Novel *FAM134B* mutations and their clinicopathological significance in colorectal cancer

Farhadul Islam¹, Vinod Gopalan¹, Riajul Wahab¹, Katherine Ting-wei Lee¹, Md. Hakimul Haque¹, Afraa Mamoori¹, Cu-tai Lu², Robert A Smith^{1,3}, Alfred K-Y Lam^{1,*}

¹Cancer Molecular Pathology, School of Medicine and Griffith Health Institute, Griffith University

²Department of Surgery, Gold Coast Hospital, Gold Coast, Queensland, Australia

³Genomics Research Centre, Institute of Health and Biomedical Innovation, Faculty of Health, Queensland University of Technology, Brisbane, Queensland, Australia

***Address for correspondence:**

Professor Alfred Lam,

Head of Pathology, Griffith Medical School, Gold Coast Campus, Gold Coast QLD 4222, Australia.

E-mail: a.lam@griffith.edu.au

Telephone +61 7 56780718; Fax +61 7 56780303

Acknowledgement: The project was supported by the student scholarship from Griffith University and funding from Menzies Health Institute Queensland.

Abstract

FAM134B is a putative tumour suppressor gene and mutations in *FAM134B* have never been reported in colorectal cancer (CRC). This study aims to identify *FAM134B* mutation sites and the clinicopathological significance of the gene in patients with CRC. Eighty-eight colorectal cancers were studied for *FAM134B* mutations by sequencing. The mutations in these cancers were correlated with the clinical and pathological parameters. Also, mRNA and protein expression of *FAM134B* in colorectal cancers were examined by polymerase chain reaction, Western blots and immunofluorescence analysis. *FAM134B* mutation was noted in 46.5% (41/88) of patients with CRC. Thirty-one novel pathogenic mutations were noted in coding and intronic sites of *FAM134B* in CRC. Also, the majority of the variants identified in *FAM134B* were single nucleotide substitution. Of the 31 mutations, 8 novel frameshift mutations showed potential of nonsense-mediated mRNA decay (NMD) in computational analysis. In addition, *FAM134B* mutations were associated with various clinical and pathological variables including sex of the patients, presence of metachronous cancer, size, T staging, presence of distant metastases and positivity of microsatellite instability (MSI) in the cancer ($p < 0.05$). *FAM134B* mRNA and protein expressions were decreased in *FAM134B* mutated cancers. To conclude, *FAM134B* mutation is common in colorectal cancer. The association of the mutation with adverse clinical and pathological parameters are in concur with the tumour suppressive property of the gene.

Keyword: *FAM134B*, colorectal cancer, cancer genetics, mutation, mRNA expression.

Introduction

FAM134B (Family with sequence similarity 134, member B) also called *JK1* belongs to a family of three genes (*FAM134A*, *FAM134B* and *FAM134C*), located at 5p15.1, which was discovered using comparative genomic hybridization analysis (Tang et al. 2001; Tang et al. 2007). Human *FAM134B* cDNA encodes a 497-amino-acids *cis*-Golgi transmembrane protein consisting of a C-terminal coiled-coil domain and two long hydrophobic segments, which are essential for long-term survival of nociceptive and autonomic ganglion neurons (Kurth et al. 2009). Expression of *FAM134B* gene is also involved with vascular dementia and allergic rhinitis (Kong et al. 2011; Melchioni et al. 2014). Another study revealed that *FAM134B* regulated endoplasmic reticulum turnover by selective autophagy (Khaminets et al. 2015). *FAM134B* is an ER-resident receptor protein that can bind to autophagy modifier microtubules-associated protein 1A/1B-light chain 3 (LC3) and gamma aminobutyric acid receptor-associated protein (GABARAP) and facilitate endoplasmic reticulum degradation (Khaminets et al. 2015). It was also reported that suppression of *FAM134B* protein in human cells resulted an expansion of the ER structure while upregulation caused endoplasmic reticulum fragmentation and lysosomal degradation (Kurth et al. 2009; Khaminets et al. 2015). In addition, a recent study demonstrated that loss of *FAM134B* was associated with increased production of Ebola virus proteins GP and VP40 and higher accumulation of nucleocapsid lattices (Chiramel et al. 2016).

In cancer, studies from our group reported that *FAM134B* is a growth related gene and plays important role in the pathogenesis of oesophageal and colorectal cancer (Kasem et al. 2014a; Kasem et al. 2014b; Kasem et al. 2014c; Islam et al. 2016; Haque et al. 2016). In colorectal cancer, 57% patients had *FAM134B* DNA copy number deletion, which was associated with advanced cancer stages (Kasem et al. 2014c). It was also reported that *FAM134B* mRNA and protein expression was significantly downregulated in patients with

colorectal cancer when compared to that of colonic adenoma (pre-invasive lesions) and non-cancer tissues samples (Kasem et al. 2014a). A functional study revealed that *FAM134* modulated the invasion and migration of colon cancer cells (Kasem et al. 2014b). Also, *FAM134B* expression is stage dependent and it suppresses the growth and proliferation of colon cancer both *in vivo* and *in vitro* (Islam et al. 2016).

Mutations in *FAM134B* lead to production of mutant FAM134B protein which is unable to act as autophagy receptor and causes sensory neuropathy in human (Kurth et al. 2009; Khaminets et al. 2015). Loss of function mutations in *FAM134B* were first reported in hereditary sensory and autonomic neuropathy type IIB (HSAN IIB) (Kurth et al. 2009). Subsequently, a number of studies noted homozygous loss of function mutations in *FAM134B* in neuronal and vascular diseases (Kong et al. 2011; Davidson et al. 2012; Murphy et al. 2012; Ilgaz et al. 2014). Also, a recent study from our group reported novel mutations in *FAM134B* in patients with oesophageal squamous cell carcinoma with advanced pathological stages (Haque et al. 2016). However, mutational status of *FAM134B* and its clinicopathogenic associations remain completely unknown in colorectal cancer. As *FAM134B* copy number deletion, reduced mRNA and protein expression were noted in advanced pathological stages of colorectal cancer, mutations may be contribute to this downregulation in cancer tissues samples. Thus, the present study aims to identify novel mutation sites and its impact on clinicopathological features in colorectal cancer.

Materials and Methods

Patients and clinicopathological parameters

Tumour tissues and matched non-cancer tissues (near the surgical resection margin) from the same patient who underwent resection of colorectal carcinomas by a colorectal surgeon (CTL) were prospectively collected during the period of 2012 to 2014 from hospitals in Queensland, Australia. The collected tissues were snap frozen in liquid nitrogen and stored at -80°C. Patients were recruited consecutively during the study period with no selection bias. Tissues samples were chosen randomly for this study and were excluded in cases of lacking of adequate histology and loss of follow up. Ethical approval for this work has been obtained from the Griffith University Human Research Ethics Committee (GU Ref No: MSC/17/10/HREC).

Histological sections were cut and stained with haematoxylin and eosin (H&E) for light microscopic examination. The pathological features of patients with colorectal cancer were analysed by examining the histological sections and were graded, classified and staged according to the World Health Organization (WHO) criteria (Hamilton et al. 2000). Only adenocarcinoma was included in the study. After reviewing, 88 patients (46 women; 42 men) with colorectal adenocarcinomas were included for the analysis. The mean age of the patients used in this study was 69 years (range, 24-91). For the cancer, the site and size (maximum length in mm) of the cancer were recorded. Proximal cancers were defined as cancer located in the caecum, ascending colon and transverse colon whereas distal cancers were defined as the cancers found in the region of descending colon, sigmoid colon and rectum. 47 % (n = 41) of colorectal cancer was located in proximal colon while 53 % (n = 47) was located in the distal colorectum. Among the 88 cases studied, 22 % (n = 19) of the carcinomas had lymph node metastasis whereas 88 % (n = 69) of cancer patients had no lymph node metastasis. There were 11 % (n = 10) stage I, 43 % (n = 38) stage II, 22 % (n =

19) stage III and 24 % (n = 21) stage IV carcinomas. A cryostat (Leica Biosystems, Mt Waverley, VIC, Australia) was used to section tissues from these patients for RNA and DNA extraction, with confirmation that it contained a representative cancer area. A schematic summary of the flow of the experiments was shown in Figure 1.

Clinical Management

Clinical management was performed by a pre-agreed standardized multi-disciplinary protocol. The cancer tissues were investigated for microsatellite instability markers (by immunohistochemistry) and *BRAF* mutation status (by Sanger sequencing) according to the clinical guidelines. The use of post-operative adjuvant therapy was based on the pathological stage of the cancer. The follow-up period was defined as the interval between the date of surgery for colorectal carcinoma and the date of death or closing date of the study. The actuarial survival rate of the patients was calculated from the date of surgical resection of the colorectal carcinomas to the date of death or last follow-up. Only cancer-related death was counted as the end point in the statistical analysis. Persistence or recurrence of the disease was also recorded.

Cell Culture

Three colon cancer cell lines (SW-480, SW-48 and HCT116) and one non-neoplastic colon epithelial cell line (FHC) were used in this study. All these cells were purchased from ATCC (American type culture collection) and maintained as previously described (Islam et al. 2016).

Extraction of DNA and RNA

Sections were cut for haematoxylin and eosin staining from the collected frozen tissues. The sections were examined by an anatomical pathologist (AKL) to confirm that cancer cells made up more than 70% of the volume of the samples. Then, the selected fresh frozen tissues were sectioned into 7 μm slices for RNA and DNA extractions. DNA was extracted and purified from selected tissue sections with Qiagen DNeasy Blood & Tissue kit (Qiagen Pty. Ltd., Hilden, Germany) following the manufacturer's guide lines. DNA from cultured cells also extracted using this kit. RNA was extracted from the fresh frozen tissue sections using miRNeasy Mini kit (Qiagen) according to the manufacturer's protocol. The purity of the extracted DNA and RNA was checked with the optical density (OD) using a nanodrop spectrophotometer and the concentrations were noted in $\text{ng}/\mu\text{L}$.

