Global Journal of Basic Science ISSN : XXXX-XXXX

Research Article

A computational approach reveals the critical role of PARK2 gene and the potential infectious pathways and HPV infection in colorectal cancer

Leena Hussein Bajrai [1,2], Thamir A. Alandijany [1,3], Arwa A. Faizo [1,3], Isra Alsaady [1,3], Raja Jeet [4], Mohammad Mobashir [1,5], and Esam Ibraheem Azhar [1,3,*]

1 Special Infectious Agents Unit, King Fahd Medical Research Centre, King Abdulaziz University, Jeddah, Saudi Arabia; lbajrai@kau.edu.sa
2 Biochemistry Department, Faculty of Sciences, King Abdulaziz University, Jeddah, Saudi Arabia
3 Medical Laboratory Sciences Department, Faculty of Applied Medical Sciences, King Abdulaziz University, Jeddah, Saudi Arabia; talandijany@kau.edu.sa
4 Department of Botany, G D College, Begusarai, Bihar, India- 851101
5 Department of Biomedical Laboratory Science, Faculty of Natural Sciences Norges teknisk-naturvitenskapelige universitet (NTNU), N-7491 Trondheim, Norway.
* Correspondence: eazhar@kau.edu.sa (E.I.A.)

Citation: Bajrai L.H., Alandijany T.A., Faizo A.A., Alsaady I., Jeet R., Mobashir M., and Azhar E.I.. A computational approach reveals the critical role of PARK2 gene and the potential infectious pathways and HPV infection in colorectal cancer. *Jour. Bas. Sci*. 2025, 1(4). 1-11.

Received: November 05, 2024
Revised: January 20, 2025
Accepted: February 02, 2025
Published: February 07, 2025

doi: 10.63454/jbs20000013

ISSN: XXXX-XXXX

Abstract: Colorectal cancer is a type of cancer that begins in the colon or rectum, and it is also known as colon cancer or rectal cancer, depending on where it begins. Colorectal cancer is one of the most common and deadly cancers, with a wide range of outcomes and medication responses. The human papillomavirus produces benign and malignant neoplasms in epidermal and mucosal epithelial cells. HPVs that are often found in anogenital malignancies, particularly cancers of the cervix and anus, are classified as high-risk or oncogenic. The PARK2 gene is one of the most important genes known to be linked to the ubiquitin proteasomal system and to play a role in cancer and other serious disorders. In order to comprehend the genotype-phenotype relationship in complicated disorders, high-throughput data (genomic and proteomic) is commonly collected. In this study, we focused on the relationship between the human papillomavirus, the PARK2 gene, gene expression patterns, altered genes, and their activities in colorectal cancer. We got the data from a publically available database for gene expression, mutations, and structural variants in human colorectal cancer for this study. We looked at it in terms of clinical importance by combining the Gene Expression Omnibus, TCGA, and ProteinAtlas datasets. According to our findings, the PARK2 gene impacts not only the ubiquitin proteasomal systems, but also key pathways, particularly those involved with the immune system, which may be directly responsible for colorectal and other types of cancer, as well as other key human disorders. Furthermore, it demonstrates that the HPV-infected genes and their estimated roles were potentially shared with genes that are important components of well-known colorectal cancer control and coordination pathways.

Keywords: PRKN gene, CRC, differentially expressed genes, enriched pathways, survival analysis, clinical relevance, HPV infection

1. Introduction

Colorectal cancer (CRC) is the most common cancer diagnosed in both men and women worldwide. CRC is a complex illness influenced by a variety of variables ranging from lifestyle to genetics and the environment[1-5]. CRC formation is characterized by a multistep phenotypic change sequence involving many molecular signaling pathways. Different stages of neoplastic transformation are linked to molecular alterations in specific genes. Genetic instability is a key aspect of CRC, which is caused by two distinct pathways. Chromosomal instability (CIN), which is defined by substantial alterations in chromosome number and structure such as deletions, gains, translocations, and other chromosomal rearrangements, is the most prevalent. Hypermutation and microsatellite instability (MSI) are caused by poor DNA mismatch repair (MMR) in the second mechanism[4-13].

CRC can originate in the colon or the rectum, and depending on where it starts, it's called colon cancer or rectal

cancer. Colon and rectal cancers are frequently lumped together since they have many characteristics. Infection with the human papillomavirus (HPV) causes benign to malignant neoplasms in epidermal or mucosal epithelial cells. Furthermore, many forms of HPVs (such as HPV16, 18, 31, and 45) are frequently discovered in many cancer types and are classified as high-risk. HPV infection is characterized by the incorporation of the viral genome into the cancer cell's or genome's genome. Furthermore, HPV6 and HPV11 are low-risk or non-oncogenic viruses[3, 8, 14-27].

Human colorectal tumors have recurrent mutations and consistent overexpression for selected genes associated with critical signalling pathways such as immune system related pathways, WNT, PI3K, TP53, TGF-ß, and cancer signalling pathways, and CCR is one of the leading causes of death among the global population. PARK2, also known as PRKN, is a crucial gene that is known to be related with the ubiquitin proteasomal system (UPS) and plays a key role in cancer and other serious disorders[2, 7, 8, 28-40]. PARK2 is an E3 ubiquitin ligase that recognizes proteins on the outer membrane of mitochondria and mediates the clearance of damaged mitochondria via autophagy and proteasomal mechanisms. It also plays a role in the ubiquitination process (by which polyubiquitin chains are formed) and directed towards degradation in proteasomes or lysosomes. It also improves cell viability by inhibiting apoptosis, both mitochondria-dependent and mitochondria-independent. Mutations are linked to mitochondrial malfunction, which can result in neuronal death in Parkinson's disease and abnormal metabolism in cancer. Mitophagy, clearance of (reactive oxygen species) ROS, cell survival (through NF-kB activation), and cancer are all recognized functions of the PRKN gene[41-48].

