The role of AT rich interactive domain 3A

in the tumorigenesis of colorectal carcinomas

Meiying Song

Department of Medical Science

The Graduate School, Yonsei University

The role of AT rich interactive domain 3A

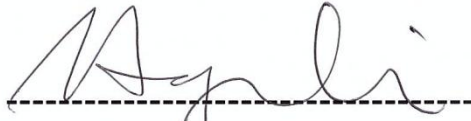# in the tumorigenesis of colorectal carcinomas

Directed by Professor Hoguen Kim

The Doctoral Dissertation
submitted to the Department of Medicine,
the Graduate School of Yonsei University
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Meiying Song

December 2012

This certifies that the Doctoral
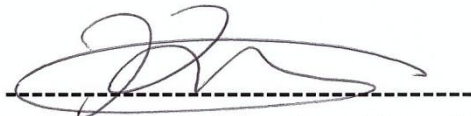Dissertation of Meiying Song is approved.

Thesis Supervisor: Hoguen Kim

Thesis Committee Member: Nam Kyu Kim

Thesis Committee Member: Tae Il Kim

Thesis Committee Member: Jae Yong Cho

Thesis Committee Member: Hyunki Kim

The Graduate School
Yonsei University

December 2012

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to all those who made invaluable contributions to my research directly or indirectly. Without their invaluable helps and generous encouragements, this thesis could not have reached its present form.

I am deeply indebted to my supervisor Professor Hoguen Kim, and Dr. Hyunki Kim working in Department of Pathology, Yonsei University. Their invaluable help, stimulating suggestion, constant encouragement, and patient guidance helped me in all the time of my research.

My colleagues from Department of Pathology, Yonsei University also gave me lots of valuable advice when I was faced with problems during the process of my research and writing this thesis. I want to thank them for all their helps.

Especially, I should give my thanks to my parents. It was their patient love and encouragement that enabled me to finish my doctorate in education at medical college of Yonsei University.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The role of AT rich interactive domain 3A
in the tumorigenesis of colorectal carcinomas

Meiying Song

Department of Medical Science
The Graduate School, Yonsei University

(Directed by Professor Hoguen Kim)

AT rich interactive domain 3A (ARID3A) is a member of the ARID family of DNA-binding proteins. Previous reports have shown that ARID3A controls the cell growth through p53-dependent manner. Recently, it has been reported that expression of ARID3A protein was 14.29-fold increased in colon cancer tissue, compared to matched normal colonic mucosa. To date, biological function of ARID3A in human disease, especially in colorectal cancer (CRC), remains largely unknown. Therefore, my research was focused on studying the role of ARID3A in the tumorigenesis and the physiological implication of CRC.

The expression of ARID3A was investigated by immunohistochemistry. ARID3A expression was detected in a subset of colorectal adenomas and carcinomas, and the location of ARID3A was mainly in the nucleus. To investigate the prognostic impact of ARID3A in CRC, tissue specimens from 690 patients with CRC were examined. Of the 690 cases, 195 tumors were strong-positive for ARID3A, 187 tumors were weak-positive and 308 tumors were negative. The expression of ARID3A in CRC

was significantly correlated with age, degree of differentiation, depth of invasion, lymph node metastasis, distant metastasis, TNM stage, status of microsatellite instability, and CEA levels. The overall survival of CRC patients with ARID3A-strong expression was significantly longer than that of patients with ARID3A-negative or weak expression. On multivariate analysis, the strong expression of ARID3A was proven to be an independent predictor for better prognosis in CRC.

Recent studies support a model in which ARID3A acts as a suppressor of lineage plasticity. I found that upregulation of ARID3A expression correlated with the reduction of transcriptional levels of several pluripotency-associated markers (*OCT4, SOX2* and *KLF4*) and colorectal cancer stem cell (CSC) markers (*CD133, CD44, CD166, CD24,* and *ALDH1*) in CRC cell lines. Among these markers, I demonstrated the inverse relationship between ARID3A and CD133 expression in protein levels.

In conclusion, my study strongly suggests that ARID3A might play an important role in colorectal carcinogenesis and can be used as a biomarker to predict the prognosis of CRC. Additionally, upregulation of ARID3A in CRC cell lines reduces expression levels of CD133, one of most important cancer stem cell markers in CRC. Further studies are necessary to delineate the mechanistic basis of these features.

The role of AT rich interactive domain 3A

in the tumorigenesis of colorectal carcinomas

Meiying Song

Department of Medical Science

The Graduate School, Yonsei University

(Directed by Professor Hoguen Kim)

I. INTRODUCTION

Colorectal cancer (CRC) is the third most common cancer in the Western world. Every year, more than one million new cases are diagnosed as colorectal cancer worldwide, and colorectal cancer is also the third most common leading cause of cancer-related death in the developed countries [1].

The classic description of colorectal carcinogenesis is the adenoma-carcinoma sequence. Most colorectal cancers develop through an ordered series of events beginning with the transformation of normal colonic epithelium to adenoma and then ultimately adenocarcinoma [2]. Multiple genetic events are required for tumor progression, and genomic instability is recognized as an essential genetic feature that accompanies the acquisition of these mutations [3]. Colorectal cancer is classified into at least three pathways of genomic and epigenomic instability: the chromosomal instability (CIN), microsatellite instability (MSI), and CpG island methylator phenotype (CIMP) pathways [3]. Most colorectal cancers arise

sporadically. About 85% of sporadic cases have chromosomal instability; an allelic imbalance at some chromosomal loci (5q, 8q, 17p, and 18q), and chromosome amplification and translocation, which together contribute to tumor aneuploidy [1, 2, 4, 5]. The remaining 15% of sporadic colorectal cancer have microsatellite instability-high (MSI-high) phenotypes [6]. This phenotype is caused by mutation or loss of function through epigenetic gene silencing of DNA mismatch-repair genes [1, 7, 8].

Prognostic markers are associated with survival, and predictive markers indicate likely benefit of treatment. Microsatellite instability and 18q imbalance are very important prognostic and predictive markers in colorectal cancer. These two markers can be used to discriminate between molecular subtypes in stage II CRC, which contribute to the risk-benefit assessment of adjuvant treatment [1]. Even though cancer biomarkers, such as microsatellite instability and 18q imbalance, have been identified, the overall 5-year survival rate of colorectal cancer is still poor. Therefore, many efforts have been made to find other molecular markers to identify high-risk disease and to select patients for adjuvant treatment.

Recently, Kang *et al.* performed proteome profiling of microsatellite stable (MSS) colorectal cancer using stable isotope labeling and mass spectrometry. 1487 proteins showed different expression levels (higher than 2-fold) between normal colonic mucosae and carcinoma tissues. As one of these proteins, ARID3A was identified and quantified by mTRAQ labeling method. The expression ratio in cancer *vs.* normal tissue was 14.29- fold for ARID3A, which means the expression levels of ARID3A is increased in CRC tissues as compared with the matched normal tissues [9]. To date, the role of ARID3A in human disease, especially in colorectal cancer, remains largely unknown.

ARID3A/BRIGHT/DRIL1/E2FBP1 is a member of the AT rich interactive domain (ARID) family of DNA-binding proteins. ARID proteins bind to A/T rich regions of DNA. There are 15 human ARID family members, which share a highly

conserved DNA-binding domain. The ARID consensus sequence spans about 100 amino acid residues, and structural studies identify the major groove contact site as a modified helix-turn-helix motif [10]. ARID family proteins are implicated in development, tissue-specific gene expression, chromatin remodeling, and cell cycle regulation [10, 11]. The human ARID family can be divided into seven subfamilies based on the degree of sequence identify between individual members [12]. Human ARID3 subfamily contains three members - ARID3A, ARID3B, and ARID3C. These members have 80% sequence similarity among the ARID domain, but are not closely related outside their DNA-binding domains [10]. The ARID3 subfamily contains an extended area of conserved 40 amino acids outside the core DNA-binding domain. This "extended ARID" sequence facilitates their sequence-specific DNA binding [11].

The two proteins in which AT-rich interactive domain was originally defined are murine Bright (B cell regulator of immunoglobulin heavy chain transcription) and *Drosophila* Dri (dead ringer). Bright is a B cell-specific transactivator cloned in a search for proteins binding to immunoglobulin heavy-chain matrix-associating regions [11]. Matrix attachment regions are AT-rich sequences, and the Bright protein was indeed found to bind preferentially to AT-rich DNA sequences [13]. At the same time, the *Drosophila* Dri was cloned in a search for novel proteins associating with homeobox domains. Homeobox domains are also AT-rich sequences, and the Dri protein was likewise found to bind preferentially to AT-rich DNA sequences in a similar oligonucleotide selection and enhancement protocol [14]. This novel DNA binding domain was designated ARID, based on the shared features of Bright and Dri. The derivation of a Bright/Dri consensus sequence led to the recognition of other ARID-containing proteins already cloned or subsequently added to the database [11]. Human homologue ARID3A has over 90% similarity to mouse Bright, and both of them associates with the transcription factor II-I (TFII-I) and active

Bruton's tyrosine kinase (Btk) to upregulate IgH transcription [15]. Moreover, ARID3A was cloned through homology with the *Drosophila* Dri [16], and later re-identified as a protein associated with E2F1 (E2FBP1) [17]. E2F1 is a transcription factor involved in the regulation of many cell cycle genes that are necessary for cell cycle entry into the S phase [18]. ARID3A heterodimerizes with E2F1 to transactivate an adenovirus E2 promoter, suggesting that it can associate with E2F1 to activate cell cycle target genes [17]. Overexpression of ARID3A in primary mouse embryonic fibroblasts (MEFs) overcame RAS-dependent senescence and led to immortalized MEFs. This escape from senescence was correlated with an increase of the E2F-1 target cyclin E1 [19]. This report suggests that ARID3A may be associated with oncogenic function. However, in other studies, ARID3A has been proposed to function as a tumor suppressor. ARID3A is induced by p53 and up-regulated following DNA damage caused by UV radiation or doxorubicin treatment in a manner dependent on endogenous p53. The overexpression of ARID3A induces cell growth arrest in U2OS cells expressing wild-type p53, but not Saos-2 cells lacking p53 [20]. Additionally, ARID3A cooperates with p53 to transcriptionally activate *p21^{WAF1}*, a p53-target gene important for cell-cycle arrest [21]. To date, whether ARID3A acts as oncogene or tumor suppressor remains controversial.

