*In silico* transcriptomic analysis of ascending colon cancer uneartths known and novel genes and gene sets regard to characteristic features of colon cancer

Can Türk

Department of Medical Microbiology, Faculty of Medicine, Lokman Hekim University, Ankara, Turkey

**Abstract**

**Objectives:** Colon cancer emerges as a serious health problem in both men and women. Cancers in the colon have different genotypes and phenotypes according to the anatomical region. Tumors in ascending colon are usually diagnosed later, but it is more malignant than the descending and transverse colon, and the survival rates of patients are lower than other regions. The purpose of this study was to determine significantly high or low expressed genes in the ascending colon tumors by comparing all genome information obtained from cancer samples of ascending, transverse and descending colon. In concordance with all this information, another aim of the study was to identify the pathways to which the genes obtained from the colon in the large intestine and to determine their relationship with each other and to correlate them with the characteristics of cancer.

**Methods:** Gene expression values for three subtypes of colon cancer as ascending, transverse, and descending were obtained from GEO (Gene Expression Omnibus) (GSE41258). Data included a total of 47 ascending, 18 transverse and 31 descending colon cancer patient samples. Linear regression analysis was performed to determine differentially expressed genes. Gene Cluster 3.0 was used in order to cluster the genes hierarchically. In addition to linear regression and hierarchical clustering, network analysis with multivariable genes was performed in Cytoscape application 3.8.2 using GeneMANIA. GSEA 4.1.0 (Gene Set Enrichment Analysis) was performed to understand the different genes among the specified groups.

**Results:** As a result of these analyses, it was determined that there were 85 genes with high expression and 139 genes with low expression in the ascending colon tumor samples. It has been shown that these genes can differentiate tumor samples in the ascending colon better than tumor samples in other colon regions.

**Conclusion:** Our findings are important for understanding the genome of ascending colon tumors; if these findings are confirmed *in vitro* and clinically, it may have potential to be revealed that the identified genes also have biomarker properties for tumors in the ascending colon.

**Keywords:** ascending colon; descending colon; transcriptomic analysis; transverse colon

**Introduction**

The large intestine is approximately 1.5–2 meters long and consists of caecum, colons and rectum. The rectum is the last part before the anus and this part is also known as the area where feces is stored. On the other hand, the colon, forms the large part of the large intestine. Colon cancers basically occur as a result of an abnormality in this part. Studies carried out at the molecular level show that the formation of colon cancer occurs through a complex mechanism influenced by many factors. Genetic factors trigger the formation of colon cancer cells as well as the effects of lifestyle factors such as smoking and eating habits. Studies show that mainly CIMP (CpG island methylator phenotype), MSI (microsatellite instability), and additionally, CIN (chromosomal instability) mechanisms play a role in cancer development.
It has been shown that approximately 20% to 30% of patients with colon cancer have abnormalities in the CIMP pathway. Hypermethylation in the promotor sequences of the cell plays an essential role in the pattern of gene expression. Basically, CpG dinucleotide sequences are known to locate in this promotor region. The trigger for colon cancer cells is that sequences are not hypermethylated properly. In addition, this imprecision also affects critical cell mechanisms such as apoptosis, invasion, angiogenesis, cell cycle regulation, DNA adhesion and repair.\[1-4\]

According to statistical studies, MSI is responsible for 15% of colon cancers. MSI occurs due to DNA incompatibility, is involved in the process of DNA replication, mutations in some genes involved in the mechanism. These mutations generally inactivate the functions of genes. It plays an essential role in the protein synthesis of MMR genes. These proteins cause a decrease in polymerase function for to recognize and correct these defects, resulting in anomalies occurring on the microsatellite during replication. Mutations that arise because of the function of the recovery system accumulate and trigger formation of colon cancer cells.\[1-4,6\]

Colon cancer is a common type of cancer among gastrointestinal cancers worldwide. The prevalence of colon cancer varies depending on age. I.e., while it is 1.6% in the 50–60 age group, it is known that this rate increases up to 3% over the age of 70. In addition, studies show that the incidence of colon cancer may vary depending on gender. It has been shown that the incidence rate in women is higher than in men.\[7-10\]

Considering the mechanism of colon cancer, the importance can be seen more clearly. Basically; CIN, MSI, and CIMP play critical roles on colon cancer. Abnormalities caused by these mechanisms increase mutagenic activity in tumor suppressors and oncogenes. Critically, these mutations lead to an increase in the number of cancer stem cells, which play an important role in the onset of tumor formation. In addition, the acceleration of mutation accumulation also accelerates the epigenetic change of cells.\[7,7-10\]

Importantly, the characteristics and effects of colon cancer may vary according to anatomical regions. It consists of three main anatomical parts, respectively, ascending colon, transverse colon and descending colon. Generally, about 45% of colon cancer is located in the left colon region. However, in recent years, studies have showed that right colon (cecum and ascending colon) cancers have reached the rate of up to 25% and the reasons for this increase include the increase of the populatio

Materials and Methods
Data Collection and Normalization
Gene expression values for three subtypes of colon cancer as ascending, transverse, and descending, respectively, were obtained from (Gene Expression Omnibus) (GSE41258, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE41258).\[15\]

GSE41258 data includes a total of 47 ascending, 18 transverse and 31 descending colon cancer patient samples. In the data gene expressions profiled by array-based. Sample codes and anatomic locations used within this data are shown in Appendix 1. The obtained raw data were normalized by RMA (Robust Multi-Array) normalization algorithm in R 3.6.1.\[18\]

Linear Regression Analysis
Linear regression analysis was performed to determine differentially expressed genes. Student’s t-test (p-value) was calculated among the ascending, and transverse and ascending, and descending groups, and as a result of this
analysis, genes with p<0.05 were selected. The analysis continued with the genes expressing these genes significantly in both groups.

**Hierarchical Clustering**

Gene Cluster 3.0 application\[^{19}\] was applied in order to cluster the determined statistically significant genes hierarchically. This clustering is based on Euclidean distance with a similarity metric limit for both genes and sequences, as well as the full link aggregation method. This methodology supports differentiating and distinguishing statistically highly variable genes.

**Network and Pathway Analysis**

In addition to linear regression and hierarchical clustering, network analysis with multivariable genes was performed in Cytoscape application 3.8.2\[^{20}\] using GeneMANIA.\[^{21}\] This app helps to better understand the correlation of statistically significant and highly variable genes by showing genetic interaction and their co-expression. Cytoscape also allows to illuminate the link between identified genes and even with each other.

In addition, the online DAVID: Bioinformatics Resources Tool was used to understand the respective pathways of these genes.\[^{22,23}\] This tool allows to show the proper pathway linked to these genes.

