

Article

Identification of prognostic biomarker signatures and candidate drugs in colorectal cancer: Insights from systems biology analysis

Md. Rezanur Rahman^{1,2,#,*}, Tania Islam^{1,#}, Esra Gov³, Beste Turanli⁴, Md. Shahjaman⁵, Nilufa Akhter Banu¹, Md. Nurul Haque Mollah⁶, Kazim Yalcin Arga³ and Mohammad Ali Moni^{7,*}

1Department of Biotechnology and Genetic Engineering, Islamic University, Kushtia, Bangladesh; rezanur12@yahoo.com (M.R.R.), taniaislam1304@gmail.com (T.I.),
2Department of Biochemistry and Biotechnology, School of Biomedical Science, Khwaja Yunus Ali University, Sirajgonj, Bangladesh; rezanur12@yahoo.com (M.R.R.)
3Department of Bioengineering, Adana Science and Technology University, Adana, Turkey; egov@adanabtu.edu.tr (E.G.).
4Department of Bioengineering, Marmara University, Istanbul, Turkey; bcalimlioglu@gmail.com (B.T.); kazim.arga@marmara.edu.tr (K.Y.A.).
5Department of Statistics, Begum Rokeya University, Rangpur, Bangladesh; shahjaman_brur@yahoo.com (M.S.).
6Laboratory of Bioinformatics, Department of Statistics, University of Rajshahi, Rajshahi, Bangladesh; mollah.stat.bio@ru.ac.bd (M.N.H.M.).
7The University of Sydney, Sydney Medical School, School of Medical Sciences, Discipline of Biomedical Science, Sydney, New South Wales, Australia; mohammad.moni@sydney.edu.au (M.A.M.)

[#]These two authors have made an equal contribution and hold joint first authorship for this work.
* Correspondence: E-mail: mohammad.moni@sydney.edu.au (M.A.M.); rezanur12@yahoo.com (M.R.R.); Tel.: +88075163682-4

Abstract: *Background and objectives:* Colorectal cancer (CRC) is the 2nd most cause of cancer related death in the world, but early diagnosis ameliorates the survival of CRC. This report directed to identify molecular biomarker signatures in CRC. *Materials and Methods:* We analyzed two microarray datasets (GSE35279 and GSE21815) to identify common differentially expressed genes (DEGs). We performed functional overrepresentation, pathway enrichment, protein-protein interaction (PPI), reporter biomolecules, survival, and drug repositioning analyses were done on common DEGs. *Results:* Total 727 up-regulated and 99 down-regulated DEGs were detected. The significantly enriched pathways PI3K-Akt signaling, Wnt signaling, ECM-interaction, cell cycles were identified. The 10 hub proteins (ADNP, CCND1, CD44, CDK4, CEBPB, CENPA, CENPH, CENPN, MYC, and RFC2) were selected as proteomic signatures from PPI network. Analyses revealed 10 reporter transcription factors (ETS1, ESR1, GATA1, GATA2, GATA3, AR, YBX1, FDX1, E2F4, and PRDM14) and 2 reporter microRNAs (miR-193b-3p and miR-615-3p) as regulatory component. The prognostic power analysis revealed that hub proteins and reporter biomolecules related with worse survival of patients in CRC. Several candidate repositioned drugs including anti-neoplastic and immunomodulating agents were identified using Connectivity map (CMap) and geneXpharma tool. *Conclusions:* This study presents biomarker signatures at protein and RNA levels with prognostic capability in CRC. We think that the molecular signatures and candidate drugs presented in this study can be potential biomarkers and therapeutic target in CRC.

Keywords: Colorectal cancer; differentially expressed genes; biomarkers; protein-protein interaction; reporter biomolecules; candidate drugs; systems biology; drug repositioning.

1. Introduction

Colorectal cancer (CRC) is frequently related with cancer death in the world [1]. The number of patients died from CRC is still increasing [2]. Like other cancers, various factors such as genetic, cellular, epigenetic alteration, and environmental factors contribute to progression and metastasis of CRC [3,4]. Despite the comprehensive studies on the mechanism of pathogenesis in CRC, the molecular mechanism is unclear. Several biomarkers (KRAS and BRAF) are used to detect the CRC, but these biomarkers are not sufficiently sensitive or specific; consequently pursuit of efficacious biomarkers is an stimulating research arena for researchers in CRC [5]. Despite different studies, the exploration of biomarkers, regulatory patterns, and the identification of therapeutic agents is a great challenge for early diagnosis, prevention, and personalized therapy [5].

The biologists employed microarray technology to find candidate biomarker genes since it helps to understand the genetic alterations and paves the way of biomarker candidates in cancer [6–8]. These biomarkers may be used to early detect the CRC and serve as novel therapeutic target. The hundreds of DEGs have been identified from microarray data, but the results are limited due to heterogeneousness of tissues [9], ignoring the interaction of biomolecules (TFs, miRNAs). These regulatory biomolecules might be attractive biomarkers with advantages. Several reports have been found that miRNAs act as key players in CRC as prognostic biomarkers [10,11].

In this study, systems-based approaches have been considered to explore the potential key biomarker genes, reporter biomolecules (i.e., TFs and miRNAs), and therapeutic agents in CRC (Figure 1). The amalgamated multi-omics analysis of biological information augments the discernment of the molecular mechanism of the diseases [12,13]. We integrated two microarray datasets with pathway, regulatory, functional, and survival analysis to identify putative novel colorectal cancer biomarkers and considering theses biomarkers as therapeutic target, candidate drugs were identified.

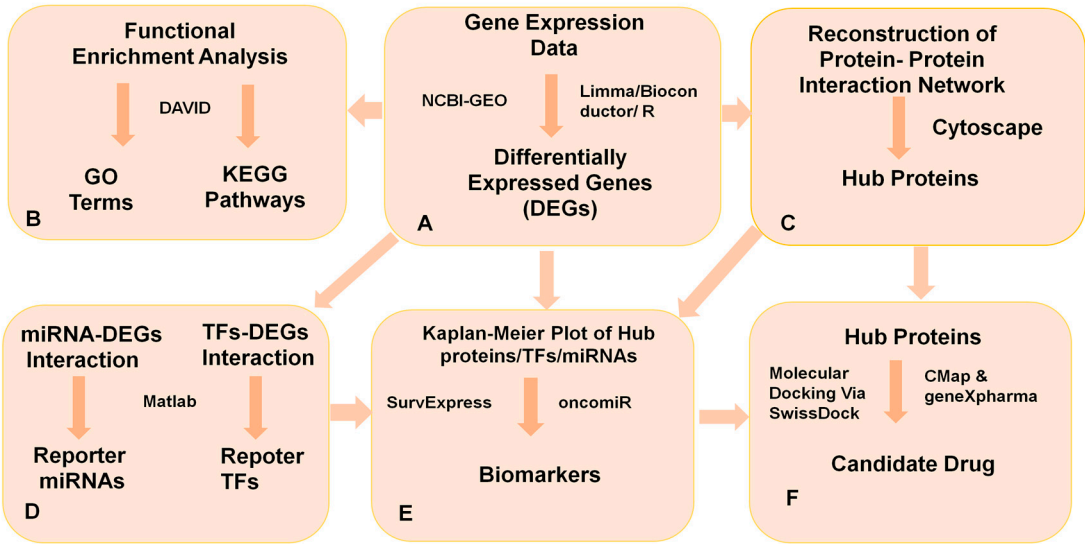


Figure 1: The integrative analytical pipeline employed in the work. (A) The colorectal cancer datasets were analyzed under Bioconductor platform in R. We used limma to detect the DEGs in CRC compared to normal samples. (B) GO terms and molecular pathways were identified by DEGs enrichment via DAVID. (C) The hub proteins were identified by PPI analysis. (D) The reporter feature algorithm used to identify reporter biomolecules transcriptional regulatory elements. (E) The survival analysis of the hub biomolecules through TCGA colorectal cancer datasets via SurvExpress and oncomiR. (F) The candidate drug molecules identified by cMap and geneXpharma.

2. Materials and Methods

2.1 High-throughput Microarray Gene Expression Datasets

We obtained GSE35279 [14] and GSE21815 [15] human CRC microarray gene expression datasets based on the Agilent-014850 Whole Human Genome Microarray platform from NCBI-GEO [16]. In total, 220 specimens (206 CRC specimens and 14 normal samples) were analyzed.

2.2 Differentially Expressed Genes from Microarray Datasets

The raw data were integrated for the analysis in Bioconductor environment (version 3.4.1) in R. Intensity values of each probe-set were log2 transformed, and a Linear Models for Microarray Data (LIMMA) was subsequently performed between CRC samples and matched normal samples. $P < 0.01$ and fold change ≥ 2 were regarded as the cut-off criteria to identify significant DEGs.