Recently, one report shows that ARID3A-deficient cells expressed several lineage markers and transcription factors associated with pluripotency. Moreover, knockdown of ARID3A directly converts a transformed human epithelial line into clones that could spontaneously differentiate *in vitro* into cells expressing markers of all three germ line lineages. Together, these data support a model in which ARID3A acts as a suppressor of lineage plasticity [22]. Furthermore, ARID3A is a direct target of the oncomir *microRNA-125b* in progenitor B-cells. In B-cell acute lymphoblastic leukemia (B-ALL) *miR-125b* is overexpressed and this microRNA targets ARID3A and maintains the progenitor characteristics of pre-BI cells by

controlling the expression of pluripotency-associated factors. The repression of ARID3A could lead to inhibition of differentiation of pre-BI cells by preventing the transcription of immunoglobulin heavy-chain genes [23].

Aberrant expression of ARID proteins is increasingly implicated in human tumorigenesis [11]. In ARID3 subfamily, ARID3B has been reported to involve in the progression of malignant neuroblastoma, breast cancer and ovarian cancer [24-26], whereas the role of ARID3A in cancer tumorigenesis and development has not been well characterized. Peeper *et al.* reported that they observed ARID3A protein was relatively highly expressed in a number of colon cancer cell lines, but whether ARID3A also occurs in primary CRC was not determined [19]. The purpose of this study is firstly to investigate the clinical significance and prognostic value of ARID3A expression in CRC, and secondly to analyze the biological function of ARID3A in CRC cells by testing tumor cell growth after up- or down-regulating its expression. Finally, the specific aim of this study is to determine whether ARID3A correlates with the expression of pluripotency-associated genes and CSC markers in CRC cells.

II. MATERIAL AND METHODS

1. Patients and samples

My study comprised of 690 patients who underwent surgery for colorectal cancer from January 2003 to December 2006 at Severance Hospital of Yonsei University. All specimens were obtained from the archives of the Department of Pathology, Yonsei University, Seoul, Korea. Authorization for the use of these tissues for research purpose was obtained from the Institutional Review Board of Yonsei University College of Medicine. Medical records of the 690 patients provided information of age, gender, and following parameters: tumor location, tumor size, tumor histology, tumor invasion, lymph node metastasis, distant metastasis, TNM stage, microsatellite instability status, and carcinoembryonic antigen (CEA) level. The median follow-up time was 64.5 months (range, 1-95 months). The mean age at diagnosis was 60.3 years (range, 12-94 years). 425 patients (61.6%) were male. Colorectal cancer tissues and matched normal mucosal tissues were frozen immediately after resection, and stored in liquid nitrogen and kept at -80℃ until use.

2. Tissue microarray and immunohistochemistry

Sections from tissue microarray (TMA) blocks were deparaffinised, rehydrated and stained with haematoxylin-eosin. Each tumor was represented in duplicate following one core for matched normal mucosa.

Immunohistochemistry was performed using a Ventana XT automated stainer (Ventana Corporation, Tucson, AZ, USA) with antibodies against ARID3A (diluted 1:100; ProteinTech Group, Inc., Chicago, Illinois, USA), and p53 (diluted 1: 50; Dako, Denmark). Sections were deparaffinized using EZ Prep solution (Ventana Corporation). CC1 standard (pH 8.4 buffer containing Tris/borate/EDTA) was used for antigen retrieval and was blocked with inhibitor D (3% H_2O_2) for 4 min at 37℃.

Slides were incubated with primary antibody for 40 min at 37˚C followed by a universal secondary antibody for 20 min, at 37˚C. Slides were incubated in streptavidin-horseradish peroxidase (SA-HRP) D for 16 min at 37˚C and then the substrate, 3,3'-diaminobenzidine tetrahydrochloride (DAB) H_2O_2, was added for 8 min followed by hematoxylin and bluing reagent counterstaining at 37˚C.

3. Evaluation of staining

ARID3A and p53 were both predominantly expressed in the nucleus of the tumor cells. The expression of ARID3A was recorded as negative, weak-positive, and strong-positive. Immunoreactivity of p53 was categorized as negative (<5% of the tumor cells) or positive (>5% of the tumor cells). The staining was evaluated by two independent observers (Hy Kim and Me Song) who were blinded to clinical and outcome data.

4. Cell lines and cultures

Ten cell lines, including cell lines originating from normal colonic mucosa (CCD18Co) and colorectal cancer (HT29, SW1116, SW480, SW620, WiDr, HCT116, LoVo, RKO, and SNUC4) were obtained from either the American Type Culture Collection (ATCC; http://www.atcc.org) or the Korean Cell Line Bank (KCLB; http://cellbank.snu.ac.kr). Cells were grown in RPMI, DMEM, or MEM supplemented with 10% FBS (Life Technologies, Grand Island, New York, USA), penicillin, and streptomycin at 37 ℃ in a humidified 5% CO_2 environment.

5. RNA extraction and semi-quantitative RT-PCR

Total RNA was isolated from cell lines using illustra RNAspin Mini RNA Isolation Kit (GE Healthcare Life Sciences, Pittsburgh, PA, USA). Up to 5×10^6 cells for each cell lines were collected by centrifugation and lysed by buffer RA1 mixed with β-mercaptoethanol, and then the lysate was filtrated through RNAspin

Mini Filter units. Then add 70% ethanol to the lysate, and load the mixture to the RNAspin Mini column. RNA was binded to the membrand of the column after centrifugation, and DNA was digested by DNase I. After washing membrane with buffer RA2 and RA3, RNA was eluted with RNase-free H_2O.

Reverse transcription was performed with 2 μg of total RNA using M-MLV reverse transcriptase (Invitrogen). RT-PCR amplification using specific primers for ARID3A. RT-PCR conditions were 94℃ for 5 min, 30 cycles of 94℃ for 30 sec, 58℃ for 30 sec, 72℃ for 40 sec, and 72℃ for 5 min. In all PCR reactions, β-actin served as an internal control for normalization. The amplified products were run on a 2% agarose gel stained with ethidium bromide. Primer sequences are shown in Table 1.

6. Real-time RT-PCR

Real-time PCR was performed using the ABI PRISM 7500 Sequence Detector (Applied Biosystems, Foster City, CA, USA) and SYBR Premix Ex TaqII (TaKaRa, Shiga, Japan), according to instructions provided by the manufacturer. The conditions for real-time PCR were 95℃ for 30 sec, 40 cycles of 95℃ for 5 sec, 60℃ for 34 sec, and dissociation stage. The relative expression levels of target genes were determined by generating a 5-point serial standard curve. The amount of target mRNA was normalized to GAPDH mRNA.

Table 1. Sequences of PCR Primers

Gene name	Forward primer (5'-3')	Reverse primer (5'-3')
ARID3A	GCGACTGGACTTACGAGGAG	CAGCATGAACAGGTCAAGG
OCT4	AGCGAACCAGTATCGAGAAC	TTACAGAACCACACTCGGAC
SOX2	AGCTACAGCATGATGCAGGAC	GGTCATGGAGTTGTACTGCAG
KLF4	TCTCAAGGCACACCTGCGAA	TAGTGCCTGGTCAGTTCATC
CD133	GGTGCTGTTCATGTTCTCCA	ACCGACTGAGACCCAACATC
CD166	TAGCAGGAATGCAACTGTGG	CGCAGACATAGTTTCCAGCA
CD44	AAGGTGGAGCAAACACAACC	ACTGCAATGCAAACTGCAAG
CD24	TGCTCCTACCCACGCAGATT	GGCCAACCCAGAGTTGGAA
ALDH1	GTTGTCAAACCAGCAGAGCA	CTGTAGGCCCATAACCAGGA
β-actin	GATTCCTATGTGGGCGACGAG	CCATCTCTTGCTCGAAGTCC
GAPDH	AAGGTGAAGGTCGGAGTCAAC	GGGGTCATTGATGGCAACAAT

7. Western blot

Whole lysates from cell lines and tissue specimens were prepared by using passive lysis buffer (Promega, Madison, WI, USA). Protein samples were separated by SDS-PAGE and transferred to PVDF membranes. Blots were blocked with TBST containing 5% skim milk, and were incubated at 4 ℃ for overnight with primary antibodies against ARID3A (diluted 1:500, ProteinTech Group, Inc., Chicago, Illinois, USA), CD133 (diluted 1:100, Cell Signaling Technology, Inc., Danvers, MA, USA), and β-actin (diluted 1:1000, Cell Signaling Technology, Inc.). After washing, the membranes were incubated with HRP-conjugated secondary antibody (Santa Cruz Biotechnology, Inc., Santa Cruz, CA, USA) for 1 h at room temperature, washed and developed with ECL-Plus (Amersham Pharmacia Biotech, Uppsala, Sweden). Proteins were quantified densitometrically using ImageJ 1.46 program.

8. Plasmid construction

Full-length cDNAs of ARID3A was obtained by PCR amplification using human ARID3A cDNA PCR primers: forward 5'-GCAAATGCGTGAATGAGCC-3', reverse 5'-CGTATCTTCCGCCATCCTG-3'. PCR reaction was performed using TaKaRa LA Taq^{TM} with GC buffer (Takara Biotechnology Co., Ltd., Dalian, China). Temperature parameters of the PCR were as follows: 1 min at 94℃ followed by 30 cycles of 94℃ for 30 sec, 60℃ for 30 sec, 72℃ for 2 min, then a 5-min final extansion at 72℃. PCR products were electrophoresed on 1% agarose gel stained with ethidium bromide, and bands corresponding to full-length size of ARID3A products were purified from agarose gel slices by GeneAll Expin^{TM} Gel SV(Biomol GmbH, Hamburg, Germany). Then the purified cDNAs (2.1 kb) were cloned directly into a TA cloning vector using RBC T&A Cloning Vector Kit (Real Biotech Corporation, Taipei, Taiwan), and the inserts were confirmed by

sequencing.

To regulate the expression of ARID3A, I selected expression vector, p3xFLAG-CMV-10 (Sigma-Aldrich, Inc., St. Louis, MO, USA). The vector was cut by Hind III and Xba I (Fermentas, M-Medical, Milan, Italy), and the full length cDNA of ARID3A was inserted at the site. p3xFLAG-ARID3A expression vectors containing the complete coding sequence of ARID3A were constructed. The upregulation of ARID3A expression was confirmed by real-time RT-PCR or Western blot.

9. Small interfering RNA and transfection

Nonspecific oligos and small interfering RNA (siRNA) targeting ARID3A were chemically synthesized by Bioneer (Daejeon, Korea). The target sequence for siRNA was as follows: (sense) 5'-GCAGCUCUACGAACUCGACTT-3', (antisense) 5'-GUCGAGUUCGUAGAGCUGCTT-3'. siRNA was transiently transfected into cells grown in 6-well plates by using lipofectamine 2000 (Invitrogen). The inhibition of ARID3A expression by siRNA targeting was evaluated by real-time RT-PCR or Western blot.

