PLOS ONE

RESEARCH ARTICLE

Identification of genes critical for inducing ulcerative colitis and exploring their tumorigenic potential in human colorectal carcinoma

Ritwik Patra[1], **Amit Kumar Dey**[2], **Suprabhat Mukherjee**[1]*

1 Integrative Biochemistry & Immunology Laboratory, Department of Animal Science, Kazi Nazrul University, Asansol, West Bengal, India, **2** Biomedical Research Centre, Translational Geroproteomics Unit, National Institute on Aging, National Institute of Health (NIH), Baltimore, MD, United States of America

* babaimbc@gmail.com, suprabhat.mukherjee@knu.ac.in

Abstract

Ulcerative colitis (UC) is a chronic relapsing inflammatory bowel disease leading to continuous mucosal inflammation in the rectum extending proximally towards the colon. Chronic and/or recurrent UC is one of the critical predisposing mediators of the oncogenesis of human colorectal carcinoma (CRC). Perturbations of the differential expression of the UC-critical genes exert an intense impact on the neoplastic transformation of the affected tissue(s). Herein, a comprehensive exploration of the UC-critical genes from the transcriptomic profiles of UC patients was conducted to study the differential expression, functional enrichment, genomic alterations, signal transduction pathways, and immune infiltration level encountered by these genes concerning the oncogenesis of CRC. The study reveals that WFDC2, TTLL12, THRA, and EPHB3 play crucial roles as UC-CRC critical genes and are positively correlated with the molecular transformation of UC to CRC. Taken together, these genes can be used as potential biomarkers and therapeutic targets for combating UC-induced human CRC.

Introduction

Inflammatory bowel disease (IBD) is the chronic inflammatory manifestations of the human intestinal mucosa [1]. It is categorized into two main types such as ulcerative colitis (UC) and Crohn's disease (CD). UC is one of the major pathogenic hallmarks of disrupted colon health characterized by sustained mucosal inflammation starting from the mucosal lining of the rectum and approaching the colon proximally [2, 3]. While CD may occur in any part of the GI tract starting from the esophagus to the colon displaying extraintestinal inflammations [4]. Irritable bowel syndrome (IBS) is an added complication of inflamed gut and it usually resulted from gut dysbiosis [5, 6]. Considering the comparative impact of CD and IBS in disrupting the normal gut function, UC is considered as the principal contributor [2]. The pathophysiology of UC includes gut dysbiosis, formation of mucosal lesions, and dysregulated

mucosal immunity with a surge of inflammatory mediators including TNF-α (tumor necrosis factor- α), IL-6 (interleukin-6), and PGE2 (prostaglandin E2) [2, 7, 8]. The proinflammatory cytokines signal the invasion of the immune cells at the inflamed site, accumulation of inflammatory mediators, exaggerated release of reactive oxygen species (ROS), diminution of the colon antioxidant capability, and the loss of intestine mucosal integrity [2, 7, 8]. Northern Europe has been reported for the highest prevalence of UC constituting 505-per-100,000 cases, while 248-per-100,000 for Canada and 214-per-100,000 cases for the USA [8]. However, the cases in the developing countries in the Middle East, Asia, and Latin America are also reported with fewer data than in Europe [8, 9]. The Global Burden of Disease (GBD) study of 2019 for IBD and cancer associated with colon and rectum are shown in **S1 Fig in S1 File**. Interestingly, occurrence of UC among the pediatric population is more extensive than the adult [10]. Patients having extensive UC display long-term persistence and relapsing episodes of inflammation have been correlated with a high risk of colorectal cancer (CRC) [11].

CRC is the third most common cancer of human and it is in the second position for cancer-associated deaths worldwide [12]. Meta-analyses on the prevalence of CRC demonstrated a cumulative probability of 3.7% of CRC among UC patients [13]. Patients having chronic UC for 10, 20, and 30 years were reported for developing CRC with a cumulative probability of 2%, 8%, and 18% respectively [13]. The pathogenesis of CRC is a multifactorial process [14]. It starts from no dysplasia to indefinite dysplasia, followed by low-grade dysplasia that further developed into high-grade dysplasia leading to carcinoma [14, 15]. The oncogenesis process involves various immunological, genetic, and epigenetic changes such as overexpression of oncogenes (e.g. *EGFR*, *KRAS*, *Cyclin D1*), and inactivation of tumor suppressor genes (e.g. *RB1* and *p53*) and mutations (missense, frameshift deletion, truncating mutation) alongside genetic instability [15]. However, the actual mechanism of the neoplastic transformation of UC to CRC is unknown. Several studies have suggested alteration of the expression pattern, epigenetic regulation, mutations and immune-dysfunction in the gene clusters comprising *COL11A1*, *GNG2*, *AGT*, *SAA1*, *ADCY5*, *LPAR1*, *NMU*, *IL8*, *CXCL12*, *GNAI1*, *CCR2*, *SFMBT2*, *LYN*, *PLCB1*, *NPSR1*, *WNT5A*, *CDC25B*, *CD44*, *RIPK2*, and *ASAP1* as influential determinants in the oncogenesis events of CRC [16–18]. Intriguingly, TGF-β, RTK-RAS, Wnt, and TP53 signaling pathways were also reported to influence the transformation of the normal colon tissue to CRC [16–20]. Therefore, exploring the new gene candidates playing a significant role in the occurrence of UC and neoplastic transformation are of major interest. The present study aims to explore the UC-critical genes from the mucosal transcriptomic profiles of pediatric and adult patients having limited to an extensive grade of UC. In addition, pathophysiological significance of these genes in the course of CRC were also deciphered using advanced bioinformatics and oncogenomics approaches.

Material and methods

Study population

In the present study, the gene expression microarray datasets were obtained from the Gene Expression Omnibus (GEO) database available in the National Centre for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/geo/) [21–24]. The GSE87473 dataset is used in this study containing the gene expression profile of the mucosal biopsies from the paediatric and/or adult patients diagnosed with moderately to severely active UC [25]. We selected 67 samples out of 127 total samples to create three groups viz., the control group comprising 21 samples from healthy/normal people, a pediatric group containing 19 samples of extensive UC patients below the age of 18 years, and an adult group (age ranging from 19 to 69 years) of 27 samples from extensive colitis patients. The extensive and limited colitis defines

the severity of UC in the patients showing pancolitis in case of extensive UC and left-sided colitis for limited UC. The criteria for selecting samples were based on the patient's age and the intensity of UC in patients. Adult patients having limited UC were not included in the study.

Data processing and analyses of the differentially expressed genes (DEGs)

The differentially expressed genes (DEGs) amongst the normal, pediatric, and adult samples having UC were explored using the GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r/) [21]. GEO2R is a simple web interface based on the R-packages, useful in comparing the groups of samples to identify and visualize the DEGs [21]. The top 250 genes were obtained using log transformation of data and p-value threshold of >0.05 that are differentially expressed in the three groups viz, normal, pediatric UC, and adult UC. After that, a separate analysis was performed to compare the overexpressed and under-expressed genes among normal vs. UC, normal vs. pediatric UC, normal vs. adult UC, and adult UC vs. pediatric UC using volcano and mean-difference plots to filter the UC-critical genes and further study.

Functional enrichment and PPI

The functional annotation of the 250 DEGs including upregulated and downregulated ones were performed using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) (https://david.ncifcrf.gov/home.jsp) online server [26]. In this study, the top 250 significant genes were used to determine the Gene Ontology (GO) which includes the biological process (BP), molecular function (MF), and the cellular component (CC). The Kyoto Encyclopedia of Genes and Genomes (KEGG) was used for the pathway enrichment analysis on DAVID. Further, the protein-protein interaction (PPI) network including both the physical and the functional networks was determined using the STRING (string-db.org) [27].

Screening of genes critical in the induction of UC

Based on the experiments/assessments described in the earlier section, a total of 24 UC-critical genes were selected based on their influential occurrence in the four major target groups such as normal vs. UC, normal vs. paediatric UC, normal vs. adult UC, and adult UC vs. paediatric UC for further studies. A total of six genes are selected from each group, of which three are selected from overexpressed genes and three from under-expressed genes. We have termed the genes that play crucial role in the induction and pathogenesis of ulcerative colitis as "colitis-critical genes."

Oncogenomics and mutational study of UC-critical genes for human colorectal carcinoma

The oncogenomics study of the aforesaid UC-critical genes was performed using cBioPortal server (https://www.cbioportal.org/) [28]. It is an open-access portal for various oncogenomics and mutational studies. The TCGA PanCancer atlas study of colon adenocarcinoma, having 594 samples was used to analyse and visualize the oncoprint of the selected genes, determination of cancer type, plots for genomic alterations, mutation, survival, and also for exploring the alterations in molecular pathways. Furthermore, the mRNA expression data (log RNAseq data by RSEM) and protein level data (MS data by CPTAC) of the selected genes are collected for both altered and unaltered samples. They were subjected to the generation of a clustered heatmap using the ClusterViz server.

Gene expression, promoter methylation, proteomics study, and survival assay

Oncomine (https://www.oncomine.org/resource/main.html), a popular cancer microarray database, was selected for studying the transcriptional expression of the UC-critical genes using the TCGA CRC dataset and also compared with the normal tissue and different cancer subtypes [29]. Taking a clue from the differential expression of the critical genes for UC at mRNA level, the functional expression of these genes was explored by employing UALCAN server (http://ualcan.path.uab.edu/index.html) [30]. We analyzed the expression profile and promoter methylation of the UC-critical genes for colon adenocarcinoma involving various clinicopathological parameters to identify biomarkers and validate the gene of interest by studying the cancer OMICS data (TCGA, MET500, and CPTAC) The protein expression profile from the CPTAC dataset of colon cancer was also explored.

In order to predict the expression of the genes in cancer and normal tissue, we examined the antibody-based protein profiling using immunohistochemistry of thin tissue sections obtained from a normal subject and colorectal cancer patients using the Human Protein Atlas (https://www.proteinatlas.org) available online [31].

In addition, the survival assay of the genes of interest was determined by constructing a Kaplan Meier (KM) plot to assess the overall survival, disease-free survival, and disease-specific survival using the UCSC-Xena server (https://xena.ucsc.edu/).

Immune cell infiltration analysis

TIMER 2.0 (http://timer.cistrome.org/) is an open-access resource for studying immune behaviour and immune infiltration analyses of different cancers [32]. Expression profiles of immune infiltration-related parameters were studied by TIMER, CIBERSORT, quanTIseq, xCell, MCP-counter, and EPIC algorithms. The TCGA COAD mRNA expression data of UC-critical genes were obtained from the cBioPortal and subsequently utilized for the analysis of infiltration level and estimation.