**Gene Set Enrichment Analysis (GSEA)**

GSEA 4.1.0 (Gene set enrichment analysis) was performed to understand the different genes among the specified groups.\[^{24}\] In this study, gene set enrichment analysis was performed among the ascending colon cancer patient groups and transverse colon cancer patient groups, ascending colon cancer patient groups and decreasing colon cancer patient groups. GSEA was performed by gene expression of GSE41258 data. As a result of gene set enrichment analysis, the enriched pathways and the most important and associated genes are determined by comparing the ascending and transverse and ascending and descending groups and their gene expression levels.

**Results**

After the normalization process of the data obtained was completed, the analyzes were continued with a total of 13,432 genes (21,225 Probe Sets). Ascending tumor samples were compared with transverse and descending tumor samples. When the genes belonging to ascending and transverse colon cancer were encountered, it was determined that a total of 1035 genes were expressed differently. This number was determined as 1531 when the increasing and decreasing subgroups were compared. 224 genes were found in common in these two groups (Figure 1). Appendix 2 includes t-test p-values (p<0.05), which expressed statistically significant.

Then, upregulated genes in the ascending cohort and downregulated in the transverse and descending cohorts, and vice versa, were identified. I.e., while MLH3 and APC genes have a higher expression value when compared to other subgroups (Figure 2a), and BAX and PMS2 genes are expressed lower (Figure 2b). The statistically significant genes cluster colon cancer subtypes (ascending, transverse and descending) in a hierarchical manner are shown in Figure 3.

Network analysis was performed to understand the network link between these 50 genes that are upregulated in ascending colon cancer and down-regulated in transverse and descending colon cancer, and vice versa. The connecting line between genes illuminates the network of these genes. The thickness of the binding line determines the binding strength of the respective genes. The thickest lines show that it has been determined that the connection between these genes has been determined by studying more precisely. In addition, the black nodes indicate the target genes given by the authors. On the other hand, gray nodes show genes associated with genes determined by GeneMANIA application in Cytoscape. The co-expression of genes is shown in Figure 4a and the genetic interaction between these genes is shown in Figure 4b. The genes that are available in the Online Mendelian Inheritance in Man (OMIM) database in the DAVID application and are statistically significantly up or down regulated in the ascending colon tumor subtype are shown in Table 1.
As a result of GSEA, it was determined that a total of 11 gene sets were enriched in the ascending tumor type, while 6 gene sets were not enriched in the same group (Appendix 3). Among these SPLICEOMAL_SNRNP_ASSEMBLY and TRANSITION_METAL_ION_HOMEOSTASIS contain the most gene sets. Therefore, they can be consid-

Table 1
The major 4 genes that are determined from OMIM database in the DAVID program and the diseases associated with these genes.

| Gene symbols | OMIM disease |
|--------------|--------------|
| APC, WNT signaling pathway regulator (APC) | Colorectal cancer, somatic, Hepatoblastoma, somatic, Desmoid disease, hereditary, Adenomatous polyposis coli, Brain tumor-polyposis syndrome 2, Gardner syndrome, Gastric cancer, somatic, Adenoma, periampullary, somatic |
| BCL2 associated X, apoptosis regulator (BAX) | Colorectal cancer, somatic, T-cell acute lymphoblastic leukemia, somatic |
| PMS1 homolog 2, mismatch repair system component (PMS2) | Mismatch repair cancer syndrome, Colorectal cancer, hereditary nonpolyposis, type 4 |
| mutL homolog 3 (MLH3) | Colorectal cancer, somatic, Endometriai cancer, susceptibility to, Colorectal cancer, hereditary nonpolyposis, type 7 |

Figure 2. (a) Expression profile of the MLH3, APC, BAX and PMS2 genes. (a) MLH3 and APC significantly upregulated in ascending colon cancer compared to transverse and descending colon cancers; (b) BAX and PMS2 genes significantly down-regulated in ascending colon cancer compared to transverse and descending colon cancers. Numbers near the genes are Probe Set IDs; (Probe Set ID: The identifier that refers to a set of probe pairs selected to represent expressed sequences on an array).
ered to have a more important roles. SPLICEOSOMAL_ SNRNP_ASSEMBLY is enriched in ascending colon tumor (Figure 5a), while the TRANSITION_METAL_ ION_HOMEOSTASIS gene set is enriched in transverse and descending colon cancer (Figure 5b).

**Discussion**

Studies have intensified in the early 2000s to reveal the molecular differences of tumors in the right and left colon regions. In a comprehensive study by Guinney et al. 25 4 subgroups with different biological behavior were identified, taking into account the many expression sequences belonging to both regions of the colon.

In the present study, high or low expressed genes were detected in tumors belonging to the ascending region compared to other colon tumors. The network between these genes as well as the pathways were determined. Accordingly, 4 genes with statistically different expression values in the ascending colon cancer samples are associated with colon cancer based on the OMIM database. These genes are *APC* (Adenomatous polyposis coli), *BAX*, *PMS2* and *MLH3*.

*APC* gene is one of the most critical genes that affect colon cancer formation. The *APC* gene is used as a negative regulator for the Wnt signaling pathway involved in colon cancer development. It also takes part in phosphorylation occurring in cells. Studies show that the *APC* gene increases the expression of the *MMP9* gene using the JNK signaling pathway. Importantly, this indicates that the

---

**Figure 3.** Hierarchical clustering of 224 statistically significant genes between groups. Genes highlighted in green represent genes with low expression, while red colored groups represent high expression.

**Figure 4.** Representation of the co-expression and interactions of 100 genes. (a) statistically significant co-expression of 100 genes both among themselves and between other genes; (b) Representation of the genetic interactions of 100 genes that have statistically significant expression, both among themselves and between other genes.
change in gene expression pattern is critical for the development of colorectal tumor cells. In other words, mutations in the APC gene play a role in the course of sporadic colorectal cancers, indicating that this gene is not only responsible for familial adenomatous polyposis (FAP). As a result of the researches, colorectal tumor formation occurs with the gradual occurrence of histological changes triggered by genetic changes and “adenoma-carcinoma” sequence as a result of mutation from tumor suppressor or oncogenic genes. These changes are known to occur as a result of loss of function resulting from mutation in the APC gene. In order to inactivate the critical function of the APC gene and trigger the formation of cancer cells, genetic instability and clonal expansion must basically occur. Because, these two changes can enable the activation of genes that support malignant transformation and tumor progression.

In parallel with the results we obtained, Du et al. showed that high expression of the APC gene is associated with poor prognosis in gastritis cancer. In our results, we determined that this gene has a higher expression in the ascending colon samples. This may be one of the reasons why colon cancers occurring in the right region have a worse prognosis than the left side.