2.3 Gene Ontology and Pathway Analysis

We used DAVID bioinformatics resources to obtain GO and KEGG pathways [17]. Fisher's exact test and $p < 0.05$ was considered for all enrichment analyses.

2.4 Protein-protein Interaction Network and Modular Analysis

We used previously reconstructed PPI network of *Homo sapiens* (Karagoz et al., 2016, Islam et al., 2018). The PPI between 21,052 proteins was employed in this report to construct a PPI subnetwork around the proteins encoded by the identified DEGs. The subnetwork was visualized and analyzed via Cytoscape (v3.4 and 2.8.3) [19]. The topological analysis (degree and betweenness) was performed to determine the hub proteins through Cyto-Hubba plugin [20,21] to detect hub proteins. The modules in the PPI sub-network were identified using MCODE plug-in [20]. The modules were further analyzed through enrichment analyses in DAVID.

2.5 Identification of Reporter Biomolecules

We identified transcriptional regulatory biomolecules (i.e., TFs, and miRNAs) from TF-DEGs and miRNA-DEGs interactions from HTRIdb [22] and miRTarbase (Release 6.0) [23] databases, as well as from our previous studies [12,24]. The reporter features algorithm [25] was implemented as described previously [12,24,26], to obtain z-scores and corresponding p values of the molecules. The p-value was corrected via Benjamini-Hochberg's method. An adjusted $p < 0.01$ was regarded significant reporter biomolecules.

2.6 Survival Analysis of Biomolecules

We performed the prognostic analysis of hub proteins, TFs and miRNAs through CRC cohort datasets from TCGA through SurvExpress [27] and OncomiR [28]. Kaplan-Meier plots and a log-rank p -value < 0.05 was regarded significant in all survival analyses.

2.7 Identification of Candidate Drugs

We used simultaneously the Connectivity Map (CMap) database [29] and geneXpharma tool [30] to identify the potential candidate drugs. CMap stores the of expression profiles from cultured human cells exposed to small molecular agents. A total of 50,304 gene-drug interactions comprising 4344 genes and 11,939 drugs were presented in geneXpharma. The hypergeometric probability test was used to find drugs associated with each disease.

3. Results

3.1 Selection of Differentially Expressed Genes

We studied microarray CRC datasets to detect DEGs dydregulated in CRC compared to normal tissues. We detected 727 up-regulated DEGs and 99 down-regulatory DEGs between the two datasets in CRC (Figure 2). Then, we used DAVID to obtain the Gene Ontology and pathways to elucidate the biological roles and significance in regard to molecular functions, biological processes, and cellular components of identified DEGs. Top 5 GO terms for up-regulated and down-regulated DEGs were summarized in Table 1.

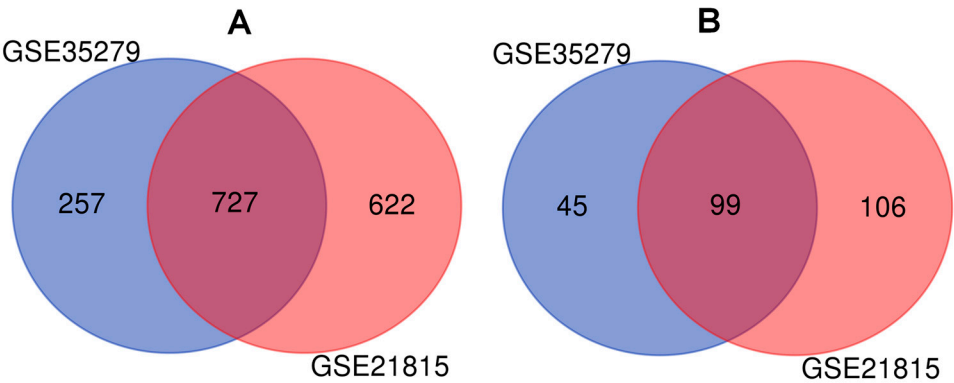


Figure 2: Identification of DEGs in CRC datasets. (A) The up-regulated genes in the CRC expression profiling datasets. (B) The down-regulated genes in the CRC expression profiling datasets.

Table 1. Functional overrepresentation of differentially expressed genes in colorectal cancer.

Gene Ontology	GO term	# of genes	Coverage (%)	P-value
Up-regulated				
Biological Process	collagen fibril organization	11	1.62	4.53E-07
	extracellular matrix organization	22	3.24	2.94E-06
	male gonad development	14	2.06	1.53E-05
	positive regulation of transcription from RNA polymerase II promoter	58	8.56	3.90E-05
	collagen catabolic process	11	1.62	5.07E-05
Cellular Component	Extracellular region	84	12.40	2.4E-05
	cytoplasm	216	31.90	5.8E-05
	extracellular space	70	10.33	0.00015
	basement membrane	11	1.62	0.000256
	extracellular matrix	23	3.39	0.000334
Molecular Function	protein binding	354	52.28	8.10E-08
	protein homodimerization activity	42	6.20	7.54E-04
	growth factor activity	15	2.21	0.00103802
	extracellular matrix binding	6	0.88	0.00146737

	amino acid transmembrane transporter activity	7	1.03	0.00442634
Down-regulated				
Biological Process	bicarbonate transport	5	4.90	5.89E-05
	one-carbon metabolic process	4	3.92	4.00E-04
	chloride transmembrane transport	5	4.90	0.0010592
	nervous system development	7	6.86	0.00262752
	regulation of chloride transport	2	1.96	0.00962447
Cellular Component	plasma membrane	31	30.39	0.01084818
	extracellular space	14	13.72	0.01352057
	integral component of membrane	36	35.29	0.01625726
	anchored component of membrane	4	3.92	0.01792487
	integral component of plasma membrane	13	12.74	0.04206787
Molecular Function	carbonate dehydratase activity	4	3.92	4.16E-05
	hormone activity	5	4.90	0.00123794
	zinc ion binding	15	14.70	0.0017798
	UDP-galactose:beta-N-acetylglucosamine	3	2.94	0.00184167
	beta-1,3-galactosyltransferase activity			
	chloride channel activity	4	3.92	0.0024553

143
144 The significant molecular pathways altered in CRC (Figure 3). The PI3K-Akt signaling
145 pathway, Wnt signaling pathway, cell cycle, lung cancer, ECM-receptor interaction, protein
146 digestion and absorption, pathways in cancer, and TGF-beta signaling pathway were up-regulated
147 CRC (Figure 3A). Contrarily, nitrogen metabolism, pancreatic secretion, axon guidance, retinol
148 metabolism, renin secretion, and chemical carcinogenesis were down-regulated in CRC (Figure 3B).
149

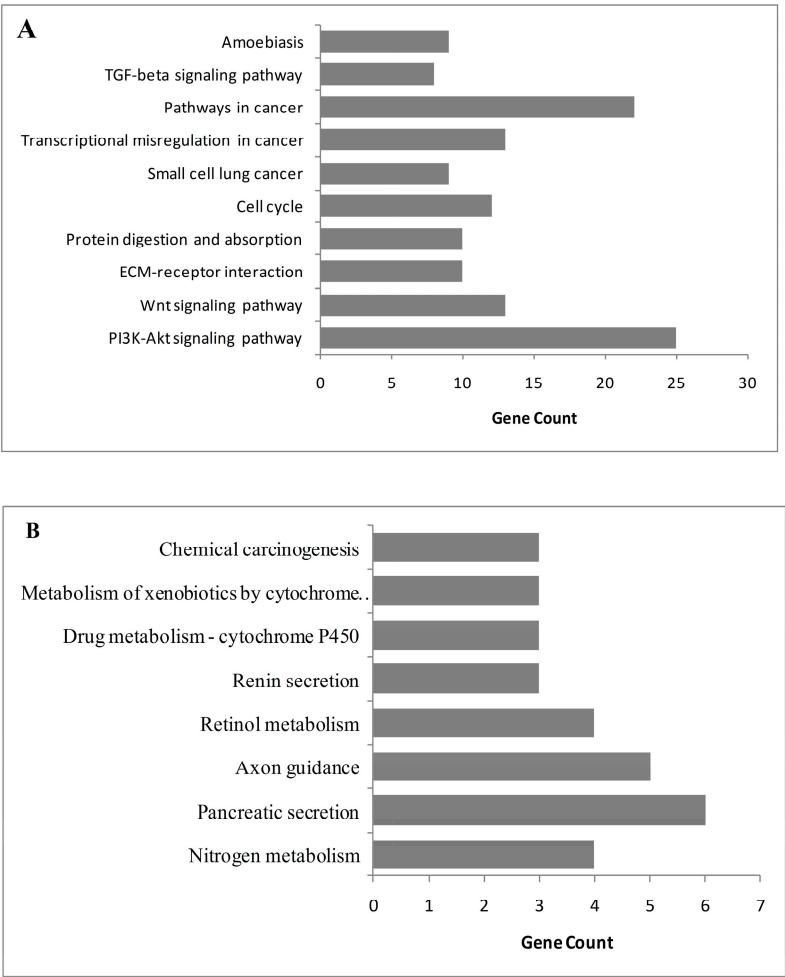


Figure 3: The significant pathway results in CRC. (A) Up-regulated pathways in CRC. (B) Down-regulated pathways in CRC.