Quantitative real time PCR (qPCR) analysis

For quantitative PCR, first-strand cDNA was generated using DyNAmo™ cDNA Synthesis Kits (Qiagen) as previously described (Gopalan et al. 2010). *FAM134B* mRNA expression changes in colorectal cancer tissues samples were examined using an IQ5 multicolour real-time polymerase chain reaction (PCR) detection system (Bio-Rad, Hercules, CA, USA). qPCR was performed in a total volume of 20 μL reaction mixture containing 10 μL of DyNAmo Flash SYBR green master mix (Bio-Rad), 1.5 μL of each 5 $\mu\text{mol/L}$ primer, 3 μL of cDNA at 50 $\text{ng}/\mu\text{L}$, and 4 μL of 0.1% diethylpyrocarbonate (DEPC) treated water as previously described (Kasem et al. 2014a). The amplification efficiencies were normalized to that of multiple housekeeping genes including beta actin, 18s and glyceraldehyde 3-phosphate dehydrogenase (GAPDH). GAPDH and β -actin were selected on the basis of consistent results. Primers that were used in this study are given in supplementary data 1. Results were presented as [inverse ratio \(expression of FAM134B normalised by internal control GAPDH and \$\beta\$ -actin expression\)](#). Fold changes were

calculated according to previously published protocol (Islam et al., 2017). A fold change of more than 2 was considered as high *FAM134B* expression and a fold change of 2 or less was considered as low *FAM134B* expression.

High-resolution melt (HRM) curve analysis

Genomic DNA extracted from 88 cancers and matched non-cancer tissues were used for HRM analysis to screen the possible mutations in *FAM134B* sequence. HRM curve analysis was performed followed by amplifying target sequences on the Rotor-Gene Q detection system (Qiagen) using the software - Rotor-Gene ScreenClust HRM Software. The exons 1-9 of *FAM134B* were PCR amplified in a total reaction volume of 10 µl comprising 5 µl of 2Xsensimix HRM master mix, 1 µl of 30ng/µl genomic DNA, diethylpyrocarbonate (DEPC, RNase free) treated water 2 µl and 1 µl of each *FAM134B* primer. The thermal cycling protocol was same as published previously (Haque et al. 2016). Each PCR run included a negative (no template) control. The melt curve data were generated by increasing the temperature from 65°C to 85°C for all assays, with temperature increase rate of 0.05°C/seconds and recording fluorescence.

Purification of PCR products and Sanger sequencing analysis

All the possible variants detected in HRM analysis were further checked with Sanger sequencing for identifying the possible mutations. Briefly, after HRM curve analysis, selected and specific PCR products was purified according to the manufacturer's protocols from the NucleoSpin[®] Gel and PCR Clean-up kit (Macherey- Nagel, Bethlehem, PA, USA). The purified PCR products were then subjected to sequencing using Big Dye Terminator (BDT) chemistry Version 3.1 (Applied Biosystems, Foster City, CA, USA) under standardised cycling PCR conditions. The generated data were analysed at the Australian

Genome Research Facility (AGRF) using a 3730xl Capillary sequencer (Applied Biosystems). Sequences were then analysed with Sequence Scanner 2 software (Applied Biosystems).

In silico analysis

The Ensembl transcript ID ENST00000306320 (transcript variant 1 for *FAM134B*) was used when required by a method as input. All the variants were analysed using freely available bioinformatics tools such as Mutation Taster with NCBI 37 and Ensembl 69 database release (Schwarz et al. 2014), PROVEAN (protein variation effect analyser) and SIFT (sorting intolerant from tolerant) to evaluate the consequences of the identified mutations. The results were also compared with ExAc and 1000 Genomes variant databases. The cut-off value used for PROVEAN and SIFT were -2.5 and 0.05 respectively, for predicting the pathogenic/non-pathogenic variants in this study.

Immunofluorescence

To examine the impact of *FAM134B* mutation on protein expression in mutated and non-mutated tissues samples of colorectal cancer, immunofluorescence analysis was performed. After being de-waxed (xylene) and re-hydrated (alcohols and water), fresh frozen sections (4 µm) were blocked with peroxidase block solution (Dako Australia Pty. Ltd, Sydney, NSW Australia). Then the slides were incubated with rabbit FAM134B (Santa Cruz, Dallas, TX, USA) (1:150) antibody at 4°C overnight. This was followed by two hours of incubation with secondary antibody labelled with fluorescein isothiocyanate (FITC) (Sigma-aldrich) at room temperature. The slides were mounted on glass slides and then observed under a confocal microscope (Nikon A1R+, Nikon Inc., Tokyo, Japan).

Western blot analysis

Total proteins were extracted from fresh frozen tissue samples with lysis buffer (Bio-Rad, Gladesville, NSW, Australia). These were followed by homogenization and quantitation by absorbance spectrometry [10]. Afterwards, total protein (30 µg) was separated by 15% SDS-PAGE (Bio-Rad) and transferred to nitrocellulose membranes (Bio-Rad). The membrane was then developed to detect protein bands according to the published protocol (Islam et al. 2016). Expression of FAM134B protein was quantified and normalized to GAPDH with Image J 1.48 software.

Statistical analysis

Comparisons between variable groups were analysed using the chi-square test, likelihood ratio and Fisher's exact test. All the data was entered into a computer database and the statistical analysis was executed using the Statistical Package for Social Sciences for Windows (version 24.0, IBM SPSS Inc., New York, NY, USA). Survival analysis was tested using Kaplan-Meier method. Significance level was taken at $p < 0.05$.

Results

***FAM134B* mRNA expression in colorectal cancer tissues**

The expression of *FAM134B* in cancer and matched non-neoplastic tissues samples was presented in Figure 2A. The expression was statistically significant ($p=0.002$) between cancer and non-cancer tissues samples (0.85 ± 0.07 versus 0.97 ± 0.03). Among the colorectal cancers, 76% ($n=67/88$) showed low levels of *FAM134B* mRNA expression whereas the remaining 24% ($n=21$) had *FAM134B* over expression (supplementary data 2). The association of *FAM134B* mRNA expression with clinicopathological parameters were also analysed. The data are presented in supplementary data 2. It was noted that colorectal adenocarcinomas with *BRAF* mutation had a higher chance of *FAM134B* down-regulation ($p=0.046$).

The median overall follow-up of patients with colorectal cancer was 47 months and the survival rates correlated with the pathological stages of cancer ($p=0.0001$) (Figure 2B). Patients with colorectal cancer expressing low *FAM134B* mRNA had shorter survival time when compared to those with high *FAM134B* mRNA expression (38 months versus 47 months) (Figure 2C). However, the difference in survival time between the groups did not reach statistical significance ($p=0.41$).

Identification of *FAM134B* mutations in colorectal tissues samples and cell lines

Variants in *FAM134B* sequence in tissues and cells samples were first suspected on the basis of distinctive melting curve of HRM analysis. The mutant HRM amplified products (exon 1 to exon 9) are presented in supplementary data 3.

In colorectal cancer, we noted 42 nucleotide variants in the complete coding (exon 1 to exon 9) region of *FAM134B* (Table 1, Figure 3). Among these, thirty six were identified as novel mutations. Also, six variants (c.744C>T, c.1164A>G, c.1112T>C, c.1149A>G and

c.1107G>C), were noted and they were previously reported in oesophageal squamous cell carcinoma (Haque et al. 2016). Of the 42 mutations, thirty four were identified as substitutional changes, while five were insertions (c.1130_1131insT, c.1160_1161insG, c.1155_1156insA, c.1176_1177A and c.654_655insGG) and three deletions (c.1148_1148delT, c.1134_1134delT and c.1138_1138delC) (Table 1).

In matched non-cancer tissues, we identified eight mutations (Table 1). In addition, in colon cancer cell lines, we noted c.1112T>C, c.1149A>G and c.944C>T variants (Figure 4 and supplementary data 4). Of the 42 variants identified, 37 were heterozygous and five were homozygous changes. Interestingly, the mutational variants c.1149A>G and c.1112T>C were found to be common in cancer, non-cancer tissues and cell lines of colon cancer origin. We also identified seven known single nucleotide polymorphisms (SNPs) in the cancers (Supplementary data 5 and 6).

We noted 32 variants in the noncoding intronic sequence of *FAM134B* in colorectal cancer (Table 2). Twenty five of them were identified as novel in colorectal cancer whereas seven (g.135916_135916delA, g.136084G>A, g.139115C>A, g.135963C>A, g.136122G>C, g.138869G>T and g.139063_139063delT) were previously described in oesophageal squamous cell carcinoma (Haque et al. 2016). Among these intronic variants, 23 were single nucleotide substitutions, one (g.135871CG>TT) a double base substitution, two were insertions and six were deletions (Table 2). A single (g.139063_139063delT) deletion, intronic variant was identified in non-cancer matched tissues whereas no mutation was noted in the intron region of *FAM134B* in colon cancer cell lines.

Results of predictions using Mutation Taster, PROVEAN and SIFT, which primarily assess the consequences of nucleotides as well as amino acids changes on protein features and functions, are listed in table 1 and table 2. More than 64% (27/42) of variants in the coding region of *FAM134B* were predicted as deleterious or damaging in colorectal cancer

(Table 1), whereas only 25% (8/32) of intronic mutations were noted as disease causing (Table 2). In non-cancer matched tissue samples, variants (p.Asp349Arg, p.Glu362Lys, p.Gly356Arg and p.Asp349Arg) were predicted to be pathogenic (Table 1). Among the known variants, p.Thr208Met and p.Val203Met were noted as damaging or deleterious (supplementary data 5). In addition, variant c.944C>T (p.Ser286Leu), which was identified in colon cancer cell lines was predicted as pathogenic (supplementary data 4).

***FAM134B* mutations associated with clinicopathological parameters in CRC**

The association of *FAM134B* mutations with clinical and pathological factors of patients with colorectal cancer is shown in Table 3. Approximately 47% (41/88) of patients with colorectal cancer had *FAM134B* mutations. Significant correlation between *FAM134B* mutations with the presence of microsatellite instability (MSI) ($p=0.001$) in colorectal cancer was detected. Also, a significant ($p=0.043$) association of metachronous colorectal cancer (defined as the presence of more than one primary carcinoma detected consequently in a single person and the second or subsequent cancer is not a recurrence or distant metastasis) (Lam et al. 2012) and *FAM134B* mutations was noted (Table 3).

