American Journal of Sciences and Engineering Research

E-ISSN -2348 – 703X, Volume 7, Issue 3,

The Significance of Genetic Analysis in Colorectal Cancer and the Bioinformatics Analysis of Gene Data with Colorectal Cancer Overview

Zeynep Kucukakcali[1], Ipek Balikci Cicek[2]

[1,2] *Department of Biostatistics and Medical Informatics, Inonu University Faculty of Medicine, 44280, Malatya, Turkey*

Abstract:

Aim: *It is important to understand the prognosis and develop treatment strategies for colorectal cancer (CRC), which has become a public health problem with an increasing incidence worldwide. Genomic studies should be accelerated for this disease which has a genetic basis. For this purpose, in this study, an open access dataset of CRC tumor tissues and CRC normal tissues was analyzed by various bioinformatics analyses and it was aimed to identify genes that are differentially regulated in CRC tumor tissues.*

Material and methods: *CRC patient dataset consisting of lncRNA expression data was used in the study. With this data set, gene expression analysis was performed with the help of the functions in the limma package. The distribution of each tissue in the data set is given by the distribution graph and expression density graph. The association of tissue types is given by UMAP graph. Genes showing up and down regulation are given with the rise volcano graph.*

Results: *According to the UMAP garage, the tissues in the dataset were completely segregated into CRC tumor tissues and CRC normal tissues. The analysis showed that many genes differed in both groups under log2FC>1 p<0.05 and log2FC<-1 and p<0.05 conditions. In the tables, 10 of the genes that were differentially regulated (up and down) in CRC tumor tissues compared to CRC normal tissues are given.*

Conclusion: *Genetic factors, environmental influences and the complex interactions between the two have a major impact on the development of colorectal cancer. Understanding the genetic basis of colorectal cancer is critical for both prevention and treatment of the disease. Genetic studies enable the development of targeted prevention strategies and personalized treatment approaches. In the future, with the further advancement of genetic research, more innovative and effective methods for the prevention and treatment of colorectal cancer are expected to be developed.*

Keywords: *Colorectal cancer, Gen expression, up and down regulation*

I. Introduction

The primary cause of cancers is the proliferation of epithelial cells that line the body's surfaces. Colon cancers, often known as CRC, are a kind of carcinoma that originates from epithelial cells in the mucosa of the colon and rectum (1). Colorectal cancers are prevalent neoplastic conditions and significant contributors to mortality. This type of cancer is a major contributor to global mortality rates (2). Colorectal cancer ranks as the third most prevalent ailment globally among both males and females, with an annual incidence exceeding one million new cases (3). The regions with the highest prevalence rates are Asia, America, Europe, Australia, and

New Zealand, particularly in emerging countries where a predominantly Western diet and lifestyle are prevalent. Regions with the lowest prevalence are typically found in underdeveloped and densely populated countries like India (4). The insidance of these tumors escalates markedly with age and can generally be regarded as a condition that primarily affects older individuals (5). Based on data from the United States, the five-year survival rate for colorectal cancer is 65%. From 2000 to 2016, the approval of cancer drugs led to a 34% decrease in mortality attributable to colorectal cancer (6).

Genomic technology, a field that utilizes information technologies to process and store its outputs, has been developed through advancements in automation and bioinformatics. By implementing appropriate genomic technology, research can be conducted in various medical disciplines such as Oncology, Pharmacology, Immunology, Biochemistry, Microbiology, and others (7). By conducting comparison investigations, it offers research possibilities in areas such as cancerization and prognosis prediction, medication response prediction and personalized drug creation, understanding the nature of immune response, and even predicting transplantation outcomes (8). Next-generation sequencing (NGS) has facilitated recent advancements in the investigation of genomic alterations in cancer research and clinical practice (9). Colorectal cancer arises due to the accumulation of genetic (gene mutations, gene amplification, etc.) and epigenetic (abnormal DNA methylation, chromatin modifications, etc.) changes that convert colonic epithelial cells into colon adenocarcinoma cells. Comprehending the factors and functions of genomic and epigenomic instability in the development of colon tumors has the capacity to result in improved methods for preventing and treating colorectal cancer (10). Consequently, numerous comprehensive investigations of colorectal cancer at the genomic level have been conducted, with a focus on analyzing gene expression data.