10. Proliferation assay

Two CRC cell lines (RKO, LoVo) were selected to measure cell viability and proliferation. Cells were plated to 6-well plate one day before transfection such that cells would be 50-60% confluent at the time of transfection. 1 µg of 3xFLAG-tagged ARID3A expression vector, as well as 100 pmol of ARID3A-targetted siRNA, were transfected into the cells for upregulating and knockdown of ARID3A expression, respectively. p3xFLAG-CMV10 empty vector and scrambled siRNA were used as control. After transfection, cells were incubated for 12 hours, and cells of each well of 6-well plate were harvested by trypsinizing and centrifugation, then counted using hemocytometer. 100 µl of cells (5000 total cells) were seeded into each well of 96-well plate in triplicates for performing proliferation assays. In this

study, I chose two methods of measurements – MTT [3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide] assay (Sigma-Aldrich) and VisionBlue[TM] Fluorescence Cell Proliferation Assay Kit (BioVision, Inc., Mountain View, CA, USA). MTT (Sigma-Aldrich) powder was dissolved in PBS to make 5 mg/ml solution, and then filtered with a 0.22 mm filter (Millipore, Carrigtwohill, Ireland). Added 25 μl of 5 mg/ml MTT to each well and incubated 4 hours at 37℃ in the dark. Then aspirated media from the wells and added Dimethyl sulfoxide (DMSO) (Sigma-Aldrich) to each well. The plate was leaved at room temperature for 10 min. Finally, the absorbance at 570 nm was recorded by a VersaMax Microplate Reader (Molecular Devices, Sunnyvale, CA, USA). To validate the results of MTT assay, VisionBlue[TM] Fluorescence Cell Viability Assay Kit (BioVision) was employed to exam the proliferation at the same conditions as described above. 10 μl of VisionBlue[TM] Reagent was added into each well of 96-well plate and mix well gently. After incubating plate for 3 hours in standard culture conditions, fluorescence intensity was measured at excitation of 550 nm and emission of 600 nm. The assay was performed at 0, 24, 48, 72, 96 hr after transfection. The growth curves were set up taking into the average results obtained from three independent experiments.

11. Cell cycle and apoptosis assay

LoVo cell line was transfected with ARID3A expression vector as described in method 10. Cells were harvested by trypsinization 72 hr after transfection, washed with ice-cold PBS twice and resuspend at 2×10^6 cells/ml. To fix the cells, gently added 4.5 ml ice-cold ethanol (70%) dropwise to tube containing 0.5 ml of cell suspension in PBS while vortexing gently and leaved on ice for 2 hour. Prepared propidium iodide (PI) staining solution as followings: added 20 ul PI (Sigma-Aldrich) stock solution (1 mg/ml) and 20 ul RNase A stock solution (1 mg/ml) to 2

ml PBS. The fixed cells were washed with in ice-cold PBS twice and resuspended with 1 ml of PI staining solution to the cell pellet and mixed well, and incubated in dark at room temperature for 30 min. For cell cycle analysis, the samples were analyzed by fluorescence activated cell sorter scan (FACS) caliber (BD Biosciences, San Jose, CA, USA).

I also analyzed apoptosis of LoVo cells treated as described above using Annexin V-FITC Apoptosis Detection Kit (Biovision, Milpitas, CA, USA). On the 3rd day after transfection, 0.5×10^6 cells were collected by centrifugation and resuspended in 500 μl of 1x Binding Buffer. Then added 5 μl of Annexin V-FITC and 5 μl of propidium iodide (50 μg/ml) to the sample, and incubated at room temperature for 5 min in the dark. Apoptosis was analyzed by FACS caliber (BD Biosciences).

12. Statistical analysis

Statistical calculations were performed with SSPS version 18.0 for Windows (SPSS Inc., Chicago, IL, USA). To analyze the association between ARID3A and each clinicopathological parameter, the Pearson Chi-square test was performed for comparison of categorical variables between ARID3A-negative/weak group and ARID3A-strong group. The student's *t*-test was used to analyze continuous data. Differences were considered significant at $p<0.05$. Overall survival analysis was performed using the Kaplan–Meier method, and the difference between groups was assessed using the log-rank test. Multivariate survival comparisons were carried out using Cox proportional hazard regression models. Estimated relative risks of dying were expressed as adjusted hazard ratios (HR) and corresponding 95% confidence intervals (95% CI).

III. RESULTS

1. Quantification and identification of ARID3A by proteomic analysis.

ARID3A was identified and quantified by mTRAQ labeling method, ARID3A showed the highest increasing ratio 14.29-fold difference, which led us to test their diagnostic potential in a follow-up study. ARID3A protein was identified by a single unique peptide match. The sequence of the peptide derived from ARID3A was AAPDEDREPESAR. This peptide was identified with XCorr value 3.412 and deltaCN 0.175 for the +3 charged ion. It was heavy-mTRAQ labeled and the TPP protein probability was 0.9981. Figure 1A shows an MS/MS spectrum of the precursor at 529.5893 m/z, which was matched to the above-mentioned peptide. ARID3A protein was quantified by three spectra. The spectra, two from the same run, the other one was from different run. The average ratio (L/H) of spectra was 0.07 and standard deviation was 0.01 (Figure 1B).
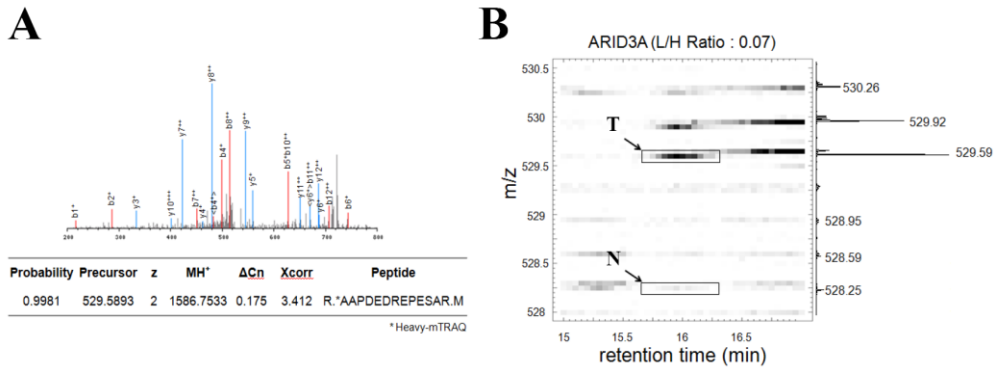
Figure 1. Quantification and identification of ARID3A. (A) MS analysis output identifying ARID3A protein. MH+, theoretical mass for the singly charged molecular ion; ΔCn, delta correlation; Xcorr, cross-correlation score. Tandem mass spectrum for the selected precursor ion is shown below. Spectrum was generated using the Trans Proteomic Pipeline. (B) Contour map of a single LC-MS run, quantification of ARID3A protein through the mTRAQ-labeled peptide AAPDEDREPESAR. Up arrow was heavy labeled peptide with high abundance intensity and down arrow was light labeled peptide with low abundance intensity.

2. Detection of ARID3A expression in colorectal cancer cell lines and tissues

The expression of ARID3A was analyzed by RT-PCR and western blot in the nine CRC cell lines (HT29, SW1116, SW480, SW620, WiDr, HCT116, LoVo, RKO and SNUC4) and one normal colon cell line CCD-18Co. Endogenous ARID3A was detected in all cell lines, and the expression levels were different among cell lines, relatively lower in HT29, WiDr, HCT116 and CCD-18Co than that in other cell lines (Figure 2A and B).

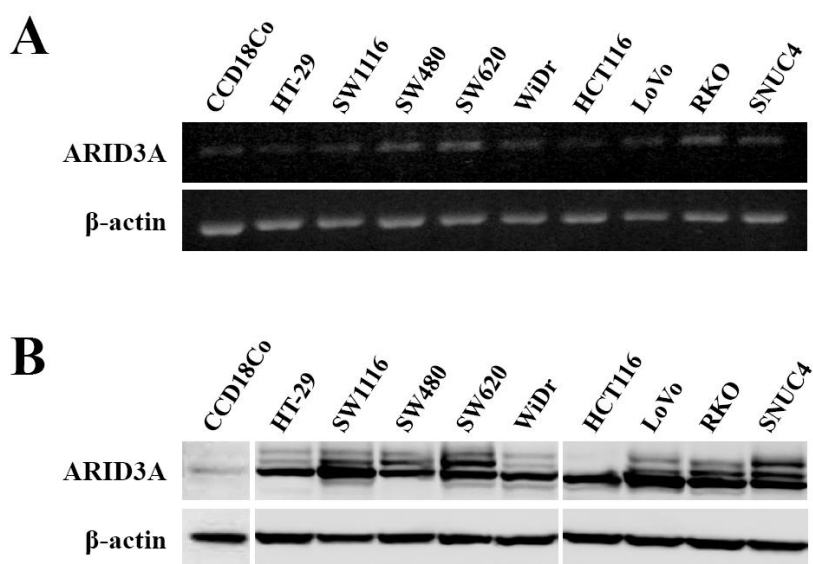**Figure 2.** ARID3A expression in colorectal cancer cell lines. RT-PCR analysis of ARID3A mRNA (A) and western blot analysis of ARID3A protein (B) expression in a normal colon fibroblast cell line, CCD18CO, and nine CRC cell lines. β-actin was used as an internal control.

