2 | The shared gene modules between CRC and IBD. A total of 48 shared hub genes (cancer compared with normal tissue; IBD compared with normal samples) were identified and imported to GeneMANIA to construct a co-expression and gene interaction network. Each circle represents a gene. Each edge (line) represents interactions between two genes. Multiple lines correspond to multiple sources, as indicated by the color coding. See also **Supplementary Tables S3, S4** for additional information.

recurrent tumor, and metastatic tumor tissues were extracted from the TCGA-COAD dataset using the UCSC Cancer Browser. Additionally, to identify the overlap between our findings and independent studies and validate our results using various microarray and RNA-seq data, GEO2R tool (<https://www.ncbi.nlm.nih.gov/geo/geo2r/>; R 3.2.3, Biobase 2.30.0, GEOquery 2.40.0, limma 3.26.8) and edgeR package (Robinson et al., 2010) were applied in the current study (**Supplementary Table S1**).

2.7 Static Analyses

R versions (3.2.0 and 4.02) and a range of R packages for statistical analyses, such as WGCNA, annotate, Limma, Affy, AgiMicroRna, ggplot2, edgeR, and pheatmap were used. The statistical tests used are specified in the main text, and Benjamini-adjusted *p*-values were used to compensate for multiple-hypothesis testing.

3 RESULTS

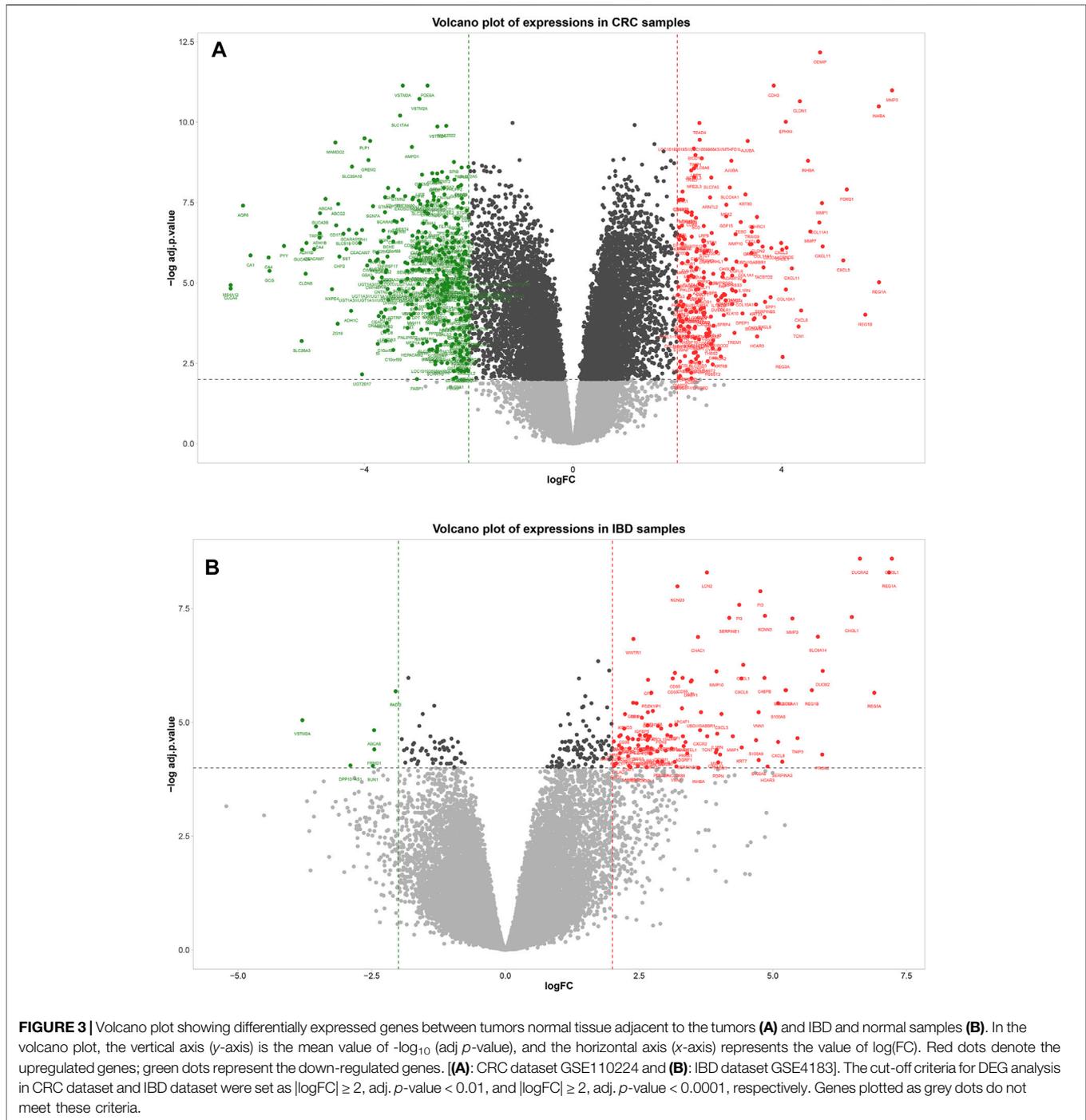
3.1 Identification of Modules in CRC and IBD Samples

To investigate gene expression's overlap and differences between IBD and CRC, we used two microarray datasets: GSE110224 and GSE4183 (**Figure 1**). For CRC, GSE110224 includes 17 CRC

samples and 17 matched adjacent normal tissue samples. For IBD, GSE4183 contains normal colon tissue samples ($n = 8$) and IBD samples ($n = 15$). First, we performed a clustering analysis, indicating that normal, CRC, and IBD samples clustered together across the two datasets (**Supplementary Figure S1**). Then, specific soft-threshold power was selected for each dataset. A weighted co-expression network was constructed for CRC and IBD, using the normal samples in each of these datasets as controls (**Supplementary Figure S2**; see Methods for details). We constructed clustering dendrograms for module illustration with different colors (**Supplementary Figure S3**). The numbers of genes included in each module in each dataset are shown in **Supplementary Table S1** and **Supplementary Table S2**. The genes in the gray "module" were not classified into any modules, as they failed to classify as a distinct co-expression module.

