AMERICAN UNIVERSITY OF BEIRUT

TET2 EXPRESSION AS A FUNCTION OF Cx43 REGULATION: AN IN-VITRO AND ARCHIVAL TISSUE STUDY OF IBD AND HUMAN COLON CANCER

by

KAWTHAR HASAN SHARAF AL DEEN

A thesis
Submitted in partial fulfillment of the requirements
for the degree of Master of Sciences
to the department of Anatomy, Cell Biology, and Physiological Sciences
of the Faculty of Medicine
at the American University of Beirut

Beirut, Lebanon
August 2020

AMERICAN UNIVERSITY OF BEIRUT

TET2 EXPRESSION AS A FUNCTION OF CX43 REGULATION: AN *IN VITRO* AND *EX VIVO* STUDY OF IBD AND HUMAN COLON CANCER

BY

KAWTHAR HASAN SHARAF AL DEEN

Approved by:

_____
Dr. Marwan El-Sabban, Professor and Vice Chair   Advisor
Department of Anatomy, Cell Biology,
and Physiological Sciences

_____
Dr. Wassim Abou-Kheir, Associate Professor   Member of Committee
Department of Anatomy, Cell Biology,
and Physiological Sciences

_____
Dr. Margret Shirinian, Assistant Professor   Member of Committee
Department of Experimental Pathology,
Immunology, and Microbiology

Date of thesis defense: September 7, 2020

AMERICAN UNIVERSITY OF BEIRUT

THESIS RELEASE FORM

Student Name: Sharaf Al Deen   Kawthar   Hasan
                    Last   First   Middle

I authorize the American University of Beirut, to: (a) reproduce hard or electronic copies of my thesis; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes:

☑ As of the date of submission

☐ One year from the date of submission of my thesis.

☐ Two years from the date of submission of my thesis.

☐ Three years from the date of submission of my thesis.

_____   Sep 17, 2020

Signature   Date

ACKNOWLEDGMENTS

AN ABSTRACT OF THEISIS OF

Kawthar Sharaf Al Deen   for   Master of science
                                                      Major: Physiology

Title: TET2 Expression as a function of Cx43 regulation: an In-vitro and Archival tissue study of IBD and Colon cancer

Inflammatory bowel disease (IBD) is a complex disease characterized by chronic inflammation of the gastrointestinal tract. There is evidence that IBD patients have higher propensity to develop colorectal cancer. The study of the junctional complexes is essential for a better understanding of the functional coordination between IECs and inflammatory cells in IBD, intestinal barrier function and its progression to colon cancer. On the epigenetic level, Ten-eleven translocation (TET) proteins oxidize 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) and are reported to be dysregulated in different tumors. We studied the expression of TET2 as a function of Cx43 regulation using an in vitro inflammation cell culture model, and using normal, IBD, and colon cancer human archived tissues. We documented TET2 functionality by determining the levels of 5-hmC epigenetic modification under the same conditions. Our data demonstrate that TET2 expression decreases upon an induced inflammation as well as upon the Cx43 upregulation, whereas TET2 expression increases when Cx43 is downregulated. In addition, 5-hmC levels mimic that of TET2 expression, thus reflecting TET2 functionality. The expression of TET2 and its epigenetic mark 5-hmC in Normal, IBD and colon cancer human colon tissues is highly downregulated in IBD and colon cancer human tissues when compared to Normal colon tissues.

Keywords: Inflammatory Bowel Disease, DNA demethylation, gap junctions, colorectal cancer.

CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

LITERATURE REVIEW

CHAPTER 1

DNA METHYLATION

A. DNA methylation an epigenetic DNA modification

Even though all cells in an organism possess the same genetic information, each cell type has its unique gene expression. In other words, epigenetic mechanisms mediate the diversity of gene expression among cells and tissues. One of the most important epigenetic modifications is DNA methylation, since it satisfies the stringent criterion of a mitotically and meiotically heritable epigenetic system. [1] DNA methylation involves the transfer of a methyl group to the C5 position of cytosine (5mC) and it mostly occurs at the CpG sites which are heavily methylated except for some DNA regions as CpG islands [2]. It is well known that DNA methylation is a dynamic process, but as cells differentiate they take an overall unique and stable pattern that is responsible for tissue-specific gene transcription.

B. DNA methyl transferases

A family of DNA methyltransferases (Dnmt)[1] are responsible for DNA methylation. Those catalyze the transfer of a methyl group from S-adenylyl methionine (SAM) to the fifth carbon of a cytosine residue to form 5-methyl cytosine (5mC) (Figure1). Dnmt1, called maintenance Dnmt, copies the methylation of the newly synthesized DNA strand upon DNA replication to maintain the original DNA methylation pattern, whereas Dnmt3a and Dnmt3b, known as de novo Dnmt,

can methylate an un-modified DNA strand. [3] The maintenance of DNA methylation can be highly conserved as the two female X chromosomes, where each one has a specific DNA methylation pattern from the early embryonic state till the end of life.[4] On the other hand, DNA methylation patterns change dynamically in cells to regulate gene expression and other cellular mechanisms. Research showed that the Dnmt is extensively expressed upon embryonic development, but when cells reach their final differentiation, Dnmt expression is reduced.[3] Among Dnmts is Dnmt2 which shows a weak methyltransferase activity and Dnmt3L that is catalytically inactive but seems to interact directly with Dnmt3a and Dnmt3b stimulating their catalytic activity.[5-7]
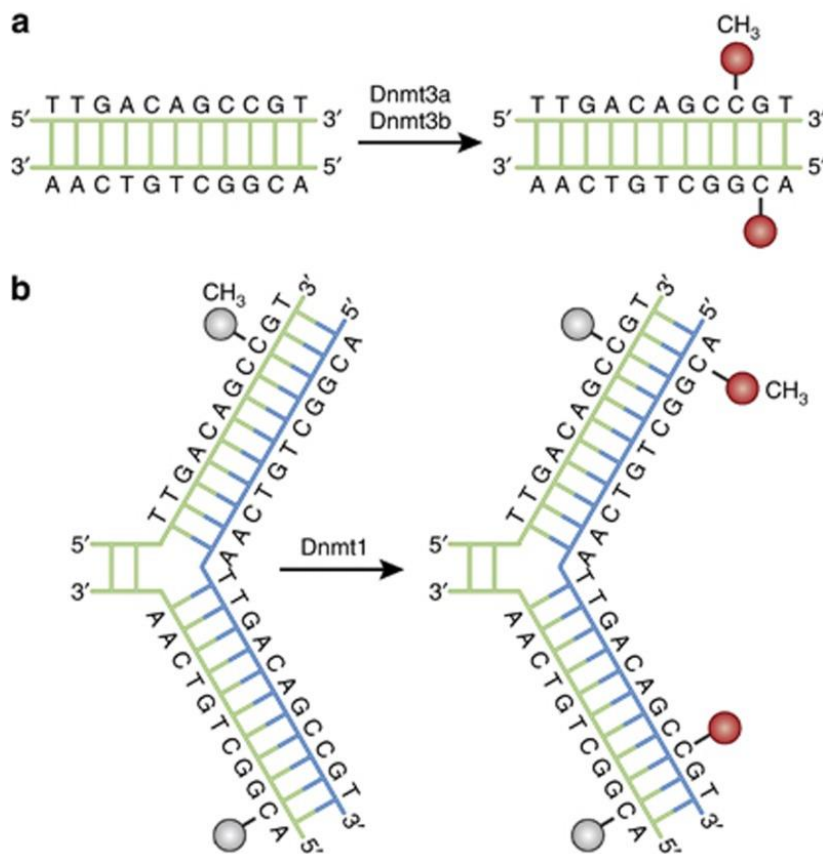
Figure 1: DNA methylation pathways.

(a)De novo Dnmt3a and Dnmt3b responsible for the methylation of naked DNA. (b) Maintenance Dnmt1 keeps the same DNA pattern during replication. During DNA replication, Dnmt1 precisely replicates the original DNA methylation pattern [2] by adding methyl groups on to daughter strands (blue)

C. DNA methylation effect on gene expression

Three different families of proteins recognize DNA methylation: the MeCP2, the MBD proteins, the UHFR proteins, and the zinc finger proteins. The MeCP2 and MBD family has a function in transcriptional repression. For example, MBD family proteins binds directly by its transcriptional repression domain (TRD) to a variety of repressor complexes at methylated DNA sequences thus acting as a transcriptional repressor.[3] Also, zinc-finger domain proteins repress transcription in a DNA- methylation dependent manner.[4] However, UHRF domain (ubiquitin-like, containing PHD and RING finger domain) maintains DNA methylation upon DNA replication by targeting Dnmt1 proteins to hemi-methylated cytosines.[5] Thus different protein domains suppress gene expression by binding methylated DNA.

At the level of gene bodies, DNA methylation is also analyzed. Evidence suggests that methylation of promoters is responsible for gene silencing[6], and that a higher DNA methylation in the gene body of dividing cells is associated with higher gene expression.[7] However, in slowly proliferative cells and non-proliferative cells such as in the brain, gene expression is not associated with gene methylation.[8] In particular, the promoters of housekeeping genes are mostly imbedded in CpG islands, that does not undergo methylation, which can explain the conserved gene expression of housekeeping genes throughout various

tissue.[9] Therefore, DNA methylation affects gene expression and the methylation of various genomic sequences influence gene expression differently depending on the underlying genetic sequence.

D. Importance of DNA methylation

DNA methylation is responsible for the regulation of different genetic processes, including embryonic development, transcription, X-chromosome inactivation, DNA repair as well as DNA replication. One of the most important roles of DNA methylation is to maintain the silencing of harmful genetic elements as transposable and viral elements that form about 45% of the human genome. This is demonstrated in the pathophysiology of different autoimmune diseases and cancer, where the hypomethylation of long interspersed nuclear elements (LINEs) autonomous retrotransposons and human endogenous retroviruses (HERVs) leads to their activation. [10] Accordingly, defects in DNA methylation may cause genome instability and thus contributes to various human diseases, including cancer.

E. DNA demethylation pathways

DNA demethylation is categorized as either passive or active demethylation. The passive DNA demethylation occurs upon inhibition or dysfunction of Dnmt1 in DNA replication, thus the newly incorporated cytosines will remain demethylated. The strong carbon-to-carbon bond is never cleaved by a direct process. Instead, the removal of the methyl group from 5mC requires series of enzymatic reactions, those reactions modify 5mC by deamination or oxidation reactions followed by the base excision repair (BER) pathway that replaces the modified-5mC with a naked cytosine. [11] [12] [13]

Activation induced deaminase (AID) deaminates cytosine to Uracil and, to a lesser extent, to thymine, thus creating G/T mismatch and inducing BER pathway to correct the base. This whole pathway seems to be incomplete, where some studies state that the methylation action of AID on C is better than that on 5mC, others suggest that the deamination of C rather than 5mC leads to a regional repair, and thus an indirect demethylation of 5mC. Interestingly, AID showed to work on single stranded DNA or RNA-DNA hetero-duplexes only.

Ten-eleven translocation (TET) methyl cytosine dioxygenases are a family of enzymes that perform oxidation of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5FC) and 5-carboxylcytosine (5caC).[14] This pathway is the major DNA demethylation pathway that is further explained in the next chapter.

CHAPTER 2

TEM ELEVEN TRANSLOCATION (TET) DIOXYGENASES

Ten-eleven translocation (TET) methyl-cytosine dioxygenases are a family of three enzymes TET1, TET2, and TET3 which perform oxidation of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5FC) and 5-carboxylcytosine (5caC).[14] It has been demonstrated that the oxidation products are in a dynamic state and are not short-term transient intermediates in the DNA methylation enzymatic pathway. Those intermediates are incorporated in the DNA with high stability and are responsible for significant and pivotal regulatory functions.

