The study on the role of chromatin architectural protein CTCF in the control of gene expression by establishing acute CTCF protein depletion system in human colorectal cancer

Moo-Koo Kang

Department of Medical Science

The Graduate School, Yonsei University

The study on the role of chromatin architectural protein CTCF in the control of gene expression by establishing acute CTCF protein depletion system in human colorectal cancer

Moo-Koo Kang

Department of Medical Science

The Graduate School, Yonsei University

The study on the role of chromatin architectural protein CTCF in the control of gene expression by establishing acute CTCF protein depletion system in human colorectal cancer

Directed by Professor Hyoung-Pyo Kim

The Master's Thesis
submitted to the Department of Medical Science
the Graduate School of Yonsei University
in partial fulfillment of the requirements for the degree
of Master of Medical Science

Moo-Koo Kang

December 2020

This certifies that the Master's Thesis of
Moo-Koo Kang is approved.

Thesis Supervisor : Hyoung-Pyo Kim

Thesis Committee Member#1 : Hyun Seok Kim

Thesis Committee Member#2 : Hosung Jung

The Graduate School
Yonsei University

December 2020

ACKNOWLEDGEMENTS

연구를 하는 내내 놀라운 경험을 허락하신 하나님께 감사드립니다.

학부 시절 안성훈 교수님의 세포생물학 강의를 통해 후성유전학을 처음 접하게 되었고, 이 때 느꼈던 두근거림이 저를 연구의 길로 이끌었습니다. 후성유전학을 주로 연구하는 연구실로 감사하게도 지도교수님이신 김형표 교수님 연구실을 알게 되었고, 2018년 초 두 달의 인턴 생활에 이어 입학까지 하게 되었습니다. 교수님께서는 좋은 연구 주제, 좋은 협력 관계, 그리고 마음껏 실험할 수 있는 환경을 경험하게 해주셨고, 이는 연구 기간 동안 정말 큰 축복이었습니다.

학부를 우수한 성적으로 졸업할 정도로 학구열이 높았고, 때문에 스스로 너무 믿었던 입학 초반, 이런 자만은 실험에서도 관계에서도 저를 좋지 않은 방향으로 이끌었습니다. 이 시기에 양철민 박사님께서는 제게 많은 실험들을 가르쳐 주시며 실험 과정에서 다양한 가능성을 제시해 주셨고, 순조로울 때는 다 제 노력의 결과라고, 실험이 풀리지 않을 때는 그래도 잘 하고 있다고 저를 흔들리지 않게 잡아 주셨습니다. 박사님을 통해 실험이나 연구를 넘어 존중을 기반으로 한 삶의 태도를 말 그대로 보고 배울 수 있었습니다.

연구는 개인의 성과물이기 보다는 연구 활동 내내 함께 했던 식구들과 더불어 저에게 도움을 주셨던 모두가 함께 만든 결과라고 당연히 생각합니다. 실험실 생활 동안 스쳐 지나가며 작은 부분에도 관심을 가지고 디스커션 해주신 실험실 식구분들이 계셨기에 부딪혔던 어려움들을 극복해 나갈 수 있었습니다.

짧은 기간에 밀도 높은 실험들을 전수해 주신 민지 누나, 제 말도 안되는 유머에 유일하게 재미있어 해주신 미경 누나, 제가 부화뇌동 할 때 넓은 어깨로 품어 주셨던 웅재 형, 경이로운 풍부한 실험 노하우와 경이로운 파이펫팅 속도를 보여주신 예은 누나, 실험실의 모든 것을 기억하는 보배 누나, 매번 저 스스로를 다시 바라보게끔 해주시는 은총이형, 모든 랩미팅이 명강의 였던 정식이형, 항상 꿍꿍이 가득하고 무심한 듯 무심한 용진이, 삼계탕을 먹는 날엔 양이 많다고 절반을 떼어 주던 수경이, 적시적소에 꿀팁을 주던 경우까지 짧은 글에 담을 수 없는 감사함이 있습니다. 함께 나눈 디스커션들, 실험실 안팎에서 나눴던 대화들, 함께 다녔던 학회들을 떠올리다 보니 본 논문보다도 더 긴 글을 적을 수 있을 듯합니다.

마지막으로 모든 과정 동안 기쁨과 슬픔 모두 진심으로 나누고 응원해주신 아버지, 어머니께 감사 인사를 올립니다. 또 어떤 상황도 다 의미가 있음을 함께 되새기며 서로 페이스 메이커가 되어주는 형에게 감사 인사를 올립니다.

석사 학위 과정을 통해 얻은 소중한 경험을 가지고 어디를 가나 도움 주신 모든 분들께 부끄럽지 않은, 자랑스러운 아들, 친구, 선배, 후배, 제자가 되겠습니다. 감사합니다.

2020년 12월

강무구 올림

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The study on the role of chromatin architectural protein CTCF in the control of gene expression by establishing acute CTCF protein depletion system in human colorectal cancer

Moo-Koo Kang

Department of Medical Science

The Graduate School, Yonsei University

(Directed by Professor Hyoung-Pyo Kim)

CCCTC-binding factor CTCF is a chromatin architecture protein that mediates long-range DNA interactions and defines topologically associated domain boundaries. In order to explore the direct role of CTCF on the transcriptional regulation in the human colorectal cancer with minimal effect of cell cycle, an auxin inducible degron (AID) system was established by integrating the miniAID-mClover3 or miniAID-Halo cassette to the endogenous CTCF locus in a human colorectal cancer cell

line HCT116. The rapid and controlled depletion of CTCF in the CTCF degron system was clearly observed at the protein level as early as 1 hour after treatment with auxin. Transcriptome analysis by RNA-seq demonstrated hundreds of genes differentially expressed in CTCF-depleted cells, where the changes of gene expression were associated with characteristics of cancer such as tumor growth and metastasis. My study demonstrated that CTCF play a critical role in the tumor progression by directly regulating gene expression program in human colorectal cancer.

Key words: CTCF, degron, chromatin structure, transcriptome, colorectal cancer

The study on the role of chromatin architectural protein CTCF in the control of gene expression by establishing acute CTCF protein depletion system in human colorectal cancer

Moo-Koo Kang

Department of Medical Science

The Graduate School, Yonsei University

(Directed by Professor Hyoung-Pyo Kim)

I. INTRODUCTION

Genetic information is contained in approximately 2 meter-long DNA, condensed in 10 micrometer-wide nucleus [1,2]. To accomplish this dynamic fold change, cells take the strategy of hierarchical folding. It is broadly known that chromatin is folded into loop/topologically-associating domain (TADs), compartments, and territories in ascending order [3-10]. These chromatin structures are dynamic and

well-regulated by diverse factors rather than maintaining static state and being formed randomly. Among various factors involved in regulating chromatin structures, CCCTC-binding factor CTCF and ring-shape protein Cohesin are well known chromatin architecture protein [11-14].

CTCF and Cohesin are known to work as complex, inducing chromatin loops through loop extrusion model, in which chromatin extruding through Cohesin is stalled by the pair of convergent-bound CTCF [15]. Through chromatin 3D landscape mediated by CTCF-Cohesin complex, various genetic elements can physically interact [16-20]. In particular, high-resolution and global sequence-based chromatin conformation capture techniques revealed that chromatin interactions involved in gene expression such as promoter-promoter and/or promoter-enhancer interactions are included in CTCF-mediated chromatin domains [21-23]. However, depletion of CTCF/Cohesin has few effects on chromatin compartment [13,14]. Hence, it seems that CTCF/Cohesin regulates loop/TAD-level chromatin conformation. Concentrating on CTCF, how the CTCF-mediated chromatin conformation regulates gene expression is one of the most important question.

Auxin-inducible degron (AID) system has been used for acute target protein depletion [24]. Shortly, F-box protein OsTIR1 existed in *Oryza sativa* is activated by plant hormone auxin and attach ubiquitin to mini AID-tagged target protein, inducing target protein degradation by proteasome (Fig. 1). Applying AID system using plant-only F-box protein to mammalian cells, it is expected that auxin treatment affects only OsTIR1-mediated target protein degradation with

minimum side effects. It is very effective tool to study lethal genes in comparison to genetic knock out or transcript-level knock down because of acute and total depletion and rescue in protein level. CTCF is also known as essential gene, inducing cell death after CTCF knock out in mouse embryonic stem cell [17,25]. There are several studies using AID system for CTCF, accessing to the connection between CTCF and chromatin structure. For example, it is reported that acute CTCF depletion induces decrease of loops and TADs but not compartments [13]. Another study shows that the loss of enhancer-promoter looping affects downregulation of *MYC* expression after acute CTCF depletion in human B-ALL cell line [26]. However, protein level CTCF depletion study in human colon cancer cells has not been done yet. Here, to elucidate the effects of chromatin architecture protein CTCF on transcription, auxin-inducible degron system for CTCF was established in human colon cancer cells. Various confirming steps such as PCR, Western blot, and FACS and phenotypic study such as apoptosis/necrosis was performed to confirm the establishment of auxin inducible degron system for CTCF in human colon cancer cell line. Further, RNA-seq was performed for comparing difference of transcriptome between with and without CTCF in human colon cancer cell.