To evaluate the expression status of ARID3A in CRC tissues, firstly I tested its expression in eight frozen CRC specimens using western blot. Comparative analysis indicated that ARID3A expression levels were markedly or weakly higher in some CRC samples than those in the adjacent normal tissues, whereas in the other CRC specimens ARID3A levels were downregulated compared with the adjacent normal tissues (Figure 3A). The result suggested that ARID3A expression was increased in a subgroup of patients with CRC. Then the expression was further validated in paraffin-embedded CRC specimens using immunohistochemistry (IHC). Typically observed expression patterns of ARID3A are shown in Figure 3B. Normal mucosa was negative for ARID3A (Figure 3B, a). In CRC tissues, ARID3A protein was observed predominantly in the nucleus of cancerous cells. In keeping with the expression patterns showed in Figure 3A, the specimens were classified as negative group, weak-positive group, and strong-positive group according to ARID3A expression. (Figure 3B, b-d)
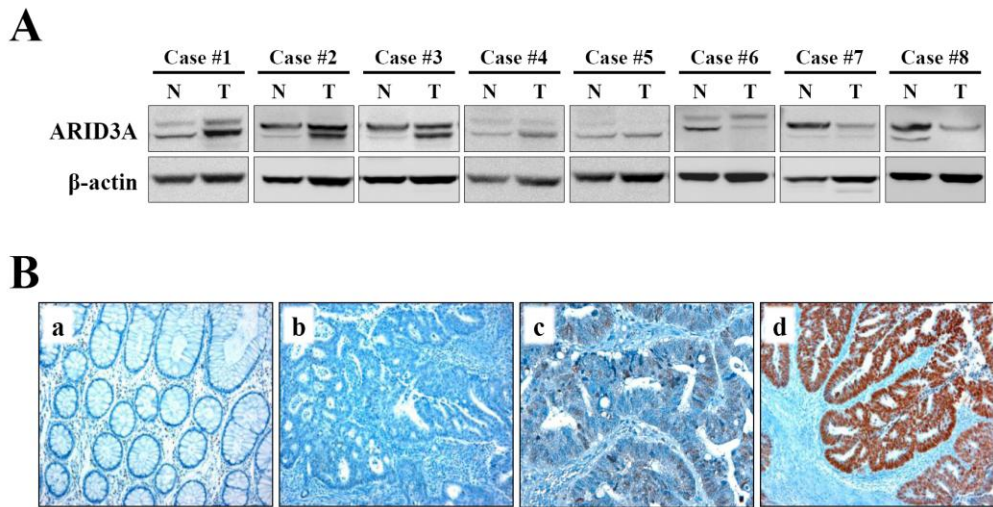
Figure 3. ARID3A expression in colorectal cancer tissues. (A) Western blot of whole tissue lysates from eight frozen specimens showing ARID3A expression in carcinomas (T) and adjacent normal tissues (N) paired from the same patient. β-actin served as an internal control. (B) Representative immunohistochemical staining of ARID3A in normal colorectal mucosa and carcinoma specimens. (a) Negative expression of ARID3A in normal mucosa. (b-d) Negative expression (b), weak expression (c), and strong expression (d) of ARID3A in carcinomas (localized in the cell nucleus). Original magnification, x200.

3. Immunohistochemical analysis of ARID3A expression in colorectal adenomas and carcinomas

To determine whether the expression of ARID3A is an early or late molecular event in colorectal tumorigenesis, the frequency of ARID3A expression was examined in 43 matched adenomas and carcinomas from the same patient by immunohistochemistry and western blot. ARID3A expression was detected in a subset of adenoma tissues, and representative images were showed in Figure 4A-B. In colorectal adenomas, 5 (11.6%) cases were negative for ARID3A, 28 (65.1%) cases and 10 (23.3%) cases showed weak-positive and strong-positive expression of ARID3A, respectively. In CRC, 5 (11.6%) cases were negative, 15 (34.9%) cases and 23 (53.5%) cases displayed weak-positive and strong-positive expression of ARID3A, respectively (Figure 4C). The frequency of strong expression of ARID3A in CRC was significantly higher than that in adenomas. In light of these findings, I concluded that the expression of ARID3A was associated with CRC tumorigenesis and progression.

A

Normal mucosae | Adenomas | Carcinomas

N1 N2 N3 | A1 A2 A3 | T1 T2 T3

B

Case No. 1 2 3

N A T N A T N A T

ARID3A

β-actin

C

□ Negative □ Weak □ Strong

% of cases

	Adenoma	Carcinoma
Strong	23.3	53.5
Weak	65.1	34.9
Negative	11.6	11.6
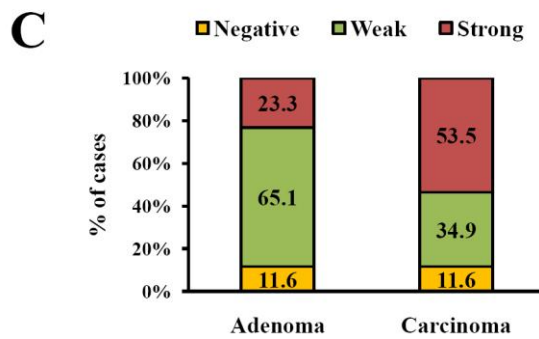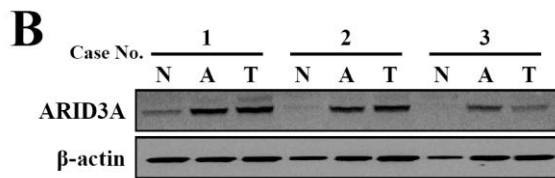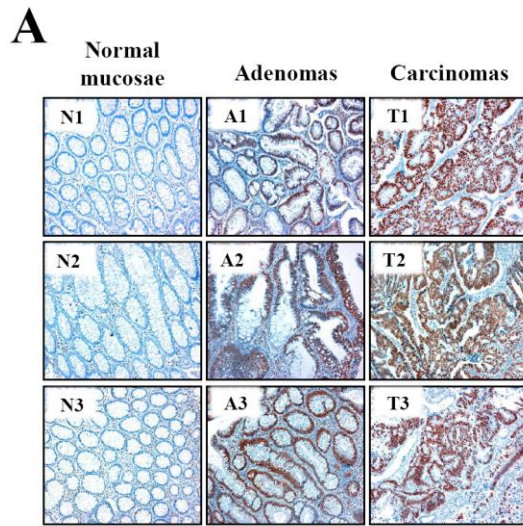
Figure 4. Representative images of ARID3A expression in three matched sets of normal, adenoma, and carcinoma tissues. Immunohistochemistry (A) and western blot (B) analysis showed that the expression of ARID3A in matched adenomas and carcinomas were upregulated, whereas ARID3A expression was negative in all normal mucosae. (C) The frequency of ARID3A expression was examined in 43 matched adenomas and carcinomas from the same patient by immunohistochemistry. N, normal mucosae; A, adenomas; T, carcinomas. Original magnification, x200.

4. Relationship between ARID3A expression and clinicopathological parameters

To analyze the clinical impact of ARID3A on the patients with CRC disease, ARID3A expression was evaluated in 690 CRC tissues represented in the TMA by immunohistochemistry. Of the 690 cases, 195 tumors (28.3%) were strong-positive for ARID3A, 187 tumors (27.1%) were weak-positive and 308 tumors (44.6%) were negative. To investigate the relationship between ARID3A and p53 status in CRC, the expression of p53 was tested by immunohistochemistry simultaneously. Among the 690 tumors, 358 tumors (51.9%) were positive for p53 expression and 332 tumors (48.1%) were negative for p53 expression.

The specimens were further grouped into two categories according to ARID3A expression: negative or weak-positive group and strong-positive group. Clinicopathological parameters related to ARID3A expression status were shown in Table 2. The ARID3A expression in CRC tissues was significantly correlated with age (P=0.019), and degree of differentiation (P<0.001). The frequency of ARID3A expression was higher in well (29.1%) and moderately (31.6%) differentiated CRC cases than those in poorly (2.7%) differentiated CRC cases. ARID3A expression was also inversely correlated with depth of invasion (P=0.001), lymph node metastasis (P=0.006), distant metastasis (P=0.004), TNM stage (P<0.001), and CEA level (P=0.003). Moreover, there was significant correlation between ARID3A and microsatellite instability status (P<0.001), ARID3A strong-positive cases were more frequently found in MSS and MSI-low CRC cases than those in MSI-high cases. Finally, statistically significant positive correlations were found between protein expression of ARID3A and p53 (P=0.045). There was no significant association between ARID3A expression and gender, tumor location or size.

Table 2. Clinicopathological features according to ARID3A expression in 690 colorectal cancers

Characteristics	Total no. (n=690)	ARID3A Nuclear Expression		P-value
		Negative / Weak (%) 495 (71.7)	Strong (%) 195 (28.3)	
Age (years)		59.69 ± 11.70	61.77 ± 9.94	**0.019**
Gender				
Male	425	299 (70.4)	126 (29.6)	0.306
Female	265	196 (74.0)	69 (26.0)	
Location				
Colon	388	283 (72.9)	105 (27.1)	0.428
Rectum	302	212 (70.2)	90 (29.8)	
Size				
≤ 5 cm	379	270 (71.2)	109 (28.8)	0.748
> 5 cm	311	225 (72.3)	86 (27.7)	
Differentiation				
Well	79	56 (70.9)	23 (29.1)	**< 0.001**
Moderate	538	368 (68.4)	170 (31.6)	
Poor	73	71 (97.3)	2 (2.7)	
T stage				
T1	11	5 (45.5)	6 (54.5)	**0.001**
T2	84	51 (60.7)	33 (39.3)	
T3	570	418 (73.3)	152 (26.7)	
T4	25	21 (84.0)	4 (16.0)	
N stage				
N0	362	246 (68.0)	116 (32.0)	**0.006**
N1	203	148 (72.9)	55 (27.1)	
N2	125	101 (80.8)	24 (19.2)	
M stage				
M0	595	415 (69.7)	180 (30.3)	**0.004**
M1	95	80 (84.2)	15 (15.8)	
TNM stage				
I	75	41 (54.7)	34 (45.3)	**< 0.001**
II	265	189 (71.3)	76 (28.7)	
III	255	185 (72.5)	70 (27.5)	
IV	95	80 (84.2)	15 (15.8)	
MSI status				
MSS/MSI-low	538	360 (66.9)	178 (33.1)	**< 0.001**
MSI-high	70	62 (88.6)	8 (11.4)	
CEA level				
≤ 5 ng/ml	461	314 (68.1)	147 (31.9)	**0.003**
> 5 ng/ml	229	181 (79.0)	48 (21.0)	
p53 expression				
Negative	332	250 (75.3)	82 (24.7)	**0.045**
Positvie	358	245 (68.4)	113 (31.6)	