Network enrichment analysis of UC-critical genes

The network enrichment of the expression profiles, KEGG prediction and GO functions of the UC-critical genes was explored by employing the NetworkAnalyst server (https://www.networkanalyst.ca/) [33]. Furthermore, PPI network (using STRING database), tissue-specific PPI (using DifferentialNet database), gene-miRNA interaction (in miRTarBase V8.0), protein-drug interaction (collected from DrugBank), and protein-chemicals interactions (from Comparative Toxicogenomics Database (CTD) were also studied.

Results

Screening and characterization of DEGs that play critical roles in inducing UC

The profile for the study of UC in the adult and paediatric patient population and the healthy samples obtained from the dataset GSE87473 were found to be statistically significant for cross-comparing. The age of the paediatric patients ranges from 6–17 years with a median value of 15 years, and the adult colitis patients were belonging to the age group of 19–69 years with a median age of 32 (**S1 Table in S1 File**).

The gene expression profile of each mucosal biopsy sample from the healthy normal, pediatric colitis patients, and adult colitis patients were found to be mean-centered and were cross comparable for the study (**Fig 1A**). The expression density curve shows the normalized
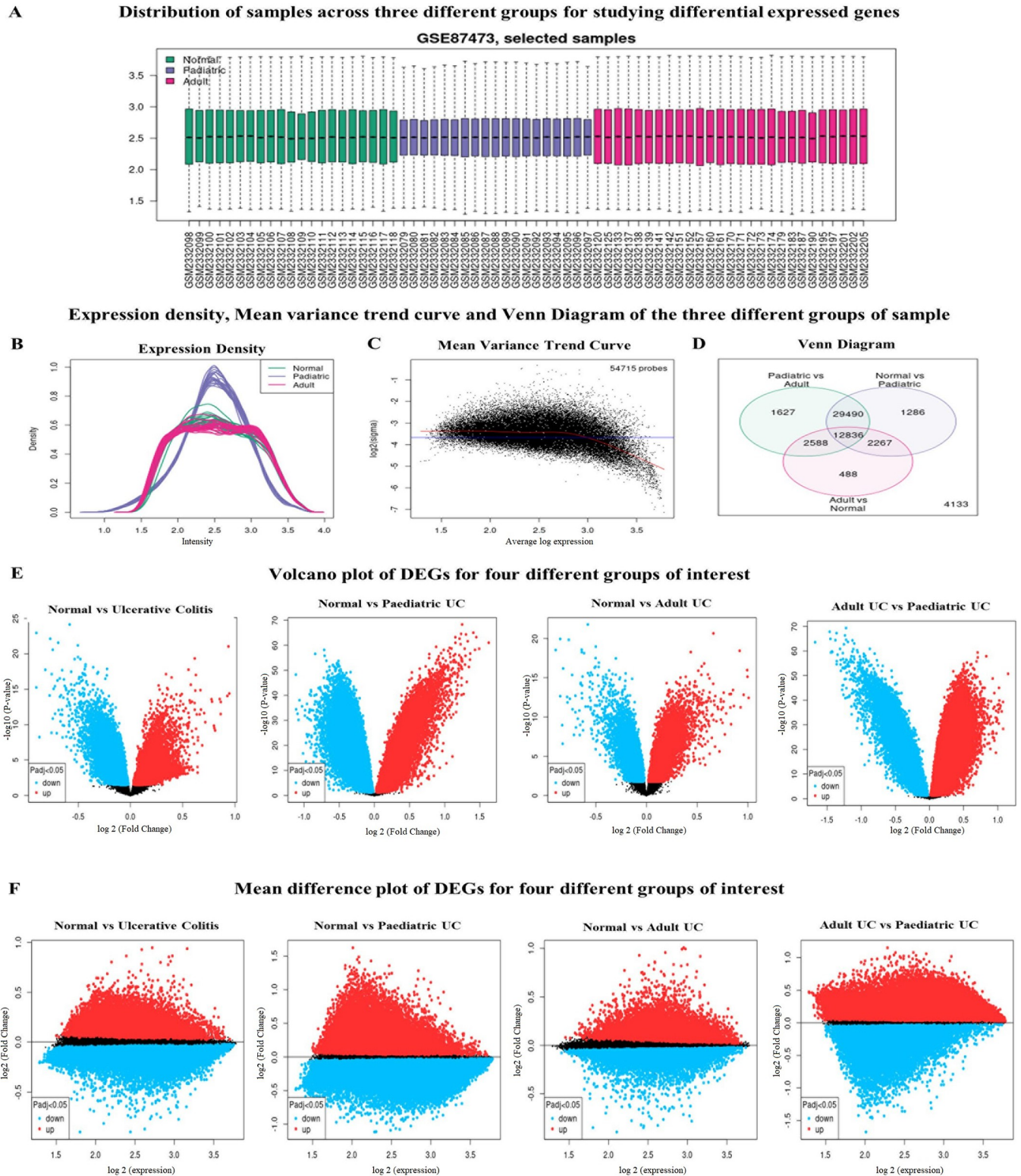
Fig 1. Identification of the differentially expressed genes in the transcriptomic profile of ulcerative colitis (UC) tissue from patients of different age groups. A. Boxplots representing the samples of dataset GSE87473, where green, blue, and red colours respectively depict the three groups of samples viz, normal, pediatric UC, and adult UC. **B.** Comparative expression density plots of the samples after log transformation of the data. **C.** Mean-variance trend plot projecting the variations in DEGs among the different samples. **D.** Venn diagram displaying the significant genes involved in inducing UC among pediatric vs adult, normal vs pediatric, and adult vs normal groups. **E.** Volcano plot and **F.** Mean-difference plots reflect the upregulated and downregulated genes influencing the induction of UC in the colon tissue of the groups of interest.

https://doi.org/10.1371/journal.pone.0289064.g001

distribution of samples over the three groups (**Fig 1B**), followed by a strong recommendation of the mean-variance trend for differential gene expression analysis (**Fig 1C**). The mean-variance trend represents the relationship between the gene expression and the mean-variance after fitting a linear model, where each point indicates the genes, and the red line is a mean-variance trend. Next, we explored the Venn diagram to determine the significant overlapping genes amongst the different sample groups included in this study (**Fig 1D**). While comparing the gene expression profiles for normal vs. pediatric UC, 1286 genes are common in this group and 488 genes in normal vs. adult UC. In the case of pediatric vs. adult analysis, the significant involvement of 1627 common genes were found for both the patient groups. Interestingly, 12836 genes were found to be overlapped among all the three study groups viz, normal vs. pediatric UC, adult UC vs normal, and pediatric UC vs adult UC.

From these 12836 genes, we have screened out the top 250 genes based on the p-value of the relative abundance of the differentially expressed mRNAs' most significant/influential genes amongst these three groups (**S2 Table in S1 File**). A clustered heatmap showing the expression of the top 250 genes among the normal, pediatric UC, and adult UC samples are shown in **S2 Fig in S1 File**. Next, to determine the alteration in expression levels of the DEGs, we used GEOquery and limma function of R packages to generate the volcano plots and mean-difference plots for normal vs UC, normal vs. pediatric UC, normal vs. adult UC, and pediatric UC vs. adult UC samples. The volcano plot depicts the upregulated (red) and downregulated (blue) genes based on the statistical significance (-log10 p-value) versus magnitude of change (log2 fold change) at an adjusted p-value cut-off of 0.05 (**Fig 1E**). Similarly, the mean-difference plot reveals the upregulated (red) and downregulated (blue) genes for the three groups based on the expression versus fold change (**Fig 1F**). Collectively, the aforesaid results indicated the significant roles of these genes in the occurrence of UC in humans.

Furthermore, the GO study for BP (Biological Pathways) shows the involvement of 234 genes out of the total 250 DEGs (95.1%), attributing higher DNA templated transcription, mRNA processing, peptidyl-serine phosphorylation, and RNA splicing as depicted in **S3 Table in S1 File**. The GO: MF analyses the broad-level molecular functions performed by the gene products of these DEGs. We have found the involvement of 231 genes (93.9%) in the GO: MF function, of which the most critical and enriched functions are protein binding, poly(A) RNA binding, protein serine/threonine kinase activity, and transcriptional regulatory region DNA binding. Similarly, the GO: CC reveals the cellular localization of the gene expression of these DEGs. In this, we found 129 genes (52.4%) expressed on the nucleus, and 47 genes (19.1%) expressed over the membrane. The KEGG pathway enrichment analyses the molecular interaction of these DEGs with the biologically important pathways of humans. We have found the involvement of 115 genes (46.7%), of which the connection of these genes is common in platelet activation, pathways in cancer, proteoglycans in cancer, and lysine degradation.

The PPI through STRING at a high confidence score of 0.7 (**S3 Fig in S1 File**) demonstrates 241 nodes and 153 edges with an average node degree of 1.27 and PPI enrichment p-value of 1.58e-06. The expected number of edges seems to be 112, with an average local clustering coefficient of 0.28.

The UC-critical DEGs were identified from the four major target groups that are normal vs UC patients containing samples from both paediatric and adult, normal vs. paediatric UC, normal vs. adult UC, and adult UC vs. paediatric UC groups each containing six significant DEGs as given in **Table 1**. The selection of the 24 UC-critical genes is based on selecting the three upregulated and downregulated genes from each group respectively using the volcano and mean-difference plots.

Table 1. Screening of top 24 genes from the expression profiles of 250 DEGs that play critical roles in the induction of UC.

Sample Type	Gene Symbol	Gene Name	Expression status	Fold Change (Log 2)	p-value (-Log 10)
Normal vs Ulcerative colitis	HNRNPK	Heterogeneous nuclear ribonucleoprotein K	Downregulated	-0.146	3.165
	TDG	Thymine DNA glycosylase		-0.211	3.175
	PMPCB	Peptidase, mitochondrial processing beta subunit		-0.229	4.373
	MAPK1	Mitogen-activated protein kinase 1	Upregulated	0.204	4.315
	ZNF655	Zinc finger protein 655		0.495	3.546
	CCL5	C-C motif chemokine ligand 5		0.296	6.322
Normal vs Paediatric UC	HMGCS2	3-hydroxy-3-methylglutaryl-CoA synthase 2	Downregulated	-0.868	13.426
	SLC51A	Solute carrier family 51 alpha subunit		-0.905	19.075
	CHP2	Calcineurin like EF-hand protein 2		-0.716	15.116
	ERGIC1	Endoplasmic reticulum-golgi intermediate compartment 1	Upregulated	0.346	19.739
	THRA	Thyroid hormone receptor, alpha		0.136	8.724
	EPHB3	EPH receptor B3		0.152	5.177
Normal vs Adult UC	PTPN21	Protein tyrosine phosphatase, non-receptor type 21	Downregulated	-0.199	9.993
	DPP10	Dipeptidyl peptidase like 10		-0.639	18.581
	PCK1	Phosphoenolpyruvate carboxykinase 1		-0.715	7.72
	MMP3	Matrix metallopeptidase 3	Upregulated	0.992	15.979
	DUOX2	Dual oxidase 2		0.918	18.440
	HSPA6	Heat shock protein family A (Hsp70) member 6		0.196	5.389
Adult UC vs Paediatric UC	PATL1	PAT1 homolog 1, processing body mRNA decay factor	Downregulated	-1.221	64.486
	KMT2C	Lysine methyltransferase 2C		-1.103	63.024
	MUC4	Mucin 4, cell surface associated		-0.874	60.869
	SCARB1	Scavenger receptor class B member 1	Upregulated	0.111	5.142
	TTLL12	Tubulin tyrosine ligase like 12		0.281	22.481
	WFDC2	WAP four-disulfide core domain 2		0.134	1.783

Genes playing critical role in developing UC also inducing the oncogenesis of CRC in human

UC is a degenerative inflammatory change in human gut mucosa and the positive correlation between UC and the occurrence of CRC has been enumerated in several literatures [11, 14, 34, 35]. Upon successful identification of the 24 UC-critical genes, we studied the oncogenic relevance of these genes in the context of neoplastic transformation of the affected tissue to CRC.

