Identification of a Novel Fusion Gene, FAM174A-WWC1, in Early-Onset Colorectal Cancer: Establishment and Characterization of Four Human Cancer Cell Lines from Early-Onset Colorectal Cancers

Soon-Chan Kim*,†,‡,1, Rumi Shin§, ¶, 1, Ha-Young Seo*†,‡, Minjung Kim‡,§,#, Ji Won Park‡,§,#, Seung-Yong Jeong‡,§,# and Ja-Lok Ku*†,‡

*Korean Cell Line Bank, Laboratory of Cell Biology, Cancer Research Institute, Seoul National University College of Medicine, Seoul 03080, Republic of Korea; †Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul 03080, Republic of Korea; ‡Cancer Research Institute, Seoul National University College of Medicine, Seoul 03080, Republic of Korea; §Department of Surgery, Seoul National University College of Medicine, Seoul 03080, Republic of Korea; ¶Department of Surgery, Seoul Metropolitan Government Seoul National University Boramae Medical Center, Seoul 07061; #Division of Colorectal Surgery, Department of Surgery, Seoul National University Hospital, Seoul 03080, Republic of Korea

Abstract

Colorectal cancer (CRC) is the third most common cancer diagnosed worldwide and represents the second most common cause of all cancer-related deaths in Korea. Although epidemiological data indicate a sharp increase in the incidence of CRC among individuals older than 50 years, more than 10% of CRCs occur before reaching 50. These are known as early-onset CRCs (EOCRCs) and are likely to be suggestive of hereditary predisposition. However, known familial CRC syndromes account for only 20% of genetic aberrations of EOCRC, and the remaining 80% are still in question. Therefore, we aimed to establish reproducible biological resources and contribute to expand the mutation database of EOCRC. Four cell lines derived from the original tumor mass of CRC patients diagnosed under age 30 years were established, and next-generation sequencing technique was used to identify the genetic features of EOCRC. We have identified one novel fusion gene, FAM174A-WWC1, and analyzed its functional role. The induction of FAM174A-WWC1 to normal fibroblast caused alternations in cellular morphology as well as intercellular expression of E-cadherin and N-cadherin. Moreover, WWC1 carrying the fused FAM174A domain not only abrogated the membrane expression of YAP1 but also significantly increased the levels of nucleic YAP1. As a result, the FAM174A-WWC1 expression increased the oncogenic capacity and invasiveness of normal fibroblasts, which suggests its role as a potential driver mutation of EOCRC.

Introduction

Colorectal cancer (CRC) is the third most common cancer diagnosed worldwide and is the second most common cause of cancer deaths in Korea [1,2]. Although epidemiological data indicate the incidence of CRC sharply rises among individuals older than 50 years, more than 10% of CRCs occur before age 50. Such early-onset CRC (EOCRC) may suggest hereditary predisposition [3–5]. However, known familial CRC syndromes account for only 20% of genetic aberrations of EOCRC, and the remaining 80% are still poorly known [5].
Epidemiological factors such as obesity, diabetes mellitus, and diet have been proposed as potential risk factors but do not fully explain the development of EOCRCs [6,7]. The clinicopathologic features of EOCRCs are known to be poorly differentiated, carrying mucinous and signet ring features, and are diagnosed at advanced stages, which partially reflect the aggressive behavior of EOCRC. Based on their distinct clinicopathological characteristics, previous studies suggested the creation of a unique molecular subgroup for classification of these tumors [8].

Structural chromosome rearrangements such as translocation, interstitial deletion, or chromosomal inversion may result in the formation of a chimeric gene from previously separate genes [9]. Such hybrid gene often generates a novel protein with functional capabilities distinct from those of the parental genes. Especially, the specificity of gene fusion in neoplastic tissues and their diverse functionalities during carcinogenesis provide critical clues for disease mechanisms underlying tumorigenesis, enabling tumor subclassification [10]. Since gene fusion is closely associated with specific tumor phenotype, which is derived from the unique chimeric domains, they represent ideal targets for anticancer treatment and risk stratification [11]. Although the effect of translocation involving coding and regulatory DNA sequences has been intensely studied in chronic myeloid leukemia [12,13] and lung cancer [14,15], its role in EOCRC has yet to be elucidated.