As a result, we developed the goal that a well-controlled, integrative, and comparative study would be more informative, and based on the importance of the PARK2 gene and its association with a number of critical functions and diseases, we decided to study the role of the PARK2 gene, and finally studied differentially expressed genes (DEGs) and enriched pathways from gene expression data (obtained from Gene Expression Omnibus (GEO)). Moreover, we have also explored different types of mutations, coexpression with respect to the PRKN gene (TCGA[33,34] and cBioPortal), PRKN expression in tissues (Protein Atlas).

2. Methods

Colorectal cancer (CRC) is among the most frequent and highly lethal disease with heterogeneous outcomes and drug response. High-throughput data (genomic and proteomic) are frequently generated with the goal to understand the genotype-phenotype relationship in the complex diseases. We have selected genome-wide mRNA expression data for CRC tumor samples. By applying computational approach and integrating experimental data, we have unraveled the critical genes and the pathways which appear to be associated with PRKN gene. Here, we applied interdisciplinary approach to unravel the role of PARK2 gene in CRC and its clinical relevance. Here, we have selected the dataset GSE21510 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE21510)[49] for CRC and processed it for normalization and log2 values of all the mapped genes as shown in the workflow Figure 1a. This dataset contains 148 samples where 104 samples are laser capture microdissection (LCM) cancer, 19 homogenized cancer, and 25 samples are homogenized normal and for differential gene expression analysis, we have compared the cancer samples with normal samples.

For HPV-infected CRC differential expression profiling we have adapted the list of DEGs from the previously published work and for clinical data cBioPortal has been used to access TCGA datasets for CRC mutated and structural variant dataset. For survival analysis, the PROGgeneV2 online web server has been used.

There are total of 148 microarray datasets (GSE21510) obtained from LCM and homogenized tissues of colorectal cancer patients. These samples expression values were normalized using robust multi-array average (RMA) method. The normalized gene expression levels were presented as log2-transformed values by RMA. In summary the steps involved for the entire study are raw file processing, intensity calculation and normalization and for normalization, GCRMA, RMA, and EB are the most commonly used approaches. Here, we have used EB for raw intensity normalization. Now, we proceed for our goal which is to understand the gene expression patterns and its inferred functions and also the impact of PARK2 gene. For differential gene expression prediction and statistical analysis, MATLAB functions (e.g., mattest) has been used and for pathway analysis, we used KEGG database and used in-house code designed to pathway and network analysis. For generating DEGs network, FunCoup2.0 has been used for all the networks throughout the work and cytoscape[63] has been used for network visualization. FunCoup predicts four different classes of functional coupling or associations such as protein complexes, protein-protein physical interactions, metabolic, and signaling pathways. Most of our coding and calculations MATLAB has been used[29, 50-63].

3. Results

3.1. PARK2/PRKN gene mainly controls genes belonging to UPS: We have mapped out all of the genes that are directly connected to the PARK2 gene in this study, as well as their connectedness to the PRKN gene and their relationships to one another. We've shown the drawing in Figure 1a, where we've summarized the study's main focus point. We explicitly stated that we collected the datasets mostly from the GEO and TCGA databases, which are mostly

linked to expression profiling. Gene expression profiling was carried out on the datasets obtained from GEO, and TCGA data was used to investigate the influence of PARK2 and its related genes in clinical samples. In the first step of our research, we identified the genes that are directly linked to PARK2 (Figure 1b). The connected functions with PARK2 genes have also been investigated, and it is obvious that the most important pathways are linked to the PARK2 gene (Table 1). Immune system pathways and major infectious disease-related pathways are among the essential pathways. Figure 1b shows a substantial number of ubiquitin proteasomal system genes, followed by key infectious disease-associated pathways and cancer signaling pathways. Based on our first findings, we believe that the PARK2 gene is significant not just in infectious disease but also in other human diseases such as cancer and brain disorders. A vast number of genes are directly linked to the PARK2 gene, and the bulk of these genes have been well-studied for their roles in human disorders like infection, inflammation, and cancer. Moreover, we have also summarized the overall alterations associated with PRKN in case of different CRCs (Figure 1c) where deep deletion and mutation appear highly dominant followed by amplification and structural variant. Before moving on to the next step, we examined the clinical significance of the PRKN gene in CRC and plotted a survival curve (Figure 2a) using PROGeneV2; survival analysis revealed that overexpression of the PARK2 gene is significant with a p-value of 0.0385, implying that overexpression of the PARK2 gene may play a role in the patient's survival. In Figure 2b, all the overall mutations in PRKN gene have been presented for this specific cancer subtype.

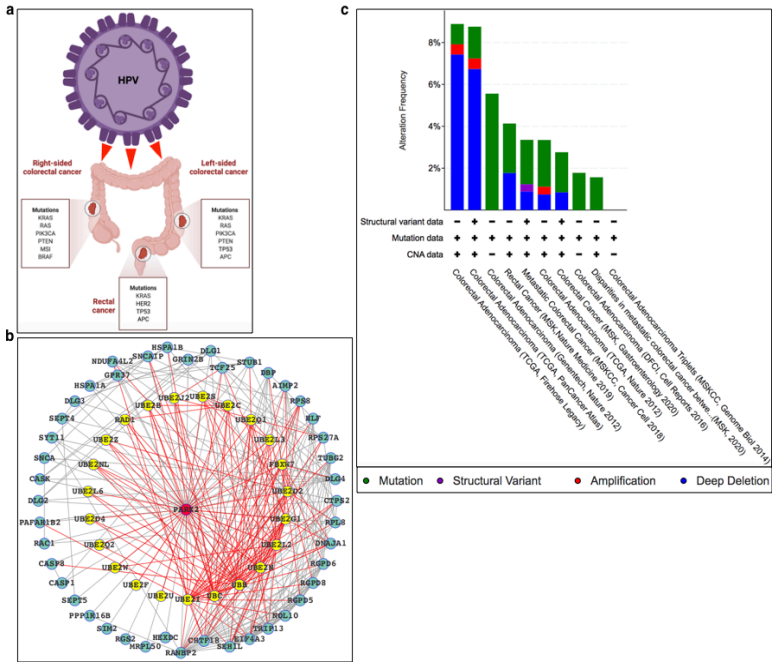

Figure 1: Colorectal cancer, PARK2 or PRKN and the associated genes. Colorectal cancer, PARK2 or PRKN and the associated genes. (a) An sketch (drawn from BioRender (www.biorender.com)) for presenting colorectal cancer and well-known biomarkers and the HPV infection and (b) genes connected with PARK2. The network shows the connectivity with PARK2 and also connectivity with PARK2 gene components. (c) To display the overall alterations (CNA, mutation, and structural variation) in different types of CRC datasets.