In another study, it was shown that the APC gene is associated with a poor prognosis in microsatellite-stable proximal colon cancer supports the findings we obtained.

Our results show that the BAX gene has significantly less expression in colon tumors. Basically, BAX protein is known to promote cell death. Thus, it can inactivate the expression of cancer cells. Studies on the importance of the BAX gene have shown that mutations in the BAX gene reduce the apoptotic index of colorectal cancer cells. It has been determined that this situation is seen in 50% of colorectal cancer cases. In addition, in similar studies, high expression of the BAX gene shows that it can be a good prognostic marker for colon cancer patients, except for the ascending colon.

One of the four basic sensitivity genes in Lynch syndrome (LS), the most common cancer syndrome in the world, is the PMS1 Homologous 2, Mismatch Repair System Component (PMS2) gene. However, unfortunately it is not known whether the decrease in the gene expression value of the PMS2 gene has an effect on the repair mechanism and, critically, how this effect may occur.

The study by Kasela et al. shows that MMR activity is significantly reduced in cells in which the PMS2 gene is knocked out. These findings suggest that low expression of the PMS2 gene in colon tumors that rise in parallel with the findings we obtained as a result of our analysis causes a decrease in DNA mismatch repair (MMR), leading to poor prognosis of the colon.

Another gene that we found to have higher expression in the colon cancer samples compared to other types of colon cancer is the MLH3 gene. Although many studies show that descending colon cancer cases have a better survival rate, it is known that colon tumors in the right region have a worse prognosis than the left. Although not statistically significant, a study by Zhao et al. on ovarian can-
cancer showed that higher expression of the MLH3 gene is associated with lower survival. In addition, MLH3 (MutL homolog 3), MSH2 (MutS homolog 2) and MSH3 (MutS homolog 3) genes are also known to be frequently seen in colorectal cancer. These genes have also been identified as potential genetic markers for personalized therapy, showing that they are associated with chemoresistance.

As a result of GSE analysis, important gene sets associated with colorectal cancer progression and metastasis were determined. The enrichment of spliceosomal snRNP assembly gene sets in the ascending colon suggests that spliceostatin A, which has the capacity to target pladieno-lide compounds and spliceosome of these types of colon cancers, may be anticancer potential drugs. Based on the fundamental role of DNA methylation in colon cancer development, the application of DNMT inhibitors for the treatment of colon cancer patients, especially patients with DNA hypermethylation, is recommended as a result of studies.

Our results showed that the gene sets of Methyl transferase activity are enriched in ascending colon tumors. This suggests that such agents may be more effective in the treatment of this type of colon cancer subgroups especially. On the other hand, our analysis showed that Transition metal ion homeostasis gene sets enriched in other colon types except for ascending colon tumors. This suggests the use of drugs that target transition metal homeostasis such as ferristatin II, clioquinol, and omeprazole in colon cancers other than ascending colon cancers.

In addition to the characteristic features of tumors that occur in different parts of the colon, many studies have shown that their response to treatment can be very different. In this study, differently expressed genes and pathways were determined by comparing the whole genome profiles of tumors in different regions of the colon. A better understanding of the biology of the tumor allows more effective treatment. This may provide more effective treatment choices in the future. Importantly, it is crucial to validate the results of this study in vitro and clinically.

Conflict of Interest
No conflicts declared.

Ethics Approval
No ethics approval needed.

Funding
This research did not receive any specific grant from funding agencies in the public or commercial.
19. de Hoon MJL, Imoto S, Nolan J, Miyano S. Open source clustering software. Bioinformatics 2004;20:1453–4.
20. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498–504.
21. Franz M, Rodriguez H, Lopes C, Zuberi K, Montejo J, Bader GD, Morris Q. GeneMANIA update 2018. Nucleic Acids Res 2018;46:W60–W4.
22. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 2009;37:1–13.
23. Subramanian A, Tamayo P, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 2005;102:15545–50.
24. Guinney J, Dietmann R, Wang X, de Reyniès A, Schlicker A, Marisa L, Roepman P, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 2005;102:15545–50.
25. Zhang L, Shay JW. Multiple roles of Apc and its therapeutic implications in colorectal cancer. Nat Rev Cancer 2017;17:548–59.
26. Zhou X, Li S, Zhao M, Zhu H, Zhu X. Prognostic values of DNA mismatch repair genes in ovarian cancer patients treated with platinum-based chemotherapy. Archives Gynecol Obstet 2018;297:147–53.
27. van Alphen RJ, Wiemer EAC, Burger H, Eskens FALM. The spliceosome as target for anticancer treatment. Br J Cancer 2009;100:228–32.
28. Cervena K, Siskova A, Buchler T, Vodicka P, Vymetalkova V. Methylation-based therapies for colorectal cancer. Cells 2020;9:1540.
29. Turk C. In silico transcriptomic analysis of ascending colon cancer unearths known and novel genes and gene sets with characteristic features of colon cancer. Anatomy 2021;15(1):11–25.

ORCID ID: C. Türk 0000-0003-1514-7294

Conflict of interest statement: No conflicts declared.

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 Unported (CC BY-NC-ND4.0) Licence (https://creativecommons.org/licenses/by-nc-nd/4.0/) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited. How to cite this article: Türk C. In silico transcriptomic analysis of ascending colon cancer unearths known and novel genes and gene sets with characteristic features of colon cancer. Anatomy 2021;15(1):11–25.
### Appendix 1

Sample codes and anatomic locations of colon cancer patients.

