MTG16 (CBFA2T3) represses E protein-dependent transcription to regulate colonic secretory cell differentiation, epithelial regeneration, and tumorigenesis

Rachel E. Brown[1,2], Justin Jacobse[3,4,5], Shruti A. Anant[6], Koral M. Blunt[5], Bob Chen[7,8], Paige N. Vega[8,9], Chase T. Jones[5], Jennifer M. Pilat[1], Frank Revetta[3], Aidan H. Gorby[5], Kristy R. Stengel[10], Yash A. Choksi[1,5,11,12], Kimmo Palin[13,14,15], M. Blanca Piazuelo[5], M. Kay Washington[3], Ken S. Lau[8,9,16], Jeremy A. Goettel[1,3,5,12], Scott W. Hiebert[10,16], Sarah P. Short[1,5,12], and Christopher S. Williams[1,2,5,11,12]

[1] Program in Cancer Biology, Vanderbilt University School of Medicine, Nashville, Tennessee, 37232, USA
[2] Medical Scientist Training Program, Vanderbilt University School of Medicine, Nashville, Tennessee, 37232, USA
[3] Department of Pathology, Microbiology, and Immunology, Vanderbilt University, Nashville, TN, USA
[4] Willem-Alexander Children's Hospital, Leiden University Medical Center, Leiden, The Netherlands
[5] Department of Medicine, Division of Gastroenterology, Hepatology, and Nutrition, Vanderbilt University Medical Center, Nashville, Tennessee, 37232, USA
[6] Vanderbilt University, Nashville, Tennessee, 37232, USA
[7] Program in Chemical and Physical Biology, Vanderbilt University School of Medicine, Nashville, Tennessee, 37232, USA
[8] Epithelial Biology Center, Vanderbilt University Medical Center, Nashville, Tennessee, 37232, USA
[9] Department of Cell and Developmental Biology, Vanderbilt University School of Medicine, Nashville, Tennessee, 37232, USA
[10] Department of Biochemistry, Vanderbilt University School of Medicine, Nashville, TN 37232, USA
[11] Veterans Affairs Tennessee Valley Health Care System, Nashville, Tennessee, 37232, USA
[12] Center for Mucosal Inflammation and Cancer, Vanderbilt University Medical Center, Nashville, Tennessee, 37232, USA
[13] Department of Medical and Clinical Genetics, University of Helsinki, Helsinki, Finland
[14] Applied Tumor Genomics Research Program, Research Programs Unit, University of Helsinki, Helsinki, Finland
[15] iCAN Digital Precision Cancer Medicine Flagship, University of Helsinki, Helsinki, Finland
[16] Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, Tennessee, 37232, USA

Grant Support

This work was supported by the National Institutes of Health (F30DK120149 to REB, F31DK127687 to PNV, F31CA232272 to JMP, R01DK103831 to KSL, U01CA215798 to KSL, R03DK123489 to JAG, R01CA178030 to SWH, K01DK123495 to SPS, F32DK108492 to SPS, R01DK099204 to CSW, P30DK058404 to the Vanderbilt Digestive Disease Research Center [DDRC], P50CA236733 to the Vanderbilt-Ingram Cancer Center [VICC] Spore in Gastrointestinal Cancer, and T32GM00734 to the Vanderbilt Medical Scientist Training Program), the U.S. Department of Veterans Affairs Office of Medical Research (1I01BX001426 to CSW and IK2BX004648 to YAC), the Crohn's and Colitis Foundation (623541 to CSW and 662877 to SPS), the Royal Netherlands Academy of Arts and Sciences (Academy Ter Meulen Grant to JJ), and the Prince Bernhard Cultural Foundation (Cultural Foundation Grant to JJ). KP was supported by grants from Academy of Finland (Finnish Center of Excellence Program 2018–2025, 312041), the iCAN Digital Precision Cancer Medicine Flagship (320185), and the Sigrid Jusélius Foundation.

YAC was also supported by the P30DK058404 Vanderbilt DDRC Pilot and Feasibility grant. The Vanderbilt Translational Pathology Shared Resource (TPSR) was supported by NCI/NIH Cancer Center Support Grant 2P30CA068485-14 and the Vanderbilt Mouse Metabolic Phenotyping Center Grant 5U24DK059637-13.

Correspondence

Christopher S. Williams; Mailing: 1075 Light Hall/MRBIV, 2215-B Garland Ave., Nashville, TN 37232; Tel: 615-322-3642; Fax: 615-343-6229; Email: christopher.s.williams@vumc.org

Disclosures

The authors declare no competing interests.

Author Contributions

REB designed experiments, performed experiments, and analyzed data. SAA, KMK, CTJ, JMP, and AHG performed experiments and analyzed data. JJ and REB designed and performed bulk RNA-sequencing (RNA-seq) data analysis, respectively. KSL, BC, and PNV designed, collected, and analyzed human and murine colon single cell RNA-seq. FR validated antibodies for and performed immunohistochemical staining. MBP and MKW performed histological analyses and provided expertise in gastrointestinal pathology. JJ, RB, and KP analyzed RNA-seq from human tumors. CSW, SPS, SWH, KSL, JAG, KRS, and YAC provided expertise regarding experimental models and experimental design. REB wrote the manuscript. All authors edited and approved the manuscript.

Acknowledgments

We thank Barbara Fingleton, Stephen Brandt, Matthew Stier, and other members of the Williams and Hiebert Labs for their critical insights into this manuscript. We would also like to thank the TPSR, Digital Histology Shared Resource (DHSR), Vanderbilt Technologies for Advanced Genomics (VANTAGE), and Advanced Computing Center for Research and Education (ACCRE) for aid with histology, imaging, RNA-seq, and data storage/processing, respectively. Additional computational resources were provided by the ELIXIR node, hosted at the CSC–IT Center for Science, Finland.

Abstract

Aberrant epithelial differentiation and regeneration pathways contribute to colon pathologies including inflammatory bowel disease (IBD) and colitis-associated cancer (CAC). MTG16 (also known as CBFA2T3) is a transcriptional corepressor expressed in the colonic epithelium. MTG16 interaction partners include E box-binding basic helix-loop-helix transcription factors (E proteins). MTG16-deficient mice exhibit worse colitis and increased tumor burden in inflammatory carcinogenesis. In this study, we sought to understand the role of MTG16 colonic epithelial homeostasis and the mechanisms by which MTG16 protects the epithelium in colitis and CAC. We demonstrated that MTG16 deficiency enabled enteroendocrine cell differentiation from secretory precursor cells at the expense of goblet cells. Transcriptomic analysis implicated dysregulated E protein function in MTG16-deficient colon crypts. Using a novel mouse model with a point mutation that disrupts MTG16:E protein complex formation (*Mtg16*^{P209T}), we established that enteroendocrine:goblet cell balance was dependent on MTG16:E protein interactions and that the shift in lineage allocation was associated with enhanced expression of *Neurog3*, the central driver of enteroendocrine lineage specification. Furthermore, *Mtg16* was upregulated in the previously described *Ascl2*$^{+}$, de-differentiating cells that replenish the stem cell compartment in response to colon injury. *Mtg16* expression was also increased in dextran sulfate sodium (DSS)-treated mouse colon crypts and in IBD patients compared to unaffected controls. We determined that the effects of MTG16 in regeneration are also dependent on its repression of E proteins, as the colonic epithelium failed to regenerate following DSS-induced injury in our novel mutant mouse model. Finally, we revealed that uncoupling MTG16:E protein interactions contributes to the enhanced tumorigenicity in *Mtg16*$^{-/-}$ colon in the azoxymethane(AOM)/DSS-induced model of CAC. Collectively, our results demonstrate that MTG16, via its repression of E protein targets, is a key regulator of cell fate decisions during colonic differentiation and regeneration.

Introduction

The colonic epithelium is a complex, self-renewing tissue comprised of specialized cell types with diverse functions, including: stem cells, which self-renew and generate all epithelial lineages; colonocytes, which absorb water and electrolytes; goblet cells, which secrete mucus to protect from luminal insults; and enteroendocrine cells, which secrete hormones[1–3]. The complexity and diverse functions of the secretory lineage, including goblet and enteroendocrine cells, has only recently been appreciated[4–13]. Secretory lineage dysregulation now been implicated in inflammatory bowel disease (IBD), colorectal cancer (CRC), obesity, diabetes, and the gut-brain axis[9,14–20]. Furthermore, recent studies have demonstrated that colonic epithelial regeneration in response to injury occurs through the opposite process: de-differentiation of committed precursor cells[21–23]. Aberrant regenerative programs may lead to uncontrolled cell proliferation and dysplasia[24]. Thus, these processes must be tightly controlled by coordinated networks of transcriptional activators and repressors[3].

Myeloid translocation gene 16 (MTG16, also known as CBFA2T3) is a transcriptional corepressor crucial for the regulation of stem cell dynamics, differentiation, and cancer[25–29]. MTG16 bridges transcription factors and epigenetic regulators such as histone deacetylases to repress transcription of target loci[27,30,31]. While MTG16 may be expressed in a given cell type, its repression targets vary depending on the expression gradients of other components of the repression complex[25]. Indeed, MTG16 may even be displaced resulting in an activation complex and gene transcription[32]. MTG16 has previously been shown to repress stem cell genes and pan-secretory genes in the homeostatic small intestine (SI)[33]. Furthermore, our group has determined the effects of global MTG16 deficiency in colonic phenotypes, including increased injury in dextran sulfate sodium (DSS)-induced colitis[34] and increased tumor burden in the azoxymethane (AOM)/DSS model of inflammatory carcinogenesis[35,36]. However, the precise mechanisms by which MTG16 controls colonic homeostasis, regeneration, and tumorigenesis remain unknown.

One family of transcription factors repressed by MTG16 are basic helix-loop-helix (bHLH) transcription factors[25,28,37–39]. Ubiquitously expressed (class I) and tissue-specific (class II) bHLH transcription factors homo- and heterodimerize to control stem cell dynamics, lineage allocation, and differentiation[40–42]. The class I bHLH transcription factors, also known as E proteins because they bind to consensus Ephrussi box (E box) DNA sequences, are E12 and E47 (splice variants of *E2A*), HEB, and E2-2[28,38,40]. However, the functions of E proteins have mainly been described outside of the colon[40,42–49].

In this study, we define the topology of *Mtg16* expression in the colon and discover novel functions of MTG16 in the colonic epithelium. We identify marked upregulation of E protein transcriptional signatures in the setting of MTG16 deficiency and use a novel point mutant mouse model to determine that uncoupling MTG16:E protein interactions drives these phenotypes. We investigate this regulatory relationship in homeostatic secretory lineage allocation and test its functional impact on epithelial regeneration in murine models of IBD, where *Mtg16* expression is increased, and colitis-associated carcinoma (CAC). Analogously, *MTG16* expression is upregulated in biopsies from patients with active IBD but returns closer to baseline with disease remission. Overall, we demonstrate novel, context-specific roles for MTG16 in the colon dependent on its interaction with E proteins.

Materials and Methods

Mouse models and animal care. *Mtg16*$^{-/-}$ mice were previously generated and validated[26]. *Mtg16*^{P209T} mutant mice were CRISPR-generated as recently described[38]. Littermate-matched wild type (WT) and *Mtg16*$^{-/-}$ and littermate-matched WT and homozygous *Mtg16*^{P209T} mutant (*Mtg16*$^{T/T}$) mice were bred using heterozygous x heterozygous breeding schemes and housed in the same facility throughout the duration of the experiments. Mice were provided standard rodent chow (#5L0D, LabDiet) *ad libitum* and kept on a 12:12 light-dark cycle. All animal experiments were carried out in accordance with protocols approved by the Vanderbilt Institutional Animal Care and Use Committee (IACUC). Approximately equal numbers of 8-12-wk-old male and female mice were used for all experiments.