5. Evaluation of ARID3A as a potential prognostic marker for colorectal cancer

At the last follow-up, 174 of 690 patients (25.2%) had died from CRC and 516 of 690 patients (74.8%) remained alive or lost follow up. A statistically significant difference of overall survival was found between CRC patients with tumors showing strong ARID3A expression and tumors with negative or weak ARID3A expression. Kaplan-Meier survival curves showed that the overall survival of CRC patients with ARID3A-strong expression was significantly longer than that of patients with ARID3A-negative or weak expression (P=0.002; Figure 5A). The favorable prognostic effect of the strong ARID3A expression was also found in the subgroup analysis using stage II and III CRC (P=0.039; Figure 5B). To determine whether p53 status affects the clinical impact of ARID3A in CRC, I divided 690 patients into p53-negative group and p53-positive group, which represent p53 wild-type and p53 mutation, respectively. In p53-negative group, patients with ARID3A strong-positive showed better prognosis compared with ARID3A-negative or weak group (P=0.003; Figure 5C), whereas in the p53-positive group, there was no significant correlation between ARID3A expression and survival rates (P=0.101; Figure 5D). These results suggested that the effect of ARID3A on the survival rate of CRC might differ by p53 status.
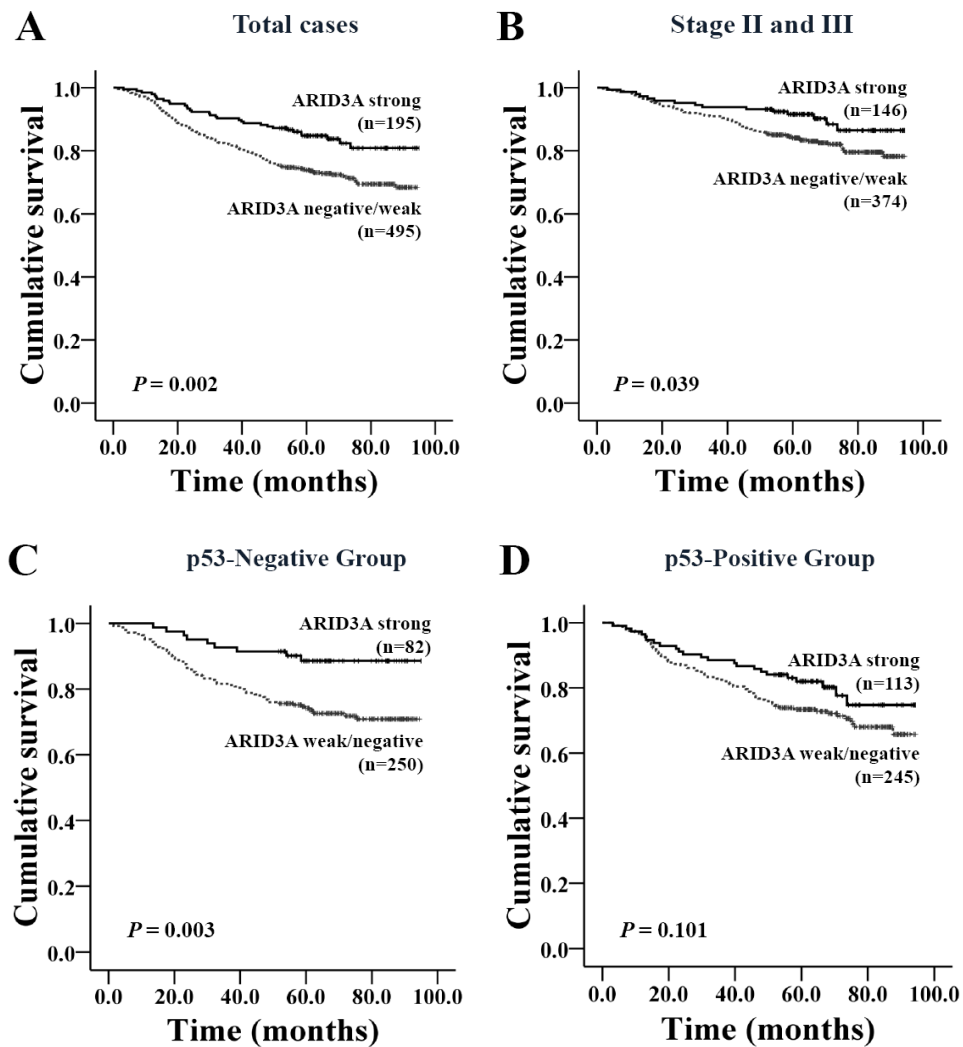
Figure 5. Kaplan-Meier analysis of survival in patients with colorectal cancer. Overall survival rates of all CRC patients (A), and CRC patients with stage II and III disease (B) according to ARID3A expression. (C) Correlation of ARID3A expression with overall survival among patients with tumors of negative p53 expression. (D) Correlation of ARID3A expression with overall survival among patients with tumors of positive p53 expression.

Univariate analysis based on clinicopathologic features showed that differentiation grade (*P*=0.028), depth of invasion (*P*=0.013), lymph node metastasis (*P*<0.001), distant metastasis (*P*<0.001), TNM stage (*P*<0.001), CEA level (*P*<0.001) and the expression status of ARID3A (*P*=0.002) were statistically significant risk factors affecting the overall survival of CRC patients (Table 3).

Four clinicopathological factors, including differentiated grade, TNM stage, CEA level and ARID3A expression status, were selected, and a multivariate analysis of these individuals was performed using the Cox Proportional Hazards Model. The results revealed that TNM stage, CEA level and ARID3A expression status were the three most important prognostic factors, and ARID3A strong expression was proven to be an independent predictor for better prognosis in CRC (95% CI: 0.44-0.96; *P*=0.031).

Table 3. Relationship between clinicopathologic factors and overall survival by univariate Cox proportional hazards regression analysis

Characteristics	no.	HR (95% CI)		P-value
Differentiation grade				
Well and Moderate	617	1.00 (-)		
Poor and others	73	1.609	(1.054-2.459)	0.028
T stage				
T1 and T2	95	1.00 (-)		
T3 and T4	595	2.004	(1.160-3.460)	0.013
N stage				
N0	362	1.00 (-)		
N1/N2	328	3.34	(2.400-4.648)	< 0.001
M stage				
M0	595	1.00 (-)		
M1	95	12.288	(9.016-16.747)	< 0.001
TNM stage				
I and II	340	1.00 (-)		
III and IV	350	5.342	(3.639-7.843)	< 0.001
CEA level				
≤ 5 ng/ml	461	1.00 (-)		
> 5 ng/ml	229	2.715	(2.015-3.659)	< 0.001
ARID3A expression				
Negative/Weak	495	1.00 (-)		
Strong	195	0.545	(0.372-0.800)	0.002

Table 4. Multivariate Cox regression analysis of overall survival

Characteristics	no.	HR (95% CI)		P-value
ARID3A expression				
Negative / Weak	495	1.00 (-)		
Strong	195	0.653	(0.444-0.961)	0.031
CEA level				
≤ 5 ng / ml	461	1.00 (-)		
> 5 ng /ml	229	2.044	(1.509-2.767)	< 0.001
TNM stage				
I and II	340	1.00 (-)		
III and IV	350	4.627	(3.138-6.821)	< 0.001

6. Relationship between ARID3A expression and proliferation of colorectal cancer cell lines

To evaluate the possible role of ARID3A in the proliferation of human CRC cells, upregulation of ARID3A with expression vector and knockdown of endogenous ARID3A with siRNA silencing experiments were performed respectively in RKO cell lines. The effect of ARID3A gene regulation was confirmed by western blot (Figure 6A, C). Proliferation assay showed that cell growth rates did not have significant differences between the ARID3A-overexpressing and the control cells (Figure 6B). Similar results were observed between the ARID3A-inhibiting cells and the mock transfected cells (Figure 6D). In another CRC cell line, LoVo, proliferation assay indicated that when ARID3A was overexpressed, there was a trend that the cell growth of LoVo was reduced as compared to the control cells (Figure 7A). To analyze whether ARID3A expression affects the cell cycle or apoptosis in CRC cells, cell cycle assay and apoptosis assay were also performed in LoVo cells. As show in figure 6, neither cell cycle nor apoptosis assay showed any significant differences between the ARID3A-overexpressing cells and the control cells (Figure 7B-C).
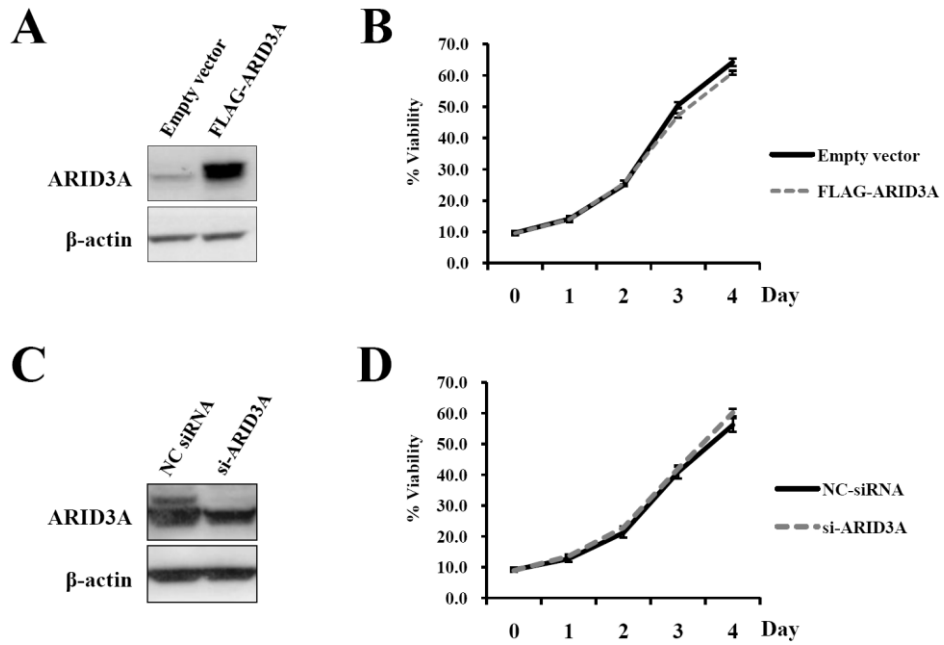
Figure 6. Proliferation assay in RKO colon cancer cell line. Upregulation of ARID3A was confirmed by western blot (A), and cell proliferation was quantitated by fluorescence cell viability assay (B). Endogenous ARID3A was inhibited with siRNA (C), and cell proliferation was quantified using viability assay (D). Error bars represent mean ± SD from three independent experiments.
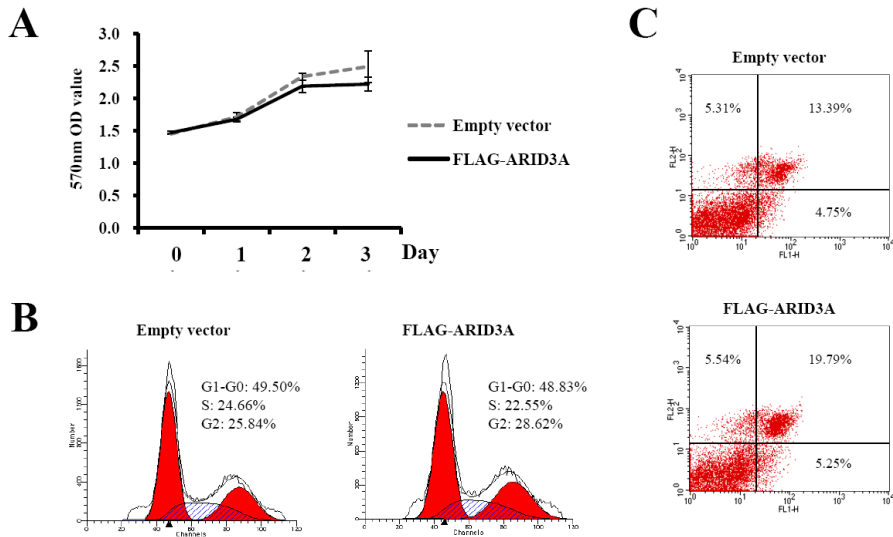
Figure 7. Upregulation of ARID3A expression in LoVo colon cancer cell lines. (A) MTT assay showed that there was a trend that the cell growth was reduced when ARID3A was overexpressed after transient transfection. However, there was no significant difference between the ARID3A-overexpressing cells and the control cells when performing cell cycle (B) and apoptosis (C) assays. Error bars represent mean ± SD from three independent experiments.