| Sample code          | Anatomic location      | Sample code          | Anatomic location      |
|----------------------|------------------------|----------------------|------------------------|
| GSM1012286_00620AR1  | Ascending colon        | GSM1012591_C019AR1   | Ascending colon        |
| GSM1012308_02500AR1  | Ascending colon        | GSM1012592_C019HR1   | Ascending colon        |
| GSM1012320_03283AR1  | Ascending colon        | GSM1012627_C0323AR1   | Ascending colon        |
| GSM1012326_03519AR2  | Ascending colon        | GSM1012628_C0323H    | Ascending colon        |
| GSM1012327_03519HR2  | Ascending colon        | GSM1012631_C0330AR3   | Ascending colon        |
| GSM1012350_04276AR2  | Ascending colon        | GSM1012634_C0334AR3   | Ascending colon        |
| GSM1012351_04276HR2  | Ascending colon        | GSM1012645_C0487AR3   | Ascending colon        |
| GSM1012358_04800AR4  | Ascending colon        | GSM1012646_C0487HR3   | Ascending colon        |
| GSM1012359_04800HR4  | Ascending colon        | GSM1012307_02308AR1   | Descending colon       |
| GSM1012361_05025AR4  | Ascending colon        | GSM1012325_03465AR3   | Descending colon       |
| GSM1012366_05629AR4  | Ascending colon        | GSM1012336_03706AR2   | Descending colon       |
| GSM1012368_05786AR2  | Ascending colon        | GSM1012337_03706HR2   | Descending colon       |
| GSM1012369_05786HR3  | Ascending colon        | GSM1012346_04176AR2   | Descending colon       |
| GSM1012371_05885AR2  | Ascending colon        | GSM1012347_04176HR3   | Descending colon       |
| GSM1012384_06404AR2  | Ascending colon        | GSM1012374_06220AR2   | Descending colon       |
| GSM1012399_08018AR2  | Ascending colon        | GSM1012381_06706AR2   | Descending colon       |
| GSM1012416_09468AR2  | Ascending colon        | GSM1012385_06997AR3   | Descending colon       |
| GSM1012423_10194AR2  | Ascending colon        | GSM1012389_07427AR2   | Descending colon       |
| GSM1012426_10264AR3  | Ascending colon        | GSM1012408_09054AR2   | Descending colon       |
| GSM1012447_14475AR3  | Ascending colon        | GSM1012421_09811AR2   | Descending colon       |
| GSM1012461_A1516AR4  | Ascending colon        | GSM1012429_10630AR3   | Descending colon       |
| GSM1012463_A1716AR4  | Ascending colon        | GSM1012445_13321AR3   | Descending colon       |
| GSM1012468_A2367AR4  | Ascending colon        | GSM1012446_13357AR3   | Descending colon       |
| GSM1012494_A5135AR2  | Ascending colon        | GSM1012456_A0702AR4   | Descending colon       |
| GSM1012495_A5135AR2_ez | Ascending colon     | GSM1012469_A2434AR3   | Descending colon       |
| GSM1012496_A5135DR2  | Ascending colon        | GSM1012474_A3536AR4   | Descending colon       |
| GSM1012499_A5320AR3  | Ascending colon        | GSM1012478_A4248AR4   | Descending colon       |
| GSM1012539_C0123AR3  | Ascending colon        | GSM1012544_C0128AR3   | Descending colon       |
| GSM1012540_C0123HR1  | Ascending colon        | GSM1012545_C0128BR1   | Descending colon       |
| GSM1012547_C0134AR1  | Ascending colon        | GSM1012565_C0151AR1   | Descending colon       |
| GSM1012548_C0134HR1  | Ascending colon        | GSM1012577_C0168AR3   | Descending colon       |
| GSM1012549_C0136AR1  | Ascending colon        | GSM1012580_C0172AR1   | Descending colon       |
| GSM1012550_C0136HR1  | Ascending colon        | GSM1012581_C0172HR1   | Descending colon       |
| GSM1012551_C0136KR1  | Ascending colon        | GSM1012595_C0200AR3   | Descending colon       |
| GSM1012552_C0136UR1  | Ascending colon        | GSM1012603_C0230BH   | Descending colon       |
| GSM1012572_C0157AR1  | Ascending colon        | GSM1012615_C0273A    | Descending colon       |
| GSM1012573_C0157H    | Ascending colon        | GSM1012617_C0283AR3   | Descending colon       |
| GSM1012589_C0193AR1  | Ascending colon        | GSM1012629_C0329AR1   | Descending colon       |
| GSM1012590_C0193HR1  | Ascending colon        | GSM1012630_C0329HR1   | Descending colon       |
Appendix 1 [Continued]
Sample codes and anatomic locations of colon cancer patients.

| Sample code          | Anatomic location | Sample code          | Anatomic location |
|----------------------|-------------------|----------------------|-------------------|
| GSM1012297_00990AR1  | Sigmoid colon     | GSM1012561_C0147AR1  | Sigmoid colon     |
| GSM1012303_02184AR2  | Sigmoid colon     | GSM1012562_C0147AR3  | Sigmoid colon     |
| GSM1012304_02184HR2  | Sigmoid colon     | GSM1012563_C0147HR1  | Sigmoid colon     |
| GSM1012310_02679AR1  | Sigmoid colon     | GSM1012569_C0154AR1  | Sigmoid colon     |
| GSM1012311_02679BR1  | Sigmoid colon     | GSM1012576_C0159AR3  | Sigmoid colon     |
| GSM1012314_02815AR1  | Sigmoid colon     | GSM1012578_C0170AR1  | Sigmoid colon     |
| GSM1012315_02815HR1  | Sigmoid colon     | GSM1012579_C0171AR1  | Sigmoid colon     |
| GSM1012317_03023AR1  | Sigmoid colon     | GSM1012584_C0180AR1  | Sigmoid colon     |
| GSM1012318_03023HR1  | Sigmoid colon     | GSM1012585_C0180HR1  | Sigmoid colon     |
| GSM1012319_03156AR1  | Sigmoid colon     | GSM1012586_C0181AR3  | Sigmoid colon     |
| GSM1012354_04494AR4  | Sigmoid colon     | GSM1012587_C0186AR3  | Sigmoid colon     |
| GSM1012355_04494HR4  | Sigmoid colon     | GSM1012588_C0192A    | Sigmoid colon     |
| GSM1012367_05708AR2  | Sigmoid colon     | GSM1012593_C0198AR1  | Sigmoid colon     |
| GSM1012375_06265AR2  | Sigmoid colon     | GSM1012594_C0198HR1  | Sigmoid colon     |
| GSM1012379_06657AR2  | Sigmoid colon     | GSM1012611_C0257AR3  | Sigmoid colon     |
| GSM1012380_06657HR3  | Sigmoid colon     | GSM1012618_C0285AR1  | Sigmoid colon     |
| GSM1012386_07061AR2  | Sigmoid colon     | GSM1012619_C0295AR3  | Sigmoid colon     |
| GSM1012387_07145AR2  | Sigmoid colon     | GSM1012620_C0297AR1  | Sigmoid colon     |
| GSM1012397_07939AR2  | Sigmoid colon     | GSM1012621_C0297HR1  | Sigmoid colon     |
| GSM1012401_08061AR2  | Sigmoid colon     | GSM1012624_C0312AR3  | Sigmoid colon     |
| GSM1012405_08168AR2  | Sigmoid colon     | GSM1012625_C03156H   | Sigmoid colon     |
| GSM1012407_08792AR2  | Sigmoid colon     | GSM1012631_02832AR1  | Transverse colon  |
| GSM1012409_09077AR2  | Sigmoid colon     | GSM1012632_03531AR2  | Transverse colon  |
| GSM1012411_09297AR2  | Sigmoid colon     | GSM1012633_03657AR2  | Transverse colon  |
| GSM1012413_09394AR2  | Sigmoid colon     | GSM1012633_03657HR2  | Transverse colon  |
| GSM1012424_10216AR3  | Sigmoid colon     | GSM1012634_03862AR2  | Transverse colon  |
| GSM1012428_10512AR3  | Sigmoid colon     | GSM1012634_03862HR2  | Transverse colon  |
| GSM1012439_12292AR3  | Sigmoid colon     | GSM1012635_04388AR2  | Transverse colon  |
| GSM1012442_12847AR3  | Sigmoid colon     | GSM1012636_05424AR4  | Transverse colon  |
| GSM1012462_A1644AR4  | Sigmoid colon     | GSM1012692_07632AR2  | Transverse colon  |
| GSM1012492_A4947AR4  | Sigmoid colon     | GSM1012393_07662AR2  | Transverse colon  |
| GSM1012502_A5627AR3  | Sigmoid colon     | GSM1012396_07925AR2  | Transverse colon  |
| GSM1012503_A5627BR3  | Sigmoid colon     | GSM1012410_09185AR2  | Transverse colon  |
| GSM1012526_C0101AR3  | Sigmoid colon     | GSM1012437_12237AR3  | Transverse colon  |
| GSM1012527_C0104AR3  | Sigmoid colon     | GSM1012467_A2226AR4  | Transverse colon  |
| GSM1012531_C0112AR1  | Sigmoid colon     | GSM1012507_A6141AR   | Transverse colon  |
| GSM1012535_C0115AR3  | Sigmoid colon     | GSM1012529_C0111AR1  | Transverse colon  |
| GSM1012541_C0124AR3  | Sigmoid colon     | GSM1012530_C0111HR1  | Transverse colon  |
| GSM1012553_C0137AR1  | Sigmoid colon     | GSM1012570_C0155AR3  | Transverse colon  |
| GSM1012554_C0137HR1  | Sigmoid colon     | GSM1012571_C0156AR3  | Transverse colon  |