Structural and/or functional alteration in the UC/cancer critical genes is known to be the major cause behind the transcriptional abnormalities as well as dysregulation of vital signaling pathways that signal the induce/promote tumorigenesis leading to the transformation of UC to CRC. Herein, the oncogenomics of UC-critical genes revealed genetic alteration in 429 samples out of a total study sample of 594 (72%). The expression of genes was analysed both at mRNA and protein levels to generate the clustered heatmaps within an expression abundance scale between +3 to -3 with a mean-centered of 0 (**Fig 2A and 2B**).

The UC-critical genes from our study infer an alteration frequency of 85.25% in the mucinous adenocarcinoma of the colon and rectum following 69.84% in colon adenocarcinoma and 69.68% for rectal adenocarcinoma (**Fig 2C**). These CRC subtypes mainly include multiple alterations, mutations, deep deletions, amplifications, structural variants, low and/or high mRNA, and protein expressions (**S4 Table in S1 File**). The oncoprint (**Fig 2D**) is the concise graphical summary of alterations associated with each UC-critical gene. We found the highest frequency of alterations among Mucin 4 (*MUC4*), Phosphoenolpyruvate carboxykinase 1
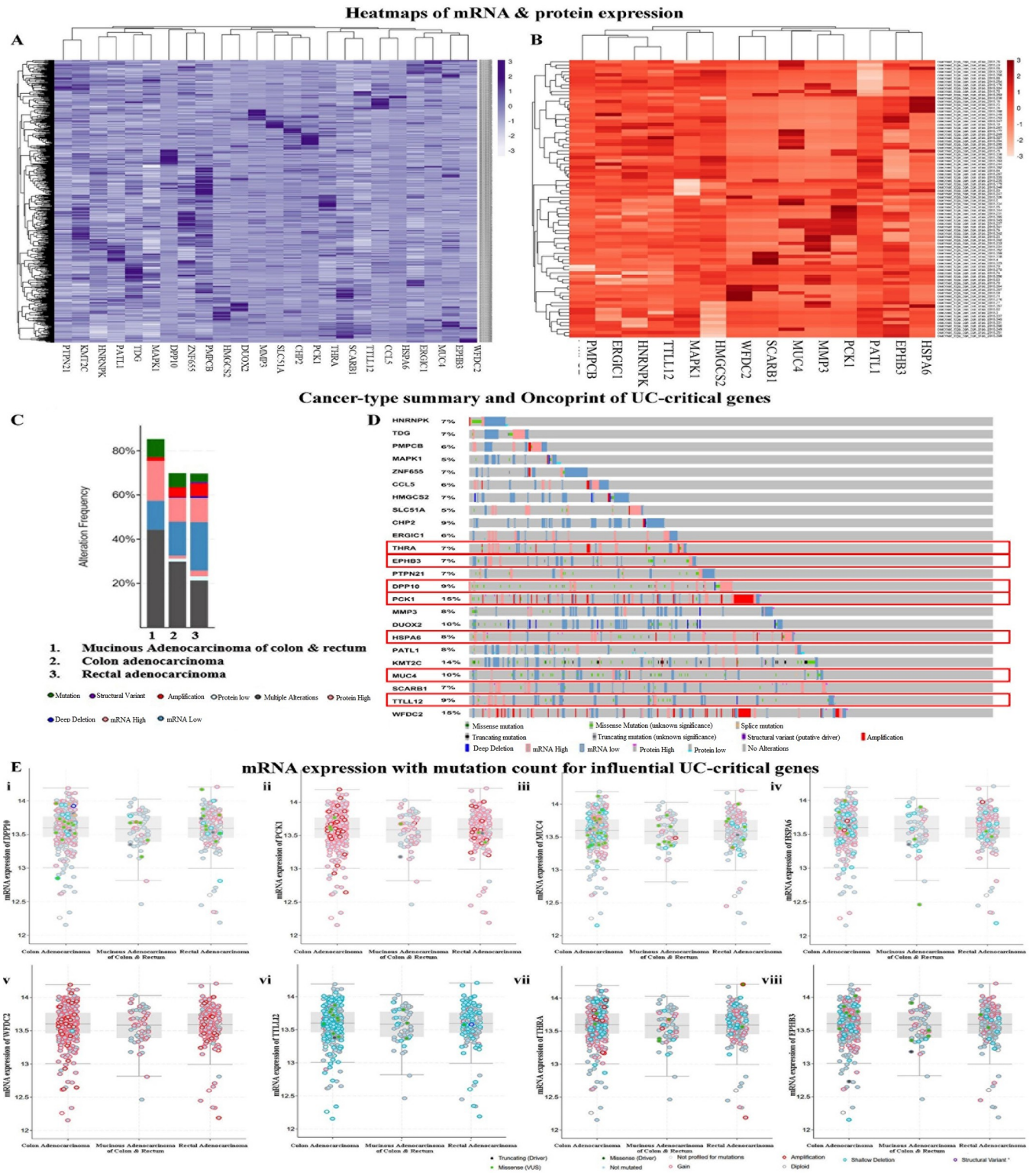
Fig 2. Determination of the significance of the UC-critical genes in the neoplastic transformation of UC to CRC in humans. Heatmaps illustrate the expression of UC-critical genes at **A.** Transcriptional (mRNA) level, **B.** Translational (protein) level. **C.** Types of alterations associated with the different colorectal cancer subtypes resulted from the altered expression of the 24 UC-critical genes. **D.** Oncoprint represents the different molecular alterations in each UC-critical gene concerning the development of CRC. Eventually, 8 genes out of 24 were found to have a higher impact on UC to CRC transition. **E.** Mutational counts in mRNA expression of **i.** DPP10, **ii.** PCK1, **iii.** MUC4, **iv.** HSPA6, **v.** WFDC2, **vi.** TTLL12, **vii.** THRA, and **viii.** EPHB3.

https://doi.org/10.1371/journal.pone.0289064.g002

(*PCK1*), Dipeptidyl peptidase like 10 (*DPP10*), Tubulin tyrosine ligase like 12 (*TTLL12*), Thyroid hormone receptor, alpha (*THRA*), WAP four-disulfide core domain 2 (*WFDC2*), Erythropoietin-Producing Hepatoma (EPH) receptor B3 (*EPHB3*), and Heat shock protein family A (*Hsp70*) member 6 (*HSPA6*). The plot of mRNA expression of these genes for the mutation count shows a greater value in the colon and rectal adenocarcinoma. Herein, *PCK1* and *WFDC2* have more frequent amplifications, while *MUC4*, *TTLL12*, *THRA*, *EPHB3*, *DPP10*, and *HSAP6* exhibit greater shallow deletions and missense mutations (**Fig 2E**). These genes could be interpreted as influential mediators of the oncogenesis of CRC.

Mutations in the genes critical for UC/UC-CRC and prediction of the relative influence of these mutations on the survivability of the affected subjects

The Wnt signaling pathway is a vital mediator of epithelial tissue repair and homeostasis, which plays an intrinsic role in UC and its alteration resulted in hyperactivation of the signaling pathway leading to carcinoma [36]. Previous studies reported that the Wnt signaling pathway is altered in over 92% of CRC cases, with the highest mutation in *APC* [37]. In support of the previous studies, our result indicates the alteration of *APC* (68%) induced by UC-critical genes (**S4A Fig in S1 File**). Similarly, for the RTK-RAS pathway, our data reveal an alteration of *KRAS* (41.1%), *NRAS* (12.1%), and *BRAF* by 13.3% (**S4B Fig in S1 File**). Previous studies have shown that *KRAS* and *NRAS* are mutated in 40% of cases of CRC, resulting in conformational changes inhibiting the GTPase activity of RAS-GAP [38]. The mutation in *BRAF* increases the kinase activity causing over-activation of *RAS* and amplification of the *MAPK* signaling pathway leading to the generation of CRC [38]. Similarly, in colon tissue, the TGF-β signaling cascade alleviates epithelial cell proliferation and induces apoptosis by binding ligands to the *TGFBR2* receptors, followed by phosphorylation and downstream activation of SMAD proteins [39]. It is reported that mutations in SMAD protein, especially SMAD4, have a 30–40% chance of developing CRC [39]. Herein, the UC-critical genes altered the SMAD proteins (*SMAD4* by 18.2%, *SMAD2*, and *SMAD3* by 9.1% and 10.3%, respectively) and TGFBR2 (8.4%) and might elevate the cellular proliferation and transformation to CRC (**S4C Fig in S1 File**). Lastly, our study on alteration of the p53 pathway signifies alteration frequency of 55.7%, which infers loss of function of signaling cascade resulting in disrupting the cell cycle arrest, apoptosis, metastasis and promoting the progression of CRC (**S4D Fig in S1 File**).

We have examined the survival assay and the KM plots describing the effect of UC-critical genes on the survivability of the altered and unaltered groups. The overall survival of the altered group is lower than that of the unaltered group over time, with the least survival of 40% (**S5A Fig in S1 File**). The disease-specific survival is also lower for the altered group, with a minimum survival rate of 70% (**S5B Fig in S1 File**). However, in the case of disease-free survival, the unaltered group shows a significantly lower survival than the altered group (**S5C Fig in S1 File**). It clearly indicates that the alteration in UC-critical genes lowers the survival in CRC and is supported by the progression-free survival plot (**S5D Fig in S1 File**).