7. Relationship between ARID3A and CD133 expression

To investigate whether ARID3A expression is associated with pluripotency-associated markers in human CRC cells, I selected three pluripotency markers (*OCT4*, *SOX2* and *KLF4*) to test their mRNA expression in LoVo cells and SW620 cells which were transfected with pFLAG and pFLAG-ARID3A vector, respectively. I also determined whether ARID3A correlates with colorectal cancer stem cell markers in CRC cells. *CD133*, *CD44*, *CD166*, *CD24* and *ALDH1*, which are regarded as important cancer stem cell markers in colon cancer, were selected to compare their mRNA expression levels between ARID3A-overexpressing and control cells as described above. As shown in Figure 8, ARID3A overexpression was confirmed by real-time PCR. The results showed that mRNA levels of these markers were reduced in cells with ARID3A overexpression compared with the control cells. Among these stemness-related markers, *OCT4* and *CD133* transcription were significantly suppressed in LoVo and SW620 cell lines after upreglating ARID3A. These novel findings indicated that ARID3A might regulate the transcription of cancer stem cell markers in CRC.
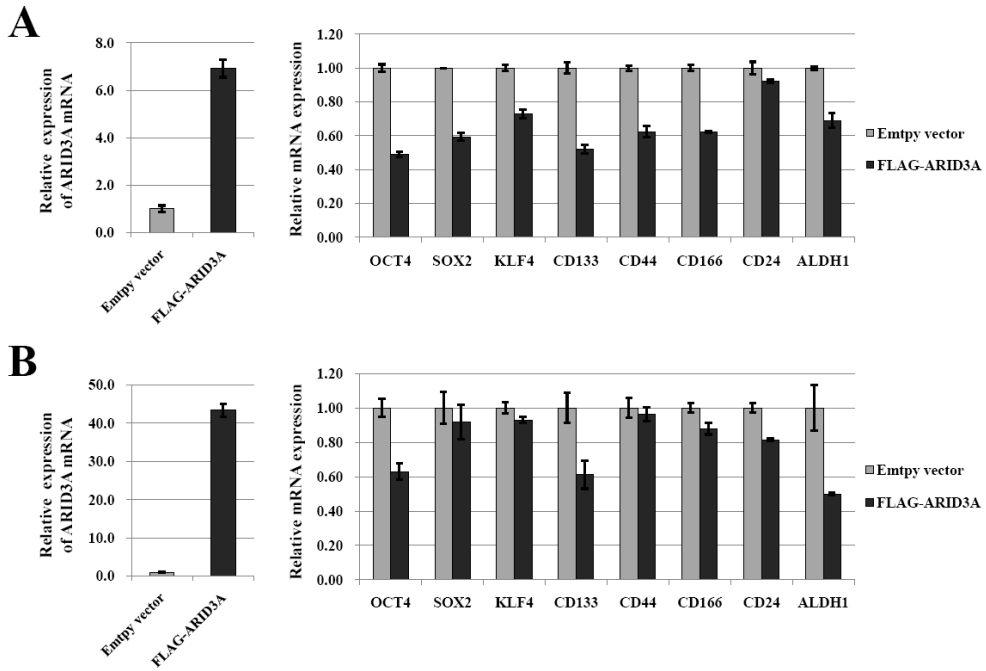
Figure 8. Overexpression of ARID3A reduces the transcription levels of pluripotency-associated markers and cancer stem cell markers. Upregulation of ARID3A in LoVo (A) and SW620 (B) cells were analyzed by real-time PCR (left). Relative mRNA levels of pluripotency-associated markers (*OCT4*, *SOX2* and *KLF4*) and cancer stem cell markers (*CD133*, *CD44*, *CD166*, *CD24*, and *ALDH1*) were determined by real-time PCR in vector-infected and ARID3A-infected LoVo (A) and SW620 (B) cell lines (right). Error bars represent mean ± SD from three independent experiments. Expression levels were normalized for GAPDH.

Furthermore, in order to analyze the relationship between ARID3A and cancer stem cell markers in protein level, I focused on CD133, one of the most important cancer stem cell markers in CRC. I investigated the expression of CD133 protein in CRC cell lines by western blot (Figure 9A). CD133 was highly expressed in HT29 and WiDr cell lines, and moderately expressed in LoVo and SW620 cells. In keeping with qPCR results, expression levels of CD133 protein were reduced after overexpression of ARID3A in LoVo and SW620 cells (Figure 9B-C). Reversely, when ARID3A was repressed by siRNA inhibition in SW620 cells, the levels of CD133 protein were significantly upregulated as compared to the control cells (Figure 9D). The results strongly suggested that ARID3A was associated with CD133 expression in CRC cells.
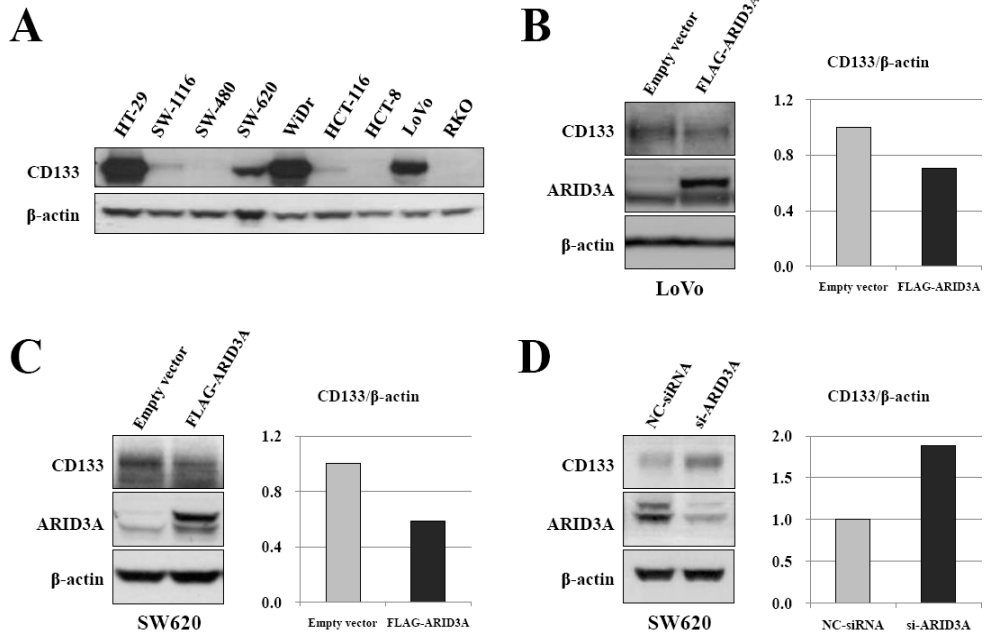
Figure 9. ARID3A regulates the CD133 expression in colorectal cancer cells. (A) The CD133 expression was analyzed in nine CRC cell lines by western blot. Upregulation of ARID3A in LoVo (B) and SW620 (C) cells was associated with reduction of CD133 expression, whereas ARID3A inhibition in SW620 cells (D) was associated with upregulation of CD133 protein levels. (B-D) Western blot analysis of CD133 expression was quantified densitometrically using the ImageJ 1.46 program. β-actin was used as an internal control.

IV. DISCUSSION

In this study, I tried to explore the role of ARID3A in CRC, and demonstrated that strong expression of ARID3A was related to a better prognosis in colorectal cancer. I investigated the expression of ARID3A by immunohistochemistry in tissue microarray specimens of 690 patients with CRC. ARID3A was completely negative in normal mucosae, but was strongly expressed in a subset of CRC tissues. On Kaplan-Meier survival analysis, patients with ARID3A-strong expression have better survival rate than those with ARID3A-negative or weak expression. Interestingly, the survival curves of latter two groups almost overlap (data not shown), suggesting that ARID3A may affect the CRC cell functions when it is highly expressed in the nucleus, whereas weak expression of ARID3A does not exert significant biological functions. On the basis of the result, I classify the CRC patients into two groups - ARID3A-negative or weakly positive and ARID3A-strongly positive group. The strong ARID3A expression was significantly correlated with well-to-moderate differentiation, and MSS or MSI-low type CRC. The strong expression of ARID3A also had an inverse correlation with tumor invasion, lymph node metastasis, distant metastasis, TNM stage, status of microsatellite instability and CEA levels. The present results reveal that strong ARID3A expression is associated with longer survival in CRC patients, and the multivariate Cox regression analysis demonstrate that strong expression of ARID3A is an independent indicator of good prognosis for patients with CRC. These observations suggest that ARID3A may involve in tumorigenesis and progression of CRC.

