

Colon Cancer Progression Is Reflected to Monotonic Differentiation in Gene Expression and Pathway Deregulation Facilitating Stage-specific Drug Repurposing

MARILENA M. BOURDAKOU¹, GEORGE M. SPYROU^{2,3} and GEORGE KOLIOS¹

¹Laboratory of Pharmacology, Department of Medicine, Democritus University of Thrace, Alexandroupolis, Greece;

²The Cyprus Institute of Neurology and Genetics, Nicosia, Cyprus;

³The Cyprus School of Molecular Medicine, Nicosia, Cyprus

Abstract. *Background/Aim:* Colon cancer is one of the most common cancer types and the second leading cause of death due to cancer. Many efforts have been performed towards the investigation of molecular alterations during colon cancer progression. However, the identification of stage-specific molecular markers remains a challenge. The aim of this study was to develop a novel computational methodology for the analysis of alterations in differential gene expression and pathway deregulation across colon cancer stages in order to reveal stage-specific biomarkers and reinforce drug repurposing investigation. *Materials and Methods:* Transcriptomic datasets of colon cancer were used to identify (a) differentially expressed genes with monotonicity in their fold changes (MEGs) and (b) perturbed pathways with ascending monotonic enrichment (MEPs) related to the number of the participating differentially expressed genes (DEGs), across the four colon cancer stages. Through an in silico drug repurposing pipeline we identified drugs that regulate the expression of MEGs and also target the resulting MEPs. *Results:* Our methodology highlighted 15 MEGs and 32 candidate repurposed drugs that affect their expression. We also found 51 MEPs divided into two groups according to their rate of DEG content alteration across colon cancer stages. Focusing on the target MEPs of the highlighted repurposed drugs, we found that one of them, the neuroactive

ligand-receptor interaction, was targeted by the majority of the candidate drugs. Moreover, we observed that two of the drugs (PIK-75 and troglitazone) target the majority of the resulting MEPs. *Conclusion:* These findings highlight significant genes and pathways that can be used as stage-specific biomarkers and facilitate the discovery of new potential repurposed drugs for colon cancer. We expect that the computational methodology presented can be applied in a similar way to the analysis of any progressive disease.

Colon cancer is one of the most common cancer types and the second leading cause of cancer-related deaths in the United States (1). The most significant factor for patient survival, prognosis and treatment is tumor staging (2, 3). As a rule, the earlier the stage of colon cancer, the smaller it's metastatic potential. It is well known that cancer metastases are the major cause of colon cancer-related deaths and there are no available drugs to confine the spread to other organs (4).

Despite the efforts to identify significant features that could be used as important indicators for the better understanding of the colon cancer progression (5-9), there is an increased trend in the appearance and death rates from colon cancer. Furthermore, the molecular characterization of colon cancer stages and identification of predictive biomarkers remain a challenge (10). Few studies in the literature have focused on the comparison and identification of genes that may be associated with clinical stages of colon cancer (11-13) and to our knowledge, there are no computational methodologies to facilitate the investigation of deregulation of molecular mechanisms during colon cancer progression. In the present study we developed a novel computational methodology for the analysis of stage-specific alterations in differential gene expression and pathway deregulation of colon cancer. Specifically, we used publicly available transcriptomic datasets of colon cancer and normal samples accompanied with their clinical information, in order to identify monotonically expressed genes (MEGs) *i.e.*, differentially expressed genes

This article is freely accessible online.

Correspondence to: Professor Dr. George Kolios, Laboratory of Pharmacology, Faculty of Medicine, Democritus University of Thrace, Dragana, Alexandroupolis, 68100, Greece. Tel: +30 2551030368, Fax: +30 2551030541, e-mail: gkolios@med.duth.gr

Key Words: Monotonically expressed genes, monotonically enriched pathways, drug repurposing, colon cancer progression, colon cancer staging.

Table I. Detailed summary of the transcriptomic colon cancer datasets used in this study. The number of normal samples (N) and the number of samples from each stage (I-IV) is presented for each dataset.

Gene expression omnibus	Dataset id	Normal	Stage I	Stage II	Stage III	Stage IV
	GSE39582 (56)	19	33	264	205	60
	GSE71187 (57)	12	13	28	52	6
	GSE21815 (58, 59)	9	12	27	16	11
	GSE21510 (60)	25	15	46	39	25
	GSE35279 (61)	5	13	30	17	14
The cancer genome atlas	TCGA-COAD*	41	81	187	132	66
	TCGA-READ**	10	30	51	51	24

*TCGA-COAD: The Cancer Genome Atlas – Colon Adenocarcinoma. **TCGA-READ: The Cancer Genome Atlas – Rectal Cancer.

(DEGs) with ascending or descending trend in their differential expressions (log 2-fold changes-log2FC) during colon cancer progression. Through an *in silico* drug repurposing pipeline, we pointed out candidate repurposed drugs which have been experimentally found to significantly change the expression of the majority of these MEGs in colon cancer cell lines. We further explored alterations in the revealed molecular mechanisms that participate in cancer progression by investigating the monotonically enriched pathways (MEPs) *i.e.*, the perturbed pathways with ascending enrichment related to the participating DEGs across the four colon cancer stages.

Our computational approach yielded in 15 MEGs and 32 candidate repurposed drugs that affect their expression. Furthermore, we found 51 MEPs that were divided into two groups according to their perturbation rate across stages *i.e.*, rate of DEG content alteration across colon cancer stages. We also investigated which of the candidate repurposed drugs also target the resulting MEPs and we found two drugs (PIK-75 and troglitazone) that target most of them. On the other hand, one MEP, namely the neuroactive ligand-receptor interaction, is targeted by the majority of the candidate repurposed drugs from our analysis.

Overall, our findings highlight genes and pathways monotonically deregulated across colon cancer stages that can be used as stage-specific biomarkers and as a stepping stone in the direction of drug repurposing against colon cancer. This methodology provides insights for further experimentation and can be applied in a similar way to the analysis of any progressive disease.

Materials and Methods

Data selection and preprocessing. We collected transcriptomic datasets that contain colon cancer samples from all stages (Stage I-IV) and normal samples from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>) (14) and from The Cancer Genome Atlas database (TCGA) (<https://www.cancer.gov/tcga>). We

utilized the GEO datasets for our analysis (as reference sets) and the TCGA datasets as validation sets. In GEO, when looking for transcriptomic datasets with samples in all four colon cancer stages accompanied with healthy ones, we found only five independent microarray datasets whose accession numbers and the number of samples are presented in Table I. We used Linear Models for Microarray Data- *LIMMA* R package (15) to perform the differential expression analysis in each dataset comparing samples from each colon cancer stage with normal ones. Finally, we selected the DEGs by applying selection thresholds of adjusted p-value <0.05 and absolute log2FC≥1. The number of DEGs from each comparison is presented in Table II.

Furthermore, we retrieved RNA-seq data from TCGA using TCGAbiolinks R package (16) in order to use them as validation set. Specifically, we downloaded raw-counts from TCGA Colon Adenocarcinoma (TCGA-COAD) and Rectum Cancer (TCGA-READ). Concerning staging information, we separated the samples based on their clinical records. Finally, each subset was statistically analyzed in order to find the differentially expressed genes comparing samples from each colon cancer stage and normal ones. For the differential expression analysis we used DESeq2 R package (17) and we set as threshold 10 total counts per gene to filter out very low expressed genes. Moreover, the same selection criteria of absolute log2FC≥1 and the adjusted p-value <0.05 were utilized. The description of the TCGA datasets and the total number of DEGs from each analysis are presented in Table I and Table II respectively.

Finding and scoring the monotonically expressed genes. For each dataset, we were interested in the DEGs with a monotonic trend (constantly increasing or decreasing trend) in their log 2-fold changes across colon cancer progression, noting them as MEGs. Based on this concept, we examined the monotonically over-expressed genes with an ascending trend in their log2FC and the monotonically under-expressed genes with a descending trend in their log2FC across colon cancer stages.

To further evaluate these MEGs, we investigated their linear correlation in terms of log2FC across the four colon cancer stages. To measure this correlation, we used linear regression to fit the log fold changes of each MEG with line and we subsequently calculated that line's slope and coefficient of determination (R-squared) *i.e.*, how close the log2FC of each MEG are to the fitted regression line. Moreover, we calculated the distance between the extreme log2FC of Stage IV and Stage I for each MEG and by combining those

Table II. The total number of over- and under-expressed genes per colon cancer stage for each dataset from the Gene Expression Omnibus Database (GEO) and The Cancer Genome Atlas (TCGA).

