Genomics 112 (2020) 3416–3426

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Detection of molecular signatures and pathways shared in inflammatory bowel disease and colorectal cancer: A bioinformatics and systems biology approach

Md. Al-Mustanjid[a], S.M. Hasan Mahmud[b,*], Md. Rejaul Islam Royel[a], Md. Habibur Rahman[c], Tania Islam[d], Md. Rezanur Rahman[e], Mohammad Ali Moni[f,*]

[a] Department of Software Engineering, Faculty of Science and Information Technology, Daffodil International University, Dhaka 1207, Bangladesh
[b] School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
[c] Department of Information and Communication Technology, Mawlana Bhashani Science and Technology University, Tangail, Bangladesh
[d] Department of Biotechnology and Genetic Engineering, Faculty of Biological Sciences, Islamic University, Kushtia 7003, Bangladesh
[e] Department of Biochemistry and Biotechnology, School of Biomedical Science, Khwaja Yunus Ali, University, Enayetpur, Sirajganj 6751, Bangladesh
[f] WHO Collaborating Centre on eHealth, UNSW Digital Health, School of Public Health and Community Medicine, Faculty of Medicine, UNSW Sydney, Australia

ARTICLE INFO

Keywords:
Inflammatory bowel disease
Colorectal cancer
Systems biology
Protein-protein interaction
DEGs-transcription factors interaction
DEGs-microRNAs interaction
Protein-drug interaction
Gene-disease association

ABSTRACT

Emerging evidence indicates IBD is a risk factor for the increasing incidence of colorectal cancer (CRC) development. We used a system biology approach to identify common molecular signatures and pathways that interact between IBD and CRC and the indispensable pathological mechanisms. First, we identified 177 common differentially expressed genes (DEGs) between IBD and CRC. Gene set enrichment, protein-protein, DEGs-transcription factors, DEGs-microRNAs, protein-drug interaction, gene-disease association, Gene Ontology, pathway enrichment analyses were conducted to these common genes. The inclusion of common DEGs with bimolecular networks disclosed hub proteins (LYN, PLCB1, NPSR1, WNT5A, CDC25B, CD44, RIPK2, ASAP1), transcription factors (SCD, SLC7A5, IKZF3, SLC16A1, SLC7A11) and miRNAs (mir-335-5p, mir-26b-5p, mir-124-3p, mir-16-5p, mir-192-5p, mir-548c-3p, mir-29b-3p, mir-155-5p, mir-21–5p, mir-15a-5p). Analysis of the interaction between protein and drug discovered ASAP1 interacts with cysteine sulfonic acid and double oxidized cysteine drug compounds. Gene-disease association analysis retrieved ASAP1 also associated with pulmonary and bladder neoplasm diseases.

1. Introduction

Inflammatory bowel disease (IBD) is a gastrointestinal disorder that causes chronic inflammation of the digestive tract and long-term inflammation damages the gastrointestinal tract. The two major types of IBD are Ulcerative colitis (UC) which causes inflammation at the large intestine, and Crohn's disease which generally causes inflammation in any portion of the digestive tract that mostly impacts the back half of the small bowel. The UC is considered as the third highest risk factors of CRC development. Eaden et al. [1] analyzed 116 studies involving around 55,000 UC patients indicated 1700 patients developed CRC, and the rate of diagnosis of CRC increased over time. In two large nationwide studies of Japanese and Danish people, it was observed the poor survival rate and higher histologic grade in UC-associated CRC than sporadic CRC [1]. Europe and North America have recorded the largest incidence of IBD [2]. Colorectal cancer (CRC) is also the 3rd most prevalent disease of the digestive systems.

The prolonged inflammation and extensive colitis are involved in the processes of transformation of cells into malignant cells in CRC [3]. S.Brackman and his team members investigated some patients who have carried out IBD symptoms from the last 10 years. Within the period, they found 12% of the patients are developed CRC, and 21% of the patients directly diagnosed with IBD [4]. Several studies showed that there are several convincing pieces of evidence for pathological and epidemiological links between IBD and CRC from the population-based studies which revealed that IBD is a strong significant risk factor of CRC and vice versa [5–11].

Analysis of transcriptomes is commonly used to identify the candidate biomarkers for various disease including CRC [12–14] and IBD [15,16]. Despite some important efforts on decoding key signaling molecules in CRC and IBD, underlying biomarkers and pathways commonly shared by IBD and CRC are still not discovered. Therefore,

* Corresponding authors.
E-mail addresses: hasan.swe@daffodilvarsity.edu.bd (S.M.H. Mahmud), m.moni@unsw.edu.au (M.A. Moni).

https://doi.org/10.1016/j.ygeno.2020.06.001
Received 21 January 2020; Received in revised form 3 May 2020; Accepted 2 June 2020
Available online 12 June 2020

we have engaged a systems biology approach to uncover common candidate biomarkers and underlying shared mechanisms with their origins between IBD and CRC. We studied two microarray datasets of IDB and CRC from NCBI-GEO to identify commonly dysregulated genes. To determine shared ontologies and pathways, the common DEGs were used to carry out ontological and functional enrichment study. Further, based on such findings, we have also scrutinized i) protein-protein interaction network deciphered by common genes, ii) DEGs and miRNAs interaction networks to discern transcriptional and post-transcriptional regulatory factors, ii) protein-drug interaction network to identify potential drugs, iii) gene-disease association network to acquire the link of other associated diseases with the common DEGs. Also, we have identified crucial pathways, hub proteins, and transcriptions of interest. Thus, these inclusive systems biology processes enable us to unearth potential biomolecular signatures between IBD and CRC. This study provides new insights into the underlying pathology and clarifies potential mechanisms shared between IBD and CRC as well as further biological study.

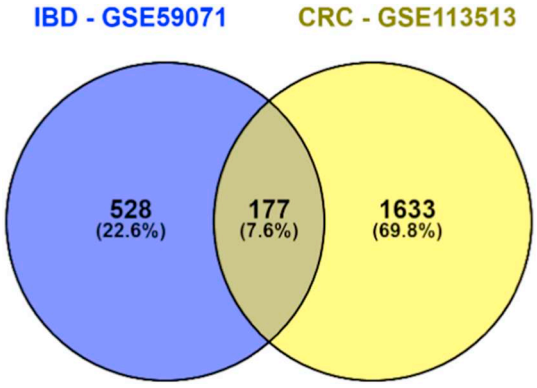

Fig. 2. Two high-throughput gene expression datasets comprising colorectal cancer dataset (GSE113513) and inflammatory bowel disorder (GSE59071) were analyzed. The jointed analysis showed 177 common differentially expressed genes between colorectal cancer and inflammatory bowel disorder datasets.

2. Methodology

In this study, we employed a methodological context to identify the underlying biological resources and processes of interrelation exchanged between IBD and CRC (Fig. 1). Datasets for gene expression were analyzed and significant DEGs in IBD and CRC were identified. Common DEGs were obtained between IBD and CRC and this common gene set was used to investigate functional and enrichment analysis. Gene ontological processes and signaling pathway enrichment experiments were led by the Enrichr tool. DEGs-TFs and DEGs-miRNA interaction analysis were performed to elicit reporter biomolecules using the NetworkAnalyst web platform. Again, protein-drug interaction and gene-disease association were also explored using the NetworkAnalyst. After that, using the String database, the PPI network was generated and visualized using the Cytoscape. Besides, the topological analysis was conducted using the Cytoscape to detect hub proteins.

2.1. Datasets and identification of differentially expressed genes

We studied gene expression microarray datasets to capture genetic interrelations at the molecular level between IBD and CRC. For this study, we attained the gene expression microarray dataset from the NCBI-GEO database (https://www.ncbi.nlm.nih.gov/geo/) [17]. We evaluated two datasets with accession IDs GSE59071 for IBD and GSE113513 for CRC. IBD dataset (GSE59071) is the microarray gene expression data of endoscopic-derived intestinal mucosal biopsies from IBD and control patients. This dataset samples were recorded from endoscopy of colon ulcerative colitis (UC) patients, Crohn's disease patients, and 11 controls for taking biopsies in the most impacted fields but at a distance from ulcerations. While CRC dataset (GSE113513) is the microarray gene expression data of cancerous and matched non-cancerous tissues of CRC patients undergoing surgical resection. This dataset samples were accumulated from fourteen pairs of CRC primary lesions and non-cancerous surrounding tissue.

The datasets were examined to detect DEGs in IBD and CRC comparatively to their matched controls in the R (version 3.6.1) and the Bioconductor environment. The datasets for gene expression were normalized by log2 transformation and employed statistical method using the R package “Limma” with Benjamini-Hochberg correction to control the rate of false discovery. The significant DEGs was identified stand on P value $< .01$ and $|\text{logFC}| \geq 1$. Common genes were pinpointed from the two datasets as shown in Fig. 2 with Venn analysis using the web tool-Jvenn [18]Figure 2.

2.2. Gene ontology and pathway enrichment analysis

Gene ontology and pathway enrichment analysis were performed to find out the functional biological terms and signaling pathways of common DEGs. The EnrichR (https://amp.pharm.mssm.edu/Enrichr) - a widespread gene set enrichment web tool [19] was used to carry out the analyses. Gene ontology (GO) as well as functional process is divided into three categories: biological process, cellular component, and molecular functions. KEGG, Reactome, and Biocarta databases were selected as the source of pathway annotations. The p-value $< .01$ was determined as a significant margin for all analyses.