3.2 Identification of Module-Trait Associations

We evaluated module-module correlations and eigengenes for each module to determine how modules were associated with the disease. Our weighted gene co-expression network analysis (WGCNA) revealed that the blue module, derived from the

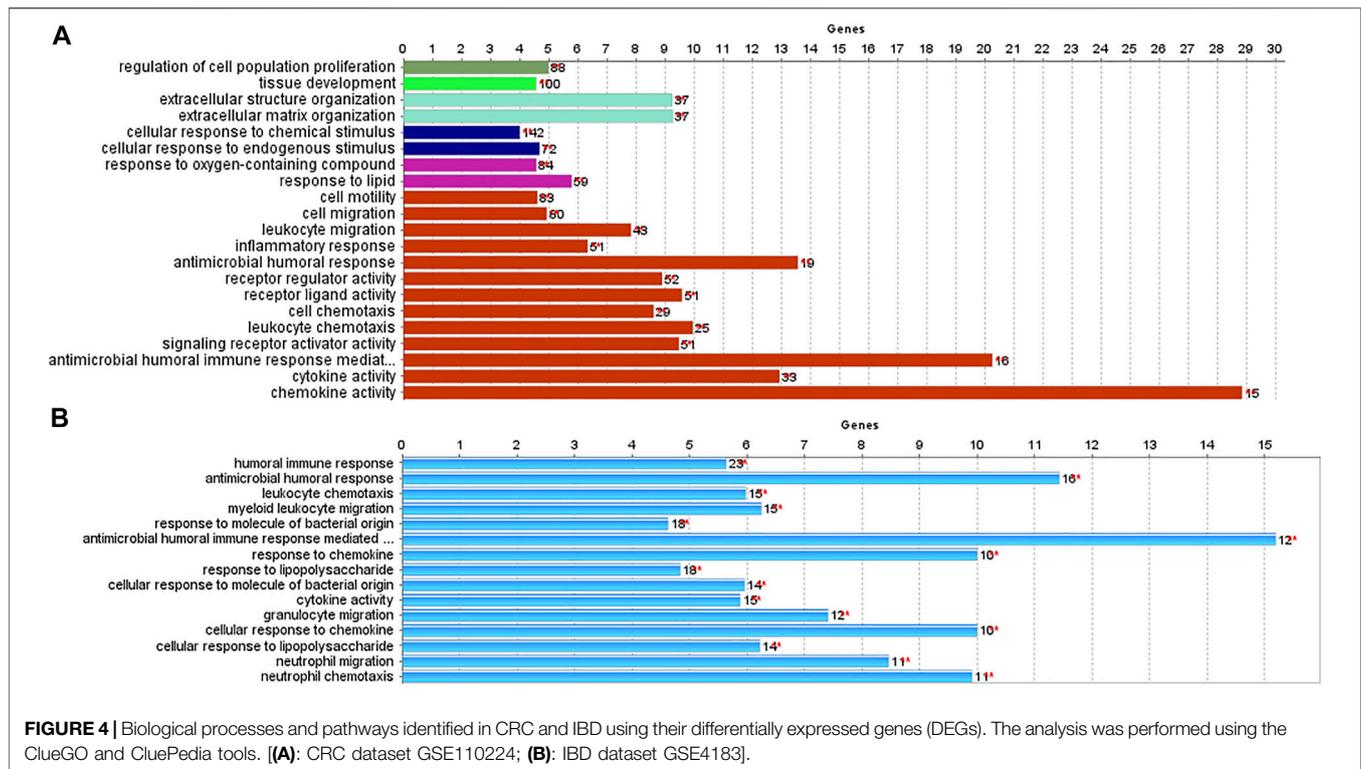


CRC dataset (GSE110224), had a negative correlation. This was the most significant module in the CRC dataset. However, the turquoise module in the GSE4183 showed a positive correlation (Supplementary Figure S4, Supplementary Table S1, and Supplementary Table S2). The correlation between module membership (MM) and gene significance (GS) of the selected modules was highly statistically significant ($R \geq 0.95$, $p < 10^{-200}$ for blue module and $R \geq 0.93$, $p < 10^{-200}$ for the turquoise module; Pearson correlation; Supplementary Figure S5). The illustrated

genes in each module are highly significant and are most correlated with IBD and CRC.

3.3 Mutual Hub Genes From the Most Significant Modules Identify Common Pathways

Next, we identified the most significant shared gene modules in the two datasets by assessing the overlap of the identified genes.



Altogether, 201 mutual gene modules were identified between the two datasets. This network comprised 48 genes (**Supplementary Table S3, Figure 2**). Pathway analysis using Reactome identified IL-4 and IL-13 signaling as significantly enriched (adjusted p -value = 0.0003) (**Supplementary Table S4**).

3.4 Identification of Genes Differentially Expressed in CRC and IBD Datasets

A total of 573 genes were identified as differentially expressed genes (DEGs) in the CRC samples using thresholds of adjusted p -value < 0.01 and $|\log_{2}FC| \geq 2$, including 201 upregulated and 372 down-regulated genes in the tumor compared to the adjacent tissue. In the IBD dataset, 116 genes were detected as DEGs using the same thresholds, including 110 upregulated and six down-regulated genes in IBD compared to normal samples. The volcano plots in **Figure 3** show the upregulated and downregulated genes in CRC and IBD samples. Both DEGs shared the most important genes involved in inflammation and cancer, such as *CXCL3*, *MMP7*, *MMP3*, *INHBA*, *CXCL2*, *CEMIP*, *DUOX2*, *GZMB*, *CXCL1*, *CXCL11*, *CXCL8*, *REG1A*, *VSTM2A*, *REG1B*, *CXCL5*, *CXCL6*, *REG3A*, *MMP10*, *SLC6A14*, *MMP1*, *LCN2*, *S100A9*, etc.

3.5 GO and KEGG Analysis of DEGs

We used the *ClueGO* and *CluePedia* tools to investigate the biological processes and pathways of identified DEGs. Biological functions and pathways associated with CRC were receptors to chemokine activity, antimicrobial humoral immune response mediated by an antimicrobial peptide, extracellular matrix organization, and extracellular structure organization. For IBD,