The aim of this study is to identify potential genes that may be associated with CRC tumor tissues by bioinformatics methods using open-access gene expression data obtained from CRC tumor tissues and CRC normal tissues. For this purpose, an open-access dataset of CRC patients was analyzed by various bioinformatics analyses and genes that differ in CRC tumor tissues compared to CRC normal tissue (showing up-down regulation) were identified and summarized in various graphs and table.

II. Material and Methods

Dataset

The dataset used in the study is an open access dataset created to investigate important lncRNAs that play a role in tumor formation and progression in colorectal cancer. The dataset used in the study was obtained from the National Center for Biotechnology Information (NCBI). The current code of the dataset in NCBI is GSE102340 and the dataset includes 6 CRC tumor tissues and 6 CRC normal tissues.

Long non-coding RNA lar (lncRNA)

lncRNAs are RNA molecules that exceed 200 nucleotides in length and lack open reading frames. They interact with DNA, mRNA, proteins, and miRNA structures, and regulate gene expression at different levels of function, including epigenetic, transcriptional, post-transcriptional, and translational, through diverse mechanisms (11). According to reports, lncRNAs engage with DNA, RNA, and other transcriptional molecules, playing a role in several biological processes and contributing to the development of diseases. They play a crucial role in gene expression, gene silencing, heat shock response, imprinting, embryogenesis, and various other biological processes. Furthermore, alterations in long non-coding RNAs (lncRNAs) have been linked to a multitude of diseases, such as viral infections, cancer, and neurodegenerative disorders. Furthermore, it has been documented that any disruption in the control of long non-coding RNAs (lncRNAs) has a significant impact on various essential physiological functions, including the initiation of blood vessel formation, cell growth, resistance to programmed cell death, and avoidance of mechanisms that restrict tumor development. Nevertheless, there have been limited research endeavors about lncRNAs (12). The recognition that problems in gene regulation are linked to numerous diseases has underscored the significance of lncRNA-related research. Long non-coding RNAs (lncRNAs) have a regulatory function in gene regulation at multiple levels. Any disturbance in this regulatory function contributes to the occurrence, advancement, and progression of a diverse range of disorders. Abnormalities in the sequencing and three-dimensional structures of lncRNAs, as

well as abnormalities in DNA-protein interaction and expression levels, have been found to be closely associated with the development of most diseases. This is because lncRNAs play a functional role in various physiological and pathological processes. Due of these factors, long non-coding RNAs (lncRNAs) have been suggested as potential targets for therapy and as indicators for numerous significant disorders (13).

Bioinformatics and gene expression analysis

Bioinformatics encompasses the systematic collection, storage, organization, preservation, analysis, and presentation of knowledge obtained from the application of theoretical and practical concepts in disciplines such as biology, medicine, behavioral sciences, and health sciences. Furthermore, the main objective of this undertaking is to investigate and enhance computational tools and methodologies, in order to broaden the use and manipulation of data obtained from research efforts or the use of established protocols. Obtained by rigorous intellectual research or by following established procedures. Bioinformatic analyses are performed by selecting an appropriate database and utilizing a tool that enables the execution of bioinformatic analysis, in accordance with the particular biological question, molecule, or structure being investigated. The data collected and insights obtained from the studies are combined, and the ensuing assessments are carefully evaluated in relation to the existing literature (14).

Changes in the physiological condition of an organism or cell always result in commensurate adjustments in the gene expression pattern. Therefore, the assessment of gene expression is highly significant in all areas of biological study. The DNA microarray technology, currently in the developmental stage, is used to study gene expression. This is accomplished through the process of hybridization, in which mRNA molecules are attached to a highly concentrated array of immobilized target sequences. Each of these target sequences corresponds to a unique gene. The influence of chemical compounds on the control of gene expression can offer valuable information on both functional and toxicological characteristics. Conducting investigations on clinical samples, including both healthy individuals and those with illnesses, has the potential to reveal previously unknown biomarkers (15).