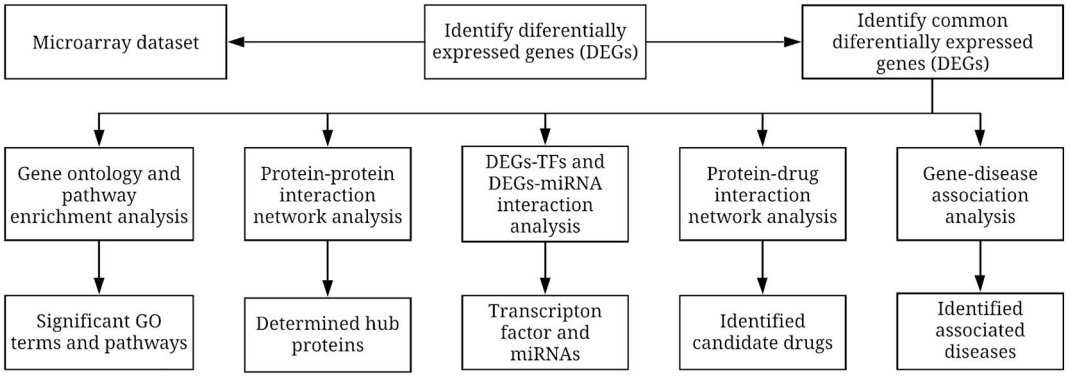

Fig. 1. Flowchart of analytical steps in this study.

2.3. Identification of transcription factors and miRNAs interact with common DEGs

Transcription factors (TFs) target gene interactions were analyzed to identify TFs that bind with DEGs at the regulatory regions. JASPAR is a freely accessible database for multi-species TFs profiles in six taxonomic groups [20]. NetworkAnalyst is an internet-based tool for meta-analysis of gene expression data and attains understanding into biological processes, functions, and presumptions [21]. We recognized significant TFs from JASPAR database based on topological analysis using NetworkAnalyst [21]. miRNAs target gene interactions were deployed to detect miRNAs that attempt to bind on a gene transcript to influence protein expression negatively [22]. Tarbase [23] and mirTarbase [22] are the largest databases of miRNA-target interactions with experimental validation. We characterized significant miRNAs from both Tarbase nd mirTarbase predicated on topological analysis through NetworkAnalyst.

2.4. PPI network construction

Using the STRING Protein-Protein Interaction database (version 11.0) (https:/string-db.org/) [24], the PPI network of proteins encoded by common DEGs was constructed to represent how our identified DEGs, as well as proteins, interact with each other physically and functionally. STRING provides evidence of interaction focused on observations and verified experimental results including text-mining, experiments, databases, co-expression, neighborhood, gene fusion, and co-occurrence information [24]. In STRING, annotation of protein interaction varies with confidence scores (low: 0.1, medium: 0.4, high: 0.7 and highest: 0.9). Due to a large number of common DEGs, we set the highest score confidence criteria to generate the PPI network sharply. Cytoscape software (v.3.7.1) - an open-source platform for network visualization [25] was used to visualize the PPI network and the network has been regarded as undirected. Based on topological assessment (node degree > 15), strongly connected proteins, as well as hub proteins, were spotted [26,27].

2.5. Protein-drug interaction analysis

DrugBank (version 5.0) is an intriguing web-based database of comparative drug information. At the same time, it also provides information about the effects of drugs at the protein expression level [28]. To locate prospective drugs from the DrugBank database that may interact with our common DEGs, we performed protein-drug interactions via NetworkAnalyst [21].

2.6. Gene-disease association analysis

DisGeNET is an extensive repository of gene-disease associations that incorporates connections from many sources covering various biomedical characteristics of diseases. It concentrates in general on the present insight into human genetic diseases, including mendelian, complicated, and ecological diseases [29]. Gene-Disease association was analyzed to acknowledge associated diseases and their chronic conditions with common DEGs through NetworkAnalyst [21].

3. Results

3.1. Identification of common DEGs between IBD and CRC

We analyzed the gene expression microarray data from NCBI-GEO to explore the reciprocal interrelation between IBD and CRC. The detection of dysregulated genes with this type of study compares control with the case for each disease. We screened significant DEGs considering p-value < .01 and $|\log FC| \geq 1$. We identified 705 significant DEGs in the IBD dataset where 519 and 186 DEGs were significantly up and down-regulated. We also identified 1810 significant DEGs in CRC in the same manner where 989 DEGs were up and 821 DEGs were down-regulated. We used the Jvenn tool to conduct cross-comparison to distinguish common DEGs between IBD and CRC. In this way, we discovered 177 DEGs as common between IBD and CRC (Fig. 1 reflects the overall cross comparative assessment of two datasets to achieve mutual DEGs between IBD and CRC). Further, this identified common gene-set was applied to obtain further experiments.

To get insights into the biological significance of the identified DEGs, we performed GO analysis and obtained enriched GO terms. Enrichr was used to conduct the gene ontology analysis of common genes within 3 categories (biological process, cellular component, molecular function) and the GO database was taken as an annotation source of ontological processes. Significant ontological processes have been identified through statistical analysis. Table 1 lists the top 10 terms in the biological process, molecular functions, and cellular components category are summarized in Table 1.

3.2. Pathway enrichment analysis

Pathway enrichment was assessed using Enrichr to acquire pathways enriched by the common DEGs. KEGG and Reactome databases were selected for preferred pathway sources, and significant pathways have been identified through p-value < .01. The top pathways collected from selected databases are enlisted in Table 2.

3.3. Determination of hub proteins

We investigated the protein-protein interactions around the proteins encoded by the DEGs to understand the signaling molecule (Fig. 3). We utilized the STRING PPI database for evaluating the PPI network among proteins encoded by common DEGs. The PPI network is made up of 632 nodes and 761 edges. The topological analysis of the PPI network determined eight central hub proteins (LYN, PLCB1, NPSR1, WNT5A, CDC25B, CD44, RIPK2, and ASAP1). The topological parameters and involvement of hub proteins in IBD and CRC or other forms were summarized in Table 3.

3.4. Determination of regulatory signatures

To get insights into the regulatory molecules of the hub protein or DEGs, we employed a network based approach to decode the regulatory TFs and miRNAs, the DEGs-TFs, and DEGs-miRNA interactions networks were assessed to detect transcriptional and post-transcriptional regulatory signatures of common DEGs. Fig. 4 illustrates the interactions between DEGs and TFs; Fig. 6 also shows interactions between DEGs and miRNAs. Five TFs namely SCD, SLC7A5, IKZF3, SLC16A1, SLC7A11, and ten miRNAs namely mir-335-5p, mir-26b-5p, mir-124-3p, mir-16-5p, mir-192-5p, mir-548c-3p, mir-29b-3p, mir-155-5p, mir-21–5p, mir-15a-5p were extracted from both interaction networks. Table 4 outlines all the definitions and literature analytical proofs affiliated with TFs and miRNAs in IBD and CRC.

3.5. Recognition of protein-drug interactions

To identify candidate drugs that might possibly influence IBD and CRC at the same time as the interaction between protein and drug was investigated. Protein-drug interaction [42,43] analysis is essential in understanding the structural characteristics that are necessary for receptor sensitivity [44]. Protein-drug interaction analysis uncovered the interaction of the drug with a hub protein. From Fig. 5 it appears that two characterized namely cysteine sulfonic acid and double oxidized cysteine drug compounds interact with the hub proteins namely CDC25B.

Table 1
Ontological analysis of common DEGs between IBD and CRC. Top 10 terms of each category are listed.