Table 4 lists the relationship of clinicopathological variables with the mutations detected in different exons of *FAM134B*. Approximately 8% (7/88) of colorectal cancer showed *FAM134B* mutation in exon 4. Variants identified in exon 4 (11/42) were significantly correlated with sex of the patients, presence of metachronous cancer, MSI positivity and T-stage ($p=0.013$, 0.032, 0.024 and 0.014 respectively) of the cancer (Table 4). More than 5% (5/88) of the cancer patients had *FAM134B* mutations positive in exon 5 (5/42). These mutations were significantly associated with positivity of MSI ($p=0.004$) (Table 4). Exon 7 mutation was detected in 12.5% (11/88) of the cancers. The mutation in the exon was associated with distant metastases ($p=0.048$) and larger size (>40 mm) of the

tumour ($p=0.044$) (Table 4). *FAM134B* mutations were frequently seen in exon 9 as it was noted in 20.45% (18/88) of the cancers. The mutation were significantly correlated with age ($p=0.031$) of the patients and location of the cancer ($p=0.031$) (Table 4). Changes in cDNA, protein sequence, splice site alterations, frameshift due to mutations were related with pathological variables (supplementary data 7).

Finally, we analysed the relationship between the prevalence of *FAM134B* mutations and patient survival (supplementary data 8). Patients with *FAM134B* mutations in exon 4 had lower survival rates (39 months versus 48 months) when compared to those with no mutations (supplementary data 8A). Patients with cancers positive for exon 9 mutations had a mean survival time of 42 months whereas patients without mutation had a longer survival time of 47 months (supplementary data 8B). In addition, patients with deletion mutations in introns had shorter survival times (41 months versus 47 months) when compared to the non-mutated patients (supplementary data 8C). Nevertheless, the difference between these groups did not reach statistical significance level ($p>0.05$).

***FAM134B* mutations inhibit protein expression**

FAM134B protein expression in mutation positive cancers ($n=10$) and non-mutated cancers ($n=10$) were examined by immunofluorescence microscopy. *FAM134B* protein expression in representative mutation positive cancer samples (7/10) was remarkably lower when compared to that of mutation negative samples (Figures 5A & B). A number of pathogenic mutations were detected in the selected cancer samples (sample ID: 8, 20, 62, 96, 99, 117 and 125), which exhibited reduced protein expression. Also, very low *FAM134B* mRNA expression was noted in these samples (Figure 5C). In addition, colon cancer cells (in cell lines from SW480, SW48 and HCT116) having deleterious *FAM134B* mutations, expressed lower levels of *FAM134B* protein and mRNA when compared to that of non-

neoplastic colon epithelial (FHC) cells in immunofluorescence microscopic analysis (supplementary data 9).

The reduction of FAM134B protein expression in mutated cancer tissues were confirmed by Western blot analysis in representative samples. Figures 5D & E presented the results from mutated cancer tissues (n=7) and non-mutated matched non-cancer tissues (n=2). Five patients (5/7) with colorectal cancer (sample ID 20, 62, 99, 117 and 125) harbouring multiple mutations showed significant ($p=0.001$) reduction of FAM134B protein expression when compared with the cancers having no mutation (Figures 5C & D). The protein band intensity of FAM134B expression in cancer samples (20, 62, 99, 117 and 125) were 0.126 ± 0.02 , 0.137 ± 0.03 , 0.058 ± 0.02 , 0.07 ± 0.02 and 0.066 ± 0.01 respectively, whereas protein band intensity in the non-cancer tissues samples (sample ID 1 and 17) were 0.899 ± 0.12 and 0.0908 ± 0.08 respectively. Colon cancer cell lines (SW480, SW48 and HCT116) showed notable downregulation of FAM134B protein expression when compared to that of non-neoplastic colon epithelial (FHC) cell line (Supplementary data 10).

Discussion

Using the data of the ICGC (international cancer genome consortium), *FAM134B* mutation can be identified in many cancers (<https://dcc.icgc.org/>). *FAM134B* mutation was noted in approximately 7% of human cancers. In the current study, we determined that approximately 47% (41/88) of colorectal cancers had *FAM134B* mutation. In these 41 patients, we have identified 42 exonic variants in colorectal cancer tissues and colon cancer cell lines, 36 of which were novel mutations (Table 1). Among the 42 exonic variants, 29 were missense mutations which change the amino-acid sequences of FAM134B protein (Table 1). In computational analysis, we noted that eight frameshift mutations (p.Asp349ArgfsX13, p.Pro351HisfsX3, p.Thr359AsnfsX3, p.Met357AsnfsX5, p.Glu364ArgfsX9, p.Gln190GlyfsX63, p.Asn355MetfsX11, and p.Pro351HisfsX3) have the potential to cause nonsense-mediated mRNA decay (NMD), consequently can produce strongly truncated protein. However, further functional studies with the identified variants (especially with predicted pathogenic variants) are needed to confirm their roles in producing NMD or generating truncating protein. The remaining five variants (p.Glu354Glu, p.Thr359Thr, p.Gly358Gly, p.Gly211Gly and p.Phe178Phe) were identified as silent or synonymous mutations (no change in amino acid) in colorectal cancer. Mutations in *FAM134B* associated with non-malignant conditions were previously detected in exon 1 and intron 7. By contrast in colorectal cancer, most of the mutations were located in either of exons 4, 5, 7 and 9. Similarly, in oesophageal cancer, *FAM134B* mutations were also detected in either of exons 4, 5, 7 and 9 (Haque et al. 2016). Thus, variants in *FAM134B* appear to be unique in cancer and might be playing a key role in carcinogenesis.

In the current study, *FAM134B* mutations in exon 4 were significantly associated with the sex of the patients ($p=0.013$) with colorectal cancer. About 15% of female populations (7/46) showed *FAM134B* exon 4 mutations in patients with colorectal carcinoma. Mutations

in tumour suppressors such as *p53*, *BRCA1*, *BRCA2*, *WT1* or *ATM* genes, which maintain genetic stability, are more common in women and correlations were noted in different cancers (Buchholz et al. 1999; Lai et al. 2002; Rivera et al. 2007). For example, Rivera *et al* has reported that heterozygous mutations of an uncategorized gene *FAM123B* in female patients were associated with X chromosome mediated inactivation of tumour suppressor *WT1* in Wilms tumours (Rivera et al. 2007). Thus, *FAM134B* exon 4 mutations in female patients with colorectal cancer may be involved with the modulation of other regulatory gene(s) in the activation of genes on the X chromosome which may in turn lead to the inactivation of the *FAM134B* gene by a single-hit mechanism. Therefore, further research using genome-wide scanning with large series of cohort is needed to unveil the exact role of *FAM134B* in maintaining genetic integrity, especially in the female population.

FAM134B mutations (exon 4) were noted to be correlated with T-staging of CRC ($p=0.014$). Late stage cancers more often had *FAM134B* exon 4 mutations when compared to early stages of cancer. Similarly, patients with larger tumour ($>40\text{mm}$) in their colorectum exhibited a high frequency of *FAM134B* mutation in exon 7 ($p=0.044$) (Table 4). Also, mutations in exon 7 correlated with presence of distant metastasis in colorectal cancer ($p=0.048$). In addition, we observed that patients with cancer having exon 9 mutations or exon 4 mutations had poorer prognosis (supplementary data 8A and 8B). Furthermore, it was noted disease causing mutations (deletions) in the introns had an impact on patients' survival time. Mutation positive populations had shorter survival compared to that of non-mutated patients (supplementary data 8C). Small sample size and relatively shorter follow up data may be the limiting factor for getting statistically significant difference of survival rates between the groups of mutated and non-mutated populations in the present study.

However, the multiple clinical correlations we noted in this study imply that *FAM134B* mutations can be used in predicting tumour progression and aggressiveness. All

these results concur with the tumour suppressive and poor prognostic predictive roles of *FAM134B* in colorectal cancer. In our previous study we noted that low *FAM134B* mRNA and protein expression in colorectal carcinoma tissues were significantly associated with larger tumour size and advanced cancer stages. In the present study, mutations of *FAM134B* were linked with larger tumour size, advanced pathological stages, distant metastasis and poorer clinical outcomes. Therefore, mutations in *FAM134B* sequences could drive the lower mRNA and protein expressions in colorectal cancer.

Microsatellite instability (MSI) occurs in 15% of colorectal cancer and results from the inactivation of DNA mismatch repair genes family by aberrant methylation and /or by inducing somatic mutations (Pritchard and Grady 2011; Pakneshan et al. 2013). Metachronous colorectal carcinoma often has positive microsatellite instability (Lam et al. 2012). In this study, we noted that patients with metachronous colorectal cancers or MSI positive cancers showed more frequent *FAM134B* mutations. It could be postulated that *FAM134B* may be involved in the maintenance of genetic integrity and protect against unstable genotype formation.

The clinical features and molecular pathways are often different between cancers from proximal and distal portions of the large intestine (Lam et al. 2008a; Lam et al. 2008b; Saleh et al. 2008; Gopalan et al. 2010; Lam et al. 2011; Alqurashi et al. 2013). The present study noted the largest number of variants (18) of *FAM134B* in exon 9 and the frequency of those mutations were associated with the sites of the cancer in the colorectum (p=0.031). Approximately one fourth of the cancer from proximal colon showed *FAM134B* exon 9 mutations.

Eight out of 25 novel variants detected in the noncoding region of *FAM134B* in colorectal cancer have the potential to affect the protein features by modifying the splice sites of the gene. *In silico* [<http://www.mutationtaster.org/>] analysis have revealed that these

variants can make the existing splice site either stronger or weaker leading to change protein structure and functions. They also have the potential to activate additional splice sites or deactivate the existing splice sites completely. Thus, mutations in these sites could have the potentials to affect gene functions, likely as previously reported in neurological disease (Kurth et al. 2009).

