PINCH2, a subset of IPP complex, a cancer metastasis-related copy number variant with paracrine activity in colon cancer

Chan Hee Park

Department of Medical Science

The Graduate School, Yonsei University

PINCH2, a subset of IPP complex, a cancer metastasis-related copy number variant with paracrine activity in colon cancer

Directed by Professor Hyun Cheol Chung

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

Chan Hee Park

December 2010

This certifies that the Doctoral Dissertation of Chan Hee Park is approved.

Thesis Supervisor : Hyun Cheol Chung

Thesis Committee Member #1 : Won Ho Kim

Thesis Committee Member#2 : Nam Kyu Kim

Thesis Committee Member#3 : Young Nyun Park

Thesis Committee Member#4 : Sun Young Rha

The Graduate School
Yonsei University

December 2010

ACKNOWLEDGEMENTS

<TABLE OF CONTENTS>

LIST OF FIGURES

LIST OF TABLES

<ABSTRACT>

PINCH2, a subset of IPP complex, a cancer metastasis-related copy number variant with paracrine activity in colon cancer

Chan Hee Park

Department of Medical Science
The Graduate School, Yonsei University

(Directed by Professor Hyun Cheol Chung)

Up to now, genomic variability has been reported in human genomes. Copy number variants (CNVs) represent a copy number change involving a DNA fragment that is ~1 kilobases (kb) or larger. Although almost 40% of cancer-related genes are interrupted by CNVs, very few studies have reported on the correlation between CNVs and cancer. Genome-wide arrays provided a more comprehensive analysis of the entire genome. We screened colon cancer metastasis-related CNVs using array-CGH, and investigated the functional mechanism of the selected CNV with *in silico* and *in vitro* assay. We used 41 normal colon tissues from stage III colon cancer patients composed of 22 patients who had been disease-free for over 5 years and 19 relapsed patients within 3 years after surgery. In total, 1,359 CNVs showed different aberration patterns between patients whose cancer recurred and those whose cancer did not recur. From 1,359 CNVs, we selected 89 cancer metastasis-related CNVs consisting of cancer migration/invasion pathways, cellular movement-related pathways, and cell-to-cell signaling and interaction pathways. Noticeably, PINCH2 showed copy number amplification and up-regulation of mRNA

expression in the non-recurred group. PINCH2 is known to regulate cancer cell movement participating in the IPP complex (ILK, PARVA, and PINCH), which is related to cellular movement pathway. In the *in vitro* experiment, five colon cancer cells had a lower protein and mRNA expression of PINCH2 as compared to the three normal epithelial colon cells. PINCH2 had a higher binding affinity with ILK than PINCH1 in both *in silico* and *in vitro* analyses. The PINCH2-ILK complex reduced the expression of MMP-9 metalloprotenase with decreased cell migration and invasion activity. Also, the migration rates of both PINCH2 suppressed HCT-116 cells (autocrine) and cells treated with conditioned media (CM) of PINCH2 suppressed HCT-116 cells (paracrine) increased. In conclusion, PINCH2, a cancer metastasis-related CNV, is associated with cancer cell migration and invasion as an autocrine and paracrine activity in cancer by regulating IPP complex formation.

Key words : Copy number variants, Array-CGH, Metastasis, PINCH2, IPP complex, Colon cancer

PINCH2, a subset of IPP complex, a cancer metastasis-related CNV with paracrine activity in colon cancer

Chan HeePark

Department of Medical Science
The Graduate School, Yonsei University

(Directed by Professor Hyun Cheol Chung)

I. INTRODUCTION

Up to now, tremendous genomic variability has been reported in human genomes. Millions of single nucleotide polymorphisms (SNPs), genomic duplication, and deletion have been revealed as genomic diversities in human populations. SNPs are being discovered with the determination of each personal genome sequence (1, 2). Usually, genomic rearrangement such as genomic duplications and deletions are determined with large-scale copy number variants (LCVs), copy number polymorphisms (CNPs), intermediate-size variants (ISVs), and copy number variants (CNVs) (3-6). Following molecular mechanisms are suggested for these genomic rearrangements : (a) gene dosage, (b) gene interruption, (c) gene fusion, (d) position effects, (e) unmasking of recessive alleles or functional polymorphism, and (f) potential transfection effects (7, 8).

Generally, CNV represents a copy number change involving a DNA fragment that is ∼1 kilobase (kb) or larger. CNVs appear to have a much higher de novo locus-specific mutation rate than SNPs from PCR assays using pooled sperm DNA (9). CNVs are now thought to cover at least 10% of the human

genome. Therefore, a CNV map for the human genome is being continuously regarding about their location, copy number, gene content, frequency, and approximate breakpoints of numerous CNVs. With the advantage of genome-wide analysis tools, several studies have revealed that CNVs are widespread in human genomes and represent a significant source of genetic variation (10, 11). Currently, over 38,000 CNVs and many other structural variations (SVs) have been reported (12, 13) and are now considered to be important factors in diseases pathogenesis.

Genomic rearrangement leads to phenotypic and cellular context changes depending on its gene dosage or cellular function (14). The copy number change can be positively or negatively correlated with gene expression levels, for example, deletion of a transcriptional repressor contributes to elevated gene expression. By gene ontology categories, few genes within CNVs regions were known to be related to environmental sensor functions such as cell adhesion, chemical stimuli, and neurophysiological processes (12, 14). These genomic change give rise to a variety of diseases called 'genomic disorder' (15-18). Several studies have investigated and supported the potential effect of CNVs on human disease e.g. PLP1 (19), MECP2 (20), and LIS1 (21).

Amongst the current CNV-related publications in the Database of Genomic Variants (DGVs), many important tumor suppressor genes and oncogenes with diverse functions that include induction of apoptosis, cell cycle checkpoint control, and DNA repair were found (22, 23). Although nearly 40% of cancer-related genes are interrupted by CNVs through analysis of the DGVs, very few studies reported the relationship between cancer and CNVs (14, 22, 24, 25). Therefore, investigation of this relationship is important for understanding of cancer biology and prognosis.

Up to now, there are no robust cancer prognotic markers. Conventionally, the tumor node metastasis system (TNM stage) and cancer differentiation have

been used as prognostic markers. Previously studied cancer biomarkers mainly targeted oncogene, tumor-suppressor gene, and cellular mechanism-related gene such as cell proliferation (i.e. cyclin A, cyclin D, and cyclin E), cellular adhesion (i.e. E-cadherin, β-catenin, laminin-5, ICAM-1, and CD44), cell invasion (i.e. MMPs family, uPA, and CTSL2) (26, 27), and tumor progression (28, 29). Additionally, Ki-67 (30), STK15 (25), survivin (31), and CCND1 (32) are known to be related with cancer recurrence. TP53, K-RAS, and loss of heterozygosity (LOH) of chromosome 18q were reported as molecular markers in colorectal cancer (CRC) (33-35). However, these biological and clinicopathological parameters are not satisfactory to predict some of aspects (36, 37), and thus, new biomarker identification is necessary to maximize the ability to predict for clinical practice.

Angiogenesis is one of the crucial steps of cancer metastasis that supplies oxygen and food sources to cancer cells. Several studies have reported that tumor angiogenesis was promoted by fibroblasts and stromal cells in the cancer microenvironment (38, 39). Myofibroblasts play an especially crucial role in wound repair, contractile ability, angiogenesis promoting, and stimulating epithelial cell growth through the production of ECM and the secretion of growth factors. It has been determined through extensive study that the survival mechanisms of prostate cancer cells depended on the interactions between tumor cells and the local microenvironment (39-41).

Cancer metastasis is regarded as a result of migration and the invasion process of cancer cells. As various cancer cells survive in a complex microenvironment, cancer cell migration and invasion is affected by surrounding environments such as the extracellular matrix (ECM), diffusible growth factors and cytokines, and a variety of non-epithelial cell types (38). The concept of field effect emphasizes the importance of presence of histologically normal but genetically abnormal tissues surrounding tumor cells.

Based on this effect, the cancer microenvironment is being changed having cancer-like genetic alterations. According to this cancer-like genetic alteration by field effect, normal cells are implicated in the multi-step carcinogenesis through cell proliferation, cell apoptosis, and cell migration (42, 43). For instance, the use of mouse models of tumorigenesis reveals that stromal cells, notable inflammatory cells, vascular cells, and fibroblasts actively support tumor growth (38, 44, 45).

Genomic disorder of microenvironment cells play a central role in local or distance recurrence of primary tumor (42, 43) and can be an important prognostic factor for cancer progression. These results suggest that such genomic disorders in the microenvironment may still remain after resection of the primary tumor and able to leads local recurrence or distance metastasis (46). Therefore, the discovery of a marker of a field effect may be useful for early detection and risk assessment in various cancers (47).

Generally, malignant cells are characterized in histopathology by detection of morphological observation but not, genetic change at the molecular level. Since prognosis predictability of patients based on the morphological changes is limited, detection of molecular change is necessary to increase prediction accuracy. Rapid and cost-effective detection methods that can reliably detect such molecular changes without impacting the patients' quality of life are required. Array-based comparative genomic hybridization (array - CGH) supplies a genome-wide screening for genomic imbalances with an increased resolution by densely spotted DNA clones or oligonucleotides on solid glass supports (10, 48-51). Genome-wide arrays provided a more comprehensive analysis of the entire genome and have been employed to overcome regional genomic biases inherent with targeted array-CGH platforms (52). Array-CGH provides several advantages such as higher resolution, dynamic range, and higher throughput, which has been used to detect genetic change in patients

with chromosomal aberrations (53, 54). The advantages of array are that it is easy to assess and that it shows good correlation of gene expression levels with detected genetic alterations such as amplification and deletion. Analyses of chromosomal imbalances by array-CGH have also contributed to defining the role of genomic alterations in cancer (55, 56).

In this study, we selected colon cancer metastasis-related CNVs using the array-CGH platform from normal colon tissues around the tumor, which had systemic metastasis. We also defined the functional mechanism of the selected CNVs, PINCH2, in metastasis using *in silico* and *in vitro* assay systems as a key molecule in tumor metastasis and progression.

II. MATERIALS AND METHODS

1. Tissue samples

Normal tissues were obtained from colon cancer patients who received curative surgical resection at the Yonsei University College of Medicine, Yonsei University Health System in Seoul, Korea from Febuary 1997 to January 2006. At the time of surgery, tissues were obtained and immediately snap-frozen in liquid nitrogen and stored at -80℃. The patients with colon cancer was done by TNM stage (AJCC 6th Edition). Age, gender, clinical characteristics, and histopathologic findings were summarized from retrospective medical record reviews. For the purpose of cohort comparison, we categorized patients with stage III by recurrence within 3 years, and patients without recurrence at least 5 years after the operation. Among 19 recurred patients, 13 patients died during the follow up period; no patients died in group without cancer recurrence. After selecting 19 patients with systemic relapse, we selected age, gender, and tumor type matched 21 non-relapsed patients from patient database.

2. Cell lines and culture

A total of 6 human cell lines including SNU-638 (gastric cancer as positive cells for PINCH2), CCD-18 Co (normal fibroblast), CCD-841 CoN (normal colon epithelial), CCD-841 CoTr (normal colon epithelial, SV40 transformed CCD-841 CoN), HCT-116 (colon cancer), and DLD-1 (colon cancer) were obtained from the Korean Cell Line Bank (KCLB, Seoul, Korea) and American Type Culture Collection (ATCC, Rockville, Maryland, USA). The cells were cultured in proper media provided by the manufacturers and according to their instructions supplemented with 10% FBS (Omega Scientific, Tarzana, CA, USA), and incubated at 37℃ in a 5% CO2 humidified

atmosphere. The conditioned medium was obtained Opti-MEM I with no growth factor added, and was replaced every 3 days.

3. Extraction of DNA and RNA

All experiments using patient tissues were performed upon approval of the internal review board of the Yonsei University Health System. The tissues used in these experiments were ground to a powder and the genomic DNA (gDNA) was prepared using the phenol/chloroform/isoamylalcohol method (57). Briefly, the tissues were minced in liquid nitrogen and incubated at 42℃ for 12-24 hours with 1 ㎖ lysis buffer [10 mM Tris PH 7.6, 10 mM EDTA, 50 mM NaCl, 0.2% SDS, 200 ㎍/㎖ Portenase K]. We also added 1 ㎖ of phenol/chloroform/isoamylalcohol (Invitrogen, Carlsbad, CA, USA) to the lysis buffer. The gDNA was precipitated with 10 M ammonium acetate, glycogen and 100% EtOH.

Total RNA was extracted from each cell line or tissues using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. Briefly, 1 ㎖ Trizol was added to tissue or cell for lysating and homogenization by inverting. After addition of 200 ㎕ chloroform, homogenate samples were centrifuged 13,000 rpm for 30 minutes at 4℃, and the aqueous phases were collected. Total RNA was precipitated with the same volume of isopropanol for 30 minutes at -30℃ and centrifuged 13,000 rpm for 30 minutes. The pellets were washed once with 1 ㎖ 70% EtOH and suspended in RNase free water. The quantity and quality of gDNA and RNA were checked by ND-100 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) and gel electrophoresis. The extracted gDNA and total RNA were stored in a -70℃ deep freezer.

4. Oligonucleotide microarray CGH (array-CGH) assay

In this study, we used 60 base pairs in situ synthesized oligonucleotide arrays designed and produced by Agilent Technologies Co. (Santa Clara, CA, USA). This was a custom array containing 43,024 probes with coding and non-coding sequences. Each probe represented a specific gene and had 43 Kb of median resolution length based on build NCBI 36 (human genome assembly UCSC hg 18, March 2006). As a reference samples, we used commercial human male and female gDNA (Promega, Madison, WI, USA). All array hybridization was performed according to the Agilent's recommended protocols. Briefly, 5 μg gDNA was digested with restriction enzymes Alu I (Promega, Madison, WI, USA) and Rsa I (Promega, Madison, WI, USA) and labeled fluorescently using the Agilent DNA Labeling kit (Agilent, Santa Clara, CA, USA). Patients' gDNAs were labeled with Cy3 and commercial gDNAs as a reference were labeled with Cy5 dUTP, respectively. The labeled DNA were denatured and pre-annealed with Cot-1 DNA and Agilent blocking reagent prior to hybridization for 40 hours at 20 rpm in an Agilent hybridization oven. Washing procedures were followed by standard procedures. The hybridized arrays were scanned at 5μm resolution using an Agilent scanner (Agilent, Santa Clara, CA, USA) and the image was analyzed with Feature Extraction Software 9.1.1.1 (Agilent, Santa Clara, CA, USA).

