Genomic analyses identify significant molecules and biological processes in colorectal cancer cells with DNA damage

Hongmei Guo, Mason Zhang, Hanming Gu*, James Liu¹

¹SHU-UTS SILC School, Shanghai University, Shanghai, China

*Corresponding author: Hanming Gu, SHU-UTS SILC School, Shanghai University, Shanghai, China

laygmp@gmail.com

Abstract

Colorectal cancer is a major cause of cancer deaths in the US. DNA damage is considered to be a novel target for the treatment of colorectal cancer. However, the molecular mechanisms and functions are still unclear. In this study, we aim to identify the significant molecules and signaling by analyzing the RNA-seq data. The GSE189366 was created by the BGISEQ-500 (Homo sapiens). The KEGG and GO analyses indicated the p53 signaling pathway and Hippo signaling pathway are major affected processes in colorectal cancer by DNA damage. Furthermore, we identified ten key interactive molecules including CDK1, STAT3, MDM2, CCNB1, CCNA2, CDKN1A, PCNA, AURKA, PLK1, and CDC6. Our study may provide potential drug targets for colorectal cancer.

Introduction

Colorectal cancer is the third most cancer in the US¹. The total incidence of colorectal cancer has decreased, but the incidence of colon and rectal cancer in young people has increased². In recent studies, male sex and aging adult show a strong relationship with colorectal cancer³. In addition, hereditary and environmental risks are involved in the development of cancer progression⁴. Positive family history showed 10-20% of patients with colorectal cancer⁵. Young adult colorectal cancers are related to hereditary colorectal cancer syndromes and the reason is still not clear⁶. Though a number of genome studies have identified cancer susceptibility genes that are closely related to colorectal cancer risk, most factors are still unclear and subject to further study.

Colorectal cancer is caused by a complex of molecular events, for example, the mismatch repair alterations account for 5-15% of all colorectal cancer patients⁷. Colorectal cancer contains two subtypes: microsatellite unstable or microsatellite stable⁸. One study showed that patients with microsatellite stable cancers can be treated by the checkpoint inhibitors such as pembrolizumab or nivolumab⁹. With the need for better improvement, genomic alterations in the DNA damage response pathway are emerging as novel targets for treatment across different cancers¹⁰.

In our study, we analyzed the impacts of DNA damage on colorectal cancer by using the RNA-seq data. We identified a number of DEGs and functional biological processes.

We also performed the gene enrichment and created the protein-protein interaction (PPI) network to figure out the relationships among the DEGs.

Methods

Data resources

Gene dataset GSE189366 was obtained from the GEO database. The data was produced by the BGISEQ-500 (Homo sapiens) (Cancer Biology Laboratory, Molecular Biology, Ariel University, RamatHa Golan Street 65', Ariel, Israel). The analyzed dataset includes 2 of controls and 6 of topoisomerase (Top1) inhibitor irinotecan treated HCT116 cell lines.

Data acquisition and processing

The data were organized and conducted by R package as previously described¹¹⁻¹⁴. We used a classical t-test to identify DEGs with $P < 0.001$ and fold change ≥ 1.5 as being statistically significant.

The Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) KEGG and GO analyses were conducted by the R package (ClusterProfiler) and Reactome. $P < 0.05$ was considered statistically significant.

Protein-protein interaction (PPI) networks

The Molecular Complex Detection (MCODE) was used to create the PPI networks. The significant modules were produced from constructed PPI networks and String networks. The pathway analyses were performed by using Reactome (<https://reactome.org/>), and $P < 0.05$ was considered significant.

Results

Identification of DEGs in colorectal cancer after DNA damage by irinotecan

To determine the effects of DNA damage and mutations on cancer cells, we analyzed the RNA-seq data from colorectal cancer with irinotecan treatment. A total of 1010 genes were identified with the threshold of $P < 0.0001$. The top increased and

decreased genes were indicated by the heatmap and volcano plot (Figure 1). The top ten DEGs were listed in Table 1.