Single-cell RNA-sequencing (scRNA-seq) and analysis. InDrop scRNA-seq and analysis were performed on 2 independent (discovery and validation) cohorts of human colon samples as previously described[24]. These data are publicly available through the Human Tumor Atlas Network (https://data.humantumoratlas.org/)[50]. scRNA-seq of colon crypts isolated from WT mice was performed as previously described[51–53]. These data (#GSE145830, #GSE114044) are available in the Gene Expression Omnibus (GEO)[54,55].

RNAscope *in situ* hybridization. RNAscope[56], high-resolution RNA *in situ* hybridization, was performed using the RNAscope Multiplex Fluorescent V2 Assay (#323100, ACDBio) according to the manufacturer's protocol. Slides were boiled in Target Retrieval Reagents for 15 min and treated with Protease IV for 30 min. Probes were specific for Mm-*Cbfa2t3* (#43601-C2, ACDBio), Mm-*Muc2* (#315451, ACDBio), and Mm-*Neurog3* (#422401, ACDBio). 1:750 TSA Cy3 (#NEL704A001KT, PerkinElmer) and TSA Cy5 (#NEL75A001KT, Perkin Elmer) were used for probe visualization.

Chromogenic and immunofluorescent immunohistochemistry (IHC). SI and colon were dissected, Swiss-rolled, and fixed in 10% neutral-buffered formalin (NBF) for 24 h at RT. Formalin-fixed tissue was paraffin-embedded (FFPE) by the Vanderbilt Translational Pathology Shared Resource (TPSR). For baseline WT and *Mtg16*$^{-/-}$ analyses, the TPSR constructed 2 tissue micro-arrays (TMAs) using the TMA Grand Master automated arrayer (PerkinElmer). Briefly, 2-mm cores were punched from individual FFPE blocks containing WT and *Mtg16*$^{-/-}$ SI and colon (4 cores per Swiss roll). Following construction, H&E slides were cut from the TMAs and evaluated for quality. The TPSR also performed all hematoxylin and eosin (H&E) and periodic acid-Schiff (PAS) staining.

For chromogenic and immunofluorescent IHC, 5-µm FFPE sections were deparaffinized, rehydrated in graded ethanol concentrations, and permeabilized using TBS-T (Tris-buffered saline with 0.05% Tween-20) followed by antibody-specific antigen retrieval (**Table S1**). Next, for chromogenic IHC, endogenous peroxidases were quenched in 0.03% H_2O_2 with sodium azide for 5 min. Slides were then incubated in primary antibody (**Table S1**) O/N at RT. Detection was performed by 30-min. incubation with Dako Envision+ System HRP-Labeled Polymers (#K4003, Agilent) followed by a 5-min incubation with DAB. For immunofluorescent staining, slides were blocked in 5% normal goat serum (#01-620-1, ThermoFisher), incubated in primary antibody (**Table S1**) O/N at 4 °C, washed in TBS-T, and incubated in the appropriate secondary antibodies (# A-11011, #A-11077, #A-21131, Invitrogen**)** for 2 h at RT. 5-Ethynyl-2'-deoxyuridine (EdU) was visualized using the Click-iT EdU Cell Proliferation Kit for Imaging (#C10037, ThermoFisher). Nuclear counterstaining was performed using ProLong Gold Antifade Mountant with DAPI (#P36931, ThermoFisher).

Imaging and quantification. Slides were imaged using a Nikon Eclipse E800 microscope with NIS-Elements Basic Research Software or by the VUMC Digital Histology Shared Resource (DHSR). For chromogenic IHC, whole slides were imaged by the DHSR using a Leica SCN400

Slide Scanner (Leica Biosystems). For fluorescent staining, whole slides were imaged by the DHSR using an Aperio Versa 200 automated slide scanner (Leica Biosystems) and visualized using Aperio ImageScope (v12.4.3) (Leica Biosystems). Quantification was performed by an observer blinded to genotype by counting positive cells in well-oriented crypts or crypt-villus units.

Colon crypt isolation. Murine colon was harvested, splayed longitudinally, and divided into proximal colon (defined as the first 3 cm from the cecum) and distal colon (defined as the first 3 cm from the anus). Colon segments were washed in phosphate-buffered saline (PBS) without calcium or magnesium and minced. Tissue fragments were incubated in chelation buffer (2 mM EDTA in PBS without calcium or magnesium) with nutation for 90 min at 4 °C and washed with PBS followed by 2-min cycles of gentle shaking in shaking buffer (43.3 mM sucrose and 54.9 mM sorbitol in PBS without calcium or magnesium) to free colon crypts. Intact crypts were collected by centrifugation at 150 *x g*, 4 °C for 12 min for RNA isolation and bulk RNA-sequencing (RNA-seq) as described below.

Murine colitis models. The DSS colitis model of epithelial injury and regeneration was performed as previously described[57]. Briefly, littermate-matched WT and mutant mice were co-housed from the time of weaning when possible. At 8-12 wk old, bedding was mixed to normalize microbiome 2 wk prior to experiments. Lixits were capped to accustom mice to drinking from water bottles 1 wk prior to experiments. Water bottles were replaced with 2% (w/v) DSS (#DS1004, Gojira Fine Chemicals) for 5 days and then replaced with water for 4 days of recovery prior to sacrifice. Mice were weighed daily. For histologic analysis, colons were dissected, Swiss-rolled, fixed in 10% NBF, and processed by the Vanderbilt TPSR as described above. H&Es were scored for histologic injury and regeneration by a gastrointestinal pathologist blinded to genotype (MKW) as described in **Table S2**. Distal colon crypts were collected as described above for RNA isolation

and RNA-seq as described below. For EdU experiments, mice were intraperitoneally (*i.p.*) injected with 50 mg/kg EdU (#NE08701, Biosynth Carbosynth) dissolved in PBS 1 h prior to sacrifice.

Murine inflammatory carcinogenesis model. Bedding was mixed, and 8-12-wk old mice were co-housed and accustomed to drinking from water bottles as described above. Mice were intraperitoneally injected with 10 mg/kg AOM and exposed to three 4-day cycles of 3% DSS (#DB001, TdB Labs). Each DSS cycle was followed by a 16-day recovery period. Mice were weighed daily throughout the experiment. At sacrifice, colons were dissected, flushed with PBS, splayed longitudinally, and imaged using a Nikon SMZ1270 dissecting scope. Tumors were identified by vascular, non-dimpled appearance from these images in conjunction with direct visualization. Tumors were measured macroscopically in 2 perpendicular dimensions using calipers, and tumor volume was calculated using the previously validated formula $W^2*L/2$, where W was the shorter caliper measurement, and L was the longer caliper measurement[58]. Representative tumors from each colon were dissected, incubated in RNA*later* (#R0901, Sigma-Aldrich) at 4 °C for 24 h, and stored at -80 °C. Histologic injury was assessed by MKW (**Table S2**). Tumor number and degree of dysplasia was evaluated as previously described[59] by a gastrointestinal pathologist blinded to genotype (MBP).

RNA isolation, bulk RNA-seq, and analysis. Murine colon crypts were collected as described above followed by resuspension in TRIzol Reagent (#15596018, Invitrogen) and homogenization by 5x shearing through a 25G needle. Murine tumors were mechanically homogenized in TRIzol. Samples were centrifuged for 10 min at 16,200 x *g*, 4 °C to remove insoluble material. Cleared homogenate was chloroform-extracted. RNA was isolated from the aqueous layer using the RNeasy Mini Kit (#74106, Qiagen) with on-column DNase treatment (#79254, Qiagen). Library preparation and paired-end 150 bp RNA-seq using the Illumina NovaSeq6000 were performed by Vanderbilt Technologies for Advanced Genomics (VANTAGE). Raw reads in FASTQ format

were trimmed with fastp (v0.20.0) with default parameters[60]. Quantification was performed using Salmon (v1.3.0)[61] against a decoy transcriptome generated from *Mus musculus* GENCODE (v21)[62]. Further analysis was done in R (v3.6.3) as described previously[63]. Briefly, quantification files were imported with tximeta (v1.4.5)[64]. Genes with counts ≤ 1 were omitted. Differential expression analysis (DEA) was performed on raw transcript counts using DESeq2 (v1.26.0)[65] and annotated with AnnotationDbi (v1.46.1)[66]. Gene Set Enrichment Analysis (GSEA)[67] was performed on DESeq2-normalized transcript counts using the GSEA software (v4.1.0) with 1,000 gene permutations and gene set randomization. Gene sets were found in the Molecular Signatures Database (MSigDB) (v7.2)[68] or derived from the literature (**Table S3**). Figures were generated using GSEA (v4.1.0), the R packages pheatmap (v1.0.12), RColorBrewer (v1.1-2), and ggplot2 (v3.2.1), Meta-Chart (https://www.meta-chart.com/venn#/display), and Inkscape (v1.0.1).

***In silico* analysis of chromatin immunoprecipitation-sequencing (ChIP-seq) data, IBD datasets, and TCGA data.** Raw and processed ChIP-seq data from Baulies *et al.*[33] (#GSE124186) and Murata *et al.*[21] (#GSE130822) were retrieved from the GEO[54,55]. Data were visualized and 5'-CATNNG (E box) sequences were annotated using the Broad Institute Integrated Genomics Viewer (v2.9.1)[69].

Metadata for rectal biopsies from treatment-naïve pediatric UC patients (n = 206) and unaffected controls (n = 20) in the Predicting Response to Standardized Pediatric Colitis Therapy (PROTECT) study[70,71] (#GSE109142) were also retrieved from the GEO[54,55]. SRR files were downloaded using the NCBI SRA Toolkit and the Advanced Computing Center for Research and Education (ACCRE) and processed as described above for bulk murine RNA-seq, with the reads quantified against a decoy-aware transcriptome generated from the GENCODE human reference genome (v29)[62]. DEA was performed on raw counts using DESeq2 (v1.26.0)[65].

Raw counts from RNA-seq of biopsies from adult UC patients (n = 14 with active disease, 14 in remission) and unaffected patients (n = 16)[72] (#GSE128682) were downloaded from the

GEO[54,55]. Pairwise DEA was performed on raw counts using DESeq2 (v1.33.1)[65]. Microarray data from adult IBD patients (n = 5 UC, 20 Crohn's colitis, 8 unaffected controls)[73] (#GSE9686) were analyzed using limma in Geo2R[55].

Raw feature counts generated from The Cancer Genome Atlas (TCGA) PanCancer Atlas[74,75] by Rahman *et al*.[76] (#GSE62944) were retrieved from the GEO[54,55]. DEA (n = 41 normal colon samples, 482 colon adenocarcinomas) was performed in R (v4.1.0) using DESeq2 (v1.33.1)[65].

Bulk RNA-seq of human CAC and CRC samples. RNA-seq of human samples collected from 9 regional hospitals in Finland was performed and analyzed as described previously[77]. Briefly, RNA was TRIzol-extracted from human tumor samples (n = 47 CRC, 17 CAC) and underwent HiSeq LncRNA-Seq library preparation and paired-end sequencing using Illumina HiSeqXTen. Raw sequences were mapped onto the human transcriptome (Ensembl release 79) using Salmon (v0.12.0)[61]. Gene-level quantification with variance stabilization was performed using DESeq2 (v1.18.1)[65] followed by limma (v3.34.9) correction of sequencing batch effects[78].