Table 1. PARK2/PRKN-linked genes and the associated pathways.

Genes	Pathways associated with PARK2 genes
AIMP2, CASK, CASP1 CASP8, CHTF18, CTPS2 DBP, DLG1, DLG2 DLG3, DLG4, DNAJA1 EIF4A3, FBXW7, GPR37 GRIN2B, HEXDC, HLF HSPA1A, HSPA1B MRPL50, NDUFA4L2 NOL10, PAFAH1B2 PARK2, PPP1R16B RAC1, RAD1, RANBP2 RGPD5, RGPD6, RGPD8 RGS2, RPL8, RPS27A RPS8, SEH1L, SEPT4 SEPT5, SIM2, SNCA SNCAIP, STUB1, SYT11 TCF25, TRIP13, TUBG2 UBB, UBC, UBE2B, UBE2C UBE2D2, UBE2D4, UBE2F UBE2G1, UBE2I, UBE2J2 UBE2L2, UBE2L3, UBE2L6 UBE2N, UBE2NL, UBE2Q1 UBE2Q2, UBE2S, UBE2U UBE2W, UBE2Z	p53 signaling pathway, Ubiquitin mediated proteolysis Protein processing in endoplasmic reticulum Phagosome, PI3K-Akt signaling Apoptosis, Wnt signaling, Axon guidance,VEGF signaling, Osteoclast differentiation Hippo Signaling, Focal adhesion, Adherens and tight junction Antigen processing and presentation Toll-like receptor signaling, NOD-like receptor signaling Cytosolic DNA-sensing, Natural killer cell mediated cytotoxicity TCR and BCR signaling, Fc epsilon RI signaling Fc gamma R-mediated phagocytosis TNF signaling, Leukocyte transendothelial migration Circadian entrainment and Long-term potentiation Neurotrophin signaling, Regulation of actin cytoskeleton Estrogen and Oxytocin signaling Alzheimer's, Parkinson's, Huntington's, Chagas, Prion disease Amyotrophic lateral sclerosis (ALS) Bacterial invasion of epithelial cells Epithelial cell signaling in Helicobacter pylori infection Chemokine signaling, NF-kappa B signaling pathway Sphingolipid signaling pathway, Neuroactive ligand-receptor interaction, Cell cycle, Spliceosome, Pathways in cancer Colorectal and Pancreatic cancer, Renal cell carcinoma Metabolism, Ribosome, RNA transport, mRNA surveillance PPAR, MAPK, Ras, RAP1 signaling cGMP-PKG and cAMP signaling

3.2. PARK2 gene network and gene expression profiling unravels the critical pathways associated with CRC: We highlighted the genes associated with UPS (yellow color) in Figure 1b after analyzing the PARK2 gene and the directly associated genes with it, as well as the clinical significance of the overexpressed PARK2 gene in CRC. The other genes belong to different pathways, primarily p53, Wnt, immune, and cancer pathways. Table 1 shows a list of pathways in which the majority of the pathways are a major cause of CRC and other types of cancers, with UPS genes being the most common cause of infection, inflammation, and malignancies, including colorectal cancer. As previously stated, we used a gene expression dataset from GEO that included 104 tumor samples and 44 normal samples, and we examined the enriched pathways for this CRC dataset. We produced the venn diagram for differentially expressed genes and their enriched pathways in the instance of cancer for the selected GEO dataset because there are two types of cancer cells (homogenized and LCM). In this case, the majority of the genes (3987) are shared between the two conditions, 806 are homogenized condition specific, and 2513 are LCM specific genes, whereas in the case of enriched pathways, almost all of the enriched (149) pathways are shared, with the exception of 6 LCM specific pathways (Figure 3a and 3b).

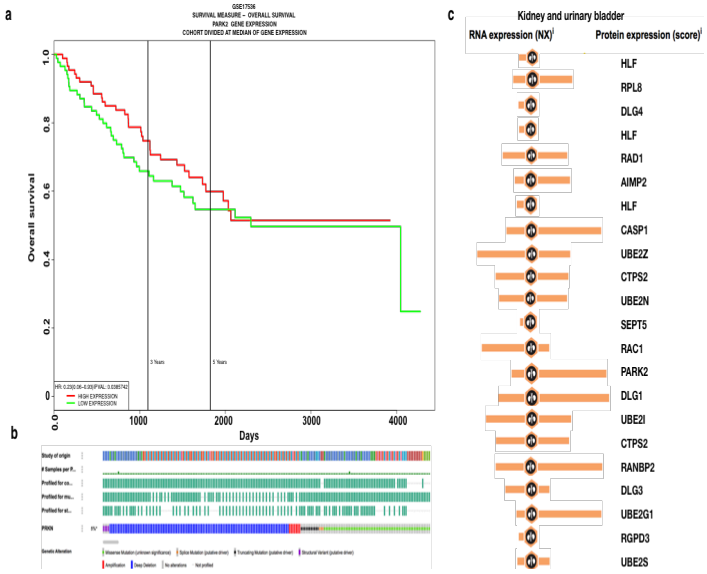

Figure 2: Clinical relevance of PARK2 gene in terms of survival in colorectal cancer. Clinical relevance of PARK2 gene in terms of survival in colorectal cancer. (a) The clinical significance of the PARK2 gene in CRC was analyzed and plotted a survival curve using PROGeneV2 where survival analysis revealed that overexpression of the PRKN gene is significant with a p-value of 0.0385. (b) PRKN gene aberrations based on TCGA datasets for CRC and the Figure was generated by using cBioPortal. (c) RNA and protein expression in case of top-ranked genes (based on differential expression profiling) by using ProteinAtlas database.