5. CNVs selection and ontology analysis

After self-self and dye-swap hybridization experiment using male and female commercial reference genomic DNA, we determined cut-off values of genetic aberration with < -0.58 (deletion) and > 0.58 (amplification) $\log_2$(R/G) ratio from X chromosome genes which represented the 1.5-fold difference in reference samples (51, 58). The total aberration frequency for all probes was analyzed in all tissue samples. In this study, we determined CNVs as an aberrant gene to have more than 20% frequency in all samples (7, 22, 59, 60).

The biological roles of the selected CNVs were investigated using Core Analysis in IPA 5.5 and various biological functions were predicted as major candidate functions. To interpret the biological significance of genes located at altered chromosomal regions of interest, a gene ontology analysis was performed using Ingenuity Pathways Analysis (Ingenuity Systems®, Redwood City, USA), Ensembl Web site database (http://www.ensembl.org/) and the DAVID (http://apps1.niaid.nih.gov/david/) web database. To reduce the false-positive rate of biological validation, probes at the intronic sequence were excluded and multiple exonic probes were averaged. Finally, we selected metastasis-related CNVs for functional study by the following criteria; a) CNV having 30% or more frequency of difference of aberration between recurred and non-recurred groups, 2) CNV having no known function in cell proliferation or apoptosis c) CNV having a functional association with secretary protein, and d) CNV located at a higher level of the cellular movement-related signal pathway cascade.

6. Quantitative real-time PCR, RT-PCR and northern-blotting

To validate the mRNA expression and DNA copy number of the selected genes, we used quantitative real-time PCR and real-time RT-PCR, respectively. cDNA was synthesized as follows; 4 ㎍ of total RNA from each sample was mixed with 2 ㎍ of oligo-dT primer (5'-GGCCAGTAATTGTAATACGACTCACTATAGGGAGGCGG-3', Genotech, Daejeon, Korea) and incubated at 65℃ for 10 minutes. After adding SuperScript II (Invitrogen, Carlsbad, CA, USA), 5X first strand buffer (Invitrogen, Carlsbad, CA, USA), 100 mM DTT (Invitrogen, Carlsbad, CA, USA), and 10 mM dNTP mix (Invitrogen, Carlsbad, CA, USA) to the RNA/oligo-dT mixture, a reverse transcription process was performed at 42℃ for 1.5 hours. The remaining RNA was hydrolyzed by incubation at 65℃ for

15 minutes in 0.1 N NaOH. Quantitative PCR and RT-PCR was performed using QuantiTech SYBR Green PCR (QIAGEN, CA, USA). Each reaction was run in triplicates on Rotor Gene 2072D (Corbett Research, New South Wales, Australia) real-time PCR machine. Expression and DNA copy values for each gene were determined using a standard curve constructed from Human Genomic DNA (Promega, Madison, WI, USA). Non-template-control wells without cDNA or DNA were included as a negative control. The house-keeping gene β-actin was selected for normalization of the standard curve for real-time PCR. The primer set for PCR and RT-PCR amplification are designed 700 to 1000bp product size within the exon sequence as GNLY (forward 5'-CTGAGCCCTCTCACCTTGTC-3', reverse 5'-GCGGCTAAAATCTTGCTTGA-3'), PINCH2 (forward 5'-GGACAAGCATGTGCAGTCAC-3', reverse 5'-TGGTCCGTACAAACAAATACAAA-3'), TGFB3 (forward 5'-GGAAGCAGTAATTGGTGTCCA-3', reverse 5'-AAGGATCACCACAACCCTCA-3'), MMP2 (forward 5'-ACTGCTGGCTGCCTTAGAAC-3', reverse 5'-GTGAACAGGGGAACCATCAC-3'). The house-keeping gene GAPDH (foward 3'-CCATGGAGAAGGCTGGGG-5', reverse 3'-CAAAGTTGTCATGGATGACC-5') was used as a positive control in RT-PCR. Independent real-time PCR experiments were performed in triplicates, and the mean±standard deviation (SD) of the $\log_2$-transformed data were compared with those from the microarray experiment. Correlations among mRNA expression levels, DNA copy number and array-CGH data of each gene were compared using correlation coefficients.

For evaluation of mRNA expression, we used non-radioactive northern-blotting method using DIG Nonradioactive Nucleic Acid Labeling and Detection System (Roche, Mannheim, Germany). Briefly, 2ug of total RNA were

separated with 1.5% MOPS-formaldehyde (33%) agarose gels and separated RNAs were transfered to a positively charged nylon membrane, using sterile 20X SSC as transfer buffer. Fixation of the nucleic acids to the membrane were accomplished using UV-cross linking (at 120 mJ) and briefly wash the membrane in water, and air dry it. For the hybridization, 100 ng/ml of labeled PINCH2 probe was used. Denature the probe for 5 min at 100°C, then immediately chill on ice and hybridized the blot at 68°C overnight in a sealed plastic bag (Roche, Mannheim, Germany). After hybridization, each membrane was immediately washed with low stringency wash buffer (2X SSC, 0.1% SDS) at room temperature and pre-warmed high stringency wash buffer (0.1X SSC, 0.1% SDS) at 68°C. Chemiluminescent detection were used CDP-Star™ (Roche, Mannheim, Germany) substrate and the signal intensity was analyzed using Image-J image analysis software (NIH, MD, USA).

7. Protein expression and western-blotting

Protein was extracted with a NP-40 lysis buffer (20 mM Tris/HCl, pH 8.0, 150 mM NaCl, 1% Non-idet P40, 10% [v/v] glycerol, 2.5 mM EDTA, 100 uM Na3VO4, 1% aprotinin, 1% leupeptinand 1 mM PMSF) and stored in a -70℃ deep freezer. Extracted protein was separated by SDS-PAGE using 9.4% polyacrylamide gels. The proteins were transferred onto PVDF membranes (GE Healthcare, Piscataway, NJ, USA) and analyzed by immunoblotting using antibodies against PINCH2, 1: 500 (Santa Cruz Biotechnology, Santa Cruz, CA, USA), PINCH1, 1:1000 (abcam, Cambridge, MA, USA), MMP-9, 1:750 (abcam, Cambridge, MA, USA), MMP-11, 1:750 (Abnova, Taipei, Taiwan), ILK, 1:1000 (SIGMA, St. Louis, MO, USA), PARVA, 1:1000 (Abnova, Taipei, Taiwan), AKT, 1:1000(Cell Signaling Tec., Danvers, MA, USA), phospho-Akt, 1:750 (Cell Signaling Tec., Danvers, MA, USA) and B-actin, 1:2000 (Delta Biolabs, Gilroy,CA, USA). Secondary

anti-mouse IgG-HRP (GE Healthcare, Piscataway, NJ, USA) and anti-goat IgG-HRP (Santa Cruz Biotechnology, Santa Cruz, CA, USA) were used at 1:2000 and 1:1000 dilutions, respectively. The membrane was developed by ECL Western blotting reagents (Amersham Pharmacia Biotech, Buckinghamshire, UK) and the signal intensity was analyzed using Image-J image analysis software (NIH, MD, USA).

8. Immunoprecipitation (IP) for protein-protein interaction

For IP analysis, a 50 ㎕ protein G Magnetic bead (Millipore, Billerica, MA, USA) and the target antibody (described in methods 7) were incubated at room temperature for 30 minutes with continuous mixing. After washing, the immobilized capture antibody (described in methods 7) was bound with 500 ㎍ of lysate protein at 4℃ overnight. The incubated samples were washed 3 times with 500 ㎕ of PBS containing 0.1% Tween-20 surfactant. After the last washing, the tube was removed from the magnetic rack and the NP-40 lysis buffer for denaturing or native elution was added.

9. Target suppression with small interfering RNA (siRNA) and harvest of conditioned media (CM)

Twenty-four hours prior to transfection, $5x10^4$cells suspended in growth medium without antibiotics and seeded in a plate. The cell density was 30-50% confluent at the time of transfection. Transfection of siRNA was carried out with Lipofectamine RNAiMAX (Invitrogen, Carlsbad, CA, USA). Stealth siRNAs (Invitrogen, Carlsbad, CA, USA) targeting PINCH2 were mixed with Opti-MEM I Reduced serum Medium (Invitrogen, Carlsbad, CA, USA). This mixture was then mixed with equal volume of Lipofectamine RNAiMAX in Opti-MEM I Reduced Serum Medium. After 20 minutes of incubation, the final mixtures were added to each well of the plate for a final

siRNA concentration of 100 to 200 pmol. To verify mRNA and protein repression by siRNA, mRNA was extracted after 24, 48, 72, 96 and 120 hours and protein was extracted after 24 and 48 hours of incubation with the 100 and 200 pmol concentration of siRNA. The sequence of 3 set siRNA is (RNA)-GGUGCUACGAGAAGUUCCCGCUGGA, (RNA)-GCGAUGUGGUGUCGGCCCUCAACAA) and (RNA)-CCACGGGCAAUAUGUCGGACGCCUU. Stealth RNAi negative Control Duplex (Invitrogen, Carlsbad, CA, USA) was used for non-targeting siRNAs (NT siRNA) as a negative control.

For migration and invasion experiment, we collected conditioned media (CM) with each siRNA treated cell culture as follows; 1) HCT-116 cells cultured with confluent conditions in Opti-MEM at 48 hours after siRNA treatment, 2) 2 ㎖ of Opti-MEM media were harvested after 48 hours, 3) harvested Opti-MEM media were centrifuged 1500g for 10 minutes at 4℃, 4) the aqueous phases were collected and stored at -70℃ before usage.

10. Enzyme-linked immunosorbent assay (ELISA)

MMP-9 and MMP-11 levels from the CM of the control and siRNA-treated HCT-116 cells were quantified with ELISA (R&D Systems, Minneapolis, MN, USA) according to the manufacturer's protocol. In brief, CM was concentrated by times using a Microcon-30 (Amicon, Bedford, MA, USA) and 200 ㎕ of concentrated CM was added directly into the wells in duplicates. After antibody conjugation and luminescence reaction, the procedure was completed with stop solution, and we measured the optical densities of each well at 450 ㎚ and 570 ㎚. The final optical density was calculated by subtracting the readings at 570 ㎚ from the readings at 450 ㎚.

11. Prediction of protein structure

The protein structures of ILK 6-171 (PDB 3XIE) and PINCH1 LIM1 6-70 (PDB 3XIE) were obtained from the Protein Data Bank (PDB, http://www.rcsb.org). We compared the PINCH1 and PINCH2 protein sequences in the GenBank database using Blast sequence similarity search to evaluate the PINCH2 protein structure. The protein structures were composed of helical structures, beta-sheet arrows, and coil-and-turn structures. The LIM1 domain structure of PINCH2 was predicted using analysis software based on the LIM1 domain of PINCH1 (PDB 3XIE). The binding affinity (ZRANK score) and strength (R Dock score) of PINCH1-ILK and PINCH2-ILK were calculated with the ZDOCK program (Accelrys, CA, USA) as follows; 1) define binding site (ligand-based or cavity-based), 2) generate ligand conformations, 3) dock each conformation (align shapes of ligand to binding site), 4) save the top docked structures (diverse poses), 5) apply scoring function to each docked structure for the best binding mode (binding affinity prediction). Finally, we selected the highest rank of binding affinity and strength score (Accelrys, San Diego, CA, USA).

12. Invasion assay using matrigel and trans-well filter

We conducted a modified version of the standard trans-filter for invasion assay. The lower sides of the trans-well filters (diameter, 6.5 ㎜; pore size, 8 ㎛; BD Biosciences, San Jose, CA, USA) were coated with 8 ㎍/㎕ Matrigel (BD Biosciences, San Jose, CA, USA) and the trans-well filters were placed on a 24-well plate (Becton Dickinson, Franklin Lakes, NJ, USA) containing medium supplemented with 0.1% bovine serum albumin (Amresco, Solon, OH, USA). Cells ($2X10^4$) were added to the upper compartment of a trans-well chamber and allowed to migrate for 36 hours at 37℃. After 36 hours, non-migrated cells on the upper side of membrane were removed with a cotton swab, and migrated cells on the bottom surface of the membrane

were fixed in 3.7% paraformaldehyde in PBS and stained with crystal violet for 10 minutes at room temperature. Cell migration was quantified by counting the number of cells from three inserts at each point.

13. Wound-healing migration assay

For the measurement of cell migration during wound-healing, cells ($2x10^5$) were seeded in individual wells of a 6-well culture plate. When the cells reached a confluent state, cell layers were wounded with a plastic micropipette tip with a large orifice. The medium and debris were aspirated away and replaced by 0.1mL of fresh, serum-free medium. The confluent monolayer of cells was scratched with a sterile pipette tip to form a wound area and cultured with 5% serum medium. The migrated cells were photographed every 3 hours in assay with autocrine activity or 6 hours in assay with paracrine activity after wounding by phase contrast microscopy in each condition. For evaluation of wound closure, three randomly selected points along each wound were marked and the horizontal distance of the migrating cells from the initial wound was measured and calculated using Image-J image analysis software (NIH, MD, USA).

14. Statistical methods

Quantitative data are represented as the mean ± SD of three independent experiments. Data were analyzed using descriptive statistics. To test for differences in variables, we used a student t-test. Statistical analysis was conducted at a 95% confidence interval. $p \leq .05$ were considered statistically significant. The data were analyzed using SPSS version 13 (SPSS Inc, Chicago, Ill, USA).