The importance of CTCF in cancer cells is broadly studied [27-29], but the certain evidences in colorectal cancer cells have not been shown. Here, with inducible protein level complete depleting system, the list of genes regulated by CTCF in human colon cancer cell were extracted by RNA-seq. Some of which are

associated with growth and metastasis of colon cancer but these genes did not exhibit consistent trend, suggesting CTCF may directly regulate gene expression involved in properties of colon cancer in the manner of chromatin conformation change.

II. MATERIALS AND METHODS

1. Cell line

Human colon cancer cell lines including HCT116 CMV-OsTIR1 and HCT116 Tet-On OsTIR1 was provided as thankful gift from prof. Kanemaki. M. Cells were cultured at 37℃ in 5% CO_2 in RPMI (SH30027.01, Hyclone, Logan, Utah, USA) supplemented with 10% FBS (SH30071.03, Hyclone, Logan, Utah, USA) and 100U/ml Penicilin and 100μg/ml Strptomycin (SV30010, Hyclone, Logan, Utah, USA).

2. Chemical usage

For antibiotics selection in acquiring CTCF-AID-Halo clones, final 0.7mg/ml G418 (A1720, Sigma Aldrich, St Louis, Missouri, USA) and 0.1mg/ml Hygromycin (H3274, Sigma Aldrich, St Louis, Missouri, USA) were co-treated for 11-13 days. 500μM 3-indole acetic acid (IAA) was used for auxin-inducible target protein degradation (I5148, Sigma Aldrich, St Louis, Missouri, USA).

3. Plasmid constructs

Repair template for knocking in Dendra2 to N-terminal of RPB1 gene was provided as thankful gift from prof. Won-Ki Cho (Dendra2-RPB1-RT). Repair templates for AID tagging on C-terminal of CTCF gene were constructed on the basis of LITMUS28 (N3628, NEB, Ipswich, Massachusetts, USA). To insert

CTCF homology arm into LITMUS28 vector, KpnI-HF (R3142, NEB, Ipswich, Massachusetts, USA) and SpeI-HF (R3133, NEB, Ipswich, Massachusetts, USA) were used. Both CTCF left and right homology arm were gained from genomic DNA of HCT116 CMV-OsTIR1 by PCR with 250bp length each from stop codon, containing KpnI and SpeI restriction site attached by PCR primer (Table 1).

Site-directed mutagenesis on stop codon region was performed for inserting BamHI site by silent mutation and for preventing sgRNA re-cutting (Table 1). To make complete repair templates, mAID-mClover3-NeoR/Hygro tags were digested by BamHI-HF (R3136, NEB, Ipswich, Massachusetts, USA) from pMK289 (mAID-mClover-NeoR) (72827, Addgene, Watertown, Massachusetts, USA) and pMK290 (mAID-mClover-Hygro) (72828, Addgene, Watertown, Massachusetts, USA), and ligated with digested vector using T4 DNA ligase (M0202, NEB, Ipswich, Massachusetts, USA).

According to methods above, final CTCF Homology arm sequence for gene editing using CRISPR/Cas9 is following:

CTAAAAGACCCTTGTGATTCTTGGGGCTTTAATGGACCATTTGTTCTG
TCTGTGCTCTTCTTTGCCAGCAACAGCTATCATTCAGGTTGAAGACC
AGAATACAGGTGCAATTGAGAACATTATAGTTGAAGTAAAAAAAGA
GCCAGATGCTGAGCCCGCAGAGGGAGAGGAAGAGGAGGCCCAGCC
AGCTGCCACAGATGCCCCCAACGGAGACCTCACGCCCGAGATGATC
CTCtcCATGATGGACCGGgGATccCGGAGCCTTGTGCGTCGCCAGGACT
TCTCTGGGCTGTGTTTAAACGGCCCGCATCTTAATTTTTCTCCCTTCT

TTCTTTTTTTGGCTTTGGGAAAAGCATCATTTTACCAAACATACCGAG
AACGAAAACTTCAAGGATGATGTTAGAAAAAAATGTGATTTAACTA
GAACTTGCTGTCTGATGTTAGCAAATCATGGAATGTTCTGAGTCCCT
GAGGGTTTACTGTGAAGTGCTGAGGACAGTG.

The original sequence of CTCF gene was gained from chromosome 16: 67,596,310-67,673,086 located in GRCh37 CTCF (ENSG00000102974). Underlined tcC shows silent mutation for preventing re-targeting of sgRNA and gGATcc for BamHI site mutation on stop codon. Each small letter represents changed sequence: tcC for AGC and gGATcc for TGATGG.

For exchanging mClover3 to HaloTag, vector containing HaloTag gene was given as gift from prof. Won-Ki Cho, KAIST, South Korea (EF1α-MCP-Halo). HaloTag gene was amplified by PCR using primer containing XbaI and NheI restriction site (Table 1). mClover3 of CTCF-mAID-mClover3-NeoR construct was exchanged with HaloTag using XbaI (R0145, NEB, Ipswich, Massachusetts, USA) and NheI (R0131, NEB, Ipswich, Massachusetts, USA). Then CTCF left arm-mAID-mClover3 of CTCF-mAID-mClover3-Hygro construct was exchanged with CTCF left arm-mAID-HaloTag of CTCF-mAID-HaloTag-NeoR construct by using MluI (R0198, NEB, Ipswich, Massachusetts, USA) and PciI (R0655, NEB, Ipswich, Massachusetts, USA).

Table 1. Oligo primer sequences for repair templates

Target	Orientation	Sequence
CTCF homology arm	Forward	GAAACTAGTCTAAAAGACCCTTGTGATTCTTGG
	Reverse	GAAGGTACCCACTGTCCTCAGCACTTCAC
Site-directed mutagenesis	Forward	GATCCTCAGTATGATGGACCGGGGATCCCGGAGCCTTGTGCGTCGCC
	Reverse	GGCGACGCACAAGGCTCCGGGATCCCCGGTCCATCATACTGAGGATC
Amplifying HaloTag gene	Forward	GAAGCTAGCATGGCAGAAATCGGTA
	Reverse	GAATCTAGAGCCGGAAATCTCGAGC

4. CRISPR/Cas9 gene editing for tagging Pol II and CTCF

For tagging Dendra2 on RPB1, sgRNA for RPB1 containing vector was provided as gift from prof. Won-Ki Cho (pSpCas9(BB)-2A-Puro (PX459) V2.0, 62988, Addgene, Watertown, Massachusetts, USA). Both repair template and sgRNA containing vector are co-transfected to about $1x10^6$ cells using FuGene HD Transfection Reagent (E2300, Promega, Madison, Wisconsin, USA). Knock in cells were sorted by flow cytometry.

For knocking in AID tag on C-terminal of CTCF gene, single guide RNA for targeting CTCF gene is designed using software in benchling (https:/s/www.benchling.com/) (Table 2). Negative control sgRNA followed previous study [30]. Single strand sgRNAs are annealed and pX330 vector are cut with BbsI-HF (R3539, NEB, Ipswich, Massachusetts, USA), following ligation with T4 DNA Ligase. For sgRNA annealing, 100μM of each oligo strand and T4 Ligase (PNK) (M0201, NEB, Ipswich, Massachusetts, USA) are mixed, incubating with 37°C for 30 minutes, 95°C for 5 minutes, and ramping down until 25°C. Repair templates (CTCF-mAID-mClover3-NeoR/Hygro or CTCF-mAID-Halo-NeoR/Hygro) and one sgRNA-containing vector (pX330-U6-Chimeric_BB-CBh-hSpCas9, 42230, Addgene, Watertown, Massachusetts, USA) were co-transfected by using FuGENE HD Transfection Reagent. 500ng of each repair templates and 400ng of sgRNA-containing vector were transfected to about $5x10^4$ cells.