The regulaltory mechanisms controlling ARID3A expression are not fully understood. Previous studies found that *ARID3A* is a p53 response gene and a functional p53-binding site is located in the second intron of the *ARID3A* gene [20, 27].

p53 can induce and upregulate expression of ARID3A following DNA damage. The exogenously expression of ARID3A can suppress cell growth in U2OS cells expressing wild-type p53 but not in Saos-2 cells lacking p53, suggesting that ARID3A may participate in tumor growth suppression mediated by p53 [20]. Recently, a report showed that ARID3A cooperates with p53 to transcriptionally activate $p21^{WAF1}$ in response to DNA damage [21]. Thus, it is of great interest to investigate whether the prognostic impact of ARID3A expression is dependent on p53 status in CRC. p53 expression is classified as the p53-negative group and p53-positive group, which represent the wild-type and the mutated p53, respectively. In this study, I cannot find a correlation between p53 expression and prognosis (data not shown), which is consistent with some previous studies [28-31]. In p53-negative group, overall survival is significantly better in patients with ARID3-strong positive CRC than in those with ARID3A-negative or weak carcinomas. However, in p53-positive group, ARID3A expression does not statistically correlated with prognosis. These findings were clinical evidence, supporting the previous reports that ARID3A requires the cooperative activity of p53 to exert its function. Additionally, a combination of ARID3A and p53 might be useful as prognostic markers in CRC.

Recent studies imply a functional link between ARID3A and p53. ARID3A is identified to interact with two factors that compose promyelocytic leukemia bodies (PML-NBs): speckled protein 100 (Sp100) and a conjugation enzyme for small ubiquitin-like modifier (SUMO)-1, UbcH9 [32]. Furthermore, p53 also co-localizes and physically interacts with PML-NBs, which is important for p53-dependent apoptosis [33-36]. PML-NBs play crucial roles in the maintenance of cellular integrity [37], whereas increases in the size and number of PML-NBs strongly suppress cell cycle progression and subsequently induce premature senescence [34]. Interestingly, in human fetal lung primary fibroblast cells (TIG-3), silencing of ARID3A expression leads to PML-NB accumulation, resulting in growth suppression and

PML mediated premature senescence, and ARID3A is required for maintenance of survival potential of the cells [32]. Moreover, ARID3A has been reported to associate with E2F-1, one of E2F transcription factor critical for G1-to-S transition [17], which suggests that ARID3A participates in Rb/E2F pathway. Peeper *et al.* revealed that ARID3A can bypass activated RAS-induced senescence in primary murine fibroblasts by deregulating the Rb/E2F1 pathway [19], but authors also mentioned that ARID3A induces premature senescence in primary human cells, probably resulting from the deregulation of Rb/E2F1 pathway, which apparently has different effects in murine and human cells. These reports and the present results support the notion that ARID3A has both oncogenic and tumor suppressing capabilities according to the different intrinsic properties of cells and its cooperating factors.

To analyze the effect of ARID3A on the cell growth in CRC cells, two CRC cell lines (LoVo and RKO) expressiong wild-type p53 were transfected with ARID3A expression vector or with ARID3A-specific siRNAs. The present results indicate that the expression status of ARID3A does not closely correlate with proliferation of CRC cells. A similar result was reported that down-regulation of ARID3A in p53-wildtype and p53-null HCT116 colon carcinoma cells shows no significant effect on colony formation [38]. Together with these results, it seems that ARID3A probably participates in other biological processes than cell growth regulation. Further studies, such as establishment of stable ARID3A expression cell lines instead of transient transfection, are still required to validate the present results and hypothesis.

ARID3A has additional functions in regulation of developmental plasticity. Recent studies support a model in which ARID3A acts as a suppressor of lineage plasticity [22, 39]. Bright/ARID3a-deficient cells from two mouse models express a number of pluripotency-associated gene products, expand indefinitely, and spontaneously differentiat into cells of multiple lineages. Direct knockdown of

ARID3A in the human embryonic kidney epithelial cell line 293T is sufficient to induce expression of pluripotency-associated markers - *OCT4*, *SOX2*, *C-MYC*, and *KLF4* [22]. It has been suggested that a normal role for ARID3A may be to suppress pluripotency early in development. More recently, another study revealed that ARID3A is a direct target of the oncomir *microRNA-125b* in progenitor B-cells, and a significant correlation between the increased expression of ARID3A and the reduction of *OCT4*, *SOX2*, *KLF4* and *NANOG* transcripts is observed [23]. Furthermore, ARID3A is critically required for embryonic transforming growth factor-beta (TGFβ)-signaling in *Xenopus* [40], and it is upregulated by TGFβ in human lung fibroblasts [41]. Recently, inhibitors of TGFβ signaling were shown to enhance somatic cell reprogramming by eliminating the need for exogenous Sox-2 [42], and ARID3A inhibition may be a consequence of TGFβ inhibition in that system.

Currently, the pluripotency genes, such as *OCT4*, *SOX2*, *NANOG*, *LIN28*, *KLF4*, and *C-MYC*, are regarded as promising surrogate markers of cancer stem cell [43]. These genes seem to facilitate a shift toward an undifferentiated state. Their expression has been associated with poor prognosis, relapse, and distant recurrence, as well as with resistance to conventional chemotherapy/radiotherapy [43]. Among these pluripotency-associated markers, I first selected three markers - *OCT2, SOX2*, and *KLF4*, and investigated whether ARID3A expression is associated with the expression of these markers. In two CRC cell lines, LoVo and SW620, ARID3A was overexpressed. At four days after transfection, mRNA levels of pluripotency markers (*OCT2, SOX2*, and *KLF4*) were examined by real-time PCR. In keeping with previous reports, mRNA levels of three markers in ARID3A-upregulated cells were reduced compare with those in control cells. *OCT4* mRNA expression was significantly inhibited in both two cell lines, whereas *SOX2* and *KLF4* mRNA levels was slightly downregulated in SW620 cell line. This result suggests that ARID3A could regulate the transcription of pluripotency-associated markers in colorectal

cancer cells.

Stem cells of the gastrointestinal tract represent the natural target of tumourigenic mutations. The hypothesis of stem cell-driven tumourigenesis in colorectal cancer has received substantial support from the recent identification and phenotypic characterisation of a subpopulation of colon cancer cells able to initiate tumor growth and to reproduce human colon carcinomas faithfully in mice [44]. Based on the previous findings, I questioned whether ARID3A also has relation with cancer stem cell markers in CRC. Several molecules have been proposed as CSC markers including *CD133*, *CD44*, *CD24*, *CD166* and *ALDH1* [43], and I performed transfection study to test mRNA expression of these markers in SW620 and LoVo cells. The results showed that upregulation of ARID3A resulted in the reduction of transcriptional levels of CSC markers, which suggests a role for ARID3A in CRC cells may be to suppress the cancer stem cell development. Among the five markers, *CD133* mRNA expression was markedly downregulated in two cell lines. I further confirmed this result by western blot analysis, and found that ARID3A had an inverse correlation with *CD133* protein in consistent with qRCR data. *CD133* has correlation with poor prognosis, recurrence, metastasis, and resistant to therapy in colorectal cancer [45]. Therefore, the finding that ARID3A can repress CD133 might contribute to the favorable outcomes of patients with colorectal cancer. Additional studies are needed to determine if ARID3A directly represses the stem cell markers or acts indirectly via other pathways.

V. CONCLUSION

In order to elucidate the role of ARID3A in the tumorigenesis of colorectal cancer, I investigated the expression characteristics of ARID3A and its biological significance. ARID3A is highly expressed in a subgroup of colorectal cancer tissues. The overall survival of CRC patients with ARID3A-strong expression was significantly longer than that of patients with ARID3A-negative or weak expression. On multivariate analysis, the strong expression of ARID3A was proven to be an independent predictor for better prognosis in CRC. The ARID3A expression is inversely correlated with several puleripotency-associated markes and cancer stem cell markers in colorectal cancer. Among them, upregulation of ARID3A in CRC cells reduces expression levels of CD133, one of most important cancer stem cell markers in CRC. Further studies are necessary to delineate the mechanistic basis of these features.

REFERENCES

1. Cunningham D, Atkin W, Lenz HJ, et al. Colorectal cancer. Lancet 2010; 375: 1030-47.

2. Vogelstein B, Fearon ER, Hamilton SR, et al. Genetic alterations during colorectal-tumor development. N Engl J Med 1988; 319: 525-32.

3. Pino MS, Chung DC. The chromosomal instability pathway in colon cancer. Gastroenterology 2010; 138: 2059-72.

4. Lothe RA, Peltomaki P, Meling GI, et al. Genomic instability in colorectal cancer: relationship to clinicopathological variables and family history. Cancer Res 1993; 53: 5849-52.

5. Vogelstein B, Fearon ER, Kern SE, et al. Allelotype of colorectal carcinomas. Science 1989; 244: 207-11.

6. Ionov Y, Peinado MA, Malkhosyan S, Shibata D, Perucho M. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. Nature 1993; 363: 558-61.

7. Kane MF, Loda M, Gaida GM, et al. Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. Cancer Res 1997; 57: 808-11.

8. Cunningham JM, Christensen ER, Tester DJ, et al. Hypermethylation of the hMLH1 promoter in colon cancer with microsatellite instability. Cancer Res 1998; 58: 3455-60.

9. Kang UB, Yeom J, Kim HJ, Kim H, Lee C. Expression profiling of more than 3500 proteins of MSS-type colorectal cancer by stable isotope labeling and mass spectrometry. J Proteomics 2012; 75: 3050-62.

10. Wilsker D, Probst L, Wain HM, Maltais L, Tucker PW, Moran E. Nomenclature of the ARID family of DNA-binding proteins. Genomics 2005; 86: 242-51.

11. Wilsker D, Patsialou A, Dallas PB, Moran E. ARID proteins: a diverse family of DNA binding proteins implicated in the control of cell growth, differentiation, and development. Cell Growth Differ 2002; 13: 95-106.

12. Patsialou A, Wilsker D, Moran E. DNA-binding properties of ARID family proteins. Nucleic Acids Res 2005; 33: 66-80.

13. Herrscher RF, Kaplan MH, Lelsz DL, Das C, Scheuermann R, Tucker PW. The immunoglobulin heavy-chain matrix-associating regions are bound by Bright: a B cell-specific trans-activator that describes a new DNA-binding protein family. Genes Dev 1995; 9: 3067-82.

14. Gregory SL, Kortschak RD, Kalionis B, Saint R. Characterization of the dead ringer gene identifies a novel, highly conserved family of sequence-specific DNA-binding proteins. Mol Cell Biol 1996; 16: 792-9.

15. Rajaiya J, Nixon JC, Ayers N, Desgranges ZP, Roy AL, Webb CF. Induction of immunoglobulin heavy-chain transcription through the transcription factor Bright requires TFII-I. Mol Cell Biol 2006; 26: 4758-68.