III. RESULTS

1. Patients characteristics

The clinical information for colon cancer patients was obtained from Yonsei Cancer Center (YCC), Yonsei University Health System (YUHS), which normal tissues from cancer patients used to select the recurrence classifier. A total of 41 normal colon tissues from stage III colon cancer patients were used (22 disease-free over 5 years and 19 recurred within 3 years after surgery). All cancer patients had complete demographic, gender, stage, survival, and recurrence information. The demographics for the recurred and non-recurred patients have been summarized in Table 1. All of the non-recurred patients survived, but 68.4% of recurred patients died of cancer.

Table 1. Patient characteristics

		Non-recurred patient (N=22)	recurred patient (N=19)	Total
Age (years)				
	median(range)	58.5 (26-77)	63 (39-83)	59.1 (26-83)
Sex				
	male	14	13	27
	female	8	6	14
Location				
	A-colon	8	8	16
	D-colon	2	2	4
	S-colon	9	8	17
	T-colon	3	1	4
Survival				
	Alive	22	6	28
	Death	0	13	13

A-colon, Ascending colon; D-colon, Descending colon; S-colon, Sigmoid colon; T-colon, Transverse colon

2. Screening of CNVs from 41 normal colon tissues

The amplification and deletion of genes in these samples were determined based on a ± 0.58 $\log_2$ ratio, which represents the 1.5 fold difference from reference samples. As a whole, 28.3% (12,204/43,024) of the probes had genetic aberration patterns called “copy number variants (CNVs)” having higher than 20% frequency. These genetic aberration probes were located on 1,881 non-coding and 10,323 coding regions. Among them, 326 probes showed more than 50% of aberration frequencies and eight probes had more than 80% aberration frequencies. In the oligonucleotide microarray, genes were immobilized with multiple probes at different regions of the same gene; generally, these multiple probes showed a same genetic aberration pattern. We selected these genes having multiple aberrated probes as a copy number variation region (CNVR). BOLA2 had amplified patterns in 14 probes located on chromosome 16, 29581455~30104842 with a 523.9KB size that was known as CNVR in the DGV database. DMD, CSMD1, ALCAM and NBPF1 genes showed same genetic aberration patterns in more than seven probes (Table 2).

Table 2. List of CNVR showing same aberration pattern in more than 5 different probes

Symbol	Cytoband	Start	Stop	Num. probes	Size of CNVs	AF(%)*
DAB1	1p32-p31	57211225	58258018	5	1046kb	29.3
NBPF1	1p36.13	142444242	142592976	7	148kb	29.3
ALK	2p23	29357766	30017573	6	659kb	45.1
ALCAM	3q13.1	106568590	106776928	7	208kb	38.3
TLL1	4q32-q33	167152764	167381050	5	228kb	26.8
MSH3	5q11-q12	79986845	80204770	5	217kb	35.1
MAD1L1	7p22	1867470	2041488	5	174kb	34.6
CSMD1	8p23.2	2795154	3876999	7	1081kb	30.7
PARD3	10p11.21	34666203	35101832	6	435kb	37.8
HCCA2	11p15.5	1447673	1733237	5	285kb	30.7
HIF1A	14q21-q24	61240176	61268981	5	28kb	32.2
RORA	15q22.2	58610250	59094212	5	483kb	27.3
XYLT1	16p12.3	17157632	17237747	5	801kb	32.2
BOLA2	16p11.2	29581455	30104842	14	523kb	36.8
TBCD	17q25.3	78348926	78491962	5	143kb	29.3
PEPD	19q12-q13.2	38588072	38702049	5	113kb	33.7
C20orf133	20p12.1	14079677	15113480	5	1033kb	30.7
AK123035	21q22.12	35413601	35984858	5	571kb	32.2
DMD	Xp21.2	30956060	32797968	9	1841kb	30.4
DACH2	Xq21.3	85333848	85877220	5	543kb	26.8
PCDH11X	Xq21.3	90979997	91325767	5	345kb	28.8

*, AF; aberration frequencies from 41 normal colon tissues

CNVs; copy number variants, CNVR; copy number variation region

3. Selection of cancer metastasis-related CNV

We determined the metastasis-related CNVs, which showed higher than 30% difference of aberration frequency of single copy change level (1.5-fold different). The frequency of alteration was compared between the recurred group and non-recurred group within 12,204 screened CNVs. Consequently, a total of 1,352 CNVs (1,395 probes, approximately 3.2% of the probes on the microarray) were recognized as metastasis-related CNVs (Figure 1). The selected 1,352 CNVs showed various genetic aberration patterns such as 205 CNVs (221 probes) of amplification patterns (Figure 1A), 207 CNVs (214 probes) of deletion pattern (Figure 1B) and 945 CNVs (960 probes) of complex patterns (amplification or deletion) (Figure 1C). In these metastasis-related CNVs, 39 CNVs (43 probes) appeared to have more than two probes. 1,352 CNVs were assigned to three main gene ontology (GO) categories; biological process, molecular functions, and cellular components. Many of the probes were assigned to multiple GO categories and pathways. Finally, 89 CNVs having a biological function of cellular movement were filtered out, (Figure 2). These CNVs contained cancer-related genes such as MMP1, ERBB2, MPK8, and known CNVs such as PINCH2, L1CAM and GNLY. Among them, 58% (52/89) of the gene product, protein, were found to be located in the extracellular space and plasma membrane supporting secretary proteins (Table 3).

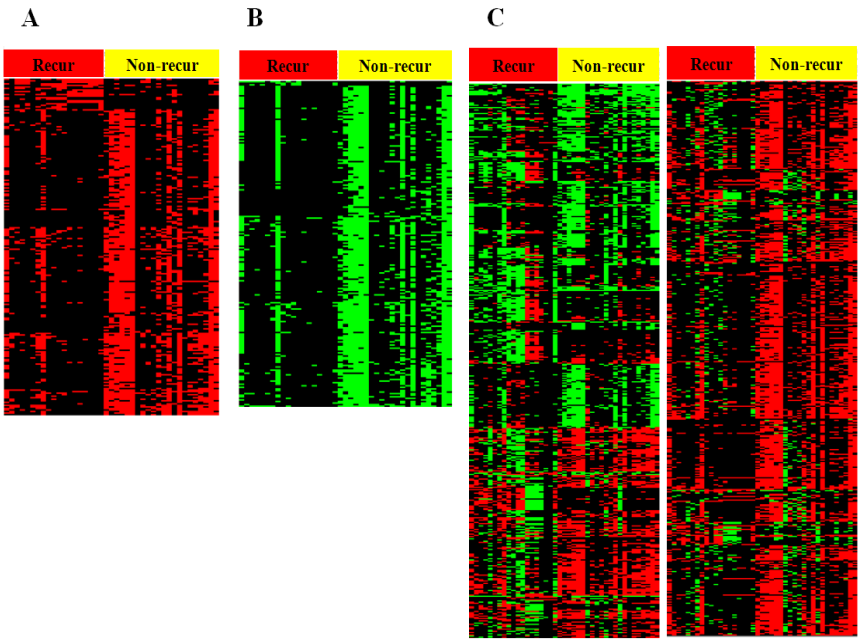


Figure 1. Heat-map with selected 1,359 metastasis-related CNVs. A. Amplification (gain) pattern of 221 probes. B. Deletion (loss) pattern of 214 probes. C. Complex pattern of 960 probes showed various patterns of genetic aberration such as gain, loss, and both. Each red, green and black section denotes amplification, deletion, and no change in copy number.

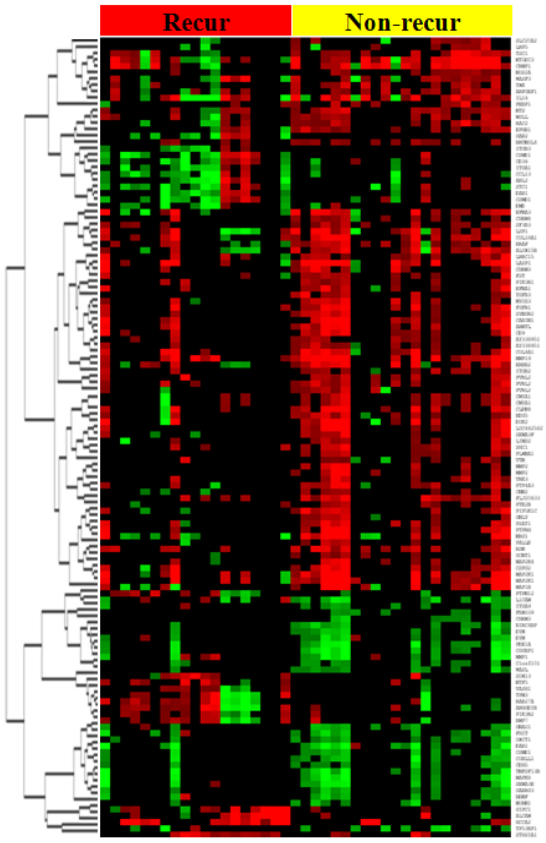


Figure 2. Heat-map with selected 89 cellular movement related CNVs. From web databases and gene ontology analysis software, we selected 89 CNVs known as cell migration and invasion-related genes. Each red, green, and black section denotes amplification, deletion, and no change in copy number.

Table 3. Gene lists and cellular location of the gene products from selected 89 cellular movement-related CNVs

Symbol	Location*	Type**
ABL2	Cytoplasm	kinase
ALCAM	Plasma Membrane	unknown
ALOX15B	Cytoplasm	enzyme
ARHGDIB	Cytoplasm	unknown
ARHGEF1	Cytoplasm	unknown
BDNF	Extracellular Space	growth factor
BMP7	Extracellular Space	growth factor
BRAF	Cytoplasm	enzyme
CCL13	Extracellular Space	cytokine
CCL26	Extracellular Space	chemokine
CD36	Plasma Membrane	transmembrane receptor
CD9	Plasma Membrane	unknown
CDH5	Plasma Membrane	unknown
CHRM3	Plasma Membrane	G-protein coupled receptor
CHRM4	Plasma Membrane	G-protein coupled receptor
CLDN4	Plasma Membrane	transmembrane receptor
CNN2	Cytoplasm	unknown
COL18A1	Extracellular Space	unknown
COL4A1	Extracellular Space	unknown
CRMP1	Cytoplasm	enzyme
CSMD1	Plasma Membrane	unknown
DAB1	Cytoplasm	unknown
DCLK1	Cytoplasm	kinase
DOK2	Plasma Membrane	unknown
EFNA1	Plasma Membrane	unknown
EPHB1	Plasma Membrane	kinase
ERBB2	Plasma Membrane	kinase

FGFR1	Plasma Membrane	kinase
FHIT	Cytoplasm	enzyme
FST	Extracellular Space	unknown
GAB2	Cytoplasm	unknown
GCNT1	Cytoplasm	enzyme
GIPC1	Cytoplasm	unknown
GNAO1	Plasma Membrane	enzyme
GNLY	Cytoplasm	unknown
IL16	Extracellular Space	cytokine
ITGA1	Plasma Membrane	unknown
ITGA9	Plasma Membrane	unknown
ITGB2	Plasma Membrane	unknown
ITGB3	Plasma Membrane	transmembrane receptor
L1CAM	Plasma Membrane	unknown
LASP1	Cytoplasm	transporter
LRP5	Plasma Membrane	unknown
LRRC15	Plasma Membrane	unknown
LSP1	Cytoplasm	unknown
MAP1B	Cytoplasm	unknown
MAP2K1	Cytoplasm	kinase
MAP2K4	Cytoplasm	kinase
MAPK8	Cytoplasm	kinase
MGLL	Plasma Membrane	enzyme
MMP1	Extracellular Space	unknown
MMP19	Extracellular Space	peptidase
MMP2	Extracellular Space	peptidase
NOS2	Cytoplasm	enzyme
NRG1	Extracellular Space	growth factor
NTF4	Extracellular Space	growth factor
NTS	Extracellular Space	unknown

PALLD	Cytoplasm	unknown
PDE4D	Cytoplasm	enzyme
PEBP1	Cytoplasm	unknown
PIK3R1	Cytoplasm	kinase
PIK3R2	Cytoplasm	kinase
PINCH2	Cytoplasm	unknown
PIP5K1C	Plasma Membrane	kinase
PLXNA1	Plasma Membrane	transmembrane receptor
PTK2B	Cytoplasm	kinase
PTP4A3	Plasma Membrane	phosphatase
PTPN12	Cytoplasm	phosphatase
PTPRH	Plasma Membrane	phosphatase
PVRL2	Plasma Membrane	unknown
RAC2	Cytoplasm	enzyme
S1PR2	Plasma Membrane	G-protein coupled receptor
SEMA3E	Extracellular Space	unknown
SEMA3F	Extracellular Space	unknown
SHC1	Cytoplasm	unknown
ST8SIA1	Cytoplasm	enzyme
STC1	Extracellular Space	kinase
TGFB3	Extracellular Space	growth factor
TNFSF13B	Extracellular Space	cytokine
TNR	Plasma Membrane	unknown
TPM3	Cytoplasm	unknown
TSC1	Cytoplasm	unknown
TXK	Cytoplasm	kinase
VASH1	Extracellular Space	unknown
VCAN	Extracellular Space	unknown
VTN	Extracellular Space	unknown
WASF3	Cytoplasm	unknown

WASL	Cytoplasm	unknown
XIRP1	Plasma Membrane	unknown

*, subcellular localization of the active form of the protein;

**, protein family or molecule class

4. Validation of array-CGH using quantitative real-time assay

We performed quantitative real-time PCR (qPCR) and RT-PCR (qRT-PCR) assays to validate the CNVs profiles and expression profiles identified by microarray hybridization analysis, respectively, from 10 normal colon tissues (5 recurred, 5 non-recurred samples) (Figure 3). We analyzed known CNVs (GNLY and PINCH2) and known cancer-related genes (TGFB3 and MMP3). To enhance the specificity of the PCR primer set, primers were designed to anneal near the sequences corresponding to the probes in coding sequence. The β-actin house-keeping gene was adopted as a control gene for the quantification of the total amount of DNA and mRNA. The qPCR data was consistent with that of array-CGH result, and the qRT-PCR experiments showed the same expression patterns to DNA copy change (Figure 3). By array-CGH and qPCR assays, selected CNVs showed genetic aberration more than single copy levels (1.5-folds, 0.58 $\log_2$ ratio) between the recurred and non-recurred groups. In addition, mRNA expression level was up-regulated by 64-fold (GNLY, MMP2) and 11.3-fold (PINCH2, TGFB3) in the non-recurred group as compared to the recurred group ($p<0.01$), suggesting that mRNA expression level is correlated with DNA copy number.