Table 2. Oligo DNA sequences for sgRNA

sgRNA Target	Orientation	Sequence
CTCF	Forward	caccgTCAGCATGATGGACCGGTGA
	Reverse	aaacTCACCGGTCCATCATGCTGAc
Negative control	Forward	caccgCTGATCTATCGCGGTCGTC
	Reverse	aaacGACGACCGCGATAGATCAGc

5. Cloning

Cloning was performed for amplifying plasmids at each step of retaining plasmid constructs. As competent cell, DH5α and Stbl3 were used. All plasmids constructs were put in competent cells by transformation. For transformation, proper amount of plasmid was applied to competent cells and incubated in ice for 20 minutes. After that, competent cells were incubated at 42℃ for 45 seconds for heat shock, putting DNAs into the competent cells, following 2 minutes ice incubation. Next, LB media is added and incubated in 37℃ for 45 minutes which is for stabilization. After cell down, the transformed competent cells were spread on the solid LB plate containing ampicillin. Colonies were confirmed with colony PCR and extracted plasmids were confirmed by enzyme digestion and sequencing. For digestion confirm, several sets of enzymes against target plasmids were selected and compared with virtual digestion.

6. Establishment of auxin inducible degron cell line

Previous study for establishment of auxin inducible degron cell line were followed in this study [24]. 24 hours after transfection of repair templates and sgRNA-containing vector, cells were transferred to 10cm plate with 1/100 to 1/1,000 dilution. 2 days after transfer, antibiotics were treated for 11-13 days, refreshing every 3 days. When the colonies appear, pick them using pipette tip and transfer to 96-well plate for expansion. Cells were splitted at first expansion step for PCR genotyping. For Genotyping, cells were lysed with Direct PCR

(F170, Thermo Scientific, Waltham, Massachusetts, USA) including proteinase K. Genotype-confirmed clones were expanded until 10cm plate and confirmed functionally by Western blotting.

7. Polymerase chain reaction (PCR) for repair templates and genotyping

A. PCR for homology arm and site-directed mutagenesis

Homology arms used in repair templates and mutation applied plasmids by site-directed mutagenesis were retained by PCR. For PCR, 10X MG Taq buffer (MP00102, MG Med, Seoul, South Korea), 2mM dNTP (MP00102, MG Med, Seoul, South Korea), each primer (Table 1), DMSO (D2650, Sigma Aldrich, St Louis, Missouri, USA), MG Taq polymerase (MP00102, MG Med, Seoul, South Korea), MG DNA Pfu Polymerase (MP00412, MG Med, Seoul, South Korea) and DNA from cells were used. The conditions for amplifying DNAs are as follows; after 6 minutes 95℃ for pre-denaturing DNAs and 2 minutes in ice, 95℃ additional denaturation, 30 amplification cycles (95℃ 30 seconds denaturation, 30 seconds annealing (following Tm value of primers), and 72℃ 1 minute/kb extension), and 72℃ final extension step are followed. For homology arm, amplified DNA was loaded on proper concentration of agarose gel after mixed with 6X DNA loading dye (6X LoadingSTAR, A750, Dynebio, Seongnam-si, Gyeonggi-do, South Korea) for electrophoresis. For cloning step,

prep the size of interest from gel using Expin Gel SV (102-102, GeneAll, Seoul, South Korea).

B. PCR for genotyping

Every clone acquired after transfection were tested their genotype by PCR. First, Dendra2-RPB1 clones were tested with primer set targeting outside of both left and right homology arms (Table 3). Second, CTCF-AID-Halo clones were tested with two sets of primers, targeting inside of NeoR or Hygro to outside of right homology arm (Table 3). The conditions for genotyping are as follows; after 1minute 30 seconds 95℃ for pre-denaturing DNAs, 30 amplification cycles of 95℃ 30 seconds denaturation, 63℃ 30 seconds annealing, and 72℃ 1 minute/kb extension are followed. Last, after 72℃ 2 minutes final extension step, amplified DNA was loaded on proper concentration of agarose gel after mixed with 6X DNA loading dye (6X LoadingSTAR, A750, Dynebio, Seongnam-si, Gyeonggi-do, South Korea).

Table 3. Oligo primer sequences for PCR genotyping

Target	Orientation	Sequence
RPB1	Forward	CCGTAACCTCTGCCGTTCAG
	Reverse	CCTGGCCCGTCACCCATAAG
NeoR-CTCF	Forward	AATGGGCTGACCGCTTCCTC
	Reverse	CTCCAAGCAAAGGGAGTCAG
Hygro-CTCF	Forward	ACTGTCGGGCGTACACAAAT
	Reverse	CTCCAAGCAAAGGGAGTCAG

8. Western blot

For Western Blotting, cell pellets up to 2 million cells were lysed in 200 μl T-PER Tissue Protein Extraction Reagent (78510, Thermo Scientific, Waltham, Massachusetts, USA), following incubation for 30 minutes at 4˚C and centrifuged for 10 minutes at 14,000 rpm at 4˚C. Supernatant was transferred to new 1.5ml tube and mixed with 4X Laemmli Sample buffer (1610747, BIO-RAD, Hercules, California, USA) containing β-mercaptoethanol. Samples were boiled for 10 minutes at 100˚C before use. Protein was loaded on proper concentration of poly-acrylamide gel and ran in homemade running buffer at 100V using mini Trans-Blot Cell (M1793030, BIO-RAD, Hercules, California, USA). Proteins were transferred on PVDF membranes (IPVH00010, Millipore, Burlington, Massachusetts, USA) for 2 hours at 300mA using mini Trans-Blot Cell in transfer buffer containing 20% methanol. Membranes were blocked with 5% skim milk in PBS-T for 1 hour at room temperature and incubated overnight at 4˚C with primary antibodies diluted in PBS-T 2% skim milk (1:1,000 anti-CTCF, 2899S, CST, Danvers, Massachusetts, USA; 1:1,000 anti-a-tubulin, sc-32293, Dallas, Texas, USA). After washing with PBS-T 4 times for 5 minutes each, membrane was incubated with secondary antibodies (1:5,000 anti-mouse IgG, HRP-linked Antibody, #7076, CST, Danvers, Massachusetts, USA; 1:2,000 anti-rabbit IgG, HRP-linked Antibody, #7076, CST, Danvers, Massachusetts, USA) for 1 hour at room temperature. Washing membrane again same as above, substrates for detecting different target proteins were applied (ECL, 32106, Thermo Scientific,

Waltham, Massachusetts, USA; ECL plus, 32134, Thermo Scientific, Waltham, Massachusetts, USA). For detection, LAS4000 was used.

9. Flow cytometry

A. Fluorescence activating cell sorting (FACS) for sorting Dendra2-RPB1 knock in cells

For sorting Dendra2-RPB1 knock in cells, BD LSR Aria II was used for FACS. The process was two steps; first, sort all GFP positive cells and culture for few days. Second, sort single GFP positive cell on 96-well plate for acquire single clone.

B. Analyzing GFP positive CTCF-mAID-mClover3 cells

For detecting GFP positive cells, BD LSR FORTESSA was used. For samples involved in each condition, cells were detached from culture plate, washed and resuspended in FACS buffer (homemade). Cells expressing mClover3 were detected by 488nm laser.

C. Apoptosis/Necrosis analysis

For Apoptosis/Necrosis analysis, Annexin V Apoptosis Detection Kit APC (88-8007-72, eBioscience, San Diego, California, USA) was used, following manufacturer's guide. Briefly, cells in each condition were detached from

culture plate, washed and resuspended in 1X Binding buffer. Annexin V-APC staining was applied on 1-5x10^5 cells for 20 minutes and washed with FACS buffer (homemade). Propidium Iodide staining was applied on each sample 15 minutes before detection. For detection, BD LSR FORTESSA was used.

10. Library preparation for RNA-sequencing

Total RNA was extracted with TRIzol LS reagent (10296028, Thermo Scientific, Waltham, Massachusetts, USA) from each condition followed by RNA precipitation in isopropanol, resuspended in RNase-free water. Library for RNA-sequencing was prepared using the NEBNext Ultra II RNA Library Kit for Illumina (E7760, NEB, Ipswich, Massachusetts, USA) and NEBNext Poly(A) mRNA isolation module (E7490, NEB, Ipswich, Massachusetts, USA) and NEBNext Multiplexed Oligos (E6440, NEB, Ipswich, Massachusetts, USA) as described in the manufacturers' protocol, with 1,000ng of total RNA as input. Quakity check for total RNA was performed by gel electrophoresis.

11. RNA-seq data analysis

Paired-end libraries (101bp) were sequenced on the Illumina platform. Raw reads were trimmed using trim_galore (v0.6.4). Reference genome data was downloaded from GENCODE v19. Trimmed reads were aligned to the hg19 reference genome using STAR (v2.6.0). Annotations of genes were obtained from GENCODE V19, and genes described as 'level 3' were filtered out. Gene

quantification was carried out using RSEM (v1.3.3).

Differential Expression Analysis was performed based on read counts using DESeq2 R package (v1.26.0). Differentially Expressed Genes were defined by $\log_2$-fold-chang(lfc) > 1, padj < 0.05.