Using paraffin-embedded cancer tissues, we have previously reported the loss of copy number and reduced expression of *FAM134B* in colorectal adenocarcinomas (Kasem et al. 2014a; Kasem et al. 2014c). The present study reported *FAM134B* mRNA expression changed in fresh colorectal adenocarcinoma tissues for the first time. Among the *FAM134B* mutation positive colorectal cancers, more than a third of them showed low *FAM134B* mRNA expression. We noted a significant reduction of *FAM134B* mRNA in cancer tissues when compared to their matched non-neoplastic tissues. These findings were analogous to our previous *FAM134B* expression profiling results (Kasem et al. 2014a; Kasem et al. 2014b; Kasem et al. 2014c; Islam et al. 2016).

BRAF mutation is an independent negative prognostic factor in patients with advanced and recurrent colorectal cancer (Pakneshan et al. 2013). In the current study, colorectal cancer with *BRAF* mutations showed reduced *FAM134B* expression suggesting *FAM134B*'s concordant downstream interactions with *BRAF* associated cancer pathways. It was reported that *BRAF* mutational activation can inactivate the tumour inhibitor *PTEN* and generate many different types of cancers by inducing hyper-activation of PI3K/AKT and mitogen-activated protein kinases (MAPK)-signalling (Rahman et al. 2013; Rahman et al. 2014; Rahman et al. 2015a; Rahman et al. 2015b; Rahman et al. 2016). Taken together, down regulation of *FAM134B* may lead *BRAF* mediated activation (or vice versa *BRAF* mutation may reduce *FAM134B* expression) of PI3K/AKT and MAPK signalling and promote colorectal cancer pathogenesis. Further studies incorporating large numbers of

samples as well as functional studies are needed to investigate the association between *BRAF* mutation and *FAM134B* expression.

We have investigated FAM134B protein expression in *FAM134B* mutated and non-mutated cancer by immunofluorescence staining and Western blot analysis. We noted significant reduction/loss of FAM134B protein expression in *FAM134B* mutated colorectal cancer (Figure 5). It was reported that homozygous nonsense mutations (p.S309X, p.Q145X) lead to nonsense-mediated decay resulting in complete absence of the protein or to a truncated and non-functional gene product occurred in autonomic neuronal disorders (HSAN II) (Kurth et al. 2009). Also, a frameshift (p.P7GfsX133) and an intronic mutation were identified in HSAN patients, which are likely subject to nonsense-mediated decay and loss of protein function. Also, in some colorectal cancers, *FAM134B* had shown mutant variants (p.Thr359Asn, p.Asp349Arg and p.Pro351His), which have the potential to mediate nonsense-mediated decay resulting in loss or reduced expression of FAM134B. Similarly, the variant noted in colon cancer cell lines, p.Ser286Leu, can alter the splice site at the location of the variant, which may be involved in the down regulation of the gene product. Thus, the pathogenic variants detected in the current study could be the possible cause of *FAM134B* down-regulation in colon cancer tissues and cell lines. Further functional studies on particular mutations are needed to confirm the impact of specific mutations on function and structure of the protein.

In conclusion, this study has for the first time reported the existence of multiple novel *FAM134B* mutation sites in colorectal carcinomas. The downregulation of *FAM134B* mRNA and protein in colorectal carcinoma might be driven by these mutations. Multiple novel mutant variations in the coding sequence of *FAM134B* were noted and each of these mutations may attribute to the altered expression or structural and functional changes of the gene, which in turn could play an important role in the pathogenesis of colorectal cancer.

The variants c.1149A>G, c.1112T>C, c.1164A>G, c.826A>T and c.1130_1131insT are the most frequent and common variants identified in colorectal cancer tissues and cell lines. In addition, the identified *FAM134B* mutations in colorectal cancer were associated with clinical and pathological variables, implying the potential clinical significance of the gene in the cancer.

Conflict of interest: There is no conflict of interest in publishing this paper

References

1. Alqurashi N, Gopalan V, Smith RA, Lam AK (2013) Clinical impacts of mammalian target of rapamycin expression in human colorectal cancers. *Hum Pathol* 44: 2089-2096.
2. Buchholz TA, Weil MM, Story MD, Strom EA, Brock WA, McNeese MD (1999) Tumor suppressor genes and breast cancer. *Radiat Oncol Investig* 7:55-65.
3. Chiramel AI, Dougherty JD, Nair V, Robertson SJ, Best SM (2016) FAM134B, the selective autophagy receptor for endoplasmic reticulum turnover, inhibits replication of Ebola virus strains Makona and Mayinga. *J Infect Dis* 214:s319-s325.
4. Davidson G, Murphy S, Polke J, Laura M, Salih M, Muntoni F, Blake J, Brandner S, Davies N, Horvath R, Price S, Donaghy M, Roberts M, Foulds N, Ramdharry G, Soler D, Lunn M, Manji H, Davis M, Houlden H, Reilly M (2012) Frequency of mutations in the genes associated with hereditary sensory and autonomic neuropathy in a UK cohort. *J Neurol* 259:1673-1685.
5. Gopalan V, Smith RA, Nassiri MR, Yasuda K, Salajegheh A, Kim SY, Ho YH, Weinstein S, Tang JC, Lam AK (2010) GAEC1 and colorectal cancer: a study of the relationships between a novel oncogene and clinicopathologic features. *Hum Pathol* 41:1009-1015.
6. Hamilton SR, Vogelstein B, Kudo S, Riboli E, Nakamura S, Hainaut P (2000) Carcinoma of the colon and rectum. In: Hamilton SR, Aaltonen LA (eds) WHO classification of tumors: Pathology and genetics of tumors of digestive system, IARC Press, Lyon, France, pp. 105-119.

7. Haque MH, Gopalan V, Chan KW, Shiddiky MJ, Smith RA, Lam AK (2016) Identification of novel FAM134B (JK1) mutations in oesophageal squamous cell carcinoma. *Sci Rep* 6:29173.
8. Ilgaz Aydinlar E, Rolfs A, Serteser M, Parman Y (2014) Mutation in FAM134B causing hereditary sensory neuropathy with spasticity in a Turkish family. *Muscle Nerve* 49:774-775.
9. Islam F, Gopalan V, Wahab R, Smith RA, Qiao B, Lam AK (2017) Stage dependent expression and tumor suppressive function of FAM134B (JK1) in colon cancer. *Mol Carcinog* 56: 238-249.
10. Kasem K, Gopalan V, Salajegheh A, Lu CT, Smith RA, Lam AK (2014a) The roles of JK-1 (FAM134B) expressions in colorectal cancer. *Exp Cell Res* 326:166-173.
11. Kasem K, Sullivan E, Gopalan V, Salajegheh A, Smith RA, Lam AK (2014b) FAM134B (FAM134B) represses cell migration in colon cancer: a functional study of a novel gene. *Exp Mol Pathol* 97:99-104.
12. Kasem K, Gopalan V, Salajegheh A, Lu CT, Smith RA, Lam AK (2014c) FAM134B (FAM134B) gene and colorectal cancer: a pilot study on the gene copy number alterations and correlations with clinicopathological parameters. *Exp Mol Pathol* 97:31-36.
13. Khaminets A, Heinrich T, Mari M, Grumati P, Huebner AK, Akutsu M, Liebmann L, Stolz A, Nietzsche S, Koch N, Mauthe M, Katona I, Qualmann B, Weis J, Reggiori F, Kurth I, Hübner CA, Dikic I (2015) Regulation of endoplasmic reticulum turnover by selective autophagy. *Nature* 522:354-358.
14. Kong M, Kim Y, Lee C (2011) A strong synergistic epistasis between FAM134B and TNFRSF19 on the susceptibility to vascular dementia. *Psychiatr Genet* 21:37-41.

15. Kurth I, Pamminer T, Hennings JC, Soehendra D, Huebner AK, Rotthier A, Baets J, Senderek J, Topaloglu H, Farrell SA, Nürnberg G, Nürnberg P, De Jonghe P, Gal A, Kaether C, Timmerman V, Hübner CA (2009) Mutations in FAM134B, encoding a newly identified Golgi protein, cause severe sensory and autonomic neuropathy. *Nat Genet* 41:1179-1181.
16. Lai H, Lin L, Nadji M, Lai S, Trapido E, Meng L (2002) Mutations in the p53 tumor suppressor gene and early onset breast cancer. *Cancer Biol Ther* 1:31-36.
17. Lam AK, Gopalan V, Carmichael R, Buettner PG, Leung M, Smith R, Lu CT, Ho YH, Siu S (2012) Metachronous carcinomas in colorectum and its clinicopathological significance. *Int J Colorectal Dis* 27:1303-1310.
18. Lam AK, Gopalan V, Nassiri MR, Kasim K, Dissanayake J, Tang JC, Smith RA (2011) Altered JS-2 expression in colorectal cancers and its clinical pathological relevance. *Mol Oncol* 5:475-481.
19. Lam AK, Ong K, Giv MJ, Ho YH (2008a) p16 expression in colorectal adenocarcinoma: marker of aggressiveness and morphological types. *Pathology* 40: 580-585.
20. Lam AK, Ong K, Ho YH (2008b) Aurora kinase expression in colorectal adenocarcinoma: correlations with clinicopathological features, p16 expression, and telomerase activity. *Hum Pathol* 39:599-604.
21. Melchioti R, Puan KJ, Andiappan AK, Poh TY, Starke M, Zhuang L, Petsch K, Lai TS, Chew FT, Larbi A, Wang de Y, Poidinger M, Rotzschke O (2014) Genetic analysis of an allergic rhinitis cohort reveals an intercellular epistasis between FAM134B and CD39. *BMC Med Genet* 15:73.