Statistical methods. Statistical analyses of human CRC vs. CAC data were performed using limma (v3.34.9)[78]. Statistical analyses of all other bulk RNA-seq were performed using DESeq2 (v1.26.0, v1.33.1)[65]. Statistical analyses of microarray data were performed using limma in Geo2R[55]. In these studies, $p_{adj} < 0.05$ was considered significant. Statistical analyses for GSEA were performed using the GSEA program (v4.1.0). A false discovery rate (FDR) q-value < 0.25 was considered significant. All other statistical analyses for all other data were performed in GraphPad Prism (v9.0.1) as indicated in each figure legend. Data are displayed as arithmetic mean ± SEM unless otherwise noted, and $p < 0.05$ was considered significant.

Data availability. All code is available upon reasonable request.

Results

***MTG16* is predominantly expressed in secretory cells in human and murine colon.** To understand the topography of *MTG16* expression in the human colon, we queried single-cell RNA-seq (scRNA-seq) datasets generated from 2 independent cohorts of normal human colon samples[24]. Interestingly, we observed that *MTG16* was predominantly expressed in secretory cells (**Fig. 1A, S1**). To understand if this expression pattern was evolutionarily conserved, we queried our inDrop scRNA-seq generated from WT murine colon crypts[52,53] and similarly found that *Mtg16* was expressed in the goblet and enteroendocrine cell clusters (**Fig. 1B**). Confirming these findings, RNAscope for *Mtg16* and *Muc2* (the predominant mucin expressed by colonic goblet cells[11]) mRNA demonstrated co-localization in the murine colonic epithelium (**Fig. 1C**). Collectively, these data suggested a role for MTG16 in the colonic secretory lineage.

MTG16-deficient mice have skewed secretory lineage allocation in the colon, distinct from the SI. Transcriptional corepressors, including MTG16, regulate differentiation and cell-fate decisions[25–28,37–39,79,80]. Prior studies have established that *Mtg16*$^{-/-}$ mice have fewer goblet cells in the SI but retain normal numbers of enteroendocrine cells[33,81]. In agreement with these studies, *Mtg16*$^{-/-}$ mice had fewer goblet cells, but no increase in SYP^{+} enteroendocrine cells, in the SI (**Fig. S2A-B**). Tuft cells, which had not previously been investigated, were decreased per crypt-villus unit in the SI (**Fig. S2C**). We next assessed secretory lineage allocation in the colon at baseline in WT and *Mtg16*$^{-/-}$ mice. Similar to the SI, *Mtg16*$^{-/-}$ mice had fewer goblet cells per colon crypt (**Fig. 1D**). However, immunohistochemistry (IHC) for synaptophysin (SYP) demonstrated more enteroendocrine cells per crypt in *Mtg16*$^{-/-}$ colon (**Fig. 1E**). Tuft cells were decreased in the colons of *Mtg16*$^{-/-}$ mice (**Fig. 1F**). Together, these data implicated a colon-specific role for MTG16 in orchestrating cell type allocation within the secretory lineage.

Loss of MTG16 distorts the transcriptomic balance of enteroendocrine and goblet cell lineages. Prior studies have demonstrated subtle differences in secretory cell differentiation along the length of the colon[6,11,82–84]. Thus, we stratified our human scRNA-seq datasets by proximal and distal colon and examined *MTG16* expression. *MTG16* was consistently expressed in goblet cells, but its expression pattern in other secretory cell types varied between proximal and distal colon and between each cohort (**Fig. S3**). We next queried the *Tabula Muris*, a publicly available murine colon scRNA-seq dataset that specifies whether certain clusters originate from the proximal or distal colon[85], for *Mtg16* expression (**Fig. S4A-B**). In the proximal colon, *Mtg16* was enriched in all goblet cell clusters. In the distal colon, *Mtg16* was enriched in *Lgr5*$^{+}$ amplifying, undifferentiated cells and most goblet cells, except those located at the top of the crypt (**Fig. S4B**). Bulk RNA-seq of proximal and distal WT colon revealed higher expression of *Mtg16* in the proximal colon, potentially reflecting the higher density of goblet cells in the proximal colon (**Fig. S4C**). We confirmed this *Mtg16* expression pattern by RNAscope (**Fig. S5**).

Due to these differences, we performed bulk RNA-seq of colon epithelial isolates from the proximal and distal colon separately. In both colon segments, goblet-associated genes were downregulated, and enteroendocrine-associated genes were upregulated (**Fig. 1G-H**). However, goblet cell-associated gene downregulation was more pronounced in the proximal colon (**Fig. 1G**), while enteroendocrine cell-associated gene upregulation was more striking in the distal colon (**Fig. 1H**). Closer inspection of the distal *Mtg16*$^{-/-}$ colon revealed marked upregulation of key enteroendocrine cell bHLH transcription factors (*Neurog3*, *Neurod1*) and hormone products (e.g., *Chga, Gcg, Sct*) (**Fig. 2A**). Due to this novel phenotype, we chose to further investigate the distal colon RNA-seq by Gene Set Enrichment Analysis (GSEA) using gene sets derived from the literature (**Table S3**). GSEA demonstrated significant enrichment of all stages of enteroendocrine cells during differentiation, except the earliest enteroendocrine progenitor cells in which *Neurog3* expression has not yet been induced[86], in *Mtg16*$^{-/-}$ distal colon crypts (**Fig. 2B**). Concurrently, the goblet cell gene signature was significantly de-enriched in the absence of MTG16 (**Fig. 2B**).

Notably, our RNA-seq did not show significant differences in the stem cell genes *Lgr5* and *Ascl2*, distinct from *Mtg16*$^{-/-}$ SI[33]. GSEA for gene sets representing stem cells and other undifferentiated early progenitor cells showed enrichment of signatures representing *Bmi1*$^{+}$ "label-retaining" stem cells and *mTert*$^{+}$ stem cells, recently shown to be enteroendocrine cells expressing *Chga*[3,8] (**Fig. 2C**). As one of the genes driving these enrichment scores was *Chga*, we used immunofluorescent IHC to stain for CHGA. While CHGA^{+} cells were increased overall, they were not localized to the crypt base (including the +4/+5 cell position) (**Fig. 2D**). Additionally, we did not observe enrichment of any gene sets derived from the literature (**Table S3**) or the MSigDB[68] representing *Lgr5*$^{+}$ crypt base columnar (CBC) cells (**Fig. 2C**, **Fig. S6A**). Altogether, these data indicate that MTG16 exerts control over the balance between goblet and enteroendocrine cells during secretory lineage allocation in the colon.

MTG16 is required for attenuation of E protein activity in the colon. We next sought to illuminate potential mechanisms by which MTG16 promotes the goblet lineage and restricts the enteroendocrine lineage. We and others have previously shown that MTG16 negatively regulates major transcription factor effectors of the WNT and Notch pathways in the hematopoietic system[29,32] and SI[33]. We thus performed GSEA for pathways important in colonic epithelial homeostasis, including WNT and Notch. However, GSEA did not identify global canonical WNT or Notch pathway upregulation in *Mtg16*$^{-/-}$ colon crypts, despite upregulation of the Notch ligand *Dll3* (**Fig. 1H, 2A, S6B**). Additionally, several WNT pathway inhibitors (e.g., *Wif1*, *Cxxc4*) and the Notch pathway inhibitor, *Neurl1a,* were significantly upregulated (**Fig. 2A**). Thus, it appeared unlikely that the effects of MTG16 on colonic differentiation were due to interactions with canonical WNT and Notch effectors.

We then investigated whether target genes of known MTG16-interacting factors, including E proteins[25,28,37–39], were dysregulated in *Mtg16*$^{-/-}$ colon crypts. E proteins coordinate lineage-specific differentiation by interacting with cell type-specific bHLH transcription factors[40].

Furthermore, E proteins such as E2A and HEB are known to be expressed in SI crypts and adenomas[87]. Here, we observed enrichment of all E protein gene signatures, including E2A, its splice variants E12 and E47, and HEB, by GSEA (**Fig. 2E**). Gene sets associated with upregulation of ID1 and ID2, HLH factors that counteract E protein activity and increase stemness in intestinal homeostasis and cancer[40,88–93], were not enriched. These data demonstrated a strong induction of E protein-mediated transcription in the absence of repression by MTG16.

MTG16-driven goblet cell differentiation is dependent on its repression of the E protein transcription factors E2A and HEB. We hypothesized that MTG16 drives goblet cell differentiation from secretory progenitors by attenuating E protein-mediated enteroendocrine cell differentiation. We tested this hypothesis *in vivo* by selectively disrupting the MTG16:E protein interface. To do so, we leveraged a CRISPR-generated mouse with a point mutation resulting in a proline-to-threonine substitution in MTG16 (*Mtg16*^{P209T}) that impairs MTG16:E2A and MTG16:HEB binding and transcriptional repression[38]. We then characterized colonic secretory cell populations in homozygous *Mtg16*^{P209T} mutant (*Mtg16*$^{T/T}$) mice compared to WT littermate-matched controls (**Fig. 3A-C**). Similar to *Mtg16*$^{-/-}$ mice, *Mtg16*$^{T/T}$ mice had fewer PAS^{+} goblet cells and more CHGA^{+} enteroendocrine cells per crypt (**Fig. 3A-B**). However, unlike in the *Mtg16*$^{-/-}$ mouse colon, the number of tuft cells per crypt was unchanged (**Fig. 3C**). We next performed bulk RNA-seq on proximal and distal colon crypt isolates. In both colon segments, goblet cell genes were downregulated, while enteroendocrine cell genes were upregulated, in a pattern resembling that of *Mtg16*$^{-/-}$ colon (**Fig. 3 D-E**). The distal colon exhibited significant upregulation of *Neurog3* (**Fig. 3E**). Additionally, GSEA of the overall enteroendocrine signature, key enteroendocrine progenitor genes, *Bmi1*$^{+}$ and *m-Tert*$^{+}$ reserve stem cell signatures, and E protein signatures all largely phenocopied the *Mtg16*$^{-/-}$ distal colon (**Fig. 3F-H**). As expected, we also observed upregulation of E2A and HEB transcriptional targets (**Fig. 3H**).

We next compared results obtained from *Mtg16*$^{-/-}$ and *Mtg16*$^{T/T}$ distal colon crypts. Here, we observed that 22 of the 30 genes upregulated in *Mtg16*$^{T/T}$ distal colon crypts were also upregulated in *Mtg16*$^{-/-}$ colon (**Fig. 4A-B, Table S4**). Many of these genes (e.g., *Neurog3, Insm1, Insrr, Pex5l, Syt13, Syngap1, Kcnh6, and Agt*) were enteroendocrine-associated genes. A publicly available MTG16 ChIP-seq dataset from SI crypts[33] exhibited significant MTG16 occupancy proximal to 11 of the 22 genes (**Fig. 4A-B, Table S4**), including a large peak at a site enriched with E box motifs upstream of *Neurog3* (**Fig. 4C**), suggesting these genes are repressed by MTG16. Finally, *Neurog3*$^{+}$ enteroendocrine progenitor cells were markedly increased in *Mtg16*$^{-/-}$ and *Mtg16*$^{T/T}$ colon crypts by RNAscope (**Fig. 4D**). Altogether, these data indicate a newly discovered MTG16:E protein-dependent mechanism of colonic secretory lineage allocation, specifically at the stage of NEUROG3-mediated commitment to the enteroendocrine cell lineage.