Expression of WFDC2, TTLL12, THRA, and EPHB3 genes and their epigenetic regulation preferentially promote the transformation of UC to CRC

We found highly significant transcriptional expression of *WFDC2*, *TTLL12*, *THRA*, and *EPHB3* genes in CRC patients (237 samples). Previously, we noted that these genes exert significant impact over the genomic alterations and mutational count in CRC initiated from UC. Thus, *WFDC2*, *TTLL12*, *THRA*, and *EPHB3* were termed as potential UC-CRC critical genes

and were selected for further analysis. Low level of *WFDC2*, *TTLL12*, and *THRA* mRNAs and significantly high level of *EPHB3* mRNA were detected in colon adenocarcinoma than normal tissue (**Fig 3A**). The expression profile of these UC-CRC critical genes for different subtypes of CRC showed overexpression of *WFDC2* and *EPHB3* across rectal mucinous carcinoma (**Fig 3Bi and 3Biv**) but decreased significantly for the mRNA expression of *TTLL12* and *THRA* (**Fig 3Bii and 3Biii**).

Next, we studied the functional expression of the four UC-CRC critical genes concerning various clinicopathological parameters. The box-plot of the transcriptional expression profile of these genes showed a broad range of expression patterns in the different cancer stages than the normal samples (**Fig 3C**). However, the expression of *WFDC2* was found to be downregulated in the cancer stages than that of normal tissues (**Fig 3Ci**). On the other hand, the expression of *TTLL12*, *THRA*, and *EPHB3* was found to be upregulated in the cancerous tissue than normal. But the expression profiles revealed a gradual decrease in the level of mRNA and/or protein with the progression of cancer towards the advanced stages (**Fig 3Cii–3Civ**). The epigenetic regulation of the genes facilitated by promoter methylation, histone modification, and alteration in the miRNA regulation results in gene silencing and plays a vital role in the pathogenesis of CRC [40]. Herein, the box-plot of promoter methylation depicts the level of DNA methylation ranging from 0 (unmethylated) to 1 (fully methylated), where the cut-off range for hypermethylation and hypomethylation is 0.7–0.5 and 0.3–0.25 respectively (**Fig 3D**). With progressing cancer stages, WFDC2, TTLL12, and EPHB3 exhibit hypomethylation with the lowering of the level of promoter methylation. However, hypermethylation and elevation of THRA mRNA were recorded for advanced stages of CRC (**Fig 3D**).

The translational expression profiles of CRC-critical genes are represented by the box-plots with a z-value signifying the standard deviation from the median in the CRC samples (**Fig 3E**). The protein expression profile of WFDC2 suggested a significant reduction of the expression of this gene at the protein level in the primary tumor than that of the healthy normal tissue (**Fig 3Ei**). In contrast, an increased expression of TTLL12 and EPHB3 at the translational level was noted in the transformed tissue (**Fig 3Eii and 3Eiii**) and the same was also supported by the translational expression of *WFDC2*, *TTLL12*, *THRA*, and *EPHB3* genes through immunohistochemistry (IHC) data (**Fig 4A**).

KM plots reveal that higher expression of WFDC2 leads to a lower survival probability with a minimum survival rate of >50% for disease-specific and disease-free survival (**Fig 4B**). TTLL12 signifies poor survival at the lower expression in overall, disease-specific, and disease-free survival (**Fig 4C**). At a level of 50% survivability, higher expression of THRA indicates lower survival for both overall and disease-specific survival (**Fig 4D**). However, EPHB3 shows a poor prognosis at the lower expression for overall as well as disease-specific survival (**Fig 4E**). Collectively, lower expression of UC-CRC critical genes was associated with poor prognosis of CRC.

WFDC2, TTLL12, THRA, and EPHB3 genes promote infiltration of innate and adaptive immune cells around the transformed colon tissue

The relative level of infiltration and proportion of different infiltrating immune cells in the various individual samples are presented in **Fig 5**. The infiltration level of neutrophils was found highest in the samples following the abundance order neutrophil> NK-cells> macrophages/monocytes> CD8+ T cells (**Fig 5A**). Next, we compared the expression of the UC-CRC critical gene expression and immune cell assembly. The expression of WFDC2 was found to be positively correlated with the infiltration of CD4+ T cells and B cells while negatively correlated with the infiltration of neutrophils (**Fig 5B**). Infiltration of CD4+ T cells, macrophages, neutrophils, and NK cells all were positively correlated with the expression of TTLL12 and THRA
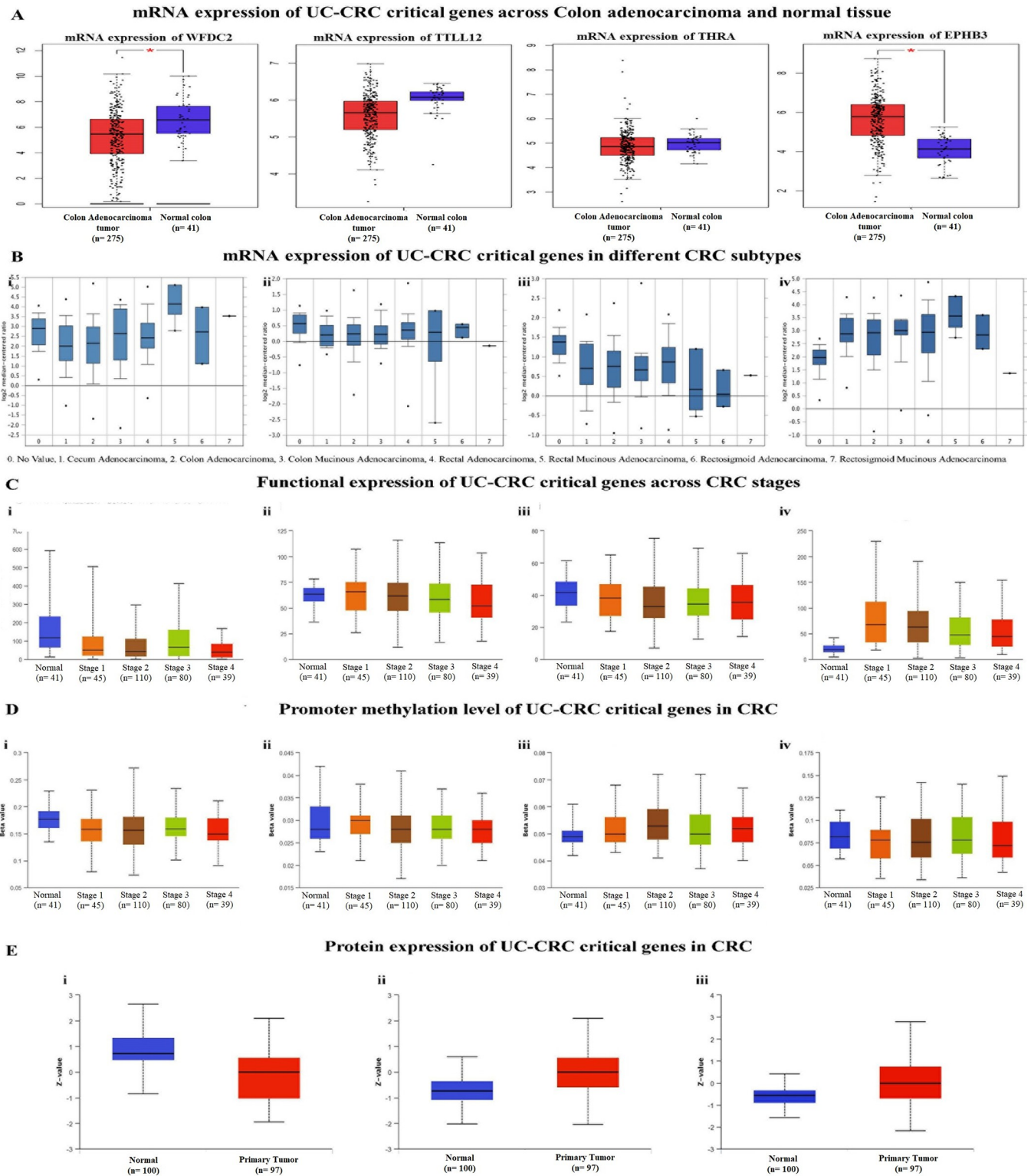
Fig 3. Transcriptional and translational expression profiles of UC-CRC critical genes. A. Box-plot depicts the mRNA level expression of WFDC2, TTLL12, THRA, and EPHB3 in normal and colon adenocarcinoma tissues. **B.** mRNA expression for different CRC subtypes for **i.** WFDC2, **ii.** TTLL12, **iii.** THRA and **iv.** EPHB3. **C.** Functional transcriptional expression of **i.** WFDC2, **ii.** TTLL12, **iii.** THRA, and **iv.** EPHB3 in different CRC stages. **D.** Level of promoter methylation over different cancer stages of **i.** WFDC2, **ii.** TTLL12, **iii.** THRA, and **iv.** EPHB3. **E.** Translational expression of **i.** WFDC2, **ii.** TTLL12, and **iii.** EPHB3 in normal and cancerous colon tissue.
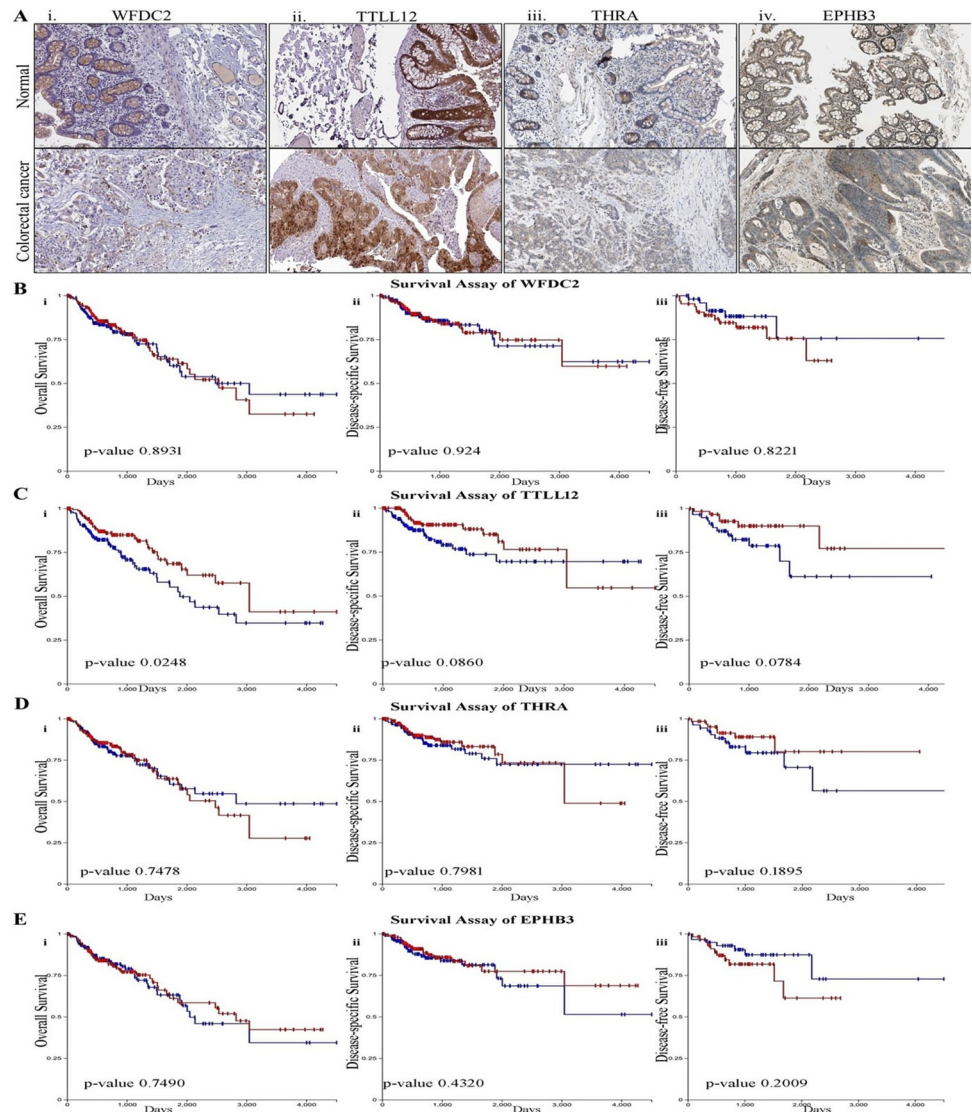
Fig 4. Determining the impact of UC-CRC critical genes on tissue level expression and survivability. A. Immunohistochemistry of normal and colorectal cancer tissue depicting the tissue level expression of **i.** WFDC2, **ii.** TTLL12, **iii.** THRA and **iv.** EPHB3. KM plots depict the overall survival, disease-specific survival, and disease-free survival of **B.** WFDC2, **C.** TTLL12, **D.** THRA, **E.** EPHB3 mRNA. Red and blue colours respectively indicate the higher and lower expression.