In mammalian genomes, the relatively stable epigenetic modifications are known to be 5mC and 5hmC where as 5fC and 5caC are extremely rare and directly removed by thymine DNA glycosylase (TDG). [15]

A. TET nomenclature

TET proteins nomenclature was after the finding of the ten-eleven translocation (t(10;11)(q22;q23)) in rare cases of acute myeloid and lymphocytic leukemia. In this translocation the mixed-lineage leukemia 1 (*MLL1*) gene located on human chromosome 10 was fused with *TET1* gene on human chromosome 11.[16]

B. TET proteins structure and domains

TET proteins are multi-domain proteins with a molecular mass that ranges between 180 and 230 kDa. The structure of TET family proteins have a common catalytic C-terminal CD domain (figure 3), which consists of a conserved double-stranded β-helix dioxygenase domain (DSBH), and a cysteine-rich (Cys) region. [17] Besides, the DSBH domain contains metal-binding sites for α-ketoglutarate (α-KG) and Fe which functions as a co-substrate for the TET enzymatic activity and iron-dependent dioxygenase activity, respectively. [18, 19] The TET2 binding to DNA was described as a crystal structure, this structure enables the two zinc fingers to bring the DSBH fold and the Cys-rich domains together to form a compact catalytic domain.[20] On the other hand, the amino (N)-terminal region of TET1 and TET3 harbors a conserved CXXC-type domain which is known to bind the un-methylated cytosine residues in DNA. In TET2 case, the *CXXC* domain is split from *TET2* to form a distinct gene, *IDAX/CXX4*, at the 5'-end of *TET2*. Interestingly, this gene is transcribed in the opposite direction and the Idax CXXC domain binds DNA sequences at un-methylated CPGs, localized at promotors, which interact directly with the catalytic domain of TET2 and results in caspase activation and TET2 protein degradation thus TET2 downregulation. [21]
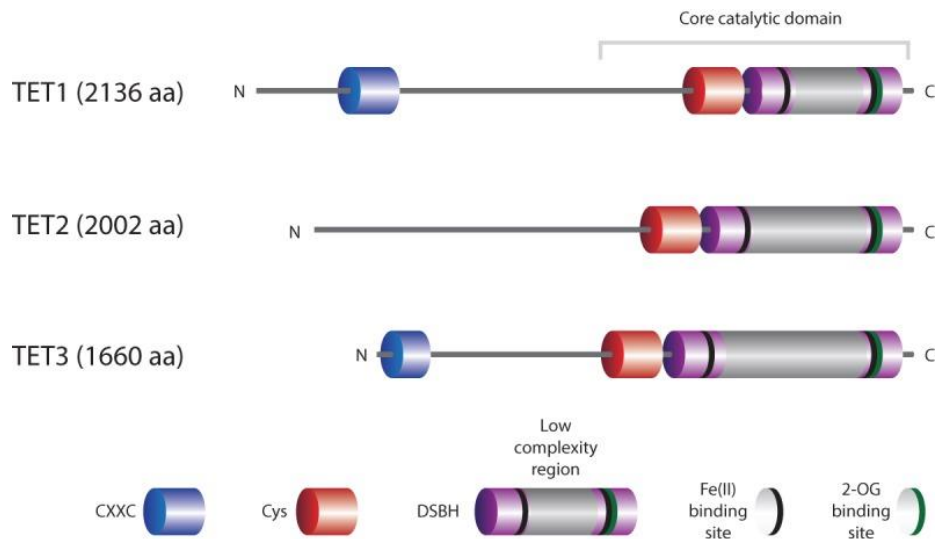
Figure 2: TET proteins structural domains.

All three TET enzymes share the same C-terminal core catalytic domain consisting of DSBH domain, Fe and 2-OG cofactor binding sites, and a cysteine rich domain. It also contains a large low-complexity region with unknown function. TET1 and TET3 have an N-terminal CXXC domain that binds DNA and facilitate its binding to genomic target sites.[22] This CXXC domain is involved in TET1 and TET3 binding to CpG sites where studies showed that TET1 is highly enriched on CpG islands associated with promoters and TET3 occurs at genomic elements in developing embryos. [23] Whereas, the structure of TET2 doesn't reveal its direct binding to DNA, and thus it is potentially recruited to DNA by a distinct CXXC domain. TET2 is known to bind at gene bodies, regulatory elements and transcription factors.[24, 25]

19

C. Mechanisms of TET mediated demethylation

The conversion of 5mC and its oxidized derivatives back to cytosine can be either "passive" or "active" demethylation. TET- mediated 5hmC deposition may induce passive-replication-dependent DNA demethylation on the newly synthesized DNA strand. Other articles have shown that DNA methylation maintenance may be established by the binding of Dnmt1 to 5hmC through its interaction partners UHRF1 and UHRF2.[26] For instant, the overall passive demethylation at 5hmC is still not fully understood.

In active DNA demethylation, several pathways have been proposed, but the pathway involving thymine–DNA–glycosylase (TDG)-catalyzed base excision and DNA base excision repair (BER) is believed to be the major pathway (Figure 3). TDG, a DNA mismatch repair enzyme, binds and excises mismatched pyrimidines in G: U and G: T base pairs. This excision yields an abasic site that is replaced by a cytosine by BER. [27-29] It is demonstrated that TDG has no effect on 5hmC but works on 5fC and 5caC. [27, 29, 30] This suggests that 5fC and 5c are constantly being removed by TDG/BER-dependent pathway and that 5hmC is a stable cytosine modification that remains in the genomic DNA for an extended period of time.
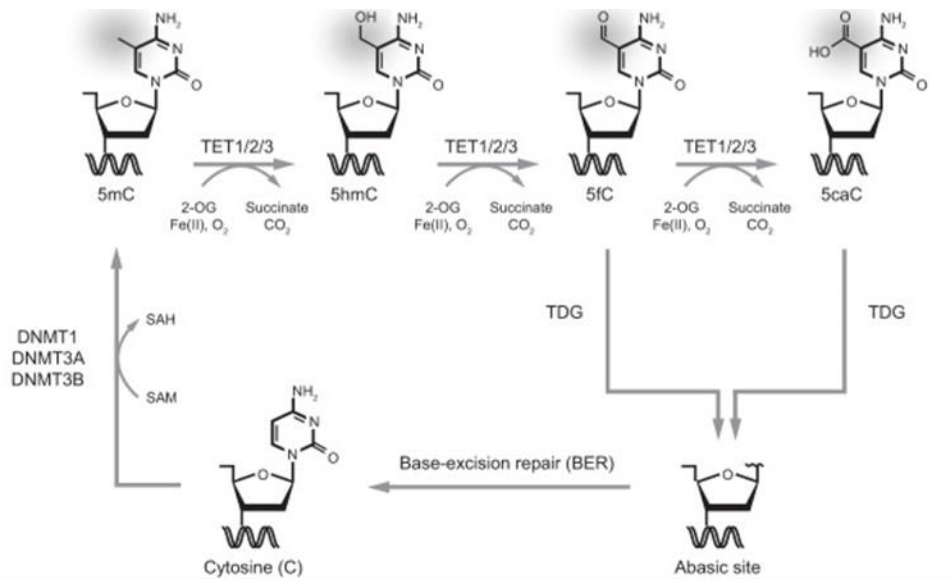
Figure 3: DNA demethylation involving thymine–DNA–glycosylase (TDG)-catalyzed base excision and DNA base excision repair (BER)

D. *TET2* mutations in Hematological Malignancies

TET2 mutations were first reported in hematological malignancies in MDS patients with chromosome 4q24 abnormalities. [31] It was reported that 26% of MDS patients had mutations in the coding region of *TET2*. [32] Other studies reported that patients with AML, MDS, or myeloproliferative neoplasm, have mutations or deletions in *TET2* gene.[33] [34]In addition, *TET2* mutations are present in chronic myelomonocytic leukemia (CMML) and malignant lymphoma. [1] These findings suggest that the mutations disrupt the catalytic activity of TET2 and constitutes a critical background for the initiation and development of hematological malignancies.

E. THE ENZYMATIC ACTIVITY OF TET FAMILY PROTEINS IN TUMORS

Mutations or decreased expression of TET proteins have been observed in a wide variety of human solid tumors. Missense and truncating mutations in TET genes are reported in nearly all types of tumors. Some studies showed that up to 20% of patients carry mutations in one or more of the TET genes (http://www.cBioPortal.org). Other studies showed that TET mutations in solid tumors are mostly missense mutations with unknown significance on TET activity. In addition, lower 5hmC levels were reported as a general hallmark of many cancer types, including gastric, prostate, liver, lung, and breast cancer as well as glioblastoma and melanoma [35] [36] [37]

The mechanisms that lead to a decline in TET activity is not fully known, but several ones were proposed. One mechanism is the overexpression of TET2-interacting proteins IDAX/CXXC4 and RINF/CXXC5 which are known to have a negative effect on TET2 stability and function. [38] Also, several microRNAs have been reported to target TET proteins and these miRNAs are known to be overexpressed in cancer solid tumors. [39] [40] [41] Other researches stated that the mutations in different metabolic genes as *IDH1/2*, *SDH*, and *FH* that are found in different solid tumors may simultaneously inhibit TET proteins thus impact DNA methylation. [42]

F. 5hmC levels in normal and tumor tissues

It is well known that the main function of TET proteins is to catalyze the oxidation of 5mC into 5hmC thus the 5hmC level reflects the action of TET proteins. Reports showed that 5hmC levels are reduced in hematological malignancies and a broad spectrum of tumors including liver, breast, colon, and prostate compared to normal tissues. [43, 44] 5hmC have shown to be

abundant in the human brain tissue where 5hmC levels are about 1% of all cytosines, which

signifies the important role of 5hmC in the mammalian brain. [45] Other studies have shown that

5hmC levels are enriched in terminally differentiated neurons, whereas less differentiated

neurons have lower 5hmC levels. Taken all together, 5hmC levels differ by tissue type and

differentiation stage and tumor type.

CHAPTER 3

THE INTESTINAL MUCOSAL BARRIER

The Mucosal barrier has the ability to provide adequate containment of undesirable luminal microorganisms and molecules while providing its main function in nutrient and water absorption. This mucosal barrier is well known to be altered in IBD and colon cancer.

A. Components of the Intestinal Barrier

1. The epithelium

This physical barrier, the epithelium, is a continuous monolayer of columnar epithelial cells that are highly organized into crypts and villi. It is composed of 4 major Intestinal epithelial cells (IECs): absorptive enterocytes, mucus-producing goblet cells, antimicrobial peptide producing Paneth cells, and hormone-producing enteroendocrine cells. [36] Paneth cells and Stem cells reside at the base of the crypt whereas other differentiated cells migrate towards the tip of the villus which makes the crypt a secretory compartment whereas the villus an absorptive one. There also exists specialized epithelial M cells for antigen sampling and endocytosis of different molecules that are then rapidly transported to the underlying Payer's patches, lymphoid tissue.[46] The epithelium is covered by a layer of mucus, AMP, and Secretory IgA. IECs regulates the intestinal barrier thus are responsible for maintaining the mucosal homeostasis. The Dysregulation and alteration of this epithelial barrier has been linked to different conditions, including IBD .[47]

2. *Microbiota:*

The microbiota colonizes the gastrointestinal tract and forms a mutual relationship with the host. It is an important component of the Intestinal barrier and affects it directly and indirectly. The microbiota stimulates epithelial proliferation by the secretion of IL-8, one of the main chemokines, and IL-22 [48] , it also has a role in immunostimulation, relocation of tight junction (TJ) proteins, and induction of heat shock proteins. [49] [50] The host-friendly bacteria is responsible for dietary fiber digestion and provides the host with short-chain fatty acids which are a source of energy and have an anti-inflammatory role. [51] Alteration of the normal microbiota lead to inflammation and it is a known feature of IBD.

3. *The mucus layer:*

The mucus layer forms a physical and biochemical barrier and acts as the first line of defense in the gastrointestinal tract. It is composed of an outer loose mucus layer and an inner adherent mucus layer that is largely devoid of bacteria and separates them from IECs.[52] Mucus is mainly composed of mucins (MUCs) that are divided into two groups, secretory-gel forming mucins (MUC2, 5AC, 5B, and 6) that are produced by goblet cells and membrane-associated MUCs (MUC1, 3, 4, 13, and 17), secreted by both goblet and absorptive cells in their apical membrane. MUC2 forms the major mucus component forming the effective mucus layer. [53]

Recent studies have been shown that IBD patients have altered mucus layer and mucin secretions.[54]

4. *Antimicrobial Peptides and sIgA*

AMPs are secreted by Paneth cells, enterocytes and immune cells into the crypt's space and the mucus layer. They include defensins, cathelicidin, regenerating islet-derived protein, secretory leukocyte protease inhibitor, bactericidal/permeability-increasing protein, lysozyme, lactoferrin, and also chemokines such as chemokine (C–C motif) ligand (CCL) 14 and CCL15. [55] AMPs are part of the antimicrobial defense in the intestinal barrier. For instance, defensin alpha5 stimulates IL-8 secretions, but cathelicidin has wound-healing and anti-apoptic effects in IECs. [55] sIgA is secreted by B lymphocytes and then secreted into the lumen as dimers. It is the main antibody isotype found in the intestinal mucosa. sIgA neutralizes bacterial epitope expression and maintain intestinal homeostasis, in which the absence of IgA induces inflammation.[56]

B. Overview of Inflammatory bowel disease

Inflammatory bowel disease (IBD) is a complex disease characterized by chronic inflammation of the gastrointestinal tract. It develops as a result of complex interactions between individual's genetic background, environmental factors, a dysregulated immunity, as well as

alterations in intestinal microbiota composition. [57] It is well recognized that IBD patients are at higher risk of developing colorectal cancer .[58]

1. *Inflammatory bowel disease Subtypes*

Chronic IBD is mainly divided into two subtypes; Ulcerative colitis (UC) and Crohn's disease (CD). Both UC and CD cause digestive disorders and inflammation in the gastrointestinal tract and share similar symptoms including diarrhea, abdominal pain, rectal bleeding and weight loss.[59] Ulcerative colitis is continuous inflammation of the inner mucosal layer of the colon and usually starts in the rectum. It is associated with bloody stool and severe pain. On the other hand, Crohn's disease is characterized by a discontinuous, transmural ulcerations in all layers of the bowel wall and through the entire gastrointestinal tract, but it mostly involves the terminal ileum and the colon. Rectal bleeding is less common in CD but it exists in severe cases. "Indeterminate colitis" is a term describing indistinguishable cases of patients with inflamed colon.[60, 61]