22. Murphy SM, Davidson GL, Brandner S, Houlden H, Reilly MM (2012) Mutation in FAM134B causing severe hereditary sensory neuropathy. *J Neurol Neurosurg Psychiatry* 83:119-120.
23. Pakneshan S, Salajegheh A, Smith RA, Lam AK (2013) Clinicopathological relevance of BRAF mutations in human cancer. *Pathology* 45:346-356.
24. Pritchard CC, Grady WM (2011) Colorectal cancer molecular biology moves into clinical practice. *Gut* 60:116-129.
25. Rahman MA, Salajegheh A, Smith RA, Lam AK (2014) BRAF inhibitor therapy for melanoma, thyroid and colorectal cancers: development of resistance and future prospects. *Curr Cancer Drug Targets* 14:128-143.
26. Rahman MA, Salajegheh A, Smith RA, Lam AK (2013) B-Raf mutation: a key player in molecular biology of cancer. *Exp Mol Pathol* 95:336-342.
27. Rahman MA, Salajegheh A, Smith RA, Lam AK (2016) Inhibition of BRAF kinase suppresses cellular proliferation, but not enough for complete growth arrest in BRAF V600E mutated papillary and undifferentiated thyroid carcinomas. *Endocrine* 54:129-138.
28. Rahman MA, Salajegheh A, Smith RA, Lam AK (2015a) MicroRNA-126 suppresses proliferation of undifferentiated (BRAF(V600E) and BRAF(WT)) thyroid carcinoma through targeting PIK3R2 gene and repressing PI3K-AKT proliferation-survival signalling pathway. *Exp Cell Res* 339:342-350
29. Rahman MA, Salajegheh A, Smith RA, Lam AK (2015b) Multiple proliferation-survival signalling pathways are simultaneously active in BRAF V600E mutated thyroid carcinomas. *Exp Mol Pathol* 99:492-497.
30. Rivera MN, Kim WJ, Wells J, Driscoll DR, Brannigan BW, Han M, Kim JC, Feinberg AP, Gerald WL, Vargas SO, Chin L, Iafrate AJ, Bell DW, Haber DA (2007)

An X chromosome gene, WTX, is commonly inactivated in Wilms tumor. *Science* 315:642-45.

31. Saleh S, Lam AK, Ho YH (2008) Real-time PCR quantification of human telomerase reverse transcriptase (hTERT) in colorectal cancer. *Pathology* 40:25-30.
32. Schwarz JM, Cooper DN, Schuelke M, Seelow D (2014) MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods* 11:361-362.
33. Tang JC, Lam KY, Law S, Wong J, Srivastava G (2001) Detection of genetic alterations in esophageal squamous cell carcinomas and adjacent normal epithelia by comparative DNA fingerprinting using inter-simple sequence repeat PCR. *Clin Cancer Res* 7:1539-1545.
34. Tang WK, Chui CH, Fatima S, Kok SH, Pak KC, Ou TM, Hui KS, Wong MM, Wong J, Law S, Tsao SW, Lam KY, Beh PS, Srivastava G, Chan AS, Ho KP, Tang JC (2007) Oncogenic properties of a novel gene JK-1 located in chromosome 5p and its overexpression in human esophageal squamous cell carcinoma. *Int J Mol Med* 19:915-923.

Figure 1. Schematic representation of the methodological flow used in this study. Inclusion criteria- patients with colorectal carcinomas (stage I to stage IV) underwent for surgical resection by a single surgeon. Excluding criteria- patients who loss the follow up and lacking of good histology.

Figure 2. Altered FAM134B mRNA expression and survival distribution in colorectal cancer. A) Scatter plot of mRNA expression level of *FAM134B* in colorectal cancer and matched non-neoplastic tissues. Paired T-test showed significant difference in expression of *FAM134B* mRNA (inverse ratio of *FAM134B* normalized by *GAPDH* or β -actin) in cancer and matched non-neoplastic tissues ($p=0.002$). B) Survival rates of patients with colorectal cancer. Patients with low level of *FAM134B* mRNA in had shorter survival time when compared to those with high *FAM134B* mRNA.

Figure 3. Novel exonic *FAM134B* mutations in colorectal cancer Comparison of representative wildtype and mutants sequences of *FAM134B* in colorectal cancer. The presence of variants was identified by normalized HRM melt curves and by direct sequencing. **A)** Homozygous variant c.1157T>A (ATG>AAG; p.Met357Lys) and frameshift variant c.1160_1161insG (ACA>AAC; p.Thr359AsnfsX3) in exon 9 of *FAM134B*. **B)** Heterozygous variant c.826A>T (AGT>TGT; p.Ser276Cys) in exon 7 of *FAM134B*. **C)** Synonymous variant c.744C>T (CTC>CTT; p.Leu219Leu) in exon 5. **D)** Heterozygous variant in exon 4 c.641T>G (ATG>AGG; p.Met185Arg) in *FAM134B*.

Figure 4. Novel mutations in *FAM134B* detected colon cancer cell lines. Mutations sites of *FAM134B* in colon cancer cell lines were confirmed with Sanger sequencing followed by normalized characteristics HRM curve analysis. Comparison of representative mutant and wildtype sequences of *FAM134B* in colon cancer cells. **A)** Variant c.1112T>C (GTT>GCT) in *FAM134B* was noted in SW48 and HCT116 cells. **B)** GAA>GAG (c.1149A>G) was noted in SW480 and SW48 colon cancer cells in *FAM134B* sequence. **C)** Variant c.944C>T (TCG>TTA) was detected in all three (SW480, SW48 and HCT116) colon cancer cells

Figure 5. *FAM134B* mutation lead reduced/loss of mRNA and protein expression in cancer. Mutated colorectal cancer showed reduced *FAM134B* mRNA and protein expression when compared to those cancers with no mutation. **A)** Representative image of mutated cancer tissues expressing reduced level of FAM134B protein. **B)** Cancers with no *FAM134B* mutation expressed higher level of FAM134B protein. Representative images are shown as 10x magnification. Scale bar 50 μ m. Green fluorescence represents FAM134B protein

expression and blue indicate DAPI (counterstain) expression. **C)** *FAM134B* mutation leads reduced mRNA expression in colorectal cancer (mutant; T) when compared to cancers with no mutation (wildtype; WT) tissues. **D)** Absence or very weak FAM134B protein bands were noted in mutated cancer when compared to the non-mutated samples. **E)** Bar graph represents the protein band intensity (normalized to GAPDH) obtained from three independent measurements. Results are shown as mean $\pm$ standard deviation from three independent tests. Level of significance was *** $p > 0.001$ when compared with cancers with no *FAM134B* mutation. N: cancer with no *FAM134B* mutation; T: cancer with *FAM134B* mutation.

Table 1. Mutations detected in coding region of *FAM134B* (*JK1*) in colorectal cancer

Sample ID	Exons	mRNA expression	Change in DNA sequence	Change in Codon	Change in protein sequence	Frameshift	Protein features	In silico prediction		
								PROVEAN	SIFT	Mutation taster
6	Exon 9	Very low	c.1149A>G c.1162A>G c.1188T>A c.1112T>C	GAA>GAG ACA>GCA CTT>CTA GTT>GCT	p.Glu354Glu p.Thr359Ala p.Leu367Leu p.Val342Ala	No No No No	Might get lost (downstream of altered splice site) Might get lost (downstream of altered splice site) Might get lost (downstream of altered splice site) Might get lost (downstream of altered splice site)	Neutral Neutral Neutral Neutral	Tolerated Tolerated Tolerated Tolerated	Polymorphism Polymorphism Polymorphism Polymorphism
8		Low	c.1149A>G c.1171G>A c.1130_1131insT	GAA>GAG GAA>AAA GAT>AGA	p.Glu354Glu p.Glu362Lys p.Asp349ArgfsX13	No No Yes	Might get lost (downstream of altered splice site) Might get lost (downstream of altered splice site) Strongly truncated protein, might cause NMD	Neutral Deleterious Deleterious	Tolerated Damaging Damaging	Polymorphism Diseases causing Diseases causing
20		Low	c.1157T>A c.1160_1161insG c.1188T>A c.1102A>T	ATG>AAG ACA>AAC CTT>CTA AGT>TGT	p.Met357Lys p.Thr359AsnfsX3 p.Leu367Leu p.Ser339Cys	No Yes No No	Protein structure (might be) affected Strongly truncated protein, might cause NMD Might get lost (downstream of altered splice site) Protein structure (might be) affected	Neutral Neutral Neutral Deleterious	Tolerated - Neutral Damaging	Polymorphism Diseases causing Polymorphism Polymorphism
27		Low	c.1148_1148delA c.1115T>C	AAT>ATG TTT>TCT	p.Asn355MetfsX11 p.Phe343Ser	Yes No	Truncated protein, might cause NMD Might get lost (downstream of altered splice site)	Deleterious Deleterious	Damaging Damaging	Diseases causing Diseases causing
28		Low	c.1149A>G c.1155_1156insA	GAA>GAG AUG>AAT	p.Glu354Glu p.Met357AsnfsX5	No Yes	Might get lost (downstream of altered splice site) Truncated protein, might cause NMD	Neutral Neutral	Tolerated Tolerated	Polymorphism Diseases causing
37		Normal	c.1149A>G c.1164A>G c.1112T>C	GAA>GAG ACA>ACG GTT>GCT	p.Glu354Glu p.Thr359Thr p.Val342Ala	No No No	Might get lost (downstream of altered splice site) Protein feature (might be) affected Might get lost (downstream of altered splice site)	Neutral Neutral Neutral	Tolerated Tolerated Tolerated	Polymorphism Diseases causing Polymorphism
42		Very low	c.1149A>G c.1164A>G c.1112T>C	GAA>GAG ACA>ACG GTT>GCT	p.Glu354Glu p.Thr359Thr p.Val342Ala	No No No	Might get lost (downstream of altered splice site) Protein feature (might be) affected Might get lost (downstream of altered splice site)	Neutral Neutral Neutral	Tolerated Tolerated Tolerated	Polymorphism Diseases causing Polymorphism
57		Very low	c.1149A>G c.1164A>G c.1130_1131insT	GAA>GAG ACA>ACG GAT>AGA	p.Glu354Glu p.Thr359Thr p.Asp349Arg fsX13	No No Yes	Might get lost (downstream of altered splice site) Protein feature (might be) affected Strongly truncated protein, might cause NMD	Neutral Neutral Deleterious	Tolerated Tolerated Damaging	Polymorphism Diseases causing Diseases causing
60		High	c.1149A>G c.1112T>C	GAA>GAG GTT>GCT	p.Glu354Glu p.Val342Ala	No No	Might get lost (downstream of altered splice site) Might get lost (downstream of altered splice site)	Neutral Neutral	Tolerated Tolerated	Polymorphism Polymorphism
62		Very low	c.1149A>G c.1164A>G c.1098G>C c.1107G>C	GAA>GAG ACA>ACG CGA>CCC GAG>GAC	p.Glu354Glu p.Thr359Thr p.Arg337Pro p.Glu340Asp	No No No No	Might get lost (downstream of altered splice site) Protein feature (might be) affected Protein structure (might be) affected No protein features affected	Neutral Neutral Deleterious Neutral	Tolerated Tolerated Tolerated Damaging	Polymorphism Polymorphism Diseases causing Diseases causing
72		Low	c.1149A>G	GAA>GAG	p.Glu354Glu	No	Might get lost (downstream of altered splice site)	Neutral	Tolerated	Polymorphism
84		Low	c.1149A>G c.1112T>C	GAA>GAG GTT>GCT	p.Glu354Glu p.Val342Ala	No No	Might get lost (downstream of altered splice site) Might get lost (downstream of altered splice site)	Neutral Neutral	Tolerated Tolerated	Polymorphism Polymorphism
92		Low	c.1176_1177insA	GAA>AGA	p.Glu364Arg fsX9	Yes	Protein structure (might be) affected	Deleterious	Damaging	Diseases causing
96		Low	c.1153G>C c.1164A>G	GGC>CGC ACA>ACG	p.Gly356Arg p.Thr359Thr	No No	Might get lost (downstream of altered splice site) Protein feature (might be) affected	Deleterious Deleterious	Tolerated Damaging	Diseases causing Diseases causing