Enrichment analysis of DEGs in colorectal cancer after DNA damage by irinotecan

To further understand the potential mechanisms of DNA damage in cancer cells, we introduced the KEGG and GO analyses (Figure 2). We identified the significant KEGG signaling pathways including “Human papillomavirus infection”, “p53 signaling pathway”, “Axon guidance”, “Hippo signaling pathway”, “Cell cycle”, “Breast cancer”, “Gastric cancer”, “Basal cell carcinoma”, “Melanogenesis”, and “Circadian rhythm”. We also identified the biological processes of GO including “Positive regulation of catabolic process”, “Positive regulation of cellular catabolic process”, “Positive regulation of kinase activity”, “Regulation of autophagy”, “Positive regulation of protein kinase activity”, “Signal transduction by p53 class mediator”, “Gland morphogenesis”, “Positive regulation of autophagy”, “Semaphorin–plexin signaling pathway involved in axon guidance”, and “Semaphorin–plexin signaling pathway involved in neuron projection guidance”. We identified the cellular components of GO “Spindle”, “Organelle outer membrane”, “Outer membrane”, “Mitochondrial outer membrane”, “Site of polarized growth”, “Spindle pole”, “P–body”, “Lysosomal lumen”, “Protein kinase complex”, and “Striated muscle thin filament”. We identified molecular functions of GO including “Catalytic activity, acting on a tRNA”, “Aminoacyl–tRNA ligase activity”, “Ligase activity, forming carbon–oxygen bonds”, “Extracellular matrix binding”, “Frizzled binding”, “Cyclin binding”, “Tropomyosin binding”, “Semaphorin receptor activity”, “Deoxycytidine deaminase activity”, and “Cytidine deaminase activity”.

PPI network analysis

To explore the relationships among the DEGs, we created the PPI network by using 947 nodes and 3532 edges (combine score > 0.4) with the Cytoscape software. Table 2 indicated the top ten genes with the highest degree scores. The top two significant modules were shown in Figure 3. We further analyzed the PPI and DEGs with Reactome map (Figure 4) and identified the top ten significant processes including

“TP53 Regulates Transcription of Cell Cycle Genes”, “TP53 Regulates Transcription of Cell Death Genes”, “Transcriptional Regulation by TP53”, “TP53 Regulates Transcription of Death Receptors and Ligands”, “TP53 Regulates Transcription of Genes Involved in G2 Cell Cycle Arrest”, “TP53 Regulates Transcription of Genes Involved in Cytochrome C Release”, “G0 and Early G1”, “Transcription of E2F targets under negative control by p107 (RBL1) and p130 (RBL2) in complex with HDAC1”, “NR1H2 & NR1H3 regulate gene expression linked to lipogenesis”, and “TP53 Regulates Transcription of Genes Involved in G1 Cell Cycle Arrest” (Supplemental Table S1).

Discussion

Colorectal cancer is the third most common cancer in the US and the third most reason for cancer-related death¹⁵. Targeted therapy is to use the biological cancer features such as immune cells and other cells to inhibit the progression of cancer¹⁶. However, it is difficult to completely kill the cancer cells due to the complex microenvironment. This study is to help understand the complex biology of colorectal cancer through genomic analysis.

Based on our analysis, we found that the p53 signaling pathway and Hippo signaling pathway are major affected processes in colorectal cancer by DNA damage. As a tumor suppressor, p53 is majorly found in colorectal cancers. Hak-Su Goh et al found cancer patients with p53 mutation showed a poorer prognosis than those patients without mutations. Moreover, cancer patients showed responses to therapies are largely dependent on p53 mutations¹⁷. Shuxia Liu et al also found the anti-p53 antibody in the blood can be considered a diagnostic biomarker for colorectal cancer, which has the potential to distinguish the context of colorectal cancer from healthy controls or benign diseases¹⁸. The hippo pathway contains a number of proteins that regulate the growth of different tissues during development and regeneration such as cancer. The deregulation of the Hippo pathway has been shown in a wide range of different cancers including lung, colorectal, liver and prostate cancers, and it is closely related to the poor prognosis¹⁹. Besides these significant signaling pathways, the circadian rhythm is also indicated a great impact during the progression of colorectal cancer. Circadian clocks

regulated various physiology to maintain homeostasis such as metabolism, immune, and senescence²⁰⁻³⁰. Disruption of the circadian clocks plays essential roles in downstream targets and diseases including cancers. Recently finding showed targeting the core clock genes a novel approach in cancer therapy³¹.