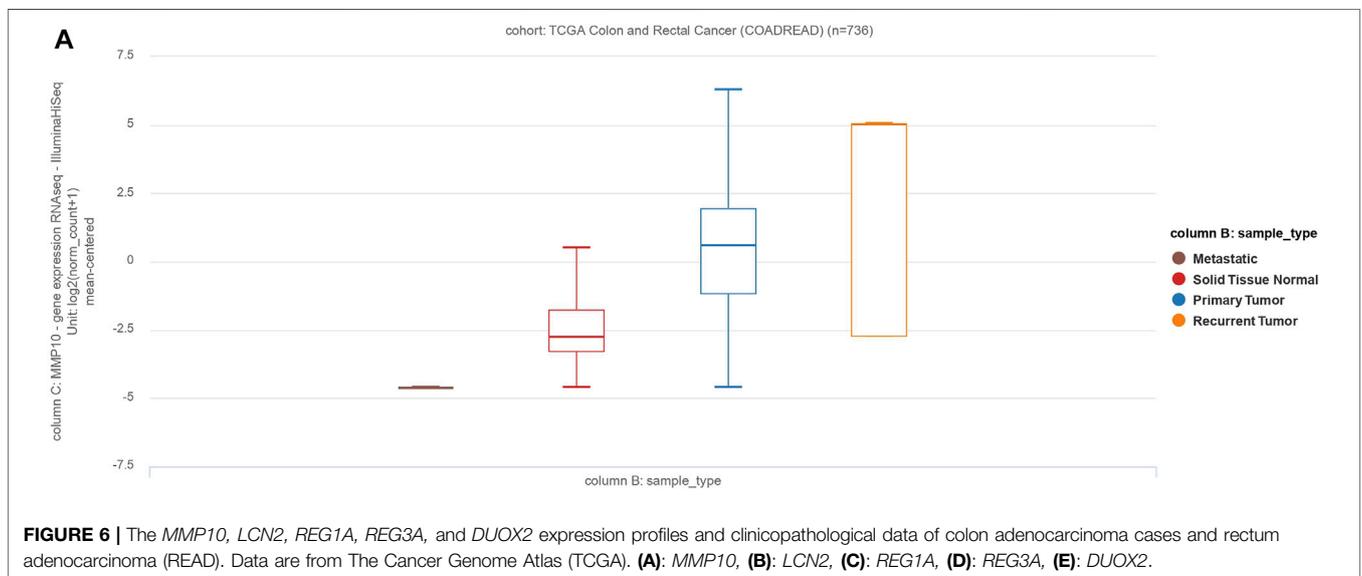
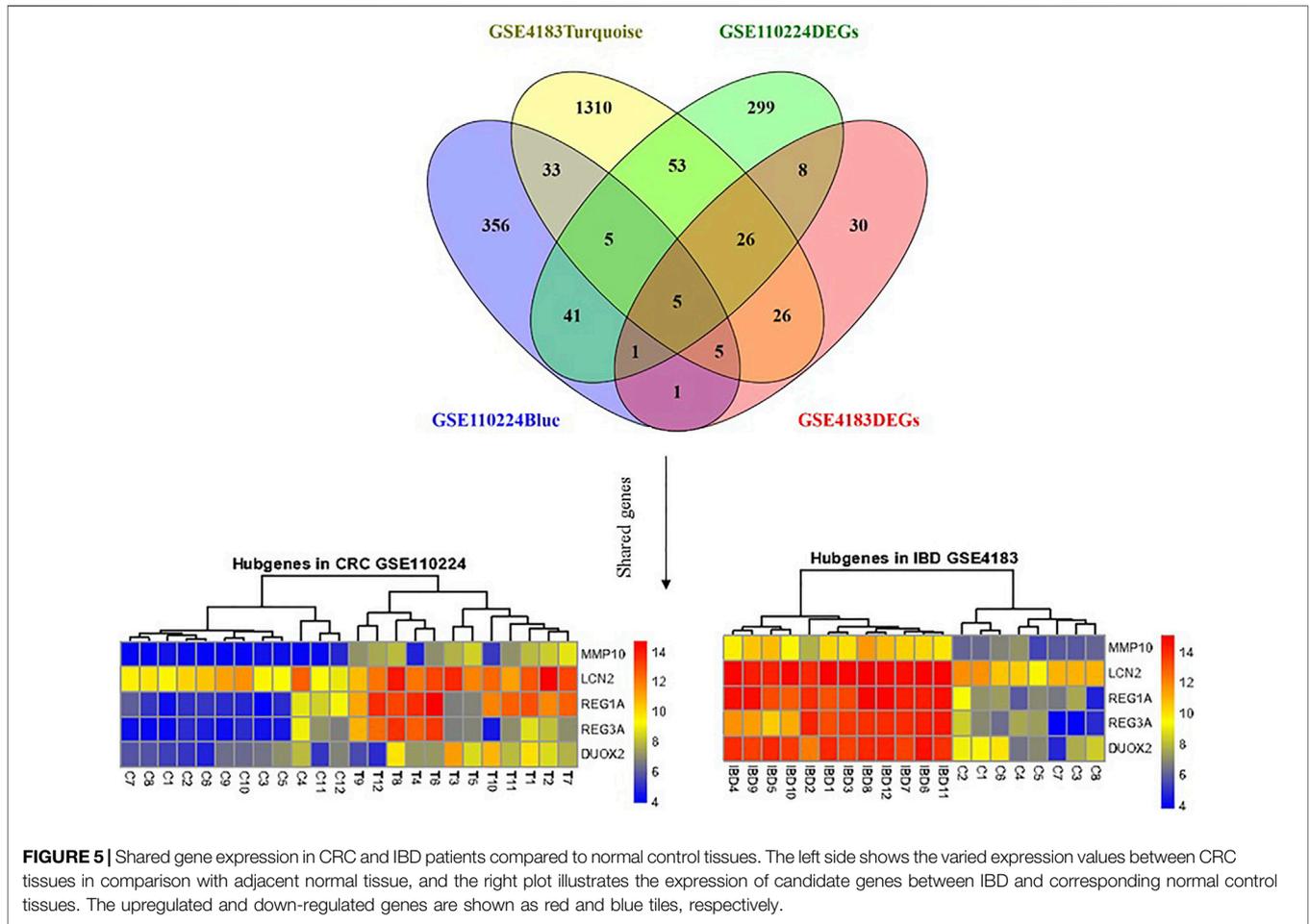
enriched pathways included inflammatory response, antimicrobial humoral response, and neutrophil activation. As demonstrated in **Figure 4**, antimicrobial humoral response and cytokine activation were shared pathways in CRC and IBD datasets, further supporting the key role of inflammation in CRC and IBD.