2. *Genetic factors contributing to IBD pathogenesis*

Although only a minor risk for IBD can be explained by genetic factors alone, the genome-wide association studies (GWAS) has identified 240 susceptible loci in the human genome [62]. The first found gene to be associated with CD was *NOD2* located within IBD1 on chromosome 16. NOD2 protein is known to play an immunity role by activating nuclear factor kappa B (NF-κB) and recognizing bacterial receptors (LPS).[63] On the

other hand, IBD2 carrying advilin gene is associated with UC extensively. In addition, mutations in *HLA, IL23*R, *ATG16L1, IRGM, TNFSF15, NCF2/NCF4* and *PTPN22* have been associated with chronic IBD.[64] Those mutations directly affect autophagy, phagocytosis, Paneth cell failure as well as signaling pathways of the adaptive immunity, such as IL-10 mutations. [65] Actually, the risk of developing IBD is only 20% higher than any other individual who is not a carrier of a particular mutation. Which keeps genetic testing of IBD an un-useful test for IBD diagnosis. [62]

3. *Epigenetic modification in IBD*

Different studies showed altered epigenetic modifications in IBD tissues. In correspondence to DNA methylation, it has been reported that accelerated DNA methylation and cell turnover was found in locally inflamed colonic epithelium. Actually, increased DNA methylation results in genetic instability leading to cancer development. It has been found that genes already related to carcinogenesis are differentially methylated in the colonic mucosa of UC patients, such as*CDKN2a/p16INK4A, CDKN2a/p14ARF, CDH1, MLH1, HPP1* and *MYOD1*.[66] [67] Epigenome-wide methylation association studies (EWAS) reported 50 significant altered methylation sites in CD patients. Some of those gene have a critical role in immune responses, such as *IL21R, S100A13, FASLG, MAPK13, RIPK3* or *PRF1*.[68] In addition, GWAS analysis revealed a relation between IBD and Polymorphism in *DNMT3a*. Moreover, higher expression of DNMT3b and DNMT1 were reported. [2]

C. Cancer secondary to chronic intestinal inflammation

Chronic intestinal inflammation is the primary risk factor for cancer including CRC, small bowel adenocarcinoma, intestinal lymphoma, and cholangiocarcinoma (Table 1). Different sporadic molecular mechanisms are described leading to colon cancer in IBD patients. These include mutations in the adenomatous polyposis coli (*APC*) gene, aneuploidy, DNA methylation, microsatellite instability (MSI), activation of the oncogene *k-ras*, activation of *COX-2*, and mutation in tumor suppressor genes *DCC/DPC4*, and eventual loss of *p53* function.[69].

The immune system plays an important role in the initiation and prognosis of carcinogenesis. Indeed, the inflammatory microenvironment of IBD mimics that of cancer at the level of immune cells, epithelial cells, stromal cells, cytokines, as well as chemokines. These mediators include tumor necrosis factor alpha (TNF-α), ILs-1, 6, 12, 13, 17, 22, and 23.[70]

Table 1: Cancer secondary to chronic intestinal inflammation [71]

Cancer type	Standardized incidence ratio
Colorectal cancer[3]	5.7 (95%CI: 4.6-7.0)
Small bowel adenocarcinoma[20]	27.1 (95%CI: 14.9-49.2)
Intestinal lymphoma[36]	17.51 (95%CI: 6.43-38.11)
Anal cancer[60]	Data not available
Cholangiocarcinoma[23]	916.63 (95%CI: 297.88-2140.99) in UC

CHAPTER 4

CONNEXIN 43

A. The Connexin Family

Gap junctions are intercellular plasma membrane channels that facilitate the intercellular exchange of ions and small molecules (less than 20kDa) between the cytosol and adjacent cells. This intercellular communication is critical for functional coordination between cells and cellular development, where coupled cells respond to a ligand synchronously and differentiate together to acquire the nearly same fate. In vertebrates, gap junctions are composed of a family of proteins called Connexins

1. *Connexin nomenclature*

The protein nomenclature employs an abbreviation beginning with Cx for Connexin followed by a number that resembles the predicted molecular weight from the cDNA sequence of the Connexin (for example Cx43). [72] Connexin corresponding genes are identified with a symbol starting with *GJ* for gap junction, followed by the subfamily symbol corresponding to the differences between Connexin sequences (α, β, γ, δ, and ε or *GJA*, *GJB*, *GJC*, *GJD*, and *GJE*) (Table 2). [73]

Table 2: The Connexin protein and gene families in humans. [74]

Gene symbol	Protein name
GJB1	CX32
GJB2	CX26
GJB3	CX31
GJB4	CX30.3
GJB5	CX31.1
GJB6	CX30
GJB7	CX25
GJA1	CX43
GJA3	CX46
GJA4	CX37
GJA5	CX40
GJA8	CX50
GJA9	CX59
GJA10	CX62
GJC1	CX45
GJC2	CX47
GJC3	CX30.2/31.3
GJD2	CX36
GJD3	CX31.9
GJD4	CX40.1

2. *Connexin gene structure*

The human Connexin family comprises 21 confirmed connexin genes that have tissue and developmental- specific functions. The Connexin genes show a similar structure, most of them are made of two exons (Ex1, Ex2), one in-between intron, and 3' and 5' untranslated region (UTR). Ex2 is uninterrupted, however in Cx45 three exons with the coding region in Ex3. Other connexin genes, like Cx36 gene has its coding region in Ex1 and Ex2, interrupted by an intron. Other connexins are transcribed by different promotes (Ex1A and Ex1B) resulting in different transcripts over different tissues (Figure 4). [75]
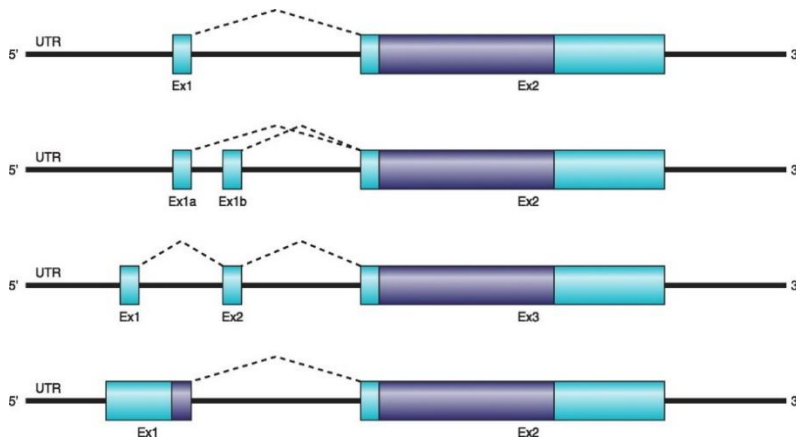
Figure 4: Connexin gene structure.[75]

3. *Connexin protein structural domains*

The membrane topology of Connexins includes tetra-span transmembrane (TM) domains connected by two Extracellular loops (ECL1 and ECL2) and one intracellular loop (IL), in addition to cytoplasmic amino (NT) - and carboxy-termini (Figure2). Despite the highly conserved homology of Cxs, there are important differences in the length and the amino acid sequence of CT and ICL domains. Connexin-based channels are formed by the oligomerization

of connexin proteins forming heteromeric (HC consists of more than one Cxs type) and heterotypic channels (two homomeric channels each made by a different Cxs type). Research have showed that the NT region is responsible for the oligomerization compatibility between Cxs. While, the CT domain encloses many phosphorylation sites for GJ regulation and protein-protein interactions. [75] [76] (Figure5)
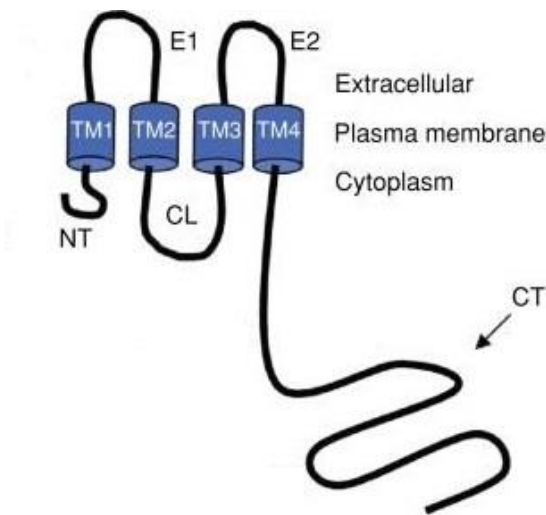
Figure 5: Connexin protein structural domains[77]

4. *Connexin transcriptional modifications*

a. Transcription factor regulation

Transcription factors control connexin gene expression which consists of two or more exons as mentioned earlier. This transcription is controlled by a combination of tissue specific transcription factors. An interesting example is the heart, in which Cx40, Cx43, and Cx45 are differently expressed in the heart chambers and in the conducting system. Cx43 is the sole

connexin of the ventricular myocardium, whereas Cx43 and Cx40 are co-expressed in the atria.[78] Connexins are also regulated within very short time frames in fundamental physiological functions. As an example is the upregulation of Cx43 in the uterus upon labor, which is necessary for the synchronization of forceful coordinated uterine contractions. [79]

b. Epigenetic regulation at the level of DNA methylation

Cx43 mRNA expression levels are also regulated by histone acetylation, DNA methylation and micro-RNAs. DNA methylation plays a crucial role in the normal regulation of connexin expression. An example is the expression of Cx43 instead of Cx32 upon the methylation of Cx32 promoter in the liver epithelium.[80] DNA methylation is also involved in the pathophysiological conditions, where connexin expression in cancer is altered differently.[81, 82] [83]

5. Connexin post-translational modification

Most Connexins are modified post-translationally through phosphorylation, hydroxylation, acetylation, ubiquitination, disulfide binding, and nitrosylation. But the most extensively studied modification of Connexin in undoubtedly phosphorylation. Using different available techniques, it was demonstrated that Cx phosphorylation is on the CT tail except for Cx26. This step is important for the correct oligomerization of connexins into connexons thus forming functional channels and precise traffic of channels to the membrane.[84]

B. Connexin channel

1. Connexin channel Structure

Six Connexins oligomerize to form hemi-channel, called Connexon which then forms a Gap junction by docking another hemichannel from adjacent cells. Gap junctions can be either formed of six identical Cx subunits (Homomeric) or can be formed of six different Cx subunits (Heteromeric). Gap junction channels may be homotypic, when two identical connexons assemble, or heterotypic, when two different connexons assemble between adjacent cells. (Figure 6) [85]. In addition, individual connexons can function as hemi-channels and allow the diffusion of molecules from ECM to the cytosol.
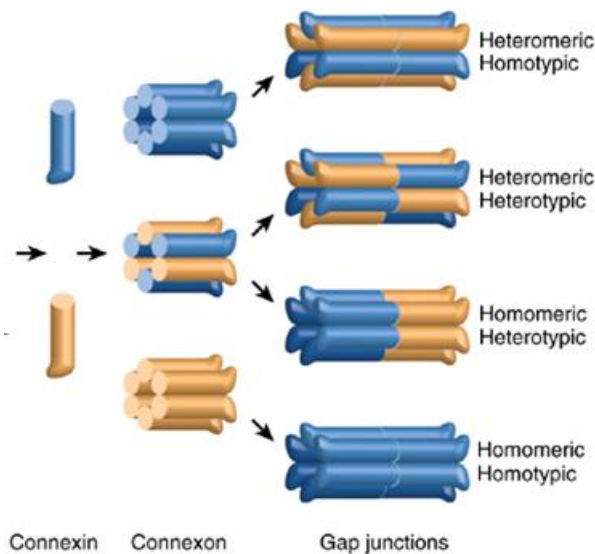

Figure 6: Gap junction types [85]

2. *Connexin channel Function*

At the molecular level, Cx43 have three main functions: Cell- cell communication between adjacent cells, hemichannel activity, and direct connexin-protein interaction through its C-terminal to activate different signaling pathways. Each function plays and important role in normal and patho-physiology.

At the physiological level, Gap junctions play a crucial role in cellular development, embryonic growth, differentiation and regeneration. Research had mentioned their essential role in electrical coupling in cardiac tissue, uterine contraction, alveolar differentiation, bone formation, lens growth, the inner ear, gastrointestinal tract and central nervous system.

C. Connexin 43 in colonic epithelium and cancer

In the gastrointestinal tract, gap junctions are highly expressed especially in the inner circular layer of the intestinal smooth muscle to coordinate the contractile activity of the gut. In the normal colonic epithelium, three major connexins have been shown at the protein level, Cx26, Cx32 and Cx43. The most studied isoform in the contest of intestinal inflammation and motility is Cx43. In Hireschprung's disease the gut at certain parts is aganglionic and lack normal motility. It was found that Cx43 expression is severely reduced or absent leading to decreased levels of gap junctions, thus explaining the disturbance of colon motility.[86] At the level of colorectal cancer, loss of Connexin expression or relocalization from the plasma membrane to the cytosol was suggested. Therefore, Gap junctions lose their main function in colorectal cancer.[87]

AIM OF THE STUDY

The aim of the study is to explore the variation of TET2 protein expression in conditions of colonic inflammation and its transitional state to colon cancer. This was accomplished by an *in-vitro* model of IBD, established by treating IECs with inflammatory mediators or by overregulating and downregulating Cx43 in IECs which showed to be altered upon inflammation. This was also studied by screening TET2 expression in IBD and colon cancer archived human colon tissues.