3.2 Analysis of Protein-Protein Interaction Network to Identify Central Hub Proteins

To identify central hub proteins, a PPI sub-network was constructed of corresponding proteins of the DEGs, and topological analysis was performed. We detected the presence of 10 hub proteins (ADNP, CCND1, CD44, CDK4, CEBPB, CENPA, CENPH, CENPN, MYC, and RFC2). These hub proteins may play significant key roles in signal transduction during the progression of CRC (Table 2).

Two functional modules were revealed from the PPI. Module 1 had IPO5, RBP2, and RAN nodes. In module 2 CENPN, CENPA, and CENPH nodes were found. The protein import into nucleus, intracellular protein transport, and translocation enriched in biological process by the DEGs of module 1; the nuclear pore and viral process were enriched in cellular processes. The biological processes enriched by the DEGs of module 2 was sister chromatid cohesion, kinetochore assembly, and CENP-A containing nucleosome assembly. The chromosome, centromeric region, nucleoplasm and cytosol were cellular component enriched by the DEGs of module 2.

3.3 Identification of Regulatory Biomolecules

The significant transcriptional regulatory components were obtained. We detected 10 common TFs (ETS1, ESR1, GATA1, GATA2, GATA3, AR, YBX1, FOXP3, E2F4, and PRDM14) in CRC. Moreover, 10 common reporter miRNAs (miR-16-5p, miR-26b-5p, miR-124-3p, let-7b-5p, miR-92a-3p, miR-192-5p, miR-155-5p, miR-93-5p, miR-193b-3p, and miR-17-5p) were selected between the two datasets (Table 3).

Table 2. Summary of hub proteins in colorectal cancer

Symbol	Description	Feature
<i>Hub proteins</i>		
ADNP	activity dependent neuroprotector homeobox	Stimulatory and inhibitory effect on the growth of tumor cells
CEBPB	CCAAT/enhancer-binding protein beta	Involved in immune and inflammatory responses
CCND1	Cyclin D1 (afflicted with cancers colonic adenocarcinomas , myeloma)	Cell cycle regulatory protein
CD44	CD44 molecule	Required in cell-cell interactions, migration
CDK4	Cyclin Dependent Kinase 4	Cyclin D1 activates <i>CDK4</i> , which causes proliferation of cellular division.
CENPA	Centromere protein A (afflicted with colorectal cancer)	Central role in the assembly of kinetochore
CENPH	Centromere Protein H (afflicted with colorectal cancer)	Central role in assembly of kinetochore proteins
RFC2	Replication factor C subunit 2	Encodes activator 1 small subunits family
MYC	MYC Proto-Oncogene	Regulator gene contributes to formation of many human cancers
CENPN	Centromere Protein N	Involved in cell cycle process

Table 3. Summary of reporter biomolecules in colorectal cancer.

Symbol	Description	Feature
<i>Reporter Transcription Factors</i>		
AR	Androgen receptor	Involved in prostate cancer
GATA1	GATA Binding Protein 1	Transcriptional activator or repressor
GATA2	GATA Binding Protein 2 (afflicted with colorectal cancer)	Transcriptional activator
GATA3	GATA Binding Protein 3	Transcriptional activator
E2F4	E2F Transcription Factor 4	Controls of cell cycle
ETS1	ETS Proto-Oncogene 1	Involved in tumorigenesis
YBX1	Y-Box Binding Protein 1	Aberrant expression is associated with cancer
PRADM14	PR/SET Domain 14	Involved in breast cancer
ESR1	Estrogen Receptor 1	Involved in breast cancer
FOXP3	Forkhead Box P3 (afflicted with colorectal cancer)	DNA binding
<i>Reporter microRNAs</i>		
miR-193b-3p	MicroRNA 193	Afflicted with CRC and epidermal squamous cell carcinoma
miR-615-3p	MicroRNA 615	Afflicted with CRC
miR-16-5p	MicroRNA 16	Potential biomarkers in gastric cancer
miR-26b-5p	MicroRNA 26	Afflicted with CRC
let-7b-5p	MicroRNA 7	Afflicted with CRC
miR-92a-3p	MicroRNA 92	Afflicted with CRC
miR-124-3p	MicroRNA 124	Afflicted with CRC, gastric and breast cancer
miR-484	MicroRNA 484	Afflicted with CRC
miR-192-5p	MicroRNA 192	Afflicted with CRC
miR-93-5p	MicroRNA 93	Afflicted with head and neck cancer

3.4 Survival Analysis of Biomolecules

We performed the survival analysis of biomolecules (i.e., 10 hubs, 10 TFs, and 10 miRNAs) using datasets from TCGA. The differential expression levels were represented by the box-plot and the Kaplan-Meier plots represents the survival probabilities of the biomolecules. The 10 hub proteins were found statistically significant in their prognostic capability (log-rank $p < 0.05$) (Figure 4). The 10 reporter TFs were found significant in their prognostics capability assessment (Figure 5). However, the survival analysis of the reporter miRNAs showed miR-193b-3p and miR-615-3p were statistically significant in survival analysis. The prognostic power of these two miRNAs showed in Figure 6.