Dataset id	Stage I		Stage II		Stage III		Stage IV	
	Over	Under	Over	Under	Over	Under	Over	Under
GSE39582	1,335	1,138	1,251	1,123	1,256	1,149	1,245	1,138
GSE71187	3,421	3257	4,037	2,932	3,504	2,859	5,798	3,445
GSE21815	5,821	332	8,893	547	8,632	548	8,024	614
GSE21510	2,843	1,924	2,519	1,882	2,728	1,927	2,652	1,791
GSE35279	899	125	6475	480	6990	461	6,126	496
Colon adenocarcinoma (COAD)	9,618	4,327	9,390	4,351	8,522	4,017	8,205	4,476
Rectal cancer (READ)	5,448	4,739	5,672	4,544	5,778	4,461	5,421	4,500

measures we scored and ranked the MEGs of each dataset using a weighted sum according to the following equation [1]:

$$S_i = w_1 * |\Delta FC_{I,IV}|_i + w_2 * |a_i| * R^2_i, \quad i = 1, \dots, N \text{ MEGs} \quad [1]$$

Where, $|\Delta FC_{I,IV}|_i$ is the absolute distance of the log2FC between Stage IV and Stage I, $|a_i|$ is the absolute value of slope and R^2_i the coefficient of determination for each MEG. This weighted sum with w_1 , w_2 weights, combines the actual total increase or decrease of the log2FC across stages with the trend parameters calculated through the linear regression. In the present study we selected to give precedence in the actual total change in the log2FC by setting w_1 and w_2 0.6 and 0.4, respectively.

According to these scores, each MEG in each dataset acquired a ranking. The rank of each MEG was normalized according to equation [2]:

$$NR_{ij} = \frac{g_{i,j}}{N_j} \quad [2]$$

where $g_{i,j}$ is the rank of the i^{th} gene in j^{th} dataset and N_j the total number of MEGs in j^{th} dataset.

Finally, we combined the normalized ranks from the five datasets and we calculated a final score as a weighted sum for the monotonically over- and under-expressed genes separately, using the following equation [3]:

$$FS_i = w_1 * \left(1 - \frac{\sum_1^N NR_{ij_i}}{N}\right) + w_2 * f_i, \quad i = 1, \dots, N \text{ MEGs} \quad [3]$$

where f_i the number of appearances of MEG i in the datasets. In our study, the weights w_1 and w_2 were set to 0.4 and 0.6, respectively giving a precedence in the score to MEGs found in more datasets. *In silico* drug repurposing. The selected MEGs were used separately as input in an *in silico* drug repurposing tool called Drug Gene Budger

(<https://maayanlab.cloud/DGB/>) (18). DGB is a web-based tool which returns small molecules that are predicted to maximally affect the expression of the genes of interest. Users can query the genes that they want to reverse their expression and DGB results a ranked list of small molecules which have been experimentally found to produce the desired expression effect. The experimental data of DGB have been extracted from the LINCS L1000 dataset (19), the Connectivity Map (CMap) dataset (20), and the GEO database (14). In this study we selected the results from LINCS L1000 and more specifically the significant small molecules with $q\text{-value} < 0.05$ and absolute $\log_2 FC \geq 1$, that reverse the expression of MEGs in the three colon cancer cell lines (HT29, SW620 and SW948) that take part in L1000 data.

Monotonically enriched pathways (MEPs). To further investigate the alterations in molecular mechanisms related to the colon cancer progression, we explored the perturbed pathways with ascending monotonic enrichment related to the number of the involved DEGs across the four colon cancer stages. For this reason, we constructed a consensus gene signature for each colon cancer stage by examining the common over- and under-expressed genes from the differential expression analysis of the five GEO datasets. Then we found the molecular mechanisms in which the common DEGs of each stage are involved. Specifically, by parsing the biological pathways from Kyoto Encyclopedia of Genes and Genomes – KEGG (21), we found all genes that are involved in each pathway. Finally, we investigated the pathways with an ascending monotonic enrichment in the number of the participating DEGs from Stage I to Stage IV.

MEPs network construction. MEPs from all clusters were used as input in the PathIN - an integrated web tool that provides an easy and flexible way for rapidly creating pathway-based networks, at several functional biological levels: genes, compounds and reactions (<http://bioinformatics.cing.ac.cy/PathIN>). We used PathIN to map the MEPs on the reference pathway-to-pathway network parsed from the KEGG database.

Results

MEGs found in colon cancer progression. We found the following numbers of monotonically over-/under-expressed genes per dataset: GSE39582 (146/151), GSE21815 (668/51),

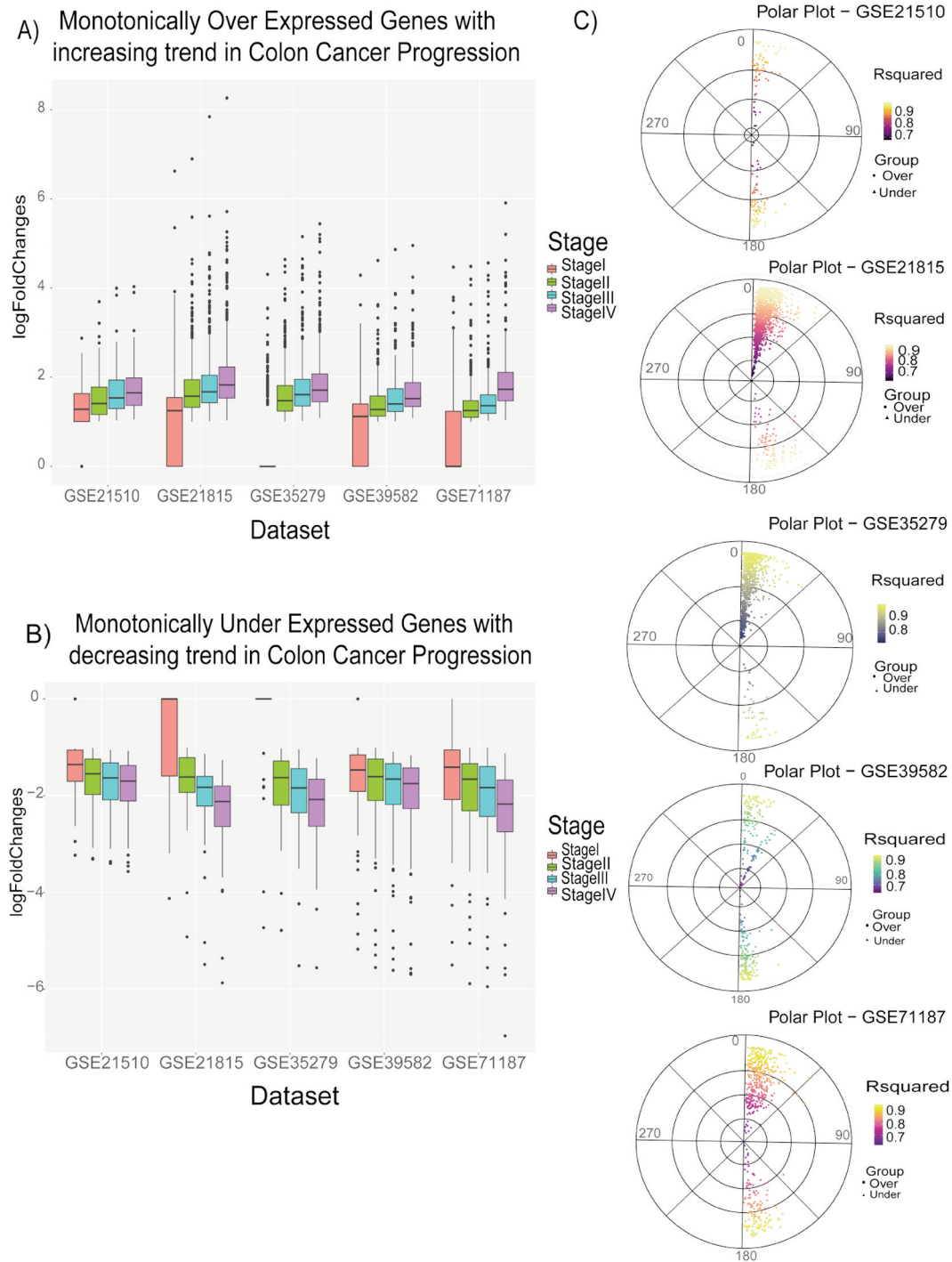


Figure 1. (A) Boxplot with monotonic over-expressed genes for the five Gene Expression Omnibus (GEO) datasets. (B) Boxplot with monotonic under-expressed genes for the five GEO datasets. (C) Polar plot with the monotonic over- and under-expressed genes of each GEO dataset. Degrees represent the corresponding slopes of monotonically expressed genes (MEGs) and color the R-squared values.