We see that a high number of the enriched pathways correspond to the PARK2 gene network's related pathways (Supplementary data S1). We have mapped the PARK2 genes (Figure 1b) with the DEGs for the selected dataset and top-ranked genes/proteins expression in tissue have been analyzed by using Protein Atlas database (Figure 2c). based on protein atlas database, it appears that PRKN, CASP1, DLG1, EIF4A3, STUB1, and PPP1RR16B genes with higher expression level. Further, we observe that there are 21 genes that belong to the DEGs list (Figure 3b) and have fold change greater than 2.0 (up regulated) or less than -2.0 (down regulated) with p-values less than 0.05, similar to comparative (associated with PARK2 gene network and enriched) pathways analysis. Following that, we showed the top 30 genes up and down regulated in the CRC dataset, along with their fold changes and p-values (Figure 3c and 3d), and we noticed that the fold for down-regulated genes is comparatively low (Figure 3d), implying that there are more genes whose expression is severely suppressed as a result of cancer in CRC. Based on this analysis we could clearly see that irrespective of homogeneity or LCM, most of source of CRC cause are common in terms of genes and finally, the pathways (Supplementary data S1). After analyzing the DEGs and relevant data, we also analyzed the co-expression pattern (Figure 3e) with respect to the PRKN gene where the TCGA and cBioPortal have been used and among these top-rated (based on the spearman correlation), a large number of the co-expressed genes belong to ubiquitin proteasomal, infection, and inflammatory systems while the other genes are well-established the part of cancer associated pathways including CRC.

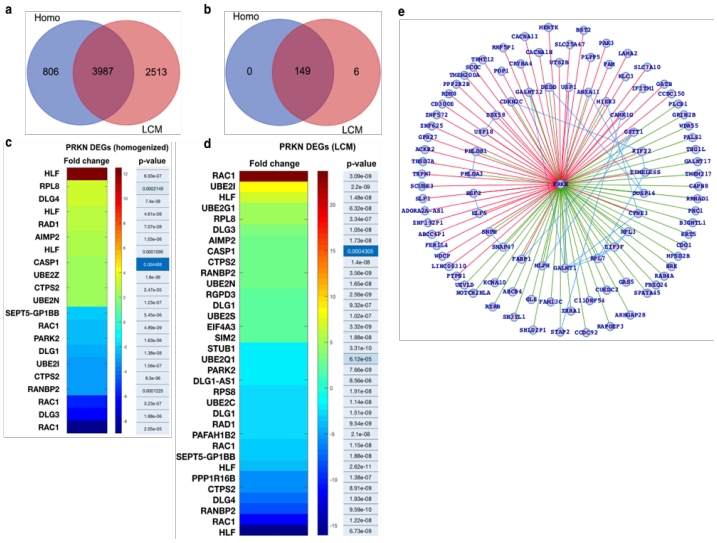

Figure 3: Differential expression profiling and enriched pathways. (a) Venn diagram for DEGs in both the conditions (Homogenized and LCM). (b) Venn diagram for the enriched pathways in both the conditions (Homogenized and LCM). (c) Top ranked DEGs with the respective fold changes and the p-values (homogenized). (d) Top ranked DEGs with the respective fold changes and the p-values (LCM). (e) Co-expression network of PRKN gene generated by using cBioPortal TCGA database. The selected genes show very strong correlation (spearman's correlation) either +1.0 (green color edge) or -1.0 (red color edge). The other connections are to show the connection based on network database (FunCoup).

3.3. Clinical relevance of PARK2 gene and the linkage of HPV-infected CRC DEGs: In order to understand and unearth more detail information about PRKN and its associated genes, we went to in depth analysis by including more clinical samples from TCGA database. Here, we have analyzed the overall frequency of all the PARK2 genes in colon, rectal, and colorectal adenocarcinoma and also percentage of individual gene in colorectal adenocarcinoma where we observe the huge impact in terms of higher mRNA expression (Figure 4a). Even at individual level, PARK2 gene itself shows higher mRNA expression among 10% of the patients. Among the highly altered expression are RAC1 (19%), UBE2I (13%), UBE2Q1 (15%), RPL8 (14%), TUBB1 (11%), and more shown in Figure 4b. In the first

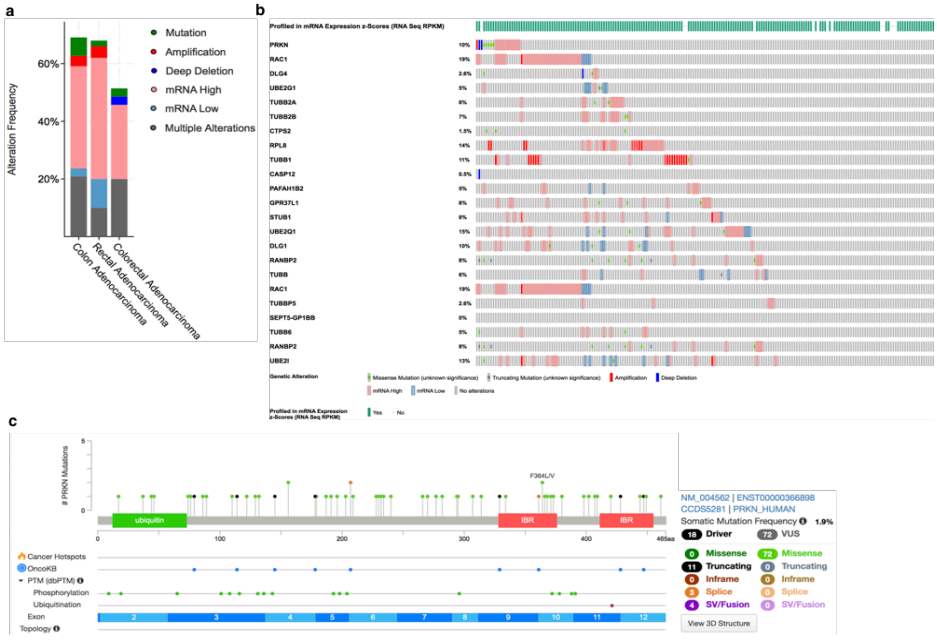

Figure 4: Over-expression of PRKN and associated genes for clinical human samples (TCGA database). (a) Overall changes in colon, rectal, and colorectal adenocarcinoma. (b) Individual representation of over-expression of PRKN associated genes. (c) Different types of mutations associated with PRKN gene in case of CRC (result generated from TCGA dataset by using cBioPortal).