https://doi.org/10.1371/journal.pone.0289064.g004

(**Fig 5C and 5D**). Interestingly, EPHB3 showed a positive correlation with the infiltration of CD4+ T cells and macrophages, while a negative correlation with neutrophils was also noted (**Fig 5E**).

Interactome and signaling crosstalk of WFDC2, TTLL12, THRA, and EPHB3 facilitate proposition of new therapeutic targets for treating UC/ UC-CRC

The KEGG pathways enrichment network demonstrates the involvement of pathways related to the Toll-like receptor (TLR) signaling pathway, TNF signaling pathway, chemokine signaling, and prion diseases (**Fig 6Ai**). The Reactome networks show the influence of phospho-
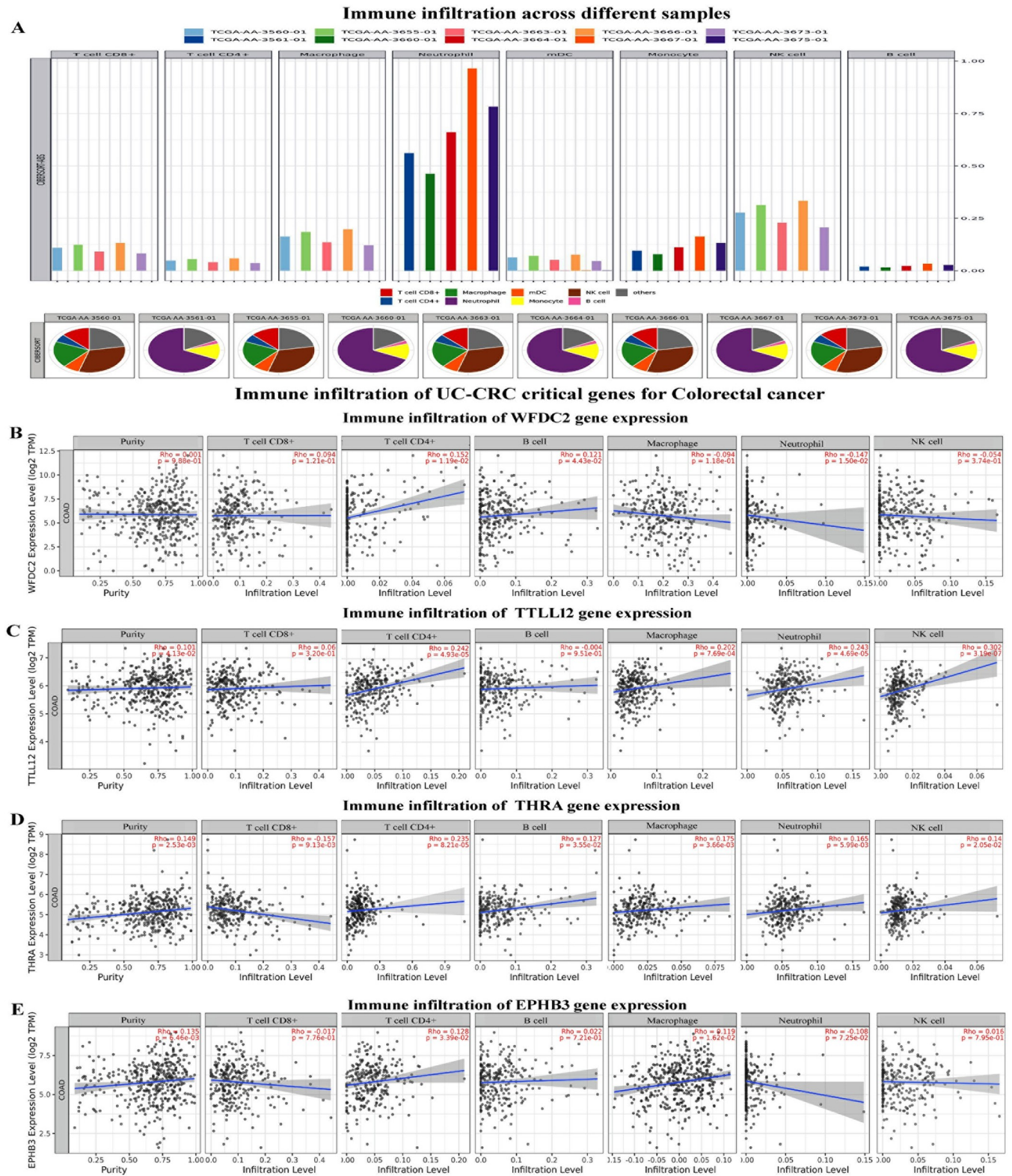
Fig 5. Prediction of the immune estimation and immune infiltration in human CRC. A. Level of infiltration and relative abundance of different immune cells in the CRC tissue samples. Correlation between the translational expression of **B.** WFDC2, **C.** TTLL12, **D.** THRA, and **E.** EPHB3 and the level of immune infiltration in colon adenocarcinoma.

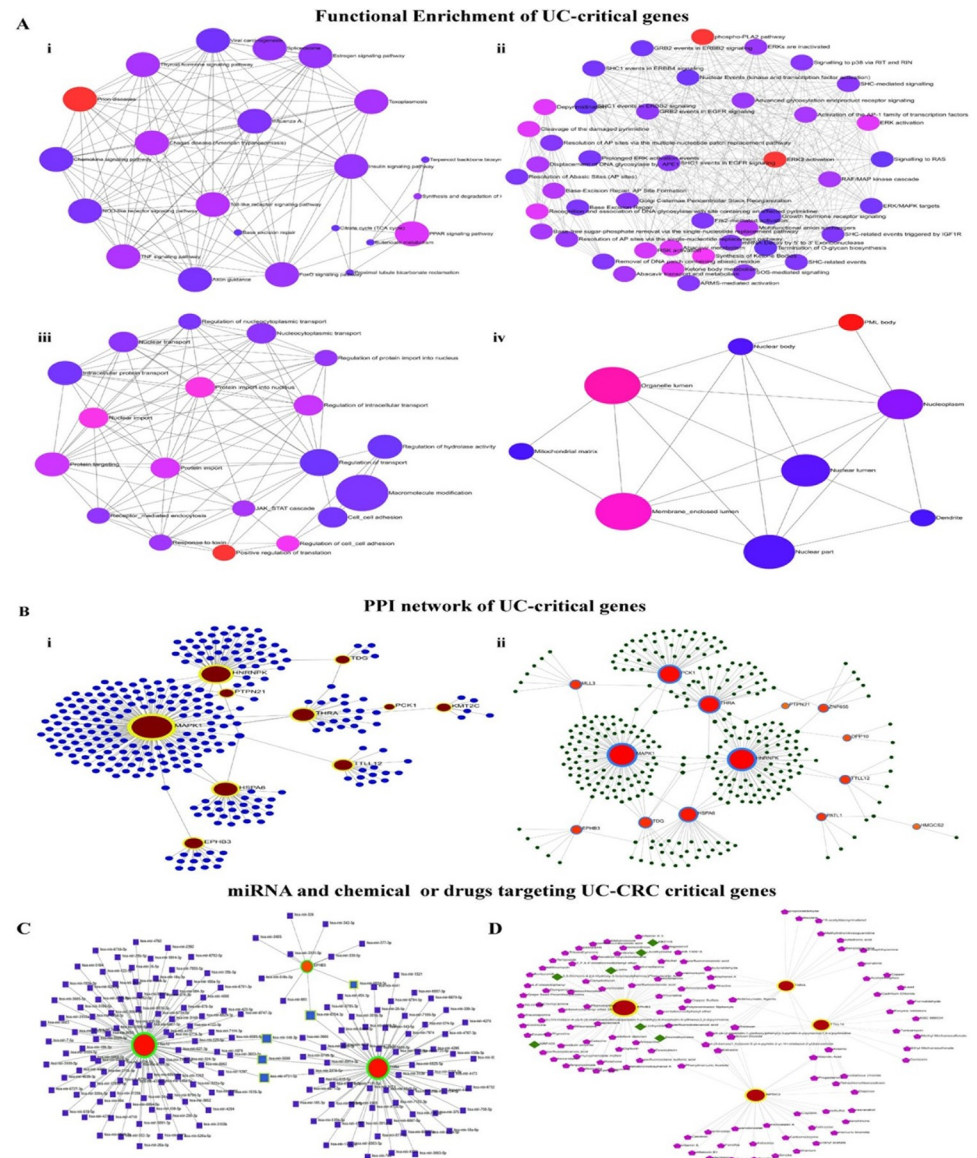https://doi.org/10.1371/journal.pone.0289064.g005

Fig 6. Network enrichment analyses of UC and UC-CRC-critical genes. A. Global enrichment network for UC-critical genes. **i.** KEGG pathways depicting the possible signaling cascades, **ii.** Reactome pathways represent the interaction and signaling-crosstalk amongst the signaling molecules. Analysis of Gene Ontology in **iii.** biological processes and **iv.** cellular component for predicting the types of biological events mediated by the genes of interest. **B.** Protein-protein interactions (PPI) network of UC-critical gene products depicting **i.** Generic PPI networks and **ii.** sigmoid colon tissue-specific PPI networks demonstrating the regulating microenvironment. Exploring the molecular regulators of UC-CRC critical genes. **C.** Analysis of gene-miRNA interaction network. **D.** Protein-drug and protein-chemical compounds interaction union network.

https://doi.org/10.1371/journal.pone.0289064.g006

PAL2 pathways and ERK2 activation with the UC-critical genes (**Fig 6Aii**). On the other hand, GO: BP shows significance over positive regulation of translation, cell-cell adhesion and migration, cell recognition, and protein transport (**Fig 6Aiii**). The cellular localization of the UC-critical gene expression was found primarily in the organelle and membrane-enclosed lumen (**Fig 6Aiv**).