III. RESULTS

1. Establishment of Dendra2-RPB1 knock in cell line on HCT116 expressing OsTIR1

Dendra2 is one of photoactivable fluorescence protein, which expresses green fluorescence and is converted to express red fluorescence after activated with 405nm light [31]. This characteristic is key factor for super resolution imaging [32-34]. For various applications, Dendra2-tagged Pol II cell in colon cancer cell line HCT116 expressing OsTIR1 was established. For tagging Dendra2 to N-terminal of RPB1, the largest subunit of Pol II, CRISPR/Cas9 system was used (Fig. 2A). Because RPB1 has long amino acid repeats at C-terminal, which has important function in transcription, N-terminal was targeted for tagging. After transfection with sgRNA and repair template including homology arms, GFP positive cells were sorted by FACS (Fluorescence activated cell sorting) in two steps: first, sort all GFP positive cells and incubate cells for few days, and then sort single GFP positive cells for acquiring single clones (Fig. 2B). This process is required because repair template for N-terminal tagging has start codon for transcription, inducing transient expression itself. After the first sorting, culturing cells for a few passages dilutes transiently-expressing cells and the population of knock in-stable cells mainly remains. Through genotyping by PCR, one homozygous clone was gained (Fig. 2C). Next, for discriminating Dendra2-tagged Pol II protein size, WT (250kDa) and Dendra2 fused RPB1 (276kDa)

was confirmed by Western blotting (Fig. 2D). With these processes, the establishment of Dendra2-RPB1 knock in cell line was confirmed in molecular level.

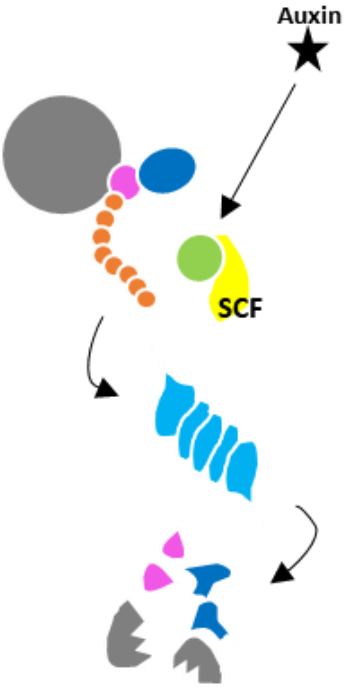

Figure 1. Scheme of auxin inducible degron system. OsTIR1 (Green), E3 ligase existing in *Oryza sativa*, is activated by auxin, and attach ubiquitins (scarlet) to miniAID tag (magenta) fused with target protein (grey) and another tag (deep blue). Ubiquitin-tagged target protein is degraded by proteasome (light blue) rapidly.

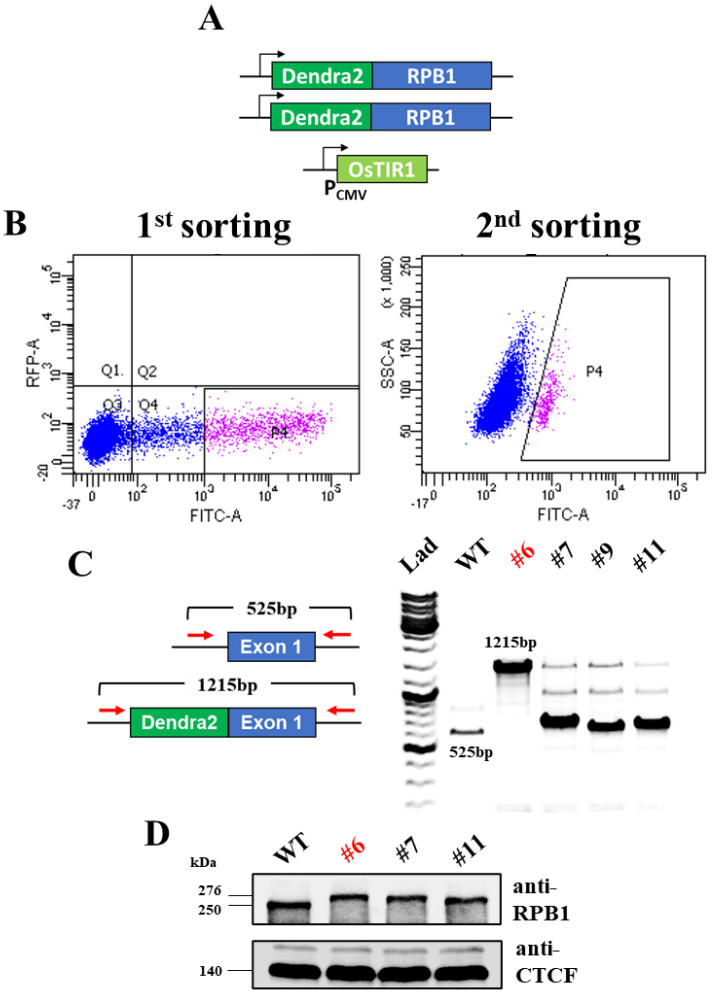

Figure 2. Establishment of Dendra2-RPB1 knock in cell line on HCT116 expressing OsTIR1. (A) Genetic scheme of Dendra2-RPB1. (B)Fluorescence activating cell sorting (FACS) analysis for sorting GFP positive cells representing Dendra2 expression. All GFP positive cells were sorted at first, and only single GFP positive cells were sorted at second. (C) Genotyping of Dendra2-RPB1 clones by PCR. A primer set targeting the first exon discriminate both wild type

and Dendra2-tagging RPB1. Wild type RPB1 has 525bp band, and Dendra2-RPB1 clones has 1215bp band. (D) Confirming protein size of Dendra2-RPB1 clones by Western blotting. Wild type RPB1 is approximately 250kDa, and Dendra2-RPB1 is approximately 276kDa. 140kDa CTCF protein was used for control, showing that equal amount of protein was loaded for all clones.

2. Establishing auxin inducible degron (AID) system for CTCF in human colon cancer cell line

To study one of essential gene CTCF, conventional knock out or knock down systems have limits on lethality, inducibility, and reversibility. To overcome these limits, auxin inducible degron (AID) system for CTCF was established in human colon cancer cell line HCT116 expressing OsTIR1 (Fig. 3A). Overall process followed previous methods [24]. Briefly, for establishing AID cell, two major genes are needed; F-box protein and miniAID (mAID) tag. OsTIR1, one of plant-derived F-box protein only exists in *Oryza sativa*, should be knocked-in in AAVS1 locus, known as safe harbor. mAID tag is needed to be knocked-in to N- or C-terminal of target protein. Various fluorescence proteins can be fused with mAID tag, facilitating detection of target protein. mAID tag was fused with mClover3, which is one of GFP variants, to C-terminal of CTCF gene, because it is reported that N-terminal of CTCF is important to interaction between CTCF and Cohesin [35]. Two repair templates having different antibiotics resistance gene were used for CRISPR-Cas9 mediated knock in, identifying editing both alleles. After transfection of repair templates and sgRNA and antibiotics selection, genotyping was done by PCR during expanding each clone 96-well plate to 10cm plate (Fig. 3B). Approximately 150 colonies were identified after selection and 48 colonies of them had double positive in genotype. Among them, two clones were used for next steps. To confirm the function of AID system, Western blotting was done (Fig. 3C). Wild type CTCF is

approximately 140kDa and mAID-mClover3 tagging adds 30kDa. After 500μM auxin treatment at each time point, gradual degradation of protein was identified. Based on process for establishing AID system instructed above, another CTCF-AID cell, expressing CTCF fused with HaloTag instead of mClover3, was established. HaloTag has no fluorescence itself, but can bind with various HaloTag ligands not only fluorescence dye but also biotin, antibodies, optimal beads for HaloTag, etc. Because of its broad application, HaloTag is used in various research. HCT116 with Dendra2-RPB1 expressing OsTIR1 was used as parental cell for CTCF-AID-HaloTag knock in (Fig. 4A). As same as above procedure, genotype of each clone was confirmed by PCR (Fig. 4B). 42 clones were acquired after selection, and 19 clones were identified as knock in for both alleles. After choosing two clones for next steps, function of AID system was confirmed by Western blotting (Fig. 4C). mAID-HaloTag adds 41kDa than wild type proteins. Larger size of tagged CTCF than WT CTCF was depleted only after 6 hours-auxin treatment. Strikingly, CTCF protein level was rescued 24 hours after wash off.