Categories	GO ID	Term	P-Value	Associated Genes
Biological Process	GO:0046688	Response to copper ion	4.38×10^{-11}	MT2A;MT1M;MT1F;MT1G;MT1X;MT1H;LOXL2;MT1E
	GO:0071294	Cellular response to zinc ion	1.74×10^{-10}	MT2A;MT1M;MT1F;MT1G;MT1X;MT1H;MTE
	GO:0006882	Cellular zinc ion homeostasis	2.82×10^{-10}	MT2A;MT1M;MT1F;SLC39A5;MT1G;MT1X;MT1H;MT1E
	GO:0055069	Zinc ion homeostasis	2.82×10^{-10}	MT2A;MT1M;MT1F;SLC39A5;MT1G;MT1X;MT1H;MT1E
	GO:0071280	Cellular response to copper ion	3.95×10^{-10}	MT2A;MT1M;MT1F;MT1G;MT1X;MT1H;MTE
	GO:0072503	Cellular divalent inorganic cation homeostasis	1.69E-09	MT2A;CCL23;EPHX2;MT1M;MT1F;SLC39A5;MT1G;MT1X;MT1H;CCL19;MT1E
	GO:0071276	Cellular response to cadmium ion	6.47E-09	MT2A;MT1M;MT1F;MT1G;MT1X;MT1H;MT1E
	GO:0046686	Response to cadmium ion	1.05E-08	MT2A;MT1M;MT1F;MT1G;MT1X;MT1H;MT1E
	GO:0010043	Response to zinc ion	1.33E-08	MT2A;MT1M;MT1F;MT1G;MT1X;MT1H;MT1E
	GO:0046916	Cellular transition metal ion homeostasis	1.51E-07	MT2A;MT1M;MT1F;SLC39A5;MT1G;MT1X;MT1H;ABCG2;MT1E
Molecular Function	GO:0046914	Transition metal ion binding	2.90E-06	APOBEC3G;MT1M;MT1X;RASGRP1;MT2A;RNF125;TDP2;SCD;ENPP2;MT1F;MT1G;MT1H;TIMP1;PCK1;MT1E
	GO:0046872	Metal ion binding	9.89E-06	EPHX2;MT1M;MT1X;RASGRP1;MT2A;SCIN;TDP2;CHP2;ENPP2;MT1F;CALU;MT1G;MT1H;PCK1;MT1E
	GO:0008270	Zinc ion binding	1.72*E*-05	MT2A;RNF125;APOBEC3G;MT1M;ENPP2;MT1F;MT1G;MT1X;MT1H;TIMP1;RASGRP1;MT1E
	GO:0008009	Chemokine activity	5.82E-05	CCL23;CXCL11;CXCL1;CCL19;CXCL3
	GO:0042379	Chemokine receptor binding	7.87E-05	CCL23;CXCL11;CXCL1;CCL19;CXCL3
	GO:0048407	Platelet-derived growth factor binding	1.07E-04	COL1A1;PDGFRA;COL4A1
	GO:0070888	E-box binding	2.51E-04	BHLHE40;TCF21;TCF4;ARNTL2
	GO:0045236	CXCR chemokine receptor binding	4.23E-04	CXCL11;CXCL1;CXCL3
	GO:0043425	Bhlh transcription factor binding	9.27E-04	BHLHE40;TCF21;TCF4
	GO:0015179	L-amino acid transmembrane transporter activity	9.94E-04	SLC36A1;SLC7A5;SLC7A11;SLC25A20
Cellular Function	GO:0042613	MHC class II protein complex	9.63E-08	HLA-DMA;HLA-DMB;HLA-DPB1;HLA-DQA1;HLA-DPA1
	GO:0042611	MHC protein complex	4.00E-07	HLA-DMA;HLA-DMB;HLA-DPB1;HLA-DQA1;HLA-DPA1
	GO:0005887	Integral component of the plasma membrane	4.75E-06	SLC26A2;SLC22A5;CD3G;SLC7A11;CLDN1;SLC4A4;PCDH18;P2RY8;CYSLTR1;EDNRB;TSPAN7;ENPP2;SLC39A5;PAG1;HLADQA1;CD55;HLADPA1;SLC17A4;PDGFRA;CD52;SLC16A1;P2RY14;CD2;SLC6A6;SLC7A5;MEP1A;GPR183;CD27;FXYD5;PLPP1;CD44
	GO:0045334	Clathrin-coated endocytic vesicle	0.001071293	CEMIP;HLA-DPB1;HLA-DQA1;HLA-DPA1
	GO:0071556	Integral component of lumenal side of endoplasmic reticulum membrane	0.002320851	HLA-DPB1;HLA-DQA1;HLA-DPA1
	GO:0030669	Clathrin-coated endocytic vesicle membrane	0.003333285	HLA-DPB1;HLA-DQA1;HLA-DPA1
	GO:0005764	Lysosome	0.00453247	SLC36A1;HLA-DMA;HLA-DMB;NAAA;HLA-DPB1;CTSV;RNF152;TINAG;HLA-DQA1;HLA-DPA1
	GO:0005765	Lysosomal membrane	0.004572072	SLC36A1;HLA-DMA;VNN1;HLA-DMB;HLA-

(*continued on next page*)

Table 1 (*continued*)

Categories	GO ID	Term	P-Value	Associated Genes
	GO:0098852	Lytic vacuole membrane	0.004907253	DPB1;RNF152;HLA-DQA1;HLA-DPA1 SLC36A1;HLA-DMA;HLA-DMB;HLA-DPB1;RNF152;HLA-DQA1;HLA-DPA1
	GO:0030665	Clathrin-coated vesicle membrane	0.006067109	HLA-DPB1;CD3G;HLA-DQA1;HLA-DPA1

3.6. Gene-disease association

One of the conditions for two diseases to be linked to or related to each other is that they must have at least one or more common genes [70,71]. The disease therapeutic design techniques begin with the unlocking of the association between genes and diseases [72]. Gene-disease association analysis has shown that ASAP1 hub protein is actively engaged in other diseases. Fig. 7 represents that ASAP1 hub protein is involved in tuberculosis, pulmonary, and bladder neoplasm diseases.

4. Discussion and conclusion

Much research is being undertaken to understand the IBD and CRC, but the association between the two diseases and how the IBD might influence the development of CRC is still not clearly understood. In the present study, we employed a systems biology approach to get insight into the transcriptional dysregulations and signaling molecules that show a common pattern of regulation in IBD and CRC. The statistical analysis of the IBD and CRC transcriptomics revealed 177 DEGs which show similar expression patterns in IBD and CRC. These common genes were evaluated by Gene Ontology (GO) and pathways analysis to get insight into the biological significance in the pathogenesis of IBD and CRC. GO is a generic theoretical model in the context of gene regulation that outlines the functions of genes and their relations. It progresses gradually through the acquisition of biological knowledge concerning gene functions and their regulations based on various ontological categories and linguistic relations between classes [73]. Enrichr was used to conduct the gene ontology analysis of common genes within 3 categories (biological process, cellular component, molecular function) and the GO database was taken as an annotation source of ontological processes. According to GO [74], biological processes are molecular activities, the cellular component being the cellular structure where the gene regulates its function, and the molecular term being a description of activities at the molecular level. The pathway-based analysis is a modern tactic that helps to understand and uncover how biologically or molecularly complex diseases are linked to one another. The pathway is the best way of getting an organism's reactions which are triggered by internal changes [75] [76]. Some of the identified pathways are of particular attention, for instance, cell adhesion a molecule (cams) was significantly enriched pathways according to our analysis. It has been demonstrated the crucial roles of cams in the metastasis of CRC [77]. The immune systems pathway was consistently enriched in this study. The immune system is one of the critical components in the tumor microenvironment. Therefore, immune systems pathways enriched in the present study suggesting the importance of tumor microenvironment in CRC [78]. Overall, the GO and pathways analysis explained the molecular mechanisms of CRC to some extent.

Protein-protein interaction is used to decode the key drivers in signaling molecules in molecular networks [79–81]. Topological analysis of the PPI networks reveals the most significant hub proteins. Therefore, in order to get insights into the signaling molecules, we conducted PPI analysis of the proteins encoded by the shared DEGs in IBD and CRC. The PPI analysis showed eight hub proteins that might be considered as the candidate key genes and drug targets in IBD and CRC (Table 3). The hub protein LYN is a proto-oncogene, and engaged in the activation of CD24-induced ERK1/2 in CRC and proposed to be involved in ulcerative colitis-associated CRC [30,31]. The hub protein PLCB1 is overexpressed in CRC tissues [33]. Two isoforms of hub genes NPSR1 has been reported to be altered in IBD and influence cancer-associated pathways [33,34]. The WNT5A is implicated an increased risk of ulcerative colitis and suggested as potential biomarkers of CRC [35–37]. The CDC25B is prognostic CRC markers [38]. CD44 is reported that it has a significant impact on tumor progression and metastasis [10,39]. The RIPK2 is implicated in the development of urothelial cancer [40], but its role in IBD or CRC is not understood yet. ASAP1 has been reported that its dysregulation is associated with poor

Table 2
Pathway enrichment analysis of common DEGs between IBD and CRC. The top pathways of each database are listed.

Category	Pathway Term	Adj. P-Value	Genes in the pathway
KEGG	Hematopoietic cell lineage	4.55×10^{-06}	CD2;HLA-DMA;HLA-DMB;ITGA4;HLA-DPB1;CD3G;HLA-DQA1;CD44;CD55;HLA-DPA1
	Mineral absorption	2.28×10^{-06}	MT2A;MT1M;MT1F;MT1G;MT1X;MT1H;TRPM6;MT1E
	Cell adhesion molecules (cams)	7.11×10^{-06}	CD2;HLA-DMA;HLA-DMB;ITGA4;CLDN8;ITGA8;HLA-DPB1;ITGAL;CLDN1;HLA-DQA1;HLA-DPA1
	Staphylococcus aureus infection	1.16×10^{-05}	C1QB;HLA-DMA;HLA-DMB;HLA-DPB1;ITGAL;HLA-DQA1;HLA-DPA1;C2
	Viral myocarditis	5.38×10^{-05}	HLA-DMA;HLA-DMB;HLA-DPB1;ITGAL;HLA-DQA1;CD55;HLA-DPA1
Reactome	Response to metal ions	1.85×10^{-09}	MT2A;MT1M;MT1F;MT1G;MT1X;MT1H;MT1E
	Metallothioneins bind metals	9.25×10^{-10}	MT2A;MT1M;MT1F;MT1G;MT1X;MT1H;MT1E
	Immune system	9.63×10^{-04}	IFITM3;PHLPP2;C1QB;IFITM1;WIPF1;NEDD4L;CD3G;CTSV;ITGAL;CLU;RASGRP1;C2;MT2A;FGF7;HLA-DMA;HLA-DMB;IFI16;IGKC;SLAMF7;TRIM22;PAG1;HLA-DQA1;CD55;HLA-DPA1;LYN;PDGFRA;ITGA4;RIPK2;RNF125;IRF4;HLA-DPB1;CD27;CD44
	Phosphorylation of CD3 and TCR zeta chains	7.28×10^{-04}	HLA-DPB1;CD3G;PAG1;HLA-DQA1;HLA-DPA1
	Cell surface interactions at the vascular wall	9.86×10^{-04}	LYN;CD2;SLC7A5;SLC16A1;ITGA4;SLC7A11;ITGAL;CD44