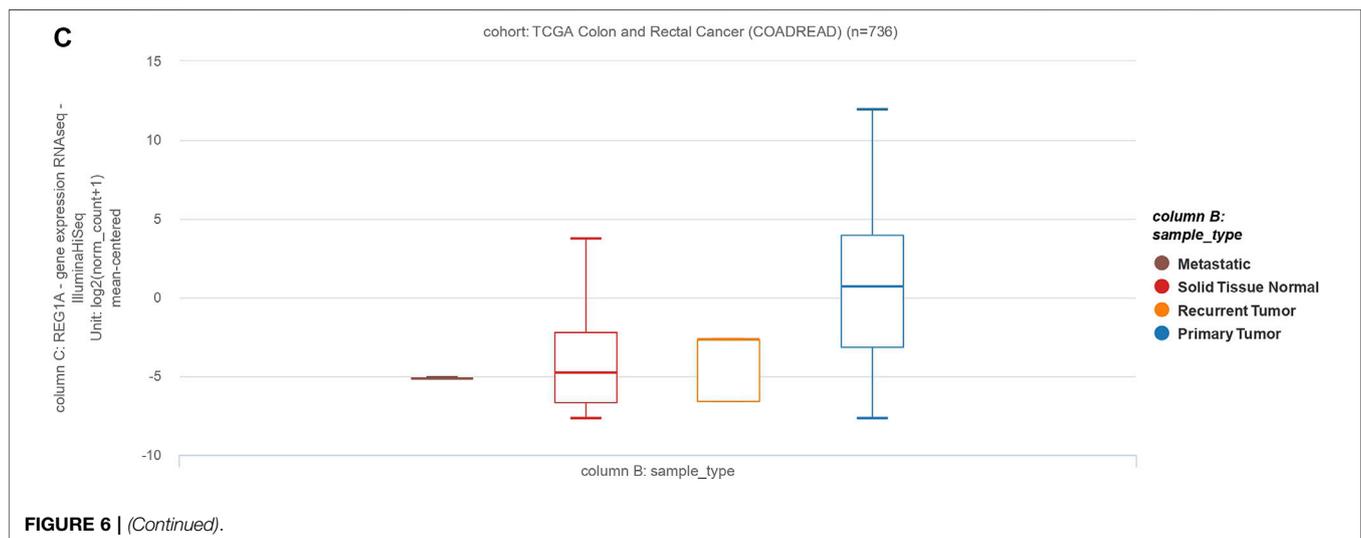
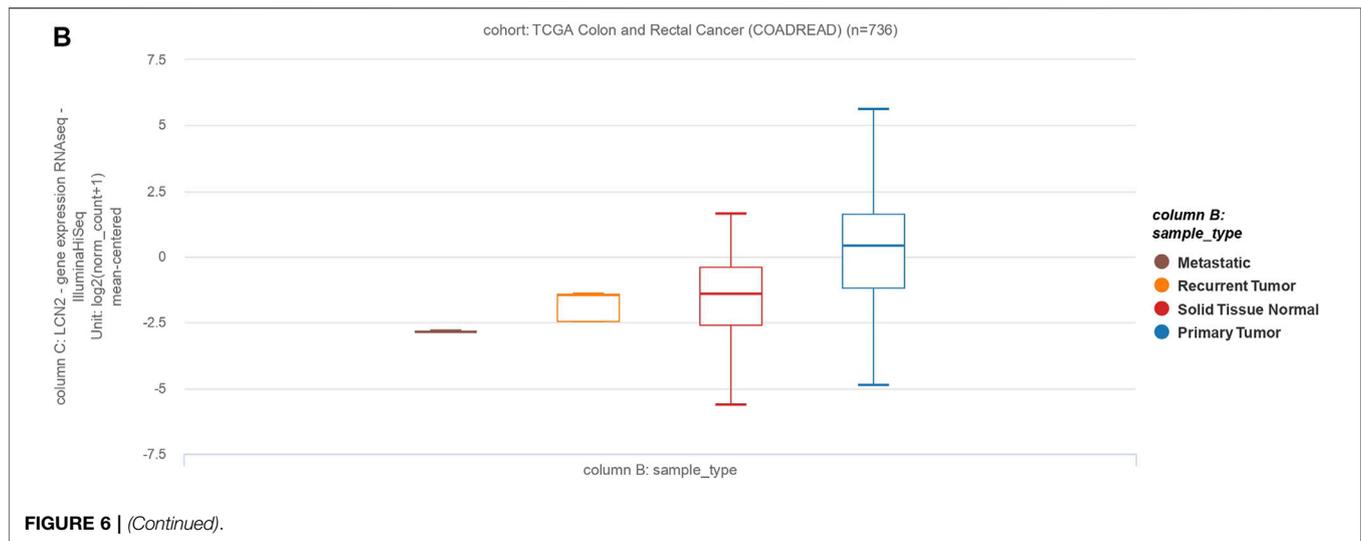
3.6 Common Genes Detection in all Datasets and Illustration by Heatmap

MMP10, *LCN2*, *REG1A*, *DUOX2*, and *REG3A* genes were considered mutual genes between the two primary analyzed datasets. As visualized by a Venn diagram, five candidate genes were defined as hub genes (module genes) and DEGs in both CRC and IBD: These were the mutual genes that were upregulated in both the CRC and the IBD samples, each compared to the respective internal normal control tissue groups (**Figure 5**). Notably, despite the small number of genes, Reactome pathway analysis of these genes identified the antimicrobial peptide, IL-4, and IL-13 signaling, and metal sequestration by antimicrobial protein pathways as statistically significantly enriched (adjusted p -values < 0.05) (**Supplementary Table S5**). Notable, our identification of IL-4 and IL-13 signaling validated our previous identification of this pathway based on the most significant modules (**Section 3.3** and **Supplementary Table S4**). Thus, our data further support the importance of IL-4 and IL-13 signaling in both CRC and IBD.

3.7 External Validation of Common Genes in CRC

To further validate our *in-silico* study, we compared our results to TCGA Colon and Rectum adenocarcinoma data. We identified a





total of five common genes (*MMP10*, *LCN2*, *REG1A*, *DUOX2*, and *REG3A*) whose expression was significantly upregulated. Here, we validate that these are significantly upregulated ($p < 0.05$) in various primary CRCs, recurrent CRCs, and metastatic CRCs compared to normal tissues (**Figure 6**). Thus, this validates our findings on an independent dataset from TCGA.