GSE21510 (73/71), GSE71187 (419/133) and GSE35279 (899/45) that are presented in the box plots of Figure 1A and B. Moreover, as described in the Materials and Methods

section we investigated the linear relationships between their log2FC across the four colon cancer stages, by calculating the regression coefficient (Slope) and the coefficient of

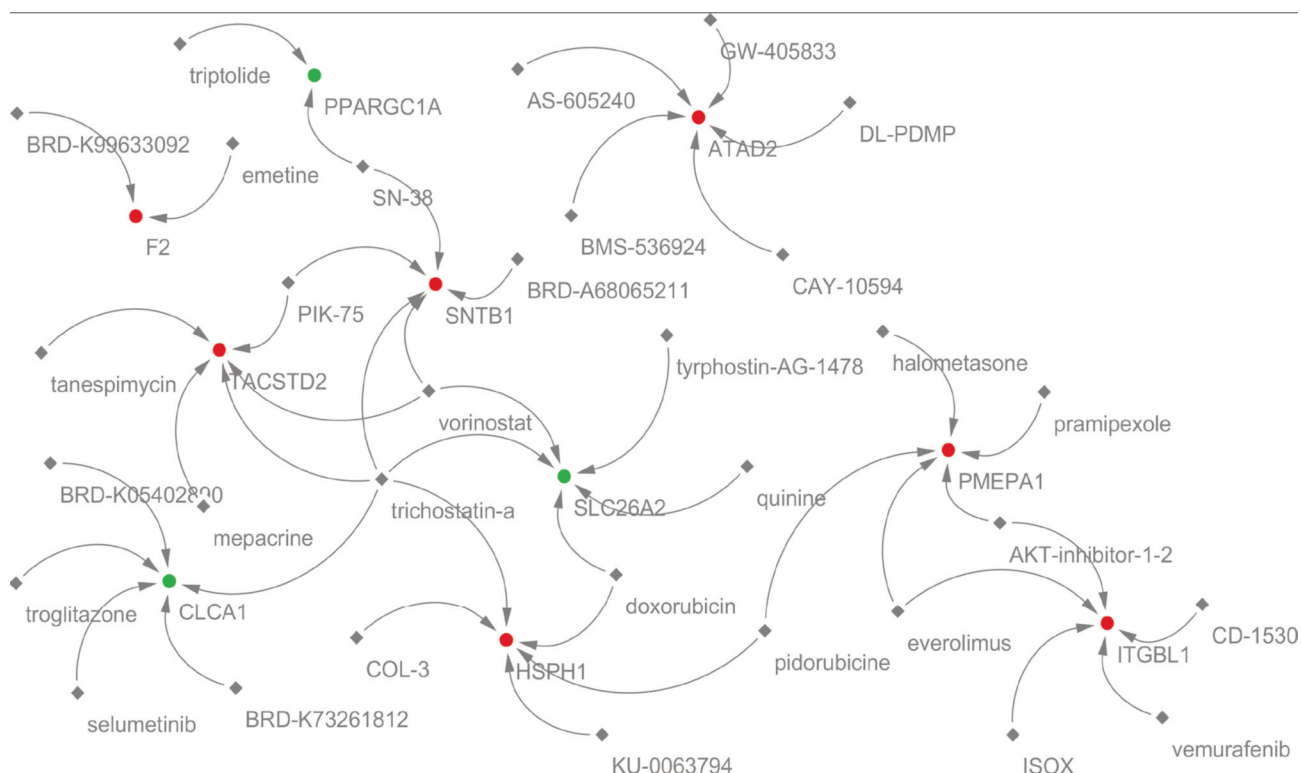


Figure 2. Bipartite network with the resulting monotonically expressed genes (MEGs) and the corresponding candidate repurposed drugs that are predicted to influence their expression. Monotonically over- and under-expressed genes are presented as circular nodes with red and green color respectively and drugs with grey rhombus-shaped nodes.

determination (R-squared) in a linear regression model. The corresponding metrics for each MEG and for each dataset are presented in a polar coordinate system in Figure 1C.

For each MEG, we calculated a final weighted score FS_i that combines the normalized ranks and the frequency of occurrence of each MEG across the five GEO datasets. Following the sorting of MEGs according to their FS_i score, we highlighted those that surpassed the following criteria: 1) their FS_i is above 0.5 and 2) they are found as MEGs with the same trend in at least one TCGA validation set. Following these filtering criteria, we kept 12 MEGs with ascending monotonicity (HOXB8, SNTB1, ATAD2, KLRG2, ITGBL1, RDH12, LEMD1, TACSTD2, F2, RELL2, PMEPA1 and HSPH1) and 3 MEGs with descending monotonicity (PPARGC1A, SLC26A2 and CLCA1) respectively.

Candidate repurposed drugs for the highlighted MEGs. The 15 selected MEGs were separately used as input to the DGB drug repurposing tool. We found candidate repurposed drugs that regulate 7 out of 12 over-expressed MEGs and all the 3 under-expressed MEG. From these drugs we kept only those that were experimentally found to regulate the resulting MEGs

in colon cancer cell lines. Then we used as filtering criteria the q -value <0.05 and the absolute $\log_2FC \geq 1$ and we selected for each MEG the maximum top 5 drugs, based on the \log_2FC that quantify the observed change in the expression of each gene. With this procedure, we ended up with 32 unique drugs presented in the bipartite network of Figure 2.

MEPs found in colon cancer progression. We investigated the common over- and under-expressed genes across stages and datasets. For the case of Stage I, 151 common over-expressed genes and 51 under-expressed genes were found across the 5 datasets. Following the same procedure, we concluded to 367 common over- and 141 common under-expressed genes for Stage II, 360 and 127 for Stage III and 488 and 139 for Stage IV respectively. Finally, we ended up with 51 MEPs, shown in Figure 3. For each MEP, we normalized the number of DEGs in each stage with the total number of genes that are involved in this pathway. We then investigated the groups of pathways with similar deregulation across stages. Specifically, the 51 MEPs were grouped into 4 clusters according to their monotonic enrichment in the four colon cancer stages (Figure 3). The four clusters have been found

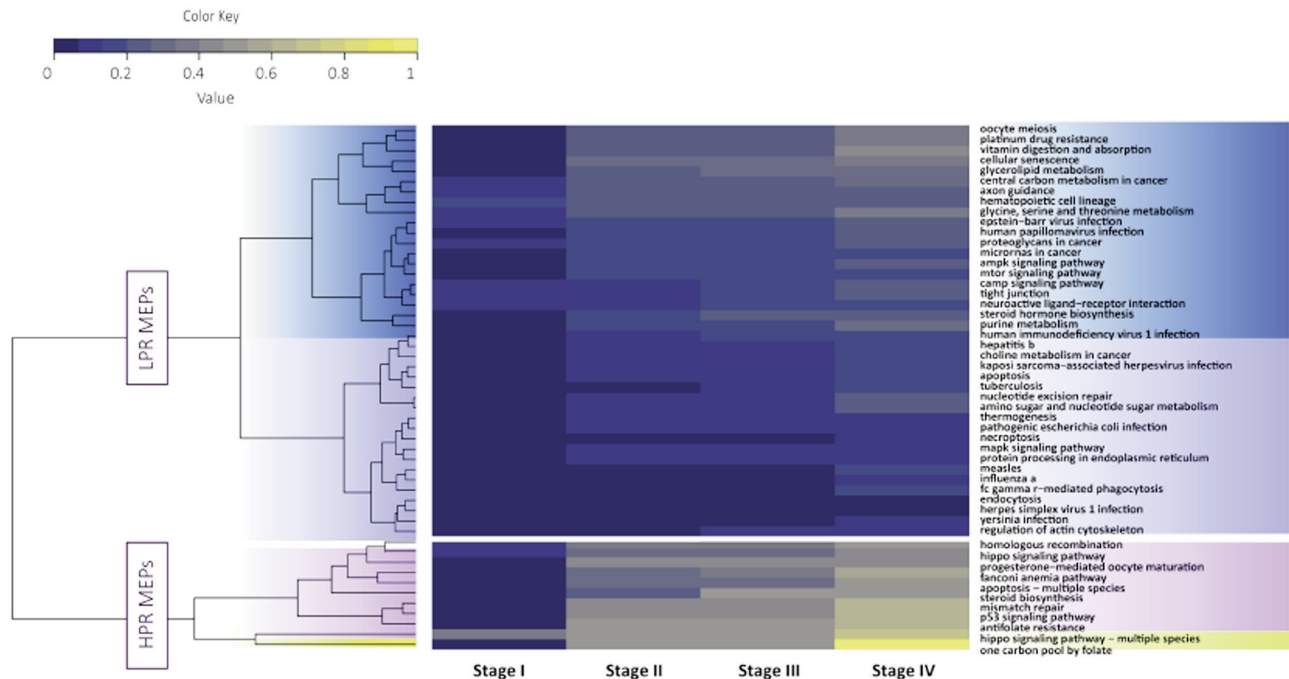


Figure 3. Heatmap visualization of monotonically enriched pathways (MEPs). Color scale indicates the normalized number of differentially expressed genes (DEGs) in each stage.

based on the normalized number of DEGs. The color scale in the presented heat map represents a transformation of the normalized number of DEGs in the scale of [0, 1].