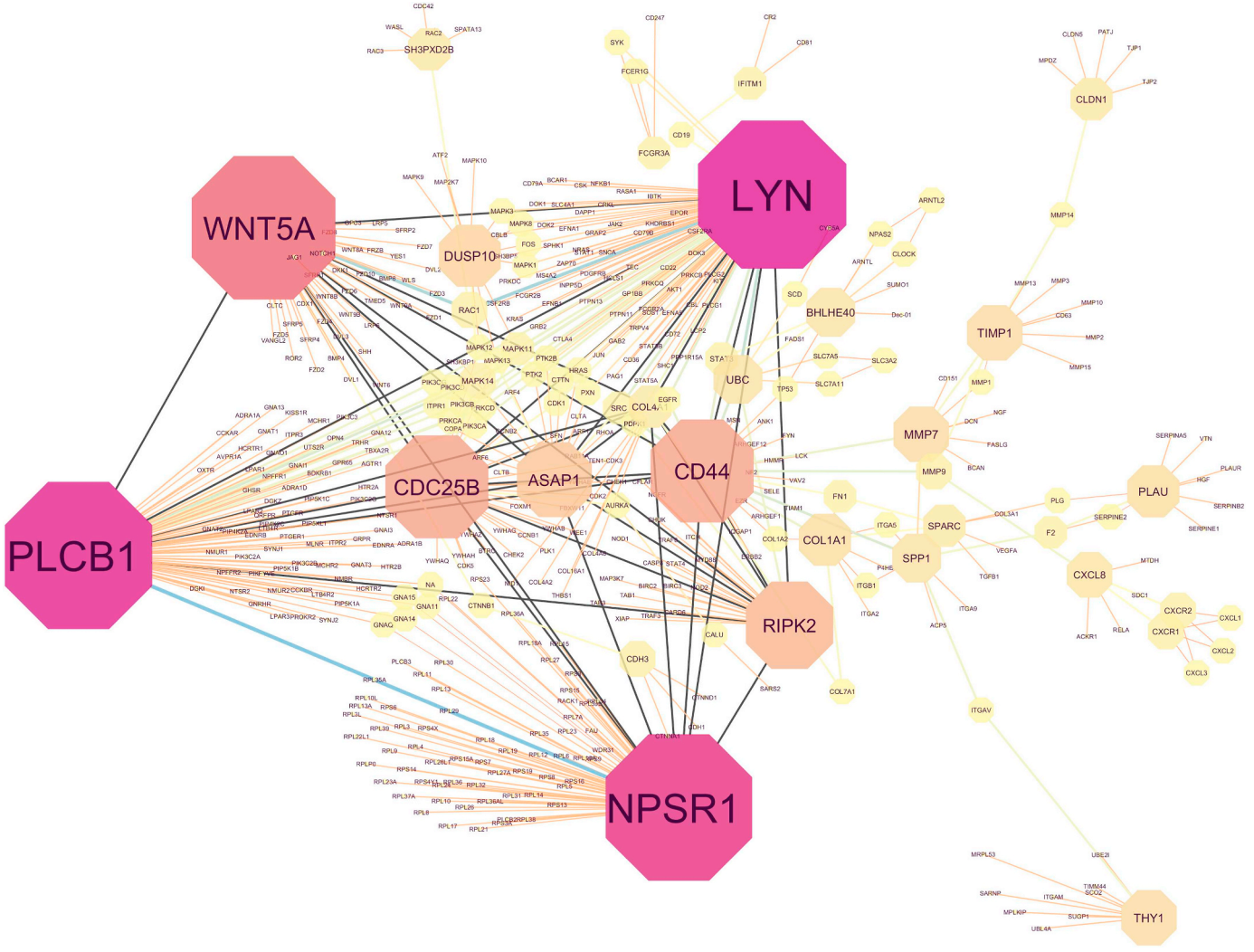

Fig. 3. Protein-protein interaction network of common DEGs between IBD and CRC. In figure, the nodes represent DEGs and edges represent the interactions between nodes. The PPI network has 632 nodes and 761 edges. The PPI network was generated using String and visualized in Cytoscape. The selection of hub proteins is based on the topological parameters (degree > 15).

survival outcomes in CRC [41].

The genes are regulated at the transcriptional and post-transcriptional levels. The significant regulatory molecules TF and miRNAs were evaluated to identify the regulatory molecules of DEGs. Among the identified TFs, SCD was reported as new biomarkers for the early risk of CRC [45]. SLC7A5 has been demonstrated as a therapeutic target for bladder cancer [46], but its association with IBD and CRC is not understood. The IKZF3 has been recognized as the key molecule in IBD [47]. SLC16A1 has been involved in ulcerative colitis and CRC [48,49]. The SLC7A11 has been associated with ulcerative colitis [50].

Table 3
Overview of hub proteins obtained from PPI.

Symbol	Degree	Description	Aspect	References
LYN	91	LYN proto-oncogene, Src family tyrosine kinase	Engaged in activation of CD24-induced ERK1/2 in colorectal cancer; involved in ulcerative colitis-associated colorectal cancer	[30,31]
PLCB1	82	phospholipase C beta 1	Overexpressed in CRC tissues	[32,33]
NPSR1	72	neuropeptide S receptor 1	Gene polymorphism associated with IBD and also involved in 2 major isoforms (NPSR1-A and NPSR1-B) expression alteration in IBD patients; influencers of cancer-related pathways	[34,35]
WNT5A	40	Wnt family member 5A	Wnt low expression of myofibroblast involved in increasing the risk of cancer in ulcerative colitis; Methylated differently in the cell lines of colon cancer; identified as a potential biomarker of CRC	[35–37]
CDC25B	25	cell division cycle 25B	novel distinct colorectal carcinoma prognostic marker	[38]
CD44	25	CD44 molecule (Indian blood group)	Its variants create a connection between cancer-associated glycosylation modifications and CD44 splicing with oncofetal carbohydrate antigens expression; significant impact in human tumor progression and metastasis	[10,39]
RIPK2	20	receptor interacting serine/threonine kinase 2	Implicated in the amplification and development of urothelial cancer risk	[40,41]
ASAP1	16	ArfGAP with SH3 domain, ankyrin repeat and PH domain 1	Its expression is highly upregulated in a multitude of tumors compared to ordinary tissue in human cancer that is also involved in the poor prognosis of patients with colorectal cancer	[41]

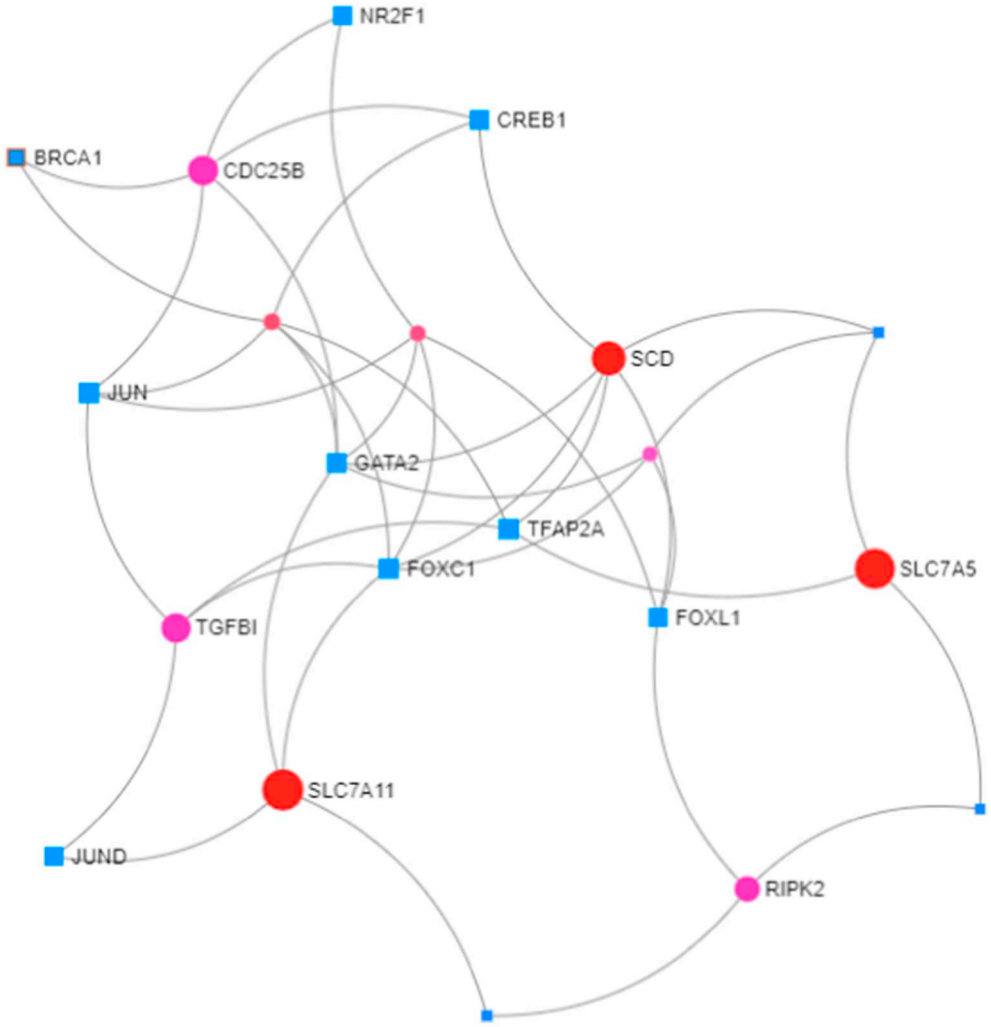

Fig. 4. DEGs-TFs interaction network of common DEGs generated via NetworkAnalyst based on medium confidence score cutoff. The nodes represented with circles indicate TFs and the nodes represented with squares interact with TFs.

Table 4
Top regulatory signatures of common DEGs predicted from DEGs-TFs and DEGs-miRNAs interaction network.