In this study, four cell lines derived from patients with EOCRC were established and analyzed via next-generation sequencing to identify molecular and genetic features. We also have identified one novel fusion established and analyzed via next-generation sequencing to identify specific gene fusion in neoplastic tissues and their diverse functionalities during carcinogenesis provide critical clues for disease mechanisms underlying tumorigenesis, enabling tumor subclassification [10]. Since gene fusion is closely associated with specific tumor phenotype, which is derived from the unique chimeric domains, they represent ideal targets for anticancer treatment and risk stratification [11]. Although the effect of translocation involving coding and regulatory DNA sequences has been intensely studied in chronic myeloid leukemia [12,13] and lung cancer [14,15], its role in EOCRC has yet to be elucidated.

In this study, four cell lines derived from patients with EOCRC were established and analyzed via next-generation sequencing to identify molecular and genetic features. We also have identified one novel fusion gene, FAM174A_WWC1, and analyzed its functional role.

Materials and Methods

Establishment of Cell Lines and Cell Culture Condition

Cell lines were established from pathologically proven colorectal carcinomas. Detailed procedure was described previously [16]. The locations and stages of original tumors of newly established EOCRC cell lines are listed in Table 1.

DNA Fingerprinting Analysis

Total DNA was isolated from cell pellet by using QIAamp DNA Mini Kit (Qiagen, Hilden, Germany), and RNA was extracted using the TRizol (Life Technologies, CA, USA) and RNaseq Plus Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer’s protocol. For cDNA synthesis, detailed procedure was described previously [17]. The STR profiles of established EOCRC cell lines are listed in Table 2.

Growth Properties and Morphology In Vitro

Suspensions of 5×10⁵ cells were seeded to 35-mm tissue culture dishes containing a culture medium. The number of cells was counted in triplicate at 24-hour intervals for 10 days using WST-1 assay. The doubling time of the cells was calculated from the growth phase. Mycoplasma contamination was tested by the 16S-rRNA-gene-based polymerase chain reaction (PCR) amplification method using e-Myc Mycoplasma PCR Detection Kit (Intron Biotechnology, Kyonggi, Korea).

Nucleic Acid Isolation and Complementary DNA Synthesis

Genomic DNA was extracted from the cell lines using QIAamp DNA Mini Kit (Qiagen, Hilden, Germany), and RNA was extracted using the TRizol (Life Technologies, CA, USA) and RNaseq Plus Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer’s protocol. For cDNA synthesis, detailed procedure was described previously [17].

Whole Exome Sequencing Analysis

Total DNA was isolated from cell pellet using QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer’s protocol. Exonic sequences were enriched in samples using the Illumina Exome Enrichment protocol, and the captured libraries were sequenced using Illumina HiSeq 2000 Sequencers (Illumina, CA, USA). The reads were mapped to the human genome assembly UCSC hg19 using Burrows-Wheeler Alignment (BWA version 0.5.8, http://bio-bwa.sourceforge.net/). Single nucleotide polymorphisms and small insertions/deletions were called with SAMtools (version 0.1.7, http://samtools.sourceforge.net/). The annotated variants were filtered against the variations reported in dbSNP132, the 1000 Genomes project November 2010 edition, and the NHLBI Exome Sequencing Project (http://evs.gs.washington.edu/EVS/) databases. Computational tools SIFT [18] and PolyPhen-2 [19] were used to predict the impact of missense mutations, while MaxEntScan [20]

---

**Table 1. Clinicopathologic Characteristics of Four EOCRC Cell Lines**

| Cell Line | Growth Pattern | Doubling Time (h) | Cell Morphology | Sex | Age | Family History | Degree of Differentiation | Stage | Pathology | Location | Recurrence |
|-----------|----------------|-------------------|-----------------|-----|-----|----------------|--------------------------|-------|-----------|----------|------------|
| NCC-375   | Adherent       | 35                | Fusiform        | F   | 24  | None           | M/D                      | T3N2M1| adc       | Rectal cancer with liver and lung metastasis | Recur. |
| SNU-1460  | Adherent       | 48                | Polygonal       | M   | 25  | None           | M/D                      | T4aN0| adc       | S-colon cancer | NED       |
| SNU-1826  | Adherent       | 45                | Polygonal       | M   | 15  | None           | P/D                      | T4aN0| adc       | Cecal cancer | NED       |
| SNU-2446  | Adherent/floating | 82            | Round/Polygonal | F   | 21  | None           | T4bN2                    | Mucinous carcinoma | D-colon cancer with p-seeding | PD       |

GF, grandfather; adc, adenocarcinoma; M/D, moderately differentiated; P/D, poorly differentiated.
and NNSplice (http://www.fruitfly.org/seq_tools/splice.html) were used to predict the effect of splice variants.