Bioinformatics analysis phase

This study conducted gene expression studies using transcriptomic data derived from both CRC tumor tissues and CRC normal tissues. The investigation employed the limma package, a software tool accessible in the R programming language that facilitates expression analysis (16). Limma, or Linear Models for Microarray Analysis, is a software suite specifically created to evaluate gene expression microarray data. The main goal is to utilize linear models to examine certain experiments and detect differential expression. The packet's functionalities can be utilized in many gene expression techniques, such as microarrays, RNA-seq, and quantitative PCR. The Limma software utilizes Empirical Bayes techniques to produce consistent conclusions, even in settings with a small number of sequences. The bioinformatic study produced Lof2FC, a metric that measures the fold change in gene expression disparities. This metric arranges the genes in a descending order based on their level of importance. Genes that exhibit increased expression are identified by applying a threshold of log2 fold change (log2FC) greater than 1, while genes that show decreased expression are identified by using a threshold of log2FC less than -1.

The data distribution in the study was shown using box plots and density plots. The graphs illustrate samples with comparable characteristics, shown by the use of consistent colors. The research choose to employ the Uniform Manifold Approximation and Projection (UMAP) graph to visually depict the interconnections among the samples being studied. The volcano plot was chosen as the preferred method for visualizing genes that exhibit differential expression, including both upregulation and downregulation. The volcano plot depicts the logarithmic relationship between significance and fold-change, with the y-axis representing significance and the x-axis representing fold-change on a logarithmic base of 2. This visual depiction enables the quick identification of genes that demonstrate differential expression. The graph depicts gene expression levels, where the color red indicates genes that are up-regulated, blue indicates genes that are down-regulated, and black indicates genes that show no significant difference in expression. Furthermore, alongside the Volcano plot, we employed the Mean Difference (MD) plot, an additional graphical representation that clearly depicts

genes exhibiting differential expression between groups. The MD plot visually displays the log2 fold change of differentially expressed genes in relation to the average log2 expression levels. In the volcano plot, the coloring distinguishes between genes that are up-regulated and those that are down-regulated. The coloration in the volcanic graph exhibits a resemblance to that of this graph.

III. Results

Figure 1 and Figure 2 display distribution plots of 6 tumor tissues and 6 normal tissues, respectively, which were utilized in the investigation. In the graphs, C denotes tumor tissues and N denotes normal tissues. It is used to display the distribution of values of the samples selected in both graphs. In the graphs the samples are colored according to the groups. These plots are used to check data normalization before differential expression analysis and the expression density plot is often preferred to complement the boxplot.