In this study, we also discovered the interactive molecules during DNA damage-induced colorectal cancer. J K Buolamwini et al found that CDK1 inhibitors can control the cell cycle in cancers, which has been demonstrated in preclinical studies and clinical trials³². Yan Lin et al found the STAT3 and IL6 signaling in colorectal cancer show an effect on tumor-infiltrating immune cells in the tumor microenvironment³³. GPCR and its related-RGS signaling pathways are associated with several critical human diseases including cancer, arthritis, and aging-related diseases³⁴⁻⁴³. Stéphane Pelletier et al found GPCR agonists stimulate tyrosine phosphorylation of STAT3 proteins in a Rac-dependent manner, which is critical in the development of cancers³⁶. Rami M Elshazli et al found MDM2 (rs2279744) indicated relationships with colorectal cancer risks among Asians and Africans under a recessive model by stratification analysis⁴⁴. Gennadi V Glinsky et al found cancer cells indicate a stem cell-like expression profile such as increased levels of certain cell cycle marker proteins CCNB1 and CCNA2⁴⁵. Kaori Shima et al found CDKN2A promoter methylation and gene silencing are related to the CpG island methylator phenotype in colorectal cancer⁴⁶. Recently, targeting PCNA is considered to be as an effective method to repress the proliferation of cancer cells⁴⁷. AURKA is a serine/threonine kinase, which shows higher expression in cancer tissues according to TCGA database⁴⁸. Xiang Ding et al found PLK1 is a critical gene in colorectal cancer, which is positively correlated with the development of cancer⁴⁹. Chang Yu et al found CDC6 may serve as a prognostic biomarker in colorectal cancer⁵⁰.

In summary, our study discovered the significant effects of DNA damages on colorectal cancer. The P53 signaling pathway and Hippo signaling pathways are the major affected processes during the development of colorectal cancer. Our study provides valuable insights for the treatment of colorectal cancer.

Author Contributions

Hongmei Guo, Mason Zhang: Methodology and Writing. Hanming Gu, James Liu: Conceptualization, Writing- Reviewing and Editing.

Funding

This work was not supported by any funding.

Declarations of interest

There is no conflict of interest to declare.

Figure 1. Heatmap and volcano plot of HCT116 cancer cells after DNA damage by irinotecan

(A) Heatmap of significant DEGs. Significant DEGs ($P < 0.01$) were used to construct the heatmap.

(B) Volcano plot for DEGs of HCT116 cancer cells after DNA damage by irinotecan. The most significantly changed genes are highlighted by grey dots.

Figure 2. KEGG and GO analyses of DEGs in HCT116 cancer cells after DNA damage by irinotecan

(A) KEGG analysis, (B) Biological processes, (C) Cellular components, (D) Molecular functions.

Figure 3. The PPI network analyses of DEGs in HCT116 cancer cells after DNA damage by irinotecan

The cluster (A) and cluster (B) were constructed by MCODE.

Figure 4. Reactome map representation of the significant biological processes in HCT116 cancer cells after DNA damage by irinotecan

Table 1

Entrez gene	Gene symbol	Fold-change	Regulation
Top 10 down-regulated DEGs			
60489	APOBEC3G	-7.424536552	Down
3872	KRT17	-6.811169422	Down
634	CEACAM1	-6.058997335	Down
9333	TGM5	-5.764499265	Down
124975	GGT6	-5.584480624	Down
353135	LCE1E	-5.377672374	Down
1308	COL17A1	-5.143544779	Down
10814	CPLX2	-5.1191221	Down
256158	HMCN2	-5.089356824	Down
3579	CXCR2	-4.985685897	Down
Top 10 up-regulated DEGs			
7425	VGF	3.801006634	up
79365	BHLHE41	3.523496772	up
1628	DBP	2.874281498	up
9957	HS3ST1	2.521391179	up
100533182	SEC24B-AS1	2.443627402	up
101926975	LINC01844	2.352385673	up
7857	SCG2	2.338057417	up
80119	PIF1	2.246041274	up
100507283	MSANTD2-AS1	2.219586241	up
6857	SYT1	2.180446776	up