			c.1134_1134delT	CCA>CAT	p.Pro351His fsX3	Yes	Strongly truncated protein, might cause NMD	Deleterious	Damaging	Diseases causing
99		Very low	c.1149A>G c.1130_1131insT c.1107G>C	GAA>GAG GAT>AGA GAG>GAC	p.Glu354Glu p.Asp349Arg fsX13 p.Glu340Asp	No Yes No	Might get lost (downstream of altered splice site) Strongly truncated protein, might cause NMD No protein features affected	Neutral Deleterious Damaging	Tolerated Damaging Damaging	Polymorphism Diseases causing Diseases causing
117		Very low	c.1154G>C c.1161A>C c.1149A>G	GGC>GCC GGA>GGC GAA>GAG	p.Gly356Ala p.Gly358Gly p.Glu354Glu	No No No	Protein feature (might be) affected Protein feature (might be) affected Might get lost (downstream of altered splice site)	Neutral Neutral Neutral	Tolerated Tolerated Tolerated	Diseases causing Diseases causing Polymorphism
121		Low	c.1149A>G c.1171G>A c.1112T>C c.1130_1131insT	GAA>GAG GAA>AAA GTT>GCT GAT>AGA	p.Glu354Glu p.Glu362Lys p.Val342Ala p.Asp349Arg fsX13	No No No Yes	Might get lost (downstream of altered splice site) Might get lost (downstream of altered splice site) Might get lost (downstream of altered splice site) Strongly truncated protein, might cause NMD	Neutral Deleterious Neutral Deleterious	Tolerated Damaging Tolerated Damaging	Polymorphism Diseases causing Polymorphism Diseases causing
125		Low	c.1138_1138delC c.1153G>C c.1171G>A	CCA>CAT GGC>CGC GAA>AAA	p.Pro351His fsX3 p.Gly356Arg p.Glu362Lys	Yes No No	Truncated protein, might cause NMD Might get lost (downstream of altered splice site) Might get lost (downstream of altered splice site)	Deleterious Deleterious Deleterious	Damaging Tolerated Damaging	Diseases causing Polymorphism Diseases causing
8	Exon 7	Low	c.933A>T	GAA>GAT	p.Glu282Asp	No	No protein features affected	Deleterious	Damaging	Diseases causing
15		Low	c.826A>T	AGT>TGT	p.Ser276Cys	No	Might get lost (downstream of altered splice site)	Neutral	Tolerated	Polymorphism
23		Low	c.826A>T	AGT>TGT	p.Ser276Cys	No	Might get lost (downstream of altered splice site)	Neutral	Tolerated	Polymorphism
40		Low	c.868C>T	CCT>TCT	p.Pro290Ser	No	Protein features (might be) affected	Deleterious	Damaging	Diseases causing
43		Low	c.841A>T	AGT>TGT	p.Ser281Cys	No	Protein features (might be) affected	Deleterious	Damaging	Diseases causing
79		Normal	c.826A>T	AGT>TGT	p.Ser276Cys	No	Might get lost (downstream of altered splice site)	Neutral	Tolerated	Polymorphism
89		Normal	c.826A>T	AGT>TGT	p.Ser276Cys	No	Might get lost (downstream of altered splice site)	Neutral	Tolerated	Polymorphism
104		Low	c.868C>T	CCT>TCT	p.Pro290Ser	No	Protein features (might be) affected	Deleterious	Damaging	Diseases causing
115		Low	c.857C>T	TCA>TTA	p.Ser286Leu	No	Protein features (might be) affected	Deleterious	Damaging	Diseases causing
117		Very low	c.826A>T c.857C>T	AGT>TGT TCA>TTA	p.Ser276Cys p.Ser286Leu	No	Might get lost (downstream of altered splice site) Protein features (might be) affected	Neutral Deleterious	Tolerated Damaging	Polymorphism Diseases causing
122	Exon 5	Low	c.841A>T	AGT>TGT	p.Ser281Cys	No	Protein features (might be) affected	Deleterious	Damaging	Diseases causing
1		Low	c.735G>T	GGG>GTG	p.Gly216Gly	No	Might get lost (downstream of altered splice site)	Neutral	Tolerated	Diseases causing
41		Low	c.744C>T	CTC>CTT	p.Leu219Leu	No	No abrogation of potential splice sites	Neutral	Tolerated	Diseases causing
85		Low	c.744C>T	CTC>CTT	p.Leu219Leu	No	No abrogation of potential splice sites	Neutral	Tolerated	Diseases causing
90		Normal	c.704T>A	TTT>TAT	p.Phe206Tyr	No	No protein features affected	Deleterious	Damaging	Diseases causing
105	Exon 4	High	c.682C>A c.717G>A	CTG>AUG GGG>GGA	p.Leu199Met p.Gly211Gly	No	Protein features (might be) affected Might get lost (downstream of altered splice site)	Neutral Neutral	Damaging Tolerated	Diseases causing Diseases causing
16		Low	c.574G>A c.604G>A c.615G>A c.641T>G	GAA>AAA GAA>AAA ATG>ATT ATG>AGG	p.Glu163Lys p.Glu173Lys p.Met176Ile p.Met185Arg	No	Protein features (might be) affected Might get lost (downstream of altered splice site) Protein features (might be) affected Protein features (might be) affected	Neutral Deleterious Neutral Deleterious	Tolerated Damaging Tolerated Damaging	Polymorphism Diseases causing Diseases causing Diseases causing
37		Normal	c.608C>T c.621C>T	TCA>TTA TTC>TTT	p.Ser174Leu p.Phe178Phe	No	Might get lost (downstream of altered splice site) No abrogation of potential splice sites	Deleterious Neutral	Tolerated Tolerated	Diseases causing Diseases causing
41		Low	c.574G>A	GAA>AAA	p.Glu163Lys	No	Protein features (might be) affected	Neutral	Tolerated	Polymorphism

			c.600T>G c.615G>A c.643T>G c.654_655insGG	ATT>ATG ATG>ATT TCT>GCT CAG>GGG	p.Ile171Met p.Met176Ile p.Ser186Ala p.Gln190Gly fsX63	Yes	Might get lost (downstream of altered splice site) Protein features (might be) affected Protein features (might be) affected Protein structure affected might be caused NMD	Neutral Neutral Neutral Deleterious	Damaging Tolerated Tolerated Damaging	Diseases causing Diseases causing Diseases causing Diseases causing
70		High	c.571G>A c.574G>A c.641T>G	GAT>AAT GAA>AAA ATG>AGG	p.Asp162Asn p.Glu163Lys p.Met185Arg	No	Protein features (might be) affected Protein features (might be) affected Protein features (might be) affected	Neutral Neutral Deleterious	Tolerated Tolerated Damaging	Diseases causing Polymorphism Diseases causing
85		Low	c.574G>A c.577A>G	GAA>AAA AGA>GGA	p.Glu163Lys p.Arg164Gly	No	Protein features (might be) affected Might get lost (downstream of altered splice site)	Neutral Neutral	Tolerated Tolerated	Polymorphism Diseases causing
91		Low	c.641T>G	ATG>AGG	p.Met185Arg	No	Protein features (might be) affected	Deleterious	Damaging	Diseases causing
96		Low	c.641T>G	ATG>AGG	p.Met185Arg	No	Protein features (might be) affected	Deleterious	Damaging	Diseases causing
Mutation detected in matched non-cancer tissues samples										
8	Exon 9	-	c.1149A>G c.1130_1131insT	GAA>GAG GAT>AGA	p.Glu354Glu p.Asp349Arg fsX13	No Yes	Might get lost (downstream of altered splice site) Strongly truncated protein, might cause NMD	Neutral Deleterious	Tolerated Damaging	Polymorphism Diseases causing
20		-	c.1157T>A c.1188T>A c.1112T>C	ATG>AAG CTT>CTA GTT>GCT	p.Met357Lys p.Leu367Leu p.Val342Ala	No No No	Protein structure (might be) affected Might get lost (downstream of altered splice site) Might get lost (downstream of altered splice site)	Neutral Neutral Neutral	Tolerated Tolerated Tolerated	Polymorphism Polymorphism Polymorphism
84		-	c.1171G>A	GAA>AAA	p.Glu362Lys	No	Might get lost (downstream of altered splice site)	Deleterious	Damaging	Diseases causing
96		-	c.1162A>G	ACA>GCA	p.Thr359Ala	No	Might get lost (downstream of altered splice site)	Neutral	Tolerated	Polymorphism
117		-	c.1153G>C	GGC>CGC	p.Gly356Arg	No	Might get lost (downstream of altered splice site)	Deleterious	Tolerated	Diseases causing
121		-	c.1130_1131insT	GAT>AGA	p.Asp349ArgfsX13	Yes	Strongly truncated protein, might cause NMD	Deleterious	Damaging	Diseases causing