section of this work we have shown the relevance of PRKN gene in terms of expression and after expression analysis we have also decided to investigate the relevance of it in case of mutation and mutational dataset from TCGA database for the same cancer CRC. In case of colorectal cancer, we observe that there are multiple types of alterations which happens and dominantly for PARK2 and their associated genes (Figure 4). In addition to it, we have also analyzed all the different types of mutations (such driver, missense, PTM, splice, and SV/fusion) associated with PRKN gene in its different regions (such as PRKN domains) in CRC (Figure 4c).

To explore the linkage of HPV-infected CRC DEGs, we have collected the DEGS from the previously published work[18,40], collected the CRC mutated genes, CRC structural variant genes, and the DEGs of GSE21510 and compared all these genes and the inferred pathways. Figure 5 was prepared to display the linkage and we observe that there were 12 genes common with HPV-infected CRC DEGs, mutated, and structural variant genes list, 29 genes shared with mutated list, 56 genes shared between mutated and structural variant list (Figure 5a). Similar to it, 35 pathways common with HPV-infected CRC DEGs, mutated, and structural variant genes list, 38 pathways shared with HPV-infected CRC and mutated list, and no pathways shared between mutated and structural variant list (Figure 5b). Moreover, the list of mutated genes and structural variants genes were presented in Figure 5c and 5d where it appears that most of the structural variant genes were common with mutated genes. Finally, the comparative study of HP-infected CRC DEGs and non-HPV CRC DEGs shows that overall they share 24 genes with each other and three genes were shared between HP-infected CRC DEGs and CRC-homogenized DEGs, and nine genes shared between HP-infected CRC DEGs and CRC LCM DEGs (Figure 5e).

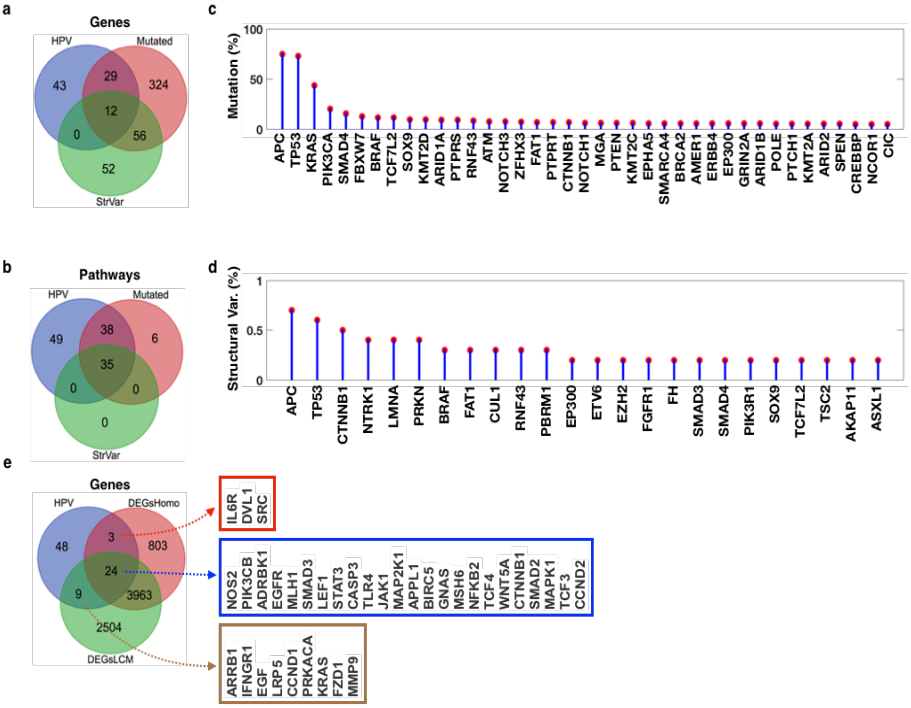

Figure 5: Linkage of HPV-infected CRC genes with CRC DEGs and mutated genes and also the functions. (a) and (b) Genes and pathways common and specific to HPV-infected CRC and CRC. (c) Highly mutated genes (>= 5.0%) in case of CRC (TCGA database via cBioPortal). (d) Genes with higher (>0.1%) structural variants in CRC (TCGA database via cBioPortal).

4. Dicussion: CRC is globally a leading cause of death and is considered as a disease with heterogeneous outcomes and drug response. With the advancement in technology, high-throughput data (genomic and proteomic) are frequently generated at very high accuracy with the goal to understand the genotype-phenotype relationship in the complex diseases. This carcinogenesis is a complex and multistep process. It may involve environmental and lifestyle factors, genetic aberration and viral infection. In general, the genetic aberrations, tumor suppressor and MMR genes inactivation or oncogenes activation are known to be involved in cell growth, proliferation, and differentiation and thus are implicated in CCR development[1, 3, 14, 28, 29, 64-66].

There are a number of genes and pathways known to be associated with CRC as a cause and possible future biomarkers or targeting pathways and PARK2 gene is one of them. PARK2 gene is known to be associated with UPS, known to be associated with a list of critical pathways, and play critical roles in cancer and other severe diseases. It is a E3 ubiquitin ligase which plays critical roles in ubiquitination and de-ubiquitination process, also enhances cell survival by suppressing apoptosis. In short, we could say that the known functions associated with PARK2 gene are mitophagy, clearance of ROS, cell survival, and tumorigenesis. Although, there are a number of works which are related to PARK2 gene and gene expression data but different from previous studies, we applied interdisciplincary approach to unravel the role of PARK2 gene in CRC and its clinical relevance and the dataset is microarray mRNA data i.e., GSE21510. In this work, we have studied and presented the results from microarray expression data, TCGA mRNA expression data, mutational data, HPV-infected CRC expression data, network-level understanding, and finally, the clincal relevance of the data followed by survival plot of overexpression PARK2 gene CRC[43-48].