**RNA Sequencing Analysis**

Total RNA was isolated from cell lysate using TRIzol (Qiagen, Hilden, Germany) and Qiagen RNeasy Kit (Qiagen, Hilden, Germany). Sequencing libraries were prepared using the Illumina TrueSeq Stranded Total RNA Library Prep Kit. Fifty-one million reads were obtained from the cell lysates. Following base-calling and alignment with the Tuxedo Suite, the rejected reads were analyzed using FusionMap, ChimeraScan, and Defuse with default parameters for RNA and alignment to GRCh37.72. The output was filtered to include in-frame fusions, with at least one rescued read and two unique seed reads and excluded known and recurrent artifacts. Sequencing of RNA isolated from EOCRC cell lines was used to identify RNA sequences spanning two different regions in a chromosome. The in-frame fusion with the highest read score spanned the FAM174A gene on chromosome 5q21.1 and WWCl on chromosome 5q34. The exon RPKM track displays RPKM calculated for each exon and where the library size normalization factor was considered as the number of reads aligned to exons alone. The fusion read spanning exon 2 in FAM174A and exon 10 in WWCl was also confirmed with IGV.

**Verification of FAM174A-WWC1 in Cancer Cell Line Sample**

Based on previously described gene fusion isoforms involving the FAM174A and WWCl genes, the oligonucleotide primers were designed to encompass breakpoint junctions of FAM174A-WWC1 using Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primerblast) (FAM174A Forward: GTC AGG ACG GTC AGG ATG AGA, WWC1 Reverse: GCT TCC TCC AGT TCT CTC ACA A). The PCR cycling conditions were: 94°C for 5 minutes; followed by 35 cycles of 94°C for 30 seconds, 55°C for 45 seconds, and 72°C for 1 minute; and a final extension of 72°C for 10 minutes using Intron PCR Kit (Intron). Using primers complimentary to the sequence in exon 2 of FAM174A and exon 10 of WWCl, reverse transcription followed by polymerase chain reaction (RT-PCR) resulted in a product with a sequence suggesting in-frame junction.

**Cloning of FAM174A-WWC1**

After confirming the FAM174A-WWC1 fusion gene by RT-PCR, cloning primers were designed to amplify the complete coding sequence of the FAM174A-WWC1 gene fusion. (FAM174A forward: CAC CAT GAA GGC CTC GCA GTG CTG C, WWCl reverse: GAC GTC ATC TGC AGA GAG AGC TGG GAT). The PCR cycling conditions were: 94°C for 5 minutes, followed by 3 cycles of 94°C for 30 seconds and 72°C for 3 minutes, and a final extension of 72°C for 10 minutes using Q5 Hot Start High-Fidelity Taq (NEB). The amplified FAM174A-WWC1 sequence was cloned into a pENTR/SD/D/TOPO vector according to the manufacturer’s recommendations. The fusion junction was verified by Sanger sequencing. The entry vector was then subcloned into a plenti6.2/V5-DEST lentiviral expression vector according to the manufacturer’s instructions.

**Immunocytochemistry**

Cells were seeded on chambered cover glass with a desirable cell confluence. The chambered cover glass was designed to be hydrophilic, with no ECM component treated before seeding. Cells were washed three times with cold DPBS for 15 minutes, 72 hours after cell seeding. The cells were fixed and permeabilized with BD Cytofix/Cytoperm. Cells were washed with the washing solution and treated with DPBS containing 2% FBS for an hour for blocking. After cells were washed with cold DPBS, they were incubated with KIBRA and V5 Tag antibody diluted in 0.05% of PBST for 1.5 hours at room temperature. Subsequently, cells were washed with 0.05% of PBST, followed by incubation with Alexa 488 and Alexa 568 secondary antibodies diluted in 0.05% of PBST for 1 hour at room temperature, and an additional 30 minutes with 1× DAPI and Rhodamine-conjugated Phalloidin diluted in distilled water. The cells were washed three times with DPBS and imaged under an LSM800 Confocal Laser Scanning Microscope equipped with ZEN software to analyze the cells. The intensity of each channel was fixed to compare the target protein expression between the samples. Digital resolution, scan speed, and the average number of images were set to 1024 × 1024, 40 seconds per one channel, and 8, respectively. The images were focused at the very bottom of the fixed cells to investigate the protruding regions of the cell.