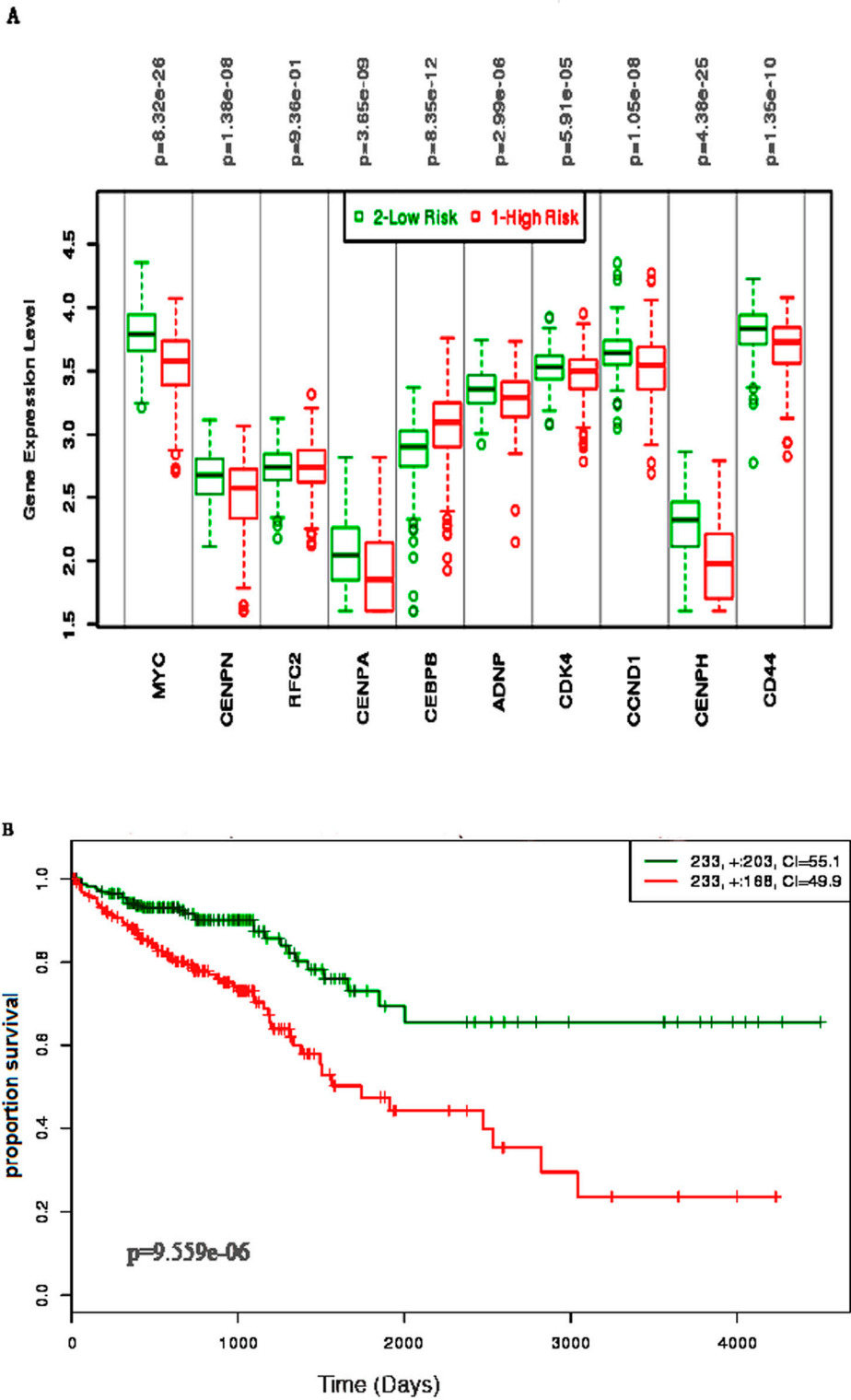
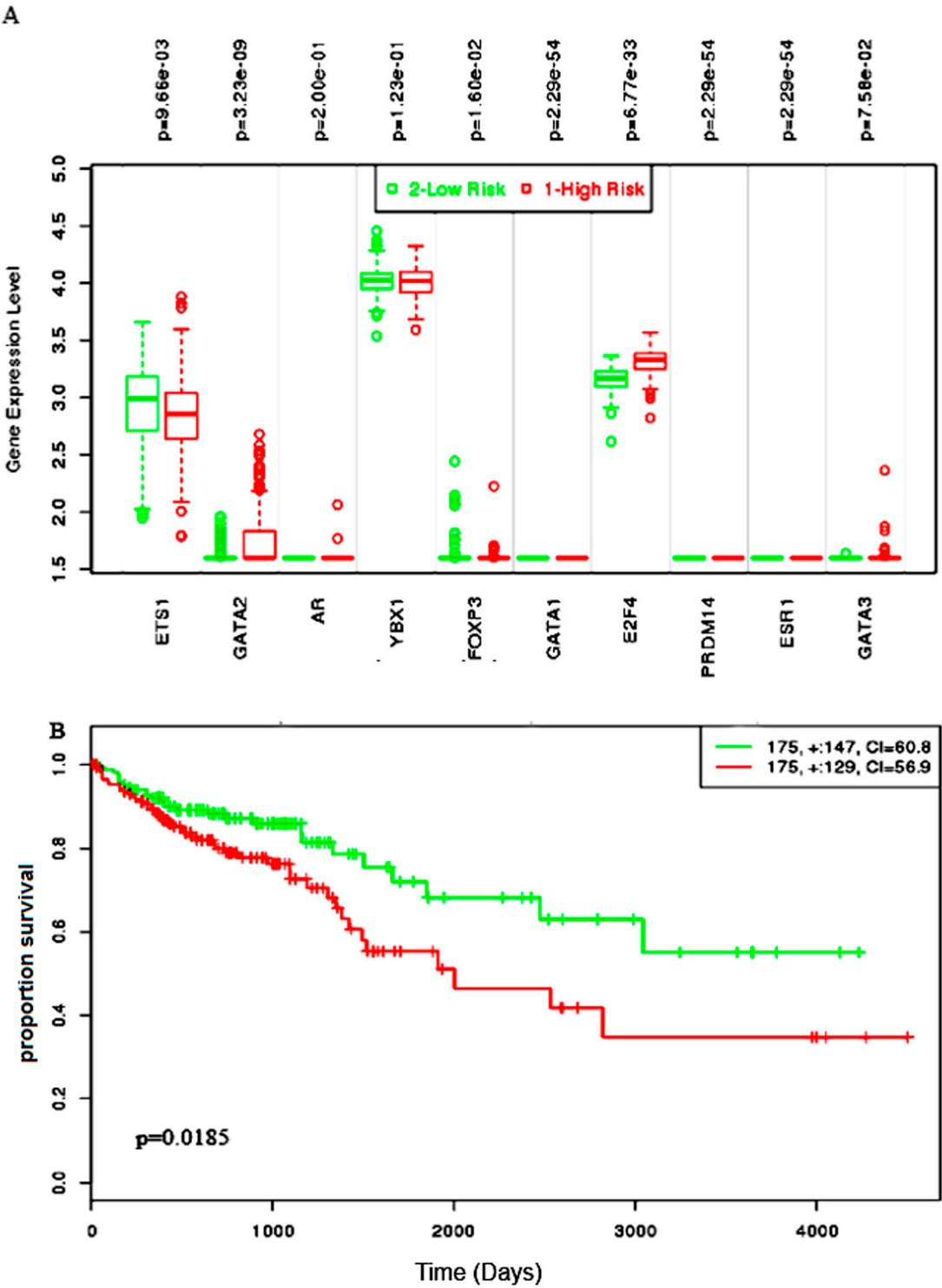


Figure 4: The survival analysis of the hub genes in the prognosis of colorectal cancer. (A) The box plot represents the differential expression of the 10 hub genes. (B) Kaplan-Meier plot represents the prognostic ability of the 10 hub gene signatures in CRC.

212



213

214 Figure 5: The survival assessment of the reporter TFs in the prognosis of colorectal cancer. (A) The box
215 plot represents the differential expression of the 10 TFs. (B) Kaplan-Meier plot represents the prognostic power
216 of the 10 TFs signatures in CRC.

217

218

219

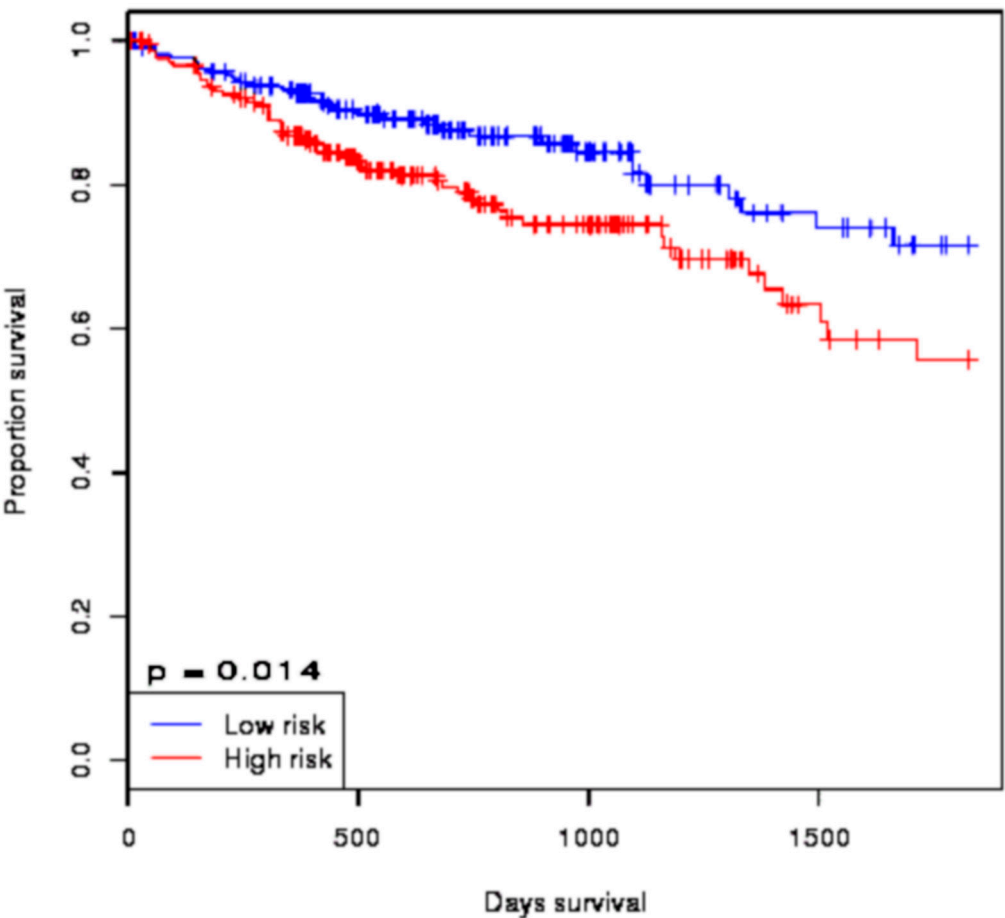


Figure 6: The survival analysis of the reporter miRNAs in colorectal cancer. Kaplan-Meier plot represents the prognostic ability of miRNA signatures (miR-193b-3p and miR-615-3p) in CRC.