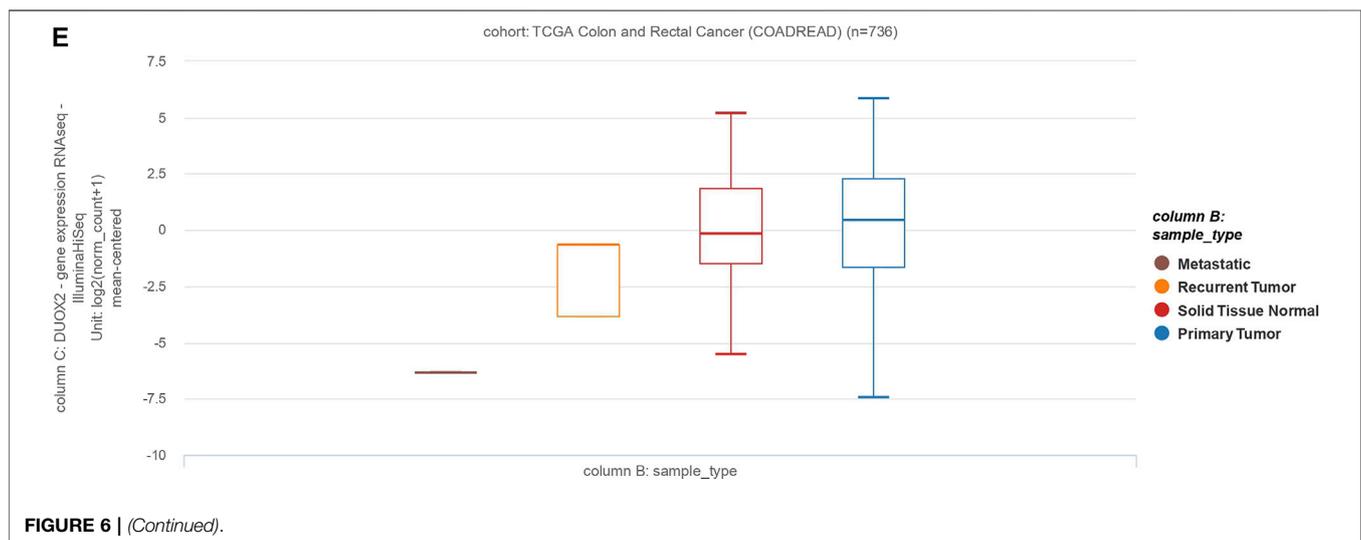
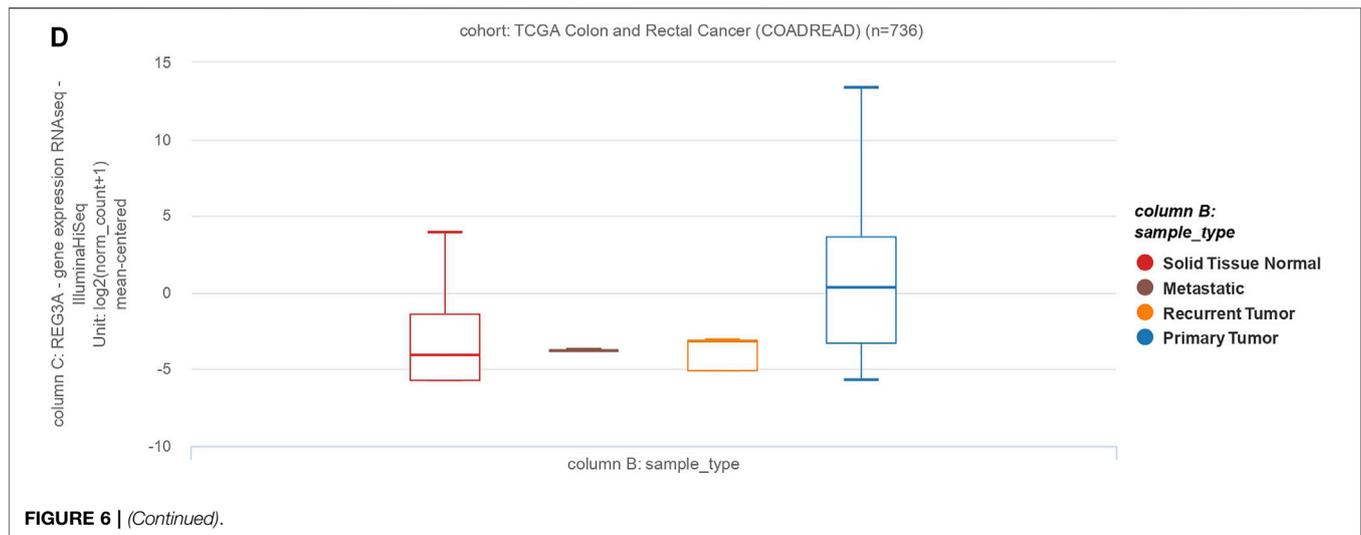
3.8 External Validation of Common Genes in IBD and its Clinical Impact

To validate the role of our five identified genes signature in IBD and CRC, we also validated their gene expression levels in 15 independent datasets, including UC, CD, CRC, and corresponding controls (GSE75970, GSE41328, GSE25070, GSE184093, GSE156451, GSE54986, GSE113513, GSE164541, GSE100179, GSE156732, GSE22619, GSE134025, GSE59071, GSE179285, and GSE102133) (**Supplementary Table S1** and **Figure 7**).

Our five identified genes signature was also significantly upregulated in these samples, suggesting a potential biomarker gene signature for identifying IBD patients at risk of developing CRC.

4 DISCUSSION

Patients with inflammatory bowel disease (IBD), such as Crohn's disease and ulcerative colitis, are at increased risk of developing colorectal cancer (CRC). Still, our knowledge about common and distinct pathways in these diseases remains elusive. Despite the reported correlation between IBD and CRC according to previous studies, some of the following research suggests otherwise (Söderlund et al., 2009; Jess et al., 2012b; Lutgens et al., 2013; Castaño-Milla et al., 2014). Despite disagreement regarding the specific incidence of CRC in patients with inflammatory bowel



disease (IBD), the overall link between IBD and an elevated risk of developing CRC is well documented and widely accepted in clinical practice (Eaden et al., 2001; Jess et al., 2012a).

Therefore, we aimed to identify overlapping and divergent pathways in this study. We demonstrate that weighted gene co-expression network analysis (WGCNA) is suitable for this study. WGCNA is a systematic biological method used to characterize the patterns of gene connections between diverse datasets. This enables the identification of potential biomarker genes or therapeutic targets based on the interconnection of genes and the relationship between genes and clinicopathological features of patients (Liao et al., 2020).

In recent years, systems biology techniques, such as WGCNA, have been mainly used to analyze novel biomarkers in different diseases, such as adrenocortical cancer, breast cancer, and CRC (Keller et al., 2019; Shi et al., 2020). It has been reported that IBD is associated with an increased risk of increase risk of CRC

development (Keller et al., 2019). However, work published to date has aimed to identify therapeutic targets for IBD or CRC individually, while shared and distinct genetic alterations between CRC and IBD—such as by using WGCNA—have not been a focus. Herein, we used WGCNA for the assessment of gene expression to describe hub modules and genes associated with the pathogenesis of IBD and CRC. In addition, we identified a signature of common significantly upregulated genes (*MMP10*, *LCN2*, *REG1A*, *DUOX2*, and *REG3A*), revealing the shared importance of IL-4 and IL-13 signaling pathway in both IBD and CRC.

Ulcerative colitis and Crohn's disease are common forms of IBD. Ulcerative colitis is restricted to the mucosal surface of the colon, while Crohn's disease may more broadly affect the gastrointestinal tract from the mouth to the anus with transmural inflammation (Kaistha and Levine, 2014). Our study has demonstrated that the most enriched pathways in

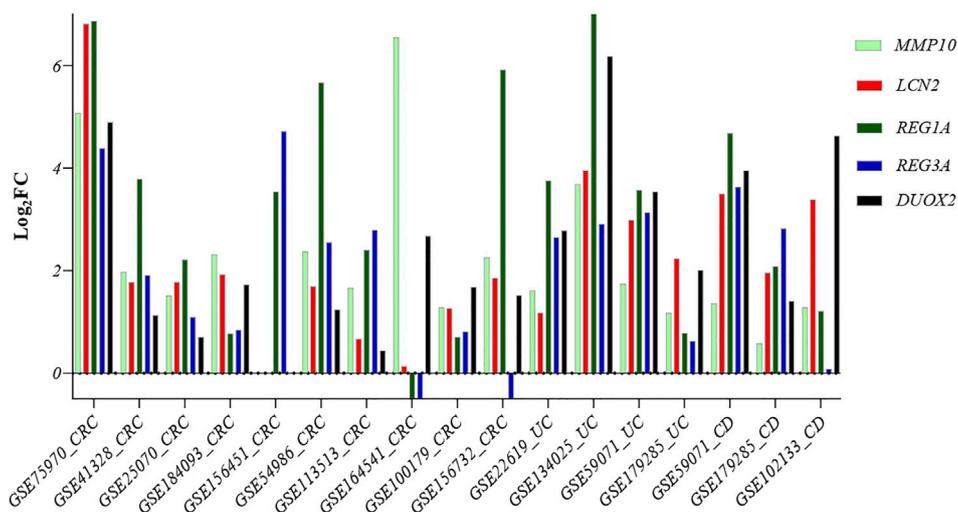


FIGURE 7 | Comparison of the expression level of the five common genes in samples from different diseases: Crohn's disease, ulcerative colitis, and CRC patients. Abbreviations: CRC: Colorectal cancer, UC: Ulcerative colitis, CD: Crohn's disease.