Symbol	Description	Aspect	References
TFs			
SCD	Stearoyl-CoA Desaturase	Defined as new biomarkers for early risk of CRC	[45]
SLC7A5	Solute Carrier Family 7 Member 5	Addressed as a forecast and a therapeutic target for bladder urothelial carcinoma	[46]
IKZF3	IKAROS Family Zinc Finger 3	Recognized as significant CD, UC, and IBD-associated loci	[47]
SLC16A1	Solute Carrier Family 16 Member 1	Involved as butyrate transporter in Ulcerative Colitis; associated with CRC	[48,49]
SLC7A11	Solute Carrier Family 7 Member 11	Tied up in transmembrane transporter activity between CD and UC;	[50]
microRNA			
mir-335-5p	MicroRNA 335	Performed a critical position in the sharing of metabolic pathways between UC and CRC; participated in the development of gastric cancer and pancreatic cancer	[51,52]
mir-26b-5p	MicroRNA 26	Identified as a biomarker for the development of cancerogenesis and gastrointestinal diseases; recognized considerably different with distinct metastatic potential between CRC cell line SW620 and SW480; It has performed a defensive role in the molecular etiology of human breast cancer by promoting apoptosis	[53–55]
mir-124-3p	MicroRNA 124	Involved in liver cancer progression and the apoptosis pathway of patients with colorectal cancer	[56] [57]
mir-16-5p	MicroRNA 16	The burden of CRC and IBD	[58–60]
mir-192-5p	MicroRNA 192	Identified as the aide of SRPX2 expression regulates colon cancer; affiliated with IBD	[61] [62]
mir-548c-3p	MicroRNA 548	Involved in CRC progression	[63]
mir-29b-3p	MicroRNA 29	Detected in CRC as substantially dyregulated	[64]
mir-155-5p	MicroRNA 155	Engaged in regulating Warburg effect in gastric cancer cells	[65,66]
mir-21–5p	MicroRNA 21	Associated with CRC	[67]
mir-15a-5p	MicroRNA 15	Associated with CRC; adversely regulated epithelial junctions in pediatric patients with IBD via Cdc42	[68,69]

Fig. 5. Protein-drug interactions. The interaction between a hub protein and its drugs is depicted.

MicroRNAs are short lengths (~22 nt) RNA molecules that target mRNAs and block the translation of mRNAs. In this way, the miRNAs regulate the DEGs. Nowadays, miRNAs are becoming attractive biomarkers in complex diseases including cancers. Among the identified regulatory miRNAs, mir-335-5p has been proposed as a critical position in the shared metabolic pathways between UC and CRC and also been participated in the development of gastric and pancreatic cancer [51,52]. Mir-26b-5b has been proposed as biomarkers of cancers and gastrointestinal disorders [53–55]. Other identified regulatory miRNAs namely mir-124-3p, mir-16-5p, mir-192-5p, mir-548c-3p, mir-29b-3p, mir-155-5p, mir-21–5p, mir-15a-5p were reported as associated with CRC and gastrointestinal diseases [56–69]. The gene-disease association network analysis revealed the comorbid diseases associated with the hub genes. The ASAP1 associated with bladder cancer and pulmonary tuberculosis. Finally, to identify new drugs targeting with the hub proteins, we studied the protein-drugs interactions and identified cysteine sulfonic acid and double oxidized cysteine drug compounds interact with the hub protein CDC25B. Further investigations are suggested to decipher the roles of these drugs in the treatment of IBD and CRC.

We have utilized bioinformatics and systems biology analyses to study transcriptomes in IBD and CRC to identify key candidate genes and their regulatory molecules. The assessment of networks of DEGs with protein-protein interactions network, DEG-TFs, and DEGs-miRNAs interaction networks analysis revealed a set of genes, TFs, miRNAs. The PPI analysis showed hub proteins namely, LYN, PLCB1, NPSR1, WNT5A, CDC25B, CD44, RIPK2, and ASAP1. The transcriptional and/or post-transcriptional interaction networks analysis showed TFs (SCD, SLC7A5, IKZF3, SLC16A1, SLC7A11), and miRNAs (mir-335-5p, mir-

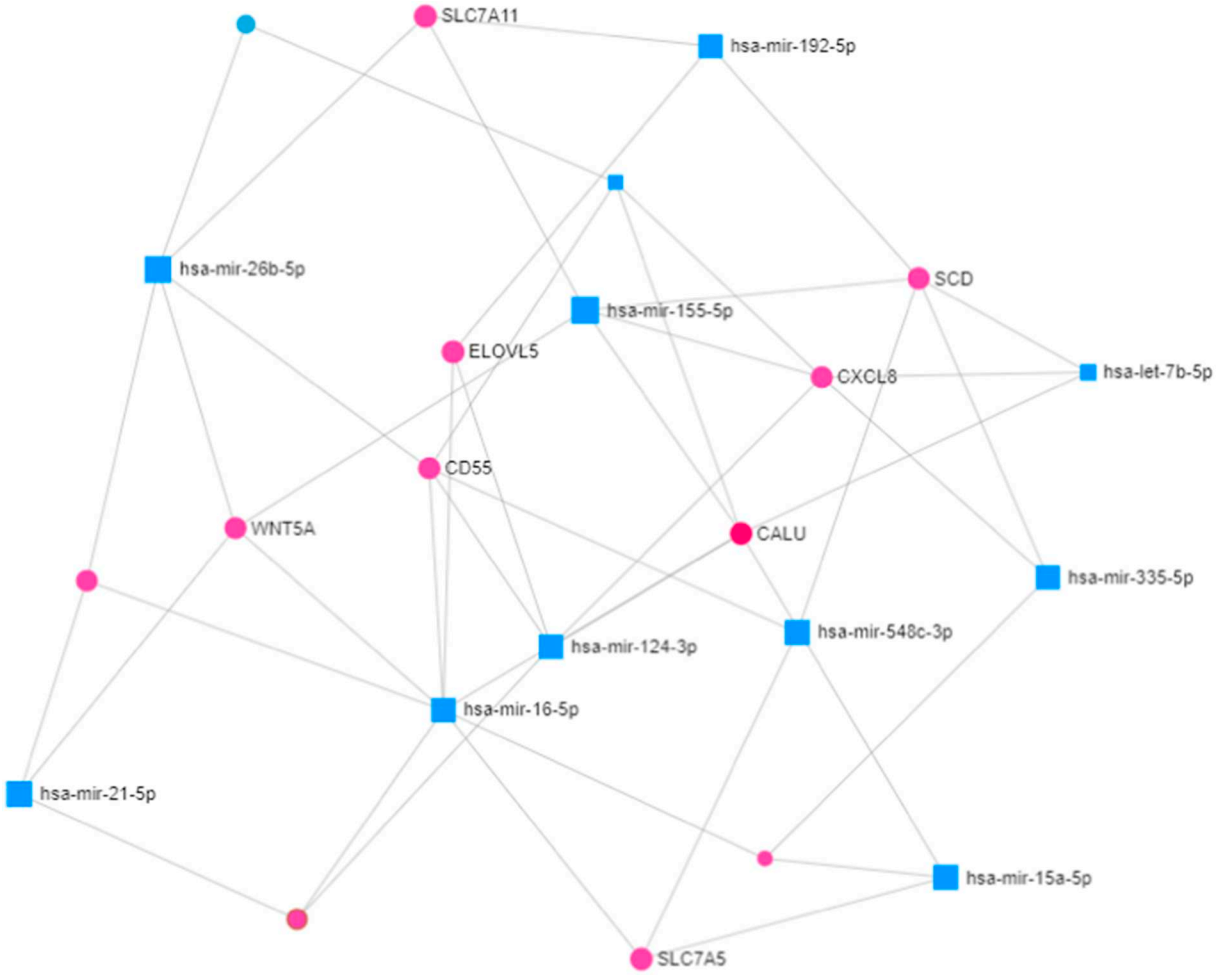

Fig. 6. DEGs-miRNAs interaction network of common DEGs generated via NetworkAnalyst based on medium confidence score cutoff. The circular nodes symbolize the DEGs and the shape of the square circle represents miRNAs.

Fig. 7. Gene disease association network between one hub gene and its associated diseases. The network was generated through NetworkAnalyst.

26b-5p, mir-124-3p, mir-16-5p, mir-192-5p, mir-548c-3p, mir-29b-3p, mir-155-5p, mir-21–5p, mir-15a-5p) in IBD and CRC. The drug molecule cysteine sulfonic acid and double oxidized cysteine interact with the hub protein CDC25B suggesting the further investigation of the molecules for the treatment of IBD and CRC.

References

[1] J.A. Eaden, K.R. Abrams, J.F. Mayberry, The risk of colorectal cancer in ulcerative colitis: A meta-analysis, Gut. 48 (2001) 526–535, https://doi.org/10.1136/gut.48.4.526.

[2] S.C. Ng, H.Y. Shi, N. Hamidi, F.E. Underwood, W. Tang, E.I. Benchimol, R. Panaccione, S. Ghosh, J.C.Y. Wu, F.K.L. Chan, J.J.Y. Sung, G.G. Kaplan, Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies, Lancet. 390 (2017) 2769–2778, https://doi.org/10.1016/S0140-6736(17)32448-0.

[3] A. Hnatyszyn, S. Hryhorowicz, M. Kaczmarek-Ryś, E. Lis, R. Słomski, R.J. Scott, A. Pławski, Colorectal carcinoma in the course of inflammatory bowel diseases, Hered. Cancer Clin. Pract. 17 (2019) 18, , https://doi.org/10.1186/s13053-019-0118-4.