Overall, we conclude that PARK2 gene appears to be associated with a bunch of important genes which control critical pathways associated with colorectal cancer, more cancers, and other human diseases. In terms of clinical aspect, its overexpression is important for patients' survival. The critical pathways associated with colorectal cancer are regulation of actin cytoskeleton, insulin signaling, Akt signaling, focal adhesion, and neuroactive ligand-receptor interaction pathways. For better understanding and clarity, we have presented a supplementary file which includes all the genes and the enriched pathways in both the conditions with their respective data (fold changes and p-values). Here, we clearly see all the enriched pathways with their p-values and the p-values are as low as 10e-20. Most of the highly (top-ranked) enriched pathways are well-known to be associated with human cancer. Among the highly interesting and critical cancer associated pathways are p53, ErbB, AKT, apoptosis, TGF, VEGF, cell-cycle, junctions, JAK-STAT pathways, etc[5, 6, 14, 16, 28, 35, 38, 65, 67-69].

To tailor CRC diagnosis and treatment, it is critical and essential to first prepare model the level of heterogeneity by defining subtypes of patients with homogeneous biological, level of complexity, predict promising biomarkers, and clinical characteristics and second match these subtypes to cell lines for which extensive pharmacological data is available, thus linking targeted therapies to patients most likely to respond to treatment. Further, in terms of future perspectives, the work could be combined preparing the signaling and gene regulatory networks and incorporate the current datasets to predict the dynamics of proteins and genes.

5. Conclusions: Based on the combined study and results, we conclude that PARK2 gene appear to be associated with important genes which control critical pathways directly or indirectly associated with colorectal cancer and other human diseases. In terms of clinical aspect, its overexpression is important for patients survival. The critical pathways associated with colorectal cancer are major immune signaling pathways, regulation of actin cytoskeleton, insulin signaling, Akt signaling, focal adhesion, and neuroactive ligand-receptor interaction pathways.

Author Contributions: Conceptualization, L.H.B., T.A.A., A.A.F., I.A., M.M., R.J., and E.I.A.; methodology, L.H.B., T.A.A., A.A.F., I.A., M.M., R.J., and E.I.A.; software, E.I.A. and M.M.; validation, L.H.B., T.A.A., M.M., and E.I.A.; formal analysis, L.H.B., M.M., S.A., E.I.A.; investigation, L.H.B., T.A. A., A.A.F., I.A., M.M., and E.I.A.; resources, M.M., and E.I.A.; data curation, M.M., and E.I.A.; writing—original draft preparation, L.H.B., T.A.A., A.A.F., I.A., M.M., R.J., and E.I.A.; writing—review and editing, L.H.B., T.A.A., A.F., M.M., and E.I.A.; visualization, L.H.B., M.M., and E.I.A.; supervision, M.M., and E.I.A.; project administration, E.I.A.; funding acquisition, M.M. and E.I.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, Saudi Arabia, grant number FP-5-42 and The APC was funded by FP-5-42.

Acknowledgments: We are thankful to DSR, KAU and for providing us the resources and the facility to carry out the work to Special Infectious Agents Unit, King Fahd Medical Research Centre, King Abdulaziz University, Jeddah, Saudi Arabia, Medical Laboratory Sciences Department, Faculty of Applied Medical Sciences, King Abdulaziz University, Jeddah, Saudi Arabia, Biochemistry Department, Faculty of Sciences, King Abdulaziz University, Jeddah, Saudi Arabia, and King Fahd Medical Research Center, King Abdulaziz University, P. O. Box 80216, Jeddah 21589, Saudi Arabia.

Availability of data and materials: We have used the publicly available dataset GSE21510 which are freely available. Furthermore, the generated results and the analyzed outcomes during the current study are available in the manuscript and if any readers need any further information regarding coding, then corresponding author (E.I.A.) will be happy to supply.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the related data are supplied in this work or have been referenced properly.

References

1. Bray, F., et al., *Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.* CA Cancer J Clin, 2024. **74**(3): p. 229-263.
2. Dang, H.X., et al., *The clonal evolution of metastatic colorectal cancer.* Science Advances, 2020. **6**(24): p. eaay9691.
3. Dekker, E., et al., *Colorectal cancer.* The Lancet, 2019. **394**(10207): p. 1467-1480.