**2D and 3D Migration Assay**

The 2D Gap closure migration assay was performed using a Cytoselect 24-Well Wound Healing Assay kit according to the manufacturer’s protocol. Briefly, the NIH3T3 Control and NIH3T3 FAM174A-WWC1 cells were diluted to 2 × 10^5 cells/ml in RPMI1640 treated with 10% of FBS and 1% penicillin streptomycin. Once the inserts provided in the assay kit were placed into each well of the 24-well tissue culture plate, 500 μl of previously diluted cell solution was dispensed to each side of the insert to obtain a total cell solution of 1 ml per each well. After 24 hours of incubation at 37°C in an atmosphere of 5% CO2 and 95% air, inserts were removed. Immediately, cells were washed with complete media twice. The wound closure was recorded under phase contrast microscope at 24-hour intervals for 5 days. The closed wound area was calculated with ImageJ version 1.8.0. The whole procedures were performed in triplicate, and the average closed wound area was analyzed using GraphPad Prism version 7.00 for Windows.

The 3D migration assay was performed using the AIM 3D cell culture chips according to the manufacturer’s protocol. Briefly, cells were diluted to 2 × 10^5 cells/ml in RPMI1640 without FBS. The cell solution was mixed with a reduced growth factor basement membrane matrix with 1:1 ratio, and 10 μl of the mixture was filled into a gel inlet channel. After incubation for 30 minutes in a 37°C incubator, the media channels were hydrated with 120 μl of RPMI1640 mixed with 10% of FBS and 1% of penicillin streptomycin, followed by further incubation for 24 hours at 37°C in an atmosphere of 5% CO2 and 95% air. Visual inspection and images were obtained with a phase contrast microscope using the 4× objective with 24-hour intervals. Images were analyzed using ImageJ version 1.8.0. The entire procedures were performed in triplicate, and the average closed wound area was analyzed using GraphPad Prism version 7.00 for Windows.

**3D Spheroid Formation Assay and 3D Spheroid Cell Invasion Assay**

The 3D spheroid formation assay was performed using Cultrex 96-Well 3D Spheroid BME Cell Invasion Assay kit according to the manufacturer’s protocol. Briefly, the reduced growth factor basement membrane matrix was thawed on ice overnight in a 4°C refrigerator.
HEK293 cell aggregates were dissociated into a single cell suspension with trypsin. Cells were diluted to $1 \times 10^6$ cells/ml in RPMI1640 combined with 10% of FBS and 1% penicillin streptomycin. The cell solution was mixed with the reduced growth factor basement membrane matrix, and 50 μl of the solution was dispensed per well of the 3D culture–containing 96-well spheroid formation plate (provided). The plate was centrifuged at 200×g for 3 minutes at room temperature in a swinging bucket rotor and incubated at 37°C in a tissue culture incubator for 72 hours to promote spheroid formation. After 24 hours, in ice, 50 μl of invasion matrix (provided) was added per well of the 96-well spheroid formation plate. The plate was centrifuged at 300×g at 4°C for 5 minutes in a swinging bucket rotor to eliminate bubbles and position spheroids within the invasion matrix towards the middle of the well. The plate was transferred to a tissue culture incubator set at 37°C for 1 hour to promote gel formation of the invasion matrix. The images were obtained after 72 hours using a phase contrast microscope equipped with the 4× objective.

The 3D spheroid cell invasion assay was conducted by replacing the reduced growth factor basement membrane matrix with 10× Spheroid Formation ECM (provided) to artificially induce the spheroid formation. We also substituted the 96-well spheroid formation plate for 96 hang-in-drop plate to retain cell aggregation. Once the spheroid was formed, it was transferred to a 96 U-bottom plate with 50 μl of invasion matrix (provided). After 1 hour, 100 μl of warm (37°C) cell culture medium was added containing chemoattractant. The spheroid in each well was photographed every 24 hours for 5 days under phase contrast microscope using the 4× objective. Images were analyzed using ImageJ version 1.8.0. All the procedures were performed in triplicate, and the intact form of the spheroid and the rate of disseminated cells were analyzed with circularity estimation using Plugin from GraphPad Prism version 7.00 for Windows.

Results

General Characteristics of the Cell Lines

On in vitro cultivation, three cell lines (NCC-375, SNU-1460, and SNU-1826) were cultured as monolayers of substrate-adherent cells, and one cell line (SNU-2446) formed floating and adherent aggregates. The majority of tumor cells displayed a polygonal shape and had round-to-oval nuclei with prominent single-to-double nucleoli (Figure 1). Each cell line was passaged at least three times prior to characteristic analysis. Population doubling times ranged from 35 to 82 hours. A summary of clinicopathologic information is provided in Table 1. All cell lines were confirmed to be free of bacterial and mycoplasma contamination (Supplementary Figure 1). Fifteen tetranucleotide repeat loci and the gender-determining marker amelogen were heterogeneously distributed in each cell line, without cross-contamination (Table 2).