The PPI network for the UC-critical genes obtained through the STRING interactome shows 10 seed genes having 292 nodes and 304 edges (**Fig 6Bi**). However, the tissue-specific

PPI network for colon sigmoid tissue collected from the DifferntialNet database showed the involvement of 14 significant genes connected by 310 nodes and 333 edges (**Fig 6Bii**).

We further extended our study to assess the gene-miRNA interaction and network association of protein-drug and/or chemical compounds for UC-CRC critical genes. The gene-miRNA interaction studied through miRTarBase database indicates 3 associated genes having 180 nodes and 182 edges of which hsa-mir-4731-5p, hsa-mir-5698, hsa-mir-6764-3p, hsa-mir-6824-3p, and hsa-mir-140-3p shows the highest betweenness. The term betweenness measures the number of shortest paths going through the nodes. Nodes that have the higher value of betweenness acts as important bottleneck in any network. They were designated as possible regulatory miRNA regulating the UC-CRC critical genes (**Fig 6C**). Moreover, the network union between the protein-drug interactions and protein-chemical compounds interactions displayed 126 nodes and 148 edges of which levothyroxine, liotrix, dextrothyroxine, {3,5-Dichloro-4-[4-Hydroxy-3-(Propan-2-Yl) phenoxy] phenyl} acetic Acid, NRP409, and KB2115 are actively interacting drugs with EPHB3. The chemical compounds belong to the various class of phenols, organophosphate and chlorides, carboxylic groups, and heavy metals (**Fig 6D**).

Discussion

Gut dysbiosis following disruption of immune-homeostasis resulting in the damages in the mucosal lining i.e. UC, induction of epithelial proliferation and dysplasia are the major contributors in inducing CRC [34, 41]. In this context, the present in-silico study aims to dissect out the mechanistic insights of the transformation of UC to CRC, with a particular emphasis on the identity of the crucial genes involved in this process. As shown in the available literatures, multiple genes have been found to play a pivotal role in the pathogenesis of UC and subsequent neoplastic transformation into CRC [18, 20]. For example, *COL11A1*, *GNG2*, *AGT*, *SAA1*, *ADCY5*, *LPAR1*, *NMU*, *IL8*, *CXCL12*, *GNAI1*, and *CCR2* were hub genes and overexpressed in CRC while methylation of *KCNJ12*, *VAV3-AS1*, and *EVC* are associated with stratification of colon cancer stages [16–18]. However, most of the studies are on the pathogenesis of either UC or CRC and therefore our understanding of the molecular linkage between UC-to-CRC is still under the shed. In this regard, we adopted modern omics approaches for identifying the new gene clusters associated with UC and UC to CRC progression. Our study started with an initial aim to explore the differential expression of UC-critical genes from microarray data available in GEO database on transcriptomic profiles of the mucosal biopsy of colon tissue from the different age groups of patients including paediatric and adult. We have detected differential expression of 250 in the three groups of study population viz, normal, paediatric UC, and adult UC (**Fig 1**, **S2 Table** in **S1 File**). Next, we analysed the GO functions for deciphering the molecular influence of these gene products which revealed that transcriptional regulations, mRNA splicing, and protein serine/threonine kinase activity mediated by these DEGs alongside perturbation in Wnt signaling, PI3K-Akt signaling pathway, and p53 signaling pathways most likely to contribute in the pathogenesis of UC. From the expression profiles of all the 250 genes, a total of 24 genes were eventually selected based on their relative expression profiles (**Table 1**, **Fig 1**). In fact, 12 upregulated genes and the same number of downregulated genes having an influential connection over the occurrence of UC in the selected study population (**Table 1**). These 24 genes were primarily considered as critical genes for the occurrence of UC and were investigated for their impact on transformation of UC to CRC.

We observed that alteration in the pattern of transcriptional and translational expression of the 24 genes do exhibit a strong influence in promoting the transformation of UC to different subtypes of CRC (**Fig 2**). Intriguingly, these genes were found manipulating the vital signaling

pathways associated with the development of CRC. Herein, TLR- and TNF-α signaling pathways were inferred as the major signaling pathways triggered by the genes and induction of inflammatory responses through TLR/TNFR activation could be linked with the tumorigenesis process. The involvement of TLR- and TNF-α signaling pathways were shown to be involved in inducing initial level of inflammation in the oncogenesis events of CRC especially in nodal metastasis [42–44] and these inferences corroborate with our observations. To better understand the direct alliance between the immune cells and transformation of UC to CRC, we have studied the possible level of the infiltrating immune cells in and around the transformed tissue. Neutrophils and NK cells were found as the major infiltrating leukocytes toward/around the tumor. Neutrophils are known to promote ROS generation and trigger leukocytes to generate anti-inflammatory responses, altering cell death and apoptosis, thus affecting the inflammatory and tissue-remodelling responses [45]. In contrast, the NK cells possess a cytolytic function and downregulate the production of IFN-γ and PD-1 to inhibit gut carcinogenesis [46]. However, clinical correlation studies need further extensive efforts to establish the beneficial and detrimental effects of the immune cell assembly around the cancerous tissue in the gut. Considering our observations, one can conclude that UC-critical genes viz. *PCK1*, *WFDC2*, *MUC4*, *TTLL12*, *THRA*, *EPHB3*, *DPP10*, and *HSAP6* possesses a definite influence over the oncogenesis of colorectal carcinoma (**Fig 2**).

Hitherto, genomic alterations, dysregulation of signaling pathways, and recruitment of immune mediators by the UC-critical genes have been reported to be associated with the tumorigenesis of colon tissue [17, 20, 47]. However, a comprehensive study on the involvement of genes mediating oncogenesis, their mutation status, epigenetic regulation, and the linkages with the immune networks is not available to date. In this direction, examining the transcriptional expression and epigenetic regulations of the most influential UC-critical genes in CRC will further enlighten the understanding of the molecular trajectory of UC-to-CRC. Herein, we have found *WFDC2*, *TTLL12*, *THRA*, and *EPHB3* as significant mediators for the transition of UC into CRC and are denoted as the vital UC-CRC-critical genes (**Figs 3–5**). *WFDC2* encoded protein product is an essential constituent of the gut mucosa and it acts as a component of the extracellular matrix to prevent the invasion of harmful microorganisms [48]. Notably, this gene is associated with carcinogenesis in the lungs and ovary [49]. On the other side, *TTLL12* has been reported to be a negative regulator of innate immune response in the gut that maintains immune homeostasis in the healthy gut [50]. Perturbation in the function of *TTLL12* has been reported to promote tumorigenesis and metastasis in several cancers (pancreatic ductal adenocarcinoma, prostate, and ovarian cancer) other than CRC [51]. Similarly, no studies have documented this gene as a mediator of UC. *THRA* encodes the receptor for triiodothyronine (T3) in the gut [52], and malfunctioning of this gene is known to cause hyperplasia, hypertrophy, and loss of function, the cardinal signs of tumorigenesis [53]. While *EPHB3* encodes the Eph receptor comprising the transmembrane tyrosine kinase receptors [54] and regulates the migration of cellular components of the human gut [55]. A study on *EPHB3* knockout mice revealed a clear indication of its association with the induction of carcinogenesis [56]. This is the maiden report hypothesizing the role of *WFDC2*, *TTLL12*, *THRA*, and *EPHB3* in the induction of UC and the subsequent development to CRC. Our data clearly indicate that higher transcriptional expression of *WDFC2* and *EPHB3* promote colorectal mucinous adenocarcinoma within UC tissue, whereas upregulation of *THRA* and *TTLL12* mRNA signal rectal adenocarcinoma (**Fig 3**). It is further validated by the higher functional expression and promoter methylation over the different cancer stages as compared to the normal samples. Promoter hypermethylation causes transcriptional silencing of several tumor suppressor genes, while upon hypomethylation it activates the transcription of protooncogenes and other key protein coding genes that infers genomic instability and metastasis [57].
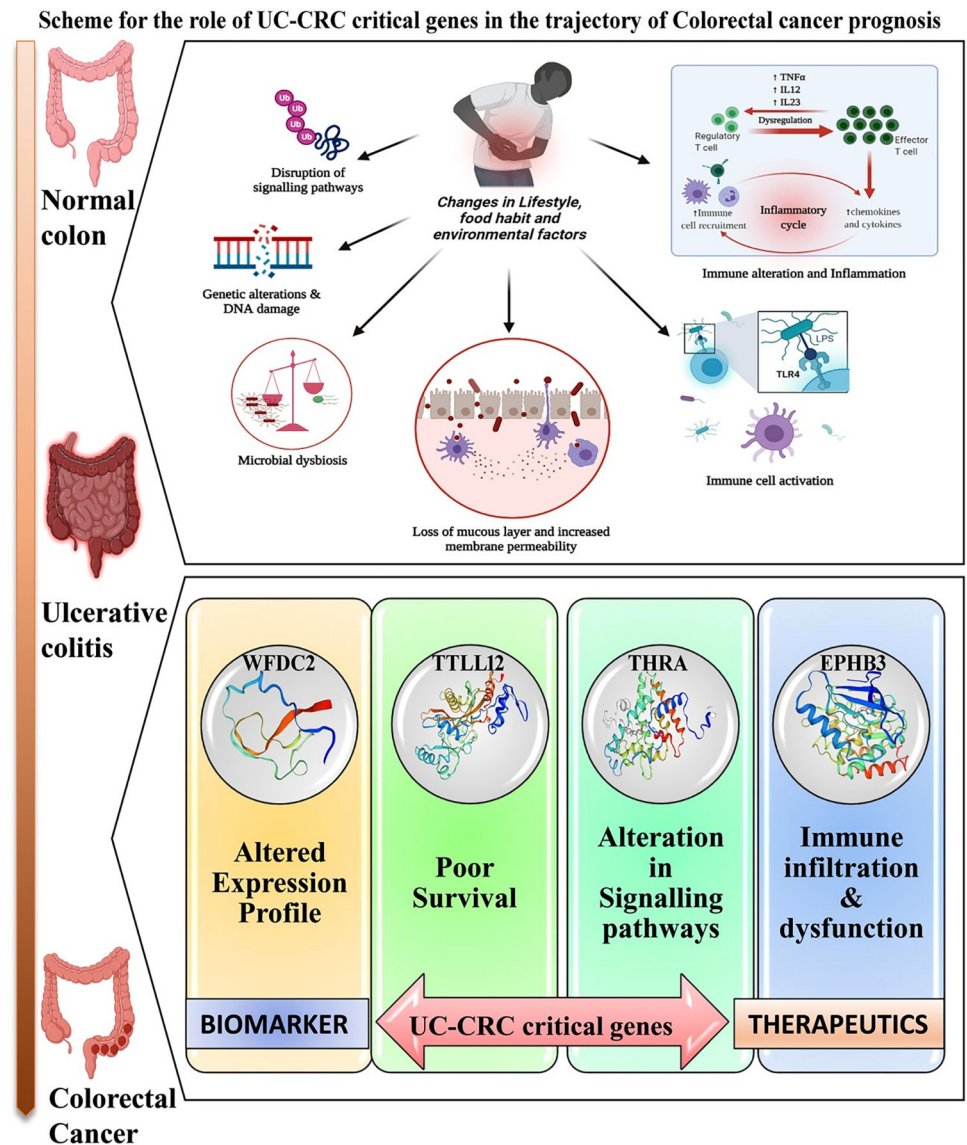
Fig 7. Scheme depicting the molecular trajectory of the development of UC and UC to CRC.