3.5 Identification of Candidate Drugs through Drug Repositioning

The hub proteins and TFs were regarded as drug targets. We identified potential drugs based on the transcriptome signatures guided using drug repositioning tool geneXpharma and CMap database. We considered only the common drugs between two databases for CRC. Statistical evaluation revealed 45 candidate drugs targeting 6 proteins (Table 4). The drugs were classified according to the anatomical sites and development stages (Figure 7). Among the 10 hub proteins considered as a drug target, 3 hub proteins i.e., CCND1, CDK4, MYC were targeted by 9 drugs (Table 4). Contrarily, among the 10 reporters TFs, 3 reporter TFs were targeted by 23 drugs (Table 4). The repositioned drugs were classified based on the Anatomical Therapeutic Chemical classification system and found that 16.12% were antineoplastic, 22.58 % were antineoplastic and immunomodulating agents. The hormones and contraceptives agents were (9.67%) were following the antineoplastic and immunomodulating agents (Figure 7). The repositioned drugs were analyzed and found that 49% of drugs were approved (Figure 7), whereas 48% were still under investigation and 3% were in the experimental stage.

Table 4. Selected repositioned drugs in colorectal cancer

Target	Repositioned Drug	Drug Class/Status/Description
<i>Hub protein</i>		
CCND1	Gefitinib	Antineoplastic Agents/ Approved, Investigational/ used in the treatment of cancer
	Hydrocortisone	Anti-Inflammatory Agents/Approved/ used in the treatment of inflammation, allergy, collagen diseases, asthma, and some neoplastic conditions
	Irinotecan	Antineoplastic Agents/Approved, Investigational/ used in the treatment of colorectal cancer
	Letrozole	Antineoplastic Agents/Approved, Investigational/ introduced for treatment of breast cancer
	Lidocaine	Anesthetics/Approved/ A local anesthetic and used as an antiarrhythmia agent
	Methotrexate	Antimetabolites, Antineoplastic/Approved/ antineoplastic antimetabolite with immunosuppressant properties
	Sirolimus	Antineoplastic and Immunomodulating Agents/Approved, Investigational/ a potent immunosuppressant and possesses both antifungal and antineoplastic properties
	Tamoxifen	Antineoplastic and Immunomodulating agents/Approved/ for the treatment and prevention of breast cancer
CDK4	Gefitinib	Antineoplastic Agents/ Approved, Investigational/ used in the treatment of cancer
	Lidocaine	Anesthetics/Approved/ local anesthetic and used as an antiarrhythmia agent
	Sirolimus	Antineoplastic and Immunomodulating Agents/Approved, Investigational/ a potent immunosuppressant and possesses both antifungal and antineoplastic properties
MYC	Gefitinib	Antineoplastic Agents/ Approved, Investigational/ used in the treatment of cancer
	Tamoxifen	Antineoplastic and Immunomodulating Agents/Approved/ for the treatment and prevention of breast cancer
	Simvastatin	Cardiovascular System/Approved/ a lipid-lowering agent
<i>Reporter TFs</i>		
GATA3	Azathioprine	Antineoplastic and Immunomodulating agents/ Approved/ immunosuppressive antimetabolite pro-drug
	Daunorubicin	Antineoplastic and Immunomodulating Agents/ Approved/ used in treatment of leukemia and other neoplasms
	Dexamethasone	Antineoplastic Agents/Approved, Investigational, Vet approved/ for the

		treatment of endocrine disorders, rheumatic disorders, collagen diseases, dermatologic diseases
	Doxorubicin	Antineoplastic and Immunomodulating agents/Approved, Investigational/used neoplastic conditions like acute lymphoblastic leukemia,
	Mercaptopurine	antimetabolite antineoplastic agent with immunosuppressant properties/ Approved/ in the treatment of leukemia
	Methotrexate	Antimetabolites, Antineoplastic/Approved/antineoplastic antimetabolite with immunosuppressant properties
ESR1	Clomifene	Estrogen Agonist/Antagonist/Approved, Investigational/ used mainly in female infertility due to anovulation to induce ovulation
	Daunorubicin	Antineoplastic and Immunomodulating Agents/ Approved/ used in treatment of leukemia and other neoplasms
	Dexamethasone	Antineoplastic Agents/Approved, Investigational/ for the treatment of endocrine disorders, rheumatic disorders, collagen diseases, dermatologic diseases
	Estriol	Estradiol Congeners/Approved, Investigational/ used as a test to determine the general health of an unborn fetus
	Estrone	Hormones/Approved/ used for management of perimenopausal and postmenopausal symptoms
	Etoposide	Antineoplastic Agents/Approved/ used in the treatment of refractory testicular tumors and in patients with small cell lung cancer
	Fulvestrant	Antineoplastic and Immunomodulating Agents/Approved, Investigational/ a drug treatment of metastatic breast cancer
	Glibenclamide	Oral Hypoglycemics/Approved/ used for the treatment of non-insulin-dependent diabetes mellitus
	Imipramine	Central Nervous System agents/Approved/ antidepressant used for the relief of symptoms of depression
	Letrozole	Antineoplastic agents/Approved, Investigational/ introduced for the treatment of breast cancer
	Megestrol	Antineoplastic and Immunomodulating Agents/Approved, Investigational/used in the palliative treatment of breast cancer
	Mifepristone	Abortifacient Agents and Blood Glucose Lowering Agents/Approved, Investigational/ For the medical termination of intrauterine pregnancy. Also indicated to control hyperglycemia
	Progesterone	Contraceptive Agents/Approved, Vet approved/ Progesterone acts on the uterus, the mammary glands, and the brain
	Raloxifene	Estrogen Agonist/Antagonist/Approved, Investigational/ used to prevent osteoporosis in postmenopausal women
	Tamoxifen	Antineoplastic and Immunomodulating Agents/Approved/ for the treatment and prevention of breast cancer
	Testosterone	Androgens and Estrogens/Approved, Investigational/ In men,

		testosterone is produced primarily by the Leydig cells of the testes. Testosterone in women functions to maintain libido and general wellbeing
AR	Cyproterone	Antineoplastic Agents and Hormone Antagonists/Approved, Investigational/ It is used in the treatment of hypersexuality in males, as a palliative in prostatic carcinoma
	Flufenamic acid	Antiinflammatory and Antirheumatic /Experimental/ analgesic, anti-inflammatory, and antipyretic properties
	Flutamide	Antineoplastic Agents, Hormonal/Approved, Investigational/ For the management of metastatic carcinoma of the prostate
	Levonorgestrel	Contraceptive Agents/Approved, Investigational/ For the treatment of menopausal and postmenopausal disorders
	Mifepristone	Abortifacient Agents and Blood Glucose Lowering Agents/Approved, Investigational/ For the medical termination of intrauterine pregnancy. Also indicated to control hyperglycemia
	Spironolactone	Agents causing hyperkalemia /Approved/ Used primarily to treat low-renin hypertension, hypokalemia, and Conn's syndrome
	Testosterone	Androgens and Estrogens/Approved, Investigational/ In men, testosterone is produced primarily by the interstitial cells of the testes. Functions to maintain libido and general wellbeing in women.