Figure 1. Phase-contrast microscopy of new established early onset colorectal cancer cell lines. On in vitro cultivation, three cell lines (NCC-375, SNU-1460, and SNU-1826) grew as monolayer of substrate-adherent cells, and one cell line (SNU-2446) formed floating and adherent aggregates. The majority of tumor cells displayed a polygonal shape and had round-to-oval nuclei with prominent single-to-double nucleoli.
**Genomic Analyses**

Due to the uncharacteristically aggressive clinical course and in the absence of the prototypical FAM174A-WWC1 fusion, the histological diagnosis was in question. To more comprehensively characterize the molecular basis of disease, all the cell lines were subjected to whole exome and transcriptome analysis. The cell lines were primarily derived from pathologically proven tumor tissue. In order to integrate whole exome sequencing data from EOCRC patients with a larger database, the top 100 mutated genes in COAD cohorts under age 35 were extracted from the TCGA-GDC data portal. Intersection between gene lists from TCGA-GDC and from EOCRC patients was sorted, and the mutations that affected the protein structure were screened. The PCLO gene was mutated in all the EOCRC samples and four COAD samples obtained from TCGA-GDC. Mutations in the PCLO gene are associated with depressive disorders.[21,22] The relationship between PCLO gene and EOCRC requires further analysis. Summarized results are listed in Table 3-6.

**Verification of FAM174A-WWC1 in Patient Sample**

The novel fusion gene, FAM174A-WWC1, was detected in NCC-375 cell line using three different fusion gene detection algorithms. The reads spanning the breakpoint of the fusion gene were visualized with an IGV program, which revealed a 67,892,563-bp deletion in chromosome 5 (Figure 2A). Oligonucleotide primers were designed to encompass breakpoint junctions of FAM174A and WWC1. Using primers complimentary to the sequence in exon 2 of FAM174A and exon 10 of WWC1, RT-PCR of the tumor derived cell line resulted in a product whose sequence verified the in-frame junction (Figure 2B). Amplicons were detected in other cell lines and confirmed as nonspecific PCR products by Sanger sequencing. The amplicon containing the target fusion gene was verified with direct Sanger sequencing which showed that the exon 2 of FAM174A was connected to the exon 10 of WWC1 (Figure 2C).

**Cloning and In Vitro Assessment of FAM174A-WWC1**

We employed RT-PCR to clone the entire coding sequence of the FAM174A-WWC1 fusion transcript using RNA isolated from the NCC-375 cell line. In silico analysis of the FAM174A-WWC1 coding sequence (http://web.expasy.org/compute_pi/) predicted the expression of a ~123-kDa protein. Examination of the WWC1-encoding sequence revealed that the truncation of WWC1 resulted in the loss of its WW domains and preservation of its C2 and PDZ binding motif. The truncation of the FAM174A portion in FAM174A-WWC1 resulted in conserved exons 1-2 encoding the DUF1180 domain at which the function is unknown (Figure 3A). Murine fibroblast NIH3T3 cells and HEK293 cells were transduced with packaged lentivirus encoding FAM174A-WWC1. The expression of the fusion transcript was verified by RT-PCR with primers spanning the fusion breakpoint (Figure 3B). The expression of fusion protein was assessed by Western blotting using a V5-tag antibody, which demonstrated a protein at the predicted molecular weight of 100 kDa (Figure 3C). The induction of FAM174A-WWC1 caused morphological alternation in both HEK293 and NIH3T3 cells. The morphology of both cells was changed from stellate to round and polygonal (Figure 3D). Induction of FAM174A-WWC1 decreased the protein expression of E-cadherin and augmented N-cadherin in HEK293 cell lines as well.

### Table 3. Shared Mutation in NCC-375 and TGCA-GDC Dataset

| GENE | CHR | POS | ID       | REF | ALT | AA_CHANGE |
|------|-----|-----|----------|-----|-----|-----------|
| MUC4 | 3   | 195,518,112 | rs142781032 | T   | TGTCTCTCTGGTAACA | p.T113MLRRRP |
| FNDC1 | 6  | 159,660,779 | rs141435210 | GCCCACCCCGCGCGCGCA | G  | p.ATTRRTT1471A |
| PCLO | 7   | 82,581,488 | rs10630259 | T   | TTCA | A  | p.E2927YK |
| SYNE2 | 14 | 64,560,092 | rs2781377 | G   | A   | p.W4001* |