All AID cells established above have CMV promoter for expressing OsTIR1. However, it seems that basal level of CTCF protein is different between parental cell and AID cells (Fig. 4C). It is inferred that continuous expression of OsTIR1 may induce unexpected target protein degradation because of random activation of OsTIR1 without auxin. This phenomenon is also reported previous studies as basal level degradation [36,37]. To overcome this problem, another AID cell

expressing CTCF-AID fused with HaloTag was established with Tet-On system for expressing OsTIR1 (Fig. 5A). HCT116 Tet-On OsTIR1 cell was used as parental cell. PCR genotyping was performed after transfection of repair templates and sgRNA and selection (Fig. 5B). Among 51 clones, 34 clones had double positive genotype. Two clones were chosen and functionally confirmed by Western blotting (Fig. 5C). Notably, almost equal intensity of CTCF band in wild type samples and untreated CTCF-AID sample was identified. Also, total depletion of CTCF protein was observed after sequential treatment of doxycycline and auxin 24 hours each. After auxin wash off for rescue, total recovery of CTCF protein was confirmd in clone 4. It is noteworthy in comparison to previous rescue data (Fig. 4C). With these results, establishment of various AID system for CTCF was confirmed in human colon cancer cell line HCT116.

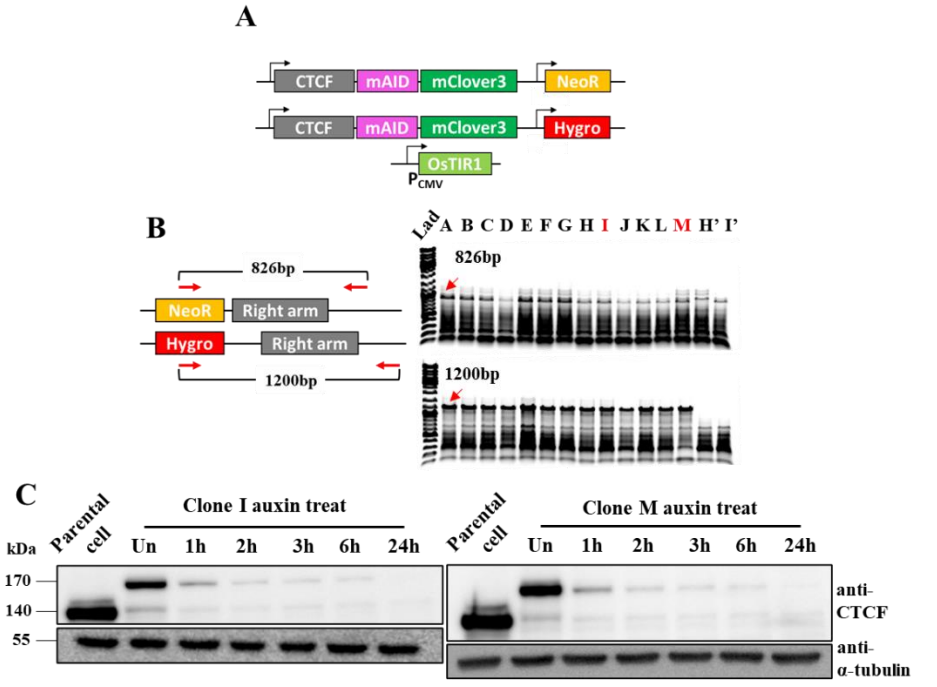

Figure 3. Establishment of auxin inducible degron system for mClover3 tagged-CTCF in OsTIR1 expressing HCT116 cell. (A) Genetic scheme of CTCF-mAID-mClover3. (B) Confirming genotype of CTCF-mAID-mClover3 clones by PCR. Two primer sets were used for detecting antibiotics resistance gene to out of right homology arm of CTCF. NeoR-out of right arm shows 826bp and Hygro-out of right arm shows 1200bp. Clone I and M were used in study. (C) Functional confirmation of CTCF-mAID-mClover3 clones by Western blotting. Wild type CTCF is approximately 140kDa, and CTCF-mAID-mClover3 is approximately 170kDa. 55kDa α-tubulin protein was used for control, showing that equal amount of protein was loaded for all clones. 500μM of auxin was used for each timepoint.

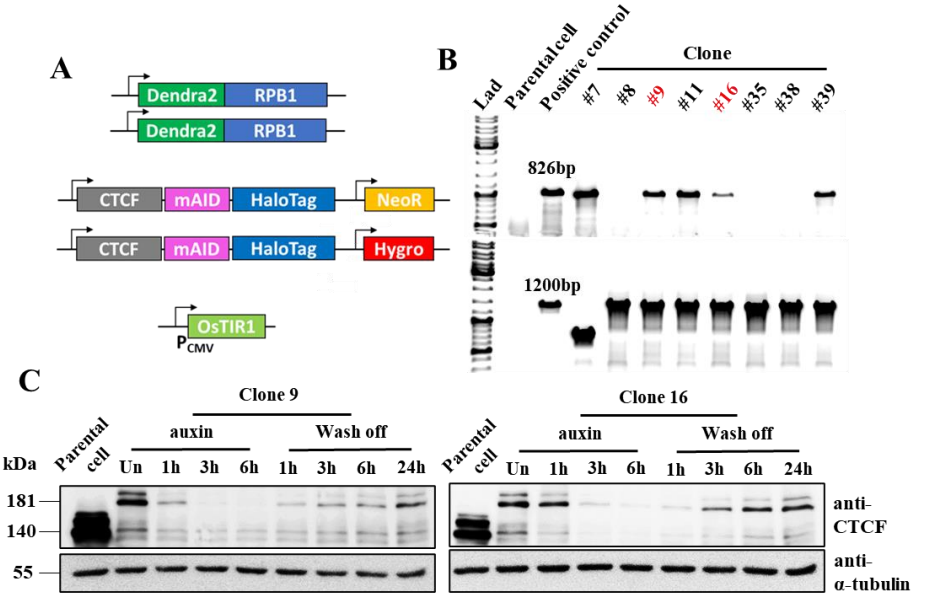

Figure 4. Establishment of auxin inducible degron system for HaloTag tagged-CTCF in OsTIR1 expressing HCT116 cell. (A) Genetic scheme of HCT116 Dendra2-RPB1 CTCF-mAID-HaloTag. (B) Confirming genotype of clones by PCR. Two primer sets were used for detecting antibiotics resistance gene to out of right homology arm of CTCF. NeoR-Right arm shows 826bp and Hygro-Right arm shows 1,200bp. (C) Confirming function of the system by Western blotting. Wild type CTCF is approximately 140kDa, and CTCF-mAID-HaloTag is approximately 181kDa. 55kDa α-tubulin protein was used for control, showing that equal amount of protein was loaded for all clones. 500μM of auxin was used for each timepoint. Wash off was performed after 6hrs auxin treatment.

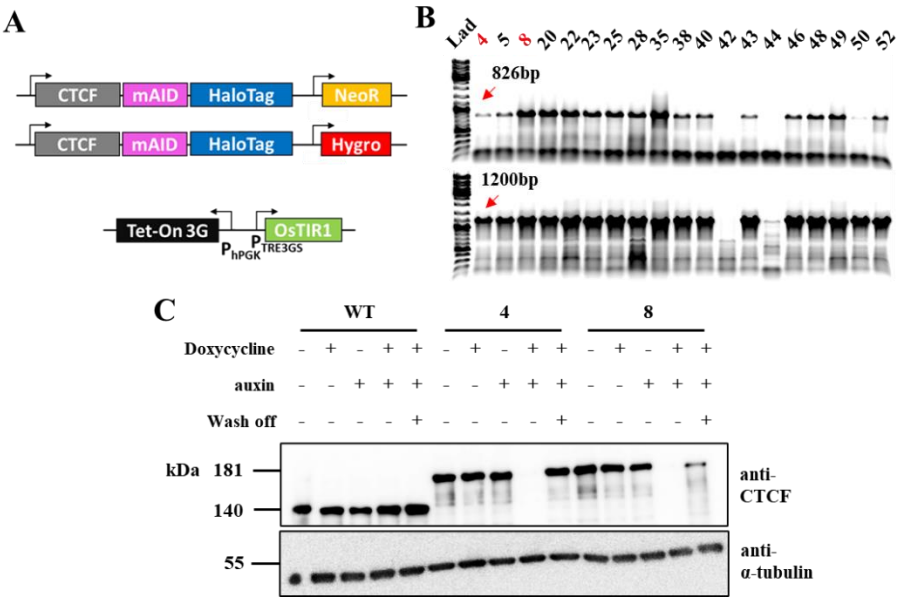

Figure 5. Establishment of auxin inducible degron system for HaloTag tagged-CTCF with tetracycline inducible OsTIR1 in HCT116 cell. (A) Genetic scheme of HCT116 Tet-On CTCF-mAID-HaloTag. (B) Confirming genotype of clones by PCR. Two primer sets were used for detecting antibiotics resistance gene to out of right homology arm of CTCF. NeoR-Right arm shows 826bp and Hygro-Right arm shows 1200bp. (C) Functional confirmation of clones after treatment of 1μg/ml doxycycline and 500μM auxin, and wash off by Western blotting. All treatment was applied for 24hrs in order of doxycycline, auxin, and wash off. Wild type CTCF is approximately 140kDa, and CTCF-mAID-HaloTag is approximately 181kDa. 55kDa α-tubulin protein was used for control, showing that equal amount of protein was loaded for all clones.