As presented in Figure 3, cluster 1 (C1) consists of 20 pathways with smooth alteration rate between Stage II – Stage III or/and Stage III – Stage IV. Cluster 2 (C2) included also 20 pathways with the lowest deregulation across all colon cancer stages and also slight or no alterations between them. For the case of cluster 3 (C3), we observed nine involved pathways with high total deregulation across the colon cancer progression and high alteration rate between Stage I and Stage II. Finally, cluster 4 (C4) also consists of two pathways with the highest perturbation rate across stages and quite high rate of alteration between Stage I and Stage II or/and Stage III and Stage IV. Observing the four clusters, we separated the MEPs into two groups: the first group includes the MEPs with low perturbation rate (LPR) across stages *i.e.*, low rate of DEG content alteration. The second group contains MEPs with high perturbation rate (HPR) across the four colon cancer stages *i.e.*, high rate of DEG content alteration (Figure 3).

Network connectivity between the MEPs. The LPR and HPR MEPs from the two groups were mapped on the KEGG pathway-to-pathway reference network with the help of

PathIN. In this network, the nodes represent the MEPs, the edges represent the connectivity found in KEGG between each pair of pathways while the edge weight represent the number of common genes between the two pathways. As shown in Figure 4, MEPs within and across groups are linked as they are sharing biochemical connections. Specifically, the majority of the LPR MEPs from the first group are strongly associated. On the other hand, subgroups from the HPR MEPs are biochemically connected, sharing also a number of common genes. Following the network statistics, the MEPs with the highest degree centralities are apoptosis and mapk signaling pathway and they belong to the LPR MEPs of the first group.

MEPs and repurposed drugs. To further enrich our analysis, we investigated the MEPs that are targets of the highlighted repurposed drugs. For this reason, we queried the corresponding gene targets of each drug in the databases CLUE –The Drug Repurposing Hub (<https://clue.io/repurposing#download-data>), the DrugBank database (<https://www.drugbank.com/>) (22) and PubChem (<https://pubchem.ncbi.nlm.nih.gov/>) (23). Then we looked in which pathways these target genes are involved. Finally, we compared these pathways with the resulting MEPs in order to find which MEPs are also targets of the resulting repurposed drugs. We applied this procedure to each group to

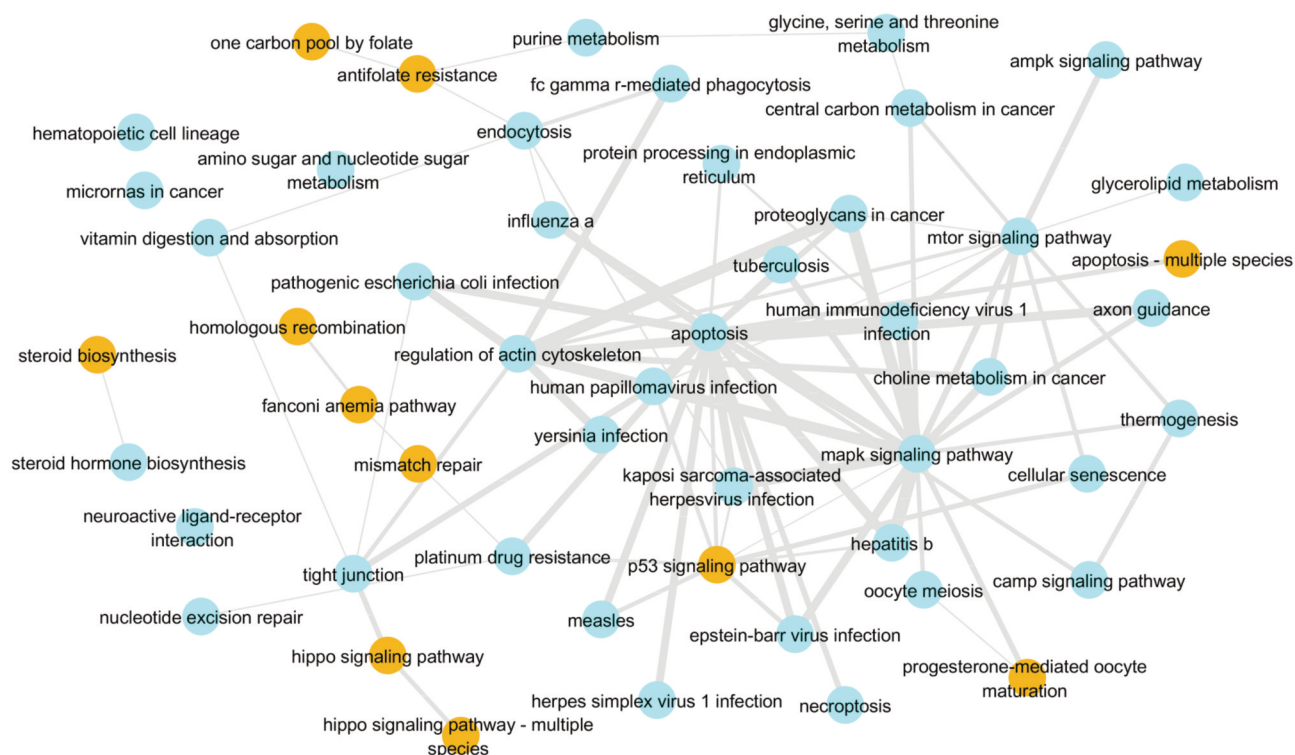


Figure 4. Pathway network from monotonically enriched pathways (MEPs) biochemical connectivity. Low perturbation rate monotonically enriched pathways (LPR MEPs) from the first group are presented with cyan color and high perturbation rate monotonically enriched pathways (HPR MEPs) from the second group with orange color. Edge width corresponds to the common genes that the two pathways are shared.

find drugs that affect MEPs with a similar tendency of deregulation during the colon cancer progression. As shown in Figure 5, 18 out of 32 repurposed drugs (PIK-75, CAY-10594, everolimus, doxorubicin, triptolide, BRD-K05402890, troglitazone, tanespimycin, tyrphostin-AG-1478, trichostatin-a, BMS-536924, CD-1530, vemurafenib, emetine, pidorubicine, vorinostat, mepacrine and selumetinib) target 33 out of 51 MEPs (ampk signaling pathway, axon guidance, camp signaling pathway, cellular senescence, central carbon metabolism in cancer, Epstein Barr virus infection, hematopoietic cell lineage, human papillomavirus infection, microRNAs in cancer, mtor signaling pathway, neuroactive ligand-receptor interaction, platinum drug resistance, proteoglycans in cancer, purine metabolism, steroid hormone biosynthesis, apoptosis, choline metabolism in cancer, endocytosis, fc gamma r-mediated phagocytosis, hepatitis b, herpes simplex virus 1 infection, human immunodeficiency virus 1 infection, influenza a, Kaposi sarcoma-associated herpesvirus infection, mapk signaling pathway, measles, necroptosis, pathogenic *Escherichia coli* infection, regulation of actin cytoskeleton, tuberculosis, yersinia infection, antifolate resistance and progesterone-mediated oocyte maturation). From these MEPs, two belong to HPR MEPs and specifically antifolate resistance and progesterone-

mediated oocyte maturation. We prioritized the candidate drugs based on the number of MEPs that they target and among them we found that the most MEPs are targeted by PIK-75 (23 target MEPs) and troglitazone (19 target MEPs) followed by tanespimycin (7 target MEPs) and CAY-10594 (4 target MEPs). The rest drugs target 1-3 MEPs. All the candidate drugs and the number of MEPs they target are presented in Table III. It is worth noting that the first two candidate repurposed drugs, PIK-75 and troglitazone target both LPR and HPR MEPs. Specifically, PIK-75 targets 22 LPR MEPs (ampk signaling pathway, axon guidance, camp signaling pathway, cellular senescence, central carbon, metabolism in cancer, Epstein Barr virus infection, human papillomavirus infection, microRNAs in cancer, mtor signaling pathway, platinum drug resistance, proteoglycans in cancer, apoptosis, choline metabolism in cancer, fc gamma r-mediated phagocytosis, hepatitis b, herpes simplex virus 1 infection, human immunodeficiency virus 1 infection, influenza a, Kaposi sarcoma-associated, herpesvirus infection, measles, regulation of actin cytoskeleton and yersinia infection) and one HPR MEP (progesterone-mediated oocyte maturation). On the other hand, 18 LPR MEPs and one HPR MEP are targeted by troglitazone (LPR MEPS: cellular senescence, Epstein Barr virus infection, hematopoietic cell