[4] S. Brackmann, S.N. Andersen, G. Aamodt, F. Langmark, O.P.F. Clausen, E. Aadland, O. Fausa, A. Rydning, M.H. Vatn, Relationship between clinical parameters and the colitis-colorectal cancer interval in a cohort of patients with colorectal cancer in inflammatory bowel disease, Scand. J. Gastroenterol. 44 (2009) 46–55, https://doi.org/10.1080/00365520801977568.

[5] M. Svrcek, J. Cosnes, E. Tiret, M. Bennis, Y. Parc, J.F. Fléjou, Expression of epidermal growth factor receptor (EGFR) is frequent in inflammatory bowel disease (IBD)-associated intestinal cancer [4], Virchows Arch. 450 (2007) 243–244, https://doi.org/10.1007/s00428-006-0338-7.

[6] J.E. Hrabe, J.C. Byrn, A.M. Button, G.K. Zamba, M.R. Kapadia, J.J. Mezhir, A matched case-control study of IBD-associated colorectal cancer: IBD portends worse outcome, J. Surg. Oncol. 109 (2014) 117–121, https://doi.org/10.1002/jso.23465.

[7] R.H.J. Golsteijn, C. Bolman, E. Volders, D.A. Peels, H. de Vries, L. Lechner, Development of a computer-tailored physical activity intervention for prostate and colorectal cancer patients and survivors: OncoActive, BMC Cancer 17 (2017) 1–19, https://doi.org/10.1186/s12885-017-3397-z.

[8] M.C. Mattar, D. Lough, M.J. Pishvaian, A. Charabaty, Current management of inflammatory bowel disease and colorectal cancer, Gastrointest. Cancer Res. 4 (2011) 53–61.

[9] Haoming Xu, Mohammad Ali Moni, Pietro Liò, Network regularised cox regression and multiplex network models to predict disease comorbidities and survival of cancer, Comput. Biol. Chem. 59 (2015) 15–31.

[10] J.M. Rhodes, B.J. Campbell, Inflammation and colorectal cancer: IBD-associated and sporadic cancer compared, Trends Mol. Med. 8 (2002) 10–16, https://doi.org/10.1016/S1471-4914(01)02194-3.

[11] H. Jiang, L. Dong, F. Gong, Y. Gu, H. Zhang, D. Fan, Z. Sun, Inflammatory genes are novel prognostic biomarkers for colorectal cancer, Int. J. Mol. Med. 42 (2018) 368–380, https://doi.org/10.3892/ijmm.2018.3631.

[12] J.B. Bramsen, M.H. Rasmussen, H. Ongen, T.B. Mattesen, M.B.W. Ørntoft, S.S. Árnadóttir, J. Sandoval, T. Laguna, S. Vang, B. Øster, P. Lamy, M.R. Madsen, S. Laurberg, M. Esteller, E.T. Dermitzakis, T.F. Ørntoft, C.L. Andersen, Molecular-subtype-specific biomarkers improve prediction of prognosis in colorectal cancer, Cell Rep. 19 (2017) 1268–1280, https://doi.org/10.1016/j.celrep.2017.04.045.

[13] A. Yamada, P. Yu, W. Lin, Y. Okugawa, C.R. Boland, A. Goel, A RNA-sequencing approach for the identification of novel long non-coding RNA biomarkers in colorectal cancer, Sci. Rep. 8 (2018) 1–10, https://doi.org/10.1038/s41598-017-18407-6.

[14] T. Matsuyama, T. Ishikawa, N. Takahashi, Y. Yamada, M. Yasuno, T. Kawano, H. Uetake, A. Goel, Transcriptomic expression profiling identifies ITGBL1, an epithelial to mesenchymal transition (EMT)-associated gene, is a promising recurrence prediction biomarker in colorectal cancer, Mol. Cancer 18 (2019) 1–7, https://doi.org/10.1186/s12943-019-0945-y.

[15] C.R. Heier, A.A. Fiorillo, E. Chaisson, H. Gordish-Dressman, Y. Hathout, J.M. Damsker, E.P. Hoffman, L.S. Conklin, Identification of pathway-specific serum biomarkers of response to glucocorticoid and infliximab treatment in children with inflammatory bowel disease, Clin. Transl. Gastroenterol. 7 (2016) e192, , https://doi.org/10.1038/ctg.2016.49.

[16] A.E. Starr, S.A. Deeke, Z. Ning, C.K. Chiang, X. Zhang, W. Mottawea, R. Singleton, E.I. Benchimol, M. Wen, D.R. Mack, A. Stintzi, D. Figeys, Proteomic analysis of ascending colon biopsies from a paediatric inflammatory bowel disease inception cohort identifies protein biomarkers that differentiate Crohn's disease from UC, Gut. 66 (2017) 1573–1583, https://doi.org/10.1136/gutjnl-2015-310705.

[17] T. Barrett, S.E. Wilhite, P. Ledoux, C. Evangelista, I.F. Kim, M. Tomashevsky, K.A. Marshall, K.H. Phillippy, P.M. Sherman, M. Holko, A. Yefanov, H. Lee, N. Zhang, C.L. Robertson, N. Serova, S. Davis, A. Soboleva, NCBI GEO: archive for functional genomics data sets—update, Nucleic Acids Res. 41 (2012) D991–D995, https://doi.org/10.1093/nar/gks1193.

[18] P. Bardou, J. Mariette, F. Escudié, C. Djemiel, C. Klopp, Jvenn: An interactive Venn diagram viewer, BMC Bioinformatics. 15 (2014), https://doi.org/10.1186/1471-2105-15-293.

[19] M.V. Kuleshov, M.R. Jones, A.D. Rouillard, N.F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S.L. Jenkins, K.M. Jagodnik, A. Lachmann, M.G. McDermott, C.D. Monteiro, G.W. Gundersen, A. Ma'ayan, Enrichr: a comprehensive gene set enrichment analysis web server 2016 update, Nucleic Acids Res. 44 (2016) W90–W97, https://doi.org/10.1093/nar/gkw377.

[20] A. Khan, O. Fornes, A. Stigliani, M. Gheorghe, J.A. Castro-Mondragon, R. Van Der Lee, A. Bessy, J. Chèneby, S.R. Kulkarni, G. Tan, D. Baranasic, D.J. Arenillas, A. Sandelin, K. Vandepoele, B. Lenhard, B. Ballester, W.W. Wasserman, F. Parcy, A. Mathelier, JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework, Nucleic Acids Res. 46 (2018) D260–D266, https://doi.org/10.1093/nar/gkx1126.

[21] J. Xia, E.E. Gill, R.E.W. Hancock, NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data, Nat. Protoc. 10 (2015) 823–844, https://doi.org/10.1038/nprot.2015.052.

[22] S.-D. Hsu, F.-M. Lin, W.-Y. Wu, C. Liang, W.-C. Huang, W.-L. Chan, W.-T. Tsai, G.-Z. Chen, C.-J. Lee, C.-M. Chiu, C.-H. Chien, M.-C. Wu, C.-Y. Huang, A.-P. Tsou, H.-D. Huang, miRTarBase: a database curates experimentally validated microRNA–target interactions, Nucleic Acids Res. 39 (2011) D163–D169, https://doi.org/10.1093/nar/gkq1107.

[23] P. Sethupathy, B. Corda, A.G. Hatzigeorgiou, TarBase: A comprehensive database of experimentally supported animal microRNA targets, RNA. 12 (2006) 192–197, https://doi.org/10.1261/rna.2239606.

[24] D. Szklarczyk, J.H. Morris, H. Cook, M. Kuhn, S. Wyder, M. Simonovic, A. Santos, N.T. Doncheva, A. Roth, P. Bork, L.J. Jensen, C. von Mering, The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible, Nucleic Acids Res. 45 (2017) D362–D368, https://doi.org/10.1093/nar/gkw937.

[25] M.A. Moni, H. Xu, P. Lio, Cytocom: a cytoscape app to visualize, query and analyse disease comorbidity networks, Bioinformatics 31 (6) (2015) 969–971.

[26] M.R. Rahman, T. Islam, T. Zaman, M. Shahjaman, M.R. Karim, F. Huq, J.M.W. Quinn, R.M.D. Holsinger, E. Gov, M.A. Moni, Identification of molecular signatures and pathways to identify novel therapeutic targets in Alzheimer's disease: insights from a systems biomedicine perspective, Genomics. (2019), https://doi.org/10.1016/j.ygeno.2019.07.018.

[27] M.R. Rahman, T. Islam, M. Shahjaman, T. Zaman, H.M. Faruquee, M.A.H.M. Jamal, F. Huq, J.M.W. Quinn, M.A. Moni, Discovering biomarkers and pathways shared by Alzheimer's disease and ischemic stroke to identify novel therapeutic targets, Medicina (B. Aires) 55 (2019) 191, https://doi.org/10.3390/medicina55050191.

[28] D.S. Wishart, Y.D. Feunang, A.C. Guo, E.J. Lo, A. Marcu, R. Grant, T. Sajed, D. Johnson, C. Li, Z. Sayeeda, N. Assempour, I. Iynkkaran, Y. Liu, A. Maciejewski, N. Gale, A. Wilson, L. Chin, R. Cummings, D. Le, A. Pon, C. Knox, M. Wilson, DrugBank 5. 0 : a major update to the DrugBank database for 2018, Nucleic Acids Res. 46 (2018) D1074–D1082, https://doi.org/10.1093/nar/gkx1037.

[29] J. Piñero, À. Bravo, N. Queralt-Rosinach, A. Gutiérrez-Sacristán, J. Deu-Pons, E. Centeno, J. García-García, F. Sanz, L.I. Furlong, DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants, Nucleic Acids Res. 45 (2017) D833–D839, https://doi.org/10.1093/nar/gkw943.