### Table 4. Shared Mutation in SNU-1460 and TGCA-GDC Dataset

| GENE | CHR | POS | ID       | REF | ALT | AA_CHANGE |
|------|-----|-----|----------|-----|-----|-----------|
| FNDC1 | 6  | 159,660,779 | n141435210 | G   | GCCCACCGCGCGCGCAGA | A  | p.ATTRRTT1471A |
| THSD7A | 7  | 11,871,469 | rs10630259 | T   | TTCA | G  | p.35RC |
| PCLO | 7   | 82,581,488 | rs10630259 | C   | GGGCGCTGCT | T   | p.E2927YK |
| NCO2 | 12 | 124,824,721 | rs61519723 | G   | GGGCGCGAG | A   | p.5115 |
| ZFHX3 | 16 | 72,821,593 | rs374146547 | A   | AGCGCGCGC | C   | p.535RC |
| ZFHX3 | 16 | 72,822,563 | rs374146547 | T   | TTGCTGCTGCT | T   | p.Q204RSK |
| TP53 | 17 | 7,578,370 | C          | T   | n/a |

### Table 5. Shared Mutation in SNU-1826 and TGCA-GDC Dataset

| GENE | CHR | POS | ID       | REF | ALT | AA_CHANGE |
|------|-----|-----|----------|-----|-----|-----------|
| MUC4 | 3   | 195,518,112 | rs142781032 | T   | TGTCTCTCTGGTAACA | p.S113MLRRRP |
| FNDC1 | 6  | 159,660,804 | rs3842694 | GCCCACCGCGCGCGCA | G  | p.ATTRRTT1471A |
| THSD7A | 7  | 11,871,469 | rs35831183 | G   | GGGCGCGAG | A   | p.35RC |
| PCLO | 7   | 82,581,493 | rs10694231 | C   | GGGCGCGAG | T   | p.5115 |
| FRMD4A | 10 | 13,698,938 | rs58913700 | T   | TGTCTCTCTGGTAACA | p.35RC |
| NCO2 | 12 | 124,834,739 | rs61519723 | C   | GGGCGCGAG | T   | p.5115 |
| ZFHX3 | 16 | 72,821,593 | rs374146547 | A   | AGCGCGCGC | C   | p.535RC |
| ZFHX3 | 16 | 72,822,563 | rs374146547 | T   | TTGCTGCTGCT | T   | p.Q204RSK |
| RTTN | 18  | 67,863,854 | rs58913700 | -   | TGTCTCTCTGGTAACA | p.35RC |
E-cadherin was expressed in NIH3T3 cell lines, and only decreased E-cadherin was observed (Figure 3E). Subcellular fractional protein extraction indicated that the fusion of WW1 with DUF1180 domain increased the levels of both total YAP1 and phosphorylated YAP1 in the nucleus (Figure 3F). The augmented levels of total YAP1 (Figure 3G) and phosphorylated YAP1 (Figure 3H) in the cytosolic and nuclear fractions were quantitatively analyzed with statistical significance.

**Table 6.** Shared Mutation in SNU-2446 and TGCA-GDC Dataset

| GENE | CHR | POS         | ID          | REF | ALT   | AA_CHANGE |
|------|-----|-------------|-------------|-----|-------|-----------|
| TTN  | 2   | 179,472,617 | .           | C   | A     | p.E17633* |
| PCLO | 7   | 82,581,488  | rs10694231  | T   | TTCA  | p.E2927DE |
| WDFY4| 10  | 50,165,232  | .           | C   | A     | p.S2679*  |

**Figure 2.** Verification of FAM174A-WWC1 in Patient Sample. (A) The novel fusion gene, FAM174A-WWC1, was detected in NCC-375 cell line with three different fusion gene detection algorisms. The reads spanning the breakpoint of the fusion gene were visualized with IGV program. (B) Oligonucleotide primers were designed to encompass breakpoint junctions of FAM174A and WWC1. (C) The amplicon containing the target fusion gene was verified with direct Sanger sequencing which represented that the exon 2 of FAM174A is connected to the exon 10 of WWC1.
Oncogenic Capacity of FAM174A-WWC1

The SNV profiling of EOCRC patients implicated oncogenic potential of the fusion gene. The tumor expressing FAM174A-WWC1 carried the lowest number of oncogene mutations (Figure 4A). To further elucidate the tumorigenic capacity of FAM174A-WWC1, cells were grown under conditions of anoikis, as tumorspheres, to assay their tumor-initiating capability and stem-like properties. FAM174A-WWC1 expression significantly increased the tumorsphere propagation (Figure 4B).