3. Apoptosis/Necrosis analysis after CTCF depletion

CTCF is well known lethal gene in mammalian cell [17,25,38]. Knock out study for lethal genes is difficult because it is hard to specify the quality of cell state after inducing knock out. Finding the point between lethal gene knock out and cell death is key factor for knock out study for essential genes. To confirm effects of acute whole CTCF depletion on HCT116, apoptosis/necrosis analysis with or without CTCF was performed in HCT116 CTCF-mAID-mClover3 clones. First, FITC signal detecting mClover3 was confirmed in each clone and condition (Fig. 6A). Both clones showed FITC signal before auxin treatment, and it was decreased to basal level of parental cell after auxin treatment. Next, apoptosis/necrosis was detected using Annexin V/PI staining (Fig. 6B). Annexin V stains phosphatidyl serine, which is located inside of cell membrane in healthy cells but move to outside of membrane during apoptosis. Propidium iodide (PI) stains nucleic acid, which is erupted after necrosis. Almost same population of Annexin V and PI single or double positive cells before and after depletion of CTCF in both clones. These results show that acute depletion of CTCF has little effect on apoptosis/necrosis in HCT116.

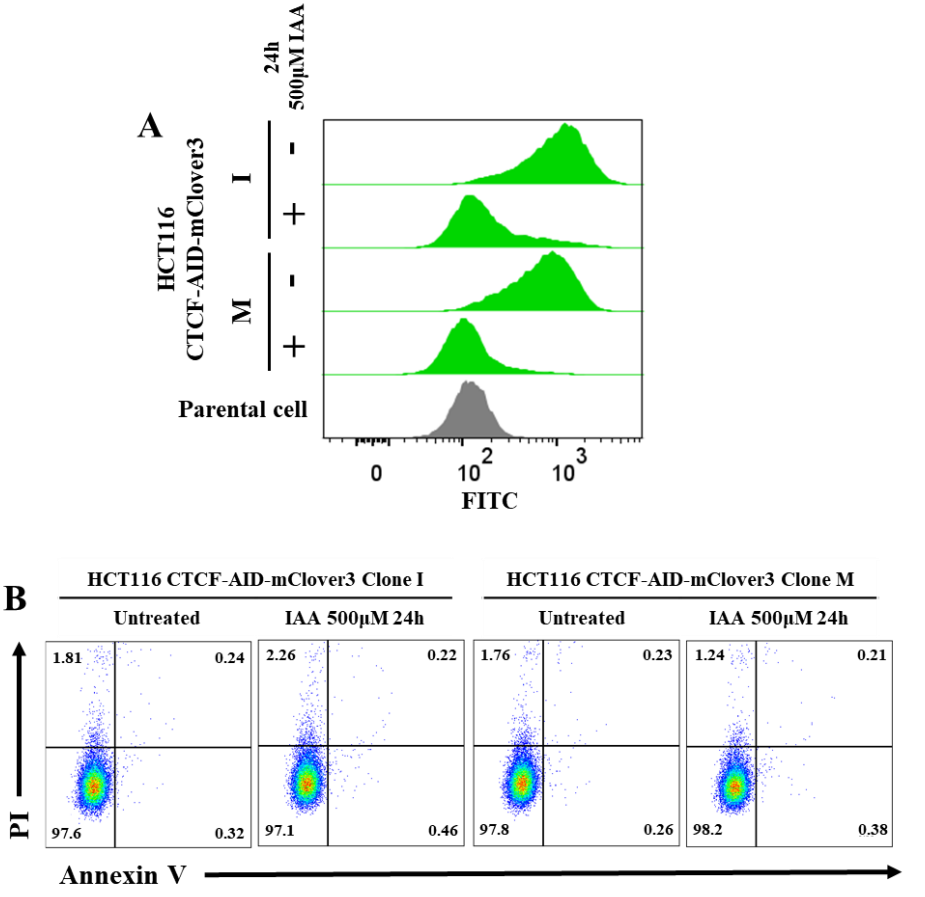

Figure 6. Apoptosis/Necrosis analysis after CTCF depletion in HCT116. (A) FACS analysis detecting FITC for comparing with and without CTCF using AID system in HCT116. Parental cell shows basal fluorescence without any fluorescence tag. (B) Apoptosis/Necrosis analysis was performed using Annexin V/PI staining. Annexin V represents apoptosis and PI represents necrosis.

4. Mining genes regulated by CTCF through analyzing differentially expressed genes (DEGs)

Next, RNA-sequencing was performed for mining differentially expressed genes (DEGs) after CTCF depletion. Overall, 264 genes were upregulated and 137 genes were downregulated after CTCF depletion (Fig. 7A). To identify terms related to up/downregulated genes, Enrichr, the website integrating public datasets, was used.

First of all, gene set representing HCT116 was acquired from Term search in Enrichr. This gene set was integrated with DEGs derived from RNA-seq, shown as volcano plot (Fig. 7B). Significantly up or downregulated genes were used as input for searching related terms. Two cutoff standards were applied: more than absolute $\log_2$ fold change 0.5 or 1. Each left panel of Fig. 7C and 7D represents 0.5 as cutoff and 1 as right panel. The upregulated results showed "Integrin-mediated cell surface interactions" at cut off 0.5 and "Epithelial-to-mesenchymal transition in colon cancer cells" at cut off 1 for upregulated gene set (Fig. 7C). Otherwise the downregulated results showed "Inhibition of metastatic potential" at cut off -0.5 and "Calcium-dependent cell-cell adhesion" at cut off -1. (Fig. 7D). Although it seems that both up and downregulated genes significantly reflect their terms (Table 4-7), it is hard to say the terms are reliable because the number of genes included in each term was too small to significantly support their terms. For example, the term "Integrin cell surface interactions" for upregulated gene set included 5 genes and "Calcium-dependent cell-cell

adhesion via plasma membrane cell adhesion molecules" for downregulated gene set only included 2 genes. Moreover, upregulated genes, such as CDH1, COL4A3, and LAMA2 represented promoting cell adhesion, while down regulated genes, such as SELL, ID1, and FOS represented promoting metastasis. With these random and conflicting results, it is implied that CTCF regulates partial and individual genes essential for characteristics of cells rather than global gene expression or specific biological patterns with coherent context.