MATERIALS AND METHODS

A. Cell lines and culture conditions

In this study, we used human intestinal epithelial cell line (IECs): HT-29, which is a
colorectal adenocarcinoma cell line widely used model of intestinal transport and pathology,
including inflammation. HT-29 cells were previously transduced in our lab with Cx-43
dendra2 lentiviral vectors so that we have HT-29 cells that overexpress Cx43 (HT-29 43D
cells) and HT-29 cells that were transfected with Cx-Sh plasmids to have HT-29 43- cells.
We also used human non-adherent, monocytic cell line (THP-1) and a communication
deficient cell line (HeLa). THP-1 cells were activated with 50ng/ml PMA for 24 hours and
after 80-90% confluency we added 1μg/ml of LPS for 3-4 hours. Cells were left to grow for
72 hours before collecting the Activated THP-1 supernatant (condition media) that was used
to induce inflammation. Cells were maintained in RPMI (HT-29, THP-1), and DMEM AQ
(HeLa). All media were supplied with 10% FBS, 100U/ml penicillin G and 100μg/ml
streptomycin. Cells were cultured at 37 $^{\circ}$C in a humidified incubator with 5% CO_2
atmosphere.

B. Construction of Cxs-Dendra2 chimeras

Cx43-Dendra2 chimeras were constructed by cloning Connexins (Cxs) into the multiple
cloning site (MCS) of pDendra2-N plasmid. Dendra2 is amonomeric green to red photo
convertible fluorescent protein. Dendra2 was ligated to the C-terminal of the Cxs using Hind
III and BamH1 restriction enzymes. cDNA covering the complete reading frame of Cx26 and
Cx43 were synthesized from 1 μ g of total cellular RNA and amplified by Phusion Flash

High-Fidelity PCR Master Mix. PCR products were separated by electrophoresis on 1% agarose gel, and visualized with ethidium bromide staining. PCR fragments corresponding to Cx26-cDNA (780bp), Cx43-cDNA (1171bp), and pDendra2N (4700bp) were excised, purified, and digested with restriction enzymes. A ligation reaction containing three molar excess of the purified Cxs-cDNA to pDendra2-N was allowed to proceed for 20 min at room temperature (RT) using T4 DNA ligase. Aliquots of the ligation reaction were transformed into DH5 α competent bacteria by heat shock. Positive colonies were identified by restriction enzymes analysis. New constructs within the plasmid were confirmed by sequencing (Applied Biosystems 3500 Genetic Analyzer, USA).

C. Transduction of IECs with Cxs-Dendra2 lentiviral vectors

For efficient delivery of Cxs-Dendra2 chimeric protein into IECs, we generated lentiviral vectors by cloning Cxs, N-terminally tagged with Dendra2 into the transfer vector pCSCW under the control of the CMV promoter. Using calcium phosphate, 293T cells were transfected with three lentiviral plasmids: pCMVDR8.91 (containing gag/pol), pVSVG2 (containing the envelope gene VSV-G) and pCSCW-Cx-Dendra2. The cell supernatant containing the recombinant lentivirus was collected 48-72 h post transfection, filtered and frozen. To determine the viral titer, HeLa cells were transduced with the produced virus and the number of fluorescent cells was quantified by flow cytometry. The titer was extrapolated from the percentage of fluorescent cells, which correlates directly to the number of transducing viral units present in the supernatant used (tu/ml). A total of 2x106 virus particles were used to transduce 2x105 IECs. Forty-eight hours later, the efficiency of transduction was determined by the percentage of fluorescent cells on the plate. Transduced cells expressing the Cx-Dendra2 fusion proteins were used for biochemical studies.

D. Fluorescence-Activated Cell Sorting (FACS)

Cultured HT-29 43D cells; previously transduced with Cx43Dendra Lentiviral vectors, were detached by trypsinization, centrifuged at 100xg to collect the pellet which was then re-suspended in complete media. Those cells were sorted using FACS Aria SORP sorter (BD biosciences). Highly positive Dendra2 cells were collected in complete media.

E. Quantitative PCR

Total RNA from confluent HT-29 cell line was extracted using QIAGEN RNeasy Kit as per the manufacturer's instructions. 1µg of the total RNA was reverse transcribed to cDNA using iQ SYBR Green Supermix in a CFX96 system (Bio-Rad Laboratories). kit. Quantitative PCR (qPCR) products were amplified using primers that recognize Cx43, TET2 and GAPDH. The primers that are used are listed in Table3.

Table 3: Quantitative PCR primers

Genes	Primer sequence
GAPDH	F: TGGTGCTCAGTGTAGCCCAG R: GGACCTGACCTGCCGTCTAG
Cx43	F:CTTCACTACTTTTAAGCAAAAGAG R: TCCCTCCAGCAGTTGAG
TET2	F: AAGGAGACCCGACTGCAACTG R: TCTTGAGAGGGTGTGCTGCTG

F. Western Blot

Cells were washed with PBS and scraped on 4 $^\circ$C in lysis buffer (0.5M Tris-HCL buffer, pH6.8, 2%SDS, 20%glycerol, phosphate and protease inhibitors). The samples were loaded onto 8% SDS-polyacrylamide gel, subjected to electrophoresis, and transferred to PVDF membrane. Membranes were blocked with 5% fat-free milk and 0.05% Tween in PBS (TPBS). For western blotting the following antibodies were used: GAPDH, TET2, Cx43, Dendra, and Actin.

G. Slot Blot

Total DNA from confluent HT-29 cell line was extracted using QIAGEN kit as per manufacturer's instructions. DNA samples were diluted to needed increasing concentrations for the same total volume 25μl and denatured at 95°C for 15mins. DNA was then directly Blotted using vacuum Slot Blot machine on N+ Nylon membrane. DNA was crosslinked using UV at optimum power. Two membranes were used for same conditions, one stained with methylene blue and the other blocked with 5% free-fat milk then incubated with 5-hmC antibody overnight.

H. Immunofluorescence on cells

HT-29 cells were cultured on coverslips, fixed with ice-cold methanol and stored at −20 °C. Cells were then washed 3X with PBS, and blocked with 5% NGS in PBS for 1 h in a humidified chamber. Cells were incubated with primary antibody for Cx43, TET2, and 5-hmC antibodies overnight at 4°C, followed by washing and incubation with IgG-conjugated secondary antibody (Alexa Fluor 488, Texas red at a concentration of 1 μg/ml) for 1 h. The

cells were then incubated with DAPI (1 µg/ml) for 10 min to stain their nuclei, washed 3X with PBS, mounted on slides using Prolong Anti-fade kit and observed using Zeiss LSM710 confocal microscope (Zeiss, Oberkochen, Germany), and images were processed using Zen 2009 (Carl Zeiss). Z-stacks of images were acquired using a 63x/1.46 Oil Plan-Apochromatic objective.

I. Immunofluorescence on paraffin embedded tissues

Tissue sections were obtained from normal and IBD and colon cancer patients. All patients identifiers were kept confidential and no use of identities were utilized in this study. Biopsies referred to as "normal" were obtained from patients presented with abdominal pain; the diagnosed biopsies ruled out IBD. After de-parafinization of slide sections, antigen retrieval was performed by incubating sections in sodium citrate buffer (pH 6.0) and incubated in a steamer for 40 min. Sections were allowed to cool for 15 min, washed 2X with deionized H2O and then were incubated with 100% ice-cold methanol at -20 °C for 2 hours. After several washes, sections were blocked with 5% NGS in PBS for 1 h in a humidified chamber. Sections were incubated with primary antibodies overnight at 4 °C, followed by washing and incubation with IgG-conjugated secondary antibody (Alexa 488, Texas red: 1 µg/ml) for 1 h. The sections were then washed 2X with PBS, mounted with Prolong Anti-fade, and observed using Zeiss LSM710 confocal microscope (Zeiss, Oberkochen, Germany), and images were processed using Zen 2009 (Carl Zeiss). Z-stacks of images were acquired using a 63x/1.46 Oil Plan-Apochromatic objective. Z-stacks of images were acquired using a 63x/1.46 Oil Plan-Apochromatic objective.

RESULTS

1. Expression and localization of TET2 in cultured intestinal epithelial cells as an effect of Cx43 regulation and in inflammatory conditions

TET2 profiling as a function of Cx43 regulation

In our previous research work, our lab showed that Cx43 protein expression is affected with inflammation. And in order to determine the effect of Cx43 regulation on TET2 expression in intestinal epithelial cells, screening for TET2 at the transcriptional, translational and cellular localization levels was performed. TET2 expression was assessed in the human intestinal epithelial cell line: HT-29, HT-29 transduced cells with Cx43-dendra, thus overexpressing Cx43, and HT-29 transfected with shCx43 plasmid, thus downregulating Cx43.

First of all, a western blot was performed to verify Cx43 downregulation in HT-29 43- cells and the expression of Cx43-Dendra in HT-29 transduced cells (Figure7A). Also we showed that TET2 expression varied as a function of Cx43 regulation (Figure7). At the transcriptional level, HT-29 43D cells expressed lower level of TET2 but HT-29 43- cells expressed higher levels compared to HT-29 cells (Figure 7A). At the protein level, western Blot results in similar conclusion (Figure 7C, D). This data was further supported by Immunofluorescence microscopy using TET2 antibody (Figure8).
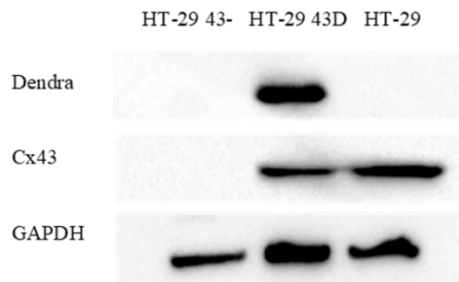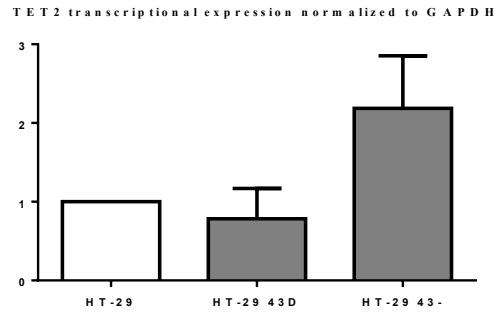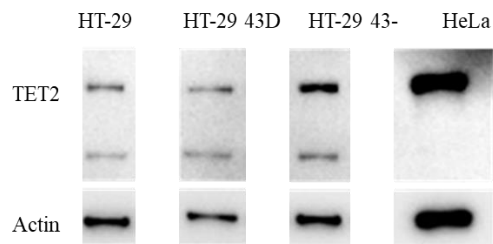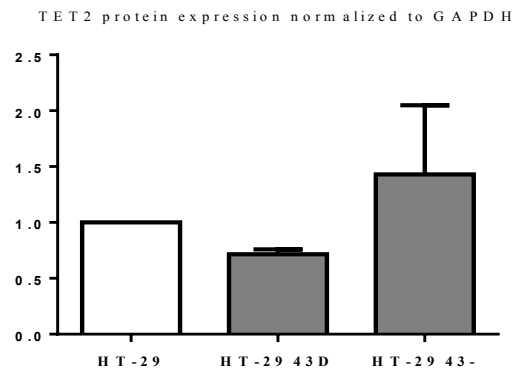
A

B

TET2 transcriptional expression normalized to GAPDH

C

D

TET2 protein expression normalized to GAPDH

Figure 7: TET2 expression as a function of Cx43 regulation in IECs. (A) Western Blot verifying the down-regulation of Cx43 in HT-29 43- cells and expression of Cx43- Dendra in Ht-29 43D cells. (B) Histogram representing the normalized expression of TET2 to GAPDH assessed by qPCR. (B) Representative western Blot of TET2 expression in IECs; HeLa as a positive control of TET2. (C) Densitometry analysis of western Blot of TET2 expression normalized to GAPDH in IECs.