https://doi.org/10.1371/journal.pone.0289064.g007

Moreover, the survival assay for the UC-CRC-critical genes infers poor prognosis of colorectal carcinoma (**Fig 4B–4E**). Recruitment of immune cells in and around the transformed colon tissue creates a tumor microenvironment that can be either tumor-inhibiting or promoting [46]. Intriguingly, immune cell assembly within/around the transformed tissue is known to be associated with the prognosis as well as survival of the patients, and such outcomes can be predicted by the immune infiltration scores. In this study, the immune infiltration of UC-CRC-critical genes discloses the positive correlation of CD4+ T cells while negatively correlated with CD8+ T cells for CRC (**Fig 5**). Amongst the various members of the adaptive immune cell repertoire reported for human CRC, CD4+ T cells and CD8+ T cells are known to possess antitumor immune responses [46]. They induce the cytolytic responses against the cancer cells, signaling B cell activation and alteration of the immune homeostasis of the tumor microenvironment [58]. B cells are reported to be linked with the antigen presentation, generation of

antibodies, and various immunosuppressive responses [46]. Herein, expression of TTLL12, THRA, and EPHB3 mRNA was found to be associated with the infiltration of macrophages and NK cells at the tumor site (**Fig 5**). Macrophages are considered as the key regulators of the inflammatory consequences of human UC and CRC [59]. The immunoregulatory roles of macrophages are majorly executed via the phenotypic switching process namely macrophage polarization between the M1/pro-inflammatory and M2/anti-inflammatory subtypes [60, 61]. M2 macrophages alleviate the colonic inflammation and thereby inhibit metastasis while M1 macrophages do the reverse [59].

Finally, our study proposes *WFDC2*, *TTLL12*, *THRA*, and *EPHB3* as the potential mediators of the oncogenic transformation of CRC from UC-associated inflamed colon tissue and these genes could be considered as significant biomarkers and/or therapeutic targets in diagnosing and treating UC-associated neoplastic transformation, especially CRC. The possible scheme for the transformation of the normal colon to CRC via UC guided by the aforesaid genes is depicted in **Fig 7**. Our findings are expected to deliver novel dimensions to the existing knowledge on the molecular transformation of UC to CRC and will open up new areas for further research and experimental validation.

Supporting information

S1 File. Contains all the supporting tables and figures.
(DOCX)

Acknowledgments

Author Contributions

Conceptualization: Ritwik Patra, Suprabhat Mukherjee.

Data curation: Ritwik Patra.

Formal analysis: Ritwik Patra, Amit Kumar Dey, Suprabhat Mukherjee.

Investigation: Ritwik Patra, Suprabhat Mukherjee.

Methodology: Ritwik Patra, Suprabhat Mukherjee.

Project administration: Suprabhat Mukherjee.

Resources: Ritwik Patra.

Supervision: Suprabhat Mukherjee.

Validation: Ritwik Patra, Amit Kumar Dey, Suprabhat Mukherjee.

Visualization: Ritwik Patra.

Writing – original draft: Ritwik Patra, Suprabhat Mukherjee.

Writing – review & editing: Ritwik Patra, Amit Kumar Dey, Suprabhat Mukherjee.

References

1. Kaplan GG. The global burden of IBD: from 2015 to 2025. Nat Rev Gastroenterol Hepatol. 2015; 12: 720–727. https://doi.org/10.1038/nrgastro.2015.150 PMID: 26323879

2. Kobayashi T, Siegmund B, Le Berre C, Wei SC, Ferrante M, Shen B, et al. Ulcerative colitis. Nat Rev Dis Prim. 2020; 6: 74. https://doi.org/10.1038/s41572-020-0205-x PMID: 32913180

3. Patra R, Padma S, Mukherjee S. An improved method for experimental induction of Ulcerative colitis in Sprague Dawley rats. MethodsX. 2023; 102158. https://doi.org/10.1016/j.mex.2023.102158 PMID: 37091959

4. Torres J, Mehandru S, Colombel J-F, Peyrin-Biroulet L. Crohn's disease. Lancet. 2017; 389: 1741–1755. https://doi.org/10.1016/S0140-6736(16)31711-1 PMID: 27914655

5. Black CJ, Ford AC. Global burden of irritable bowel syndrome: trends, predictions and risk factors. Nat Rev Gastroenterol Hepatol. 2020; 17: 473–486. https://doi.org/10.1038/s41575-020-0286-8 PMID: 32296140

6. Mukherjee S, Joardar N, Sengupta S, Babu SPS. Gut microbes as future therapeutics in treating inflammatory and infectious diseases: lessons from recent findings. J Nutr Biochem. 2018; 61: 111–128. https://doi.org/10.1016/j.jnutbio.2018.07.010 PMID: 30196243

7. Ahmedy OA, Ibrahim SM, Salem HH, Kandil EA. Antiulcerogenic effect of melittin via mitigating TLR4/TRAF6 mediated NF-κB and p38MAPK pathways in acetic acid-induced ulcerative colitis in mice. Chem Biol Interact. 2020; 331: 109276. https://doi.org/10.1016/j.cbi.2020.109276 PMID: 33002459

8. Ungaro R, Mehandru S, Allen PB, Peyrin-Biroulet L, Colombel J-F. Ulcerative colitis. Lancet. 2017; 389: 1756–1770. https://doi.org/10.1016/S0140-6736(16)32126-2 PMID: 27914657

9. Ng SC, Kaplan GG, Tang W, Banerjee R, Adigopula B, Underwood FE, et al. Population Density and Risk of Inflammatory Bowel Disease: A Prospective Population-Based Study in 13 Countries or Regions in Asia-Pacific. Am J Gastroenterol. 2019; 114: 107–115. https://doi.org/10.1038/s41395-018-0233-2 PMID: 30177785

10. Siow VS, Bhatt R, Mollen KP. Management of acute severe ulcerative colitis in children. Semin Pediatr Surg. 2017; 26: 367–372. https://doi.org/10.1053/j.sempedsurg.2017.10.006 PMID: 29126505

11. Hnatyszyn A, Hryhorowicz S, Kaczmarek-Ryś M, Lis E, Słomski R, Scott RJ, et al. Colorectal carcinoma in the course of inflammatory bowel diseases. Hered Cancer Clin Pract. 2019; 17: 18. https://doi.org/10.1186/s13053-019-0118-4 PMID: 31338130

12. Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RGS, Barzi A, et al. Colorectal cancer statistics, 2017. CA Cancer J Clin. 2017; 67: 177–193. https://doi.org/10.3322/caac.21395 PMID: 28248415

13. Eaden JA, Abrams KR, Mayberry JF. The risk of colorectal cancer in ulcerative colitis: a meta-analysis. Gut. 2001; 48: 526–535. https://doi.org/10.1136/gut.48.4.526 PMID: 11247898

14. Choi C-HR, Bakir I Al, Hart AL, Graham TA. Clonal evolution of colorectal cancer in IBD. Nat Rev Gastroenterol Hepatol. 2017; 14: 218–229. https://doi.org/10.1038/nrgastro.2017.1 PMID: 28174420

15. Itzkowitz SH, Yio X. Inflammation and Cancer IV. Colorectal cancer in inflammatory bowel disease: the role of inflammation. Am J Physiol Liver Physiol. 2004; 287: G7–G17. https://doi.org/10.1152/ajpgi.00079.2004 PMID: 15194558

16. Zhang X, Wan S, Yu Y, Ruan W, Wang H, Xu L, et al. Identifying potential DNA methylation markers in early-stage colorectal Cancer. Genomics. 2020; 112: 3365–3373. https://doi.org/10.1016/j.ygeno.2020.06.007 PMID: 32531444

17. Patra R, Mukherjee S, Das NC. Exploring the differential expression and prognostic significance of the COL11A1 gene in human colorectal carcinoma: an integrated bioinformatics approach. Front Genet. 2021; 12: 88. https://doi.org/10.3389/fgene.2021.608313 PMID: 33597969

18. Liang B, Li C, Zhao J. Identification of key pathways and genes in colorectal cancer using bioinformatics analysis. Med Oncol. 2016; 33: 111. https://doi.org/10.1007/s12032-016-0829-6 PMID: 27581154

19. Zenonos K, Kyprianou K. RAS signaling pathways, mutations and their role in colorectal cancer. World J Gastrointest Oncol. 2013; 5: 97–101. https://doi.org/10.4251/wjgo.v5.i5.97 PMID: 23799159

20. Al-Mustanjid M, Mahmud SMH, Royel MRI, Rahman MH, Islam T, Rahman MR, et al. Detection of molecular signatures and pathways shared in inflammatory bowel disease and colorectal cancer: A bioinformatics and systems biology approach. Genomics. 2020; 112: 3416–3426. https://doi.org/10.1016/j.ygeno.2020.06.001 PMID: 32535071

21. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets—update. Nucleic Acids Res. 2013; 41: D991–D995. https://doi.org/10.1093/nar/gks1193 PMID: 23193258

22. Xue G, Hua L, Zhou N, Li J. Characteristics of immune cell infiltration and associated diagnostic biomarkers in ulcerative colitis: results from bioinformatics analysis. Bioengineered. 2021; 12: 252–265. https://doi.org/10.1080/21655979.2020.1863016 PMID: 33323040