| Sample code          | Anatomic location | Sample code          | Anatomic location |
|----------------------|-------------------|----------------------|-------------------|
| GSM1012570_C0156AR3  | Transverse colon  | GSM1012571_C0156AR3  | Transverse colon  |
## Appendix 2

T-test p-values which expressed statistically significant (p<0.05).

| Probe set | Gene symbol | T-test of ascending vs transverse colon | T-test of ascending vs descending colon |
|-----------|-------------|----------------------------------------|----------------------------------------|
| 201888_s_at | IL13RA1 | 0.00016271505536615 | 0.00521747486476386 |
| 205844_s_at | VNN1 | 0.000385136080291112 | 0.0077639644507052 |
| 47530_at | C9orf156 | 0.00081241732984075 | 0.0365888700564041 |
| 206122_at | SOX15 | 0.00093693288690665 | 0.0319002678319461 |
| 212665_at | TPARP | 0.00109740862530723 | 0.0149345736824881 |
| 210219_at | SPI100 | 0.00129941458167625 | 0.02945112630853 |
| 216300_x_at | RARA | 0.001510363838347 | 0.020880829302783 |
| 209397_at | TM4SF4 | 0.0015154586859597 | 0.0033539301875218 |
| 213664_at | SLC1A1 | 0.0015657003086825 | 0.00415570175875679 |
| 215427_s_at | ZCCHC14 | 0.00169179299291247 | 0.012950331872678 |
| 210651_s_at | EPHB2 | 0.00191836064227171 | 0.00348119288028227 |
| 213823_at | HOXA11 | 0.00207814621070703 | 0.0121059077242761 |
| 214191_at | KAT1 | 0.0022316379888523 | 0.0268164312486215 |
| 205730_s_at | ABUML3 | 0.00237677641741103 | 0.0110983916886486 |
| 209320_at | ADCY3 | 0.002429040211383 | 0.041325710012813 |
| 217415_at | POLR2A | 0.00269575520034894 | 0.0443829653243419 |
| 219021_at | RNF121 | 0.00270128512746877 | 0.0241668884278377 |
| 204843_s_at | PRKAR2A | 0.00270329662456024 | 0.02295623964002 |
| 202459_s_at | LPN2 | 0.00296373853260737 | 0.0195009066395492 |
| 203139_at | DAPK1 | 0.002985384623663 | 0.0081694250979891 |
| 217165_x_at | MT1F | 0.00324876323465622 | 0.0039603427170252 |
| 218529_at | CD320 | 0.00351895978865357 | 0.018049949266918 |
| 210143_at | ANXA10 | 0.00361543881448043 | 0.0221704749647477 |
| 208559_at | PON1 | 0.00369666604226198 | 0.0122623122943415 |
| 221268_s_at | SGPP1 | 0.00378110101365417 | 0.010261200115258 |
| 202693_s_at | STK17A | 0.00384950815651231 | 0.0039142818604062 |
| 206330_s_at | SHC3 | 0.00394389390077759 | 0.024575359746245 |
| 209415_at | FZR1 | 0.0040240242963556 | 0.0182118711074509 |
| 21961_at | GCNT3 | 0.00422343917924779 | 0.047898599928003 |
| 217661_x_at | SX5 | 0.00425011261288811 | 0.0137974636054946 |
| 220220_at | LRRC7A2 | 0.00439560705223552 | 0.0083602539543393 |
| 220017_x_at | CYP2C9 | 0.00444735128361806 | 0.0013430392341892 |
| 204326_s_at | MT1X | 0.00519304170826272 | 0.0083477883935453 |
| 202631_at | OSGEP1 | 0.00526809738102296 | 0.00553805125320832 |
| 207245_at | UGT2B17 | 0.0055001172873036 | 0.0002070984407029 |
| 211837_s_at | PTCH1 | 0.0055415215804383 | 0.00671054670922654 |
| 208323_at | ANXA13 | 0.00558331241550818 | 0.00068902483317067 |
| 211612_s_at | IL13RA1 | 0.0056795911679234 | 0.00361028639239732 |
| 206396_at | SLC1A1 | 0.00595393858849176 | 0.00220879456756152 |
| 215536_at | HLA-DQB2 | 0.00630638945811806 | 0.0111442133130687 |
| 213629_at | MT1F | 0.0063543566241856 | 0.006077958592024 |
| 218902_at | NOTCH1 | 0.00657731256094427 | 0.037335712580496 |
| 33304_at | ISG20 | 0.00679460592547953 | 0.0470076157248008 |
| 215741_x_at | AKAP8L | 0.0068010240983979 | 0.0241477714028368 |
| 216671_at | MUC8 | 0.0068805329616361 | 0.0383618636348532 |
| 204487_s_at | KCNQ1 | 0.00695196777896742 | 0.0267287721809296 |
| 207392_x_at | UGT2B15 | 0.0070147429678188 | 0.000042079070386557 |
| 210126_at | PSG9 | 0.00708607742512625 | 0.0125502556385473 |
| 206461_x_at | MTP1H | 0.0070990648227567 | 0.00664244458624349 |
| 216025_x_at | CYP2C9 | 0.0081014424105605 | 0.0027835975836334 |