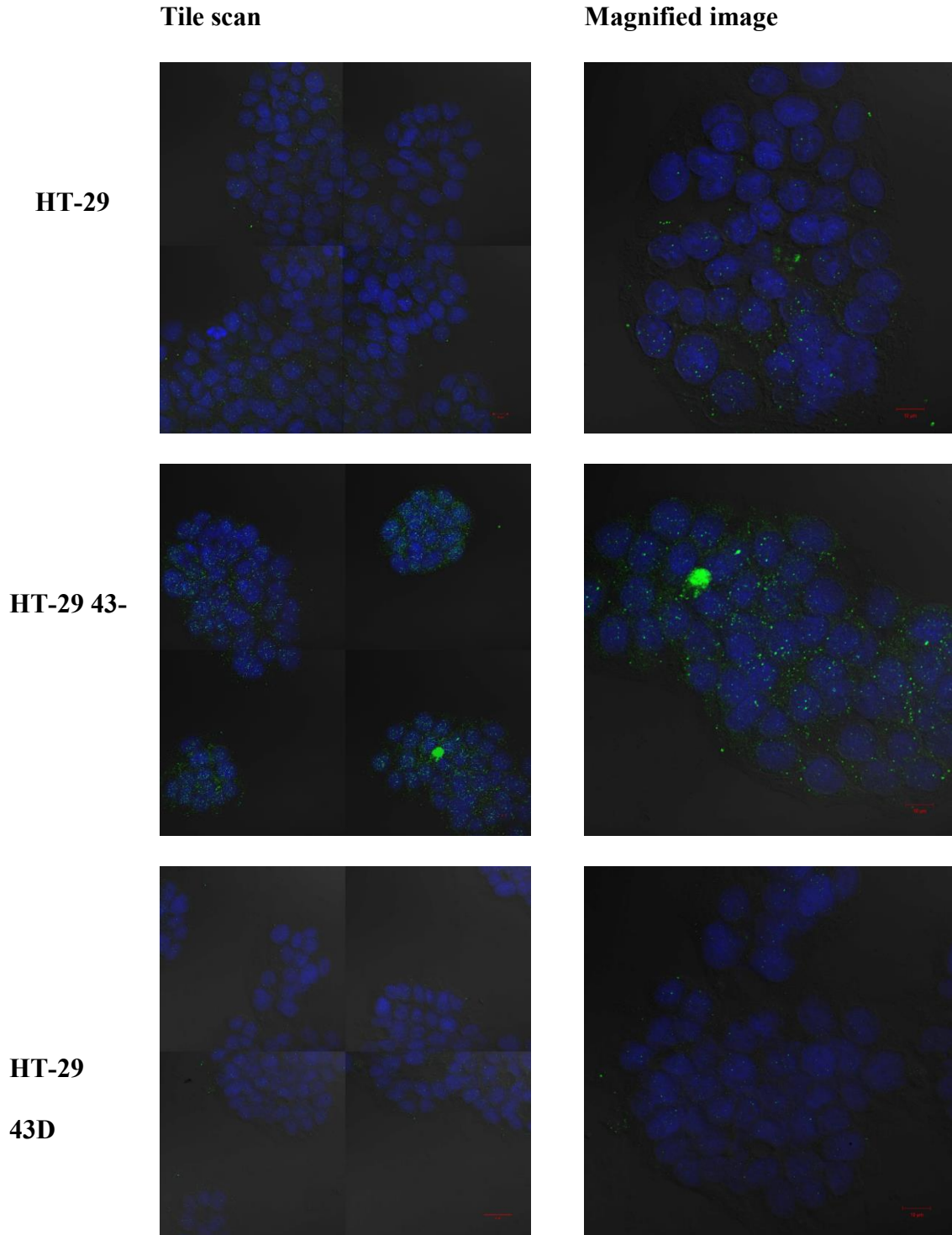**Tile scan**   **Magnified image**
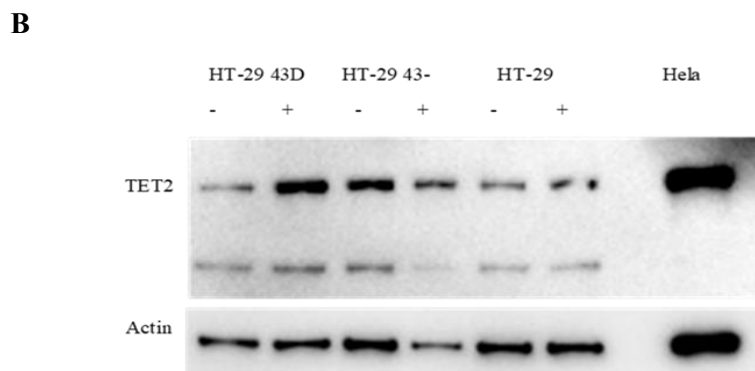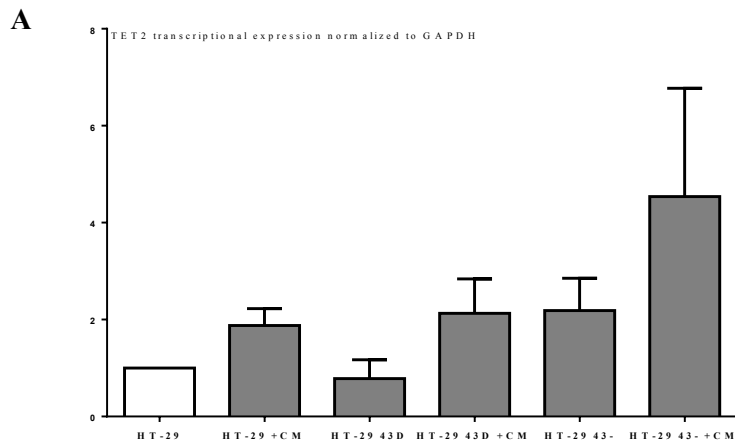

HT-29

HT-29 43-

HT-29

43D

Figure 8: Cellular expression of TET2 by immunofluorescence on HT-29, HT-29 43-, and HT-2943D. Tile scan and magnified images are represented. Images captured at a 63x/1.46 Oil Plan-Apochromatic objective. DAPI in blue, TET2 in green.

2. Regulation of TET2 expression by treatment with conditioned media from THP-1 supernatant

HT-29, HT29 43D, and HT-29 43- were treated with conditioned media from THP-1 activated cells for 24 hours. The expression levels of TET2 was also assessed at both transcriptional and protein levels. A significant increase in TET2 expression was detected in all treated cases at both transcriptional and protein levels (Figure9).

A

B

C

TET2 protein expression normalized to GAPDH

(x-axis labels: HT-29, HT-29 + CM, HT-29 43D, HT-29 43D + CM, HT-29 43-, HT-29 43- + CM)
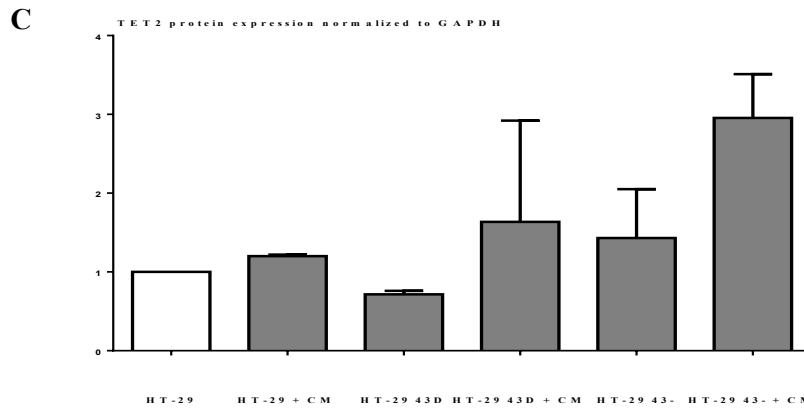
Figure 9: The effect of treatment with conditioned media for 24 hours on TET2 expression in IECs. (A) Histogram representing the normalized expression of TET2 to GAPDH assessed by qPCR. (B) Representative western Blot of TET2 expression in IECs. (C) Densitometry analysis of western Blot of TET2 expression normalized to GAPDH in IECs.

In addition, we performed an immunofluorescence assay to study the effect of inflammatory mediators on the expression of TET2 in IECs. Briefly, HT-29, HT-29 43D, and HT-29 43- cells were either treated or untreated with conditioned media from activated THP-1 cells for 24 hours. The cells were then fixed and immune-stained by TET2. TET2 expression also recorded an increase upon the treatment with Conditioned media (Figure 10).
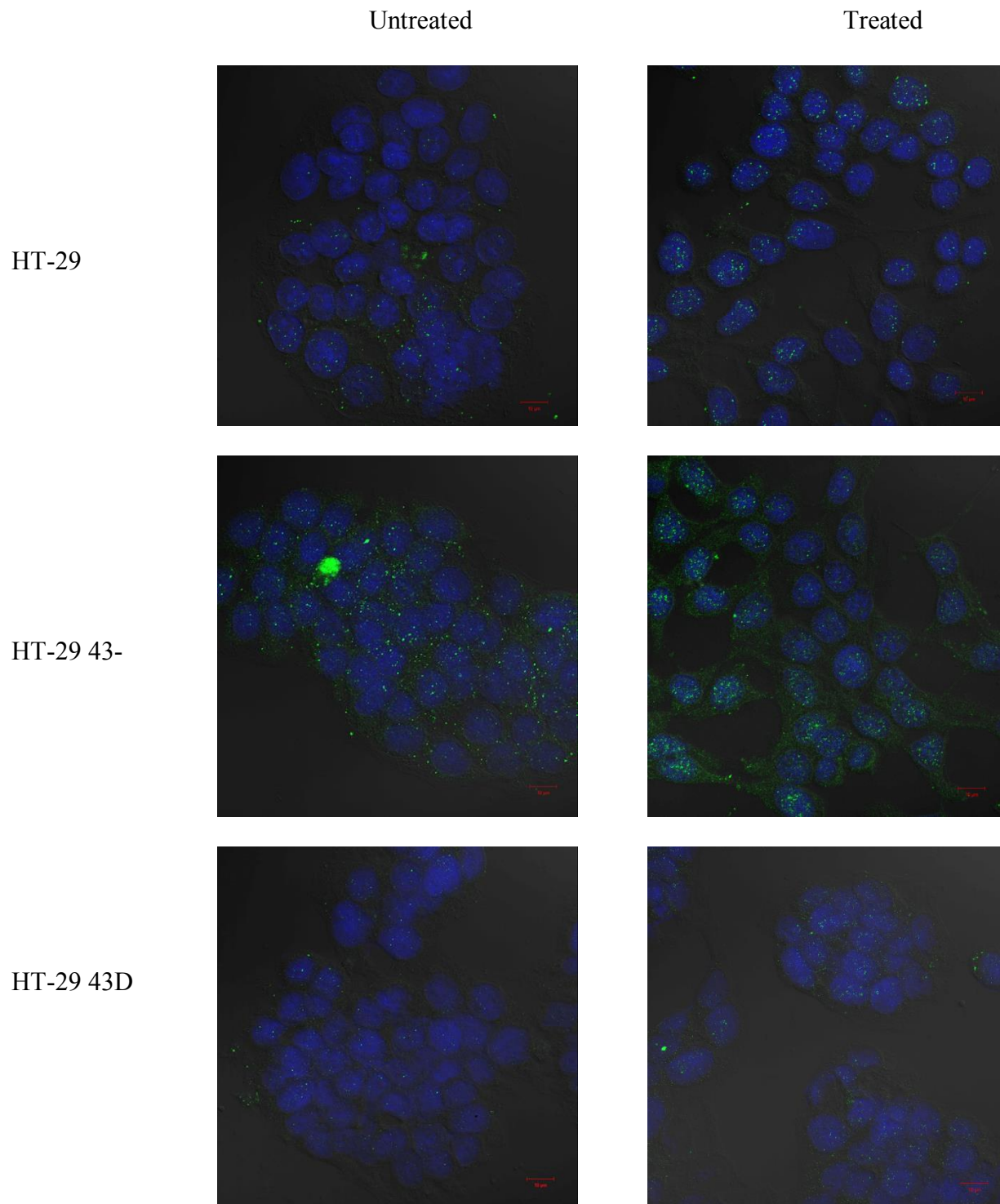
Figure 10: Cellular expression and expression of TET2 in HT29, HT29 43D, and HT-29 43- cells with and without treatment with conditioned media from activated THP-1 cells. Images captured at a 63x/1.46 Oil Plan-Apochromatic objective. DAPI in blue, TET2 in green.

3. The variation of 5-hmC epigenetic modification mimics TET2 expression

5-hmC is generated from 5-mC by TET2 protein, thus 5-hmC levels represent the functionality of TET2. We have shown that TET2 expression increases upon Cx43 down regulation and decreases when Cx43 is up-regulated. In addition, the treatment of IECs with inflammatory mediators increases TET2 expression.