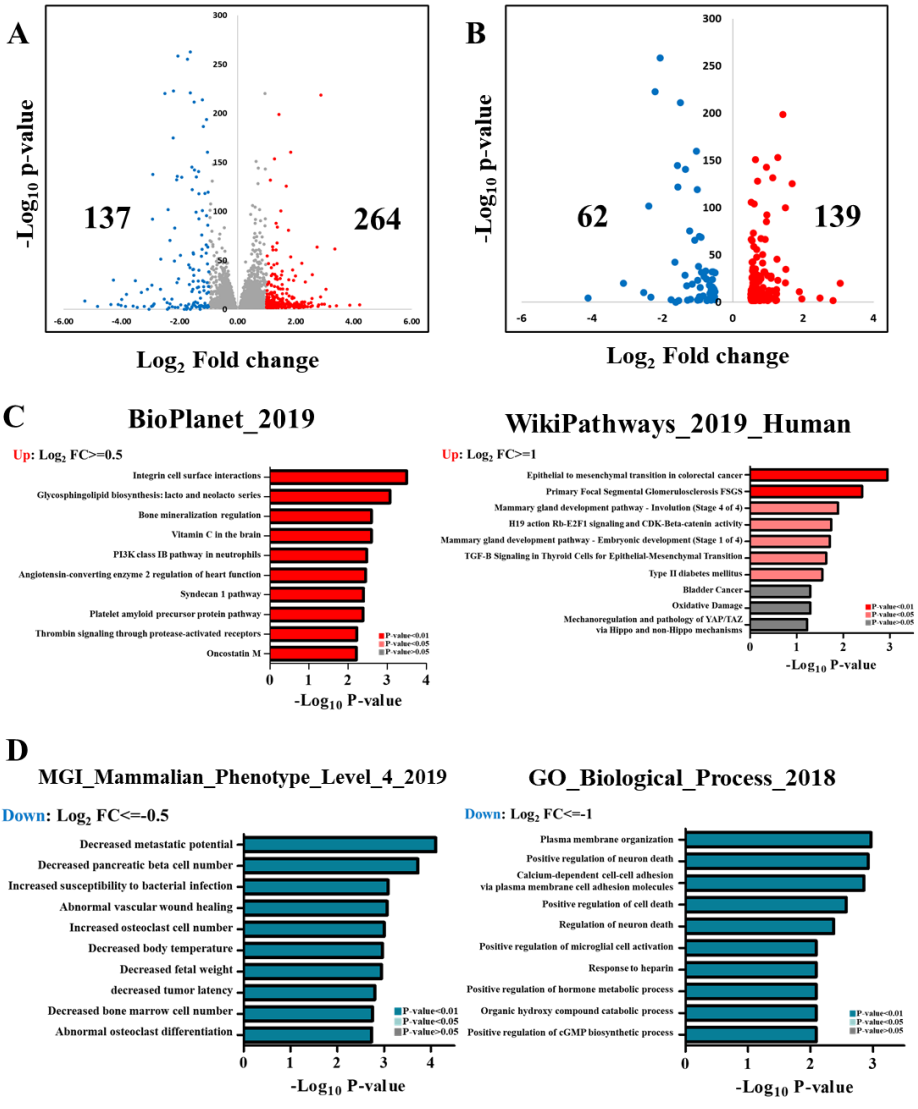

Figure 7. Differentially expressed genes (DEGs) analysis after CTCF depletion in HCT116. (A) Volcano plot using DEG set of RNA-seq. Merged data from HCT116 CTCF-mAID-mClover clone I and M are used. Log_2 fold change for x axis and -log10 p-value for y axis. Each value is calculated with DESeq2. For upregulated gene set, genes were filtered with adjusted p-value<0.05 and log_2 fold

change>=0.5 (left) or 1 (right). For downregulated gene set, genes were filtered with adjusted p-value<0.05 and $\log_2$ fold change=<-0.5 (left) or -1 (right). (B) Volcano plot for up and downregulated genes for HCT-1116 related genes, acquired from Term search in Enrichr. (C) Upregulated genes from (B) were used as input for BioPlanet 2019 and WikiPathways 2019 Human, showing biological pathways for input genes. (D) Downregulated genes from (B) were used as input for MGI Mammalian Phenotype Level 4 2019 and GO Biological Process 2018, showing related ontology for input genes. All data from Enrichr website are based on gene sets of DEGs after CTCF depletion. All lists ranked by -log10 p-value. Vivid red means p-value<0.01; pink means p-value<0.05; grey means p-value>0.05.

Table 4. Genes included in gene ontology terms from upregulated gene set after CTCF depletion (Log_2 Fold change>=0.5)

Term	P-value	Genes
Integrin cell surface interactions	$3.17x10^{-4}$	LAMA2;CDH1;COL4A4;COL4A3;ITGAX
Glycosphingolipid biosynthesis: lacto and neolacto series	$8.50x10^{-4}$	B3GNT3;ST3GAL6;FUT2
Bone mineralization regulation	$2.53x10^{-3}$	COL4A4;COL4A3
Vitamin C in the brain	$2.53x10^{-3}$	COL4A4;COL4A3
PI3K class IB pathway in neutrophils	$3.31x10^{-3}$	CYTH4;C5AR1;RAC2
Angiotensin-converting enzyme 2 regulation of heart function	$3.56x10^{-3}$	COL4A4;COL4A3
Syndecan 1 pathway	$4.01x10^{-3}$	COL4A4;COL12A1;COL4A3
Platelet amyloid precursor protein pathway	$4.13x10^{-3}$	COL4A4;COL4A3

Table 5. Genes included in gene ontology terms from upregulated gene set after CTCF depletion (Log_2 Fold change>=1)

Term	P-value	Genes
Epithelial to mesenchymal transition in colorectal cancer	$1.12x10^{-3}$	CDH1;COL4A4;COL4A3
Primary Focal Segmental Glomerulosclerosis FSGS	$3.93x10^{-3}$	COL4A4;COL4A3
Mammary gland development pathway - Involution (Stage 4 of 4)	$1.29x10^{-2}$	CDH1
H19 action Rb-E2F1 signaling and CDK-Beta-catenin activity	$1.80x10^{-2}$	CDH1
Mammary gland development pathway - Embryonic development (Stage 1 of 4)	$1.93x10^{-2}$	CDH1
TGF-B Signaling in Thyroid Cells for Epithelial-Mesenchymal Transition	$2.32x10^{-2}$	CDH1
Type II diabetes mellitus	$2.82x10^{-2}$	CACNA1A

Table 6. Genes included in gene ontology terms from downregulated gene set after CTCF depletion (Log_2 Fold change<=-0.5)

Term	P-value	Genes
Decreased metastatic potential	$7.87x10^{-5}$	SELL;ID1;FOS
Decreased pancreatic beta cell number	$1.88x10^{-4}$	PRF1;PDX1;SSTR5
Increased susceptibility to bacterial infection	$8.21x10^{-4}$	NOS3;FES;PRF1;TRPV3;FOS
Abnormal vascular wound healing	$8.56x10^{-4}$	EGR1;IRS1;NOS3
Increased osteoclast cell number	$9.87x10^{-4}$	EGR1;IRS1;ID1

Table 7. Genes included in gene ontology terms from downregulated gene set after CTCF depletion (Log_2 Fold change<=-1)

Term	P-value	Genes
Plasma membrane organization	$1.07x10^{-3}$	EHD2;DMKN
Positive regulation of neuron death	$1.20x10^{-3}$	EGR1;DKK1
Calcium-dependent cell-cell adhesion via plasma membrane cell adhesion molecules	$1.39x10^{-3}$	SELL;CDH22
Positive regulation of cell death	$2.68x10^{-3}$	EGR1;DKK1
Regulation of neuron death	$4.23x10^{-3}$	EGR1;DKK1

IV. DISCUSSION

Chromatin architecture protein CTCF regulates chromatin interactions, such as loops and TADs. These chromatin conformations seem to be closely connected with transcription. However, how the CTCF-mediated chromatin landscape regulate transcription is not well known. To identify the effect of CTCF on gene expression, various AID system was applied to CTCF; (1) CTCF-AID with green fluorescence protein mClover3, (2) Dendra2-RPB1 CTCF-AID-HaloTag, and (3) CTCF-AID-HaloTag with Tet-On OsTIR1. With this system, acute depletion of CTCF, recovery of protein, and few effects on cell viability after acute CTCF depletion were confirmed.

Random activity of OsTIR1 was observed through basal level target protein degradation, but it seems improved with Tet-On induced OsTIR1 expression system in this study. Total recovery in Tet-On CTCF-AID clone 4 may reflect lower basal expression and activation of OsTIR1 in Tet-On system in comparison to CMV-OsTIR1 based AID cells (Fig. 5C).

CTCF is well known lethal gene in mammalian cells. In previous study, it was confirmed that long term depletion of CTCF induced cell death in a majority of cells [13]. However, most of cells were alive in 24h CTCF depletion (Fig. 6B). It is inferred that acute CTCF depletion is not enough to induce cell death in differentiated cells. It means acute CTCF depletion system can be a great model for studying function of CTCF, especially regulating chromatin structure.

RNA-seq was performed to compare DEGs with and without CTCF. Among approximately 30,000 genes in human, only about 400 genes showed different expression pattern (Fig. 7A). Further, some of them seem to important in regulating progressiveness of colorectal cancer cell, mainly metastasis and epithelial-to-mesenchymal transition (Fig. 7C, 7D and Table 4-7). However, the integration of whole DEGs did not show consistent and significant biological orientation. CTCF is well known as chromatin architecture protein, regulating global interactions between genetic elements, such as “enhancer”, which regulates cell type-specific genes. With this context, it is implied that DEGs in CTCF depleted cells are directly related to chromatin interactions mediated by CTCF rather than biological meanings such as cancer or colon. Further, it can be a good example for suggesting the relationship between chromatin structure and gene expression.

V. CONCLUSION

Chromatin conformation and transcription have close relationship. However, direct relationship between them still need to be revealed. In this study, auxin inducible degron system was used for conditional knock down of CTCF protein, which is known as chromatin architecture protein. With this system, it was validated that CTCF depletion do not induce dramatic gene expression dynamics. However, several genes involved in properties of colon cancer were up or downregulated after CTCF depletion without coherent tendency. This study suggests that CTCF may directly regulate gene expression in human colon cancer cell in the manner of regulating chromatin conformation, as expected .

Additionally, emerging evidences show the relationships between chromatin 3D structure and transcription with super resolution imaging [34]. As expecting, cells stably expressing Dendra2-RPB1 and CTCF-AID-Halo established in this study will be very useful for studying spatiotemporal dynamics of nucleome containing both chromatin and transcription.