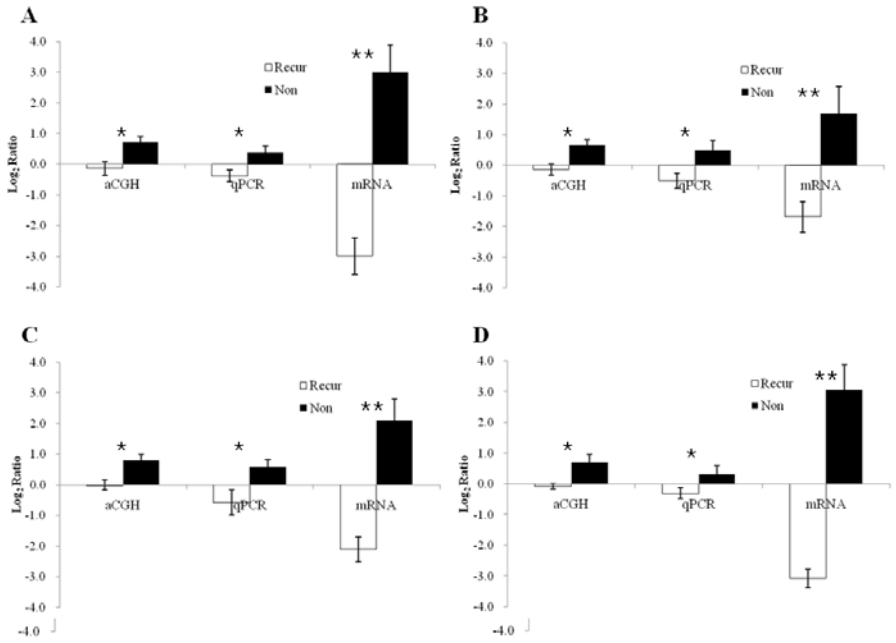


Figure 3. Confirmation of microarray data by quantitative real-time PCR and quantitative real-time RT-PCR. Four genes from two known CNVs and two cancer-related CNVs were quantitatively selected for confirmation. The quantities of each experiment were normalized by those for the β-actin gene. The genes and their identities are as follows: (A) GNLY, granulysin; (B) PINCH2, LIM and senescent cell antigen-like domains 2; (C) TGFB3, transforming growth factor, beta 3; (D) MMP2, matrix-metalloprotenase 2. The white bar shows the log-transformed ratio from the recurred group, and the black bar showed the log-transformed ratio from the non-recurred group in each experiments. The value of each gene was averaged and a student t-test was calculated for comparison between the two groups. *,$p<0.05$; **,$p<0.01$.

5. Target gene selection by ontology and pathway analysis

To characterize the 89 identified gene sets, we investigated the molecular functional pathways of 89 metastasis-related CNVs (Figure 4) and categorized them into three groups based on their major roles in cancer development. Each “ingenuity gene category” included different genes known to be connected to cell migration. The first category included genes that contribute to cell migration and invasion (Figure 4A). In this pathway, PINCH2 (LIM and senescent cell antigen-like domains2, LIMS2), which shows the copy number amplification in the non-recurred group, was found to be a key molecule. The second category (movement-related pathway) includes the mitogen-activated protein kinase (MAPK), matrix metalloproteinase 2 (MMP2), transforming growth factor, beta 3 (TGFB3), matrix metalloproteinase 19 (MMP-19), and zinc fingers and homeoboxes 2 (Raf) (Figure 4B). The third category (cell-to-cell signaling and interaction) includes several genes: activated leukocyte cell adhesion molecule (ALCAM), integrin-beta 2 (ITGB2), integrin, beta 3 (ITGB3), cell adhesion molecule (L1CAM L1) and integrin-alpha 1 (ITGA1) (Figure 4C). After ontology and pathway analysis, PINCH2 was found to regulate cancer cell migration and invasion by participating in the IPP complex (ILK, PARVA, and PINCH) (Figure 5). PINCH2 also satisfied significant criteria for functional analysis as follows: a) higher than 30% frequency difference of aberration, b) no proliferation or apoptosis function, and c) belonging to the secretary protein related pathway.

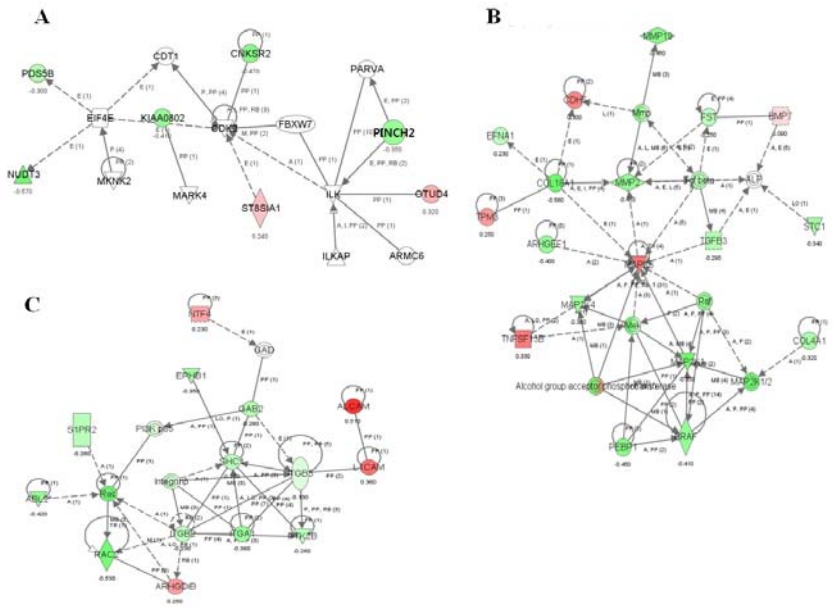


Figure 4. Designated pathways with selected 89 metastasis-related CNVs. We designed (A) an extracellular matrix (ECM) degradation-associated pathway that contained IPP (ILK, PINCH2, and PARVA) complex, (B) a cell motility-related pathway containing MAPK, MMP2, TGFB3, MMP-19, and Raf genes, and (C) a cell-to-cell signaling and interaction pathway containing ALCAM, ITGB2, ITGB3, L1CAM L1, and ITGA1. The red marker denotes amplification and the green marker denotes deletion in the recurred group. To complete the pathway, the previously reported mediator genes (white marker) have been included in the pathway.

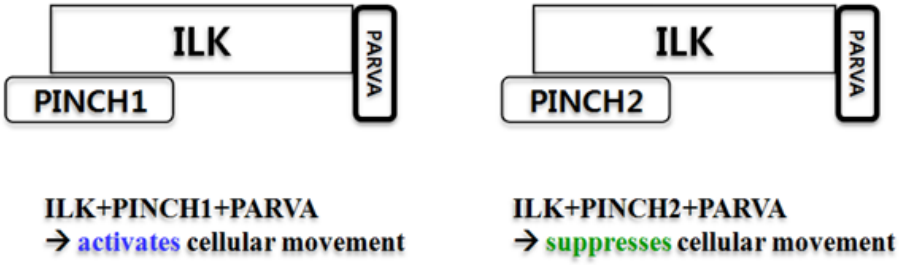


Figure 5. Diagram of ILK-PINCH-PARVA (IPP) complex. PINCH1 is known to increase cellular movement (left), whereas PINCH2 may suppress cellular movement by inhibiting IPP (ILK, PINCH1, and PARVA) complex formation (right). ILK: integrin-linked kinase; PINCH1: LIM and senescent cell antigen-lie domains 1; PINCH2: LIM and senescent cell antigen-lie domains 2; PARVA: α-parvin

6. Suppression of PINCH2 gene expression by siRNA

To observe the difference in PINCH2 expression between normal and cancer cells, we performed RT-PCR and western blotting using normal colon cells (CCD-841 CoN and CCD-18 Co) and cancerous colon cells (DLD-1 and HCT-116). Both mRNA and protein of PINCH2 were highly expressed in the normal cells as compared to the colon cancer cells (Figure 6). Notably, PINCH2 mRNA expression was 3-times down-regulated and protein was 10-times down-expressed in CCD-841 CoTr compared with CCD-841 CoN (Figure 6B). As PINCH2 was down-expressed in cancer cells, we selected HCT-116 cancer cells to perform functional analysis of PINCH2 using targeted siRNA. The PINCH2 specific sequence of the LIMS1 domain was used as a target of siRNA (Figure 7). Three anti-PINCH2 siRNAs were effectively suppressed of PINCH2 expression in both mRNA and protein levels (Figure 8). Thus, we selected the closest sequence from the PINCH2-specific LIM1 domain site, siRNA target site 1, for further biological validation. Through target site 1 suppression, mRNA and protein expression were repressed for 48 hours and PINCH2 was re-expressed for 72 hours after siRNA treatment (Figure 8B~8D). Based on these results, we selected 200 pmol of concentration and 48 hours of treatment time with siRNA site 1 for further functional experiments.

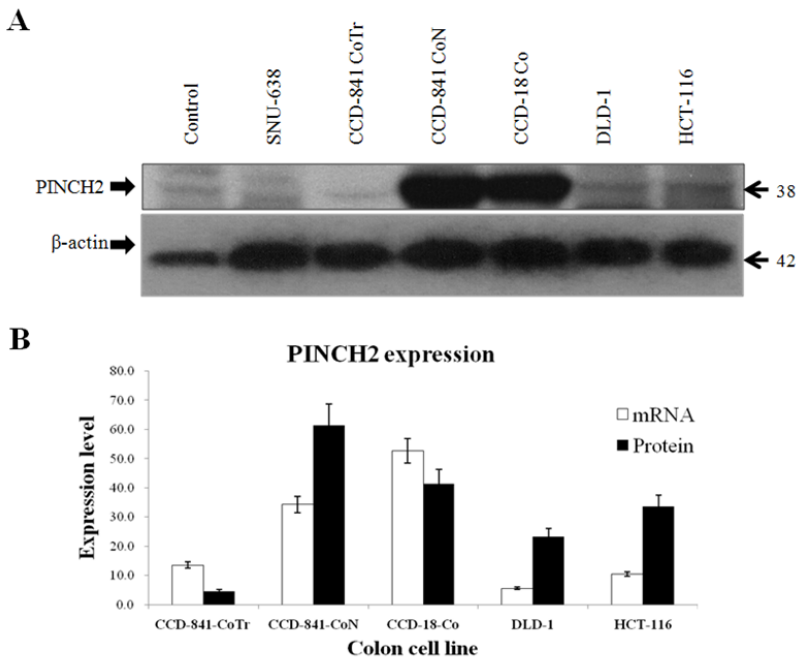


Figure 6. Comparison of PINCH2 mRNA and protein expression levels in cell lines. A. Comparison of PINCH2 protein expression levels in each cell by western blotting. B. Quantitative comparison of PINCH2 mRNA and protein expression in each cell. mRNA and protein expression of PINCH2 were higher in normal colon cells than in cancer cells (2.5-fold and 1.8-fold, respectively). In CCD-841 CoTr, PINCH2 expression was down-regulated 25% of mRNA and 55% of protein compared to CCD-841 CoN. Mouse liver tissue and SNU-638 gastric cancer cells were used as PINCH2 positive cells. Columns: expression level; white: mRNA expression; black: protein expression; bars: SD.

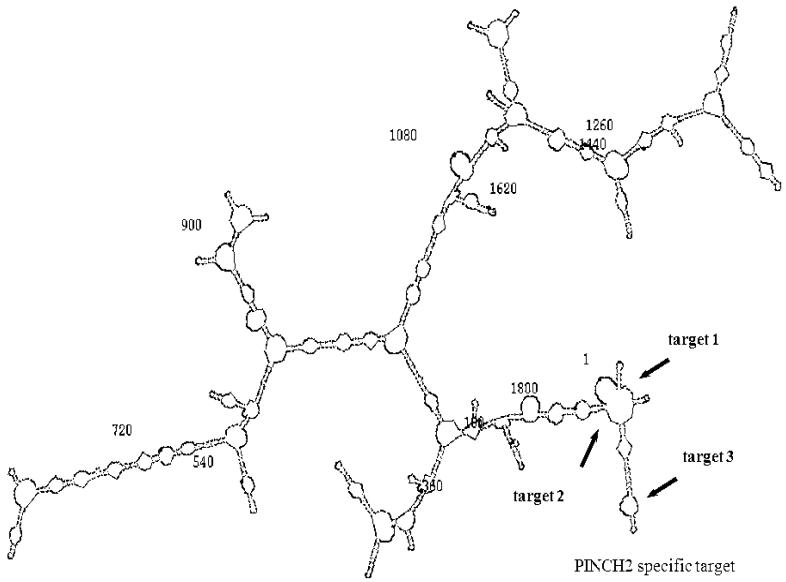


Figure 7. Secondary structure of PINCH2 mRNA. Different locations of 3 designed siRNAs within the PINCH2 specific mRNA transcript. Arrows represent each siRNA target site with PINCH2 specific LIM1 domain sequence. Target 1; (RNA)-GGUGCUACGAGAAGUUCCCGCUGGA, target 2; (RNA)-GCGAUGUGGUGUCGGCCCUCAACAA), target 3 (RNA)-CCACGGGCAAUAUGUCGGACGCCUU.