To further document the activation of TET2, 5-hmC levels were determined by Immunofluorescence and Slot Blot technique. This data shows that 5-hmC increased in down-regulated Cx43 cells and increased upon upregulation of Cx43 (Figure11, 12). Also, 5-hmC levels increased when cells where treated with inflammatory mediators for 24 hours (Figure12).
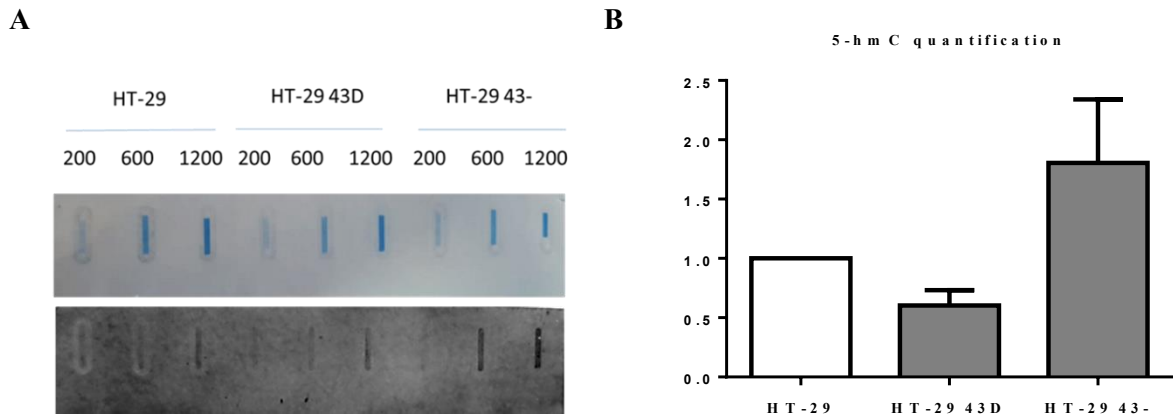
Figure 11: 5-hmC expression in IECs as an effect of Cx43 regulation and treatment. (A) Slot Blot analysis of 5-hmC expression in DNA extracted from IECs and the methylene blue corresponding to the concentrations of DNA triplicates. (B) Analysis of 5-hmC dot Blot.
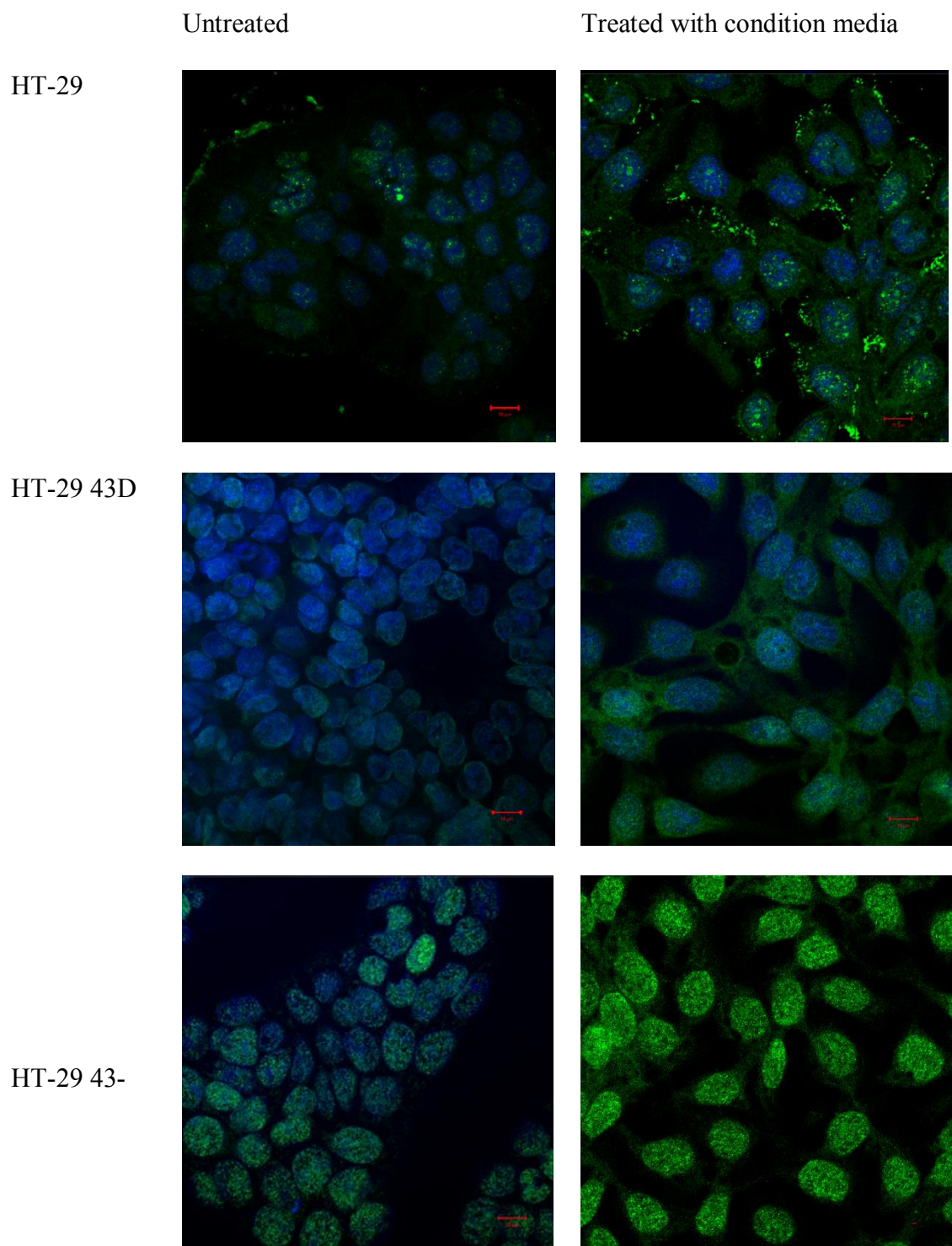
Figure 12: 5-hmC expression in IECs as an effect of Cx43 regulation and treatment by Immunofluorescence. Images captured at a 63x/1.46 Oil Plan-Apochromatic objective. DAPI in blue, 5-hmC in green.

4. Expression of CD68 in IBD and Cancer human colon tissues

As suspected, CD68 showed to increase significantly in IBD and cancer colon tissues with respect to the Normal tissues (Figure13). It can be inferred that the inflammation level in IBD, in correspondence to CD68 expression, is higher than that recorded in cancer colon tissue.

	Tile scan	**Magnified image**
Normal		
IBD		
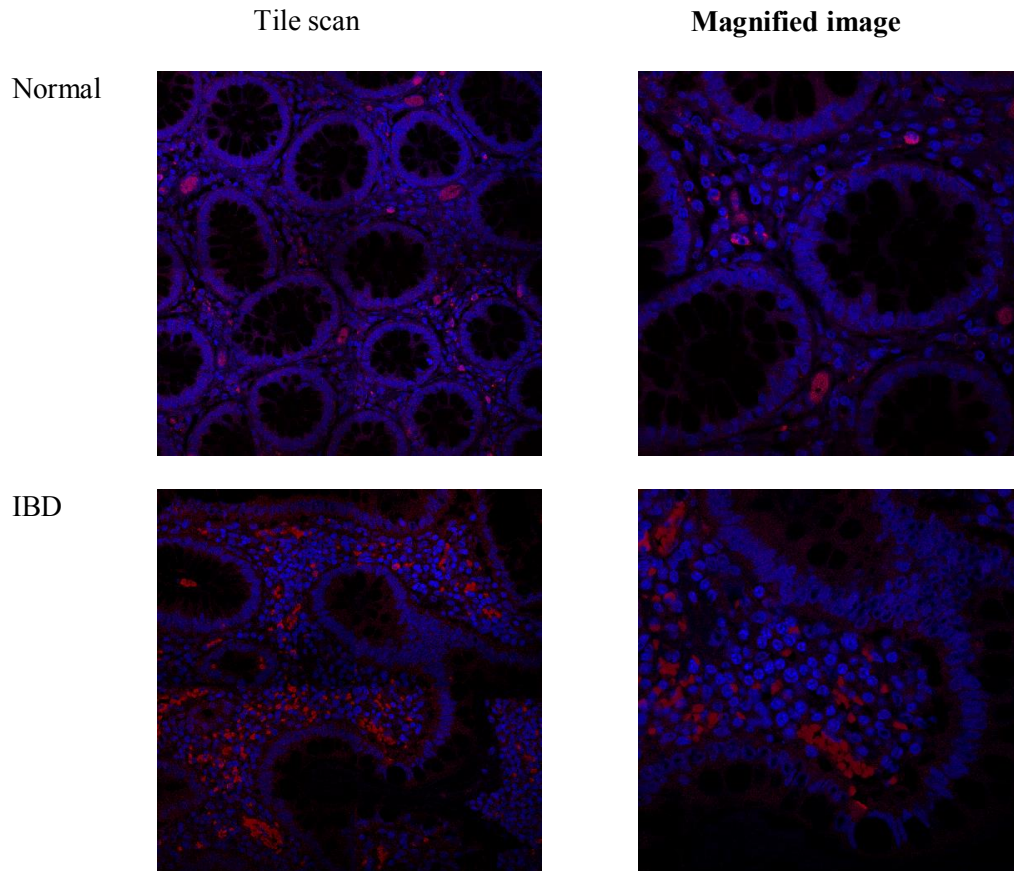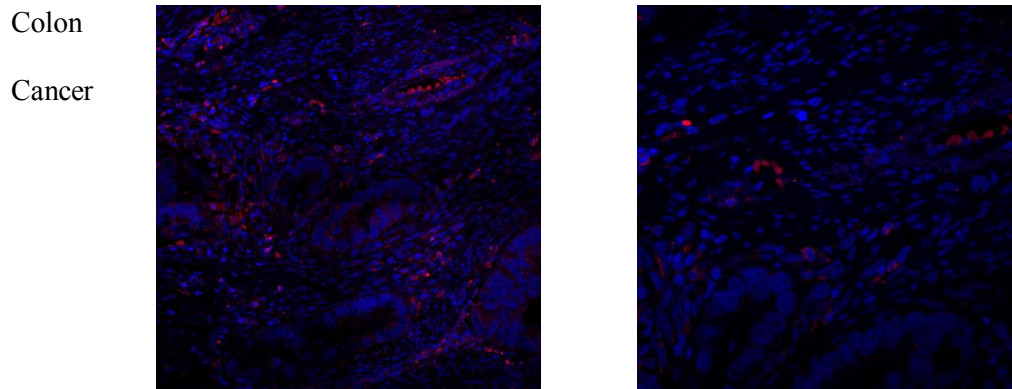
Colon

Cancer

Figure 13: CD68 expression in Normal, IBD and Cancer tissues. Both tile scan and a magnified image is represented. Images captured at a 63x/1.46 Oil Plan-Apochromatic objective. DAPI in blue, CD68 in red.

5. Altered Cx43 expression in IBD and colon cancer

Further, we assessed Cx43 expression by immunofluorescence. Cx43 was observed on the membrane of adjacent cells of the crypt in normal colon tissue (Figure 14A, B). On the other hand, Cx43 expression increases in IBD tissue, and their localization is re-distributed to the apical surface of crypt adjacent cells (Figure15). In cancer colon tissues, the expression of Cx43 is generally low and it showed to be altered with no definite criteria with respect to normal tissues (Figure15).
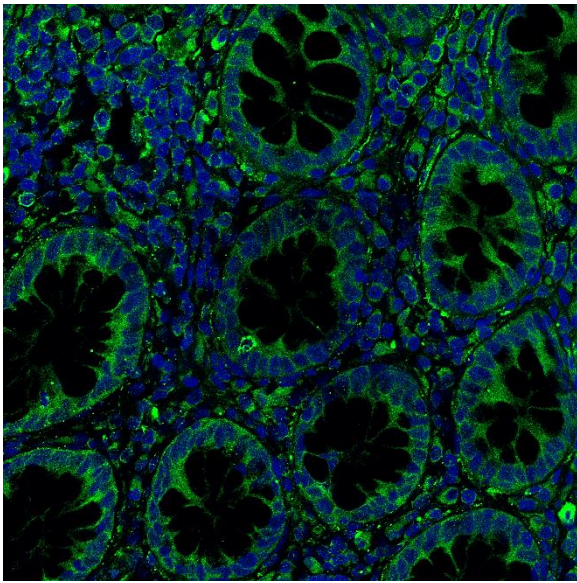
A   B
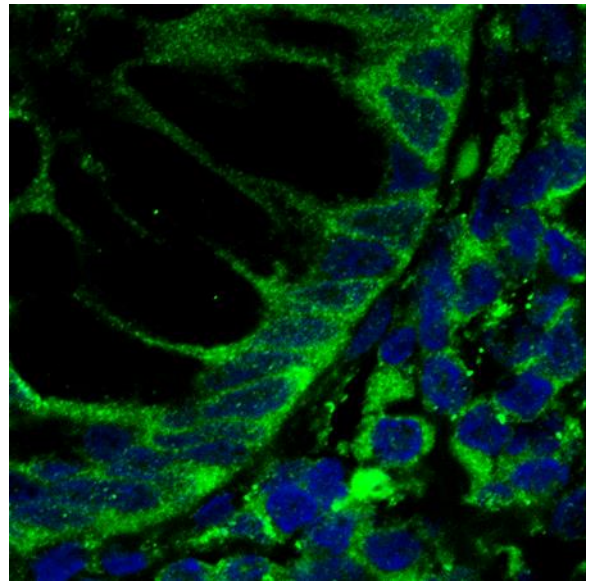
Figure 14: Cx43 expression in Normal colon crypt. (A) Shows a tile scan, (B) shows a magnified image of Cx43 expression. Images captured at a 63x/1.46 Oil Plan-Apochromatic objective. DAPI in blue, Cx43 in green.