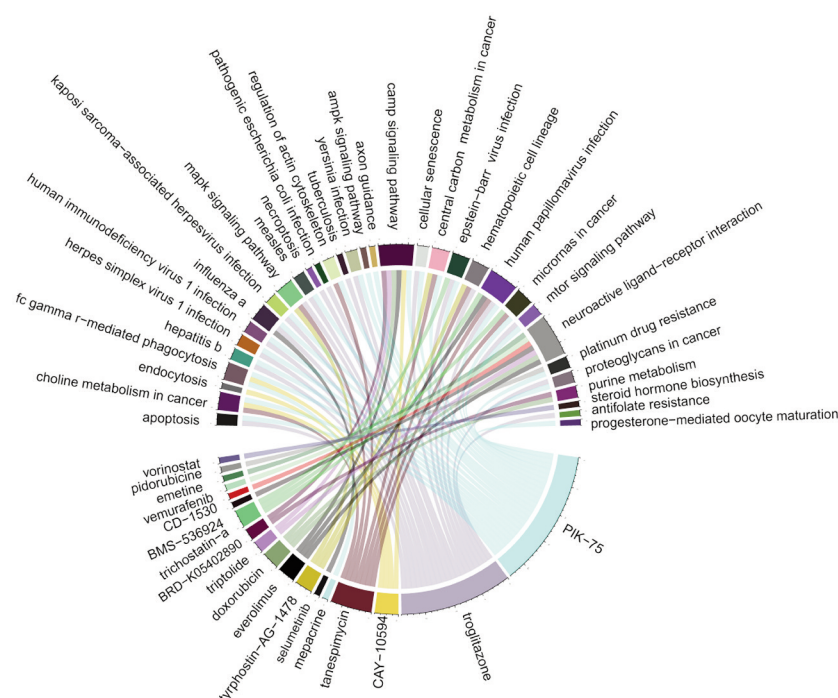


Figure 5. Circos plot that summarizes the associations between repurposed drugs and monotonically enriched pathways (MEPs).

lineage, human papillomavirus infection, mtor signaling pathway, proteoglycans in cancer, apoptosis, hepatitis b, herpes simplex virus 1 infection, human immunodeficiency virus 1 infection, influenza a, Kaposi sarcoma-associated herpesvirus infection, mapk signaling pathway, measles, necroptosis, pathogenic, *Escherichia coli* infection, tuberculosis and yersinia infection; HPR MEP: antifolate resistance). The MEPs that were targeted by the majority of the drugs are: (i) the neuroactive ligand-receptor interaction MEP (targeted by 7 repurposed drugs: BMS-536924, CD-1530, doxorubicin, emetine, everolimus, triptolide and vemurafenib), (ii) the camp signaling MEP (targeted by 6 repurposed drugs: BRD-K05402890, CAY-10594, doxorubicin, everolimus, PIK-75 and triptolide).

Discussion

Despite the enormous amount of omics data and the huge efforts to identify important molecular indicators that could be used for better understanding of molecular pathology in multistage diseases, the understanding of the molecular alterations during colon cancer progression remains a challenge. The simple way of the differential expression analysis across cancer stages is not enough to highlight related genes and molecular mechanisms that mark the phases and the progression of colon cancer.

Table III. Ranking of candidate repurposed drugs based on the number of monotonically enriched pathways (MEPs) that they target.

Candidate repurposed drugs	Number of MEP targets
PIK-75	23
Troglitazone	19
Tanespimycin	7
CAY-10594	4
Everolimus	3
Doxorubicin	3
Tyrphostin-AG-1478	3
Trichostatin-a	3
Triptolide	2
BRD-K05402890	2
BMS-536924	1
CD-1530	1
Vemurafenib	1
Emetine	1
Pidorubicine	1
Vorinostat	1
Mepacrine	1
Selumetinib	1

In this study, we analyzed five transcriptomic datasets from the Gene Expression Omnibus database (GEO), with colon cancer samples from different stages and normal samples, in order to find DEGs with monotonicity in their log₂FC during

the colon cancer progression. We found 15 MEGs, 12 over-expressed with ascending monotonicity and 3 under-expressed with descending monotonicity respectively. We searched in the literature to find the biological relevance of the resulting 15 MEGs and reached some significant observations. For the case of the 12 over-expressed MEGs, it has been reported that HOXB8 and ATAD2 are important factors for the development and progression of colon cancer and may be used as possible drug targets for colon cancer therapy (24, 25). A recent publication has also proposed the significant role of syntrophin beta 1-SNTB1 as prognostic marker for colon cancer metastasis (26). Moreover, it has been found that integrin beta-like 1 -ITGBL1 is involved in the colon cancer development and metastasis (27) while TACSTD2 and LEMD1 belong to the significant up-regulated genes of colon cancer (28, 29). For the case of under-expressed MEGs found in our analysis, it has been indicated that the lower levels of PPARGC1A and SLC26A2 could increase colon cancer risk, proliferation and propagation and they have been proposed as possible candidate targets for cancer therapy (30, 31). Finally, low expression of CLCA1 has been related to unfavorable prognosis in colon cancer and it has been indicated that could be a possible prognostic marker (32, 33). High CLCA1 expression has also been shown to suppress colon cancer belligerence (34).

Through a computational drug repurposing pipeline, we also found drugs and small molecules that affect the expression levels of the highlighted MEGs. Among the 32 resulting candidate repurposed drugs, some were found to affect the expression of more than one MEGs. Specifically, trichostatin-a that belongs to the class of organic compounds and it is used as an antifungal antibiotic, regulates the expression of three monotonically over-expressed genes (SNTB1, TACSTD2 and HSPH1) and two under-expressed (SLC26A2 and CLCA1). It has been reported that trichostatin-a has anticancer effects in the domain of cell proliferation and apoptosis and it has been suggested as possible therapy for colon cancer (35). Moreover, vorinostat, a member in the family of compounds that inhibit HDAC and also used in the management of cutaneous T cell lymphoma, was found to influence the expression of two over- and one under-monotonically expressed gene SNTB1, TACSTD2 and SLC26A2 respectively. It is also in 6 clinical trial studies for colon cancer in the ClinicalTrials.gov with National Clinical Trial number (NCT) numbers NCT00336141, NCT02316340, NCT00942266, NCT00138177, NCT00126451 and NCT01023737. SN-38, the active metabolite of irinotecan (a chemotherapeutic drug for metastatic colorectal cancer) (36), was also found as a candidate repurposed drug from our analysis regulating the expression of one over- and one under-monotonically expressed gene (SNTB1 and PPARGC1A). Doxorubicin, another antineoplastic drug of our list that also affects one over- and one under-monotonically expressed gene (HSPH1 and SLC26A2), is used to treat various types of cancer including colon cancer (37).

We further investigated the perturbed pathways with an ascending monotonic enrichment in the colon cancer progression *i.e.*, pathways with ascending monotonic enrichment in terms of their participating DEGs across the four colon cancer stages. We ended up with 51 MEPs that were mapped into four clusters forming two groups: (i) LPR MEPs with low perturbation rate (*i.e.*, low rate of DEG content alteration) across colon cancer stages and (ii) HPR MEPs with high perturbation rate respectively. We noticed that MEPs within and across the two groups are strongly associated in a pathway-to-pathway network (Figure 4). Hippo signaling pathway and tight junction pathway are strongly connected since they are sharing 27 common genes. Hippo pathway is involved in tumorigenesis in several cancer types, and it has been reported that is also connected in cancer growth and progression (38, 39). Moreover, it has become increasingly evident that tight junctions are involved in cancer metastasis and changes of their protein expression are found in several colon cancer cases (40). It is worth noting that one HPR MEP, namely the p53 signaling pathway, is associated with 9 LPR MEPs (measles, human papillomavirus infection, platinum drug resistance, apoptosis, Epstein Barr virus infection, hepatitis b, cellular senescence, mapk signaling pathway and Kaposi sarcoma-associated herpes virus infection). From them, the Epstein Barr virus infection MEP is sharing with the p53 signaling pathway the highest number of common genes. It is well known that p53 signaling pathway is responsible for the alterations of critical regulators of the cell cycle, the angiogenesis, the DNA replication and apoptosis (41). Furthermore, two LPR MEPs, the endocytosis and platinum drug resistance, were also found to be connected with two HPR MEPs, the mismatch repair and antifolate resistance respectively. It is already known that the alterations of endocytosis are involved in tumorigenesis as they affect proliferation and migration (42). Following the pathway network statistics, two LPR MEPs have the highest degree centralities: apoptosis and mapk signaling pathway. Apoptosis is one of the most deregulated pathways in colon cancer and mapk signaling has an important role in tumor growth and progression (43).