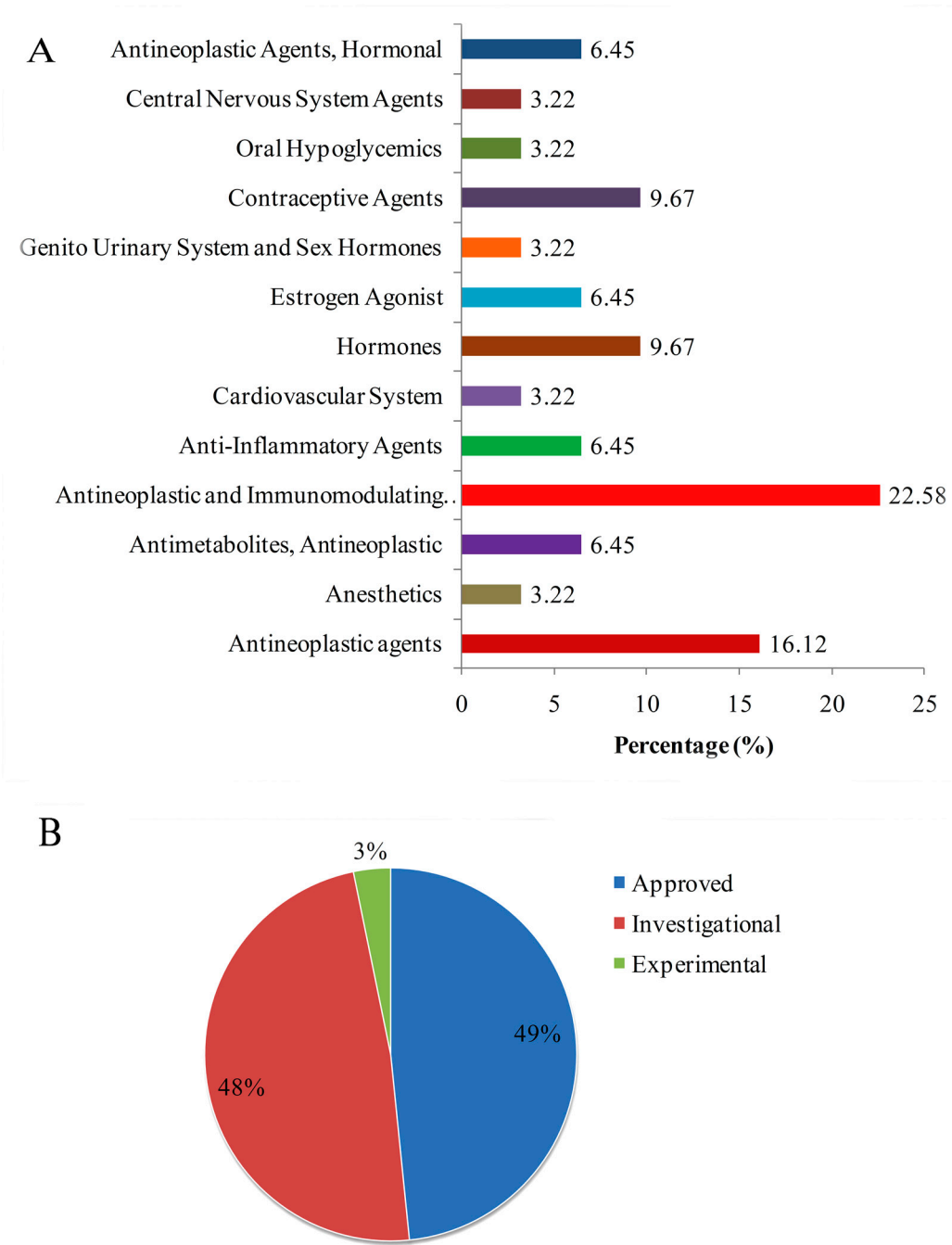


Figure 7: Drug repositioning results. (A). Classification of repurposed drugs according to drug development stages. (B) Distribution of approved drugs into anatomical therapeutic chemical drug classes.

4. Discussion

CRC is complex disease and the mechanisms of the CRC still lacking. The augmenting effect of genetic, endocrinological perturbations, and epigenetic aberrations contribute to the pathobiology of CRC [4]. The highthroughput microarray technology considered as the primary sources of candidate biomarkers [32]. Understanding the disease pathways and exploration of biomarkers requires interomics data from several levels [12]. Cosequently, we employed a system biomedicine approach to explore the in-depth mechanism of CRC in this manuscript.

We identified 727 upregulated and 99 down-regulated DEGs from microarray CRC datasets. The pathway enrichment analyses revealed significant molecular pathways. Wnt signaling pathway and inflammatory signaling pathways is implicated in CRC [33]. TGF- β pathway has dual role

behaves as tumor suppressor and tumor promoter. In different cancers, the TGF- β was proposed as a target for cancer therapy [36]. Therefore, we think these pathways altered during the progression of the CRC and they might be therapeutic target for CRC.

Analysis of the PPI provides insights into central mechanisms on the pathobiology of cancers [37]. The PPI networks were reconstructed in order to clarify the interaction among the identified DEGs. 10 hub genes came into prominence as principal contributor in PPI. The prognostic survival analysis showed that these hub genes are significantly associated with the survival outcome of the CRC patient (Figure 5). Among the hub protein, ADNP was dysregulated in CRC with high WNT activity [38]; CEBPB is afflicted with colorectal cancer and glioblastoma cells [39,40]; CCND1 dysregulations is implicated in CRC [41,42]; CD44 plays diverse roles in cancer cells; the CDK4 is the target for different cancer treatment including colorectal cancer [43,44]; CENPA is implicated in CRC [45]; CENPH was implicated in CRC [46]; RFC2 is implicated in hematologic cancers [47,48]; MYC is dysregulated in CRC [49–51]. CENPN is a protein that in humans is involved in cell cycle process showing direct binding of CENPN to CENPA [52]. The modules significantly contained the nodes (i.e., CENPA, CENPN, and CENPH) which are associated with different cancer and disease progression as discussed above.

Significant TFs regulating the DEGs were obtained by analyzing networks. Among the reporter TF, AR is dysregulated in the prostate cancer [53]; ETS is involved in a different type of cancers [54]; GATA2 is deregulated in CRC with poor survival outcomes [55]; GATA3 and GATA4 was proposed to be implicated in different cancers [56]; YBX1 and FOXP3 are markers of cancers [57–59]; the E2F4 disruption is involved in cancers [60,61]; the dysregulation of PRDM14 and ESR1 are found in breast cancers [62–64].

Expression of 500 miRNAs is mentioned in CRC [5]. Thus, we evaluated the biomarker potentiality of the miRNAs in CRC since they regulate genes involved in cell cycle [10,65,66]. We identified relevant miRNAs (miR-193b-3p and miR-615-3p) and survival analysis showed the significant potential to biomarkers in CRC. Recently, Wu et al., found that dysregulation of miR-193b-3p affects the growth of CRC via TGF-beta and regulation of SMAD signaling pathway [67]. Our pathway enrichment results also showed the dysregulation of TGF-beta signaling pathway. Moreover, miR-193b-3p is a predictive biomarkers of renal cell carcinoma [68]. The high expression of miR-615-3p was associated with pathogenesis of CRC and gastric cancer [69,70]. Researches on these miRNAs might provide therapeutic target for CRC.

Then, survival analysis of the hub genes, TFs, and miRNA clarified those hub gene signatures (MYC, CENPN, RFC, CENPA, CEBPB, ADNP, CDK4, CCND1, CENPH and CD44) showed potentiality of being prognostic biomarkers in CRC. It was found high expression of reporter TFs signatures (AR, GATA1, GATA2, GATA3, EST1, YBX1, PRADM14, ESR1, E2F4, and FOXP3) were associated with worse survival outcomes of the CRC patients. The survival analysis of the miRNAs (miR-193b-3p and miR-615-3p) showed significant prognostic power in CRC. Despite the tremendous significance of the computational finding of this present works, some basic biology experiment such as western blot, qRT-PCR, CRISPR/Cas9 gene editing may be done for further confirmation of above results.

The present study identified 45 candidate drugs (Table 4). These drugs were mostly antineoplastic, antidiabetic, and endocrinologicals drugs. These identified repositioned candidate drugs showed significant potentiality in the treatment of CRC.

5. Conclusions

We employed the systems biology analyses where transcriptome datasets were incorporated with molecular networks to reveal molecular biomarker signatures at RNA and protein levels which showed significant potentiality in CRC. The prognostic survival analysis of these reporter biomolecules performed in TCGA datasets revealed proteomic codes (MYC, CENPN, RFC, CENPA, CEBPB, ADNP, CDK4, CCND1, CENPH and CD44); TFs signatures (AR, GATA1, GATA2, GATA3, EST1, YBX1, PRADM14, ESR1, E2F4, and FOXP3); miRNAs signatures (miR-193b-3p and miR-615-3p) as prognostic biomarker signatures in CRC. In addition, candidate repositioned drugs

were described in CRC. The identified biomolecules and candidate drugs identified in this study deserve basic biology studies since they show importance as candidate biomarkers and therapeutics for precision medicine approaches to treat CRC.

Author Contributions: Conceptualization: Md. Rezanur Rahman, Kazim Yalcin Arga, and Mohammad Ali Moni; Formal analysis, Md. Rezanur Rahman, Tania Islam, Esra Gov, Beste Turanli, and Md. Shahjaman; Investigation, Md. Rezanur Rahman and Tania Islam; Methodology, Esra Gov, Md. Shahjaman, Kazim Yalcin Arga and Mohammad Ali Moni; Supervision, Nilufa Akhter Banu, Md. Nurul Haque Mollah and Mohammad Ali Moni; Writing-original draft, Md. Rezanur Rahman; Writing-review & editing, Esra Gov, Kazim Yalcin Arga and Mohammad Ali Moni

Funding: This article did not receive no external funding.

Acknowledgments: We would like to express our thanks to the Department of Biotechnology and Genetic Engineering, Islamic University, Kushtia, Bangladesh, and Laboratory of Bioinformatics, Department of Statistics, University of Rajshahi, Bangladesh for providing the bioinformatics laboratory facility. This work was supported by Islamic University research grant.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RGS, Barzi A, et al. Colorectal cancer statistics, 2017. *CA Cancer J Clin* 2017;67:177–93. doi:10.3322/caac.21395.