	Tile scan	Magnified image

Normal colon tissue

IBD colon tissue

Cancer colon tissue
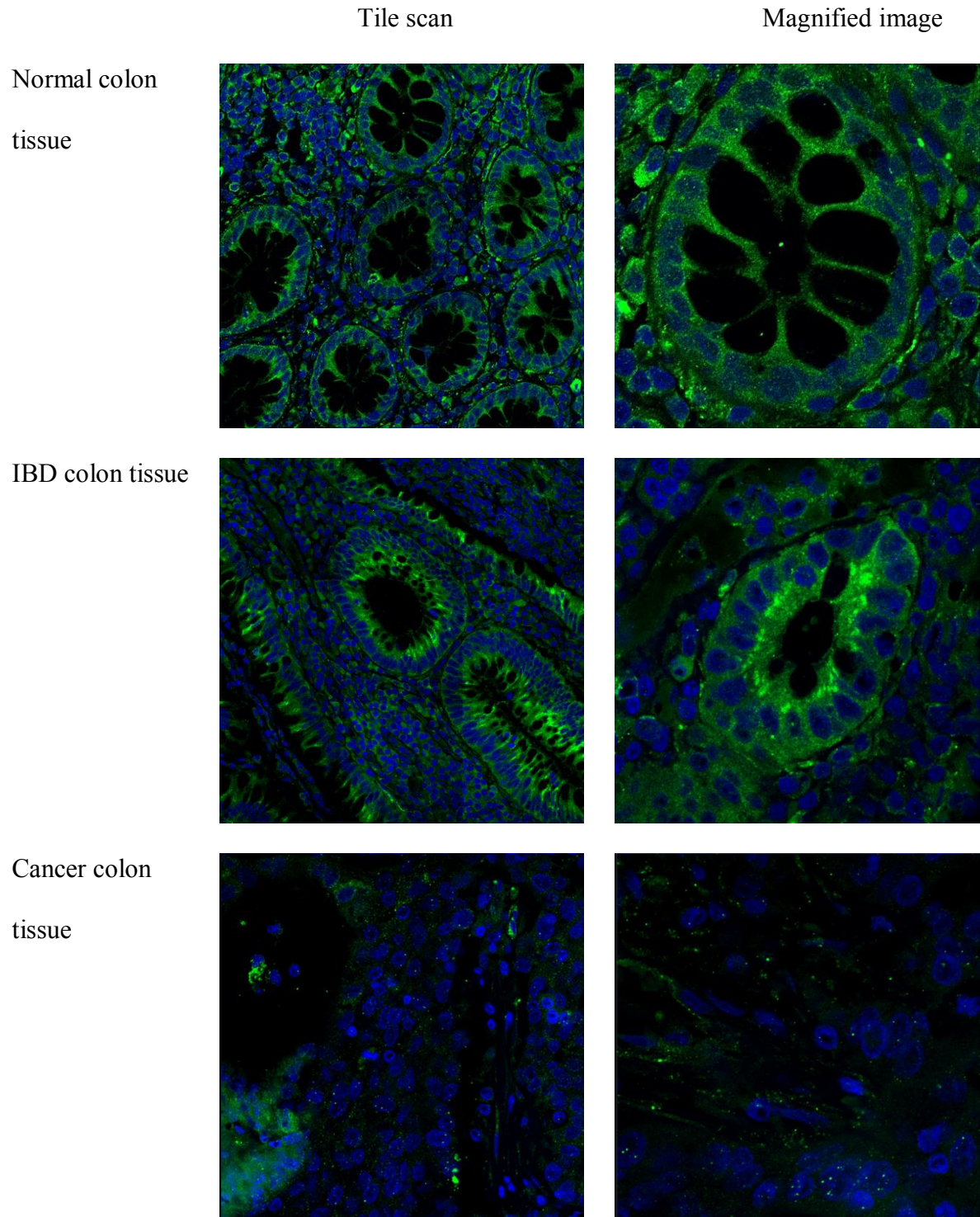
Figure 15: Cx43 expression in Normal, IBD, and colon cancer human tissues. Images captured at a 63x/1.46 Oil Plan-Apochromatic objective. DAPI in blue, Cx43 in green.

6. TET2 expression and localization in Normal human colon tissues

In order to investigate the expression of TET2 in normal colon tissues, immunofluorescence on paraffin embedded normal colon tissue sections was done. It reveals that TET2 is expressed in normal colon tissues and it is localized to both nucleus and cytoplasm. This expression is represented as green dots in the cytoplasm and the nucleus of crypt cells (Figure16). In attempt to know the difference between cytoplasmic and nuclear TET2 expression, we counted the dots corresponding to TET2 presence. For 70-76 nuclear dots a maximum of 16 cytoplasmic dots were found (Figure 16A). Which means that the nuclear expression of TET2 dominates cytosolic expression in normal colon tissue.

A   B

Figure 16: TET2 expression and localization in Normal colon tissue. (A) 2D projection image of a Z-stack of normal colon crypts and a video representing the Z-stack. (B) Two different z-stack positions of the same image showing the localization of TET2 at both cytosolic and nuclear positions. Images captured at a 63x/1.46 Oil Plan-Apochromatic objective. DAPI in blue, TET2 in green.

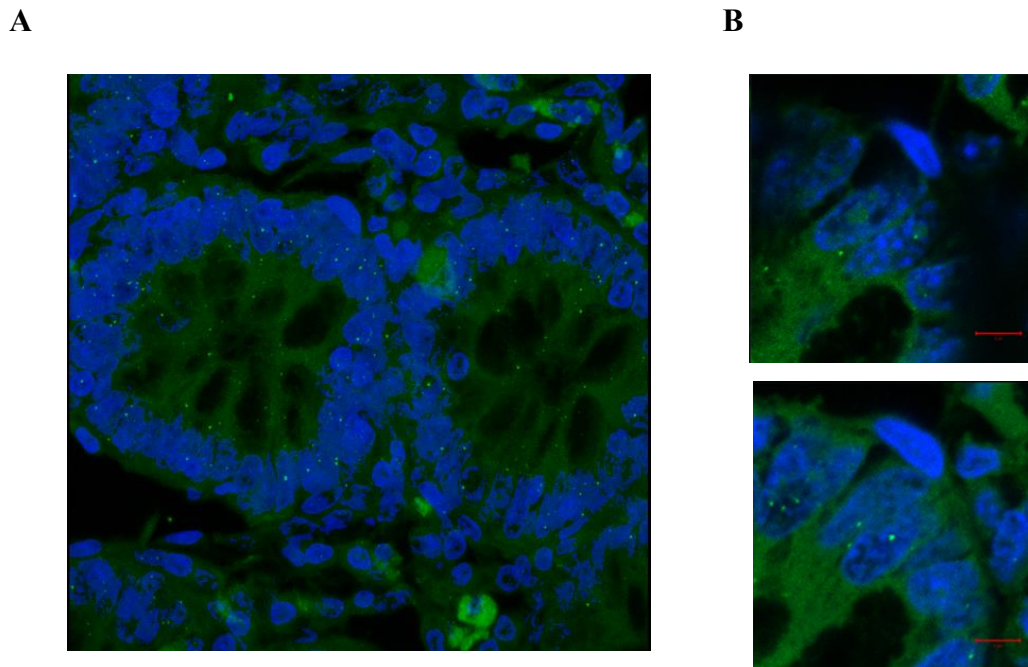## 7. TET2 expression diminished in IBD and Colon cancer tissues.

In the attempt to detect TET2 profile in IBD and colon cancer, immunofluorescence was also done on IBD and colon cancer tissues. The screening of TET2 in IBD and colon cancer showed a decrease compared to normal tissues, and it was found that the expression, if present, is localized to the nucleus (Figure16). It is important to mention that in Normal colon tissues TET2 is expressed in all crypt cells, but in IBD many cells do not even express TET2. Thus TET2 is highly down regulated in IBD and cancer colon tissue.

This data was further supported by checking TET2 functionality; 5-hmC. This was done by immunofluorescence on human colon tissues which enables us to detect 5-hmC expression in cells. 5-hmC data mimics that of TET2 data where the level of 5-hmC decreases in IBD and colon cancer (Figure17)

Tile scan TET2 expression   2D projection image TET2
                                    expression

Figure 17: TET2 cellular expression in Normal, IBD and colon cancer human tissue. Arrows shows some obvious mitotic cells in cancer tissue. Images captured at a 63x/1.46 Oil Plan-Apochromatic objective. DAPI in blue, TET2 in green.

5-hmC expression

Normal colon tissue

IBD colon tissue

Cancer colon tissue
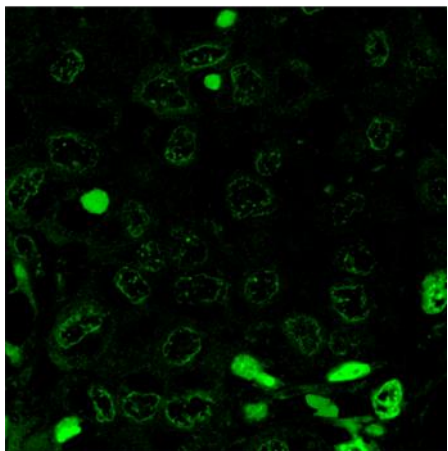
DISCUSSION AND CONCLUSION

Epigenetic modifications, including DNA methylation, had been studied widely to determine the role of epigenetics in IBD and its transformation to colorectal cancer. Research have tried to find also the correlation between TET family proteins and diseases including IBD and cancer, but no clear data was published till now. In our study, we chose to focus on the variation of TET2 expression in IBD and cancer.
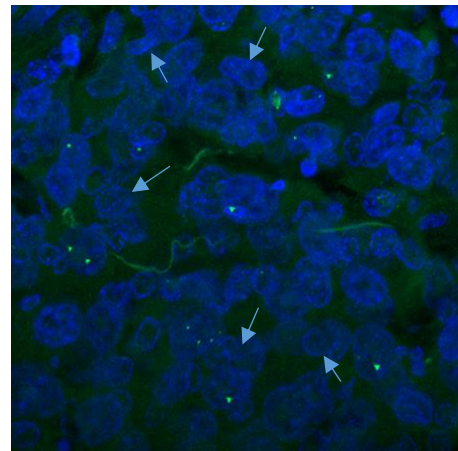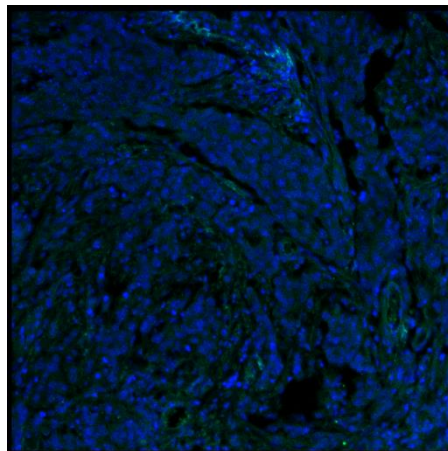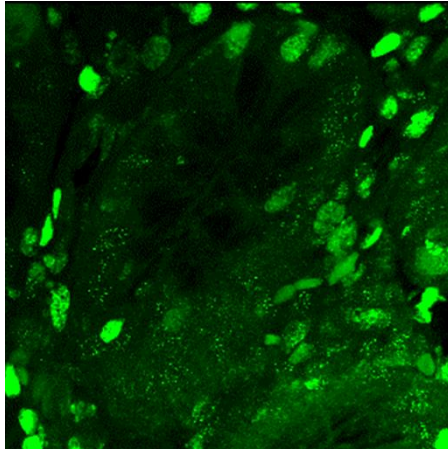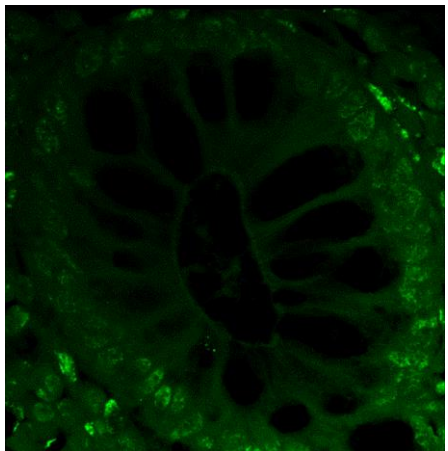
First, in an *in-vivo* inflammatory culture model we compared the expression of TET2 between HT-29, HT-29 43D, and HT-29 43- cell lines. Both transcriptional and protein expression data showed that TET2 expression increases upon the downregulation of Cx43 but TET2 expression decreases upon the over-expression of Cx43. In other experiment, IECs were cultured with activated THP-1 supernatant (condition media) to induce inflammation to IECs. The results show that TET2 expression increases upon inflammation. It is important to mention that 5-hmC data (Dot Blot and IF) mimics TET2 data. As mentioned earlier, research had stated a severely reduced or absent Cx43 expression in IBD and colorectal cancer. [86] [87] Also, it is well known that as Cx43 expression increases cell lines become more epithelial like cells were as cells lose their Cx43 they become more cancer like cells. This ensures that Cx43 down-regulation must mimic an inflammatory culture model which was translated in TET2 expression data. Thus in the *in-vitro* inflammatory culture model, TET2 expression showed to increase upon an induced inflammation.

Second, Immunofluorescence was done on archived human colon tissues. First, CD68, as an inflammation marker, showed that inflammation level is higher in IBD than in colon cancer

tissue. This was also the result of Cx43 Immunofluorescence data. Cx43 expression showed to increase in IBD but become very low in colon cancer tissue. This data ensures the relation between inflammation and Cx43 expression in colon tissue. At the level of TET2, TET2 protein expression showed a decrease in IBD and colon cancer tissue with respect to normal colon tissue. This states that TET2 protein is downregulated in IBD and colon cancer.

REFERENCES

1.   Jan, M., et al., *Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia.* Sci Transl Med, 2012. **4**(149): p. 149ra118.

2.   Franke, A., et al., *Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci.* Nat Genet, 2010. **42**(12): p. 1118-25.

3.   Nan, X., et al., *Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex.* Nature, 1998. **393**(6683): p. 386-9.

4.   Filion, G.J., et al., *A family of human zinc finger proteins that bind methylated DNA and repress transcription.* Mol Cell Biol, 2006. **26**(1): p. 169-81.

5.   Muto, M., et al., *Targeted disruption of Np95 gene renders murine embryonic stem cells hypersensitive to DNA damaging agents and DNA replication blocks.* J Biol Chem, 2002. **277**(37): p. 34549-55.

6.   Brenet, F., et al., *DNA methylation of the first exon is tightly linked to transcriptional silencing.* PLoS One, 2011. **6**(1): p. e14524.

7.   Hellman, A. and A. Chess, *Gene body-specific methylation on the active X chromosome.* Science, 2007. **315**(5815): p. 1141-3.

8.   Aran, D., et al., *Replication timing-related and gene body-specific methylation of active human genes.* Hum Mol Genet, 2011. **20**(4): p. 670-80.