4. Jung, G., et al., *Epigenetics of colorectal cancer: biomarker and therapeutic potential.* Nature Reviews Gastroenterology & Hepatology, 2020. **17**(2): p. 111-130.
5. Kouzminova, N., T. Lu, and A.Y. Lin, *Molecular Basis of Colorectal Cancer.* The New England Journal of Medicine, 2010. **362**(13): p. 1245-1247.
6. *Colorectal cancer.* Nature Reviews Disease Primers, 2015. **1**(1): p. 15066.
7. Dunn, E.F., et al., *Dasatinib sensitizes KRAS mutant colorectal tumors to cetuximab.* Oncogene, 2011. **30**(5): p. 561-574.
8. Emaduddin, M., et al., *Cell growth, global phosphotyrosine elevation, and c-Met phosphorylation through Src family kinases in colorectal cancer cells.* Proceedings of the National Academy of Sciences, 2008. **105**(7): p. 2358-2362.
9. Emaduddin, M., et al., *Odin (ANKS1A) is a Src family kinase target in colorectal cancer cells.* Cell Communication and Signaling, 2008. **6**(1): p. 7.
10. Hassan, N.Z.A., et al., *Integrated Analysis of Copy Number Variation and Genome-Wide Expression Profiling in Colorectal Cancer Tissues.* PLoS ONE, 2014. **9**(4): p. e92553.
11. Lee, A.J.X., et al., *Chromosomal Instability Confers Intrinsic Multidrug Resistance.* Cancer Research, 2011. **71**(5): p. 1858-1870.
12. Michor, F., et al., *Dynamics of colorectal cancer.* Seminars in Cancer Biology, 2005. **15**(6): p. 484-493.
13. Roepman, P., et al., *Colorectal cancer intrinsic subtypes predict chemotherapy benefit, deficient mismatch repair and epithelial-to-mesenchymal transition.* International Journal of Cancer, 2014. **134**(3): p. 552-562.
14. Alhopuro, P., et al., *Candidate driver genes in microsatellite-unstable colorectal cancer.* International Journal of Cancer, 2012. **130**(7): p. 1558-1566.
15. Barderas, R., et al., *An optimized predictor panel for colorectal cancer diagnosis based on the combination of tumor-associated antigens obtained from protein and phage microarrays.* Journal of Proteomics, 2012. **75**(15): p. 4647-4655.
16. Hacking, S., et al., *MMR Deficiency Defines Distinct Molecular Subtype of Breast Cancer with Histone Proteomic Networks.* International Journal of Molecular Sciences, 2023. **24**(6): p. 5327.
17. Jolien, T., et al., *Chemotherapy, Bevacizumab, and Cetuximab in Metastatic Colorectal Cancer.* New England Journal of Medicine, 2009. **360**(6): p. 563-572.
18. Wang, Y., et al., *Modulation of Gut Microbiota in Pathological States.* Engineering, 2017. **3**(1): p. 83-89.
19. Zhao, L., et al., *The Composition of Colonic Commensal Bacteria According to Anatomical Localization in Colorectal Cancer.* Engineering, 2017. **3**(1): p. 90-97.
20. Zhao, X., et al., *Comprehensive analysis of alfa defensin expression and prognosis in human colorectal cancer.* Frontiers in Oncology, 2023. **12**: p. 974654.
21. Adey, A., et al., *The haplotype-resolved genome and epigenome of the aneuploid HeLa cancer cell line.* Nature, 2013. **500**(7461): p. 207-211.
22. Ahmed, F., et al., *Network-based drug repurposing for HPV-associated cervical cancer.* Comput Struct Biotechnol J, 2023. **21**: p. 5186-5200.
23. Ang, K.K. and E.M. Sturgis, *Human Papillomavirus as a Marker of the Natural History and Response to Therapy of Head and Neck Squamous Cell Carcinoma.* Seminars in Radiation Oncology, 2012. **22**(2): p. 128-142.
24. Edwards, T.G., et al., *DNA Damage Repair Genes Controlling Human Papillomavirus (HPV) Episome Levels under Conditions of Stability and Extreme Instability.* PLoS ONE, 2013. **8**(10): p. e75406.
25. Patel, S. and S. Chiplunkar, *Host immune responses to cervical cancer.* Current Opinion in Obstetrics and Gynecology, 2009. **21**(1): p. 54-59.
26. Sipp, D., I.H. Frazer, and J.E.J. Rasko, *No Vacillation on HPV Vaccination.* Cell, 2018. **172**(6): p. 1163-1167.

27. Toustrup, K., et al., *Gene expression classifier predicts for hypoxic modification of radiotherapy with nimorazole in squamous cell carcinomas of the head and neck.* Radiotherapy and Oncology, 2012. **102**(1): p. 122-129.
28. Aavikko, M., et al., *WNT2 activation through proximal germline deletion predisposes to small intestinal neuroendocrine tumors and intestinal adenocarcinomas.* Hum Mol Genet, 2021. **30**(24): p. 2429-2440.
29. Anwer, S.T., et al., *Synthesis of Silver Nano Particles Using Myricetin and the In-Vitro Assessment of Anti-Colorectal Cancer Activity: In-Silico Integration.* Int J Mol Sci, 2022. **23**(19).
30. Benton, D. and J. Chernoff, *TRIMming away colon cancer: TRIM21-mediated ubiquitination as an activator of the Hippo tumor suppressor pathway.* Cell Chem Biol, 2023. **30**(7): p. 699-701.
31. Brugge, J., M.-C. Hung, and G.B. Mills, *A New Mutational aktivation in the PI3K Pathway.* Cancer Cell, 2007. **12**(2): p. 104-107.
32. Chen, X., et al., *FGF19-mediated ELF4 overexpression promotes colorectal cancer metastasis through transactivating FGFR4 and SRC.* Theranostics, 2023. **13**(4): p. 1401-1418.
33. Chen, X., et al., *Insight into the design of FGFR4 selective inhibitors in cancer therapy: Prospects and challenges.* European Journal of Medicinal Chemistry, 2023. **263**: p. 115947.
34. Diaz, Z., et al., *Next-generation biobanking of metastases to enable multidimensional molecular profiling in personalized medicine.* Modern Pathology, 2013. **26**(11): p. 1413-1424.
35. Drier, Y., M. Sheffer, and E. Domany, *Pathway-based personalized analysis of cancer.* Proceedings of the National Academy of Sciences, 2013. **110**(16): p. 6388-6393.
36. Fiordalisi, J.J., P.J. Keller, and A.D. Cox, *PRL Tyrosine Phosphatases Regulate Rho Family GTPases to Promote Invasion and Motility.* Cancer Research, 2006. **66**(6): p. 3153-3161.
37. Harvey, K.F., X. Zhang, and D.M. Thomas, *The Hippo pathway and human cancer.* Nature Reviews Cancer, 2013. **13**(4): p. 246-257.
38. Kar, G., A. Gursoy, and O. Keskin, *Human cancer protein-protein interaction network: a structural perspective.* PLoS Comput Biol, 2009. **5**(12): p. e1000601.
39. Kress, S., et al., *Expression of hypoxia-inducible genes in tumor cells.* Journal of Cancer Research and Clinical Oncology, 1998. **124**(6): p. 315-320.
40. Liao, Y., et al., *Ubiquitin specific peptidase 11 as a novel therapeutic target for cancer management.* Cell Death Discovery, 2022. **8**(1): p. 292.
41. Berger, M.F., et al., *The genomic complexity of primary human prostate cancer.* Nature, 2011. **470**(7333): p. 214-220.
42. Santiago, J.A. and J.A. Potashkin, *Network-based metaanalysis identifies HNF4A and PTBP1 as longitudinally dynamic biomarkers for Parkinson's disease.* Proceedings of the National Academy of Sciences, 2015. **112**(7): p. 2257-2262.
43. Duan, H., et al., *PARK2 Suppresses Proliferation and Tumorigenicity in Non-small Cell Lung Cancer.* Front Oncol, 2019. **9**: p. 790.
44. Hasson, S.A., et al., *High-content genome-wide RNAi screens identify regulators of parkin upstream of mitophagy.* Nature, 2013. **504**(7479): p. 291-295.
45. Sarraf, S.A., et al., *Landscape of the PARKIN-dependent ubiquitylome in response to mitochondrial depolarization.* Nature, 2013. **496**(7445): p. 372-376.
46. Trempe, J.-F., et al., *Structure of Parkin Reveals Mechanisms for Ubiquitin Ligase Activation.* Science, 2013. **340**(6139): p. 1451-1455.
47. Veeriah, S., et al., *Somatic mutations of the Parkinson's disease-associated gene PARK2 in glioblastoma and other human malignancies.* Nat Genet, 2010. **42**(1): p. 77-82.