Metastatic Potential of FAM174A-WWC1

To determine the role of FAM174A-WWC1 in cell migration and invasion, cells were seeded into a narrow channel with a basement membrane matrix containing reduced growth factors, which mimics

Figure 3. Cloning and in vitro assessment of FAM174A-WWC. (A) Mimetic diagram showing the structure of FAM174A-WWC1. The expression of the fusion transcript was verified by RT-PCR (B) with primers spanning the fusion breakpoint and by Western blotting (C) using an V5-tag antibody, demonstrating a protein at the predicted molecular weight of 100 kDa. (D) The induction of FAM174A-WWC1 caused morphologic alternation in both HEK293 and NIH3T3 cells. (E) Induced FAM174A-WWC1 decreased protein expression of E-cadherin and augmented N-cadherin in HEK293 cell lines as well. (F) Subcellular fractional protein extraction indicated that WWC1 fused with DUF1180 domain increased both total YAP1 and phosphorylated YAP1 level in nucleus. (G) Quantitative analysis of expression of YAP1 in cytoplasm, membrane, and nuclear. (H) Quantitative analysis of expression of YAP1 (p-S127) in cytoplasm, membrane, and nuclear.
extravasation of tumor cells. Compared with controls, cells expressing FAM174A-WWC1 significantly increased the number of invaded cells (Figure 5, A and B). Increased ability to invade into a membrane matrix was re-confirmed with three-dimensional spheroid invasion assay. Cell were grown as 3D cyst-like forms and seeded with invasion matrix. Consistent with previous findings, the FAM174A-WWC1 expression promoted dissemination of cells into the surrounding matrigels (Figure 5, C and D). Accordingly, FAM174A-WWC1 expression also enhanced cell migration in two dimensions (Supplementary Figure 2).

Metastatic Potential of FAM174A-WWC1 in Focal Adhesion and ECM Receptor Pathway

The intercellular localization of WWC1 and FAM174A-WWC1 was analyzed with immunocytochemistry. WWC1 and FAM174A-WWC1 were mostly colocalized in the cytoplasmic and cytoskeletal region. Nevertheless, the expression of fusion protein was mostly centered in the cytoskeletal region, which may indicate the role of fusion protein in focal adhesion and ECM binding (Figure 6A). Signal intensity analysis revealed the expression of both WWC1 and FAM174A-WWC1 in lamellipodia, with prominent FAM174A-WWC1 expression at the edge (Figure 6, B and C). The subcellular compartment of FAM174A-WWC1 was further investigated with fractional protein extraction. As in preceding, the fusion protein was mostly detected in the cytoskeletal fraction (Figure 6D). To identify genes and pathways affected by the fusion protein, RNA sequencing was performed. Consistent with the prior findings, focal adhesion and ECM receptor pathways were significantly inhibited by FAM174A-WWC1 fusion protein. Correctively, these findings suggest the potential role of fusion protein in movement (Figure 6E).

Discussion

We present four cases of EOCRC. The cases did not harbor any of the fusion oncogenes reported previously in EOCRC; however, it was found to express a novel fusion gene involving FAM174A and WWC1 in one of the patients enrolled in this study. Sequence verification and cloning of the FAM174A-WWC1 fusion gene facilitated for functional characterization of the fusion protein. We have used HEK293, human embryonic kidney cells, and NIH3T3, mouse embryonic fibroblasts, which are normal cell lines to pinpoint the tumorigenic effect of the FAM174A-WWC1. We demonstrated that the expression of FAM174A-WWC1 in HEK293 and NIH3T3 cells was sufficient for cellular transformation.

Various fusion genes including BRAF, NTRK3, RAS, and RET have been reported in CRC [23,24]; however, no study specifically elucidated gene fusions in EOCRC. WWC1 is associated with several other fusion genes. For instance, the interchromosomal in-frame fusion gene WWC1-ADRBK2 was found in BRCA-mutated breast cancer [25]. Yuri Choi et al. identified an intrachromosomal in-frame fusion gene, BOD1-WWC1, in CRC patient tissue [26]. Multiple fusions involving the WWC1 gene may be triggered by the instability of chromosome 5q arm. Chromosome 5q deletions have been associated with biological and prognostic significance of myeloid malignancies [27] and triple-negative breast cancers [28]. The FAM174A-WWC1 gene fusion resulted from a 67,892,563-bp deletion in the 5q arm, which may also reflect the loss of chromosome 5q arm in EOCRC case (Figure 2A). Due to this recurrent instability, fusion genes may carry various breakpoints. Thus, we constructed five different primer sets targeting different exons in FAM174A and WWC1 based on the identified break junction, and no variable
breakpoint existed other than exon 2 of FAM174A and exon 10 of WWC1 (data not shown). Analysis of the FAM174A-WWC1 sequence revealed that the WWC1 portion of FAM174A-WWC1 maintains the entire C2 domain and PDZ bind motif of WWC1, and lacks the WW domain (Figure 3A). WWC1 has been suggested as a key regulator of the hippo pathway, regulating organ size, cell contact inhibition, as well as tissue regeneration and tumorigenesis. WWC1 activates both large tumor suppressor (LATS1/2) kinases and Yes-associated protein (YAP) via phosphorylation on the hydrophobic motif [29,30]. Our results indicated that FAM174A-WWC1 not only lowered the membrane level of YAP1 but also significantly increased its level in the nucleus (Figure 3F), which demonstrated that the replacement of WW domains with DUF1180 domain is associated with nuclear localization of YAP1. Phosphorylation at Ser127 by LATS kinases promotes YAP translocation from the nucleus to the cytoplasm, where it is sequestered through association with 14-3-3 proteins [31]. The nucleic p-S127 YAP1 level was higher in HEK293 cells containing FAM174A-WWC1 (Figure 3F), which may imply that the fusion protein hinders the cytoplasmic localization of phosphorylated YAP1.