23. Lu J, Wang Z, Maimaiti M, Hui W, Abudourexiti A, Gao F. Identification of diagnostic signatures in ulcerative colitis patients via bioinformatic analysis integrated with machine learning. Hum Cell. 2022; 35: 179–188. https://doi.org/10.1007/s13577-021-00641-w PMID: 34731452

24. Zhang J, Wang X, Xu L, Zhang Z, Wang F, Tang X. Investigation of Potential Genetic Biomarkers and Molecular Mechanism of Ulcerative Colitis Utilizing Bioinformatics Analysis. Chiba T, editor. Biomed Res Int. 2020; 2020: 4921387. https://doi.org/10.1155/2020/4921387 PMID: 32190668

25. Li K, Strauss R, Ouahed J, Chan D, Telesco SE, Shouval DS, et al. Molecular Comparison of Adult and Pediatric Ulcerative Colitis Indicates Broad Similarity of Molecular Pathways in Disease Tissue. J Pediatr Gastroenterol Nutr. 2018; 67: 45–52. https://doi.org/10.1097/MPG.0000000000001898 PMID: 29401083

26. Jiao X, Sherman BT, Huang DW, Stephens R, Baseler MW, Lane HC, et al. DAVID-WS: a stateful web service to facilitate gene/protein list analysis. Bioinformatics. 2012/04/27. 2012; 28: 1805–1806. https://doi.org/10.1093/bioinformatics/bts251 PMID: 22543366

27. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res. 2021; 49: D605–D612. https://doi.org/10.1093/nar/gkaa1074 PMID: 33237311

28. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. Cancer Discov. 2012; 2: 401 LP– 404. https://doi.org/10.1158/2159-8290.CD-12-0095 PMID: 22588877

29. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, et al. ONCOMINE: a cancer microarray database and integrated data-mining platform. Neoplasia (New York, NY). 2004; 6: 1. https://doi.org/10.1016/s1476-5586(04)80047-2 PMID: 15068665

30. Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK, et al. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia. 2017; 19: 649–658. https://doi.org/10.1016/j.neo.2017.05.002 PMID: 28732212

31. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. Science. 2015; 347: 1260419. https://doi.org/10.1126/science.1260419 PMID: 25613900

32. Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q, et al. TIMER2. 0 for analysis of tumor-infiltrating immune cells. Nucleic Acids Res. 2020.

33. Zhou G, Soufan O, Ewald J, Hancock REW, Basu N, Xia J. NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. Nucleic Acids Res. 2019; 47: W234–W241. https://doi.org/10.1093/nar/gkz240 PMID: 30931480

34. Rogler G. Chronic ulcerative colitis and colorectal cancer. Cancer Lett. 2014; 345: 235–241. https://doi.org/10.1016/j.canlet.2013.07.032 PMID: 23941831

35. Yashiro M. Molecular Alterations of Colorectal Cancer with Inflammatory Bowel Disease. Dig Dis Sci. 2015; 60: 2251–2263. https://doi.org/10.1007/s10620-015-3646-4 PMID: 25840920

36. Moparthi L, Koch S. Wnt signaling in intestinal inflammation. Differentiation. 2019; 108: 24–32. https://doi.org/10.1016/j.diff.2019.01.002 PMID: 30718056

37. Schatoff EM, Leach BI, Dow LE. Wnt Signaling and Colorectal Cancer. Curr Colorectal Cancer Rep. 2017; 13: 101–110. https://doi.org/10.1007/s11888-017-0354-9 PMID: 28413363

38. Fearon ER. Molecular genetics of colorectal cancer. Annu Rev Pathol. 2011; 6: 479–507. https://doi.org/10.1146/annurev-pathol-011110-130235 PMID: 21090969

39. Koveitypour Z, Panahi F, Vakilian M, Peymani M, Forootan FS, Esfahani MHN, et al. Signaling pathways involved in colorectal cancer progression. Cell Biosci. 2019; 9: 97. https://doi.org/10.1186/s13578-019-0361-4 PMID: 31827763

40. Jung G, Hernández-Illán E, Moreira L, Balaguer F, Goel A. Epigenetics of colorectal cancer: biomarker and therapeutic potential. Nat Rev Gastroenterol Hepatol. 2020; 17: 111–130. https://doi.org/10.1038/s41575-019-0230-y PMID: 31900466

41. Padma S, Patra R, Sen Gupta PS, Panda SK, Rana MK, Mukherjee S. Cell Surface Fibroblast Activation Protein-2 (Fap2) of Fusobacterium nucleatum as a Vaccine Candidate for Therapeutic Intervention of Human Colorectal Cancer: An Immunoinformatics Approach. Vaccines. 2023. https://doi.org/10.3390/vaccines11030525 PMID: 36992108

42. Yesudhas D, Gosu V, Anwar MA, Choi S. Multiple Roles of Toll-Like Receptor 4 in Colorectal Cancer. Frontiers in Immunology. 2014. p. 334. https://doi.org/10.3389/fimmu.2014.00334 PMID: 25076949

43. Zhao P, Zhang Z. TNF-α promotes colon cancer cell migration and invasion by upregulating TROP-2. Oncol Lett. 2018; 15: 3820–3827. https://doi.org/10.3892/ol.2018.7735 PMID: 29467899

44. Patra R, Das NC, Mukherjee S. Toll-Like Receptors (TLRs) as Therapeutic Targets for Treating SARS-CoV-2: An Immunobiological Perspective. Adv Exp Med Biol. 2021; 1352: 87–109. https://doi.org/10.1007/978-3-030-85109-5_6 PMID: 35132596

45. Brostjan C, Oehler R. The role of neutrophil death in chronic inflammation and cancer. Cell Death Discov. 2020; 6: 26. https://doi.org/10.1038/s41420-020-0255-6 PMID: 32351713

46. Guo L, Wang C, Qiu X, Pu X, Chang P. Colorectal Cancer Immune Infiltrates: Significance in Patient Prognosis and Immunotherapeutic Efficacy. Frontiers in Immunology. 2020. p. 1052. https://doi.org/10.3389/fimmu.2020.01052 PMID: 32547556

47. Thorsteinsdottir S, Gudjonsson T, Nielsen OH, Vainer B, Seidelin JB. Pathogenesis and biomarkers of carcinogenesis in ulcerative colitis. Nat Rev Gastroenterol Hepatol. 2011; 8: 395–404. https://doi.org/10.1038/nrgastro.2011.96 PMID: 21647200

48. Parikh K, Antanaviciute A, Fawkner-Corbett D, Jagielowicz M, Aulicino A, Lagerholm C, et al. Colonic epithelial cell diversity in health and inflammatory bowel disease. Nature. 2019; 567: 49–55. https://doi.org/10.1038/s41586-019-0992-y PMID: 30814735

49. Bingle L, Cross SS, High AS, Wallace WA, Rassl D, Yuan G, et al. WFDC2 (HE4): a potential role in the innate immunity of the oral cavity and respiratory tract and the development of adenocarcinomas of the lung. Respir Res. 2006; 7: 61. https://doi.org/10.1186/1465-9921-7-61 PMID: 16600032

50. Ju L-G, Zhu Y, Lei P-J, Yan D, Zhu K, Wang X, et al. TTLL12 Inhibits the Activation of Cellular Antiviral Signaling through Interaction with VISA/MAVS. J Immunol. 2017; 198: 1274 LP– 1284. https://doi.org/10.4049/jimmunol.1601194 PMID: 28011935

51. Yang S, Liang Y, Qian H, Li Q. TTLL12 expression in ovarian cancer correlates with a poor outcome. Int J Clin Exp Pathol. 2020; 13: 239–247. PMID: 32211104

52. Tylki-Szymańska A, Acuna-Hidalgo R, Krajewska-Walasek M, Lecka-Ambroziak A, Steehouwer M, Gilissen C, et al. Thyroid hormone resistance syndrome due to mutations in the thyroid hormone receptor α gene (THRA). J Med Genet. 2015; 52: 312 LP– 316. https://doi.org/doi:10.1136/jmedgenet-2014-102936 PMID: 25670821

53. van Mullem AA, Visser TJ, Peeters RP. Clinical Consequences of Mutations in Thyroid Hormone Receptor-α1. Eur Thyroid J. 2014; 3: 17–24. https://doi.org/10.1159/000360637 PMID: 24847461

54. Ji X-D, Li G, Feng Y-X, Zhao J-S, Li J-J, Sun Z-J, et al. EphB3 is overexpressed in non-small-cell lung cancer and promotes tumor metastasis by enhancing cell survival and migration. Cancer Res. 2011; 71: 1156–1166. https://doi.org/10.1158/0008-5472.CAN-10-0717 PMID: 21266352

55. Islam S, Loizides AM, Fialkovich JJ, Grand RJ, Montgomery RK. Developmental expression of Eph and ephrin family genes in mammalian small intestine. Dig Dis Sci. 2010; 55: 2478–2488. https://doi.org/10.1007/s10620-009-1102-z PMID: 20112066

56. Herath NI, Spanevello MD, Doecke JD, Smith FM, Pouponnot C, Boyd AW. Complex expression patterns of Eph receptor tyrosine kinases and their ephrin ligands in colorectal carcinogenesis. Eur J Cancer. 2012; 48: 753–762. https://doi.org/10.1016/j.ejca.2011.07.003 PMID: 21852108

57. Luczak MW, Jagodziński PP. The role of DNA methylation in cancer development. Folia Histochem Cytobiol. 2006; 44: 143–154. PMID: 16977793

58. Borst J, Ahrends T, Bąbała N, Melief CJM, Kastenmüller W. CD4+ T cell help in cancer immunology and immunotherapy. Nat Rev Immunol. 2018; 18: 635–647. https://doi.org/10.1038/s41577-018-0044-0 PMID: 30057419

59. Cortese N, Soldani C, Franceschini B, Barbagallo M, Marchesi F, Torzilli G, et al. Macrophages in Colorectal Cancer Liver Metastases. Cancers (Basel). 2019; 11: 633. https://doi.org/10.3390/cancers11050633 PMID: 31067629

60. Mukherjee S, Mukherjee S, Maiti TK, Bhattacharya S, Babu SPS. A novel ligand of Toll-like receptor 4 from the sheath of Wuchereria bancrofti Microfilaria induces proinflammatory response in macrophages. J Infect Dis. 2017; 215: 954–965. https://doi.org/10.1093/infdis/jix067 PMID: 28453850

61. Mukherjee S, Mukherjee S, Bhattacharya S, Sinha Babu SP. Surface proteins of Setaria cervi induce inflammation in macrophage through Toll-like receptor 4 (TLR4)-mediated signalling pathway. Parasite Immunol. 2017; 39. https://doi.org/10.1111/pim.12389 PMID: 27659561