48. Zilocchi, M., et al., *Exploring the Impact of PARK2 Mutations on the Total and Mitochondrial Proteome of Human Skin Fibroblasts.* Front Cell Dev Biol, 2020. **8**: p. 423.
49. Tsukamoto, S., et al., *Clinical significance of osteoprotegerin expression in human colorectal cancer.* Clin Cancer Res, 2011. **17**(8): p. 2444-50.
50. Ahmed, S., et al., *A Network-Guided Approach to Discover Phytochemical-Based Anticancer Therapy: Targeting MARK4 for Hepatocellular Carcinoma.* Front Oncol, 2022. **12**: p. 914032.
51. Almowallad, S., R. Jeet, and M. Mobashir, *A systems pharmacology approach for targeted study of potential inflammatory pathways and their genes in atherosclerosis.* Jour. Bas. Sci., 2024. **6**(1): p. 1-12.
52. Bajrai, L.H., et al., *Gene Expression Profiling of Early Acute Febrile Stage of Dengue Infection and Its Comparative Analysis With Streptococcus pneumoniae Infection.* Front Cell Infect Microbiol, 2021. **11**: p. 707905.
53. Bajrai, L.H., et al., *Understanding the role of potential pathways and its components including hypoxia and immune system in case of oral cancer.* Sci Rep, 2021. **11**(1): p. 19576.
54. Choudhry, H., et al., *Study of APOBEC3B focused breast cancer pathways and the clinical relevance.* Jour. Bas. Sci., 2024. **2**(1): p. 1-12.
55. El-Kafrawy, S.A., et al., *Genomic profiling and network-level understanding uncover the potential genes and the pathways in hepatocellular carcinoma.* Front Genet, 2022. **13**: p. 880440.
56. Eldakhakhny, B.M., et al., *In-Silico Study of Immune System Associated Genes in Case of Type-2 Diabetes With Insulin Action and Resistance, and/or Obesity.* Front Endocrinol (Lausanne), 2021. **12**: p. 641888.
57. Helmi, N., D. Alammari, and M. Mobashir, *Role of Potential COVID-19 Immune System Associated Genes and the Potential Pathways Linkage with Type-2 Diabetes.* Comb Chem High Throughput Screen, 2022. **25**(14): p. 2452-2462.
58. Kamal, M.A., et al., *Gene expression profiling and clinical relevance unravel the role hypoxia and immune signaling genes and pathways in breast cancer: Role of hypoxia and immune signaling genes in breast cancer.* Journal of Internal Medicine: Science & Art, 2020. **1**.
59. Khan, B., et al., *Deciphering molecular landscape of breast cancer progression and insights from functional genomics and therapeutic explorations followed by in vitro validation.* Scientific Reports, 2024. **14**(1).
60. Khouja, H.I., et al., *Multi-staged gene expression profiling reveals potential genes and the critical pathways in kidney cancer.* Sci Rep, 2022. **12**(1): p. 7240.
61. Krishnamoorthy, P.K.P., et al., *In-silico study reveals immunological signaling pathways, their genes, and potential herbal drug targets in ovarian cancer.* Informatics in Medicine Unlocked, 2020. **20**: p. 100422.
62. Mobashir, M., et al., *An Approach for Systems-Level Understanding of Prostate Cancer from High-Throughput Data Integration to Pathway Modeling and Simulation.* Cells, 2022. **11**(24).
63. Warsi, M.K., et al., *Comparative Study of Gene Expression Profiling Unravels Functions Associated with Pathogenesis of Dengue Infection.* Curr Pharm Des, 2020. **26**(41): p. 5293-5299.
64. Alam, M.N., M. Almoyad, and F. Huq, *Polyphenols in Colorectal Cancer: Current State of Knowledge including Clinical Trials and Molecular Mechanism of Action.* BioMed Research International, 2018. **2018**: p. 4154185.
65. Alina, V., et al., *Cutaneous metastasis of rectal adenocarcinoma: a case report and literature review.* Pol J Pathol, 2023. **74**(3): p. 211-215.
66. Chang, L., et al., *Systematic profiling of conditional pathway activation identifies context-dependent synthetic lethalities.* Nature Genetics, 2023. **55**(10): p. 1709-1720.
67. Cho, S.-H., et al., *Attractor landscape analysis of colorectal tumorigenesis and its reversion.* BMC Systems Biology, 2016. **10**(1): p. 96.

68. Cioce, M. and V.M. Fazio, *EphA2 and EGFR: Friends in Life, Partners in Crime. Can EphA2 Be a Predictive Biomarker of Response to Anti-EGFR Agents?* Cancers, 2021. **13**(4): p. 700.

69. Hofseth, L.J., et al., *Early-onset colorectal cancer: initial clues and current views.* Nature Reviews Gastroenterology & Hepatology, 2020. **17**(6): p. 352-364.