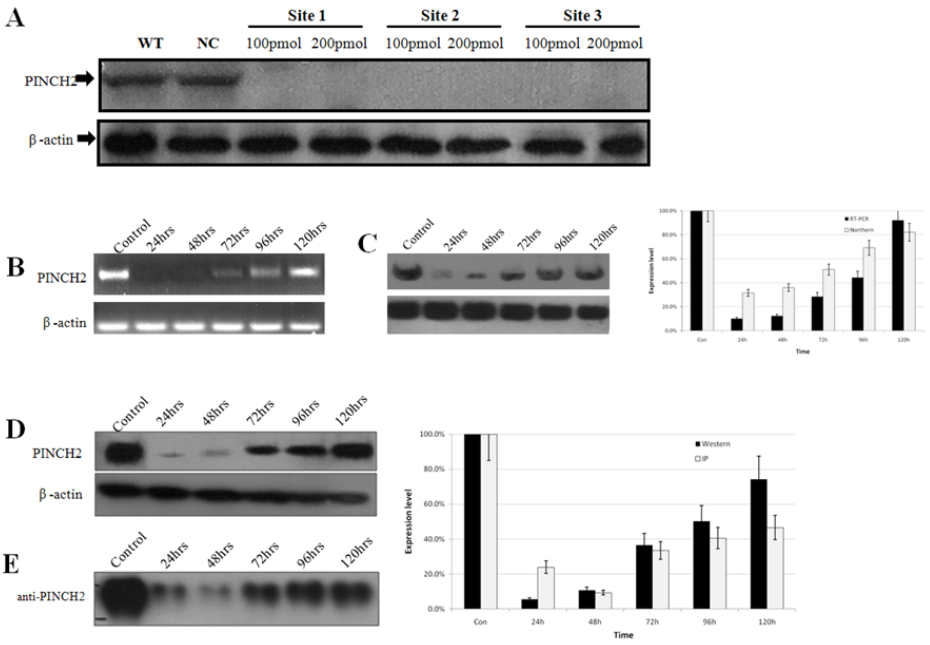


Figure 8. Suppression of PINCH2 expression by siRNA in HCT-116 cells. A. Inhibition of PINCH2 protein expression by each siRNA at 24 hours. We treated three types of siRNA sequence (site 1, site 2 and site 3) with two different concentrations (100 pmol and 200 pmol) in HCT-116 cells. B, C. mRNA expression of PINCH2 with RT-PCR (B) and northern blotting (C). PINCH2 was suppressed in all cases compared to the controls at 24, 48 hours and PINCH2 was recovered after 72 hours with 200 pmol concentration. D, E. Restoration of protein expression. Protein expression was recovered after 48 hours both western-blotting (D) and immunoprecipitation (E). WT: wild type; NC: negative control; bars: SD.

7. Comparison of *in silico* and *in vitro* binding affinities between PINCH1 and PINCH2 with ILK

With a blast sequence similarity search, we confirmed the difference in the 1^{st}, 2^{nd}, 3^{rd}, 4^{th}, 7^{th}, 18^{th}, and 46^{th} amino acid sequences of the LIM1 domain between PINCH1 and PINCH2 (Figure 9A). Then we predicted the PINCH2 protein structure based on the PINCH1 protein sequence and structure (Figure 9B). After computational modeling, we found that the zinc finger of PINCH2 has a different secondary structure shape compared to PINCH1 due to a sequence difference in 46th amino acid. Based on the differing sequence and structure, we calculated the score of computational binding affinity, or the ZRANK Score, using both PINCH1 and PINCH2 with ILK. The PINCH2 (Score=-81.019) binding affinity with ILK was higher than that of PINCH1 (Score=-65.645). But in the score of computational binding strength, R Dock Score, PINCH1 (Score=0.89698) had a higher binding strength with ILK than PINCH2 (Score=3.64549) (Figure 10). This means that binding affinity of PINCH2-ILK is higher than PINCH1-ILK although PINCH1 has a higher binding stability than PINCH2. In HCT-116 cells, PINCH2 formed IPP complex with ILK and PARVA, even if PINCH1 was expressed 2.5-fold more than PINCH2. And the IPP complex composition was changed to ILK, PARVA, and PINCH1 in PINCH2 suppressed cells (Figure 11A). After PINCH2 suppression, the amount of PINCH2-ILK complex decreased to 1/10, while the amount of PINCH1-ILK complex increased 4-fold (Figure 11B). As expected from the *in silico* prediction in natural state of HCT-116 cells, PINCH2 has a higher binding affinity than PINCH1 in the formation of IPP complex. On the contrary, a change in PINCH2 protein expression has no effect on the protein expression of ILK and PARVA (Figure 11).

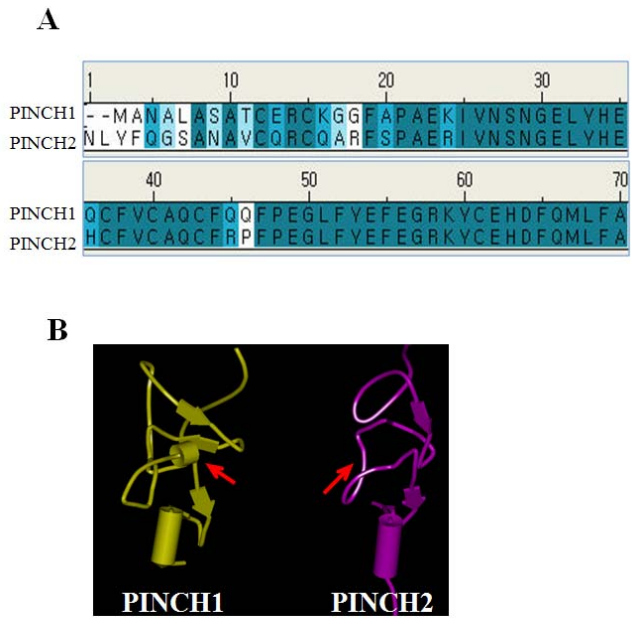
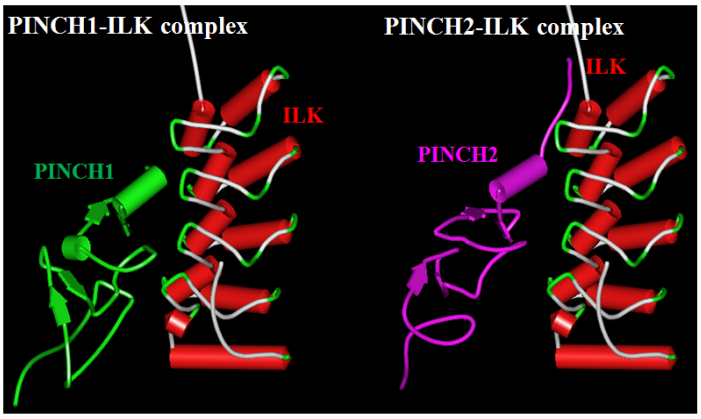


Figure 9. Difference in molecular structure of PINCH1 and PINCH2. A. Alignment of LIM1 domain sequence with PINCH1 (GeneBank Accession no. P48059) and PINCH2 (GeneBank Accession no. Q7Z4I7). The 1st, 2nd, 3rd, 4th, 7th, 18th, and 46th amino acids showed differing protein sequences between PINCH1 and PINCH2. White amino acid denoted different sequences between PINCH1 and PINCH2. B. Superposition of LIM1 domain for PINCH1 and PINCH2. The extent of the N- and C-terminal zinc fingers (zinc fingers 1 and 2, respectively) are shown. Due to the difference in the 46th amino acid, the PINCH1 zinc finger showed different secondary structure as compared to the PINCH2 zinc finger (red arrow), which is composed of 46th, 47th, and 48th amino acids. Helical structure denotes cylinders; beta-sheet arrow denotes pointing in the direction of the C-terminus; coil and turn structures denote tubes.



	PINCH1-ILK		PINCH2-ILK
Binding strength (R Dock Score)	0.89698	>	3.64549
Binding affinity (ZRANK Score)	-65.645	<	-81.019

Figure 10. Comparison of binding strength and affinity between PINCH1 and PINCH2 with ILK. In the computational binding analysis, a lower R Dock Score and a lower ZRNAK Score represent higher binding strength and binding affinity, respectively. The binding affinity of PINCH2-ILK (score= -81.019) was higher than that of PINCH1-ILK (score=-65.645), while the binding strength of PINCH1-ILK (score=0.89698) was lower in PINCH2-ILK (score=3.64549). The red structure represents the ANK (ankyrin) domain of ILK, the green structure represents the PINCH1 LIM1 domain, and the violet structure represents the LIM1 domain of PINCH2. Each score (R Dock and ZRANK Score) was calculated by computational analysis in Gene Discovery Studio software. ILK: integrin-linked kinase; Helical structure denotes cylinders; beta-sheet arrow denotes pointing in the direction of the C-terminus; coil and turn structures denotes tubes.

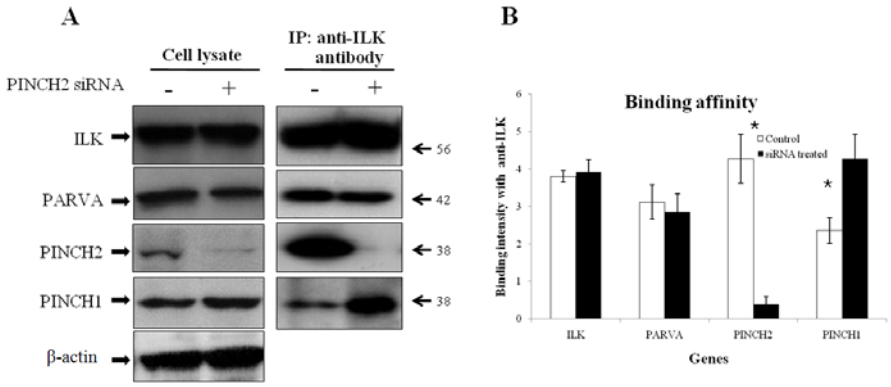


Figure 11. Changes in protein expression and binding affinities between subunits of IPP complex. A. Western blotting. After the suppression of PINCH2, there were no differences in IPP complex subunit protein expressions such as ILK, PARVA and PINCH1 in the cell lysate (left). On the contrary, after PINCH2 suppression, ILK binding affinity with PINCH2 decreased 90%, while PINCH1 binding to ILK increased 4-fold. Under normal conditions, PINCH2-ILK has a 2-fold higher binding affinity than PINCH1-ILK fold (white bar). *: $p<0.01$; Columns: binding intensity with anti-ILK; white: control group; black: siRNA treated group; bars: SD.

8. Changes of down-stream pathway gene expression with IPP complex change

We compared the protein expression of ILK down-stream pathway genes related to cell proliferation (Akt) and cellular movement (matrixmetalloprotenase -9, -11). Proliferation-related Akt and phospho-Akt expressions did not change after PINCH2 suppression (Figure 12). However, after PINCH2 suppression, MMP-9 and MMP-11 expression levels increased by 1.8 ($p<0.01$) and 1.2-fold ($p<0.01$) respectively, as compared to the pre-suppression levels (Figure 12). According to these results, PINCH2 inhibited protein expression of MMP-9 and MMP-11, whose expressions were being regulated by formation of the IPP complex with different types of PINCH1 or PINCH2. Also, we measured secretary MMP-9 and MMP-11 levels in the conditioned medium (CM) with ELISA. MMP-9 secretion was increased in PINCH2 suppressed cells (0.04 ug/ml) as compared to control cells (0.01 ug/ml), but MMP-11 was not detected in CM.

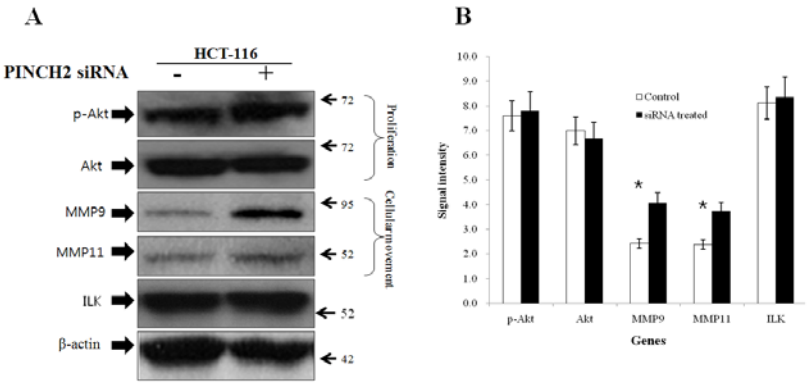


Figure 12. ILK down-stream gene expressions of IPP complex. A. Western blotting of ILK-related proteins. B. Quantitative comparison of signal intensity of the protein expression. Akt and phospho-Akt expressions did not changed after PINCH2 expression. MMP-9 protein expression increased 1.8-fold in the case of PINCH2 siRNA treatment cells compared with control cells (not treated). MMP-11 expression was increased 1.2-fold after PINCH2 siRNA treatment. *: $p<0.01$; Columns: protein expression level; white: control group; black: siRNA treated group; bars: SD.

9. Changes in autocrine activity of cell migration and invasion with PINCH2 suppression

Previously, we verified that the expressions of MMP-9 and MMP-11 were decreased after PINCH2-ILK binding complex formation. Through the wound-healing and matrigel trans-well assay with PINCH2 suppressed cells, we found that autocrine migration (1.5-fold) and invasion (1.7-fold) ability increased after PINCH2 suppression ($p<0.01$, $p<0.01$, respectively) (Figure 13). These findings suggest that the inhibition of PINCH2 expression can promote cell migration and invasion through autocrine activity.