Based on the ranking of the highlighted drugs (Table III), the top two drugs with the highest number of targeted MEPs are PIK-75 and troglitazone. These two drugs target both LPR and HPR MEPs. PIK-75 is a PI 3-kinase p110alpha inhibitor and it is known to have anti-cancer activity and anti-inflammatory properties (44). Troglitazone is an antidiabetic and anti-inflammatory drug which has potential anti-cancer properties on several cancer types including prostate cancer (45). PIK-75 is targeting 22 LPR MEPs and one HPR and troglitazone 18 LPR and one HPR MEPs also (Figure 5). Overall, 13 LPR MEPs are targeted by both drugs: cellular senescence, Epstein Barr virus infection, human papillomavirus infection, mtor signaling pathway,

proteoglycans in cancer, apoptosis, hepatitis b, herpes simplex virus 1 infection, human immunodeficiency virus 1 infection, influenza a, Kaposi sarcoma-associated herpesvirus infection, measles and yersinia infection. We can observe that 6 from these common targeted MEPs belong to the group of viral infectious diseases. It is well known that the most common oncoviruses, human papillomavirus (HPVs) and the Epstein Barr virus (EBV), contribute to many cancer types including colorectal cancer (46, 47). Moreover, it has been proposed that Epstein Barr virus infection is involved in the etiopathogenesis of inflammatory bowel diseases (48). However, the association of viral infections with colon cancer is still under consideration (49). Recent studies have shown that senescent cells are involved in various pathologic conditions and in the progression of several diseases (50). It has been also reported that cellular senescence could be a promising strategy for cancer therapy (51). The mtor pathway plays a pivotal role in cell proliferation and it is deregulated in many cancer types (52). As mentioned above, PIK-75 and troglitazone target one HPR MEP each, namely progesterone-mediated oocyte maturation and antifolate resistance respectively. It has been reported that progesterone-mediated oocyte maturation may be used as diagnostic marker and possible target for colon cancer treatment (53). Mechanisms of antifolate resistance are a pivotal cause of difficulty in cancer chemotherapy (54). Finally, we observe that the most targeted MEP is neuroactive ligand-receptor interaction signaling pathway (Figure 5). This pathway has been associated with many cancer types and more specifically is associated with the progression of bladder, prostate and renal cancer (55).

In summary, in our study we highlighted the importance of 15 MEGs as prognostic markers for colon cancer progression. Four of them (SNTB1, ATAD2, HOXB8 and ITGBL1) have been indicated to be associated with the development of colon cancer (62- 65). We further identified 51 MEPs that were mapped into two groups according to their perturbation rate across colon cancer stages. By examining their network connectivity, we highlighted the hippo signaling pathway and the tight junction pathway. It has been reported that hippo pathway components were associated with tumor differentiation, colon cancer stages and metastasis (66). On the other hand, the role of tight junction in human colon cancer development is still unknown. Through a computational drug repurposing pipeline, we also found repurposed drugs that influence the expression levels of the 15 highlighted MEGs. Among the 32 resulting candidate repurposed drugs, 18 also target the resulting MEPs. The two drugs with the highest number of MEPs that they target are the p110 α inhibitor PIK-75 and the antidiabetic drug troglitazone. Finally, the MEP that is targeted by the majority of the candidate repurposed drugs

from our analysis is the neuroactive ligand-receptor interaction. This pathway may be a promising therapeutic target for future investigation and treatment of colon cancer patients.

This study provides new computational evidence that the monotonicity of the gene differential expression and the pathway deregulation are associated with the colon cancer progression. This observation offers alternative views in drug repurposing efforts as well. Despite all the positive findings of this study, there are some possible limitations related to the database biases that may exist (a common problem in bioinformatics research), the inherent heterogeneity in biological material sampling and annotation regarding the different datasets used and the possible differences in drug selectivity on pathway targeting. Undoubtedly, this is a large-scale computational methodology and the need for further experimental validation of the findings is obvious. Nevertheless, many of our findings have been already associated to the colon cancer staging and progression and this testifies to a preliminary validity of our approach. Based on this implicit validation note, the association of the rest of the findings can be further investigated. The presented findings highlight significant genes and pathways that act as novel potential stage-specific biomarkers and facilitate stage-specific drug repurposing against colon cancer. Furthermore, this work describes a computational methodology that can be applied in a similar way to the analysis of any progressive disease.

Data Availability

For extra information and analysis please visit: <https://github.com/kmarlen1988/Colon-Cancer-Progression/>

Conflicts of Interest

The Authors declare no conflicts of interest.

Authors' Contributions

Supervision of the study: G.K. and G.M.S. Conception and design of the study: M.M.B., G.K. and G.M.S. Collection, analysis and interpretation of data: M.M.B., G.K. and G.M.S. Drafting the article and revising it critically: M.M.B., G.K. and G.M.S.

Acknowledgements

M.M.B. is funded by the State Scholarships Foundation (IKY) scholarship; This research is co-financed by Greece and the European Union (European Social Fund- ESF) through the Operational Programme "Human Resources Development, Education and Lifelong Learning" in the context of the project 'Reinforcement of Post-doctoral Researchers - 2nd Cycle' [MIS-5033021], implemented by the State Scholarships Foundation (IKY).

References

- 1 Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Butterly LF, Anderson JC, Cercek A, Smith RA and Jemal A: Colorectal cancer statistics, 2020. *CA Cancer J Clin* 70(3): 145-164, 2020. PMID: 32133645. DOI: 10.3322/caac.21601
- 2 Wang S, Guan X, Ma M, Zhuang M, Ma T, Liu Z, Chen H, Jiang Z, Chen Y, Wang G and Wang X: Reconsidering the prognostic significance of tumour deposit count in the TNM staging system for colorectal cancer. *Sci Rep* 10(1): 89, 2020. PMID: 31919408. DOI: 10.1038/s41598-019-57041-2
- 3 Kim SM, Yoon G and Seo AN: What are the most important prognostic factors in patients with residual rectal cancer after preoperative chemoradiotherapy? *Yeungnam Univ J Med* 36(2): 124-135, 2019. PMID: 31620624. DOI: 10.12701/yujm.2019.00157
- 4 Nagahashi M, Ramachandran S, Rashid OM and Takabe K: Lymphangiogenesis: a new player in cancer progression. *World J Gastroenterol* 16(32): 4003-4012, 2010. PMID: 20731013. DOI: 10.3748/wjg.v16.i32.4003
- 5 Chen Z, Lin Y, Gao J, Lin S, Zheng Y, Liu Y and Chen SQ: Identification of key candidate genes for colorectal cancer by bioinformatics analysis. *Oncol Lett* 18(6): 6583-6593, 2019. PMID: 31788116. DOI: 10.3892/ol.2019.10996
- 6 De Divitiis C, Nasti G, Montano M, Fisichella R, Iaffaioli RV and Berretta M: Prognostic and predictive response factors in colorectal cancer patients: between hope and reality. *World J Gastroenterol* 20(41): 15049-15059, 2014. PMID: 25386053. DOI: 10.3748/wjg.v20.i41.15049
- 7 Idowu MO, Laudadio J and Rizzo K: Diagnostic, prognostic, and predictive molecular biomarkers and the utility of molecular imaging in common gastrointestinal tumors. *Biomed Res Int* 2015: 890805, 2015. PMID: 26618179. DOI: 10.1155/2015/890805
- 8 Liu X, Liu X, Qiao T and Chen W: Identification of crucial genes and pathways associated with colorectal cancer by bioinformatics analysis. *Oncol Lett* 19(3): 1881-1889, 2020. PMID: 32194683. DOI: 10.3892/ol.2020.11278
- 9 Chen J, Wang Z, Shen X, Cui X and Guo Y: Identification of novel biomarkers and small molecule drugs in human colorectal cancer by microarray and bioinformatics analysis. *Mol Genet Genomic Med* 7(7): e00713, 2019. PMID: 31087508. DOI: 10.1002/mgg3.713
- 10 Koncina E, Haan S, Rauh S and Letellier E: Prognostic and predictive molecular biomarkers for colorectal cancer: updates and challenges. *Cancers (Basel)* 12(2): 319, 2020. PMID: 32019056. DOI: 10.3390/cancers12020319
- 11 Palaniappan A, Ramar K and Ramalingam S: Computational identification of novel stage-specific biomarkers in colorectal cancer progression. *PLoS One* 11(5): e0156665, 2016. PMID: 27243824. DOI: 10.1371/journal.pone.0156665
- 12 Huo T, Canepa R, Sura A, Modave F and Gong Y: Colorectal cancer stages transcriptome analysis. *PLoS One* 12(11): e0188697, 2017. PMID: 29182684. DOI: 10.1371/journal.pone.0188697
- 13 Tian S: Identification of monotonically differentially expressed genes for non-small cell lung cancer. *BMC Bioinformatics* 20(1): 177, 2019. PMID: 30971213. DOI: 10.1186/s12859-019-2775-8
- 14 Edgar R, Domrachev M and Lash AE: Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30(1): 207-210, 2002. PMID: 11752295. DOI: 10.1093/nar/30.1.207
- 15 Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W and Smyth GK: limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43(7): e47, 2015. PMID: 25605792. DOI: 10.1093/nar/gkv007
- 16 Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, Sabedot TS, Malta TM, Pagnotta SM, Castiglioni I, Ceccarelli M, Bontempi G and Noushmehr H: TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res* 44(8): e71, 2016. PMID: 26704973. DOI: 10.1093/nar/gkv1507
- 17 Love MI, Huber W and Anders S: Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15(12): 550, 2014. PMID: 25516281. DOI: 10.1186/s13059-014-0550-8
- 18 Wang Z, He E, Sani K, Jagodnik KM, Silverstein MC and Ma'ayan A: Drug Gene Budger (DGB): an application for ranking drugs to modulate a specific gene based on transcriptomic signatures. *Bioinformatics* 35(7): 1247-1248, 2019. PMID: 30169739. DOI: 10.1093/bioinformatics/bty763
- 19 Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, Gould J, Davis JF, Tubelli AA, Asiedu JK, Lahr DL, Hirschman JE, Liu Z, Donahue M, Julian B, Khan M, Wadden D, Smith IC, Lam D, Liberzon A, Toder C, Bagul M, Orzechowski M, Enache OM, Piccioni F, Johnson SA, Lyons NJ, Berger AH, Shamji AF, Brooks AN, Vrcic A, Flynn C, Rosains J, Takeda DY, Hu R, Davison D, Lamb J, Ardlie K, Hogstrom L, Greenside P, Gray NS, Clemons PA, Silver S, Wu X, Zhao WN, Read-Button W, Wu X, Haggarty SJ, Ronco LV, Boehm JS, Schreiber SL, Doench JG, Bittker JA, Root DE, Wong B and Golub TR: A next generation Connectivity Map: L1000 platform and the first 1,000,000 profiles. *Cell* 171(6): 1437-1452.e17, 2017. PMID: 29195078. DOI: 10.1016/j.cell.2017.10.049
- 20 Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet JP, Subramanian A, Ross KN, Reich M, Hieronymus H, Wei G, Armstrong SA, Haggarty SJ, Clemons PA, Wei R, Carr SA, Lander ES and Golub TR: The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 313(5795): 1929-1935, 2006. PMID: 17008526. DOI: 10.1126/science.1132939
- 21 Kanehisa M, Furumichi M, Tanabe M, Sato Y and Morishima K: KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* 45(D1): D353-D361, 2017. PMID: 27899662. DOI: 10.1093/nar/gkw1092
- 22 Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z, Assempour N, Iynkkaran I, Liu Y, Maciejewski A, Gale N, Wilson A, Chin L, Cummings R, Le D, Pon A, Knox C and Wilson M: DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res* 46(D1): D1074-D1082, 2018. PMID: 29126136. DOI: 10.1093/nar/gkx1037
- 23 Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker BA, Wang J, Yu B, Zhang J and Bryant SH: PubChem Substance and Compound databases. *Nucleic Acids Res* 44(D1): D1202-D1213, 2016. PMID: 26400175. DOI: 10.1093/nar/gkv951
- 24 Luo Y, Ye GY, Qin SL, Yu MH, Mu YF and Zhong M: ATAD2 overexpression identifies colorectal cancer patients with poor prognosis and drives proliferation of cancer cells. *Gastroenterol Res Pract* 2015: 936564, 2015. PMID: 26697062. DOI: 10.1155/2015/936564