REFERENCES

1. Piovesan A, Pelleri MC, Antonaros F, Strippoli P, Caracausi M, Vitale L. On the length, weight and GC content of the human genome. BMC Res Notes 2019;12:106.
2. Molecular biology of the cell (4th ed.): Garland Publishing; 2002.
3. Eagen KP. Principles of Chromosome Architecture Revealed by Hi-C. Trends Biochem Sci 2018;43:469-78.
4. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, et al. Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science 2009;326:289-93.
5. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 2012;485:376-80.
6. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature 2012;485:381-5.
7. Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 2014;159:1665-80.
8. Bonev B, Mendelson Cohen N, Szabo Q, Fritsch L, Papadopoulos GL,

Lubling Y, et al. Multiscale 3D Genome Rewiring during Mouse Neural Development. Cell 2017;171:557-72 e24.

9. Zhan Y, Mariani L, Barozzi I, Schulz EG, Bluthgen N, Stadler M, et al. Reciprocal insulation analysis of Hi-C data shows that TADs represent a functionally but not structurally privileged scale in the hierarchical folding of chromosomes. Genome Res 2017;27:479-90.

10. Szabo Q, Bantignies F, Cavalli G. Principles of genome folding into topologically associating domains. Science Advances 2019;5.

11. de Wit E, Vos ES, Holwerda SJ, Valdes-Quezada C, Verstegen MJ, Teunissen H, et al. CTCF Binding Polarity Determines Chromatin Looping. Mol Cell 2015;60:676-84.

12. Holwerda SJ, de Laat W. CTCF: the protein, the binding partners, the binding sites and their chromatin loops. Philos Trans R Soc Lond B Biol Sci 2013;368:20120369.

13. Nora EP, Goloborodko A, Valton AL, Gibcus JH, Uebersohn A, Abdennur N, et al. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. Cell 2017;169:930-44 e22.

14. Rao SSP, Huang SC, Glenn St Hilaire B, Engreitz JM, Perez EM, Kieffer-Kwon KR, et al. Cohesin Loss Eliminates All Loop Domains. Cell 2017;171:305-20 e24.

15. Sanborn AL, Rao SS, Huang SC, Durand NC, Huntley MH, Jewett AI, et

al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. Proc Natl Acad Sci U S A 2015;112:E6456-65.

16. Rollins RA, Morcillo P, Dorsett D. Nipped-B, a Drosophila homologue of chromosomal adherins, participates in activation by remote enhancers in the cut and Ultrabithorax genes. Genetics 1999;152:577-93.

17. Splinter E, Heath H, Kooren J, Palstra RJ, Klous P, Grosveld F, et al. CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes Dev 2006;20:2349-54.

18. Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, et al. Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature 2008;451:796-801.

19. Le Dily F, Vidal E, Cuartero Y, Quilez J, Nacht AS, Vicent GP, et al. Hormone-control regions mediate steroid receptor-dependent genome organization. Genome Res 2019;29:29-39.

20. Comoglio F, Park HJ, Schoenfelder S, Barozzi I, Bode D, Fraser P, et al. Thrombopoietin signaling to chromatin elicits rapid and pervasive epigenome remodeling within poised chromatin architectures. Genome Res 2018.

21. Chang LH, Noordermeer D. Of Dots and Stripes: The Morse Code of Micro-C Reveals the Ultrastructure of Transcriptional and Architectural Mammalian 3D Genome Organization. Mol Cell 2020;78:376-8.

22. Krietenstein N, Abraham S, Venev SV, Abdennur N, Gibcus J, Hsieh TS, et al. Ultrastructural Details of Mammalian Chromosome Architecture. Mol Cell 2020;78:554-65 e7.

23. Hsieh TS, Cattoglio C, Slobodyanyuk E, Hansen AS, Rando OJ, Tjian R, et al. Resolving the 3D Landscape of Transcription-Linked Mammalian Chromatin Folding. Mol Cell 2020;78:539-53 e8.

24. Natsume T, Kiyomitsu T, Saga Y, Kanemaki MT. Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. Cell Rep 2016;15:210-8.

25. Heath H, Ribeiro de Almeida C, Sleutels F, Dingjan G, van de Nobelen S, Jonkers I, et al. CTCF regulates cell cycle progression of alphabeta T cells in the thymus. EMBO J 2008;27:2839-50.

26. Hyle J, Zhang Y, Wright S, Xu B, Shao Y, Easton J, et al. Acute depletion of CTCF directly affects MYC regulation through loss of enhancer-promoter looping. Nucleic Acids Res 2019;47:6699-713.

27. Guo YA, Chang MM, Huang W, Ooi WF, Xing M, Tan P, et al. Mutation hotspots at CTCF binding sites coupled to chromosomal instability in gastrointestinal cancers. Nat Commun 2018;9:1520.

28. Aitken SJ, Ibarra-Soria X, Kentepozidou E, Flicek P, Feig C, Marioni JC, et al. CTCF maintains regulatory homeostasis of cancer pathways. Genome Biol 2018;19:106.

29. Fang C, Wang Z, Han C, Safgren SL, Helmin KA, Adelman ER, et al.

Cancer-specific CTCF binding facilitates oncogenic transcriptional dysregulation. Genome Biol 2020;21:247.

30. Lawhorn IE, Ferreira JP, Wang CL. Evaluation of sgRNA target sites for CRISPR-mediated repression of TP53. PLoS One 2014;9:e113232.

31. Chudakov DM, Lukyanov S, Lukyanov KA. Tracking intracellular protein movements using photoswitchable fluorescent proteins PS-CFP2 and Dendra2. Nat Protoc 2007;2:2024-32.

32. Cisse II, Izeddin I, Causse SZ, Boudarene L, Senecal A, Muresan L, et al. Real-Time Dynamics of RNA Polymerase II Clustering in Live Human Cells. Science 2013;341:664-7.

33. Cho WK, Jayanth N, English BP, Inoue T, Andrews JO, Conway W, et al. RNA Polymerase II cluster dynamics predict mRNA output in living cells. Elife 2016;5.

34. Cho WK, Spille JH, Hecht M, Lee C, Li C, Grube V, et al. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. Science 2018;361:412-5.

35. Pugacheva EM, Kubo N, Loukinov D, Tajmul M, Kang S, Kovalchuk AL, et al. CTCF mediates chromatin looping via N-terminal domain-dependent cohesin retention. Proc Natl Acad Sci U S A 2020;117:2020-31.

36. Li S, Prasanna X, Salo VT, Vattulainen I, Ikonen E. An efficient auxin-inducible degron system with low basal degradation in human cells. Nat

Methods 2019;16:866-9.

37. Xie L, Dong P, Chen X, Hsieh TS, Banala S, De Marzio M, et al. 3D ATAC-PALM: super-resolution imaging of the accessible genome. Nat Methods 2020;17:430-6.

38. Moore JM, Rabaia NA, Smith LE, Fagerlie S, Gurley K, Loukinov D, et al. Loss of maternal CTCF is associated with peri-implantation lethality of Ctcf null embryos. PLoS One 2012;7:e34915.

ABSTRACT(IN KOREAN)

인간 대장암 세포에서 급속 CTCF 단백질 제거 시스템 확립을 통한 크로마틴 구조 단백질 CTCF의 유전자 발현 조절 기능 연구

< 지도교수 김형표 >

연세대학교 대학원 의과학과

강 무 구

크로마틴 구조를 조절하는 단백질인 CCCTC-결합 인자 CTCF는 멀리 떨어진 DNA 끼리의 상호작용을 매개하고 TAD(위상적 연관 단위체)의 경계를 결정짓는다. 본 연구에서는 세포 본래의 상태에 거의 영향을 주지 않으면서 인간 대장암에서 CTCF가 전사 조절에 미치는 영향을 탐구하고자 인간 대장암 세포주인 HCT116의 CTCF 유전자에 miniAID-mClover3나 miniAID-HaloTag 카세트를 끼워 넣어 AID(옥신 유도 분해체) 시스템을 제작했다. 이렇게 제작된

CTCF degron 시스템은 식물 호르몬인 옥신을 처리한지 1시간만에 효과가 나타날 만큼 신속하며 통제 가능한 CTCF 단백질의 분해를 보여주었다. CTCF 분해 이후 RNA-seq을 통한 전사체 분석을 통해 종양 성장과 전이 등 암의 전형적인 특징과 연관된 유전자를 포함한 수백개의 유의미한 유전자 발현 변화를 확인했다. 본 연구를 통해 CTCF가 인간 대장암에서 직접적으로 유전자 발현 체계를 조절함으로써 종양의 공격성에 중요한 역할을 한다는 것을 확인했다.

핵심되는 말 : CTCF, 분해체, 크로마틴 구조, 전사체, 대장암