**Gene biomarker discovery for colon cancer by *in silico* approach**

Anatomy • Volume 15 / Issue 1 / April 2021
## Appendix 2 (Continued)

T-test p-values which expressed statistically significant (p<0.05).

| Probe set | Gene symbol | T-test of ascending vs transverse colon | T-test of ascending vs descending colon |
|-----------|-------------|----------------------------------------|----------------------------------------|
| 209150_s_at | IPO4 | 0.00813333467021513 | 0.0291253365640916 |
| 208478_s_at | BAX | 0.00842647383645954 | 0.0180985713492701 |
| 208581_x_at | MT1X | 0.008449808519362 | 0.0069586322696345 |
| 204745_x_at | MT1G | 0.00878185719233833 | 0.0082522817999293 |
| 205945_at | l6R | 0.009030501809093 | 0.015561832765579 |
| 209938_at | TADA2A | 0.00945544955869718 | 0.0030902612562343 |
| 207484_s_at | EHMT2 | 0.00948494012139873 | 0.021989383206608 |
| 217144_at | UBBP1 | 0.00958596978462974 | 0.0145424696521897 |
| 212349_at | POPUT1 | 0.00973930667339722 | 0.0006175570793332 |
| 208141_s_at | DHH | 0.00983615661677668 | 0.011910985340121 |
| 215481_s_at | PEX5 | 0.0098426572367347 | 0.0080466303700721 |
| 219705_at | QSER1 | 0.00984955869718 | 0.00676393361084865 |
| 213017_at | ABHD3 | 0.00984955869718 | 0.0047659965826437 |
| 222048_at | CRYBB2P1 | 0.00984955869718 | 0.0021989383206608 |
| 210267_at | NIPAL3 | 0.00984955869718 | 0.015561832765579 |
| 214509_at | HIST1H3I | 0.00984955869718 | 0.0006175570793332 |
| 211165_x_at | EPHB2 | 0.0104459036229018 | 0.00676393361084865 |
| 200847_at | SARAF | 0.0104459036229018 | 0.0047659965826437 |
| 216336_x_at | LOC440434 | 0.0104459036229018 | 0.0021989383206608 |
| 215090_x_at | LOC440434 | 0.0104459036229018 | 0.015561832765579 |
| 21269_at | NUDCD3 | 0.0104459036229018 | 0.0006175570793332 |
| 21285_x_at | MT2A | 0.0104459036229018 | 0.0047659965826437 |
| 216661_x_at | CYPC2C9 | 0.0104459036229018 | 0.0021989383206608 |
| 220135_at | SLCTA9 | 0.0104459036229018 | 0.015561832765579 |
| 211217_x_at | KCNQ1 | 0.0104459036229018 | 0.0006175570793332 |
| 213235_x_at | KNOPI1 | 0.0104459036229018 | 0.015561832765579 |
| 210384_at | RPL23A | 0.0104459036229018 | 0.0006175570793332 |
| 203178_at | GATM | 0.0104459036229018 | 0.0047659965826437 |
| 202840_at | TAF15 | 0.0104459036229018 | 0.0021989383206608 |
| 206342_at | KS | 0.0104459036229018 | 0.015561832765579 |
| 214421_x_at | CYPC2C9 | 0.0104459036229018 | 0.0047659965826437 |
| 206340_at | NR1H4 | 0.0104459036229018 | 0.0021989383206608 |
| 213880_at | LGR5 | 0.0104459036229018 | 0.015561832765579 |
| 218952_at | PCSK1N | 0.0104459036229018 | 0.0047659965826437 |
| 207532_at | CRYGD | 0.0104459036229018 | 0.0021989383206608 |
| 203525_s_at | APC | 0.0104459036229018 | 0.015561832765579 |
| 203011_at | IMPA1 | 0.0104459036229018 | 0.0047659965826437 |
| 212859_x_at | MT1E | 0.0104459036229018 | 0.0021989383206608 |
| 217540_at | NXPE3 | 0.0104459036229018 | 0.015561832765579 |
| 21256_at | GRM8 | 0.0104459036229018 | 0.0047659965826437 |
| 217476_at | NR1D1 | 0.0104459036229018 | 0.0021989383206608 |
| 208720_s_at | RBM39 | 0.0104459036229018 | 0.0047659965826437 |
| 211456_x_at | MT1H1L1 | 0.0104459036229018 | 0.0047659965826437 |
| 217696_at | FUT7 | 0.0104459036229018 | 0.0047659965826437 |
| 221270_s_at | QTRT1 | 0.0104459036229018 | 0.0047659965826437 |
| 212211_x_at | KS | 0.0104459036229018 | 0.0047659965826437 |
| 216842_x_at | AC007967.3 | 0.0104459036229018 | 0.0047659965826437 |
| 216218_s_at | PLCL2 | 0.0104459036229018 | 0.0047659965826437 |
| 200051_at | SART1 | 0.0104459036229018 | 0.0047659965826437 |
| 207545_s_at | LOC101928143 | 0.0104459036229018 | 0.0047659965826437 |
Appendix 2 (Continued)

Gene biomarker discovery for colon cancer by *in silico* approach

T-test p-values which expressed statistically significant (p<0.05).