[2] Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. *Gut* 2017;66:683–91. doi:10.1136/gutjnl-2015-310912.

[3] Markowitz SD, Bertagnolli MM. Molecular Basis of Colorectal Cancer. *N Engl J Med* 2013;361:2449–60.

[4] Kheirleiseid E a H, Miller N, Kerin MJ. Molecular biology of colorectal cancer : Review of the literature. *Am J Mol Biol* 2013;3:72–80. doi:10.4236/ajmb.2013.32010.

[5] Zarkavelis G, Boussios S, Papadaki A, Katsanos KH, Christodoulou DK, Pentheroudakis G. Current and future biomarkers in colorectal cancer. *Ann Gastroenterol* 2017;30:613–21. doi:10.20524/aog.2017.0191.

[6] Krizkova S, Kepinska M, Emri G, Rodrigo MAM, Tmejova K, Nerudova D, et al. Microarray analysis of metallothioneins in human diseases--A review. *J Pharm Biomed Anal* 2016;117:464–73. doi:10.1016/j.jpba.2015.09.031.

[7] Stewart JP, Richman S, Maughan T, Lawler M, Dunne PD, Salto-Tellez M. Standardising RNA profiling based biomarker application in cancer-The need for robust control of technical variables. *Biochim Biophys Acta* 2017;1868:258–72. doi:10.1016/j.bbcan.2017.05.005.

[8] Kamel HFM, Al-Amodi HSAB. Exploitation of Gene Expression and Cancer Biomarkers in Paving the Path to Era of Personalized Medicine. *Genomics, Proteomics Bioinforma* 2017;15:220–35. doi:10.1016/j.gpb.2016.11.005.

[9] Isella C, Terrasi A, Bellomo SE, Petti C, Galatola G, Muratore A, et al. Stromal contribution to the colorectal cancer transcriptome. *Nat Genet* 2015;47:312–9. doi:10.1038/ng.3224.

[10] Masuda T, Hayashi N, Kuroda Y, Ito S, Eguchi H, Mimori K. MicroRNAs as biomarkers in colorectal cancer. *Cancers (Basel)* 2017;9. doi:10.3390/cancers9090124.

[11] Michael MZ, O’ Connor SM, van Holst Pellekaan NG, Young GP, James RJ. Reduced accumulation of specific microRNAs in colorectal neoplasia. *Mol Cancer Res* 2003;1:882–91. doi:10.1002/bies.10046.

[12] Islam T, Rahman MR, Gov E, Turanli B, Gulfidan G, Haque MA, et al. Drug Targeting and Biomarkers

in Head and Neck Cancers : OMICS 2018;22:422–36. doi:10.1089/omi.2018.0048.

[13] Barabási A-L, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. Nat Rev Genet 2011;12:56–68. doi:10.1038/nrg2918.

[14] Kagawa Y, Matsumoto S, Kamioka Y, Mimori K, Naito Y, Ishii T, et al. Cell cycle-dependent Rho GTPase activity dynamically regulates cancer cell motility and invasion in vivo. PLoS One 2013;8. doi:10.1371/journal.pone.0083629.

[15] Kogo R, Shimamura T, Mimori K, Kawahara K, Imoto S, Sudo T, et al. Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. Cancer Res 2011;71:6320–6. doi:10.1158/0008-5472.CAN-11-1021.

[16] Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: Archive for functional genomics data sets - Update. Nucleic Acids Res 2013;41:991–5. doi:10.1093/nar/gks1193.

[17] Smyth GK, Ritchie M, Thorne N. Linear Models for Microarray Data User ' s Guide. Bioinformatics 2011;20:3705–6. doi:10.1093/nar/gkv007.

[18] Karagoz K, Sevimoglu T, Arga KY. Integration of multiple biological features yields high confidence human protein interactome. J Theor Biol 2016;403:85–96. doi:10.1016/j.jtbi.2016.05.020.

[19] Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. Cytoscape 2.8: New features for data integration and network visualization. Bioinformatics 2011;27:431–2. doi:10.1093/bioinformatics/btq675.

[20] Chin C, Chen S, Wu H. cyto-Hubba: A Cytoscape Plug-in for Hub Object Analysis in Network Biology. Genome Informatics ... 2009;5:2–3.

[21] Calimlioglu B, Karagoz K, Sevimoglu T, Kilic E, Gov E, Arga KY. Tissue-Specific Molecular Biomarker Signatures of Type 2 Diabetes: An Integrative Analysis of Transcriptomics and Protein-Protein Interaction Data. OMICS 2015;19:563–73. doi:10.1089/omi.2015.0088.

[22] Bovolenta LA, Acencio ML, Lemke N. HTRIdb: an open-access database for experimentally verified human transcriptional regulation interactions. BMC Genomics 2012;13. doi:10.1186/1471-2164-13-405.

[23] Hsu S Da, Tseng YT, Shrestha S, Lin YL, Khaleel A, Chou CH, et al. MiRTarBase update 2014: An information resource for experimentally validated miRNA-target interactions. Nucleic Acids Res 2014;42:78–85. doi:10.1093/nar/gkt1266.

[24] Ayyildiz D, Gov E, Sinha R, Arga KY. Ovarian Cancer Differential Interactome and Network Entropy Analysis Reveal New Candidate Biomarkers. Omi A J Integr Biol 2017;21:285–94. doi:10.1089/omi.2017.0010.

[25] Patil KR, Nielsen J. Uncovering transcriptional regulation of metabolism by using metabolic network topology. Proc Natl Acad Sci U S A 2005;102:2685–9. doi:10.1073/pnas.0406811102.

[26] Kori M, Gov E, Arga KY. Molecular signatures of ovarian diseases: Insights from network medicine perspective. Syst Biol Reprod Med 2016;62:266–82. doi:10.1080/19396368.2016.1197982.

[27] Aguirre-Gamboa R, Gomez-Rueda H, Martínez-Ledesma E, Martínez-Torteya A, Chacolla-Huaringa R, Rodriguez-Barrientos A, et al. SurvExpress: An Online Biomarker Validation Tool and Database for Cancer Gene Expression Data Using Survival Analysis. PLoS One 2013;8:1–9. doi:10.1371/journal.pone.0074250.

[28] Wong NW, Chen Y, Chen S, Wang X. OncomiR: An online resource for exploring pan-cancer microRNA dysregulation. Bioinformatics 2018;34:713–5. doi:10.1093/bioinformatics/btx627.

[29] Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, et al. The connectivity map: Using gene-expression signatures to connect small molecules, genes, and disease. Science (80-) 2006;313:1929–35. doi:10.1126/science.1132939.

[30] Turanli B, Gulfidan G, Arga KY. Transcriptomic-Guided Drug Repositioning Supported by a New Bioinformatics Search Tool: geneXpharma. *Omi A J Integr Biol* 2017;21:584–91. doi:10.1089/omi.2017.0127.

[31] Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2008;4:44–57. doi:10.1038/nprot.2008.211.

[32] Boccardi M, Gallo V, Yasui Y, Vineis P, Padovani A, Mosimann U, et al. The biomarker-based diagnosis of Alzheimer’s disease. 2—lessons from oncology. *Neurobiol Aging* 2017;52:141–52. doi:10.1016/j.neurobiolaging.2017.01.021.

[33] Guo Y, Bao Y, Ma M, Yang W. Identification of key candidate genes and pathways in colorectal cancer by integrated bioinformatical analysis. *Int J Mol Sci* 2017;18:1–15. doi:10.3390/ijms18040722.

[34] Clevers H, Nusse R. Wnt/ β -catenin signaling and disease. *Cell* 2012;149:1192–205. doi:10.1016/j.cell.2012.05.012.

[35] Manna PR, Stocco DM. The Role of Specific Mitogen-Activated Protein Kinase Signaling Cascades in the Regulation of Steroidogenesis. *J Signal Transduct* 2011;2011:1–13. doi:10.1155/2011/821615.

[36] Nagathihalli S Nagaraj Dr, PhD and Prank Datta Dr P. Targetting the Transforming Growth factor-beta Signalling Pathway in Human Cancer. *Expert Opin Investig Drugs* 2010;19:77–91. doi:10.1517/13543780903382609.Targeting.

[37] Sevimoglu T, Arga KY. The role of protein interaction networks in systems biomedicine. *Comput Struct Biotechnol J* 2014;11:22–7. doi:10.1016/j.csbj.2014.08.008.

[38] Blaj C, Bringmann A, Urbischek M, Krebs S, Blum H, Fröhlich T, et al. ADNP is a repressor of WNT signaling in colon cancer that can be therapeutically induced. *Eur J Cancer* 2016;61:S172. doi:10.1016/S0959-8049(16)61611-8.