16. Kortschak RD, Reimann H, Zimmer M, Eyre HJ, Saint R, Jenne DE. The human dead ringer/bright homolog, DRIL1: cDNA cloning, gene structure, and mapping to D19S886, a marker on 19p13.3 that is strictly linked to the Peutz-Jeghers syndrome. Genomics 1998; 51: 288-92.

17. Suzuki M, Okuyama S, Okamoto S, et al. A novel E2F binding protein with Myc-type HLH motif stimulates E2F-dependent transcription by forming a heterodimer. Oncogene 1998; 17: 853-65.

18. Bell LA, Ryan KM. Life and death decisions by E2F-1. Cell Death Differ 2004; 11: 137-42.

19. Peeper DS, Shvarts A, Brummelkamp T, et al. A functional screen identifies hDRIL1 as an oncogene that rescues RAS-induced senescence. Nat Cell Biol 2002; 4: 148-53.

20. Ma K, Araki K, Ichwan SJ, Suganuma T, Tamamori-Adachi M, Ikeda MA. E2FBP1/DRIL1, an AT-rich interaction domain-family transcription factor, is regulated by p53. Mol Cancer Res 2003; 1: 438-44.

21. Lestari W, Ichwan SJ, Otsu M, et al. Cooperation between ARID3A and p53 in the transcriptional activation of p21WAF1 in response to DNA damage. Biochem Biophys Res Commun 2012; 417: 710-6.

22. An G, Miner CA, Nixon JC, et al. Loss of Bright/ARID3a function promotes developmental plasticity. Stem Cells 2010; 28: 1560-7.

23. Puissegur MP, Eichner R, Quelen C, et al. B-cell regulator of immunoglobulin heavy-chain transcription (Bright)/ARID3a is a direct target of the oncomir microRNA-125b in progenitor B-cells. Leukemia 2012; 26: 2224-32.

24. Kobayashi K, Era T, Takebe A, Jakt LM, Nishikawa S. ARID3B induces malignant transformation of mouse embryonic fibroblasts and is strongly associated with malignant neuroblastoma. Cancer Res 2006; 66: 8331-6.

25. Akhavantabasi S, Sapmaz A, Tuna S, Erson-Bensan AE. miR-125b targets ARID3B in breast cancer cells. Cell Struct Funct 2012; 37: 27-38.

26. Joseph S, Deneke VE, Cowden Dahl KD. ARID3B induces tumor necrosis factor alpha mediated apoptosis while a novel ARID3B splice form does not induce cell death. PLoS One 2012; 7: e42159.

27. Wang B, Xiao Z, Ren EC. Redefining the p53 response element. Proc Natl Acad Sci U S A 2009; 106: 14373-8.

28. Hilska M, Collan YU, VJ OL, et al. The significance of tumor markers for proliferation and apoptosis in predicting survival in colorectal cancer. Dis Colon Rectum 2005; 48: 2197-208.

29. Watanabe T, Wu TT, Catalano PJ, et al. Molecular predictors of survival after adjuvant chemotherapy for colon cancer. N Engl J Med 2001; 344: 1196-206.

30. Elsaleh H, Powell B, Soontrapornchai P, et al. p53 gene mutation, microsatellite

instability and adjuvant chemotherapy: impact on survival of 388 patients with Dukes' C colon carcinoma. Oncology 2000; 58: 52-9.

31. Chen MF, Lee KD, Yeh CH, et al. Role of peroxiredoxin I in rectal cancer and related to p53 status. Int J Radiat Oncol Biol Phys 2010; 78: 868-78.

32. Fukuyo Y, Mogi K, Tsunematsu Y, Nakajima T. E2FBP1/hDril1 modulates cell growth through downregulation of promyelocytic leukemia bodies. Cell Death Differ 2004; 11: 747-59.

33. Fogal V, Gostissa M, Sandy P, et al. Regulation of p53 activity in nuclear bodies by a specific PML isoform. EMBO J 2000; 19: 6185-95.

34. Pearson M, Carbone R, Sebastiani C, et al. PML regulates p53 acetylation and premature senescence induced by oncogenic Ras. Nature 2000; 406: 207-10.

35. Zong RT, Das C, Tucker PW. Regulation of matrix attachment region-dependent, lymphocyte-restricted transcription through differential localization within promyelocytic leukemia nuclear bodies. EMBO J 2000; 19: 4123-33.

36. Guo A, Salomoni P, Luo J, et al. The function of PML in p53-dependent apoptosis. Nat Cell Biol 2000; 2: 730-6.

37. Lallemand-Breitenbach V, de The H. PML nuclear bodies. Cold Spring Harb Perspect Biol 2010; 2: a000661.

38. Fukuyo Y, Takahashi A, Hara E, Horikoshi N, Pandita TK, Nakajima T. E2FBP1 antagonizes the p16(INK4A)-Rb tumor suppressor machinery for growth suppression and cellular senescence by regulating promyelocytic leukemia protein stability. Int J Oral Sci 2011; 3: 200-8.

39. Webb CF, Bryant J, Popowski M, et al. The ARID family transcription factor bright is required for both hematopoietic stem cell and B lineage development. Mol Cell Biol 2011; 31: 1041-53.

40. Callery EM, Smith JC, Thomsen GH. The ARID domain protein dril1 is

necessary for TGF(beta) signaling in Xenopus embryos. Dev Biol 2005; 278: 542-59.

41. Lin L, Zhou Z, Zheng L, et al. Cross talk between Id1 and its interactive protein Dril1 mediate fibroblast responses to transforming growth factor-beta in pulmonary fibrosis. Am J Pathol 2008; 173: 337-46.

42. Woltjen K, Stanford WL. Preview. Inhibition of Tgf-beta signaling improves mouse fibroblast reprogramming. Cell Stem Cell 2009; 5: 457-8.

43. Vaiopoulos AG, Kostakis ID, Koutsilieris M, Papavassiliou AG. Colorectal cancer stem cells. Stem Cells 2012; 30: 363-71.

44. Ricci-Vitiani L, Pagliuca A, Palio E, Zeuner A, De Maria R. Colon cancer stem cells. Gut 2008; 57: 538-48.

45. Catalano V, Di Franco S, Iovino F, Dieli F, Stassi G, Todaro M. CD133 as a target for colon cancer. Expert Opin Ther Targets 2012; 16: 259-67.

ABSTRACT (IN KOREAN)

<div align="center">

AT rich interactive domain 3A 유전자가

대장암 발생과정에서의 역할규명

<지도교수 김 호 근>

연세대학교 대학원 의과학과

송 미 영

</div>

AT rich interactive domain 3A (ARID3A)는 DNA 에 결합하는 ARID 패밀리 중 한 단백질로서 예전 연구에 따르면 ARID3A 는 p53-dependent 하게 세포성장을 조절하는 작용을 가지고 있다. 예전 연구에서 ARID3A 는 대장 암조직에서의 발현이 주변 정상조직보다 14 배 이상 증가된 것으로 발견 하였다. 현재까지 인체질병, 특히 대장암에서의 ARID3A 의 생물학적 기 능에 대해 알려지지 않았다. 따라서 본 연구의 목적은 ARID3A 가 대장암 발생에서의 역할을 규명하는데 두고 있다. 면역조직화학염색 방법으로 ARID3A 의 발현을 검사해본 결과, 이 단백질은 일부 대장선종과 대장암 조직에서 발현하고, 주로 세포핵에서 발현함을 발견하였다.   ARID3A 의 발현과 대장암의 예후와 연관이 있는지를 조사하기 위하여 690 명의 대장 암 환자 조직에서의 ARID3A 의 발현특징을 확인하였다. 690 명 환자 중,

195 예에서 ARID3A 는 높게 발현하고 있었고, 187 예에서는 약하게 발현하고, 나머지 308 예에서는 발현되지 않았다. 통계분석결과, ARID3A 의 발현은 환자나이, 종양의 분화도, 침윤도, 림프전이, 원격전이, TNM 병기, 현미부수체 불안정성과 그리고 암측정인자 CEA 와 통계적으로 연관이 되어 있었다. ARID3A 가 높게 발현하는 대장암환자 군은 ARID3A 가 적게 발현하거나 발현을 하지 않는 환자 군에 비해서 생존기간이 더 길었다. 다변량분석결과, ARID3A 의 높은 발현은 독립적인 좋은 예후인자로 가능하였다.

최근에 ARID3A 가 배아줄기세포의 분화유연성을 억제한다는 보고가 나왔다. 이번 연구에서는 ARID3A 가 과발현 되었을 경우, 일부 줄기세포 관련 마커 (*OCT4, SOX2,* and *KLF4*)와 암줄기세포 마커들 (*CD133, CD44, CD166, CD24,* and *ALDH1*)의 전사 정도가 낮아지는 것을 새로 발견하였다. 또한, trasnfection 및 발현억제 (inhibition) 연구를 통하여 ARID3A 는 CD133 단백질의 발현을 억제하는 것을 발견하였다.

결론적으로, 이번 연구에서는 ARID3A 가 대장암 발생과정에서 중요한 역할을 할 가능성이 있음을 발견하였고, 또한 좋은 예후를 예측하는 인자로 가능함을 검증하였다. 또한 대장암 세포에서 ARID3A 의 과발현은 중요한 대장암 관련 줄기세포 마커인 CD133 의 발현을 억제하는 것을 발견하였다. 이러한 현상들의 기전에 대해서 보다 더 심층 깊은 연구가 필요하다.

핵심 되는 말: 대장암, ARID3A, 예후, 암줄기세포

PUBLICATION LIST

1. **Song M**, Cheong JH, Kim H, Noh SH. Nuclear expression of Yes-associated protein 1 correlates with poor prognosis in intestinal type gastric cancer. Anticancer Res 2012; 32: 3827-34.

2. Lee H, **Song M**, Shin N, et al. Diagnostic significance of serum HMGB1 in colorectal carcinomas. PLoS One 2012; 7: e34318.

3. Shin N, You KT, Lee H, Kim WK, **Song M**, et al. Identification of frequently mutated genes with relevance to nonsense mediated mRNA decay in the high microsatellite instability cancers. Int J Cancer 2011; 128: 2872-80.

4. Choi HK, Choi KC, Yoo JY, **Song M**, et al. Reversible SUMOylation of TBL1-TBLR1 regulates beta-catenin-mediated Wnt signaling. Mol Cell 2011; 43: 203-16.

5. Lee H, Shin N, **Song M**, et al. Analysis of nuclear high mobility group box 1 (HMGB1)-binding proteins in colon cancer cells: clustering with proteins involved in secretion and extranuclear function. J Proteome Res 2010; 9: 4661-70.