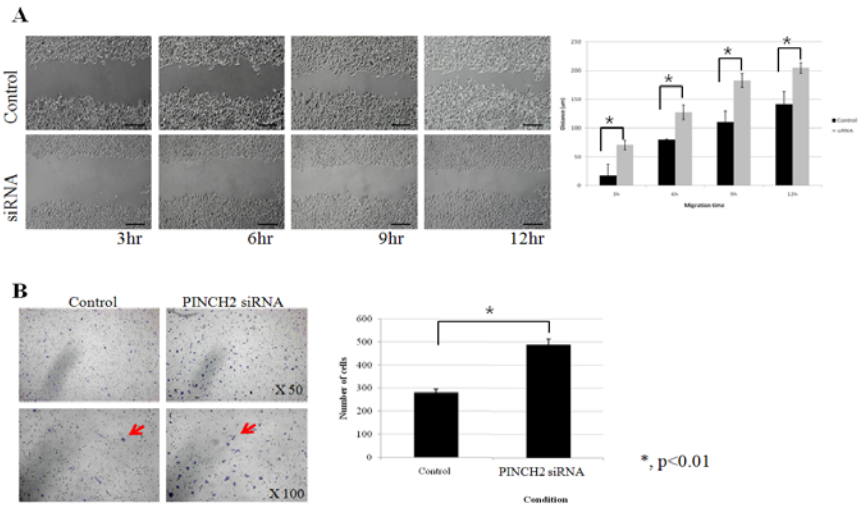


Figure 13. Autocrine activity with suppression of PINCH2 in HCT-116 cells. A. Cell migration assay with wound healing assay. Migration increased 1.5-fold in the PINCH2 knock-down cells. Cell migration distances were calculated and compared with the initial distance from the wound area. B. Invasion assay with Matrigel trans-well assay. HCT-116 cells showed increased invasion capacity with suppression of PINCH2 (1.7-fold). Columns: mean; bars: SD. *: $p < 0.01$, compared with control cells; arrows: invasive cells.

10. Changes in paracrine activity of cell migration and invasion with PINCH2 suppression

Generally, expression of MMP-9, a cell migration-related enzyme, was inhibited by ILK suppression. To evaluate the secretary effect of MMP-9 after PINCH2 suppression, we harvested CM from PINCH2 siRNA-treated HCT-116 cells (siRNA CM) and control HCT-116 cells (control CM). After 24 hours of incubation with CM, wound closure was 1.4-fold faster in siRNA CM-treated cells than control CM-treated cells (Figure 14A). Also, the invasion rate via matrigel trans-well assay increased in siRNA CM-treated cells compared with control CM treated cells as a paracrine activity (1.3-fold, $p<0.05$) (Figure 14B). These finding suggest that the inhibition of PINCH2 expression can alter the IPP complex that controls the MMPs expression in a paracrine function and eventually promotes cell migration and invasion.

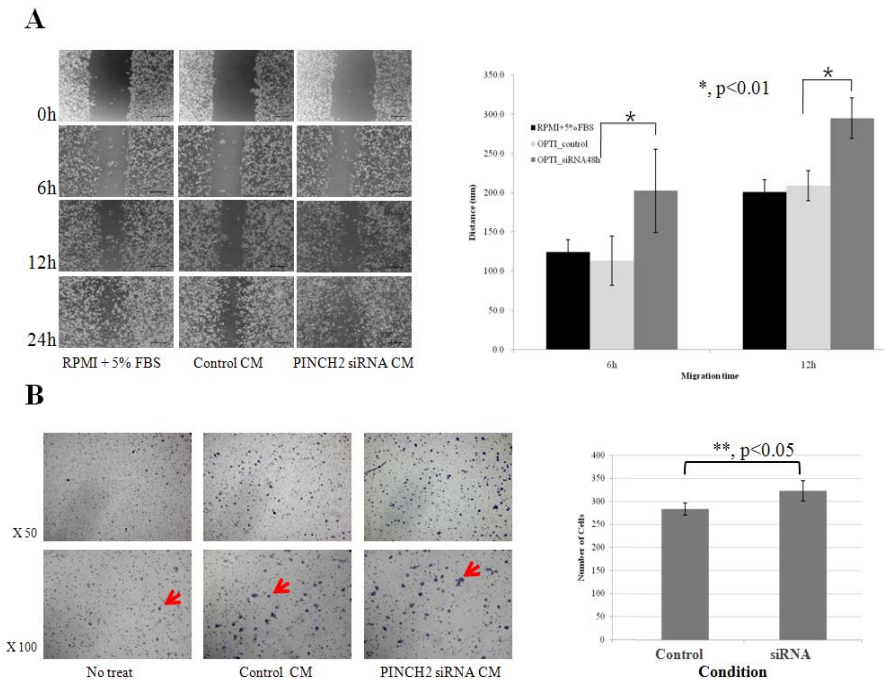


Figure 14. Paracrine activity of CM harvested from PINCH2 suppressed HCT-116 cells. A. Cell migration comparison with wound healing assay. Migration increased 1.4-fold after siRNA treated CM. Cell migration distances were calculated and compared with the initial distance from the wound area. B. Invasion assay with Matrigel trans-well assay. The invasion of HCT-116 cells increased after siRNA treated CM (1.3-fold). Control CM: conditioned media with HCT-116 cells; siRNA CM: conditioned media with PINCH2 siRNA treated HCT-116 cells; Columns: mean; bars: SD. *: $p<0.01$; **: $p<0.05$, compared with control HCT-116; arrows: invasive cells.

IV. DISCUSSION

DNA microarray is a useful tool for the study of complex genome systems and has applications in a wide variety of biological sciences (61-63). Detection of genomic alteration by array-CGH can easily be translated into both sequence and gene identification, which can provide additional information concerning complex chromosomal rearrangements and imbalances. For determine the cut-off level of genetic variation, array-CGH was accomplished by self-self hybridization with male and female genomic DNA (gDNA) as reference DNA to reduce experimental variation. Previous report, we determined the cut-off range of normal genetic aberration as 1.5-fold (-0.58 ~ 0.58 $\log_2$ ratio) (51, 52). Next, we performed sex-mismatched experiments for the determination of the cut-off range of single copy number change (2-fold difference). Based on these results, We confirmed that our microarray platform had ~1% occurrence of a technical false-positive rate and we confirmed that array-CGH is a useful method for the analysis of genomic profiling (64).

Using this cut-off level, we screened the distribution of genomic disorder (CNVs) from normal colon tissues. With 44K oligonucleotide microarray, 50~55% probes showed copy number change patterns from more than one sample. Some CNVs may be the result of multiple repeat within a species rather than a single ancestral duplication or deletion event followed by change of frequency (16). This result suggests that the frequency and patterns of genetic aberration vary widely, and definition of genetic aberration needs precise cut-off range of aberration frequency. In conventional CGH, 20~30% of aberration frequencies were determined as a range of genomic disorder and more than 50% of aberration frequency was considered to be a highly aberrant genomic region (65, 66). When we analyze the aberration frequencies of CNVs in our study, we found that 326 probes showed more than 50% (21/41

samples) of aberration frequency. Especially, 8 probes had a high genetic aberration frequencies in more than 80% (33/41 samples).

Gene duplication is known to be an important long-term evolutionary force, and thus, some lineage-specific copy number differences may contribute to the phenotypic differences among taxa, including those that distinguish humans from chimpanzees and bonobos (Pan paniscus) (16, 67, 68). Both CNVs and CNVRs can be caused by genomic rearrangements such as deletions, duplications, inversions, and translocations (69). Moreover, DNA size, orientation, and genetic distances of variation are important factors for determination of genetic function (70). Therefore, CNVR will also be an important implication as a marker of individual difference such as disease or drug sensitivity. Generally, 10~30% aberration frequencies were recommended as a cut-off frequency for CNVRs (DGV database, Database of Genomic Variants) (http://projects.tcag.ca/variation).

Although few reported CNVs or CNVRs implied important biological function, most of their functions were unknown yet (71-73). DNA copy number changes had various aberration patterns based on individual differences such as amplification or deletion of gDNA, affecting the gene expression up- or down-regulation, respectively. However, a complex pattern of CNVs (mixed with amplification and deletion in same population) makes it difficult to predict gene expression pattern and its biological function in the same population. With the selected CNVs, we found the 21 copy number variation region (CNVR), which showed genetic aberration with multiple probes. They showed CNV patterns with more than 20% frequencies and neighbored at least five probes with longer than 100kb length as a whole (Table 2). Among them, BOLA2 (bolA-like 2), DMD (dystrophin), CSMD1 (CUB and Sushi multiple domains 1), ALCAM (activated leukocyte cell adhesion molecule), and NBPF1 (neuroblastoma breakpoint family, member 1)

had CNVs in more than seven probes with a mean of 760.2kb genomic length. These CNVR were already reported in DGV database and may have important implications for individual variation with changes in gene function such as immune reaction, physiological difference, and tumor development. CNVRs and CNVs may contribute to genetic disease susceptibility. Unfortunately, so far, disease-related CNVs and CNVRs detection study would most likely be successful only in patients with monogenic diseases, such as psychiatric or neurological disorders. Selected CNVRs are related to human health, such as the cancer-relevant as well as genes potentially associated with other complex phenotypes, including congenital malformations, neurological disorder, and psychiatric disorder. Human disease-related CNVRs and CNVs that are currently known are those influencing one's susceptibility to HIV infection (72), modulating drug responses (73) and contributing to genomic microdeletion and duplication syndromes (71). Selected BOLA2 was known as a potential candidate gene in control of immune response to vaccines with DNA polymorphism (74).

To select cancer metastasis-related CNVs, we compared genomic profiling between the recurred and non-recurred groups and selected 1,359 candidates of metastasis-related CNVs with various biological functions (Figure 1). Among them, 190 CNVs were related to cancer biology and 142 CNVs were associated with neurological disease. The remaining genes have various biological functions such as cell signaling, hematologic disease, and infectious disease. In order to study the functional mechanism of CNVs, we selected 89 possible cell movement-related CNVs by gene ontology (Figure 2). For technical validation of this array-CGH for CNVs selection, we choose GNLY and PINCH2, which were known as CNVs in DGV database, as a target genes with validation. Additionally, cancer migration-related TGFB3 (75, 76) and cancer invasion-related MMP2 (77, 78) were selected as positive control

genes. According to the independent real-time PCR analysis, array-CGH and qPCR showed similar copy number change patterns and levels (within the 2-fold difference). Also, mRNA expression patterns showed the same pattern as those of DNA changes, even if expression levels were different from gene to gene (Figure 3). This suggests that the DNA copy number is correlated with mRNA expression and that array-CGH provides a reliable technique in selecting CNVs.

We categorized the 89 CNVs into the three pathways including cell migration/invasion pathways, cellular movement-related pathways, and cell-to-cell signaling and interaction pathways using Core Analysis in IPA 5.5 (Figure 4). According to the analysis, we set up several parameters of gene selection for functional analysis such as aberration frequency, gene ontology class and gene-related pathway. In these 89 CNVs, PINCH2 satisfied the recommended criteria as follows: a) amplified more in the non-recurred group (>30% frequency), b) categorized into cell migration/invasion pathway (Figure 4), c) located in both cytoplasm and nucleus and participating in IPP complex (ILK, PARVA, and PINCH) to regulate cell movement associated with MMPs (Figure 5). The IPP complex binds to integrin (79-81). Both the IPP complex and PINCH family are essential for integrin-mediated cell adhesion. The ILK-PINCH family has a role in cytoskeletal organization and adhesion signaling (82). A ternary IPP complex with PINCH1, ILK, and PARVA regulates cell migration and invasion with Rsu-1 inhibition and Ras activation. Also, dominant-negative ILK can regulate both integrin-associated rearrangements of actin filaments, cell migration, and invasion at the integrin receptor-proximal region (83-85). Therefore, IPP complex is a crucial factor for many ILK-mediated processes such as regulation of cell adhesion, cytoskeleton, and apoptosis.

There are 2 PINCH genes in mammals, PINCH1 (LIM and senescent cell

antigen-like domains 1, LIMS1) and PINCH2, which have 5 LIM domains encoding closely related proteins (43, 85). PINCH1 is a widely expressed evolutionary conserved focal adhesion protein, but PINCH2 expression has been restricted to the bladder, stomach, and intestines at these embryonic stages. Also, PINCH1 is localized to epithelial cells and smooth muscle layers, whereas PINCH2 is confined to the smooth muscle layer (86). PINCH1 binds with ILK (integrin-linked kinase) to increase cell migration or invasion without PINCH2 expression. ILK was known to be a protein related to apoptosis, proliferation, migration, and motility. It regulates the expression of Akt, GSK3B, CCND1, Gsk3, and integrin beta with PI3K/Akt/p70S6K1/Rac1 signaling pathway on actin filaments and cell migration (86-89).

There was no difference in the DNA copy number of PINCH1 between the recurred and non-recurred patients, whereas PINCH2 was amplified in the non-recurred group, suggesting that PINCH2 might suppress cell migration and invasion through IPP complex formation (Figure 5). The nucleotide and amino acid sequences of PINCH2 were similar to PINCH1 except in the LIM1 domain (Figure 9A). These LIM1 domains construct the 2 zinc finger proteins, which act in protein-protein interactions (90). From a different amino acid sequence at the 46^{th}, PINCH1 and PINCH2 had a different shape of zinc finger which was composed of 46^{th}, 47^{th}, and 48^{th} amino acid (Figure 9B). Based on the different sequences and structures between PINCH1 and PINCH2, we calculated the computational binding affinity of PINCH1-ILK, and PINCH2-ILK. PINCH2 was found to have a higher binding affinity and lower binding strength than PINCH1 for ILK (Figure 10). To compensate for the limitations of X-ray crystallography structure of PINCH1 and PINCH2 and computational binding assay, we performed an immunoprecipitation assay (Figure 11). According to *in vitro* screening, mRNA and protein expression of PINCH2 were up-regulated in normal colon epithelial cells (CCD-841 CoN) as

compared with highly aggressive transformed colon cells (CCD-841 CoTr, SV40 transformed CCD-841 CoN). Although PINCH2 was highly expressed in normal cells, this was not suitable for functional study of PINCH2 due to low transfection efficiency and slow growth rate. Thus, we selected HCT-116 cells to investigate the molecular function of PINCH2.