***MTG16* is upregulated in IBD patients with active disease.** There has been considerable interest in the role of secretory cell lineage misallocation and dysfunction in the pathogenesis of IBD, a condition defined by continuous epithelial injury and regeneration[9,11,19,94–97]. Additionally, several large RNA-seq studies of human UC biopsies have recently been published, including rectal biopsies from treatment-naïve pediatric UC patients compared to unaffected patients in the PROTECT study[70,71] and mucosal biopsies from adult UC patients with active disease, UC patients in remission, and healthy control patients[72]. We found that *MTG16* was upregulated in UC patients in both datasets (**Fig. 5A-B**). *MTG16* expression decreased with remission, although remained elevated compared to unaffected control patients (**Fig. 5B**). Lastly, we analyzed a microarray dataset of adult unaffected, UC, and Crohn's colitis patients and observed significant upregulation of *MTG16* in both UC and Crohn's colitis (**Fig. 5C**). Collectively, these data demonstrate that increased *MTG16* expression is tightly correlated with active injury and regeneration in patients with IBD.

MTG16 contributes to colon crypt regeneration following DSS-induced colitis. As stated previously, MTG16 deficiency has been shown to increase disease severity in DSS-induced colitis[34], a murine model widely used to study IBD, as it recapitulates many histologic features of human UC[98]. To see if this mouse model would recapitulate the increased *MTG16* expression in human IBD, we first assessed *Mtg16* expression following DSS-induced injury and regeneration (**Fig. 5D**). Because DSS treatment largely affects the distal colon[98], we compared distal WT colon crypt epithelial isolates post-DSS-induced injury and regeneration to WT distal colon crypts at baseline. Indeed, *Mtg16* expression was increased in regenerating colon crypts (**Fig. 5E**), similar to our observations in human IBD. Thus, we posited that MTG16 expression may be part of a critical defense and/or repair mechanism in the context of colitis and colonic regeneration.

The cellular sources of intestinal repair have been a major focus of the field over the last several decades. Recently, Murata *et al.* elegantly demonstrated a unified pathway by which any CBC progeny cell, whether a secretory or absorptive precursor, can repopulate ablated CBC cells by inducing expression of the WNT target gene and stem cell master regulator *Ascl2* to mediate de-differentiation[21]. After CBC cell ablation, 238 significantly upregulated genes defined the de-differentiating ASCL2^{+} cells compared to resting CBC cells. Interestingly, *Mtg16* was markedly (4.1x) upregulated in these cells. Furthermore, ASCL2 ChIP-seq data in the de-differentiating population yielded a peak at a second transcription start site of *Mtg16*, suggesting that ASCL2 induces *Mtg16* expression during colonic epithelial regeneration (**Fig. 5F**).

To confirm that the ASCL2-driven regeneration program was relevant to crypt regeneration following DSS-induced injury, we constructed a gene set consisting of the 238 genes upregulated in *Ascl2*$^{+}$ regenerating cells, performed GSEA, and observed a highly significant enrichment of this signature in WT distal colon crypt epithelial isolates after DSS-induced regeneration (**Fig. 5G**). Conversely, the *Ascl2*$^{+}$ regeneration signature was significantly de-enriched in *Mtg16*$^{-/-}$ colon crypts following DSS-induced injury, despite enrichment of other gene signatures associated with intestinal epithelial regeneration in the literature (**Fig. 5H-I**, **Table S3**).

Altogether, these data implicate MTG16 as a central component of ASCL2-driven regeneration and repair in the colon. Similar to our findings in colon homeostasis, E protein signatures were enriched in post-DSS *Mtg16*$^{-/-}$ colon by GSEA (**Fig. 5J-K**), leading us to hypothesize that the contribution of MTG16 to colon crypt regeneration occurs via repression of E protein transcription factors.

MTG16-mediated colonic epithelial regeneration is dependent on its repression of E protein transcription factors. To test whether MTG16-dependent regeneration occurs via an E protein-dependent mechanism, we next analyzed colonic epithelial regeneration following DSS-induced injury in *Mtg16*$^{T/T}$ mice (**Fig. 6A**). Like *Mtg16*$^{-/-}$ mice, *Mtg16*$^{T/T}$ mice exhibited significantly worse weight loss, shorter colons, and increased histologic injury (**Fig. 6B-D, Table S2**). We also observed increased extent of injury along the length of the *Mtg16*$^{T/T}$ colon (**Fig. 6E**). Next, we injected mice with EdU 1 h prior to sacrifice following DSS injury-regeneration and found that *Mtg16*$^{T/T}$ mice had fewer S-phase cells in injury-adjacent crypts (**Fig. 6F**), which actively regenerate the ulcerated colonic epithelium following DSS-induced colitis[99]. Finally, like *Mtg16*$^{-/-}$ mice, bulk RNA-seq of *Mtg16*$^{T/T}$ distal colon crypt epithelial isolates following DSS injury-repair was de-enriched for the *Ascl2*$^{+}$ de-differentiating cell signature by GSEA, again despite enrichment of other gene signatures associated with intestinal epithelial regeneration in the literature (**Fig. 6G-H, Table S3**). These data showed that MTG16 repression of E protein-mediated transcription is required for ASCL2-mediated colonic epithelial regeneration.

***MTG16* expression is decreased in dysplasia.** We previously observed decreased *MTG16* expression in sporadic CRC and CAC compared to normal colon tissue[35], but had not directly compared expression levels of *MTG16* between these different neoplastic origins. Because *MTG16* expression was upregulated in IBD, we wondered whether *MTG16* expression would be increased in IBD-driven colon cancer (CAC) compared to sporadic CRC. To do so, we first

confirmed our previous findings of decreased *MTG16* expression in sporadic CRC using DESeq2 on raw counts generated from TCGA[74–76] (**Fig. 7A**). Next, we analyzed *MTG16* expression in bulk RNA-seq of CRC and CAC tumors[100] and found no significant difference in *MTG16* expression between sporadic CRC and CAC tumors (**Fig. 7B**), indicating that *MTG16* expression is lost in colon dysplasia regardless of whether it is driven by IBD.

MTG16-mediated protection from CAC is partially dependent on its repression of E protein transcription factors. We previously showed that *Mtg16*$^{-/-}$ mice have increased mortality, tumor burden, tumor grade, and tumor cell proliferation and apoptosis when subjected to the AOM/DSS-induced inflammatory carcinogenesis CAC model[35,36]. We tested whether the mechanism driving these phenotypes was E protein-dependent by treating *Mtg16*$^{T/T}$ mice with AOM/DSS (**Fig. 7C**). *Mtg16*$^{T/T}$ mice did not have significantly higher mortality, weight loss, or colon shortening (**Fig. 7D-F**). However, *Mtg16*$^{T/T}$ mice did exhibit increased tumor multiplicity and size (**Fig. 7G-J**).

Interestingly, *Mtg16*$^{T/T}$ mice had more tumors and larger tumors in the middle and proximal colon than their WT counterparts (**Fig. 7I-K**). This was consistent with a greater extent of DSS-induced injury (**Fig. 6E**), suggesting that this phenotype could be driven by increased injury in the proximal colon. Although *Mtg16*$^{T/T}$ mice had more severe histologic injury (**Fig. 7N**), dysplasia severity did not differ (**Fig. 7O**). Additional characterization of *Mtg16*$^{T/T}$ tumors (**Fig. S7**) demonstrated no difference in proliferation (**Fig. S7A**), dissimilar to the *Mtg16*$^{-/-}$ AOM/DSS phenotype. We did observe a subtle increase in apoptotic epithelial cells per tumor (**Fig. S7B**). These phenotypes were recapitulated in tumoroids derived from distal, but not proximal, WT and *Mtg16*$^{T/T}$ tumors (**Fig. S8**). Overall, these data indicate that MTG16:E protein uncoupling recapitulates some, but not all, of the *Mtg16*$^{-/-}$ phenotypes, suggesting potential anti-tumorigenic functions of other MTG16 domains.

Discussion

In this study, we investigated the mechanism underlying context-specific functions of the transcriptional corepressor MTG16. We establish that MTG16 is a critical regulator of secretory cell differentiation and identify a previously unappreciated role for MTG16 in enteroendocrine lineage specification in the colon. MTG16 was also upregulated in active IBD, characterized by repeated cycles of injury and regeneration, and likely contributes to murine colon regeneration following DSS-induced injury. To investigate the mechanism underlying these phenotypes, we utilized a novel mouse model with a single point mutation in MTG16, *Mtg16*^{P209T}, that disrupts MTG16 binding to the E proteins E2A and HEB[38]. *Mtg16*$^{T/T}$ mice homozygous for this MTG16:E protein uncoupling mutation largely recapitulated *Mtg16*$^{-/-}$ lineage allocation and regeneration phenotypes, although the effect of MTG16 loss in dysplasia appears to only partially be due to increased E protein activity. Thus, these data indicate a novel role for MTG16:E protein complexes in the colonic epithelium.

First, we demonstrated a novel role for the transcriptional corepressor MTG16 in colonic epithelial homeostasis. MTG16 has been shown to repress stem, goblet, and enteroendocrine cell genes to drive enterocyte differentiation in the SI[33]. However, consistent with numerous studies showing important differences in SI and colon biology[3,4,11,21,101–103], *Mtg16*$^{-/-}$ colon did not exhibit increased stem cell signatures or a pan-secretory defect as observed in the SI. Instead, we observed attenuation of the goblet cell lineage and upregulation of the enteroendocrine lineage. *Neurog3*, a bHLH transcription factor required for enteroendocrine cell differentiation[86,104–106], was consistently upregulated in *Mtg16*$^{-/-}$ colon. Using GSEA with gene sets for homeostatic epithelial cell types, including cells at different stages of enteroendocrine cell differentiation, we determined that *Mtg16*$^{-/-}$ colon demonstrated enrichment of all enteroendocrine cell progenitor signatures *except* the earliest enteroendocrine precursor in which *Neurog3* expression has not yet been induced[86]. Finally, *Mtg16*$^{-/-}$ colon exhibited increased *Neurog3*$^{+}$ enteroendocrine progenitor cells by RNAscope. Although *Neurog3*$^{+}$ progenitor cells in the SI may

still differentiate into goblet and Paneth cells[102,106], they are thought to be fully committed to the enteroendocrine lineage in the colon[102]. Thus, MTG16 functions as a nexus of cell fate control in colonic enteroendocrine:goblet lineage assignment.