Table 2. Mutations detected in the noncoding region of *FAM134B (JK1)* in colorectal cancer

Sample ID	mRNA expression	DNA change	Mutation type	Prediction (Mutation taster)
12	Low	g.133485G>C g.133489G>T g.133544C>T g.133561T>G g.133759_133759delT	Substitution Substitution Substitution Substitution Deletion	Polymorphism Polymorphism Polymorphism Polymorphism Diseases causing
16	Low	g.133770G>A	Substitution	Diseases causing
37	Normal	g.133742_133742delT g.133510A>T g.133514A>T	Deletion Substitution Substitution	Polymorphism Polymorphism Polymorphism
70	High	g.133762A>G	Substitution	Diseases causing
85	Low	g.135916_135916delA g.139047_139048insC g.139063_139063delT g.135871C>T	Deletion Insertion Deletion Substitution	Polymorphism Polymorphism Diseases causing Polymorphism
1	Low	g.135871CG>TT	Substitution	Polymorphism
8	Low	g.135916_135916delA g.136084G>A g.138930_138930delA g.139042T>G g.139115C>A	Deletion Substitution Deletion Substitution Substitution	Polymorphism Polymorphism Polymorphism Polymorphism Polymorphism
41	Low	g.135871CG>TT	Substitution	Polymorphism
90	Normal	g.135916_135916delA g.135963C>A g.136094G>T	Deletion Substitution Substitution	Polymorphism Diseases causing Polymorphism
92	Low	g.135916_135916delA g.135963C>A g.136071G>A	Deletion Substitution Substitution	Polymorphism Diseases causing Polymorphism
105	High	g.135963C>A g.136071G>A	Substitution Substitution	Diseases causing Polymorphism
124	Low	g.136117C>G g.136122G>C	Substitution Substitution	Polymorphism Diseases causing
A	High	g.135916_135916delA g.136132G>A	Deletion Substitution	Polymorphism Polymorphism
15	Low	g.139067_139068insG g.138869G>T	Insertion Substitution	Polymorphism Polymorphism
23	Low	g.139063_139063delT	Deletion	Diseases causing
103	Low	g.138930_138930delA g.139041T>C g.138877T>A g.138914C>G g.138931C>A	Deletion Substitution Substitution Substitution Substitution	Diseases causing Polymorphism Polymorphism Polymorphism Polymorphism
115	High	g.139063_139063delT g.139081_139081delT g.138955C>T	Deletion Deletion Substitution	Diseases causing Polymorphism Diseases causing
117	Low	g.139081_139081delT	Deletion	Polymorphism
119	Low	g.138952C>T	Substitution	Polymorphism
123	Low	g.138930_138930delA	Deletion	Polymorphism
Mutation detected in matched non-cancer tissue sample				
8	-	g.139063_139063delT	Deletion	Diseases causing

Table 3. Correlation of *FAM134B* mutations with clinicopathological features of colorectal cancer

Features	Number	Mutation positive	Mutation negative	P-value
<u>Gender</u>				
Male	42 (47.72%)	16 (18.18%)	26 (29.54%)	0.094
Female	46 (52.3%)	25 (28.4%)	21 (23.86%)	
<u>Age</u>				
≤70	44 (50%)	23 (26.13%)	21 (23.86%)	0.196
>70	44 (50%)	18 (20.45%)	26 (29.54%)	
<u>Site</u>				
Proximal	72 (81.81%)	36 (40.9%)	36 (40.9%)	0.139
Distal	16 (18.18%)	5 (5.68%)	11 (12.5%)	
<u>Size</u>				
≤40	49 (55.68%)	19 (21.59%)	30 (34.09%)	0.076
>40	39 (44.31%)	22 (25.0%)	17 (19.31%)	
<u>Metachronous cancer</u>				
Yes	4 (4.54%)	4 (4.54%)	-	0.043
No	84 (95.45%)	37 (42.04%)	47 (53.4%)	
<u>MSI</u>				
Yes	17 (19.31%)	14 (15.9%)	3 (3.4%)	0.001
No	71 (80.6%)	27 (30.68%)	44 (50.0%)	
<u>Grade</u>				
Well	9 (10.22%)	5 (5.68%)	4 (4.54%)	0.480
Moderate	67 (76.13%)	31 (35.22%)	36 (40.9%)	
Poor	12 (13.63%)	5 (5.68%)	7 (7.95%)	
<u>Distant metastasis</u>				
Yes	20 (22.72%)	11 (12.5%)	9 (10.22%)	0.273
No	68 (77.27%)	30 (34.09%)	38 (43.18%)	
<u>Stage</u>				
<i>I</i>	10 (11.36%)	3 (3.4%)	7 (7.95%)	0.437
<i>II</i>	38 (43.18%)	20 (22.72%)	18 (20.45%)	
<i>III</i>	19 (21.59%)	7 (7.95%)	12 (13.63%)	
<i>IV</i>	21 (23.86%)	11 (12.5%)	10 (11.36%)	
<u>Disease persistence or recurrence</u>				
Yes	26 (29.54%)	15 (17.04%)	11 (12.5%)	0.132
No	62 (70.45%)	26 (26.54%)	36 (40.9%)	

Table 4. Relationship among mutations of different exons of FAM134B with clinicopathological parameters

Features	<u>Mutation in Exon 4</u>			<u>Mutation in Exon 5</u>			<u>Mutation in Exon 7</u>			<u>Mutation in Exon 9</u>		
	(+) ve	(-) ve	p-value	(+) ve	(-) ve	p-value	(+) ve	(-) ve	p-value	(+) ve	(-) ve	p-value
<u>Gender</u>												
Male	-	47.72%	0.013	2%	46%	0.544	5.68%	42.04%	0.565	9.09%	38.63%	0.482
Female	7.95%	44.31%		3%	49%		6.81%	45.45%		11.36%	40.9%	
<u>Age</u>												
≤70	3.40%	46.59%	0.500	3.4%	46.59%	0.500	7.95%	42.04%	0.260	14.7%	35.2%	0.031
>70	4.54%	45.45%		2.27%	47.72%		4.54%	45.45%		5.68%	44.31%	
<u>Site</u>												
Proximal	6.81%	75.0%	0.626	3.4%	43.18%	0.388	5.68%	40.9%	0.362	11.4%	35.2%	0.031
Distal	1.13%	6.81%		2.27%	51.13%		6.81%	46.59%		9.01%	44.31%	
<u>Size</u>												
≤40	6.81%	48.86%	0.099	2.27%	53.4%	0.392	3.4%	52.27%	0.044	9.01%	46.59%	0.209
>40	1.13%	43.18%		3.4%	40.9%		9.09%	35.22%		11.7%	32.95%	
<u>Metachronous cancer</u>												
Yes	2.27%	2.27%	0.030	1.13%	3.40%	0.212	-	4.54%	0.580	2.7%	2.7%	0.184
No	5.68%	89.7%		4.54%	90.90%		12.5%	82.95%		18.2%	77.27%	
<u>MSI</u>												
Yes	4.54%	14.77%	0.024	4.54%	14.77%	0.004	4.54%	14.77%	0.132	6.81%	12.5%	0.092
No	3.4%	77.27%		1.13%	79.54%		7.95%	72.72		13.6%	67%	
<u>Family history</u>												
Yes	2.27%	9.09%	0.179	2.27%	9.09%	0.097	-	11.36%	0.243	3.4%	7.95%	0.333
No	5.68%	82.95%		3.4%	85.22%		12.5%	76.13%		17%	71.59%	
<u>T-Stage</u>												
I & II	1.13%	11.36%	0.014	-	12.5%	0.676	1.13%	11.36%	0.828	2.27%	10.22%	0.582
III&IV	7.5%	80.68%		5.68%	81.81%		11.36%	76.13%		18.9%	69.31%	
<u>Distant metastasis</u>												
Yes	3.4%	19.31%	0.190	-	22.72%	0.266	-	22.72%	0.048	5.68%	17.0%	0.386
No	4.54%	72.72%		5.68%	71.59%		12.5%	64.77%		14.8%	62.5%	
<u>Stage</u>												
I & II	2.27%	52.27%	0.113	4.54%	50%	0.173	7.95%	46.59%	0.069	12.5%	42.0%	0.633
III&IV	5.68%	39.77%		1.13%	44.31%		4.54%	40.9%		7.95%	37.5%	

MSI- Microsatellite instability, NB- lymph-node metastasis data not shown as they are not significant for none of the exons

Supplementary Information

for

Novel *FAM134B* mutations and their clinicopathological significance in colorectal cancer

Farhadul Islam¹, Vinod Gopalan¹, Riajul Wahab¹, Katherine Ting-wei Lee¹, Md. Hakimul Haque¹, Afraa Mamoori¹, Cu-tai Lu², Robert A Smith^{1,3}, Alfred K-Y Lam^{1,*}

Supplementary data 1: List of primers used in this study for qRT-PCR and Sanger sequencing