Table 2. Top ten genes demonstrated by connectivity degree in the PPI network

Gene symbol	Gene title	Degree
CDK1	Cyclin dependent kinase 1	65
STAT3	Signal transducer and activator of transcription 3	64
MDM2	MDM2 proto-oncogene	60
CCNB1	Cyclin B1	60
CCNA2	Cyclin A2	50
CDKN1A	Cyclin dependent kinase inhibitor 1A	50
PCNA	Proliferating cell nuclear antigen	47
AURKA	Aurora kinase A	47
PLK1	Polo like kinase 1	46
CDC6	Cell division cycle 6	41

References

- [1] Rawla P, Sunkara T, Barsouk A: Epidemiology of colorectal cancer: incidence, mortality, survival, and risk factors. *Prz Gastroenterol* 2019, 14:89-103.
- [2] Loomans-Kropp HA, Umar A: Increasing Incidence of Colorectal Cancer in Young Adults. *J Cancer Epidemiol* 2019, 2019:9841295.
- [3] Abancens M, Bustos V, Harvey H, McBryan J, Harvey BJ: Sexual Dimorphism in Colon Cancer. *Front Oncol* 2020, 10:607909.
- [4] Parsa N: Environmental factors inducing human cancers. *Iran J Public Health* 2012, 41:1-9.
- [5] Henrikson NB, Webber EM, Goddard KA, Scrol A, Piper M, Williams MS, Zallen DT, Calonge N, Ganiats TG, Janssens AC, Zauber A, Lansdorp-Vogelaar I, van Ballegooijen M, Whitlock EP: Family history and the natural history of colorectal cancer: systematic review. *Genet Med* 2015, 17:702-12.
- [6] Jaspersen KW, Tuohy TM, Neklason DW, Burt RW: Hereditary and familial colon cancer. *Gastroenterology* 2010, 138:2044-58.
- [7] Testa U, Pelosi E, Castelli G: Colorectal cancer: genetic abnormalities, tumor progression, tumor heterogeneity, clonal evolution and tumor-initiating cells. *Med Sci (Basel)* 2018, 6.
- [8] Boland CR, Goel A: Microsatellite instability in colorectal cancer. *Gastroenterology* 2010, 138:2073-87 e3.
- [9] Huyghe N, Baldin P, Van den Eynde M: Immunotherapy with immune checkpoint inhibitors in colorectal cancer: what is the future beyond deficient mismatch-repair tumours? *Gastroenterol Rep (Oxf)* 2020, 8:11-24.
- [10] Alhmoud JF, Woolley JF, Al Moustafa AE, Malki MI: DNA Damage/Repair Management in Cancers. *Cancers (Basel)* 2020, 12.
- [11] Yu G, Wang LG, Han Y, He QY: clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 2012, 16:284-7.
- [12] Hanming G: nuotrophils arthritis. *Research Square* 2021.
- [13] Jing L, Letian W, Hanming G: Identification of driver genes and biological signaling for alcoholic myopathy. *Research Square* 2021.
- [14] Yuan G: Identification of biomarkers and pathways of mitochondria in sepsis patients. *bioRxiv* 2021:2021.03.29.437586.
- [15] Hagggar FA, Boushey RP: Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors. *Clin Colon Rectal Surg* 2009, 22:191-7.
- [16] Gun SY, Lee SWL, Sieow JL, Wong SC: Targeting immune cells for cancer therapy. *Redox Biol* 2019, 25:101174.
- [17] Goh HS, Yao J, Smith DR: p53 point mutation and survival in colorectal cancer patients. *Cancer Res* 1995, 55:5217-21.
- [18] Liu S, Tan Q, Song Y, Shi Y, Han X: Anti-p53 autoantibody in blood as a diagnostic biomarker for colorectal cancer: A meta-analysis. *Scand J Immunol* 2020, 91:e12829.
- [19] Harvey KF, Zhang X, Thomas DM: The Hippo pathway and human cancer. *Nat Rev Cancer* 2013, 13:246-57.
- [20] Zhu Z, Xu L, Cai T, Yuan G, Sun N, Lu C, Qian R: Clock represses preadipocytes adipogenesis via GILZ. *J Cell Physiol* 2018, 233:6028-40.
- [21] Yuan G, Hua B, Yang Y, Xu L, Cai T, Sun N, Yan Z, Lu C, Qian R: The Circadian Gene Clock Regulates Bone Formation Via PDIA3. *J Bone Miner Res* 2017, 32:861-71.
- [22] Cai T, Hua B, Luo D, Xu L, Cheng Q, Yuan G, Yan Z, Sun N, Hua L, Lu C: The circadian protein CLOCK regulates cell metabolism via the mitochondrial carrier SLC25A10. *Biochim Biophys Acta Mol Cell Res* 2019, 1866:1310-21.