IBD and CRC are receptor-ligand activity, system development, leukocyte migration, inflammatory response, and neutrophil chemotaxis. These pathways are primarily involved in the inflammation process, reinforcing that there is a correlation between IBD and CRC initiation.

Our work identifies five genes that are mutually overexpressed in IBD and CRC. As shown in **Figure 5**, *MMP10*, *LCN2*, *REG1A*, *REG3A*, and *DUOX2* are all upregulated genes in CRC and IBD patients compared to matched normal tissue samples. Notably, these five genes are from the shared modules and represent both hub genes and DEGs in CRC and IBD. As illustrated in **Figure 6**, the expression levels of the hub genes in each sample of selected datasets are significantly different between normal and patient samples, and this reveals the importance of IL-4 and IL-13 signaling in the pathogenesis of both IBD and CRC.

MMPs affect a wide range of immunological functions, including leukocyte infiltration and chemokine activity (McMahon et al., 2016). Many MMPs have been shown to be overexpressed in UC, and their increased expression levels have been associated with the exacerbated disease. *MMP10* is produced predominantly by infiltrating myeloid cells in murine and human colitis. Bone marrow transplantation experiments showed that *MMP10* derived from bone marrow is involved in colitis severity (Koller et al., 2012). Furthermore, overexpression of *MMP10*, *MMP7*, and *MMP12* has been associated with the poor prognosis of CRC patients (Klupp et al., 2016). *MMP10* overexpression is also associated with the invasion of CRC cells (Klupp et al., 2016). Our results indicate that *MMP10* is differentially expressed in primary CRC, recurrent tumors, and metastatic tumors compared to normal tissue (**Figure 6A**).

Consistent with our work, as a member of the lipocalin superfamily, *LCN2* has been shown to play an essential role in the oncogenesis and progression of various tumors. Kim et al. demonstrated that *LCN2* is differentially expressed in CRC tissues

compared to normal tissues (Kim et al., 2017). Similarly, Stallhofer et al. showed that *LCN2* expression is substantially increased in patients with active IBD compared to normal individuals (Stallhofer et al., 2015). Our bioinformatics results have shown that *LCN2* is differently expressed in primary CRC, recurrent tumors, and metastatic tumors compared to non-tumoral samples (**Figure 6B**).

REG1A is a type I subclass member of the Reg gene family. Astrosini et al. reported that *REG1A* is upregulated in early-stage CRC patients with poor clinical outcomes. Additionally, *REG1A* is a prognostic factor for CRC patients (Astrosini et al., 2008). Mao et al. demonstrated that *REG1A* was increased in the inflamed colorectal tissues of IBD-bearing patients. Additionally, abnormally expressed *REG1A* inhibited inflammatory responses, enhanced cell proliferation, and lowered epithelial apoptosis in intestinal epithelial cells (IEC). Mechanically, IL-6 and IL-22 induced a significant increase in *REG1A* transcription via activating the JAK/STAT3 signaling pathway. Additionally, the systematic introduction of *REG1A* lentivirus to mice significantly reduced DSS-induced inflammatory damage and preserved the integrity of the intestinal mucosal barrier. Collectively, their findings suggest that the new proliferative factor *REG1A*, which is regulated by IL-6/IL-22-JAK-STAT3 signaling, may represent an attractive therapeutic target for patients with inflammatory bowel disease (Mao et al., 2021).

Ying et al. have shown that increased expression of *REG3A* is associated with increased tumor size, poor tumor differentiation, low survival rate, and advanced tumor stage. Interestingly, *REG3A* knockdown substantially arrests tumor cells in the G1 phase of the cell cycle, inhibits cell proliferation, suppresses cell migration, and promotes apoptosis in CRC cells (Ye et al., 2016). Similarly, Du et al. have shown that single-chain variable fragment targeting *REG3A* (scFv-Reg3a) can inhibit cell proliferation and tumor migration in CRC. Finally, targeting

REG3A enhances chemosensitivity to 5-fluorouracil in CRC (Du et al., 2020). In line with these observations, our results have shown that the expression of *REG1A* and *REG3A* is significantly increased in primary, recurrent, and metastatic CRC relative to non-tumor adjacent normal tissue (Figures 6C,D).

A fundamental gut-epithelial innate defense response is the release of hydrogen peroxide by dual NADPH oxidase (DUOX). *DUOX2* isoenzyme is the most induced gene in IBD, a disorder defined by an imbalanced gut microbiota-immune homeostasis (Grasberger et al., 2021). According to our analysis of datasets, we found out that this gene is also significantly upregulated in CRC and IBD tissues. In line with our results, Qi et al. have shown that expression of *DUOX2* is increased in CRC tissues.

Moreover, *DUOX2* protein levels are substantially increased in stages II-IV CRC compared to stage I CRC tissues (Qi et al., 2016). Macfie et al. demonstrated that *DUOX2* and *DUOXA2* could form an enzymatic system to produce reactive oxygen species in active ulcerative colitis, and *DUOX2* is upregulated in active colitis (Macfie et al., 2014). Since dysregulated reactive oxygen species can damage the DNA, increased levels of reactive oxygen species have been associated with tumorigenicity (Derakhshani et al., 2021b).

Our pathway analysis using Reactome identified IL-4 and IL-13 signaling as significantly enriched pathways in IBD and CRC. The involvement of IL-4 has previously been reported in colitis and colorectal cancer (Jess et al., 2012a; Chung et al., 2019; Lin et al., 2019). IL-13 and IL-4 are structurally and functionally similar cytokines sharing common receptor subunits. They govern immunological responses and are also involved in developing a range of human neoplasms. Three distinct receptors have been discovered for IL-4; however, only IL-4 receptor type II (IL-4Ra/IL-13Ra1) is expressed in solid tumors. While IL-13 can also connect to three separate receptors, IL-13 receptor type I (IL-4Ra/IL-13Ra1/IL-13Ra2) and type II (IL-4Ra/IL-13Ra1) are expressed in solid tumors. After binding to its receptor, IL-4 and IL-13 can mediate tumor cell proliferation, survival, and metastasis in colorectal and gastric cancers (Song et al., 2021).