9.   Gardiner-Garden, M. and M. Frommer, *CpG islands in vertebrate genomes.* J Mol Biol, 1987. **196**(2): p. 261-82.

10. Schulz, W.A., C. Steinhoff, and A.R. Florl, *Methylation of endogenous human retroelements in health and disease.* Curr Top Microbiol Immunol, 2006. **310**: p. 211-50.

11. Mayer, W., et al., *Demethylation of the zygotic paternal genome.* Nature, 2000. **403**(6769): p. 501-502.

12. Oswald, J., et al., *Active demethylation of the paternal genome in the mouse zygote.* Curr Biol, 2000. **10**(8): p. 475-8.

13. Paroush, Z., et al., *Dynamics of demethylation and activation of the alpha-actin gene in myoblasts.* Cell, 1990. **63**(6): p. 1229-37.

14. He, Y.-F., et al., *Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA.* Science, 2011. **333**(6047): p. 1303-1307.

15. Wu, S.C. and Y. Zhang, *Active DNA demethylation: many roads lead to Rome.* Nat Rev Mol Cell Biol, 2010. **11**(9): p. 607-20.

16. Ono, R., et al., *LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23).* Cancer Res, 2002. **62**(14): p. 4075-80.

17. Bachman, M., et al., *5-Formylcytosine can be a stable DNA modification in mammals.* Nat Chem Biol, 2015. **11**(8): p. 555-7.

18. Fagarasan, S., et al., *Critical roles of activation-induced cytidine deaminase in the homeostasis of gut flora.* Science (New York, N.Y.), 2002. **298**(5597): p. 1424-1427.

19.   Tahiliani, M., et al., *Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1.* Science, 2009. **324**(5929): p. 930-5.

20.   Hu, L., et al., *Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation.* Cell, 2013. **155**(7): p. 1545-55.

21.   Ko, M., et al., *Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX.* Nature, 2013. **497**(7447): p. 122-126.

22.   Rasmussen, K.D. and K. Helin, *Role of TET enzymes in DNA methylation, development, and cancer.* Genes & development, 2016. **30**(7): p. 733-750.

23.   Zhang, H., et al., *TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine.* Cell Research, 2010. **20**(12): p. 1390-1393.

24.   Guilhamon, P., et al., *Meta-analysis of IDH-mutant cancers identifies EBF1 as an interaction partner for TET2.* Nature communications, 2013. **4**: p. 2166.

25.   Ko, M., et al., *Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX.* Nature, 2013. **497**(7447): p. 122-6.

26.   Zhou, T., et al., *Structural basis for hydroxymethylcytosine recognition by the SRA domain of UHRF2.* Mol Cell, 2014. **54**(5): p. 879-86.

27.   He, Y.F., et al., *Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA.* Science, 2011. **333**(6047): p. 1303-7.

28.   Lindahl, T. and R.D. Wood, *Quality control by DNA repair.* Science, 1999. **286**(5446): p. 1897-905.

29.   Maiti, A. and A.C. Drohat, *Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites.* J Biol Chem, 2011. **286**(41): p. 35334-8.

30.   Spruijt, C.G., et al., *Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives.* Cell, 2013. **152**(5): p. 1146-59.

31.   Delhommeau, F., et al., *Mutation in TET2 in myeloid cancers.* N Engl J Med, 2009. **360**(22): p. 2289-301.

32.   Bejar, R., et al., *Clinical effect of point mutations in myelodysplastic syndromes.* N Engl J Med, 2011. **364**(26): p. 2496-506.

33.   Chan, S.M. and R. Majeti, *Role of DNMT3A, TET2, and IDH1/2 mutations in pre-leukemic stem cells in acute myeloid leukemia.* Int J Hematol, 2013. **98**(6): p. 648-57.

34.   Itzykson, R., et al., *Clonal architecture of chronic myelomonocytic leukemias.* Blood, 2013. **121**(12): p. 2186-98.

35.   Lian, C.G., et al., *Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma.* Cell, 2012. **150**(6): p. 1135-46.

36.   Turcan, S., et al., *IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype.* Nature, 2012. **483**(7390): p. 479-83.

37.   Liu, C., et al., *Decrease of 5-hydroxymethylcytosine is associated with progression of hepatocellular carcinoma through downregulation of TET1.* PLoS One, 2013. **8**(5): p. e62828.

38.   Knappskog, S., et al., *RINF (CXXC5) is overexpressed in solid tumors and is an unfavorable prognostic factor in breast cancer.* Ann Oncol, 2011. **22**(10): p. 2208-15.

39.   Chen, Q., et al., *TET2 promotes histone O-GlcNAcylation during gene transcription.* Nature, 2013. **493**(7433): p. 561-4.

40.   Fu, X., et al., *MicroRNA-26a targets ten eleven translocation enzymes and is regulated during pancreatic cell differentiation.* Proc Natl Acad Sci U S A, 2013. **110**(44): p. 17892-7.

41.   Morita, S., et al., *miR-29 represses the activities of DNA methyltransferases and DNA demethylases.* Int J Mol Sci, 2013. **14**(7): p. 14647-58.

42.   Oermann, E.K., et al., *Alterations of metabolic genes and metabolites in cancer.* Semin Cell Dev Biol, 2012. **23**(4): p. 370-80.

43.   Huang, Y. and A. Rao, *Connections between TET proteins and aberrant DNA modification in cancer.* Trends Genet, 2014. **30**(10): p. 464-74.

44.   Ko, M., et al., *TET proteins and 5-methylcytosine oxidation in hematological cancers.* Immunol Rev, 2015. **263**(1): p. 6-21.

45.   Szwagierczak, A., et al., *Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA.* Nucleic Acids Res, 2010. **38**(19): p. e181.

46.   Clavel, T. and D. Haller, *Molecular interactions between bacteria, the epithelium, and the mucosal immune system in the intestinal tract: implications for chronic inflammation.* Current issues in intestinal microbiology, 2007. **8**(2): p. 25-43.

47.   Clayburgh, D.R., L. Shen, and J.R. Turner, *A porous defense: the leaky epithelial barrier in intestinal disease.* Laboratory investigation; a journal of technical methods and pathology, 2004. **84**(3): p. 282-291.

48.   Fukata, M. and M. Arditi, *The role of pattern recognition receptors in intestinal inflammation.* Mucosal Immunol, 2013. **6**(3): p. 451-63.

49.   Corridoni, D., et al., *Probiotic bacteria regulate intestinal epithelial permeability in experimental ileitis by a TNF-dependent mechanism.* PloS one, 2012. **7**(7): p. e42067-e42067.

50.   Hummel, S., et al., *Differential targeting of the E-Cadherin/beta-Catenin complex by gram-positive probiotic lactobacilli improves epithelial barrier function.* Appl Environ Microbiol, 2012. **78**(4): p. 1140-7.

51.   Smith, P.M., et al., *The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis.* Science, 2013. **341**(6145): p. 569-73.

52.   Matsuo, K., et al., *Histochemistry of the surface mucous gel layer of the human colon.* Gut, 1997. **40**(6): p. 782-789.

53.   Kim, Y.S. and S.B. Ho, *Intestinal goblet cells and mucins in health and disease: recent insights and progress.* Current gastroenterology reports, 2010. **12**(5): p. 319-330.

54.   Capaldo, C.T., D.N. Powell, and D. Kalman, *Layered defense: how mucus and tight junctions seal the intestinal barrier.* Journal of molecular medicine (Berlin, Germany), 2017. **95**(9): p. 927-934.

55.   Ho, S., C. Pothoulakis, and H.W. Koon, *Antimicrobial peptides and colitis.* Current pharmaceutical design, 2013. **19**(1): p. 40-47.

56.   Peterson, D.A., et al., *IgA response to symbiotic bacteria as a mediator of gut homeostasis.* Cell host & microbe, 2007. **2**(5): p. 328-339.

57.   de Souza, H.S. and C. Fiocchi, *Immunopathogenesis of IBD: current state of the art.* Nat Rev Gastroenterol Hepatol, 2016. **13**(1): p. 13-27.

58.   Ullman, T.A. and S.H. Itzkowitz, *Intestinal inflammation and cancer.* Gastroenterology, 2011. **140**(6): p. 1807-16.

59. Baumgart, D.C. and W.J. Sandborn, *Inflammatory bowel disease: clinical aspects and established and evolving therapies.* Lancet, 2007. **369**(9573): p. 1641-57.

60. Gryboski, J.D., *Ulcerative colitis in children 10 years old or younger.* J Pediatr Gastroenterol Nutr, 1993. **17**(1): p. 24-31.

61. Mehdizadeh, S., et al., *Diagnostic yield of capsule endoscopy in ulcerative colitis and inflammatory bowel disease of unclassified type (IBDU).* Endoscopy, 2008. **40**(1): p. 30-5.

62. de Lange, K.M., et al., *Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease.* Nature Genetics, 2017. **49**(2): p. 256-261.

63. Hugot, J.P., et al., *Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease.* Nature, 2001. **411**(6837): p. 599-603.

64. Knights, D., K.G. Lassen, and R.J. Xavier, *Advances in inflammatory bowel disease pathogenesis: linking host genetics and the microbiome.* Gut, 2013. **62**(10): p. 1505-10.

65. Moran, C.J., *Very early onset inflammatory bowel disease.* Semin Pediatr Surg, 2017. **26**(6): p. 356-359.

66. Azarschab, P., et al., *Epigenetic control of the E-cadherin gene (CDH1) by CpG methylation in colectomy samples of patients with ulcerative colitis.* Genes Chromosomes Cancer, 2002. **35**(2): p. 121-6.

67. Wheeler, J.M., et al., *Hypermethylation of the promoter region of the E-cadherin gene (CDH1) in sporadic and ulcerative colitis associated colorectal cancer.* Gut, 2001. **48**(3): p. 367-71.

68. Nimmo, E.R., et al., *Genome-wide methylation profiling in Crohn's disease identifies altered epigenetic regulation of key host defense mechanisms including the Th17 pathway.* Inflamm Bowel Dis, 2012. **18**(5): p. 889-99.

69. Vogelstein, B., et al., *Cancer genome landscapes.* Science, 2013. **339**(6127): p. 1546-58.

70. Francescone, R., V. Hou, and S.I. Grivennikov, *Cytokines, IBD, and colitis-associated cancer.* Inflamm Bowel Dis, 2015. **21**(2): p. 409-18.

71. Axelrad, J.E., S. Lichtiger, and V. Yajnik, *Inflammatory bowel disease and cancer: The role of inflammation, immunosuppression, and cancer treatment.* World J Gastroenterol, 2016. **22**(20): p. 4794-801.

72. Beyer, E.C., D.L. Paul, and D.A. Goodenough, *Connexin43: a protein from rat heart homologous to a gap junction protein from liver.* J Cell Biol, 1987. **105**(6 Pt 1): p. 2621-9.

73. Cruciani, V. and S.O. Mikalsen, *The vertebrate connexin family.* Cell Mol Life Sci, 2006. **63**(10): p. 1125-40.

74. Abascal, F. and R. Zardoya, *Evolutionary analyses of gap junction protein families.* Biochim Biophys Acta, 2013. **1828**(1): p. 4-14.

75. Bosco, D., J.A. Haefliger, and P. Meda, *Connexins: key mediators of endocrine function.* Physiol Rev, 2011. **91**(4): p. 1393-445.

76. Garcia, I.E., et al., *Connexinopathies: a structural and functional glimpse.* BMC Cell Biol, 2016. **17 Suppl 1**: p. 17.

77. Nielsen, M.S., et al., *Gap junctions.* Comprehensive Physiology, 2012. **2**(3): p. 1981-2035.

78.   Severs, N.J., et al., *Remodelling of gap junctions and connexin expression in diseased myocardium.* Cardiovasc Res, 2008. **80**(1): p. 9-19.

79.   Garfield, R.E., et al., *Possible role of gap junctions in activation of myometrium during parturition.* Am J Physiol, 1978. **235**(5): p. C168-79.

80.   Piechocki, M.P., R.D. Burk, and R.J. Ruch, *Regulation of connexin32 and connexin43 gene expression by DNA methylation in rat liver cells.* Carcinogenesis, 1999. **20**(3): p. 401-6.

81.   King, T.J., et al., *Correlation between growth control, neoplastic potential and endogenous connexin43 expression in HeLa cell lines: implications for tumor progression.* Carcinogenesis, 2000. **21**(2): p. 311-5.

82.   Yi, Z.C., et al., *Downregulation of connexin 43 in nasopharyngeal carcinoma cells is related to promoter methylation.* Oral Oncol, 2007. **43**(9): p. 898-904.

83.   Borinstein, S.C., et al., *Aberrant DNA methylation occurs in colon neoplasms arising in the azoxymethane colon cancer model.* Mol Carcinog, 2010. **49**(1): p. 94-103.

84.   Locke, D., et al., *Post-translational modifications of connexin26 revealed by mass spectrometry.* Biochem J, 2009. **424**(3): p. 385-98.

85.   Me, et al., *Gap Junctions: Basic Structure and Function.* Journal of Investigative Dermatology, 2007. **127**: p. 2516-2524.

86.   Nemeth, L., S. Maddur, and P. Puri, *Immunolocalization of the gap junction protein Connexin43 in the interstitial cells of Cajal in the normal and Hirschsprung's disease bowel.* J Pediatr Surg, 2000. **35**(6): p. 823-8.

87.   Sirnes, S., et al., *Connexins in colorectal cancer pathogenesis.* Int J Cancer, 2015. **137**(1): p. 1-11.