- [23] Yuan G, Hua B, Cai T, Xu L, Li E, Huang Y, Sun N, Yan Z, Lu C, Qian R: Clock mediates liver senescence by controlling ER stress. *Aging* 2017, 9:2647-65.
- [24] Fan XF, Wang XR, Yuan GS, Wu DH, Hu LG, Xue F, Gong YS: [Effect of safflower injection on endoplasmic reticulum stress-induced apoptosts in rats with hypoxic pulmonary hypertension]. *Zhongguo Ying Yong Sheng Li Xue Za Zhi* 2012, 28:561-7.
- [25] Xu L, Cheng Q, Hua B, Cai T, Lin J, Yuan G, Yan Z, Li X, Sun N, Lu C, Qian R: Circadian gene Clock regulates mitochondrial morphology and functions by posttranscriptional way. *bioRxiv* 2018:365452.
- [26] Zhu Z, Hua B, Shang Z, Yuan G, Xu L, Li E, Li X, Sun N, Yan Z, Qian R, Lu C: Altered Clock and Lipid Metabolism-Related Genes in Atherosclerotic Mice Kept with Abnormal Lighting Condition. *Biomed Res Int* 2016, 2016:5438589.
- [27] Yuan G, Xu L, Cai T, Hua B, Sun N, Yan Z, Lu C, Qian R: Clock mutant promotes osteoarthritis by inhibiting the acetylation of NFkappaB. *Osteoarthritis Cartilage* 2019, 27:922-31.
- [28] Mao SZ, Fan XF, Xue F, Chen R, Chen XY, Yuan GS, Hu LG, Liu SF, Gong YS: Intermedin modulates hypoxic pulmonary vascular remodeling by inhibiting pulmonary artery smooth muscle cell proliferation. *Pulm Pharmacol Ther* 2014, 27:1-9.
- [29] Zhu Z, Hua B, Xu L, Yuan G, Li E, Li X, Sun N, Yan Z, Lu C, Qian R: CLOCK promotes 3T3-L1 cell proliferation via Wnt signaling. *IUBMB Life* 2016, 68:557-68.
- [30] Li J, Wang W, Gu H: Identification of biological processes and signaling pathways for the knockout of REV-ERB in mouse brain. *bioRxiv* 2021:2021.11.22.469579.
- [31] Sulli G, Lam MTY, Panda S: Interplay between Circadian Clock and Cancer: New Frontiers for Cancer Treatment. *Trends Cancer* 2019, 5:475-94.
- [32] Buolamwini JK: Cell cycle molecular targets in novel anticancer drug discovery. *Curr Pharm Des* 2000, 6:379-92.
- [33] Lin Y, He Z, Ye J, Liu Z, She X, Gao X, Liang R: Progress in Understanding the IL-6/STAT3 Pathway in Colorectal Cancer. *Onco Targets Ther* 2020, 13:13023-32.
- [34] Yuan G, Huang Y, Yang ST, Ng A, Yang S: RGS12 inhibits the progression and metastasis of multiple myeloma by driving M1 macrophage polarization and activation in the bone marrow microenvironment. *Cancer Commun (Lond)* 2022, 42:60-4.
- [35] Yuan G, Yang S, Ng A, Fu C, Oursler MJ, Xing L, Yang S: RGS12 Is a Novel Critical NF-kappaB Activator in Inflammatory Arthritis. *iScience* 2020, 23:101172.
- [36] Pelletier S, Duhamel F, Coulombe P, Popoff MR, Meloche S: Rho family GTPases are required for activation of Jak/STAT signaling by G protein-coupled receptors. *Mol Cell Biol* 2003, 23:1316-33.
- [37] Yuan G, Yang S, Yang S: Macrophage RGS12 contributes to osteoarthritis pathogenesis through enhancing the ubiquitination. *Genes & Diseases* 2021.
- [38] Yuan G, Yang S, Gautam M, Luo W, Yang S: Macrophage regulator of G-protein signaling 12 contributes to inflammatory pain hypersensitivity. *Ann Transl Med* 2021, 9:448.
- [39] Fu C, Yuan G, Yang ST, Zhang D, Yang S: RGS12 Represses Oral Cancer via the Phosphorylation and SUMOylation of PTEN. *J Dent Res* 2020:22034520972095.
- [40] Yuan G, Yang S, Liu M, Yang S: RGS12 is required for the maintenance of mitochondrial function during skeletal development. *Cell Discov* 2020, 6:59.
- [41] O'Brien JB, Wilkinson JC, Roman DL: Regulator of G-protein signaling (RGS) proteins as drug targets: Progress and future potentials. *J Biol Chem* 2019, 294:18571-85.
- [42] Yuan G, Yang S, Yang S, Ng A, Oursler MJ: RGS12 is a critical proinflammatory factor in the pathogenesis of inflammatory arthritis via acting in Cox2-RGS12-NF kappa B pathway activation loop. *J Bone Miner Res: WILEY 111 RIVER ST, HOBOKEN 07030-5774, NJ USA*, 2019. pp. 147-.
- [43] Yuan G, Fu C, Yang ST, Yuh DY, Hajishengallis G, Yang S: RGS12 Drives Macrophage Activation and Osteoclastogenesis in Periodontitis. *J Dent Res* 2021:220345211045303.