Target Genes	Primer sequence (Forward and Reverse)	Amplicon size
<i>FAM134B</i> Exon 1	5'-CGGCACCCACACCCAGGCGCGCCC-3' 5'-GCACTGGGTCCCCGGGGCCCCG-3'	600 bp
<i>FAM134B</i> Exon 2	5'-GTTTCTGTGGCAGGAAGTAAACCC-3' 5'-GGACTAATTGGCTAATATGCCTAC-3'	281 bp
<i>FAM134B</i> Exon 3	5'-GTGTTAGAGATCTGAGCATTCCAC-3' 5'-CTTGGAATTTAGATTCCTGTCAC-3'	259 bp
<i>FAM134B</i> Exon 4	5'-CCAGAGGGTGTGGCCAACAGTAG-3' 5'-TGGAGAAATCTGACAAGCTG-3'	329 bp
<i>FAM134B</i> Exon 5	5'-GTCACCTATACCGGTCACTATAG-3' 5'-CTCATCCCCCTCTTCAAAC-3'	282 bp
<i>FAM134B</i> Exon 6	5'-GTGAAATACTGAAATGTACGTAGC-3' 5'-CTTTGAGGGAGATTAGCTTC-3'	346 bp
<i>FAM134B</i> Exon 7	5'-GAATATGAGAAATGTGGGGTAAG-3' 5'-GGAGTTTATTAGGAAGATCATTACAGC-3'	256 bp
<i>FAM134B</i> Exon 8	5'-CTGTCATTTTGGGGGTTTCATATGG-3' 5'-TGGTGGTAACATGTTATTTACCC-3'	314 bp
<i>FAM134B</i> Exon 9	5'-TGACCGACCCAGTGAGGA-3' 5'-GGGCAAACCAAGGCTTAA-3'	106 bp
<i>GAPDH</i>	5'- TGCACCACCAACTGCTTAGC-3' 5'- GGCATGGACTGTGGTCATGAG-3'	87 bp
<i>β-actin</i>	5'-ACCCACACTGTGCCCATCTACGA-3' 5'-CAGGAGGAGCAATGATCTTGATCTTC-3'	270 bp

Supplementary data 2: Correlation of *FAM134B* mRNA expression and clinicopathological features in colorectal carcinoma

Features	Number	High expression	Low expression	p-value
<u>Sex</u>				
Male	42 (47.7%)	9 (10.2%)	33 (37.5%)	0.398
Female	46 (52.3%)	12 (13.6%)	34 (38.7%)	
<u>Age</u>				
≤70	44 (50%)	11 (12.5%)	33 (37.5%)	0.500
>70	44 (50%)	10 (11.4%)	34 (38.6%)	
<u>Site</u>				
Proximal	72 (81.8%)	16 (18.2%)	56 (63.6%)	0.319
Distal	16 (18.2%)	5 (5.7%)	11 (12.5%)	
<u>Size (mm)</u>				
≤40	49 (55.7%)	14 (15.9%)	35 (39.8%)	0.182
>40	39 (44.3%)	7 (7.9%)	32 (36.4%)	
<u>Metachronous cancer</u>				
Positive	4 (4.5%)	1 (1.1%)	3 (3.4%)	0.671
Negative	84 (95.5%)	20 (22.8%)	64 (72.7%)	
<u>MSI</u>				
Positive	17 (19.3%)	4 (4.5%)	13 (14.8%)	0.622
Negative	71 (80.7%)	17 (19.3%)	54 (61.4%)	
<u>Histological grade</u>				
Well	9 (10.2%)	2 (2.3%)	7 (7.9%)	0.161
Moderate	67 (76.1%)	18 (20.5%)	49 (55.7%)	
Poor	12 (13.7%)	1 (1.1%)	11 (12.5%)	
<u>Lymph-vascular permeation</u>				
Positive	19 (21.6%)	3 (3.4%)	16 (18.2%)	0.272
Negative	69 (78.4%)	18 (20.5%)	51 (57.9%)	
<u>Distant metastasis</u>				
Positive	20 (22.7%)	3 (3.4%)	17 (19.3%)	0.228
Negative	68 (77.3%)	18 (20.5%)	50 (56.8%)	
<u>Stage</u>				
I	10 (11.4%)	1 (1.1%)	9 (10.2%)	0.520
II	38 (43.2%)	10 (11.4%)	28 (31.8%)	
III	19 (21.6%)	6 (6.8%)	13 (14.8%)	
IV	21 (23.8%)	4 (4.6%)	17 (19.3%)	
<u>FAM134B mutation</u>				
Positive	41 (46.6%)	9 (10.2%)	32 (36.4%)	0.445
Negative	47 (53.4%)	12 (13.6%)	35 (39.8%)	
<u>BRAF mutation*</u>				
Yes	14 (53.8%)	2 (7.7%)	12 (46.1%)	0.046
No	12 (46.2%)	6 (23.1%)	6 (23.1%)	
<u>Persistent disease/ Recurrence</u>				
Yes	26 (29.5%)	6 (6.8%)	20 (22.7%)	0.572
No	62 (70.5%)	15 (17.0%)	47 (53.5%)	

* *BRAF* screening done in 26 cancer samples

Supplementary data 3: Representative amplified PCR products of *FAM134B* for different exons (exon 1 to exon 9). The amplified products were run in 1.5% agarose gel and visualized with UV illuminator.

Supplementary data 4: Mutations detected in coding region of *FAM134B* (*JK1*) in colon cancer cell lines.

Cell lines	DNA sequence change	Change in Codon	Protein change	Protein features	<i>In silico</i> prediction		
					PROVEAN	SIFT	Mutation taster
FHC	c.1112T>C	GTT>GCT	p.Val342Ala	Protein features (might be) affected	Neutral	Tolerated	Disease causing
SW480	c.1149A>G c.944C>T	GAA>GAG TCG>TTA	p.Glu354Glu p.Ser286Leu	Might get lost (downstream of altered splice site) Might get lost (downstream of altered splice site)	Neutral Deleterious	Tolerated Damaging	Polymorphism Disease causing
SW48	c.1149A>G c.1112T>C c.944C>T	GAA>GAG GTT>GCT TCG>TTA	p.Glu354Glu p.Val342Ala p.Ser286Leu	Might get lost (downstream of altered splice site) Protein features (might be) affected Might get lost (downstream of altered splice site)	Neutral Neutral Deleterious	Tolerated Tolerated Damaging	Polymorphism Disease causing Disease causing
HCT116	c.1112T>C c.944C>T	GTT>GCT TCG>TTA	p.Val342Ala p.Ser286Leu	Protein features (might be) affected Might get lost (downstream of altered splice site)	Neutral Deleterious	Tolerated Damaging	Disease causing Disease causing

Supplementary data 5: Known single nucleotide polymorphisms (SNPs) detected in *FAM134B* in colon cancer

dsSNP ID	DNA change	Protein change	Protein features	Prediction (Mutation taster)
rs162848	g.138968C>T	No change	Might get lost (downstream of altered splice site)	Polymorphism
rs162849	g.139048T>C	No change	Might get lost (downstream of altered splice site)	Polymorphism
rs369276135	g.136003C>T	p.Thr208Met	Protein features (might be) affected	Disease causing
rs16868675	g.135983T>C	No change	Might get lost (downstream of altered splice site)	Polymorphism
rs373896313	g.135912C>T	No change	-	Polymorphism
rs143878016	g.135987G>A	p.Val203Met	Protein structure (might be) affected	Disease causing
rs2305279	g.136106G>A	No change	Protein features (might be) affected	Polymorphism

Supplementary data 6: Seven SNPs were detected in colorectal cancer samples. Among these variants, p.Thr208Met ([rs369276135](#)) and p.Val203Met (rs143878016) were identified as pathogenic as affect the protein structure.

Supplementary data 7: Correlation of clinical/pathological factors with *FAM134B* mutational effects in CRC. Consequences of *FAM134B* mutations such as cDNA change, protein amino-acids changes, splice site effects, intronic mutations were correlated with the clinicopathological factors of CRC.

Features	Substitution (P-value)	Insertion (P-value)	Deletion (P-value)	cDNA change (P-value)	Protein change (P-value)	Frameshift (P-value)	Splice site effect (P-value)	Intron mutations (P-value)
<u>Age</u>								
≤70	21 (0.06)	4 (0.64)	2 (0.24)	22 (0.04)	19 (0.01)	6 (0.5)	20 (0.06)	10 (0.60)
>70	13	4	-	13	13	5	12	10
<u>Gender</u>								
Male	13 (0.11)	3 (0.40)	-	13 (0.08)	11 (0.04)	4 (0.31)	12 (0.109)	8 (0.29)
Female	21	5	2 (0.27)	22	21	7	20	12
<u>Size</u>								
≤40	16 (0.14)	2 (0.07)	1 (0.69)	16 (0.09)	15 (0.15)	3 (0.04)	13 (0.04)	9 (0.20)
>40	18	6	1	19	17	8	19	11
<u>Site</u>								
Proximal	31 (0.06)	8 (0.18)	2 (0.66)	32 (0.05)	29 (0.08)	11 (0.09)	29 (0.08)	17 (0.48)
Distal	3	-	-	3	3	-	3	3
<u>Grade</u>								
Well	5 (0.33)	0 (0.19)	-	5 (0.57)	5 (0.34)	1 (0.08)	3 (0.68)	2 (0.72)
Moderate	25	6	-	25	23	6	24	16
Poor	4	2	2 (0.01)	5	4	4	5	2
<u>Metachronous cancer</u>								
Yes	4 (0.02)	2 (0.04)	-	4 (0.02)	4 (0.01)	2 (0.075)	3 (0.13)	1 (0.65)
No	30	6	2 (0.91)	31	28	9	29	19
<u>MSI</u>								
Yes	13 (0.001)	4 (0.04)	2 (0.03)	14 (0.001)	12 (0.001)	6 (0.001)	12 (0.001)	7 (0.04)
No	21	4	-	21	20	5	20	13

Supplementary data 8: Survival curve of patients with colorectal cancer with *FAM134B* mutation in exon 9 (A), exon 4 (B) and deletions in noncoding sequence (C). Patients with *FAM134B* mutations had poorer survival rates compared to that of non-mutated cancer patients.

Supplementary data 9: (A) Expression of FAM134B protein showed notable reduction in cancer cell (SW-480, SW-48 and HCT116) when compared to FHC. Pathogenic mutations (p.Ser286Leu and p.Val342Ala) in these cancer cells were detected, which may be contributed to the reduction of protein expression. (B) Expression of mRNA in cancer cell lines also reduced significantly (** $p > 0.01$) compared to FHC cells. *FAM134B* mRNA expression was measured as inverse ratio (*FAM134B* mRNA/internal control genes *GAPDH* or β -actin expression).

A)

B)

Supplementary data 10: In Western blot analysis, we noted that the protein expression of *FAM134B* reduced significantly in cancers cells (SW480, SW-48 and HCT116) when compared to non-neoplastic colon epithelial cell (FHC). Detected mutations in *FAM134B* coding region might be attributed to the downregulation of protein in these cancer cell lines.