In vitro studies and studies of the hematopoietic system have shown that MTG16 binds and represses various transcription factors relevant to colon biology, including N1-ICD, CSL, and TCF7L2[25,29,32]. However, in this study, we were surprised to find that the homeostatic *Mtg16*$^{-/-}$ colon did not exhibit upregulation of WNT or Notch pathways by bulk RNA-seq. Instead, *Mtg16*$^{-/-}$ colon had significant enrichment of E protein-mediated transcriptional signatures. E proteins, also known as type I bHLH transcription factors, homo- and heterodimerize with each other and heterodimerize with tissue-specific type II bHLH transcription factors to initiate transcription of a cascade of lineage-specific transcription factors[40–42]. *Tcf4* (E2A) and *Tcf12* (HEB) expression has been detected in intestinal stem cells and adenomas[87]; however, the functions of E2A and HEB have not been established in colon homeostasis. Furthermore, enteroendocrine cell lineage commitment is dependent on a cascade of type II bHLH transcription factors[107,108], including NEUROG3 and its downstream target NEUROD1[104,109], which is known to heterodimerize with E12 and E47 in other systems[110]. Additionally, E2A homodimers promote neural lineage commitment by inducing transcription of *Neurog1-3* in neurogenesis[47,49]. Thus, we hypothesized that MTG16 attenuated enteroendocrine lineage allocation by repressing E protein-mediated transcription of enteroendocrine bHLH transcription factors. We tested this hypothesis using a highly specific mouse model with a point mutation, *Mtg16*^{P209T}, that partially disrupts MTG16:E protein binding and repression[38]. Indeed, *Mtg16*$^{T/T}$ mice homozygous for this point mutation had downregulated goblet cell genes and upregulated enteroendocrine cell genes in the colon epithelium, a pattern phenocopying *Mtg16*$^{-/-}$ colon. *Mtg16*$^{T/T}$ mice also had upregulated *Neurog3* by RNA-seq, increased *Neurog3*$^{+}$ progenitors by RNAscope, and enrichment of all gene signatures representing enteroendocrine progenitors following NEUROG3 induction[86]. Additionally, mutant (*Mtg16*^{P209T}) transcript was upregulated in *Mtg16*$^{T/T}$ distal colon crypts,

consistent with autoregulation previously demonstrated in the hematopoietic system[111], potentially increasing MTG16-mediated repression of other (non-E protein) transcription factors in *Mtg16*$^{T/T}$ colon crypts. One such potential non-E protein-mediated pathway that was increased in *Mtg16*$^{-/-}$, but not *Mtg16*$^{T/T}$ colon crypts, is the WNT-planar cell polarity (PCP) pathway, a non-canonical, β-catenin-independent pathway regulating cell polarity and organogenesis[112]. This could be responsible for the slight differences between *Mtg16*$^{-/-}$ and *Mtg16*$^{T/T}$ phenotypes, as it was recently shown that enteroendocrine cells may differentiate directly from "unipotentially-primed" intestinal stem cells by the WNT-PCP pathway[113].

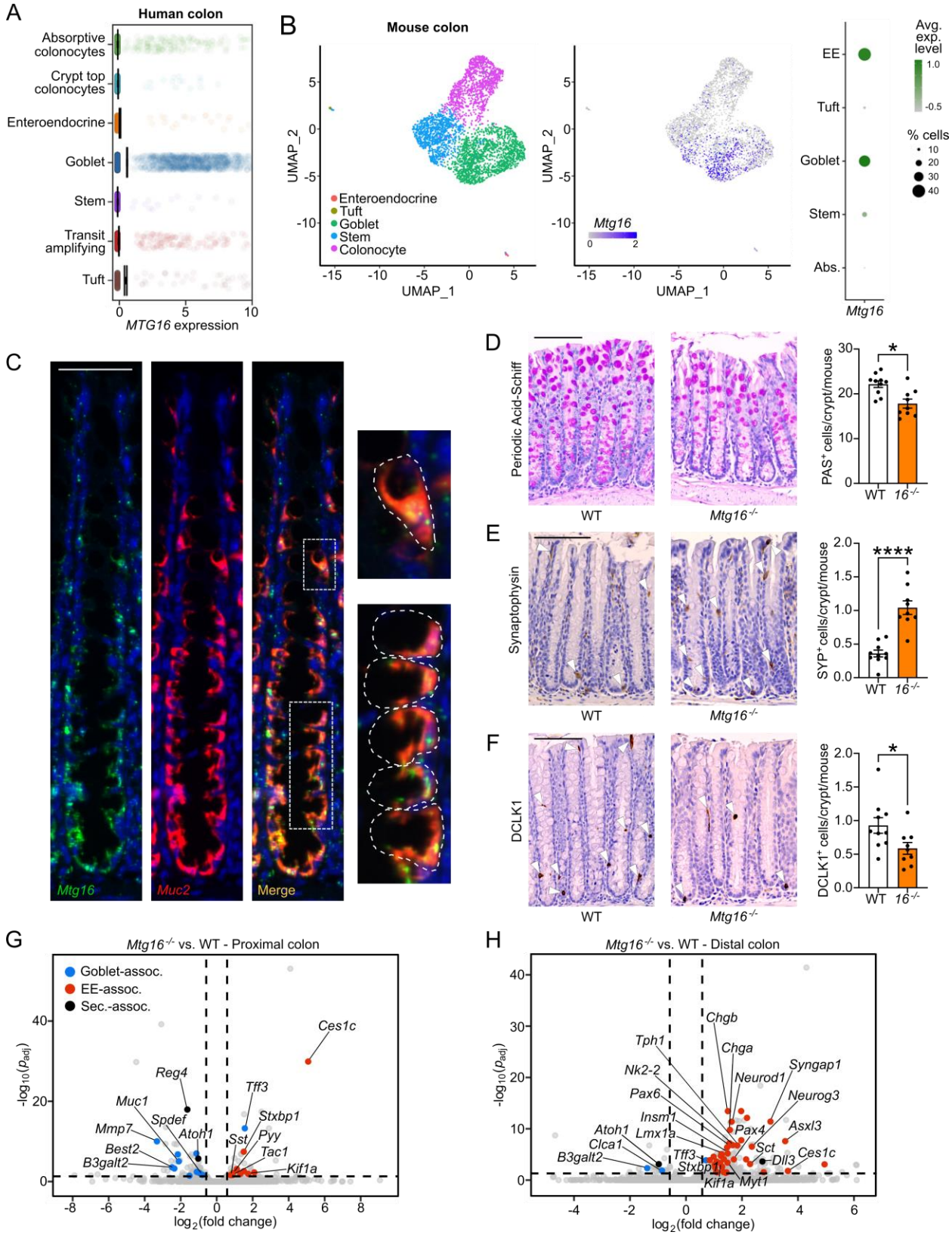
An area of considerable interest in gastrointestinal biology is the link between epithelial differentiation and regeneration[21,114,115]. Although it was long thought that dedicated "reserve" stem cells at the +4/+5 region, marked by expression of genes such as *Bmi1* and *mTert*, replenished the stem cell compartment in intestinal epithelial regeneration, recent data have indicated that these cells are most likely enteroendocrine cells with the ability to de-differentiate into stem cells following injury[8,22]. Another recent study identified a "revival" stem cell characterized by high *Clu* expression responsible for replenishing the stem cell compartment[116]. However, "committed" progenitors from both the secretory and absorptive lineages have been demonstrated to share this ability[21,117–119]. A recent study used a novel mouse model expressing FLAG-tagged ASCL2 and a fluorescent reporter from the native *Ascl2* locus, allowing the isolation of *Ascl2*$^{+}$ de-differentiating cells from the colonic epithelium following CBC cell ablation[21]. Interestingly, RNA-seq and ASCL2 ChIP-seq showed that these cells were not enriched for *Clu* or fetal epithelial genes previously thought to enhance intestinal epithelial response to injury[120,121]. In fact, *Mtg16* was one of the few genes upregulated in these regenerating cells[21]. Additionally, ChIP-seq demonstrated an ASCL2 peak at a second TSS in the *Mtg16* locus in regenerating colon epithelium[21]. Demonstrating relevance to DSS-induced colon regeneration, both *Mtg16* and the *Ascl2*$^{+}$ regenerating cell signature were upregulated in regenerating WT colon epithelium compared to baseline. Mirroring this upregulation of *Mtg16* in murine colitis, *MTG16* was

upregulated in IBD-affected colon, compared to normal colon, in multiple large patient cohorts. *MTG16* expression decreased in patients in remission, during which the colon epithelium is closer to homeostatic renewal. Finally, RNA-seq and GSEA of *Mtg16*$^{-/-}$ and *Mtg16*$^{T/T}$ colonic epithelial isolates following DSS-induced injury demonstrated enrichment of "revival" stem cell signatures and fetal epithelial signatures, but de-enrichment of the *Ascl2*$^{+}$ regenerating cell signature. These data indicate an active, E protein-mediated role for MTG16 in colon regeneration and provide a potential explanation for why *Mtg16*$^{-/-}$ mice exhibit defective colonic epithelial regeneration despite an apparent enrichment of *Bmi1*$^{+}$ and *m-Tert*$^{+}$ "reserve" stem cells at baseline. Future investigations will determine which E protein targets are being repressed by MTG16 in the *Ascl2*$^{+}$ de-differentiating cells. One possibility is that MTG16 represses *Neurog3* in these de-differentiating cells to prevent them from reversing course and committing to the enteroendocrine cell lineage.

Finally, we investigated the role of MTG16:E protein complexes in the setting of inflammatory carcinogenesis. We have previously shown that *Mtg16*$^{-/-}$ mice develop greater tumor burden than their WT counterparts with AOM/DSS treatment[35]. Although E proteins maintain lineage commitment in homeostasis, they have been shown to be both anti-tumorigenic and pro-tumorigenic in colon cancer[122,123]. Thus, we postulated that MTG16-mediated protection from tumorigenesis was mechanistically due to repression of E protein activity. We did observe an increase in histologic injury, tumor burden, and tumor size in *Mtg16*$^{T/T}$ mice compared to WT mice following AOM/DSS treatment. It is notable, however, that while *Mtg16*$^{-/-}$ tumors exhibited much higher grade of dysplasia[35], *Mtg16*$^{T/T}$ tumors did not. Thus, disabling one known function of MTG16 was not sufficient to fully replicate the *Mtg16*$^{-/-}$ phenotype. Indeed, our group recently showed that MTG16-dependent effects on CAC development is dependent on ZBTB33 (also known as Kaiso)[36]. Mouse models with different mutations in MTG16 could be used to determine additional MTG16 binding partners and repression targets responsible for suppressing carcinogenesis.

In conclusion, we discovered novel functions of MTG16 in the colonic epithelium dependent on its repression of E proteins using a novel mouse model with a point mutation that disrupts MTG16:E protein interactions. We show that MTG16:E protein complexes direct secretory lineage allocation, regeneration in response to injury, and tumorigenesis. Confirming translational relevance, we determined that *MTG16* is upregulated in patients with active IBD, but returns closer to baseline in remission. Thus, *MTG16* may be a candidate biomarker for disease activity or a target to modulate secretory lineage differentiation, in addition to its function in promoting epithelial regeneration.

621 **Figures**



622

Figure 1. MTG16 is expressed in a subset of colonic secretory cells, and its deficiency alters lineage allocation. (A) *MTG16* expression in cell populations generated from inDrop scRNA-seq of normal human colon biopsies (discovery cohort, n = 30,374 cells sequenced from 35 samples) (validation cohort in Fig. S1). **(B)** *Mtg16* expression in inDrop scRNA-seq of WT mouse colon (n = 3). EE, enteroendocrine. The average *MTG16* expression level in each cell population is represented using the color gradient, while the percentage of cells expressing *MTG16* in each cell population is represented using dot size. **(C)** RNAscope *in situ* hybridization of *Mtg16* and *Muc2* in WT mouse colon. Scale bar = 50 µm. Goblet cells are outlined in the insets at right. **(D-F)** WT and *Mtg16*$^{-/-}$ mouse colon (n = 10 WT, 9 *Mtg16*$^{-/-}$) stained for **(D)** goblet cells by periodic acid-Schiff (PAS) stain, **(E)** enteroendocrine cells by IHC for synaptophysin (SYP), and **(F)** tuft cells by IHC for doublecortin-like kinase 1 (DCLK1). Representative images at left. SYP^{+} and DCLK1^{+} cells are indicated with white arrows. Scale bars = 100 µm. *$p < 0.05$, ****$p < 0.0001$ by Mann-Whitney test. **(G-H)** Volcano plots demonstrating differentially expressed genes in bulk RNA-seq of colon epithelial isolates from **(G)** proximal (n = 2 WT, 2 *Mtg16*$^{-/-}$) and **(H)** distal (n = 4 WT, 4 *Mtg16*$^{-/-}$) mouse colon. Horizontal dashed line indicates $p_{adj} < 0.05$ by DESeq2. Vertical dashed lines indicate fold change = 1.5. Goblet-associated-genes are shown in blue, enteroendocrine (EE)-associated genes are shown in red, and secretory-associated genes are shown in black.