As three siRNAs showed similar gene suppression ability, we selected the site 1 siRNA because it is located closest to the promoter region to avoid the off-target effects of siRNA (91) (Figure 8). There was no change in PINCH1 expression after siRNA treatment despite the fact that PINCH1 has a similar structure to PINCH2 (Figure 11). Thus, we confirmed that site 1 siRNA has a PINCH2 specific suppression activity. We also confirmed that ILK-PINCH1 complex formation increased after PINCH2 suppression. This result shows that PINCH2 forms the IPP complex regardless of PINCH1 expression, as we expected from computational analysis. With siRNA treatment, mRNA expression of PINCH2 was completely suppressed until 72 hours, while PINCH2 protein expression was restored after 48 hours. Time difference in recovery patterns between mRNA and protein may result from target sequence polymorphism during gene silencing approaches such as siRNA, ribozyme, and shRNA (92, 93). Additionally, some genes have different gene regulation mechanisms, which can lead to nonspecific regulation of protein expression and RNA degradation in siRNA and shRNA assay (94, 95).

In epithelial cells, over expression of ILK leads to fibroblastoid and invasive phenotype changes with activation of E-cadherin (88, 96). Previously, MMP-9, a secretary protein, plays a role in cell migration and invasion-related activities with the extracellular matrix (ECM), expression was known to decrease in ILK suppression (87), but the relationship between PINCH2 and MMPs is still unknown. To validate the relationship between PINCH2 and MMP-9 and MMP-11, we detected the cellular protein levels of MMP-9 and

MMP-11 regardless PINCH2 expression. MMP-9 and MMP-11 protein levels increased by 1.8 ($p<0.01$) and 1.2-fold ($p<0.01$) in PINCH2 suppressed cells, respectively (Figure 12). To evaluate secretary MMP-9 and MMP-11, conditioned medium (CM) was studied after 10-fold concentration with ELISA. Secretary MMP-9 was found to increase in CM of PINCH2 siRNA treated group (0.04 ㎍/㎖) when compared with the control group (0.01 ㎍/㎖) as in Western blot analysis. However, due to the small amount of secretion, MMP-11 protein was not detected in CM of the two groups. According to wound-healing and matrigel trans-well assay, migration and invasion increased as an autocrine activity in PINCH2 suppressed cells (Figure 13). Also, wound closure rate increased in 1.4-fold and invasion rate increased in 1.3-fold in PINCH2 siRNA-treated CM-treated cells as compared to the control CM-treated cells (Figure 14). This means that the ILK-PINCH2 complex suppressed the expression of MMP-9 and -11 compared to the ILK-PINCH1 complex. As a result, the ILK-PINCH2 complex has the inhibitory mechanism of cell migration or invasion through suppression of MMP-9 expression compared to ILK-PINCH1 in both autocrine and paracrine activity.

Previously, ILK has been demonstrated to possess a number of oncogenic properties such as induction of anchorage-independent cell cycle progression, inhibition of apoptosis, and enhancement of tumorigenicity in nude mice (90, 97, 98). Generally, Akt and phospho-Akt were known to have cell proliferation-related function with IKL protein. But, the expression of Akt and phospho-Akt expression did not change after the suppression of PINCH2 (Figure 12). Our results suggested that ILK-PINCH2 complex did not affect the cell proliferation-related function. This means that the PINCH2-related IPP complex was not involved in the cell proliferation and apoptosis-related pathway.

V. CONCLUSION

We evaluated the molecular function of the PINCH2 variation in colon cancer metastasis as an autocrine and paracrine activity in normal colon tissues using high-throughput array-CGH analysis using the following steps: (a) screening of CNVs from 41 normal tissues, (b) selection of 89 cancer metastasis-related CNVs, including PINCH2, (c) *in silico* and *in vitro* verification of the molecular action mechanism of PINCH2, (d) validation of biological function of PINCH2 in cancer cell migration and invasion. As a results, PINCH2, an ILK binding protein, suppressed the MMP-9 and MMP-11 expression and the ILK-PINCH2 complex inhibited cell migration and invasion though autocrine and paracrine activity. In conclusion, metastasis-related CNVs such as PINCH2 play a role in cancer cell migration and invasion through autocrine and paracrine function in normal tissue. These molecular findings support that genetically altered environment tumor cells play a key role in tumor recurrence and progression.

REFERENCES

1. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, et al. Accurate whole human genome sequencing using reversible terminator chemistry. Nature 2008;456(7218):53-9.
2. Levy S, Sutton G, Ng PC, Feuk L, Halpern AL, Walenz BP, et al. The diploid genome sequence of an individual human. PLoS Biol 2007;5(10):e254.
3. Chance PF, Alderson MK, Leppig KA, Lensch MW, Matsunami N, Smith B, et al. DNA deletion associated with hereditary neuropathy with liability to pressure palsies. Cell 1993;72(1):143-51.
4. Carter NP. Methods and strategies for analyzing copy number variation using DNA microarrays. Nat Genet 2007;39(7 Suppl):S16-21.
5. Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU, et al. Segmental duplications and copy-number variation in the human genome. Am J Hum Genet 2005;77(1):78-88.
6. Tuzun E, Sharp AJ, Bailey JA, Kaul R, Morrison VA, Pertz LM, et al. Fine-scale structural variation of the human genome. Nat Genet 2005;37(7):727-32.
7. Zhang F, Gu W, Hurles ME, Lupski JR. Copy number variation in human health, disease, and evolution. Annu Rev Genomics Hum Genet 2009;10:451-81.
8. Lupski JR, Stankiewicz P. Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. PLoS Genet 2005;1(6):e49.
9. Lupski JR. Genomic rearrangements and sporadic disease. Nat Genet 2007;39(7 Suppl):S43-7.
10. Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, et al. Detection of large-scale variation in the human genome. Nat Genet 2004;36(9):949-51.

11. Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, et al. Large-scale copy number polymorphism in the human genome. Science 2004;305(5683):525-8.
12. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, et al. Global variation in copy number in the human genome. Nature 2006;444(7118):444-54.
13. Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, et al. Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science 2007;315(5813):848-53.
14. Wain LV, Armour JA, Tobin MD. Genomic copy number variation, human health, and disease. Lancet 2009;374(9686):340-50.
15. Braude I, Vukovic B, Prasad M, Marrano P, Turley S, Barber D, et al. Large scale copy number variation (CNV) at 14q12 is associated with the presence of genomic abnormalities in neoplasia. BMC Genomics 2006;7:138.
16. Freeman JL, Perry GH, Feuk L, Redon R, McCarroll SA, Altshuler DM, et al. Copy number variation: new insights in genome diversity. Genome Res 2006;16(8):949-61.
17. Mamtani M, Anaya JM, He W, Ahuja SK. Association of copy number variation in the FCGR3B gene with risk of autoimmune diseases. Genes Immun 2009;11(2):155-60.
18. Inoue K, Lupski JR. Molecular mechanisms for genomic disorders. Annu Rev Genomics Hum Genet 2002;3:199-242.
19. Wolf NI, Sistermans EA, Cundall M, Hobson GM, Davis-Williams AP, Palmer R, et al. Three or more copies of the proteolipid protein gene PLP1 cause severe Pelizaeus-Merzbacher disease. Brain 2005;128(Pt 4):743-51.
20. del Gaudio D, Fang P, Scaglia F, Ward PA, Craigen WJ, Glaze DG, et

al. Increased MECP2 gene copy number as the result of genomic duplication in neurodevelopmentally delayed males. Genet Med 2006;8(12):784-92.

21. Bi W, Sapir T, Shchelochkov OA, Zhang F, Withers MA, Hunter JV, et al. Increased LIS1 expression affects human and mouse brain development. Nat Genet 2009;41(2):168-77.
22. Shlien A, Malkin D. Copy number variations and cancer susceptibility. Curr Opin Oncol 2009;22(1):55-63.
23. Stransky N, Vallot C, Reyal F, Bernard-Pierrot I, de Medina SG, Segraves R, et al. Regional copy number-independent deregulation of transcription in cancer. Nat Genet 2006;38(12):1386-96.
24. Shlien A, Malkin D. Copy number variations and cancer. Genome Med 2009;1(6):62.
25. Grady WM. Genomic instability and colon cancer. Cancer Metastasis Rev 2004;23(1-2):11-27.
26. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 2004;351(27):2817-26.
27. Qin LX, Tang ZY. Recent progress in predictive biomarkers for metastatic recurrence of human hepatocellular carcinoma: a review of the literature. J Cancer Res Clin Oncol 2004;130(9):497-513.
28. Bange J, Prechtl D, Cheburkin Y, Specht K, Harbeck N, Schmitt M, et al. Cancer progression and tumor cell motility are associated with the FGFR4 Arg(388) allele. Cancer Res 2002;62(3):840-7.
29. Luangdilok S, Box C, Patterson L, Court W, Harrington K, Pitkin L, et al. Syk tyrosine kinase is linked to cell motility and progression in squamous cell carcinomas of the head and neck. Cancer Res 2007;67(16):7907-16.

30. Fernandez-Cebrian JM, Nevado Santos M, Vorwald Kuborn P, Pardo de Lama M, Martin-Cavanna J, Pacheco Martinez P, et al. Can the clinical outcome in stage II colon carcinomas be predicted by determination of molecular marker expression? Clin Transl Oncol 2007;9(10):663-70.

31. Fang YJ, Lu ZH, Wang GQ, Pan ZZ, Zhou ZW, Yun JP, et al. Elevated expressions of MMP7, TROP2, and survivin are associated with survival, disease recurrence, and liver metastasis of colon cancer. Int J Colorectal Dis 2009;24(8):875-84.

32. Corona G, Deiana M, Incani A, Vauzour D, Dessi MA, Spencer JP. Hydroxytyrosol inhibits the proliferation of human colon adenocarcinoma cells through inhibition of ERK1/2 and cyclin D1. Mol Nutr Food Res 2009;53(7):897-903.

33. Bell SM, Scott N, Cross D, Sagar P, Lewis FA, Blair GE, et al. Prognostic value of p53 overexpression and c-Ki-ras gene mutations in colorectal cancer. Gastroenterology 1993;104(1):57-64.

34. Benhattar J, Losi L, Chaubert P, Givel JC, Costa J. Prognostic significance of K-ras mutations in colorectal carcinoma. Gastroenterology 1993;104(4):1044-8.

35. Ogunbiyi OA, Goodfellow PJ, Herfarth K, Gagliardi G, Swanson PE, Birnbaum EH, et al. Confirmation that chromosome 18q allelic loss in colon cancer is a prognostic indicator. J Clin Oncol 1998;16(2):427-33.

36. Al-Mulla F, AlFadhli S, Al-Hakim AH, Going JJ, Bitar MS. Metastatic recurrence of early-stage colorectal cancer is linked to loss of heterozygosity on chromosomes 4 and 14q. J Clin Pathol 2006;59(6):624-30.

37. Muller O. Identification of colon cancer patients by molecular diagnosis. Dig Dis 2003;21(4):315-9.

38. Bhowmick NA, Neilson EG, Moses HL. Stromal fibroblasts in cancer

initiation and progression. Nature 2004;432(7015):332-7.

39. Orimo A, Weinberg RA. Stromal fibroblasts in cancer: a novel tumor-promoting cell type. Cell Cycle 2006;5(15):1597-601.

40. Koutsilieris M, Reyes-Moreno C, Choki I, Sourla A, Doillon C, Pavlidis N. Chemotherapy cytotoxicity of human MCF-7 and MDA-MB 231 breast cancer cells is altered by osteoblast-derived growth factors. Mol Med 1999;5(2):86-97.

41. Choki I, Sourla A, Reyes-Moreno C, Koutsilieris M. Osteoblast-derived growth factors enhance adriamycin-cytostasis of MCF-7 human breast cancer cells. Anticancer Res 1998;18(6A):4213-24.

42. Liang X, Sun Y, Ye M, Scimia MC, Cheng H, Martin J, et al. Targeted ablation of PINCH1 and PINCH2 from murine myocardium results in dilated cardiomyopathy and early postnatal lethality. Circulation 2009;120(7):568-76.

43. Stanchi F, Bordoy R, Kudlacek O, Braun A, Pfeifer A, Moser M, et al. Consequences of loss of PINCH2 expression in mice. J Cell Sci 2005;118(Pt 24):5899-910.

44. Coussens LM, Werb Z. Inflammation and cancer. Nature 2002;420(6917):860-7.

45. Tlsty TD. Stromal cells can contribute oncogenic signals. Semin Cancer Biol 2001;11(2):97-104.

46. Braakhuis BJ, Leemans CR, Brakenhoff RH. Expanding fields of genetically altered cells in head and neck squamous carcinogenesis. Semin Cancer Biol 2005;15(2):113-20.

47. Shen L, Kondo Y, Rosner GL, Xiao L, Hernandez NS, Vilaythong J, et al. MGMT promoter methylation and field defect in sporadic colorectal cancer. J Natl Cancer Inst 2005;97(18):1330-8.

48. Cai WW, Mao JH, Chow CW, Damani S, Balmain A, Bradley A.

Genome-wide detection of chromosomal imbalances in tumors using BAC microarrays. Nat Biotechnol 2002;20(4):393-6.

49. Hurst CD, Fiegler H, Carr P, Williams S, Carter NP, Knowles MA. High-resolution analysis of genomic copy number alterations in bladder cancer by microarray-based comparative genomic hybridization. Oncogene 2004;23(12):2250-63.

50. Oostlander AE, Meijer GA, Ylstra B. Microarray-based comparative genomic hybridization and its applications in human genetics. Clin Genet 2004;66(6):488-95.

51. Pollack JR, Perou CM, Alizadeh AA, Eisen MB, Pergamenschikov A, Williams CF, et al. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat Genet 1999;23(1):41-6.

52. Pollack JR, Sorlie T, Perou CM, Rees CA, Jeffrey SS, Lonning PE, et al. Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. Proc Natl Acad Sci U S A 2002;99(20):12963-8.