[30] N. Su, L. Peng, B. Xia, Y. Zhao, A. Xu, J. Wang, X. Wang, B. Jiang, Lyn is involved in CD24-induced ERK1/2 activation in colorectal cancer, Mol. Cancer 11 (2012), https://doi.org/10.1186/1476-4598-11-43.

[31] T. Watanabe, T. Kobunai, Y. Yamamoto, H. Ikeuchi, K. Matsuda, S. Ishihara, K. Nozawa, H. Iinuma, T. Kanazawa, T. Tanaka, T. Yokoyama, T. Konishi, K. Eshima, Y. Ajioka, T. Hibi, M. Watanabe, T. Muto, H. Nagawa, Predicting ulcerative colitis-associated colorectal cancer using reverse-transcription polymerase chain reaction analysis, Clin. Colorectal Cancer 10 (2011) 134–141, https://doi.

org/10.1016/j.clcc.2011.03.011.
[32] W.H. Jia, B. Zhang, K. Matsuo, A. Shin, Y.B. Xiang, S.H. Jee, D.H. Kim, Z. Ren, Q. Cai, J. Long, J. Shi, W. Wen, G. Yang, R.J. Delahanty, B.T. Ji, Z.Z. Pan, F. Matsuda, Y.T. Gao, J.H. Oh, Y.O. Ahn, E.J. Park, H.L. Li, J.W. Park, J. Jo, J.Y. Jeong, S. Hosono, G. Casey, U. Peters, X.O. Shu, Y.X. Zeng, W. Zheng, Genome-wide association analyses in east Asians identify new susceptibility loci for colorectal cancer, Nat. Genet. 45 (2013) 191–196, https://doi.org/10.1038/ng.2505.
[33] M. D'Amato, S. Bruce, F. Bresso, M. Zucchelli, S. Ezer, V. Pulkkinen, C. Lindgren, M. Astegiano, M. Rizzetto, P. Gionchetti, G. Riegler, R. Sostegni, M. Daperno, S. D'Alfonso, P. Momigliano-Richiardi, L. Torkvist, P. Puolakkainen, M. Lappalainen, P. Paavola-Sakki, L. Halme, M. Farkkila, U. Turunen, K. Kontula, R. Lofberg, S. Pettersson, J. Kere, Neuropeptide S receptor 1 gene polymorphism is associated with susceptibility to inflammatory bowel disease, Gastroenterology. 133 (2007) 808–817, https://doi.org/10.1053/j.gastro.2007.06.012.
[34] V. Pulkkinen, S. Ezer, L. Sundman, J. Hagström, S. Remes, C. Söderhäll, G. Dario, C. Haglund, J. Kere, J. Arola, Neuropeptide S receptor 1 (NPSR1) activates cancer-related pathways and is widely expressed in neuroendocrine tumors, Virchows Arch. 465 (2014) 173–183, https://doi.org/10.1007/s00428-014-1602-x.
[35] K.R. Hughes, F. Sablitzky, Y.R. Mahida, Expression profiling of Wnt family of genes in normal and inflammatory bowel disease primary human intestinal myofibroblasts and normal human colonic crypt epithelial cells, Inflamm. Bowel Dis. 17 (2011) 213–220, https://doi.org/10.1002/ibd.21353.
[36] Z. Wang, H. Chen, Genistein increases gene expression by demethylation of WNT5a promoter in colon cancer cell line SW1116, Anticancer Res. 30 (2010) 4537–4545.
[37] J. Ying, H. Li, J. Yu, M.N. Ka, F.P. Fan, S.C.C. Wong, A.T.C. Chan, J.J.Y. Sung, Q. Tao, WNT5A exhibits tumor-suppressive activity through antagonizing the WNT/β-catenin signaling, and is frequently methylated in colorectal cancer, Clin. Cancer Res. 14 (2008) 55–61, https://doi.org/10.1158/1078-0432.CCR-07-1644.
[38] I. Takemasa, H. Yamamoto, M. Sekimoto, M. Ohue, S. Noura, Y. Miyake, T. Matsumoto, T. Aihara, N. Tomita, Y. Tamaki, I. Sakita, N. Kikkawa, N. Matsuura, H. Shiozaki, M. Monden, Overexpression of CDC25B phosphatase as a novel marker of poor prognosis of human colorectal carcinoma, Cancer Res. 60 (11) (2000) 3043–3050.
[39] V.J.M. Wielenga, K.H. Heider, G.J.A. Offerhaus, G.R. Adolf, F.M. van den Berg, H. Ponta, P. Herrlich, S.T. Pals, Expression of CD44 variant proteins in human colorectal cancer is related to tumor progression, Cancer Res. 53 (1993) 4754–4756.
[40] M. Guirado, H. Gil, P. Saenz-Lopez, J. Reinboth, F. Garrido, J.M. Cozar, F. Ruiz-Cabello, R. Carretero, Association between C13ORF31, NOD2, RIPK2 and TLR10 polymorphisms and urothelial bladder cancer, Hum. Immunol. 73 (2012) 668–672, https://doi.org/10.1016/j.humimm.2012.03.006.
[41] T. Müller, U. Stein, A. Poletti, L. Garzia, M. Rothley, D. Plaumann, W. Thiele, M. Bauer, A. Galasso, P. Schlag, M. Pankratz, M. Zollo, J.P. Sleeman, ASAP1 promotes tumor cell motility and invasiveness, stimulates metastasis formation in vivo, and correlates with poor survival in colorectal cancer patients, Oncogene. 29 (2010) 2393–2403, https://doi.org/10.1038/onc.2010.6.
[42] S.M.H. Mahmud, W. Chen, H. Jahan, N.I. Sujan, S. Ahmed, iDTi-CSsmoteB : identification of drug – target interaction based on drug chemical structure and protein sequence using XGBoost with over-sampling technique SMOTE, IEEE Access 7 (2019) 48699–48714, https://doi.org/10.1109/ACCESS.2019.2910277.
[43] S.M.H. Mahmud, W. Chen, H. Meng, H. Jahan, Y. Liu, S.M.M. Hasan, Prediction of drug-target interaction based on protein features using undersampling and feature selection techniques with boosting, Anal. Biochem. 589 (2020), https://doi.org/10.1016/j.ab.2019.113507.
[44] M.P. Mosharaf, M.M. Hassan, F.F. Ahmed, M.S. Khatun, M.A. Moni, M.N. Mollah, Computational prediction of protein ubiquitination sites mapping on Arabidopsis thaliana, Comput. Biol. Chem. 85 (2020 Apr 1) 107238.
[45] C.A. Pickens, A. Lane-Elliot, S.S. Comstock, J.I. Fenton, Altered saturated and monounsaturated plasma phospholipid fatty acid profiles in adult males with colon adenomas, Cancer Epidemiol. Biomark. Prev. 25 (2016) 498–506, https://doi.org/10.1158/1055-9965.EPI-15-0696.
[46] J.Y. Liu, Y.B. Dai, X. Li, K. Cao, D. Xie, Z.T. Tong, Z. Long, H. Xiao, M.K. Chen, Y.L. Ye, B. Liu, J. Tan, J. Tang, Z. Xu, Y. Gan, Y.H. Zhou, F. Deng, L.Y. He, Solute carrier family 12 member 5 promotes tumor invasion/metastasis of bladder urothelial carcinoma by enhancing NF-κB/MMP-7 signaling pathway, Cell Death Dis. 8 (2017), https://doi.org/10.1038/cddis.2017.118.
[47] B.D. Ye, D.P.B. McGovern, Genetic variation in IBD: progress, clues to pathogenesis and possible clinical utility, Expert. Rev. Clin. Immunol. 12 (2016) 1091–1107, https://doi.org/10.1080/1744666X.2016.1184972.
[48] V. De Preter, I. Arijs, K. Windey, W. Vanhove, S. Vermeire, F. Schuit, P. Rutgeerts, K. Verbeke, Impaired butyrate oxidation in ulcerative colitis is due to decreased butyrate uptake and a defect in the oxidation pathway*, Inflamm. Bowel Dis. 18 (2012) 1127–1136, https://doi.org/10.1002/ibd.21894.
[49] M. Bianchini, E. Levy, C. Zucchini, V. Pinski, C. Macagno, P. De Sanctis, L. Valvassori, P. Carinci, J. Mordoh, Comparative study of gene expression by cDNA microarray in human colorectal cancer tissues and normal mucosa, Int. J. Oncol. 29 (2006) 83–94, https://doi.org/10.3892/ijo.29.1.83.
[50] O. Palmieri, T.M. Creanza, F. Bossa, O. Palumbo, R. Maglietta, N. Ancona, G. Corritore, T. Latiano, G. Martino, G. Biscaglia, D. Scimeca, M.P. De Petris, M. Carella, V. Annese, A. Andriulli, A. Latiano, Genome-wide pathway analysis using gene expression data of colonic mucosa in patients with inflammatory bowel disease, Inflamm. Bowel Dis. (2015) 1, https://doi.org/10.1097/MIB.0000000000000370.
[51] H.-Y. Huang, Y.-Y. Cheng, W.-C. Liao, Y.-W. Tien, C.-H.J. Yang, S.-M. Hsu, P.-H. Huang, SOX4 transcriptionally regulates multiple SEMA3/Plexin family members and promotes tumor growth in pancreatic cancer, PLoS One 7 (2012) e48637, , https://doi.org/10.1371/journal.pone.0048637.
[52] J. Zhang, Y. Li, C. Zhang, D. Dai, Up-regulation of CRKL by microRNA-335 methylation is associated with poor prognosis in gastric cancer, Cancer Cell Int. 17 (2017) 28, , https://doi.org/10.1186/s12935-017-0387-9.
[53] N. Benderska, A.-L. Dittrich, S. Knaup, T.T. Rau, C. Neufert, S. Wach, F.B. Fahlbusch, M. Rauh, R.M. Wirtz, A. Agaimy, S. Srinivasan, V. Mahadevan, P. Rümmele, E. Rapti, M. Gazouli, A. Hartmann, R. Schneider-Stock, miRNA-26b overexpression in ulcerative colitis-associated carcinogenesis, Inflamm. Bowel Dis. 21 (2015) 2039–2051, https://doi.org/10.1097/MIB.0000000000000453.
[54] Y. Li, Z. Sun, B. Liu, Y. Shan, L. Zhao, L. Jia, Tumor-suppressive miR-26a and miR-26b inhibit cell aggressiveness by regulating FUT4 in colorectal cancer, Cell Death Dis. 8 (2017) e2892, https://doi.org/10.1038/cddis.2017.281.
[55] X.-X. Liu, X.-J. Li, B. Zhang, Y.-J. Liang, C.-X. Zhou, D.-X. Cao, M. He, G.-Q. Chen, J.-R. He, Q. Zhao, MicroRNA-26b is underexpressed in human breast cancer and induces cell apoptosis by targeting SLC7A11, FEBS Lett. 585 (2011) 1363–1367, https://doi.org/10.1016/j.febslet.2011.04.018.
[56] M.L. Slattery, L.E. Mullany, L.C. Sakoda, R.K. Wolff, W.S. Samowitz, J.S. Herrick, Dysregulated genes and miRNAs in the apoptosis pathway in colorectal cancer patients, Apoptosis. 23 (2018) 237–250, https://doi.org/10.1007/s10495-018-1451-1.
[57] G. Koukos, C. Polytarchou, J.L. Kaplan, A. Morley-Fletcher, B. Gras-Miralles, E. Kokkotou, M. Baril-Dore, C. Pothoulakis, H.S. Winter, D. Iliopoulos, MicroRNA-124 regulates STAT3 expression and is down-regulated in colon tissues of pediatric patients with ulcerative colitis, Gastroenterology. 145 (2013), https://doi.org/10.1053/j.gastro.2013.07.001.
[58] E. Danese, A.M. Minicozzi, M. Benati, E. Paviati, G. Lima-Oliveira, M. Gusella, F. Pasini, G.L. Salvagno, M. Montagnana, G. Lippi, Reference miRNAs for colorectal cancer: analysis and verification of current data, Sci. Rep. 7 (2017), https://doi.org/10.1038/s41598-017-08784-3.
[59] B. Baran, N.-M. Ozupek, G. Calibasi-Kocal, Y. Basbinar, Oncog. Carcinog., IntechOpen (Ed.), MicroRNAs (miRNAs) in Colorectal Cancer, 2019, https://doi.org/10.5772/intechopen.80828.
[60] A. Paraskevi, G. Theodoropoulos, I. Papaconstantinou, G. Mantzaris, N. Nikiteas, M. Gazouli, Circulating MicroRNA in inflammatory bowel disease, J. Crohn's Colitis 6 (2012) 900–904, https://doi.org/10.1016/j.crohns.2012.02.006.
[61] J. Zhao, J. Xu, R. Zhang, SRPX2 regulates colon cancer cell metabolism by miR-192/215 via PI3K-AKT, Am. J. Transl. Res. 10 (2018) 483–490.
[62] C.G. Chapman, J. Pekow, The emerging role of miRNAs in inflammatory bowel disease: a review, Ther. Adv. Gastroenterol. 8 (2015) 4–22, https://doi.org/10.1177/1756283X14547360.
[63] T. Liang, L. Guo, C. Liu, Genome-wide analysis of mir-548 gene family reveals evolutionary and functional implications, J. Biomed. Biotechnol. 2012 (2012), https://doi.org/10.1155/2012/679563.
[64] M. Kara, O. Yumrutas, O. Ozcan, O.I. Celik, E. Bozgeyik, I. Bozgeyik, S. Tasdemir, Differential expressions of cancer-associated genes and their regulatory miRNAs in colorectal carcinoma, Gene. 567 (2015) 81–86, https://doi.org/10.1016/j.gene.2015.04.065.
[65] S. Han, S. Yang, Z. Cai, D. Pan, Z. Li, Z. Huang, P. Zhang, H. Zhu, L. Lei, W. Wang, Anti-Warburg effect of rosmarinic acid via miR-155 in gastric cancer cells, Drug Des. Devel. Ther. 9 (2015) 2695–2703, https://doi.org/10.2147/DDDT.S82342.
[66] J. Wan, L. Xia, W. Xu, N. Lu, Expression and function of miR-155 in diseases of the gastrointestinal tract, Int. J. Mol. Sci. 17 (2016) 709, https://doi.org/10.3390/ijms17050709.
[67] M.L. Slattery, J.S. Herrick, L.E. Mullany, W.S. Samowitz, J.R. Sevens, L. Sakoda, R.K. Wolff, The co-regulatory networks of tumor suppressor genes, oncogenes, and miRNAs in colorectal cancer, Genes Chromosom. Cancer 56 (2017) 769–787, https://doi.org/10.1002/gcc.22481.
[68] L.E. Mullany, J.S. Herrick, L.C. Sakoda, W. Samowitz, J.R. Stevens, R.K. Wolff, M.L. Slattery, miRNA involvement in cell cycle regulation in colorectal cancer cases, Genes and Cancer. 9 (2018) 860–872, https://doi.org/10.18632/genesandcancer.167.
[69] W.J. Tang, K.Y. Peng, Z.F. Tang, Y.H. Wang, A.J. Xue, Y. Huang, MicroRNA-15a - cell division cycle 42 signaling pathway in pathogenesis of pediatric inflammatory bowel disease, World J. Gastroenterol. 24 (2018) 5234–5245, https://doi.org/10.3748/wjg.v24.i46.5234.
[70] M.A. Moni, P. Liò, Network-based analysis of comorbidities risk during an infection: SARS and HIV case studies, BMC Bioinformatics. 15 (2014) 333, , https://doi.org/10.1186/1471-2105-15-333.
[71] R. Rahman, M. H, P. Peng, S. S, C. Chen, C. C, L. Lio, P. P, M. Moni, M. A, Genetic effect of type 2 Diabetes to the progression of Neurological Diseases, BioRxiv Bioinforma (2018), https://doi.org/10.1101/480400.
[72] Moni, Mohammad Ali, Pietro Liò, comoR: a software for disease comorbidity risk assessment, Journal of clinical bioinformatics 4.1 (2014) 8.
[73] H.K. Rana, M.R. Akhtar, M.B. Islam, M.B. Ahmed, P. Liò, J.M.W. Quinn, F. Huq, M.A. Moni, Genetic effects of welding fumes on the development of respiratory system diseases, Comput. Biol. Med. 108 (2019) 142–149, https://doi.org/10.1016/j.compbiomed.2019.04.004.
[74] Mohammad Ali Moni, Pietro Liò, How to build personalized multi-omics co-morbidity profiles, Frontiers in cell and developmental biology 3 (2015) 28.
[75] M. Moni, M. A, R. Rana, H. K, I. Islam, M. B, A. Ahmed, M. B, L. Lio, P. P, H. Hasan, H. Huq, Q. Quinn, Early detection of neurological dysfunction using blood cell transcript profiles, BioRxiv Bioinforma (2018) 1–8, https://doi.org/10.1101/483016.
[76] H.K. Rana, M.R. Akhtar, M.B. Islam, M.B. Ahmed, P. Lió, F. Huq, J.M. Quinn, M.A. Moni, Machine learning and bioinformatics models to identify pathways that mediate influences of welding fumes on Cancer progression, Sci. Rep. 10 (2020 Feb