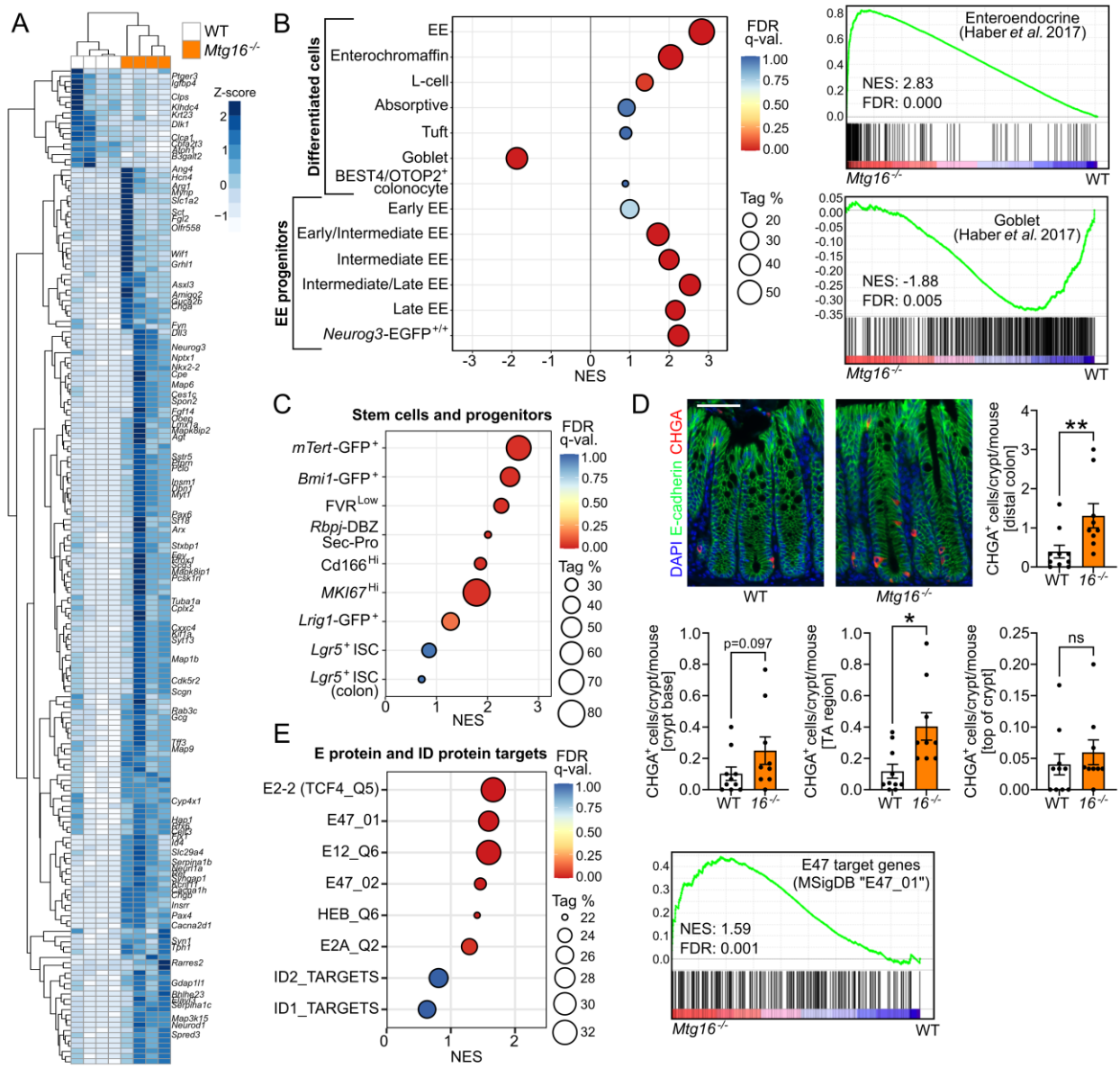


Figure 2. MTG16 deficiency is associated with increased enteroendocrine cell markers, reserve stem cell gene signatures, and E protein target upregulation in the distal colon. **(A)** Unsupervised hierarchical clustering of differentially expressed genes in distal colon RNA-seq (n = 4 WT, 4 *Mtg16*$^{-/-}$). Gene expression in individual samples is displayed by Z-score (normalized count). Genes of interest are labeled at right. **(B)** GSEA of distal colon RNA-seq using gene sets representing epithelial cell types, derived from the literature (Table S3). EE, enteroendocrine. Individual GSEA plots of enteroendocrine and goblet cell signatures displayed at right. FDR q-value < 0.25 is considered significant. Tag % is defined as the percentage of gene hits before (for positive enrichment scores) or after (for negative enrichment scores) the peak in the running enrichment score, indicating the percentage of genes contributing to the enrichment score. NES, normalized enrichment score. **(C)** GSEA of distal colon RNA-seq using gene sets representing stem and progenitor cell populations, derived from the literature (Table S3). **(D)** Staining and quantification of CHGA^{+} cells/crypt (top) stratified by location along the crypt axis (bottom). Scale bar = 50 µm. n = 10 WT, 9 *Mtg16*$^{-/-}$. *p < 0.05, **p < 0.01 by Mann-

Whitney test. **(E)** GSEA of distal colon RNA-seq using MSigDB gene sets representing E protein and ID protein transcriptional signatures. Individual GSEA plot of the E47 target gene signature at right.

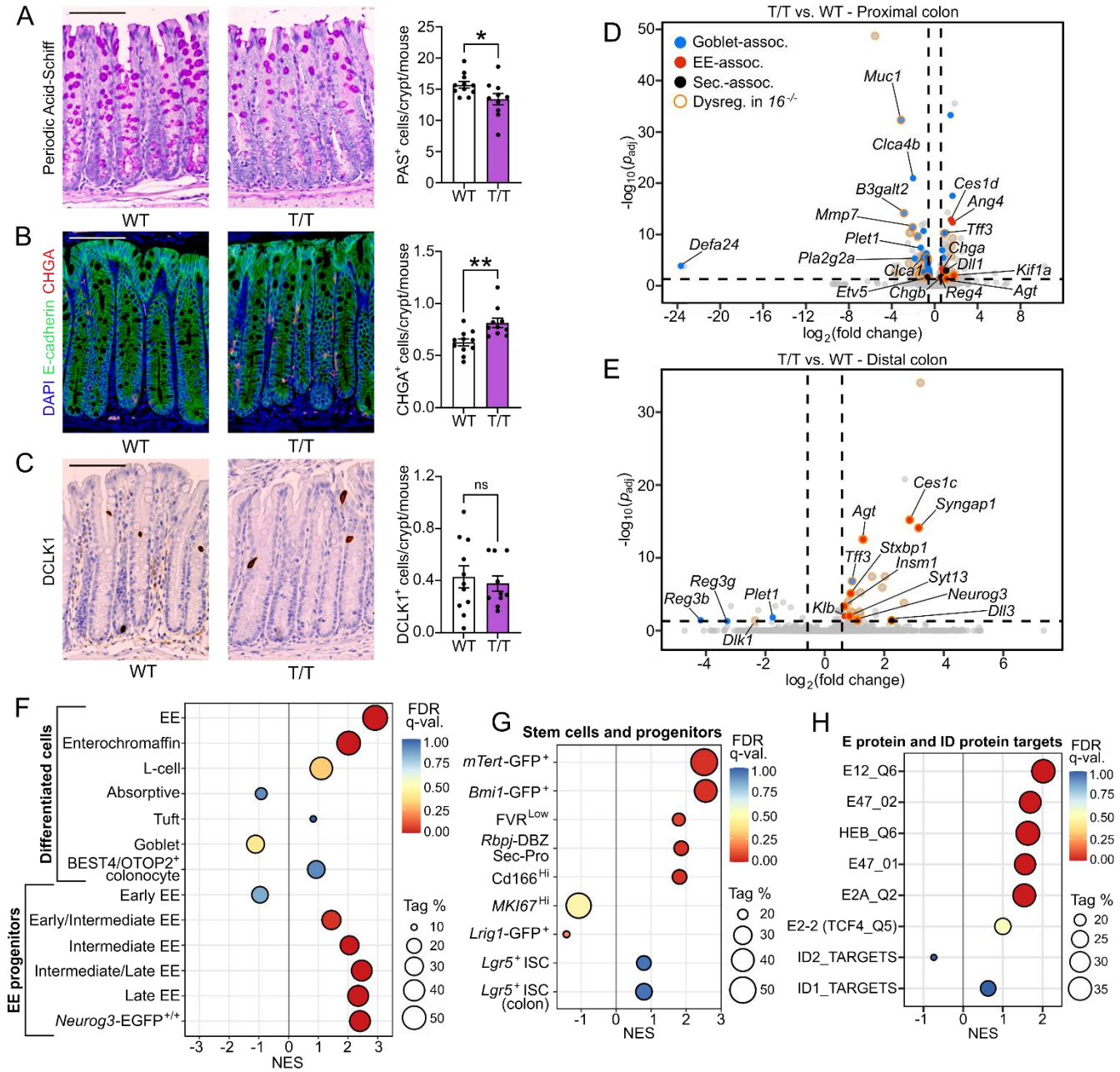
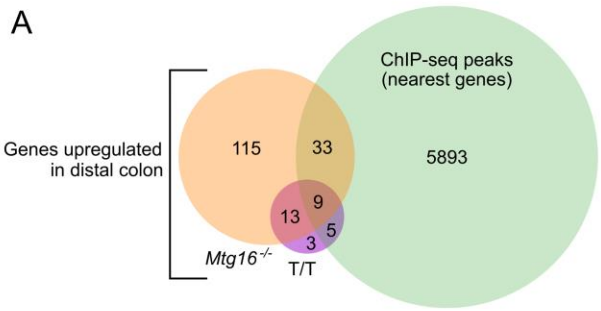


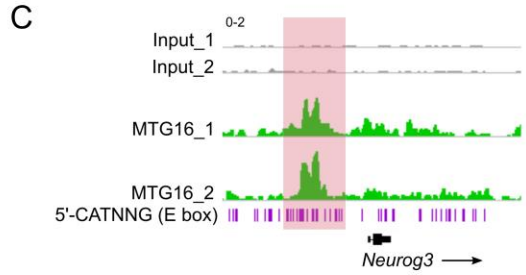
Figure 3. MTG16-driven colonic epithelial lineage allocation is dependent on its repression of the E protein transcription factors E2A and HEB. (A) WT and *Mtg16*$^{T/T}$ mouse colon (n = 11 WT, 10 *Mtg16*$^{T/T}$) stained for **(A)** goblet cells by periodic acid-Schiff (PAS) stain, **(B)** enteroendocrine cells by IHC for synaptophysin (SYP), and **(C)** tuft cells by IHC for doublecortin-like kinase 1 (DCLK1). Representative images at left. Scale bars = 100 µm. *p < 0.05, **p < 0.01 by Mann-Whitney test. **(D-E)** Volcano plots demonstrating differentially expressed genes in bulk RNA-seq of colon epithelial isolates from **(D)** proximal (n = 2 WT, 2 *Mtg16*$^{T/T}$) and **(E)** distal (n = 4 WT, 4 *Mtg16*$^{T/T}$) mouse colon. Horizontal dashed line indicates p_{adj} < 0.05 by DESeq2. Vertical dashed lines indicate fold change = 1.5. Goblet-associated-genes are shown in blue, enteroendocrine (EE)-associated genes are shown in red, and secretory-associated ("Sec.-associated") genes are shown in black. Genes that are similarly