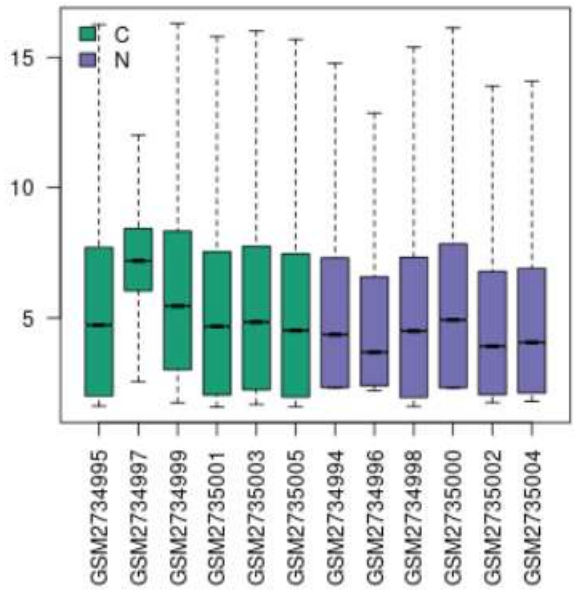

Figure 1: Distribution plot of the samples

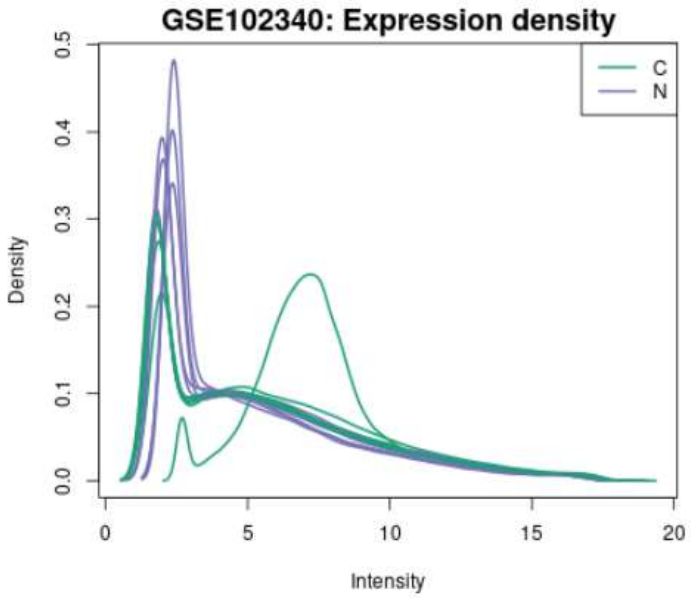

Figure 2: The expression density graph of the samples.

Figure 3 displays the UMAP graph, which clearly depicts the connections between the samples. The graph demonstrates that samples with comparable characteristics are observed to be clustered together. The graph depicts tumor tissue samples denoted by green dots and non-tumor tissue denoted by purple dots. In the graphs, C denotes tumor tissues and N denotes normal tissues.

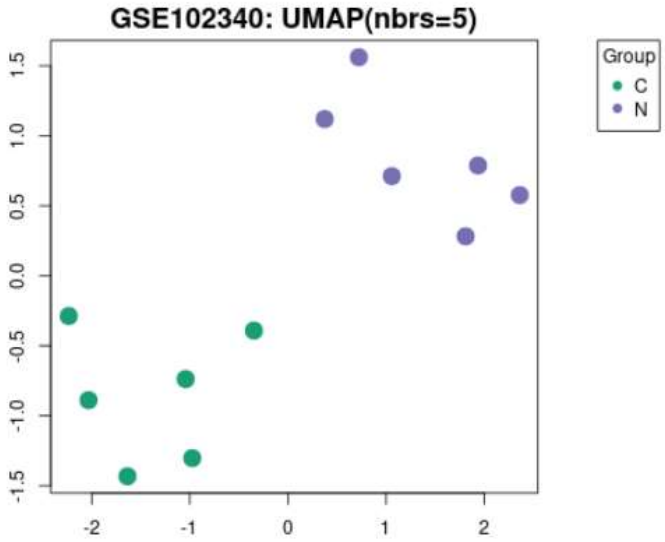

Figure 3: UMAP plot of the samples (Green Dots: Tumor tissues, Purple dots: non-tumor tissues).

Based on the results obtained by analyzing the gene expression of 62738 transcripts in the dataset, the results of the top 10 genes showing up- and down-regulation between the two groups are given in table 1 and table 2. When determining the regulation of gene expression, |log2FC| > 1.0, p 0.05 conditions were taken into account.

Table 1: Transcripts found to be up-regulated in tumorous tissue samples relative to non-tumorous tissues.

ID	adj,P,Val	P,Value	t	B	logFC	NAME
53568	3,78E-07	6,37E-12	22,01144	15,13846	6,262818	ASHG19A3A040484 Spanxd
58261	3,78E-07	1,31E-11	20,8303	14,72111	6,190817	ASHG19A3A040482 Spanxa2
37998	3,78E-07	1,81E-11	20,32269	14,52825	6,187018	ASHG19A3A041630 Spanxa1
46240	8,23E-04	6,03E-07	8,810679	6,41296	6,00268	ASHG19A3A045626 Mmp7
57119	1,51E-03	3,35E-06	7,571738	4,824871	5,971641	ASHG19A3A017482 Cst1
13257	5,80E-03	6,06E-05	5,737129	2,060316	5,83746	ASHG19A3A031308 Ly6g6d
49370	5,41E-04	2,50E-07	9,499755	7,209075	5,837019	ASHG19A3A044994 Saa2
24825	3,08E-03	1,54E-05	6,56839	3,376801	5,683801	ASHG19A3A045631 Mmp10
29354	1,74E-04	2,20E-08	11,6232	9,325916	5,663577	ASHG19A3A022082 Cldn1
53517	1,74E-04	2,22E-08	11,61659	9,320028	5,618393	ASHG19A3A003945 Cdh3

Table 2: Transcripts found to be down-regulated in tumorous tissue samples relative to non-tumorous tissues.