[39] Rask K, Thorn M, Ponten F, Kraaz W, Sundfeldt K, Hedin L, et al. Increased expression of the transcription factors CCAAT-enhancer binding protein-beta (C/EBB β) and C/EB ζ (CHOP) correlate with invasiveness of human colorectal cancer. *Int J Cancer* 2000;86:337–43.

[40] Yin J, Oh YT, Kim JY, Kim SS, Choi E, Kim TH, et al. Transglutaminase 2 inhibition reverses mesenchymal transdifferentiation of glioma stem cells by regulating C/EBP β signaling. *Cancer Res* 2017;77:4973–84. doi:10.1158/0008-5472.CAN-17-0388.

[41] Balcerczak E, Pasz-Walczak G, Kumor P, Panczyk M, Kordek R, Wierzbicki R, et al. Cyclin D1 protein and CCND1 gene expression in colorectal cancer. *Eur J Surg Oncol* 2005;31:721–6. doi:10.1016/j.ejso.2005.04.005.

[42] Porter TR, Richards FM, Houlston RS, Evans DGR, Jankowski JA, Macdonald F, et al. Contribution of cyclin d1 (CCND1) and E-cadherin (CDH1) polymorphisms to familial and sporadic colorectal cancer. *Oncogene* 2002;21:1928–33. doi:10.1038/sj/onc/1205245.

[43] Wolter F, Akoglu B, Clausnitzer A, Stein J. Downregulation of the cyclin D1/Cdk4 complex occurs during resveratrol-induced cell cycle arrest in colon cancer cell lines. *J Nutr* 2001;131:2197–203.

[44] Pek M, Yatim SMJM, Chen Y, Li J, Gong M, Jiang X, et al. Oncogenic KRAS-associated gene signature defines co-targeting of CDK4/6 and MEK as a viable therapeutic strategy in colorectal cancer. *Oncogene* 2017;36:4975–86. doi:10.1038/onc.2017.120.

[45] Tomonaga T, Matsushita K, Yamaguchi S. Overexpression and Mistargeting of Centromere Protein-A in Human Primary Colorectal Cancer Overexpression and Mistargeting of Centromere Protein-A in Human Primary 2003;63:3511–6.

[46] Tomonaga T, Matsushita K, Ishibashi M, Nezu M, Shimada H, Ochiai T, et al. Centromere protein H Is

up-regulated in primary human colorectal cancer and its overexpression induces aneuploidy. *Cancer Res* 2005;65:4683–9. doi:10.1158/0008-5472.CAN-04-3613.

[47] Maruyama T, Farina A, Dey A, Cheong J, Bermudez VP, Tamura T, et al. A Mammalian Bromodomain Protein, Brd4, Interacts with Replication Factor C and Inhibits Progression to S Phase. *Mol Cell Biol* 2002;22:6509–20. doi:10.1128/MCB.22.18.6509-6520.2002.

[48] Da Costa D, Agathangelou A, Perry T, Weston V, Petermann E, Zlatanou A, et al. BET inhibition as a single or combined therapeutic approach in primary paediatric B-precursor acute lymphoblastic leukaemia. *Blood Cancer J* 2013;3:e126-10. doi:10.1038/bcj.2013.24.

[49] Dang C V., Le A, Gao P. MYC-induced cancer cell energy metabolism and therapeutic opportunities. *Clin Cancer Res* 2009;15:6479–83. doi:10.1158/1078-0432.CCR-09-0889.

[50] Sikora K, Chan S, Evan G, Gabra H, Markham N, Stewart J, et al. c-myc oncogene expression in colorectal cancer. *Cancer* 1987;59:1289–95.

[51] Castell A, Larsson L-G. Targeting MYC Translation in Colorectal Cancer. *Cancer Discov* 2015;5:701–3. doi:10.1158/2159-8290.CD-15-0660.

[52] Hellwig D, Emmerth S, Ulbricht T, Doring V, Hoischen C, Martin R, et al. Dynamics of CENP-N kinetochore binding during the cell cycle. *J Cell Sci* 2011;124:3871–83. doi:10.1242/jcs.088625.

[53] Jenster G. The role of the androgen receptor in the development and progression of prostate cancer. *Semin Oncol* 1999;26:407–421.

[54] Seth A, Watson DK. ETS transcription factors and their emerging roles in human cancer. *Eur J Cancer* 2005;41:2462–78. doi:10.1016/j.ejca.2005.08.013.

[55] Chen L, Jiang B, Wang Z, Liu M, Ma Y, Yang H, et al. Expression and prognostic significance of GATA-binding protein 2 in colorectal cancer. *Med Oncol* 2013;30. doi:10.1007/s12032-013-0498-7.

[56] Zheng R, Blobel GA. Gata transcription factors and cancer. *Genes and Cancer* 2010;1:1178–88. doi:10.1177/1947601911404223.

[57] Prabhu L, Mundade R, Wang B, Wei H, Hartley A-V, Martin M, et al. Critical role of phosphorylation of serine 165 of YBX1 on the activation of NF-κB in colon cancer. *Oncotarget* 2015;6:29396–412. doi:10.18632/oncotarget.5120.

[58] Oda Y, Ohishi Y, Saito T, Hinoshita E, Uchiumi T, Kinukawa N, et al. Nuclear expression of Y-box-binding protein-1 correlates with P-glycoprotein and topoisomerase II alpha expression, and with poor prognosis in synovial sarcoma. *J Pathol* 2003;199:251–8. doi:10.1002/path.1282.

[59] Le Gouvello S, Bastuji-Garin S, Aloulou N, Mansour H, Chaumette MT, Berrehar F, et al. High prevalence of Foxp3 and IL17 in MMR-proficient colorectal carcinomas. *Gut* 2008;57:772–9. doi:10.1136/gut.2007.123794.

[60] Nevins JR. The Rb/E2F pathway and cancer. *Hum Mol Genet* 2001;10:699–703. doi:10.1093/hmg/10.7.699.

[61] Garneau H, Paquin MC, Carrier JC, Rivard N. E2F4 expression is required for cell cycle progression of normal intestinal crypt cells and colorectal cancer cells. *J Cell Physiol* 2009;221:350–8. doi:10.1002/jcp.21859.

[62] Zhang T, Cui G, Bi H, Shi H. PRDM14 Promotes the Migration of Human Non-small Cell Lung Cancer Through Extracellular Matrix Degradation in vitro. *Chin Med J (Engl)* 2015;128:373. doi:10.4103/0366-6999.150109.

[63] Nishikawa N, Toyota M, Suzuki H, Honma T, Fujikane T, Ohmura T, et al. Gene amplification and overexpression of PRDM14 in breast cancers. *Cancer Res* 2007;67:9649–57.

doi:10.1158/0008-5472.CAN-06-4111.

[64] Holst F, Stahl PR, Ruiz C, Hellwinkel O, Jehan Z, Wendland M, et al. Estrogen receptor alpha (ESR1) gene amplification is frequent in breast cancer. *Nat Genet* 2007;39:655–60. doi:10.1038/ng2006.

[65] Hrašovec S, Glavač D. MicroRNAs as novel biomarkers in colorectal cancer. *Front Genet* 2012;3:1–9. doi:10.3389/fgene.2012.00180.

[66] Mullany LE, Herrick JS, Sakoda LC, Samowitz W, John R, Wolff RK, et al. miRNA involvement in cell cycle regulation in colorectal cancer cases 2018;9. doi:10.18632/genesandcancer.167.

[67] Wu K, Zhao Z, Ma J, Chen J, Peng J, Yang S, et al. Deregulation of miR-193b affects the growth of colon cancer cells via transforming growth factor- β and regulation of the SMAD3 pathway. *Oncol Lett* 2017;13:2557–62. doi:10.3892/ol.2017.5763.

[68] Trevisani F, Ghidini M, Larcher A, Lampis A, Lote H, Manunta P, et al. MicroRNA 193b-3p as a predictive biomarker of chronic kidney disease in patients undergoing radical nephrectomy for renal cell carcinoma. *Br J Cancer* 2016;115:1343–50. doi:10.1038/bjc.2016.329.

[69] Schee K, Lorenz S, Worren MM, Günther CC, Holden M, Hovig E, et al. Deep Sequencing the MicroRNA Transcriptome in Colorectal Cancer. *PLoS One* 2013;8. doi:10.1371/journal.pone.0066165.

[70] Wang J, Liu L, Sun Y, Xue Y, Qu J, Pan S, et al. miR-615-3p promotes proliferation and migration and inhibits apoptosis through its potential target CELF2 in gastric cancer. *Biomed Pharmacother* 2018;101:406–13. doi:10.1016/j.biopha.2018.02.104.