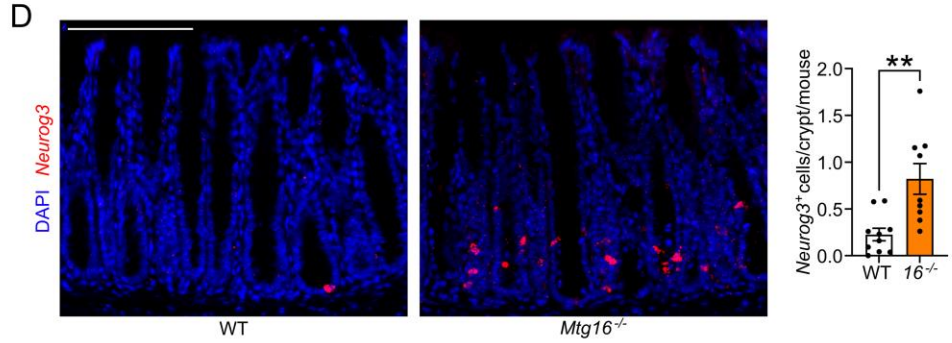
dysregulated (upregulated or downregulated in the same direction) in *Mtg16*$^{-/-}$ colonic epithelium ("Dysreg. In *16*$^{-/-}$") are outlined in orange. **(F-G)** GSEA of distal colon RNA-seq using gene sets derived from the literature to represent epithelial cell types (Table S3). **(H)** GSEA of distal colon RNA-seq using MSigDB gene sets representing E protein and ID protein transcriptional signatures. **(F-H)** FDR q-value < 0.25 is considered significant. Tag % is defined as the percentage of gene hits before (for positive enrichment scores) or after (for negative enrichment scores) the peak in the running enrichment score, indicating the percentage of genes contributing to the enrichment score. NES, normalized enrichment score.



B

Genes up in distal colon	*Mtg16*$^{-/-}$ vs. WT (Log_2FC)	*Mtg16*$^{T/T}$ vs. WT (Log_2FC)	ChIP-seq peak (nearest TSS)?
Agt	1.55	1.29	No
Ces1c	2.77	2.85	No
Dbn1	2.07	0.95	No
Dll3	2.73	2.25	Yes
Dmpk	1.80	1.26	Yes
Gfra3	1.07	0.89	No
Gnao1	1.42	1.20	No
Guca2b	1.14	0.79	Yes
Id4	1.92	1.19	No
Insm1	1.49	0.65	No
Insrr	4.30	3.22	Yes
Kcnh6	1.94	1.17	Yes
Neurog3	**2.35**	**1.09**	**Yes**
Ooep	2.83	2.66	No
Pex5l	2.65	2.02	No
Prss23	2.52	1.59	Yes
Rbp4	1.66	0.88	No
Spon2	2.84	1.93	No
Stxbp1	1.04	0.87	Yes
Syngap1	3.02	3.16	No
Syt13	1.97	0.83	No
Tff3	0.72	0.92	Yes





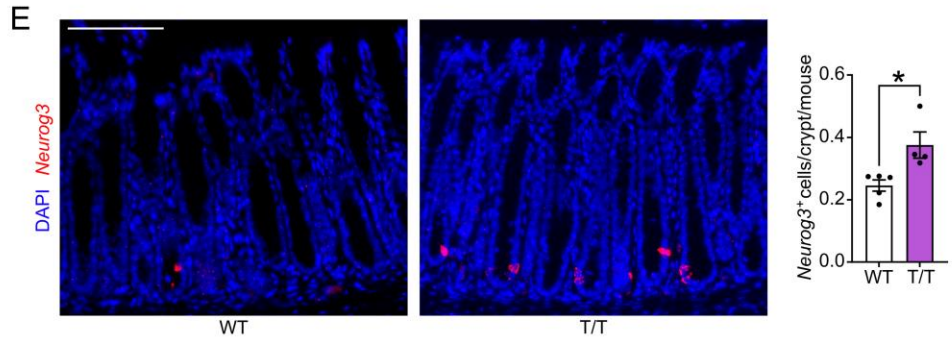


Figure 4. MTG16 represses E protein-mediated transcription of *Neurog3*. (A) Venn diagram of upregulated ($p_{adj} < 0.05$) genes from bulk RNA-seq of distal colon crypt isolates (n = 4 WT, 4 *Mtg16*$^{-/-}$, orange; (n = 3 WT, 4 *Mtg16*$^{T/T}$, purple) and genes with transcription start sites (TSS) nearest to MTG16 ChIP-seq peaks[33] (green). **(B)** Table of the 22 genes upregulated in both *Mtg16*$^{-/-}$ and *Mtg16*$^{T/T}$ colon RNA-seq indicating their log_2(fold change) (Log_2FC) in each dataset (*Mtg16*$^{-/-}$ vs. WT distal colon and *Mtg16*$^{T/T}$ vs. WT distal colon) and whether they have a TSS nearest to an MTG16 ChIP-seq peak[33]. A searchable version of this table is provided in Table S4. **(C)** MTG16 occupancy of an E box (5'-CATNNG)-rich region 5' to *Neurog3*[33]. Scale: 0-2. **(D-**

E) Staining and quantification of *Neurog3*$^{+}$ cells by RNAscope. **(D)** n = 10 WT, 9 *Mtg16*$^{-/-}$ and **(E)** n = 5 WT, 4 *Mtg16*$^{T/T}$. *p < 0.05, **p < 0.01 by Mann-Whitney test.

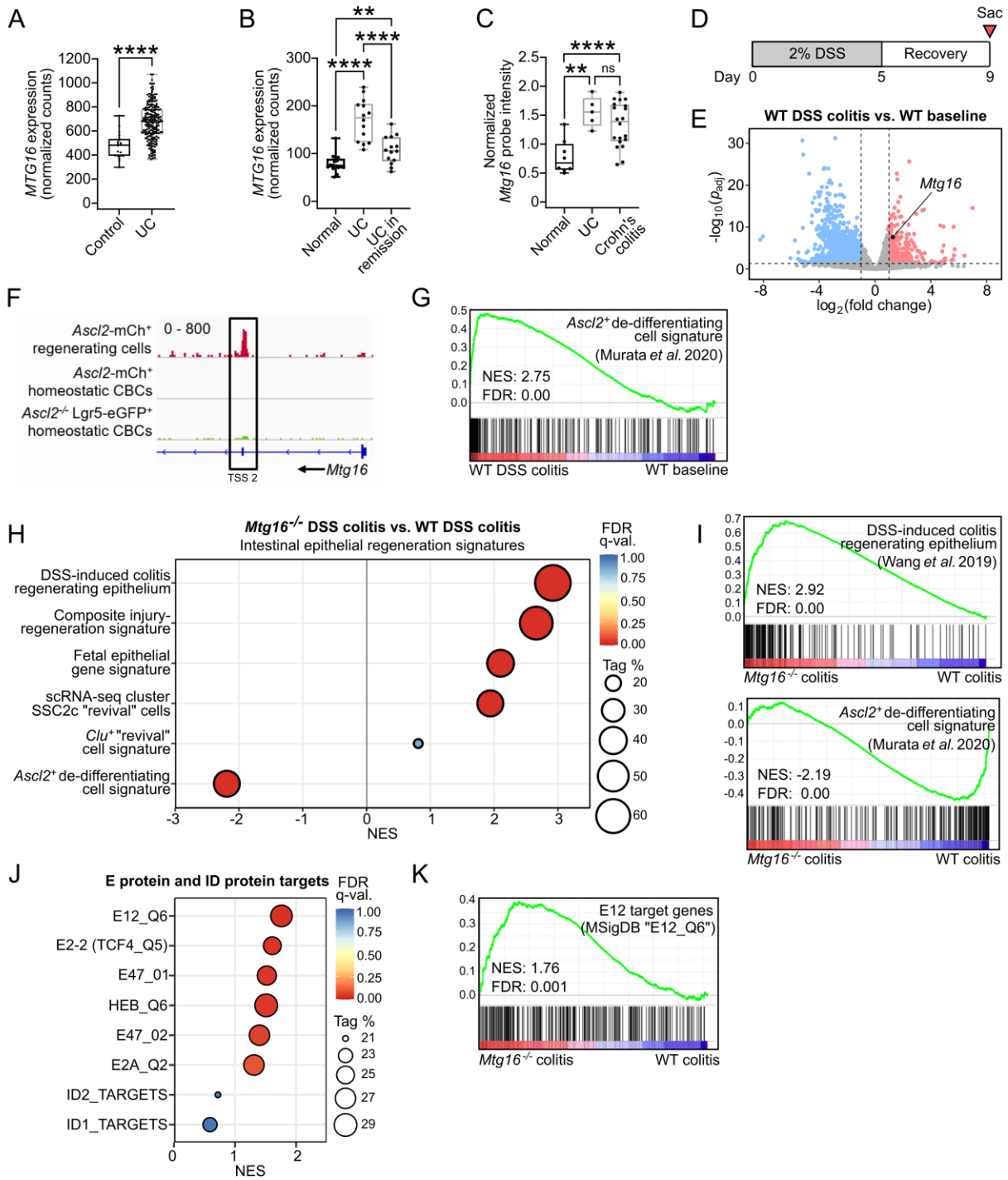


Figure 5. *Mtg16* is upregulated in experimental colitis and patients with IBD and is required for ASCL2-mediated colonic epithelial regeneration. (A) *MTG16* expression from an RNA-seq dataset generated from rectal biopsies of treatment-naïve pediatric UC patients (n = 206) vs. unaffected controls (n = 20) in the PROTECT study[70]. Transcript counts were normalized by DESeq2. ****p_{adj} < 0.0001 by DESeq2. **(B)** *MTG16* expression in an RNA-seq dataset from adult UC patients with active disease (n = 14), UC patients in remission (n = 14), and unaffected controls (n = 16)[72]. Transcript counts were normalized by DESeq2. **p_{adj} < 0.01 and ****p_{adj} < 0.0001 by pairwise DESeq2. **(C)** *MTG16* expression in a microarray analysis of

UC (n = 5) and Crohn's colitis (n = 20) patient biopsies compared to unaffected controls (n = 8)[73]. Normalized probe intensity is plotted for visualization purposes. **p_{adj} < 0.01 and ****p_{adj} < 0.0001 by limma in GEO2R. **(A-C)** The lines within each box represent the mean, the bounds of the boxes represent 25th to 75th %-iles, and the bounds of the whiskers represent the range of the data. All individual data points are shown. **(D)** Schematic of the DSS injury-repair model in which mice are treated with 2% DSS for 5 days followed by 4 days of recovery. **(E)** Volcano plot showing significantly downregulated (blue) and upregulated (red) genes in distal colon crypt isolates from WT mice following DSS-induced injury and regeneration (n = 3) compared to distal colon crypt isolates from WT mice at baseline (n = 3). Horizontal dashed line indicates p_{adj} < 0.05 by DESeq2. Vertical dashed lines indicate fold change = 2. *Mtg16* is significantly upregulated in DSS-induced colitis. **(F)** ASCL2-FLAG occupancy near the *Mtg16* transcription start site in ChIP-seq data from FACS-sorted *Ascl2*$^{+}$ regenerating cells[21]. Scale: 0-800. TSS 2, Transcription Start Site 2. **(G)** GSEA plot demonstrating significant enrichment of 238 genes upregulated in *Ascl2*$^{+}$ de-differentiating cells (shown to repopulate the colon crypt stem cell compartment following injury[21]) in distal colon crypt isolates from WT mice following DSS-induced injury and regeneration (n = 3) compared to distal colon crypt isolates from WT mice at baseline (n = 3). FDR q-value < 0.25 is considered significant. **(H-K)** GSEA performed on bulk RNA-seq of *Mtg16*$^{-/-}$ distal colon crypt isolates following DSS injury-repair (n = 3 WT, 4 *Mtg16*$^{-/-}$). **(H)** Summary of GSEA for gene sets showing significant enrichment of multiple gene sets associated with epithelial regeneration derived from the literature (Table S3), but significant de-enrichment of the *Ascl2*$^{+}$ de-differentiating cell signature in the setting of MTG16 deficiency. **(I)** Individual GSEA plots for DSS-induced colitis regenerative epithelium (left) and *Ascl2*$^{+}$ de-differentiating cell signature (right). **(J)** Summary of GSEA for MSigDB gene sets representing E protein and ID protein transcriptional signatures (Table S3). **(K)** Individual GSEA plot of the E12 target gene signature. **(H-K)** FDR q-value < 0.25 is considered significant. Tag % is defined as the percentage of gene hits before (for positive enrichment scores) or after (for negative enrichment scores) the peak in the running enrichment score, indicating the percentage of genes contributing to the enrichment score. NES, normalized enrichment score.