REFERENCES

- Astrosini, C., Roefzaad, C., Dai, Y.-Y., Dieckgraefe, B. K., Jöns, T., and Kemmer, W. (2008). *REG1A* expression Is a Prognostic Marker in Colorectal Cancer and Associated with Peritoneal Carcinomatosis. *Int. J. Cancer* 123, 409–413. doi:10.1002/ijc.23466
- Balkwill, F., and Mantovani, A. (2001). Inflammation and Cancer: Back to Virchow? *The Lancet* 357, 539–545. doi:10.1016/s0140-6736(00)04046-0
- Barabási, A.-L., Gulbahce, N., and Loscalzo, J. (2011). Network Medicine: a Network-Based Approach to Human Disease. *Nat. Rev. Genet.* 12, 56–68. doi:10.1038/nrg2918
- Castaño-Milla, C., Chaparro, M., and Gisbert, J. (2014). Systematic Review with Meta-analysis: the Declining Risk of Colorectal Cancer in Ulcerative Colitis. *Aliment. Pharmacol. Ther.* 39, 645–659. doi:10.1111/apt.12651
- Chung, D. C., Bertelsen, M., Lorenz, B., Pennesi, M. E., Leroy, B. P., Hamel, C. P., et al. (2019). The Natural History of Inherited Retinal Dystrophy Due to Biallelic Mutations in the RPE65 Gene. *Am. J. Ophthalmol.* 199, 58–70. doi:10.1016/j.ajo.2018.09.024

Exploiting systems biology to identify the common DEGs and pathways of CRC and IBD could provide the researchers with very important insights about the shared characteristics of these two nearly related diseases. This could lead to enhancing the prognosis of CRC in IBD patients and better treatment options.

5 CONCLUSION

Taken together, our results have shed new light on the molecular association between IBD and CRC. We identified a signature of five upregulated genes, *MMP10*, *LCN2*, *REG1A*, *REG3A*, and *DUOX2*, in both IBD and CRC. We also identified IL-4 and IL-13 signaling as the predominant shared pathway in IBD and CRC patients. These new insights demonstrated opportunities for the development of novel therapeutic strategies.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

ADe: the first author of the manuscript, designed the project, analyzed the data and wrote the initial version of the manuscript. NH and PD organized the data. DJ and AS wrote some part of the paper. ADu revised the paper. MA and OB edited the paper grammatically. NS and BB supervised the study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.848646/full#supplementary-material>

- Derakhshani, A., Hashemzadeh, S., Asadzadeh, Z., Shadbad, M. A., Rasibonab, F., Safarpour, H., et al. (2021a). Cytotoxic T-Lymphocyte Antigen-4 in Colorectal Cancer: Another Therapeutic Side of Capecitabine. *Cancers* 13, 2414. doi:10.3390/cancers13102414
- Derakhshani, A., Mollaei, H., Parsamanesh, N., Fereidouni, M., Miri-Moghaddam, E., Nasser, S., et al. (2020). Gene Co-expression Network Analysis for Identifying Modules and Functionally Enriched Pathways in Vitiligo Disease: A Systems Biology Study. *Iran J. Allergy Asthma Immunol.* 19, 517–528. doi:10.18502/ijaa.v19i5.4467
- Derakhshani, A., Rostami, Z., Safarpour, H., Shadbad, M. A., Nourbakhsh, N. S., Argentiero, A., et al. (2021b). From Oncogenic Signaling Pathways to Single-Cell Sequencing of Immune Cells: Changing the Landscape of Cancer Immunotherapy. *Molecules* 26, 2278. doi:10.3390/molecules26082278
- Du, P., Wang, X., Yin, T., Zhang, X., Zhang, Z., Yu, W., et al. (2020). Anti-tumor Effect of Single-Chain Antibody to Reg3a in Colorectal Cancer. *Exp. Cel. Res.* 396, 112278. doi:10.1016/j.yexcr.2020.112278
- Eaden, J. A., Abrams, K., and Mayberry, J. (2001). The Risk of Colorectal Cancer in Ulcerative Colitis: a Meta-Analysis. *Gut* 48, 526–535. doi:10.1136/gut.48.4.526

- Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., et al. (2004). Bioconductor: Open Software Development for Computational Biology and Bioinformatics. *Genome Biol.* 5, R80–R16. doi:10.1186/gb-2004-5-10-r80
- Grasberger, H., Magis, A. T., Sheng, E., Conomos, M. P., Zhang, M., Garzotto, L. S., et al. (2021). DUOX2 Variants Associate with Preclinical Disturbances in Microbiota-Immune Homeostasis and Increased Inflammatory Bowel Disease Risk. *J. Clin. Invest.* 131, e141676. doi:10.1172/JCI141676
- Haggar, F., and Boushey, R. (2009). Colorectal Cancer Epidemiology: Incidence, Mortality, Survival, and Risk Factors. *Clin. Colon Rectal Surg.* 22, 191–197. doi:10.1055/s-0029-1242458
- Hemmat, N., Derakhshani, A., Bannazadeh Baghi, H., Silvestris, N., Baradaran, B., and De Summa, S. (2020). Neutrophils, Crucial, or Harmful Immune Cells Involved in Coronavirus Infection: A Bioinformatics Study. *Front. Genet.* 11, 641. doi:10.3389/fgene.2020.00641
- Horvath, S. (2011). *Weighted Network Analysis: Applications in Genomics and Systems Biology*. Springer Science & Business Media.
- Jassal, B., Matthews, L., Viteri, G., Gong, C., Lorente, P., Fabregat, A., et al. (2020). The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* 48, D498–d503. doi:10.1093/nar/gkz1031
- Jess, T., Rungoe, C., and Peyrin-Biroulet, L. (2012a). Risk of Colorectal Cancer in Patients with Ulcerative Colitis: a Meta-Analysis of Population-Based Cohort Studies. *Clin. Gastroenterol. Hepatol.* 10, 639–645. doi:10.1016/j.cgh.2012.01.010
- Jess, T., Simonsen, J., Jørgensen, K. T., Pedersen, B. V., Nielsen, N. M., and Frisch, M. (2012b). Decreasing Risk of Colorectal Cancer in Patients with Inflammatory Bowel Disease over 30 Years. *Gastroenterology* 143, 375–381. doi:10.1053/j.gastro.2012.04.016
- Kaitha, A., and Levine, J. (2014). Inflammatory Bowel Disease: The Classic Gastrointestinal Autoimmune Disease. *Curr. Probl. Pediatr. Adolesc. Health Care* 44, 328–334. doi:10.1016/j.cppeds.2014.10.003
- Keller, D. S., Windsor, A., Cohen, R., and Chand, M. (2019). Colorectal Cancer in Inflammatory Bowel Disease: Review of the Evidence. *Tech. Coloproctol.* 23, 3–13. doi:10.1007/s10151-019-1926-2
- Kim, E. R., and Chang, D. K. (2014). Colorectal Cancer in Inflammatory Bowel Disease: the Risk, Pathogenesis, Prevention and Diagnosis. *Wjg* 20, 9872–9881. doi:10.3748/wjg.v20.i29.9872
- Kim, S. L., Lee, S. T., Min, I. S., Park, Y. R., Lee, J. H., Kim, D. G., et al. (2017). Lipocalin 2 Negatively Regulates Cell Proliferation and Epithelial to Mesenchymal Transition through Changing Metabolic Gene Expression in Colorectal Cancer. *Cancer Sci.* 108, 2176–2186. doi:10.1111/cas.13389
- Klupp, F., Neumann, L., Kahlert, C., Diers, J., Halama, N., Franz, C., et al. (2016). Serum MMP7, MMP10 and MMP12 Level as Negative Prognostic Markers in colon Cancer Patients. *BMC Cancer* 16, 494. doi:10.1186/s12885-016-2515-7
- Koller, F. L., Dozier, E. A., Nam, K. T., Swee, M., Birkland, T. P., Parks, W. C., et al. (2012). Lack of MMP10 Exacerbates Experimental Colitis and Promotes Development of Inflammation-Associated Colonic Dysplasia. *Lab. Invest.* 92, 1749–1759. doi:10.1038/labinvest.2012.141
- Korniluk, A., Koper, O., Kemona, H., and Dymicka-Piekarska, V. (2017). From Inflammation to Cancer. *Ir J. Med. Sci. (1971)* 186, 57–62. doi:10.1007/s11845-016-1464-0
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R Package for Weighted Correlation Network Analysis. *BMC Bioinformatics* 9, 559. doi:10.1186/1471-2105-9-559
- Liao, Y., Wang, Y., Cheng, M., Huang, C., and Fan, X. (2020). Weighted Gene Coexpression Network Analysis of Features that Control Cancer Stem Cells Reveals Prognostic Biomarkers in Lung Adenocarcinoma. *Front. Genet.* 11, 311. doi:10.3389/fgene.2020.00311
- Lin, X., Wang, S., Sun, M., Zhang, C., Wei, C., Yang, C., et al. (2019). miR-195-5p/NOTCH2-mediated EMT Modulates IL-4 Secretion in Colorectal Cancer to Affect M2-like TAM Polarization. *J. Hematol. Oncol.* 12, 20. doi:10.1186/s13045-019-0708-7
- Lutgens, M. W. M. D., Van Oijen, M. G. H., Van Der Heijden, G. J. M. G., Vleggaar, F. P., Siersema, P. D., and Oldenburg, B. (2013). Declining Risk of Colorectal Cancer in Inflammatory Bowel Disease. *Inflamm. Bowel Dis.* 19, 789–799. doi:10.1097/mib.0b013e31828029c0
- Macfie, T. S., Poulson, R., Parker, A., Warnes, G., Boitsova, T., Nijhuis, A., et al. (2014). DUOX2 and DUOX2 Form the Predominant Enzyme System Capable of Producing the Reactive Oxygen Species H₂O₂ in Active Ulcerative Colitis and Are Modulated by 5-aminosalicylic Acid. *Inflamm. Bowel Dis.* 20, 514–524. doi:10.1097/01.mib.0000442012.45038.0e
- Mao, H., Jia, J., Sheng, J., Zhang, S., Huang, K., Li, H., et al. (2021). Protective and Anti-inflammatory Role of REG1A in Inflammatory Bowel Disease Induced by JAK/STAT3 Signaling axis. *Int. Immunopharmacology* 92, 107304. doi:10.1016/j.intimp.2020.107304
- Mcmahan, R. S., Birkland, T. P., Smigiel, K. S., Vandivort, T. C., Rohani, M. G., Manicone, A. M., et al. (2016). Stromelysin-2 (MMP10) Moderates Inflammation by Controlling Macrophage Activation. *J.I.* 197, 899–909. doi:10.4049/jimmunol.1600502
- Ogunwobi, O. O., Mahmood, F., and Akingboye, A. (2020). Biomarkers in Colorectal Cancer: Current Research and Future Prospects. *Int. J. Mol. Sci.* 21, 5311. doi:10.3390/ijms21155311
- Qi, R., Zhou, Y., Li, X., Guo, H., Gao, L., Wu, L., et al. (2016). DUOX2 Expression Is Increased in Barrett Esophagus and Cancerous Tissues of Stomach and Colon. *Gastroenterol. Res. Pract.* 2016, 1835684. doi:10.1155/2016/1835684
- Rawla, P., Sunkara, T., and Barsouk, A. (2019). Epidemiology of Colorectal Cancer: Incidence, Mortality, Survival, and Risk Factors. *pg* 14, 89–103. doi:10.5114/pg.2018.81072
- Robinson, M. D., Mccarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data. *Bioinformatics* 26, 139–140. doi:10.1093/bioinformatics/btp616
- Shi, G., Shen, Z., Liu, Y., and Yin, W. (2020). Identifying Biomarkers to Predict the Progression and Prognosis of Breast Cancer by Weighted Gene Co-expression Network Analysis. *Front. Genet.* 11, 597888. doi:10.3389/fgene.2020.597888
- Söderlund, S., Brandt, L., Lapidus, A., Karlén, P., Broström, O., Löfberg, R., et al. (2009). Decreasing Time-Trends of Colorectal Cancer in a Large Cohort of Patients with Inflammatory Bowel Disease. *Gastroenterology* 136, 1561–1567. doi:10.1053/j.gastro.2009.01.064
- Song, X., Traub, B., Shi, J., and Kornmann, M. (2021). Possible Roles of Interleukin-4 and -13 and Their Receptors in Gastric and Colon Cancer. *Int. J. Mol. Sci.* 22, 727. doi:10.3390/ijms22020727
- Stallhofer, J., Friedrich, M., Konrad-Zerna, A., Wetzke, M., Lohse, P., Glas, J., et al. (2015). Lipocalin-2 Is a Disease Activity Marker in Inflammatory Bowel Disease Regulated by IL-17A, IL-22, and TNF- α and Modulated by IL23R Genotype Status. *Inflamm. Bowel Dis.* 21, 2327–2340. doi:10.1097/MIB.0000000000000515
- Yalchin, M., Baker, A.-M., Graham, T. A., and Hart, A. (2021). Predicting Colorectal Cancer Occurrence in IBD. *Cancers* 13, 2908. doi:10.3390/cancers13122908
- Ye, Y., Xiao, L., Wang, S.-J., Yue, W., Yin, Q.-S., Sun, M.-Y., et al. (2016). Up-regulation of REG3A in Colorectal Cancer Cells Confers Proliferation and Correlates with Colorectal Cancer Risk. *Oncotarget* 7, 3921–3933. doi:10.18632/oncotarget.6473
- Zhang, B., and Horvath, S. (2005). A General Framework for Weighted Gene Co-expression Network Analysis. *Stat. Appl. Genet. Mol. Biol.* 4, Article17. doi:10.2202/1544-6115.1128
- Zhang, M., Sun, K., Wu, Y., Yang, Y., Tso, P., and Wu, Z. (2017). Interactions between Intestinal Microbiota and Host Immune Response in Inflammatory Bowel Disease. *Front. Immunol.* 8, 942. doi:10.3389/fimmu.2017.00942

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Derakhshani, Javadrashid, Hemmat, Dufour, Solimando, Abdoli Shadbad, Duijf, Brunetti, Silvestris and Baradaran. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.