17) 1):1–5.
[77] K.A. Paschos, D. Canovas, N.C. Bird, The role of cell adhesion molecules in the progression of colorectal cancer and the development of liver metastasis, Cell. Signal. 21 (2009) 665–674, https://doi.org/10.1016/j.cellsig.2009.01.006.
[78] J.L. Markman, S.L. Shiao, Impact of the immune system and immunotherapy in colorectal cancer, J. Gastrointest. Oncol. 6 (2015) 208–223, https://doi.org/10.3978/j.issn.2078-6891.2014.077.
[79] M.R. Rahman, T. Islam, E. Gov, B. Turanli, G. Gulfidan, M. Shahjaman, N.A. Banu, M.N.H. Mollah, K.Y. Arga, M.A. Moni, Identification of prognostic biomarker signatures and candidate drugs in colorectal Cancer: insights from systems biology analysis, Medicina (B. Aires). 55 (2019) 20, https://doi.org/10.3390/medicina55010020.
[80] Utpala Nanda Chowdhury, M. Babul Islam, Shamim Ahmad, Mohammad Ali Moni, Network-based identification of genetic factors in ageing, lifestyle and type 2 diabetes that influence to the progression of Alzheimer's disease, Informatics in Medicine Unlocked 19 (2020) 100309, , https://doi.org/10.1016/j.imu.2020.100309.
[81] M.R. Rahman, T. Islam, M.A. Al-Mamun, T. Zaman, M.R. Karim, M.A. Moni, The influence of depression on ovarian cancer: discovering molecular pathways that identify novel biomarkers and therapeutic targets, Informatics Med. Unlocked. 16 (2019), https://doi.org/10.1016/j.imu.2019.100207.