| Probe set | Gene symbol | T-test of ascending vs transverse colon | T-test of ascending vs descending colon |
|-----------|-------------|----------------------------------------|----------------------------------------|
| 205208_at | ALDH1L1     | 0.018366792138689                     | 0.0041525032330587                     |
| 205221_at | HGD         | 0.018404909852041                     | 0.0027128594561777                     |
| 221820_s_at | KAT8       | 0.0186851086363866                   | 0.028753409045594                      |
| 203655_at | XRCC1       | 0.0195552969036288                    | 0.0400589045965037                     |
| 212750_at | PPP1R16B   | 0.01956461833530                    | 0.0492377107044371                     |
| 221506_s_at | TNPO2     | 0.019618462756418                   | 0.00174457407733694                    |
| 209805_at | PM52        | 0.019825195471822                    | 0.0048631699502379                     |
| 215064_at | SC5D        | 0.0199954612067164                   | 0.04009145857484                      |
| 207849_at | IL2         | 0.0207176895788813                   | 0.027646752952765                     |
| 205906_at | FOXJ1       | 0.020756339586082                    | 0.0179806461888217                     |
| 219825_at | CYP26B1     | 0.0209583654715191                   | 0.00881670973732324                    |
| 214223_at | PTPA3       | 0.021973868580162                    | 0.0479893913605289                     |
| 208126_s_at | CYP2C18    | 0.0219738271609716                   | 0.000177625442352609                   |
| 219931_s_at | KLHL12      | 0.0219738271609716                   | 0.0117095765515496                     |
| 213829_x_at | RTE1       | 0.0219376081160701                   | 0.00440303070999215                    |
| 217520_s_at | NPYA        | 0.0221274789129413                   | 0.0312121049147675                     |
| 215152_at | MYB         | 0.02264495177855                    | 0.00075685454908027                    |
| 211526_s_at | RTE1       | 0.0233096492966428                   | 0.028519092503584                      |
| 213683_at | ACSL6       | 0.0243451736397553                   | 0.0248146569085601                     |
| 208918_s_at | NADK        | 0.0243821347655757                   | 0.0399396874739623                     |
| 211866_at | SAMD14      | 0.0243821347655757                   | 0.0399396874739623                     |
| 205633_s_at | ALAS1       | 0.0248146569085601                   | 0.008214818634889686                    |
| 202695_at | STK17A      | 0.0248146569085601                   | 0.008214818634889686                    |
| 206918_s_at | CPNE1       | 0.0248146569085601                   | 0.008214818634889686                    |
| 220143_x_at | LUC7L       | 0.0248146569085601                   | 0.008214818634889686                    |
| 208078_s_at | SK1         | 0.0248146569085601                   | 0.008214818634889686                    |
| 216255_s_at | GRM8        | 0.0249146761499734                   | 0.0157547742656726                     |
| 212057_at | GSE1        | 0.0249589897515221                   | 0.048469812948672                      |
| 204600_at | EPHB3       | 0.0253505027616503                   | 0.0147809814079668                     |
| 211207_s_at | ACSL6       | 0.0253505027616503                   | 0.0147809814079668                     |
| 200647_x_at | EF3C        | 0.0253505027616503                   | 0.0147809814079668                     |
| 208739_s_at | TMEM88      | 0.0259039361300557                   | 0.019690503548056                      |
| 213588_x_at | RPL14       | 0.0260872509286091                   | 0.011147522187354                      |
| 212486_at | FYN         | 0.0260872509286091                   | 0.011147522187354                      |
| 213052_at | PRKAR2A     | 0.0260872509286091                   | 0.011147522187354                      |
| 202453_s_at | GTF2H1      | 0.026509800665638                    | 0.0346504801267993                     |
| 211082_x_at | MARK2       | 0.0266829025992098                   | 0.0425854689136332                     |
| 216076_at | L3MBTL1     | 0.0272017509712494                   | 0.0039832497474508                     |
| 203692_s_at | EZF3        | 0.02724518843137                    | 0.0156576760326496                     |
| 203060_at | PAPSS2      | 0.02789514467807                    | 0.0042850877172897                     |
| 205316_at | SLC15A2     | 0.02789514467807                    | 0.0042850877172897                     |
| 220544_at | TSK5        | 0.028706492697387                    | 0.005097903960052                      |
| 221309_at | RBM17       | 0.02907254331826                    | 0.026887155853216                      |
| 201418_s_at | SOX4        | 0.029202338514225                   | 0.012618525383067                      |
| 206092_x_at | RTE1        | 0.0293869362376055                   | 0.0011386086201966                     |
| 78330_at | ZNF335      | 0.0294420246489318                   | 0.0171741697532607                     |
| 205272_s_at | PRH1        | 0.029922012935875                   | 0.0369958217690242                     |
| 217702_at | IL27RA      | 0.0299544046093803                   | 0.0401086359899183                     |
| 209589_s_at | EPHB2       | 0.0311842634624996                   | 0.012365422902086                     |
T-test p-values which expressed statistically significant (p<0.05).

| Probe set | Gene symbol | T-test of ascending vs transverse colon | T-test of ascending vs descending colon |
|-----------|-------------|----------------------------------------|----------------------------------------|
| 211955_at | IPO5        | 0.0315578723564607                     | 0.0176122286833458                     |
| 211682_x_at | UGT2B28    | 0.031835468805484                     | 0.0064869431682457                     |
| 204087_s_at | SLCSA6      | 0.032167092191126                     | 0.0122490683721897                     |
| 203526_s_at | APC         | 0.02952672600519                      | 0.0091742281823984                     |
| 200679_x_at | HMG81      | 0.03063253557303                      | 0.0092290135374263                     |
| 213435_at | SATB2       | 0.032268565076552                     | 0.005791988193093                     |
| 200680_x_at | HMG81      | 0.03259346934992                      | 0.0327594672367312                     |
| 204016_at | LARS2       | 0.03386627531322                      | 0.007640142179958                      |
| 201741_x_at | SRSF1      | 0.033976114866492                     | 0.0072246332868078                     |
| 219017_at | ETV1        | 0.034975156635643                     | 0.0481914664756225                     |
| 221803_s_at | NRB2        | 0.034042693873365                     | 0.0008475659728556                     |
| 207470_at | BC113958    | 0.034079924012714                     | 0.0176655570248064                     |
| 200721_s_at | ACTR1A      | 0.0341968526622086                     | 0.0189529343666907                     |
| 209130_at | SNAP23      | 0.0343722782076268                     | 0.012460340970823                     |
| 219471_at | KIAA0226    | 0.034664305671111                     | 0.00420770437161399                    |
| 208209_s_at | C4BPB       | 0.035103409245856                     | 0.015300645553425                     |
| 204109_at | NYFA        | 0.035467090518342                     | 0.0188479934829926                     |
| 216032_s_at | ERGIC3      | 0.036125187029235                     | 0.01222202385168                      |
| 205556_s_at | VAMP2       | 0.036315069181276                     | 0.00326733016642112                    |
| 205459_s_at | NAP52       | 0.036356820458499                     | 0.027501096900155                     |
| 201892_s_at | IMPDH2      | 0.03684193144478                     | 0.0164254058840621                    |
| 217809_at | B2ZV2       | 0.0370411672528593                     | 0.0111530894359189                     |
| 219316_s_at | FLVCR2      | 0.037152693664439                     | 0.0182386046949649                    |
| 215930_s_at | CTAGES5     | 0.037273894038448                     | 0.009851487834859                     |
| 201716_at | SNX1        | 0.0374101133631355                    | 0.0390882393411317                     |
| 205460_at | NAP52       | 0.0377153078107637                    | 0.0043895592785321                    |
| 212198_s_at | TM9SF4      | 0.037790018920065                     | 0.031513580433592                     |
| 201791_s_at | DHCR7       | 0.037839673838822                     | 0.0105568685301129                    |
| 220354_at | MCF2L-AS1   | 0.0382621296936618                    | 0.019412214545149                     |
| 212322_at | SGPL1       | 0.0382890186278431                    | 0.026994879228522                     |
| 215852_x_at | SOGAT       | 0.038432906234485                     | 0.038601969455771                     |
| 219447_at | SLC35C2     | 0.0390051865560509                    | 0.00838972207178568                    |
| 208506_at | HIST1H3F    | 0.039378750515056                     | 0.0032765992657932                     |
| 204183_s_at | ARKB2       | 0.039397370972733                     | 0.00028760060211399                    |
| 219764_at | FZD10       | 0.0394147921859214                    | 0.0334876991598887                     |
| 205639_at | AOAH        | 0.0395018637342232                    | 0.0031126533171492                     |
| 206407_s_at | CCL13       | 0.0397475278576177                    | 0.043349304220237                     |
| 213828_x_at | H3F3A       | 0.040145315327654                     | 0.0353569437251339                     |
| 205141_at | ANKEF1      | 0.040329905997115                     | 0.015937278259481                     |
| 206411_at | MOCS3       | 0.0404920863485373                    | 0.0182634063304353                    |
| 201795_at | LBR         | 0.040652479065725                     | 0.0010026293123705                    |
| 204838_s_at | M1H3        | 0.040972516233634                     | 0.0104570300276411                    |
| 220936_s_at | H2AFJ       | 0.0409742871898888                    | 0.00740207154862594                    |
| 209134_s_at | RPS6        | 0.0411455819895845                    | 0.014310812135305                     |
| 220144_s_at | ANKEF1      | 0.0412498731136823                    | 0.0096902210987666                     |
| 213053_at | HUSS5       | 0.0412932176220598                    | 0.0150984897605612                     |
| 41577_at | PPP1R1B     | 0.0414980406250487                    | 0.0484377148658173                     |
| 220211_at | FLJ13224    | 0.041685014369416                     | 0.043598659600468                     |
Appendix 2 (Continued)