ID	adj,P,Val	P,Value	t	B	logFC	NAME
36017	6,61E-03	8,30E-05	-5,55338	1,754989	-6,93293	ASHG19A3A022063 DAGLB
43274	1,01E-02	2,16E-04	-5,0121	0,826666	-6,72144	ASHG19A3A035919 Ca1
23105	4,71E-02	5,53E-03	-3,30001	-2,32079	-5,8819	ASHG19A3A018739 Cldn8
35056	3,04E-03	1,47E-05	-6,5988	3,422928	-5,72677	ASHG19A3A007914 Myoc
21945	1,55E-03	3,66E-06	-7,51148	4,742137	-5,495	ASHG19A3A017697 Rims4
28667	5,97E-03	6,42E-05	-5,70278	2,003628	-5,40672	ASHG19A3A014560 Gcg
11714	1,89E-02	9,05E-04	-4,23782	-0,56796	-5,34815	ASHG19A3A010348 Ccbe1
59066	8,55E-03	1,40E-04	-5,25403	1,246792	-5,19165	ASHG19A3A034154 Adcyap1r1

32639	5,25E-03	4,79E-05	-5,87585	2,28742	-5,18248	ASHG19A3A019787 Sox10
8570	1,37E-02	4,37E-04	-4,62609	0,140127	-5,13873	ASHG19A3A038430 Cdkn2bas1

Figure 4: Depicts the volcano plot, which visually represents the genes that display differential expression among the several groups.

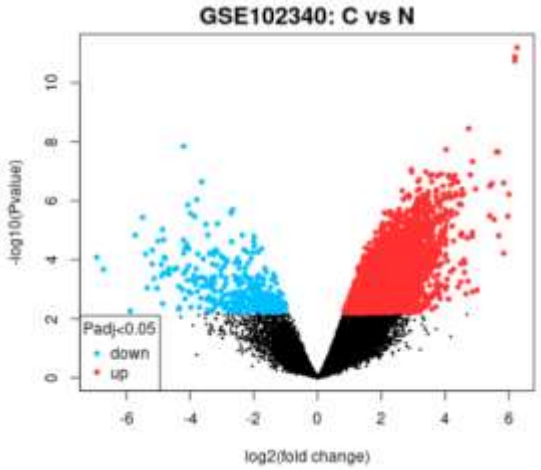

Figure 4: Volcano plot of transcripts in tumor and non-tumor tissues. (Red dots represent transcripts that increased, blue dots represent transcripts that dropped, and black dots represent transcripts whose expression level remained unchanged.)

Figure 5 shows the Mean Difference (MD) plot, another graph that visually represents genes showing differential expression between the various groups. The MD plot displays the log2 fold change in relation to the mean log2 expression levels, providing a visual representation of genes that are expressed differentially. Similarly to the volcano plot, you have the ability to hover over data points in order to view specific gene annotations. The genes that have been emphasized show a substantial difference in their expression levels. The color red indicates up-regulation, while blue indicates down-regulation. This distinction is made using a default P-value cutoff of 0.05.