- [44] Elshazli RM, Toraih EA, Elgaml A, Kandil E, Fawzy MS: Genetic polymorphisms of TP53 (rs1042522) and MDM2 (rs2279744) and colorectal cancer risk: An updated meta-analysis based on 59 case-control studies. *Gene* 2020, 734:144391.
- [45] Glinsky GV: Genomic models of metastatic cancer: functional analysis of death-from-cancer signature genes reveals aneuploid, anoikis-resistant, metastasis-enabling phenotype with altered cell cycle control and activated Polycomb Group (PcG) protein chromatin silencing pathway. *Cell Cycle* 2006, 5:1208-16.
- [46] Shima K, Nosho K, Baba Y, Cantor M, Meyerhardt JA, Giovannucci EL, Fuchs CS, Ogino S: Prognostic significance of CDKN2A (p16) promoter methylation and loss of expression in 902 colorectal cancers: Cohort study and literature review. *Int J Cancer* 2011, 128:1080-94.
- [47] Cardano M, Tribioli C, Prosperi E: Targeting Proliferating Cell Nuclear Antigen (PCNA) as an Effective Strategy to Inhibit Tumor Cell Proliferation. *Curr Cancer Drug Targets* 2020, 20:240-52.
- [48] Du R, Huang C, Liu K, Li X, Dong Z: Targeting AURKA in Cancer: molecular mechanisms and opportunities for Cancer therapy. *Mol Cancer* 2021, 20:15.
- [49] Ding X, Duan H, Luo H: Identification of Core Gene Expression Signature and Key Pathways in Colorectal Cancer. *Front Genet* 2020, 11:45.
- [50] Yu C, Chen F, Jiang J, Zhang H, Zhou M: Screening key genes and signaling pathways in colorectal cancer by integrated bioinformatics analysis. *Mol Med Rep* 2019, 20:1259-69.