T-test p-values which expressed statistically significant (p<0.05).

| Probe set | Gene symbol | T-test of ascending vs transverse colon | T-test of ascending vs descending colon |
|-----------|-------------|----------------------------------------|----------------------------------------|
| 204438_at | MRC1        | 0.0418464706452934                    | 0.00992848027255271                   |
| 40837_at  | TLE2        | 0.0423156766188777                    | 0.0062633011427297                   |
| 215103_at | CYP2C18     | 0.042370724064298                     | 0.00559948063703009                 |
| 218579_s_at | DHX35    | 0.0427011211748617                    | 0.0436691935495297                   |
| 222251_s_at | GMEB2    | 0.0430124806537605                    | 0.00725566828019381                 |
| 210357_s_at | SMOX      | 0.0433172855060324                    | 0.0338207181600041                 |
| 205129_at | NPM3       | 0.0436309133780782                    | 0.0384625268226612                  |
| 205240_at | GPM52      | 0.0439793027972249                    | 0.00063156133135152                 |
| 202576_s_at | DDX19A   | 0.0440814208639885                    | 0.018256407713014                   |
| 206650_at | IQCC       | 0.0449944650713525                    | 0.0121751394207934                  |
| 214107_x_at | LOC440434 | 0.0450720367479637                    | 0.0265633412274562                  |
| 204613_at | PLCG2      | 0.0454915719075093                    | 0.0229035737853801                  |
| 216508_x_at | HMGB1P4   | 0.0461980573409067                    | 0.0143328712291447                  |
| 205865_at | ARID3A     | 0.0463046577532617                    | 0.0029233588010168                  |
| 203909_at | SLC9A6     | 0.046675287281775                     | 0.0473501198587895                  |
| 221741_s_at | YTHDF1   | 0.0473235099369472                    | 0.0161402553890302                  |
| 211603_s_at | ETV4      | 0.0473890117318017                    | 0.0057997801630521                  |
| 219653_at | LSM14B     | 0.0476370621810747                    | 0.000580206292203339                 |
| 206170_at | ADRB2      | 0.0476421322380342                    | 0.0154418296902453                  |
| 221922_at | GPM52      | 0.0476880997134143                    | 0.0065707163768071                  |
| 210393_at | LGR5       | 0.0490924051065539                    | 0.0138545969031669                  |
| 213975_s_at | LYZ       | 0.0496410825350101                    | 0.021634858636666                   |
| 209388_at | EPHB2      | 0.049829082972293                     | 0.0316158601621646                  |
| 205362_s_at | PFDN4     | 0.0498825908777403                    | 0.000926938401448077                 |

Appendix 3

Result of GSEA showed that a total of 11 gene sets were enriched in the ascending tumor type, while 6 gene sets were not enriched in the same group.

| Enriched in ascending colon                  | Diminished in ascending colon                  |
|---------------------------------------------|-----------------------------------------------|
| HP_POSTAXIAL_FOOT_POLYDACTYLY            | GO_CARBONHYDRATE_BINDING                      |
| GO_REGULATION_OF_TELOMERE_CAPPING        | GO_REGULATION_OF_EXOCYTOSIS                   |
| HP_ABNORMALITY_OF_THE_5TH_TOE             | GO_NEGATIVE_REGULATION_OF_GLUCOSE_TRANSMEMBRANE_TRANSPORT |
| GO_N_METHYLTRANSFERASE_ACTIVITY          | GO_DEAMINASE_ACTIVITY                         |
| GO_NEGATIVE_REGULATION_OF_GENE_EXPRESSION_EPIGENETIC | GO_POSITIVE_REGULATION_OF_BLOOD_CIRCULATION |
| GO_REGULATION_OF_GENE_EXPRESSION_EPIGENETIC | GO_TRANSITION_METAL_ION_HOMEOSTASIS          |
| GO_TELOMERE_CAPPING                      |                                               |
| GO_S_ADENOSYLMETHIONINE_DEPENDENT_METHYLTRANSFERASE_ACTIVITY |                                               |
| GO_SPliceosomal_SNRNP.Assembly            |                                               |
| GO_SNRNA_PROCESSING                      |                                               |
| GO_PROTEIN_DNA_COMPLEX                    |                                               |