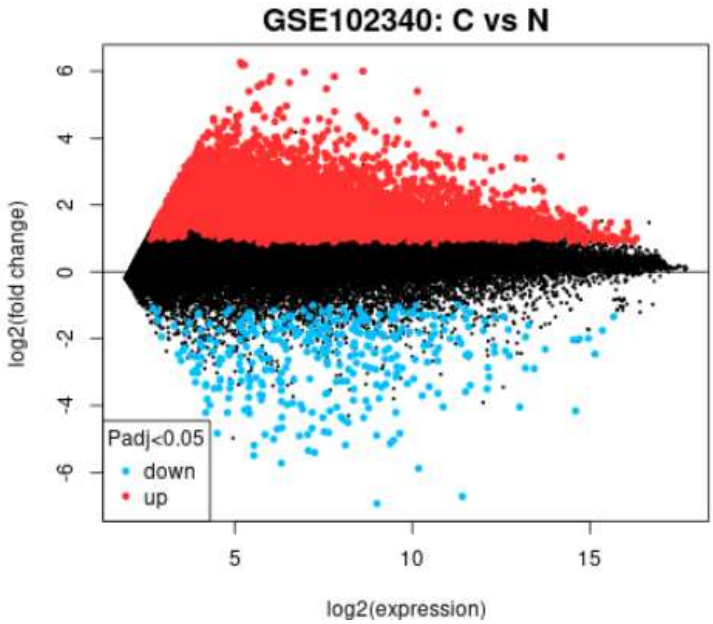

Figure 5: MD plot of transcripts in tumor and non-tumor tissues (Red dots represent transcripts that increased, blue dots represent transcripts that dropped, and black dots represent transcripts whose expression level remained unchanged)

IV. Discussion

Over the past few years, the worldwide occurrence of CRC has surged significantly. By the year 2020, it is projected that there would be approximately 1.93 million new cases of colorectal cancer (CRC) identified and 0.94 million deaths caused by CRC worldwide. These numbers represent 10% of the total global cancer cases (19.29 million new cases) and 9.4% of all cancer-related fatalities (9.96 million deaths). In 2020, colorectal cancer (CRC) is projected to be the third leading cause of cancer-related deaths globally for both men and women. It is expected to result in approximately 515,637 deaths among males and 419,536 deaths among females (17). Considerable attempts and significant progress have been made to enhance our understanding of the pathophysiology of CRC. Endoscopic resection, surgical local excision, targeted therapy, radiation therapy, ablative treatments, chemotherapy, immunotherapy, and genomic investigations have significantly improved the overall survival of the disease (17, 18). The rising prevalence of colorectal cancer (CRC) presents a significant and expanding global public health concern (17). Hence, it is imperative to devise therapeutic interventions by comprehensively elucidating the underlying physiological mechanisms of the disease. The extensive use of next-generation sequencing (NGS) has resulted in a rather comprehensive comprehension of the genomes of colorectal cancer. Nevertheless, the integration of molecular indicators into routine clinical treatment for the condition has been sluggish. Currently, there is no approved targeted therapy for colorectal cancer (CRC) that is based on the use of a positive predictive marker (18). CRC is attributed to successive genomic and epigenetic alterations, and numerous genomic investigations are underway to discover biomarkers that may have therapeutic implications for the disease. However, further research is required to identify other biomarkers.

Due to these gaps, and based on the belief that this gap can be closed with the development of genomic technologies, in the present study, an open access CRC dataset was analyzed by bioinformatics analysis and genes showing differential regulation in CRC tumor tissues compared to CRC normal tissues were identified. These genes were also visualized with volcano plot and MD plot. In addition, the distributions of the samples in the data set were examined and given with expression density graph and distrubution graph. The association of the samples in the data was visualized with the UMAP graph and it was observed that both groups differed.

When the results of bioinformatic analyses were examined, it was determined that a lot of genes showed different regulation (up or down) in CRC tumor tissues compared to CRC normal tissues. Spanxd gene showed 76.63 fold up-regulation in CRC tumor tissues compared to CRC normal tissues. Likewise, Spanxa2, Spanxa1, Mmp7, Cst1, Ly6g6d, Saa2, Mmp10, Cldn1,and Cdh3 genes had up-regulated gene expression of 73.00, 72.50, 64.00, 62.68, 56.88, 56.88, 51.26, 50.56, and 48.85 fold, respectively. DAGLB gene showed 121.93 fold down-regulation in CRC tumor tissues compared to CRC normal tissues. Likewise, Ca1, Cldn8, Myoc, Rims4, Gcg, Ccbe1, Adcyap1r1, Sox10, and Cdkn2bas1 genes had down-regulated gene expression of 105.41, 58.89, 52.70, 44.94, 42.22, 40.50, 36.50, 36.25, and 35.01 fold, respectively.