- 25 Wang T, Lin F, Sun X, Jiang L, Mao R, Zhou S, Shang W, Bi R, Lu F and Li S: HOXB8 enhances the proliferation and metastasis of colorectal cancer cells by promoting EMT via STAT3 activation. *Cancer Cell Int* 19: 3, 2019. PMID: 30622439. DOI: 10.1186/s12935-018-0717-6
- 26 Motalebzadeh J and Eskandari E: Syntrophin beta 1 (SNTB1): Candidate as a new marker for colorectal cancer metastasis. *Gene Reports* 20: 100719, 2021. DOI: 10.1016/j.genrep.2020.100719
- 27 Qi L, Song F and Ding Y: Regulatory mechanism of ITGBL1 in the metastasis of colorectal cancer. *Front Oncol* 10: 259, 2020. PMID: 32211321. DOI: 10.3389/fonc.2020.00259
- 28 Liu J, Sun GL, Pan SL, Qin MB, Ouyang R and Huang JA: Identification of hub genes in colon cancer via bioinformatics analysis. *J Int Med Res* 48(9): 300060520953234, 2020. PMID: 32961078. DOI: 10.1177/0300060520953234
- 29 Martinez-Romero J, Bueno-Fortes S, Martín-Merino M, Ramirez de Molina A and De Las Rivas J: Survival marker genes of colorectal cancer derived from consistent transcriptomic profiling. *BMC Genomics* 19(Suppl 8): 857, 2018. PMID: 30537927. DOI: 10.1186/s12864-018-5193-9
- 30 Alonso-Molero J, González-Donquiles C, Fernández-Villa T, de Souza-Teixeira F, Vilorio-Marqués L, Molina AJ and Martín V: Alterations in PGC1 α expression levels are involved in colorectal cancer risk: a qualitative systematic review. *BMC Cancer* 17(1): 731, 2017. PMID: 29121859. DOI: 10.1186/s12885-017-3725-3
- 31 Pitule P, Vycital O, Bruha J, Novak P, Hosek P, Treska V, Hlavata I, Soucek P, Kralickova M and Liska V: Differential expression and prognostic role of selected genes in colorectal cancer patients. *Anticancer Res* 33(11): 4855-4865, 2013. PMID: 24222123.
- 32 Yang B, Cao L, Liu J, Xu Y, Milne G, Chan W, Heys SD, McCaig CD and Pu J: Low expression of chloride channel accessory 1 predicts a poor prognosis in colorectal cancer. *Cancer* 121(10): 1570-1580, 2015. PMID: 25603912. DOI: 10.1002/cncr.29235
- 33 Bian Q, Chen J, Qiu W, Peng C, Song M, Sun X, Liu Y, Ding F, Chen J and Zhang L: Four targeted genes for predicting the prognosis of colorectal cancer: A bioinformatics analysis case. *Oncol Lett* 18(5): 5043-5054, 2019. PMID: 31612015. DOI: 10.3892/ol.2019.10866
- 34 Li X, Hu W, Zhou J, Huang Y, Peng J, Yuan Y, Yu J and Zheng S: CLCA1 suppresses colorectal cancer aggressiveness via inhibition of the Wnt/beta-catenin signaling pathway. *Cell Commun Signal* 15(1): 38, 2017. PMID: 28974231. DOI: 10.1186/s12964-017-0192-z
- 35 Dai L, He G, Zhang K, Guan X, Wang Y and Zhang B: Trichostatin A induces p53-dependent endoplasmic reticulum stress in human colon cancer cells. *Oncol Lett* 17(1): 660-667, 2019. PMID: 30655814. DOI: 10.3892/ol.2018.9641
- 36 Fujita K, Kubota Y, Ishida H and Sasaki Y: Irinotecan, a key chemotherapeutic drug for metastatic colorectal cancer. *World J Gastroenterol* 21(43): 12234-12248, 2015. PMID: 26604633. DOI: 10.3748/wjg.v21.i43.12234
- 37 Xiong S and Xiao GW: Reverting doxorubicin resistance in colon cancer by targeting a key signaling protein, steroid receptor coactivator. *Exp Ther Med* 15(4): 3751-3758, 2018. PMID: 29581735. DOI: 10.3892/etm.2018.5912
- 38 Han Y: Analysis of the role of the Hippo pathway in cancer. *J Transl Med* 17(1): 116, 2019. PMID: 30961610. DOI: 10.1186/s12967-019-1869-4
- 39 Snigdha K, Gangwani KS, Lapalikal GV, Singh A and Kango-Singh M: Hippo signaling in cancer: Lessons from *Drosophila* models. *Front Cell Dev Biol* 7: 85, 2019. PMID: 31231648. DOI: 10.3389/fcell.2019.00085
- 40 Wang X, Tully O, Ngo B, Zitin M and Mullin JM: Epithelial tight junctional changes in colorectal cancer tissues. *ScientificWorldJournal* 11: 826-841, 2011. PMID: 21479352. DOI: 10.1100/tsw.2011.86
- 41 Gao L, Ge C, Wang S, Xu X, Feng Y, Li X, Wang C, Wang Y, Dai F and Xie S: The role of p53-mediated signaling in the therapeutic response of colorectal cancer to 9F, a Spermine-modified Naphthalene Diimide derivative. *Cancers (Basel)* 12(3): 528, 2020. PMID: 32106543. DOI: 10.3390/cancers12030528
- 42 Poudel KR, Roh-Johnson M, Su A, Ho T, Mathsyaraja H, Anderson S, Grady WM, Moens CB, Conacci-Sorrell M, Eisenman RN and Bai J: Competition between TIAM1 and membranes balances endophilin A3 activity in cancer metastasis. *Dev Cell* 45(6): 738-752.e6, 2018. PMID: 29920278. DOI: 10.1016/j.devcel.2018.05.021
- 43 Koveitypour Z, Panahi F, Vakilian M, Peymani M, Seyed Forootan F, Nasr Esfahani MH and Ghaedi K: Signaling pathways involved in colorectal cancer progression. *Cell Biosci* 9: 97, 2019. PMID: 31827763. DOI: 10.1186/s13578-019-0361-4
- 44 Dagia NM, Agarwal G, Kamath DV, Chetrapal-Kunwar A, Gupte RD, Jadhav MG, Dadarkar SS, Trivedi J, Kulkarni-Almeida AA, Kharas F, Fonseca LC, Kumar S and Bhonde MR: A preferential p110 α /gamma PI3K inhibitor attenuates experimental inflammation by suppressing the production of proinflammatory mediators in a NF-kappaB-dependent manner. *Am J Physiol Cell Physiol* 298(4): C929-C941, 2010. PMID: 20089935. DOI: 10.1152/ajpcell.00461.2009
- 45 Chang SN, Lee JM, Oh H, Kim U, Ryu B and Park JH: Troglitazone inhibits the migration and invasion of PC-3 human prostate cancer cells by upregulating E-cadherin and glutathione peroxidase 3. *Oncol Lett* 16(4): 5482-5488, 2018. PMID: 30250621. DOI: 10.3892/ol.2018.9278
- 46 Gupta I, Al Farsi H, Jabeen A, Skenderi F, Al-Thawadi H, AlAhmad YM, Al Moustafa AE and Vranic S: High-risk human papillomaviruses and Epstein-Barr virus in colorectal cancer and their association with clinicopathological status. *Pathogens* 9(6): 452, 2020. PMID: 32521661. DOI: 10.3390/pathogens9060452
- 47 Fernandes Q, Gupta I, Vranic S and Al Moustafa AE: Human papillomaviruses and Epstein-Barr virus interactions in colorectal cancer: a brief review. *Pathogens* 9(4): 300, 2020. PMID: 32325943. DOI: 10.3390/pathogens9040300
- 48 Goetgebuer RL, van der Woude CJ, de Ridder L, Doukas M and de Vries AC: Clinical and endoscopic complications of Epstein-Barr virus in inflammatory bowel disease: an illustrative case series. *Int J Colorectal Dis* 34(5): 923-926, 2019. PMID: 30739187. DOI: 10.1007/s00384-019-03257-7
- 49 Chen H, Chen XZ, Waterboer T, Castro FA and Brenner H: Viral infections and colorectal cancer: a systematic review of epidemiological studies. *Int J Cancer* 137(1): 12-24, 2015. PMID: 25186851. DOI: 10.1002/ijc.29180
- 50 Kim YH and Park TJ: Cellular senescence in cancer. *BMB Rep* 52(1): 42-46, 2019. PMID: 30526772.
- 51 Lee S and Lee JS: Cellular senescence: a promising strategy for cancer therapy. *BMB Rep* 52(1): 35-41, 2019. PMID: 30526771.
- 52 Tan HK, Moad AI and Tan ML: The mTOR signalling pathway in cancer and the potential mTOR inhibitory activities of natural