The SNV profiling of EOCRC patients suggested the oncogenic potential of the fusion gene. The tumor expressing FAM174A-WWC1 carried the lowest number of oncogene mutations (Figure 4A), which was concurrent with a previous study suggesting tumors harboring fusion genes tended to exhibit fewer oncogene mutations [23]. Nevertheless, cancer development in patients with FAM174A-WWC1 may be influenced by oncogene and tumor suppressor gene mutations involving APC, in addition to the fusion gene. To further determine whether FAM174A-WWC1 was involved in tumorigenic capacity, cells were grown under conditions of anoikis as tumor-spheres to examine their tumor-initiating capability and stem-like properties since sphere-forming efficiency indicates tumorigenic potential [32]. FAM174A-WWC1 expression significantly increased the tumorsphere propagation (Figure 4B). It was reported that WW1/2 domains of WWC1 mediated the inhibition of tumorsphere formation [28]. The FAM174A-WWC1 not only lacks WW1/2 domains of WWC1 but also contains DUF1180 domain of FAM174A replacing WW domains, which may demonstrate that the DUF1180 domain is associated with accelerated tumor progression.
The patient harboring FAM174A-WWC1 developed multiple metastases to lungs and liver after surgery (Table 1). Two dimensional and 3D invasion assays were performed for clinicopathological analysis. Without exception, cells expressing FAM174A-WWC1 showed increased invasiveness and dissemination compared with controls (Figure 5 and Supplementary Figure 2). These data support a role for FAM174A-WWC1 in promoting metastatic dissemination, which was inconsistent with the role of naive WWC as a metastatic suppressor [28]. This finding may imply again that the replacement of WW domains with DUF1180 domain reverses the primary role of WWC1. Collectively, these data indicate that the expression of FAM174A-WWC1 augments the tumorigenic and metastatic capacity and, therefore, that its expression can confer significant advantages to EOCRC.

In conclusion, we identified and analyzed one novel fusion gene, FAM174A-WWC1, from a patient with EOCRC based on comprehensive and integrated clinicopathological analysis, NGS data, and in vitro assessments. We demonstrated the oncogenic and metastatic capacity of FAM174A-WWC1 through in vitro analysis. This fusion gene may represent a robust molecular and therapeutic target in EOCRC. Although cases of EOCRC carry an overall unfavorable prognosis, identification of young patients diagnosed with EOCRC may represent a population of high-risk cases, with tumors driven by novel genes such as FAM174A-WWC1. It would be interesting to determine the prevalence of FAM174A-WWC1 in archival samples of EOCRC and to correlate FAM174A-WWC1 status with clinical outcome.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tranon.2019.05.019.

Conflicting Interests
The authors declare that there is no conflict of interest.

Funding
This research was supported by the Korean Cell Line Research Foundation, the Priority Research Centers Program (2009-0093820) and the Basic Science Research Program (2012R1A1A3010709) through the National Research Foundation (NRF) of Korea funded by the MSIP. The first author received a scholarship from the BK21-plus education program provided by the NRF.

Ethical Approval
All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Guarantor
None.

Acknowledgement
None.

References
[1] Siegel RL and Miller KD (2016). Jemal A (2016). Cancer statistics. CA Cancer J Clin 66, 7–30.
[2] Jung KW, Won YJ, Kong HJ, and Lee ES (2018). Cancer statistics in Korea: incidence, mortality, survival, and prevalence in 2015 Cancer research and treatment : official journal of Korean Cancer Association 50, 303–316.