Further analysis and studies with the identified genes may determine that these genes may be biomarkers that may play a role in effective treatment for CRC. Drug therapies can be developed and applied with these biomarkers. Accurate and effective use of genetic biomarkers can both improve patient care and increase efficiency in healthcare systems. In the future, with the advancement of genetic research, more genes are expected to be used as biomarkers, which will lead to the further spread of personalized approaches in the medical world.

Genetic research in colorectal cancer has greatly enhanced our comprehension of the illness. Scientific research has demonstrated that hereditary variables are of utmost importance in determining the risk and advancement of colorectal cancer (19). The Genetic Epidemiology of Colorectal Cancer Consortium (GECCO) has played a crucial role in assessing genetic and environmental factors that contribute to the risk of developing colorectal cancer (20). Colorectal cancer has been linked to common genetic markers through genome-wide association studies, highlighting the importance of genetic susceptibility in the development of the illness (21). Furthermore, studies have investigated the impact of typical hereditary genetic variation on the survival of colorectal cancer, offering valuable knowledge on the genetic factors that affect patient

outcomes (22). Genetic models suggest that mutations in particular genes, such as APC, Kirsten-ras, and p53, play a role in the progression of colorectal cancer (23). Moreover, there is a correlation between the occurrence and progression of colorectal cancer and genetic vulnerability, which includes changes and polymorphisms in genes such as MIR17HG and PPARGC1A (24, 25). Research has also examined the influence of obtaining predictive genetic information on actions that could help avoid colorectal cancer, emphasizing the usefulness of genetic testing in tailoring cancer preventive strategies (26).

In conclusion, CRC is one of the most common cancers worldwide and genetic studies play an important role in understanding this disease. Genetic factors, environmental influences and the complex interactions between the two have a major impact on the development of colorectal cancer. Understanding the genetic basis of colorectal cancer is critical for both prevention and treatment of the disease. In this context, the knowledge provided by genetic studies allows the development of targeted prevention strategies and personalized treatment approaches. Genetic studies help to develop more targeted strategies for colorectal cancer prevention. Genetic testing and screening programs for high-risk individuals enable early detection and preventive measures.

Finally, understanding the genetic basis of colorectal cancer is critical for determining disease risk, developing preventive strategies and personalizing treatment approaches. Genetic studies provide in-depth knowledge about the etiology of colorectal cancer, which enables the delivery of more effective and individualized healthcare. In the future, with the further advancement of genetic research, more innovative and effective methods for the prevention and treatment of colorectal cancer are expected to be developed.

V. References

1. Günther J, Seyfert H-M, editors. The first line of defence: insights into mechanisms and relevance of phagocytosis in epithelial cells. Seminars in immunopathology; 2018: Springer.
2. Cao W, Chen H-D, Yu Y-W, Li N, Chen W-Q. Changing profiles of cancer burden worldwide and in China: a secondary analysis of the global cancer statistics 2020. Chinese Medical Journal. 2021;134(07):783-91.
3. Mattiuzzi C, Lippi G. Current cancer epidemiology. Journal of epidemiology and global health. 2019;9(4):217.
4. Sharma R. An examination of colorectal cancer burden by socioeconomic status: evidence from GLOBOCAN 2018. EPMA Journal. 2020;11(1):95-117.
5. Abualkhair WH, Zhou M, Ahnen D, Yu Q, Wu X-C, Karlitz JJ. Trends in incidence of early-onset colorectal cancer in the United States among those approaching screening age. JAMA Network Open. 2020;3(1):e1920407-e.
6. MacEwan JP, Dennen S, Kee R, Ali F, Shafrin J, Batt K. Changes in mortality associated with cancer drug approvals in the United States from 2000 to 2016. Journal of Medical Economics. 2020;23(12):1558-69.
7. Del Boccio P, Urbani A. Homo sapiens proteomics: clinical perspectives. Annali dell'Istituto Superiore di Sanita. 2005;41(4):479-82.
8. Martin DB, Nelson PS. From genomics to proteomics: techniques and applications in cancer research. Trends in Cell Biology. 2001;11(11):S60-S5.
9. Gagan J, Van Allen E. Next-generation sequencing to guide cancer therapy. Genome Med 7: 80. Link: https://bit ly/35WLrGw. 2015.
10. Grady WM, Carethers JM. Genomic and epigenetic instability in colorectal cancer pathogenesis. Gastroenterology. 2008;135(4):1079-99.
11. Zhang X, Wang W, Zhu W, Dong J, Cheng Y, Yin Z, et al. Mechanisms and functions of long non-coding RNAs at multiple regulatory levels. International journal of molecular sciences. 2019;20(22):5573.
12. Dhanoa JK, Sethi RS, Verma R, Arora JS, Mukhopadhyay CS. Long non-coding RNA: its evolutionary relics and biological implications in mammals: a review. Journal of animal science and technology. 2018;60:25.
13. Zhang X, Hong R, Chen W, Xu M, Wang L. The role of long noncoding RNA in major human disease. Bioorganic chemistry. 2019;92:103214.