- phytochemicals. *Asian Pac J Cancer Prev* 15(16): 6463-6475, 2014. PMID: 25169472. DOI: 10.7314/apjcp.2014.15.16.6463
- 53 Qi Y, Qi H, Liu Z, He P and Li B: Bioinformatics analysis of key genes and pathways in colorectal cancer. *J Comput Biol* 26(4): 364-375, 2019. PMID: 30810359. DOI: 10.1089/cmb.2018.0237
 - 54 Gonen N and Assaraf YG: Antifolates in cancer therapy: structure, activity and mechanisms of drug resistance. *Drug Resist Updat* 15(4): 183-210, 2012. PMID: 22921318. DOI: 10.1016/j.drug.2012.07.002
 - 55 He Z, Tang F, Lu Z, Huang Y, Lei H, Li Z and Zeng G: Analysis of differentially expressed genes, clinical value and biological pathways in prostate cancer. *Am J Transl Res* 10(5): 1444-1456, 2018. PMID: 29887958.
 - 56 Marisa L, de Reyniès A, Duval A, Selves J, Gaub MP, Vescovo L, Etienne-Grimaldi MC, Schiappa R, Guenot D, Ayadi M, Kirzin S, Chazal M, Fléjou JF, Benchimol D, Berger A, Lagarde A, Pencreach E, Piard F, Elias D, Parc Y, Olschwang S, Milano G, Laurent-Puig P and Boige V: Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Med* 10(5): e1001453, 2013. PMID: 23700391. DOI: 10.1371/journal.pmed.1001453
 - 57 An N, Shi X, Zhang Y, Lv N, Feng L, Di X, Han N, Wang G, Cheng S and Zhang K: Discovery of a novel immune gene signature with profound prognostic value in colorectal cancer: a model of cooperativity disorientation created in the process from development to cancer. *PLoS One* 10(9): e0137171, 2015. PMID: 26325386. DOI: 10.1371/journal.pone.0137171
 - 58 Kogo R, Shimamura T, Mimori K, Kawahara K, Imoto S, Sudo T, Tanaka F, Shibata K, Suzuki A, Komune S, Miyano S and Mori M: Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. *Cancer Res* 71(20): 6320-6326, 2011. PMID: 21862635. DOI: 10.1158/0008-5472.CAN-11-1021
 - 59 Iwaya T, Yokobori T, Nishida N, Kogo R, Sudo T, Tanaka F, Shibata K, Sawada G, Takahashi Y, Ishibashi M, Wakabayashi G, Mori M and Mimori K: Downregulation of miR-144 is associated with colorectal cancer progression via activation of mTOR signaling pathway. *Carcinogenesis* 33(12): 2391-2397, 2012. PMID: 22983984. DOI: 10.1093/carcin/bgs288
 - 60 Tsukamoto S, Ishikawa T, Iida S, Ishiguro M, Mogushi K, Mizushima H, Uetake H, Tanaka H and Sugihara K: Clinical significance of osteoprotegerin expression in human colorectal cancer. *Clin Cancer Res* 17(8): 2444-2450, 2011. PMID: 21270110. DOI: 10.1158/1078-0432.CCR-10-2884
 - 61 Kagawa Y, Matsumoto S, Kamioka Y, Mimori K, Naito Y, Ishii T, Okuzaki D, Nishida N, Maeda S, Naito A, Kikuta J, Nishikawa K, Nishimura J, Haraguchi N, Takemasa I, Mizushima T, Ikeda M, Yamamoto H, Sekimoto M, Ishii H, Doki Y, Matsuda M, Kikuchi A, Mori M and Ishii M: Cell cycle-dependent Rho GTPase activity dynamically regulates cancer cell motility and invasion in vivo. *PLoS One* 8(12): e83629, 2013. PMID: 24386239. DOI: 10.1371/journal.pone.0083629
 - 62 Liang Y, Wang B, Chen S, Ye Z, Chai X, Li R, Li X, Kong G, Li Y, Zhang X, Che Z, Xie Q, Lian J, Lin B, Zhang X, Huang X, Huang W, Qiu X and Zeng J: Beta-1 syntrophin (SNTB1) regulates colorectal cancer progression and stemness via regulation of the Wnt/ β -catenin signaling pathway. *Ann Transl Med* 9(12): 1016, 2021. PMID: 34277816. DOI: 10.21037/atm-21-2700
 - 63 Luo Y, Ye GY, Qin SL, Yu MH, Mu YF and Zhong M: ATAD2 overexpression identifies colorectal cancer patients with poor prognosis and drives proliferation of cancer cells. *Gastroenterol Res Pract* 2015: 936564, 2015. PMID: 26697062. DOI: 10.1155/2015/936564
 - 64 Wang T, Lin F, Sun X, Jiang L, Mao R, Zhou S, Shang W, Bi R, Lu F and Li S: HOXB8 enhances the proliferation and metastasis of colorectal cancer cells by promoting EMT via STAT3 activation. *Cancer Cell Int* 19: 3, 2019. PMID: 30622439. DOI: 10.1186/s12935-018-0717-6
 - 65 Qi L, Song F and Ding Y: Regulatory mechanism of ITGBL1 in the metastasis of colorectal cancer. *Front Oncol* 10: 259, 2020. PMID: 32211321. DOI: 10.3389/fonc.2020.00259
 - 66 Liang K, Zhou G, Zhang Q, Li J and Zhang C: Expression of hippo pathway in colorectal cancer. *Saudi J Gastroenterol* 20(3): 188-194, 2014. PMID: 24976283. DOI: 10.4103/1319-3767.133025

Received July 16, 2021

Revised September 3, 2021

Accepted September 16, 2021