53. Shaw-Smith C, Redon R, Rickman L, Rio M, Willatt L, Fiegler H, et al. Microarray based comparative genomic hybridisation (array-CGH) detects submicroscopic chromosomal deletions and duplications in patients with learning disability/mental retardation and dysmorphic features. J Med Genet 2004;41(4):241-8.

54. Moon HJ, Yim SV, Lee WK, Jeon YW, Kim YH, Ko YJ, et al. Identification of DNA copy-number aberrations by array-comparative genomic hybridization in patients with schizophrenia. Biochem Biophys Res Commun 2006;344(2):531-9.

55. Weiss MM, Kuipers EJ, Postma C, Snijders AM, Siccama I, Pinkel D, et al. Genomic profiling of gastric cancer predicts lymph node status and survival. Oncogene 2003;22(12):1872-9.

56. zur Hausen A, van Grieken NC, Meijer GA, Hermsen MA, Bloemena E, Meuwissen SG, et al. Distinct chromosomal aberrations in Epstein-Barr virus-carrying gastric carcinomas tested by comparative genomic hybridization. Gastroenterology 2001;121(3):612-8.

57. J. Sambrook EFFaTM. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, ; 1989.

58. Chen QR, Bilke S, Khan J. High-resolution cDNA microarray-based comparative genomic hybridization analysis in neuroblastoma. Cancer Lett 2005;228(1-2):71-81.

59. Komura D, Shen F, Ishikawa S, Fitch KR, Chen W, Zhang J, et al. Genome-wide detection of human copy number variations using high-density DNA oligonucleotide arrays. Genome Res 2006;16(12):1575-84.

60. Orozco LD, Cokus SJ, Ghazalpour A, Ingram-Drake L, Wang S, van Nas A, et al. Copy number variation influences gene expression and metabolic traits in mice. Hum Mol Genet 2009;18(21):4118-29.

61. Detours V, Dumont JE, Bersini H, Maenhaut C. Integration and cross-validation of high-throughput gene expression data: comparing heterogeneous data sets. FEBS Lett 2003;546(1):98-102.

62. Ramaswamy S, Ross KN, Lander ES, Golub TR. A molecular signature of metastasis in primary solid tumors. Nat Genet 2003;33(1):49-54.

63. Rhodes DR, Barrette TR, Rubin MA, Ghosh D, Chinnaiyan AM. Meta-analysis of microarrays: interstudy validation of gene expression profiles reveals pathway dysregulation in prostate cancer. Cancer Res 2002;62(15):4427-33.

64. Park CH, Jeong HJ, Choi YH, Kim SC, Jeong HC, Park KH, et al. Systematic analysis of cDNA microarray-based CGH. Int J Mol Med 2006;17(2):261-7.

65. Choi JS, Kim SR, Jeon YW, Lee KH, Rha HK. Identification of DNA copy number aberrations by array comparative genomic hybridization in patients with ruptured intracranial aneurysms. J Clin Neurosci 2009;16(2):295-301.

66. Qin YR, Wang LD, Fan ZM, Kwong D, Guan XY. Comparative genomic hybridization analysis of genetic aberrations associated with development of esophageal squamous cell carcinoma in Henan, China. World J Gastroenterol 2008;14(12):1828-35.

67. Locke DP, Segraves R, Carbone L, Archidiacono N, Albertson DG, Pinkel D, et al. Large-scale variation among human and great ape genomes determined by array comparative genomic hybridization. Genome Res 2003;13(3):347-57.

68. Wilson GM, Flibotte S, Missirlis PI, Marra MA, Jones S, Thornton K, et al. Identification by full-coverage array CGH of human DNA copy number increases relative to chimpanzee and gorilla. Genome Res 2006;16(2):173-81.

69. Lee JA, Lupski JR. Genomic rearrangements and gene copy-number alterations as a cause of nervous system disorders. Neuron 2006;52(1):103-21.

70. McCarroll SA, Altshuler DM. Copy-number variation and association studies of human disease. Nat Genet 2007;39(7 Suppl):S37-42.

71. Inoue K, Osaka H, Thurston VC, Clarke JT, Yoneyama A, Rosenbarker L, et al. Genomic rearrangements resulting in PLP1 deletion occur by nonhomologous end joining and cause different dysmyelinating phenotypes in males and females. Am J Hum Genet 2002;71(4):838-53.

72. Gonzalez E, Kulkarni H, Bolivar H, Mangano A, Sanchez R, Catano G, et al. The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. Science 2005;307(5714):1434-40.

73. Ouahchi K, Lindeman N, Lee C. Copy number variants and pharmacogenomics. Pharmacogenomics 2006;7(1):25-9.

74. Glass EJ. Genetic variation and responses to vaccines. Anim Health Res Rev 2004;5(2):197-208.

75. Nawshad A, Medici D, Liu CC, Hay ED. TGFbeta3 inhibits E-cadherin gene expression in palate medial-edge epithelial cells through a Smad2-Smad4-LEF1 transcription complex. J Cell Sci 2007;120(Pt 9):1646-53.

76. Xia W, Mruk DD, Lee WM, Cheng CY. Differential interactions between transforming growth factor-beta3/TbetaR1, TAB1, and CD2AP disrupt blood-testis barrier and Sertoli-germ cell adhesion. J Biol Chem 2006;281(24):16799-813.

77. Peng L, Xing X, Li W, Qu L, Meng L, Lian S, et al. PRL-3 promotes the motility, invasion, and metastasis of LoVo colon cancer cells through PRL-3-integrin beta1-ERK1/2 and-MMP2 signaling. Mol Cancer 2009;8:110.

78. Papi A, Bartolini G, Ammar K, Guerra F, Ferreri AM, Rocchi P, et al. Inhibitory effects of retinoic acid and IIF on growth, migration and invasiveness in the U87MG human glioblastoma cell line. Oncol Rep 2007;18(4):1015-21.

79. Ho B, Bendeck MP. Integrin linked kinase (ILK) expression and function in vascular smooth muscle cells. Cell Adh Migr 2009;3(2):174-6.

80. Medici D, Nawshad A. Type I collagen promotes epithelial-mesenchymal transition through ILK-dependent activation of NF-kappaB and LEF-1. Matrix Biol 2009;29(3):161-5.

81. Thiery JP. Epithelial-mesenchymal transitions in development and pathologies. Curr Opin Cell Biol 2003;15(6):740-6.

82. Qian Y, Zhong X, Flynn DC, Zheng JZ, Qiao M, Wu C, et al. ILK

mediates actin filament rearrangements and cell migration and invasion through PI3K/Akt/Rac1 signaling. Oncogene 2005;24(19):3154-65.

83. Dougherty GW, Jose C, Gimona M, Cutler ML. The Rsu-1-PINCH1-ILK complex is regulated by Ras activation in tumor cells. Eur J Cell Biol 2008;87(8-9):721-34.

84. Liang X, Sun Y, Schneider J, Ding JH, Cheng H, Ye M, et al. Pinch1 is required for normal development of cranial and cardiac neural crest-derived structures. Circ Res 2007;100(4):527-35.

85. Li S, Bordoy R, Stanchi F, Moser M, Braun A, Kudlacek O, et al. PINCH1 regulates cell-matrix and cell-cell adhesions, cell polarity and cell survival during the peri-implantation stage. J Cell Sci 2005;118(Pt 13):2913-21.

86. Braun A, Bordoy R, Stanchi F, Moser M, Kostka GG, Ehler E, et al. PINCH2 is a new five LIM domain protein, homologous to PINCHand localized to focal adhesions. Exp Cell Res 2003;284(2):239-50.

87. Assi K, Mills J, Owen D, Ong C, St Arnaud R, Dedhar S, et al. Integrin-linked kinase regulates cell proliferation and tumour growth in murine colitis-associated carcinogenesis. Gut 2008;57(7):931-40.

88. Troussard AA, Costello P, Yoganathan TN, Kumagai S, Roskelley CD, Dedhar S. The integrin linked kinase (ILK) induces an invasive phenotype via AP-1 transcription factor-dependent upregulation of matrix metalloproteinase 9 (MMP-9). Oncogene 2000;19(48):5444-52.

89. Wong RP, Ng P, Dedhar S, Li G. The role of integrin-linked kinase in melanoma cell migration, invasion, and tumor growth. Mol Cancer Ther 2007;6(6):1692-700.

90. Chiswell BP, Zhang R, Murphy JW, Boggon TJ, Calderwood DA. The structural basis of integrin-linked kinase-PINCH interactions. Proc Natl Acad Sci U S A 2008;105(52):20677-82.

91. Fedorov Y, Anderson EM, Birmingham A, Reynolds A, Karpilow J, Robinson K, et al. Off-target effects by siRNA can induce toxic phenotype. RNA 2006;12(7):1188-96.

92. Lewin AS, Drenser KA, Hauswirth WW, Nishikawa S, Yasumura D, Flannery JG, et al. Ribozyme rescue of photoreceptor cells in a transgenic rat model of autosomal dominant retinitis pigmentosa. Nat Med 1998;4(8):967-71.

93. Chi JT, Chang HY, Wang NN, Chang DS, Dunphy N, Brown PO. Genomewide view of gene silencing by small interfering RNAs. Proc Natl Acad Sci U S A 2003;100(11):6343-6.

94. Bridge AJ, Pebernard S, Ducraux A, Nicoulaz AL, Iggo R. Induction of an interferon response by RNAi vectors in mammalian cells. Nat Genet 2003;34(3):263-4.

95. Sledz CA, Holko M, de Veer MJ, Silverman RH, Williams BR. Activation of the interferon system by short-interfering RNAs. Nat Cell Biol 2003;5(9):834-9.

96. Novak A, Hsu SC, Leung-Hagesteijn C, Radeva G, Papkoff J, Montesano R, et al. Cell adhesion and the integrin-linked kinase regulate the LEF-1 and beta-catenin signaling pathways. Proc Natl Acad Sci U S A 1998;95(8):4374-9.

97. Attwell S, Roskelley C, Dedhar S. The integrin-linked kinase (ILK) suppresses anoikis. Oncogene 2000;19(33):3811-5.

98. Wu LS, Wang LD, Chen PW, Chen LJ, Tzen JT. Genomic cloning of 18 kDa oleosin and detection of triacylglycerols and oleosin isoforms in maturing rice and postgerminative seedlings. J Biochem 1998;123(3):386-91.

<ABSTRACT(IN KOREAN)>

대장암에서 IPP 복합체의 아단위인 PINCH2 유전자 복제수 변이의 주변분비 활성에 의한 암전이 기전연구

<지도교수 정현철>

연세대학교 대학원 의과학과

박 찬 희

지금까지 인간 게놈에서 다양한 유전적 변이들이 보고되었다. 유전자 복합수 변이 (CNV) 는 1kb 이하 또는 그 이상의 DNA 조각들에서 변이 차이를 나타내는 것을 말한다. 아직까지 CNV와 암과의 연관관계에 대한 연구는 미흡하지만, 40% 정도의 암과 관련된 유전자가 CNV에 속하는 것으로 보고되고 있다. Genome-wide 분석을 통하여 DNA의 복잡성과 연관이 있는 인간 게놈에서 넓게 분포하고 있는 유전자 복합수 변이 (CNVs)에 관한 연구가 가능하게 되었다. 본 연구에서는 array-CGH 실험 방법을 이용하여 대장암의 전이와 연관이 있는 CNV 유전자를 선별하고, 선별된 CNV 유전자에 *in silico, in vitro* 실험을 통하여 기능적 기전을 연구하고자 하였다. 실험에는 22명의 재발이 없는 그룹과 19명의 재발이 있는 그룹으로 구성된 41명의 3기 대장암 환자의 정상 대장 조직으로부터 획득한 DNA를 이용하였다. array-CGH 실험을 통하여, 재발그룹과 재발하지 않는 그룹에서 차이를 나타내는 1,359개의 CNVs 를 선별하였다. 선별된 1,359개의 CNVs 중에서 89개의 암세포의 침윤, 세포의 이동, 그리고 세포간의 신호절달 기전과 관련된 CNVs 유전자들을 선별하였다. 이 중에서 DNA와 mRNA가 동일하게 변하며, 세포의 이동과 침윤의 기전에만 작용하는

것으로 알려진 PINCH2 유전자를 선별하였다. 선별된 PINCH2 유전자는 MMP-9과 같은 분비 단백질을 조절하는 ILK-PARVA-PINCH 복합체(IPP 복합체)의 기능을 조절하는 것으로 알려져 있다. *in vitro* 실험을 통하여 정상 대장 세포에 비하여 대장암 세포에서 mRNA의 발현과 단백질의 발현이 낮은 것을 확인하였다. 데이타베이스 분석과 세포 실험을 통하여 PINCH2 유전자의 IPP 복합체에서, PINCH1이 PINCH2에 비하여 ILK와 결합하는 결합력은 높지만, 친화력은 PINCH2가 PINCH1보다 강한 것을 확인하였다. 뿐만 아니라, PINCH2가 ILK와 결합함으로써, 세포의 이동능과 침윤능에 중요한 역할을 하는 MMP-9 유전자의 발현이 감소하는 것을 관찰하였다. 특히, PINCH2 유전자를 억제하였을 경우, IPP 복합체의 기능 조절을 통하여 대장암 세포인 HCT-116 세포의 이동능과 침윤능이 증가하였으며, PINCH2 발현이 억제된 세포주로부터 획득한 조건부 배양액에서도 역시 주변분비 작용을 통하여 암세포의 이동능과 침윤능이 증가하였다. 결론적으로 암의 전이와 연관이 있을 것으로 판단되는 CNV 유전자인 PINCH2가 IPP 복합체의 결합을 통한 자가분비, 주변분비 작용으로 세포의 이동능과 침윤능을 조절함을 확인하였다.

핵심되는 말 : 유전자 복제변이, Array-CGH, 전이, 대장암, PINCH2, IPP 복합체