14. Akalın PK. Introduction to bioinformatics. Molecular nutrition & food research. 2006;50(7):610-9.
15. van Hal NL, Vorst O, van Houwelingen AM, Kok EJ, Peijnenburg A, Aharoni A, et al. The application of DNA microarrays in gene expression analysis. Journal of biotechnology. 2000;78(3):271-80.
16. Smyth GK. Limma: linear models for microarray data. Bioinformatics and computational biology solutions using R and Bioconductor: Springer; 2005. p. 397-420.
17. Xi Y, Xu P. Global colorectal cancer burden in 2020 and projections to 2040. Translational Oncology. 2021;14(10):101174.
18. Mondaca S, Yaeger R. Colorectal cancer genomics and designing rational trials. Annals of Translational Medicine. 2018;6(9).
19. Dimou N, Kim A, Flanagan O, Murphy N, Díez-Obrero V, Shcherbina A, et al. Probing the Diabetes and Colorectal Cancer Relationship Using Gene – Environment Interaction Analyses. 2022.
20. Lindström S, Finucane H, Bulik-Sullivan B, Schumacher F, Amos CI, Hung RJ, et al. Quantifying the Genetic Correlation Between Multiple Cancer Types. Cancer Epidemiology Biomarkers & Prevention. 2017;26(9):1427-35.
21. Cho YA, Lee J, Oh JH, Chang HJ, Sohn DK, Shin A, et al. Genetic Risk Score, Combined Lifestyle Factors and Risk of Colorectal Cancer. Cancer Research and Treatment. 2019;51(3):1033-40.
22. He Y, Timofeeva M, Li X, Din FVN, Blackmur JP, Vaughan-Shaw PG, et al. A Comprehensive Study of the Effect on Colorectal Cancer Survival of Common Germline Genetic Variation Previously Linked With Cancer Prognosis. Cancer Epidemiology Biomarkers & Prevention. 2019;28(11):1944-6.
23. Smith G, Carey FA, Beattie J, Wilkie MJV, Lightfoot T, Coxhead J, et al. Mutations in APC, Kirsten-Ras, and P53—alternative Genetic Pathways to Colorectal Cancer. Proceedings of the National Academy of Sciences. 2002;99(14):9433-8.
24. Chen P, Bai Y, Yue-min Y, Chen Y, Pang J, Zhu H, et al. Association Between Polymorphisms of <i>MIR17HG </I>and Risk of Colorectal Cancer in the Chinese Han Population. Molecular Genetics & Genomic Medicine. 2019;7(6).
25. Cho YA, Lee J, Oh JH, Chang HJ, Sohn DK, Shin A, et al. Genetic Variation in PPARGC1A May Affect the Role of Diet-Associated Inflammation in Colorectal Carcinogenesis. Oncotarget. 2016;8(5):8550-8.
26. Kim JS-M, Coyte PC, Cotterchio M, Keogh L, Flander L, Gaff C, et al. The Impact of Receiving Predictive Genetic Information About Lynch Syndrome on Individual Colonoscopy and Smoking Behaviors. Cancer Epidemiology Biomarkers & Prevention. 2016;25(11):1524-33.