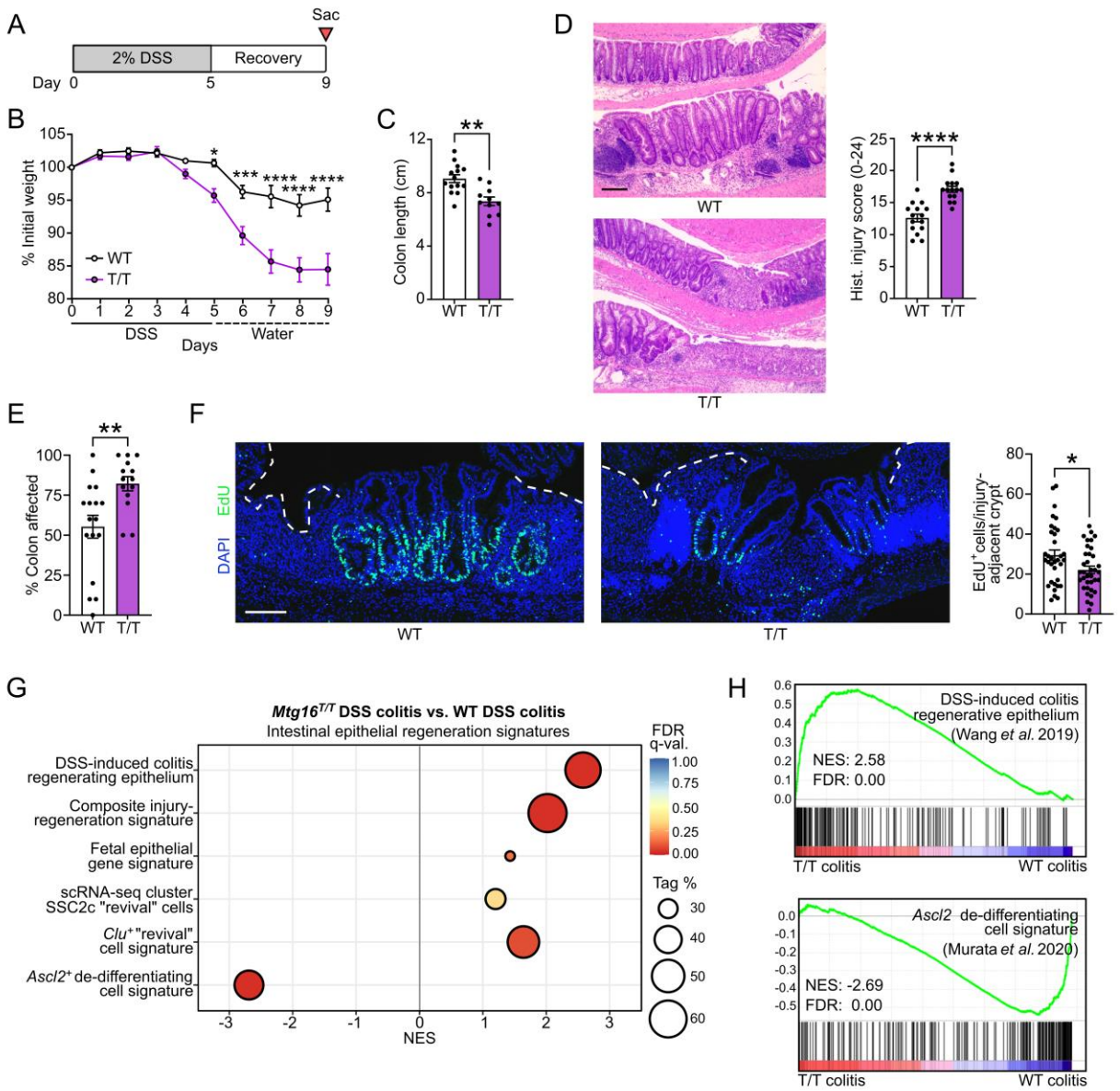


Figure 6. MTG16-mediated colonic epithelial regeneration is dependent on its repression of E protein transcription factors. (A) Schematic of the DSS injury-repair model in which mice are treated with 2% DSS for 5 days followed by 4 days of recovery. **(B-D)** Colitis severity evaluated using daily weight loss **(B),** colon length **(C)**, and histologic injury score **(Table S2)** assessed by a gastrointestinal pathologist blinded to genotype **(D)**. **(E)** Percent of colon with histologic findings, estimated from the distal end, as a measure of proximal extent of disease. **(B-E)** Data are combined from 2 independent experiments (n = 15 WT, 16 *Mtg16*$^{T/T}$). **(F)** EdU-positive cells in injury-adjacent, regenerating crypts at the time of sacrifice on day 9 (n = 30-50 crypts from n = 5 WT, 5 *Mtg16*$^{T/T}$ mice). Representative images at left. Dashed white line = ulcerated epithelium. Scale bar = 100 µm. **(B-F)** *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$ by 2-way ANOVA followed by Sidak's multiple comparison tests **(B)** or Mann-Whitney test **(C-F)**. **(G-H)** GSEA performed on bulk RNA-seq of distal *Mtg16*$^{T/T}$ colon crypt isolates following DSS injury-repair (n = 3 WT, 3 *Mtg16*$^{T/T}$) (gene sets described in Table S3). **(H)** Individual GSEA plots for DSS-induced colitis regenerative epithelium (top) and *Ascl2*$^{+}$ de-

differentiating cell signature (bottom). **(G-H)** FDR q-value < 0.25 is considered significant. Tag % is defined as the percentage of gene hits before (for positive enrichment scores) or after (for negative enrichment scores) the peak in the running enrichment score, indicating the percentage of genes contributing to the enrichment score. NES, normalized enrichment score.

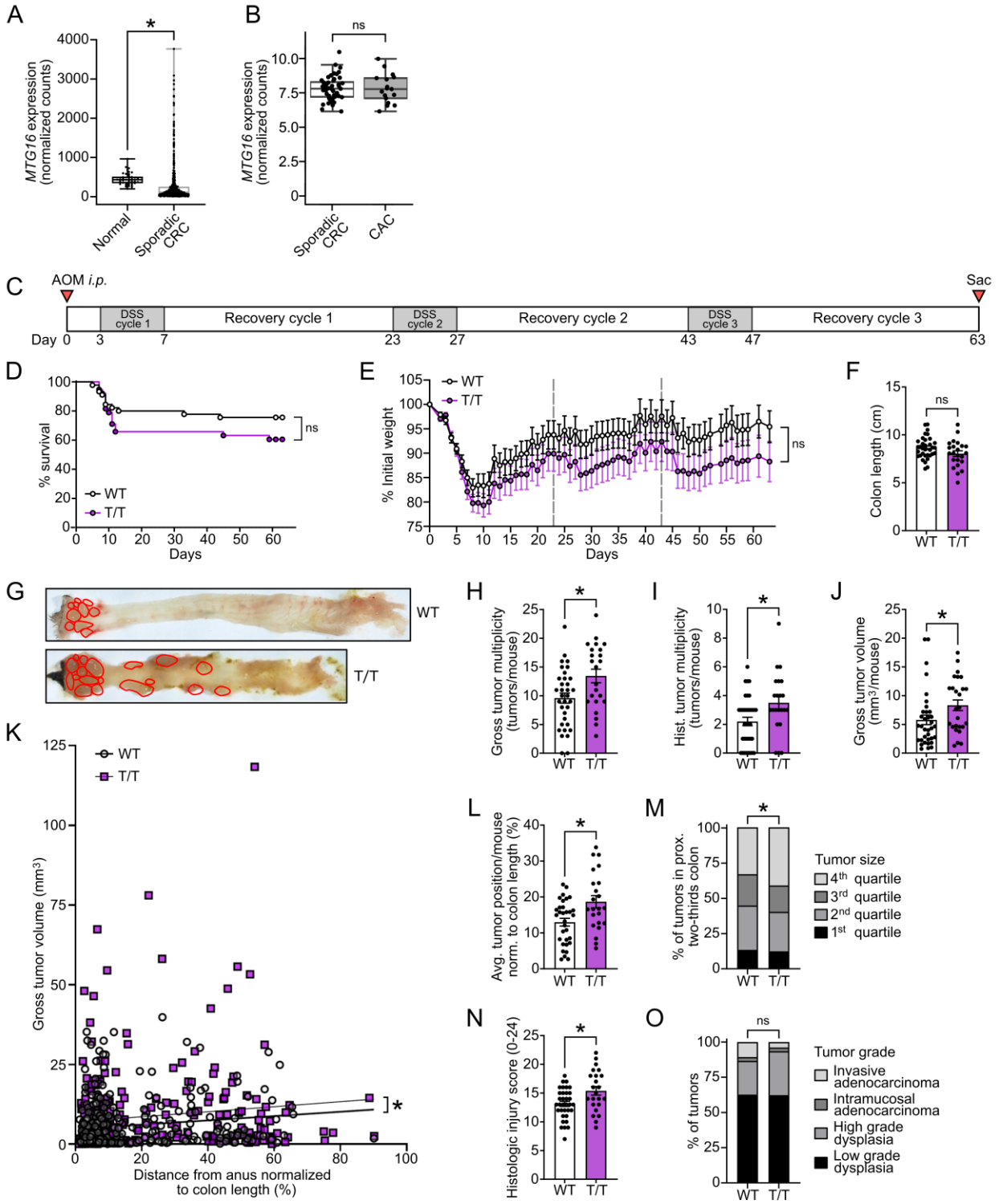


Figure 7. MTG16-mediated protection from tumorigenesis is partially dependent on repression of E protein activity. (A) *MTG16* expression in sporadic CRC (n = 482) compared to normal colon tissue (n = 41) queried from TCGA data analyzed by Rahman *et al.*[76]. Counts were normalized by DESeq2. The lines within each box represent the mean, the bounds of the boxes represent the 25th to 75th %-iles, and the bounds of the whiskers represent the range of

the data. All individual data points are shown. *p_{adj} < 0.05 by DESeq2. **(B)** *MTG16* expression queried from sporadic CRC (n = 47) and CAC (n = 17) tumors. *MTG16* expression was normalized by DESeq2 and limma batch correction. The box extends from the lower to upper quartile values of the data, with a line at the median. The lower whisker is at the lowest datum above Q1 - 1.5*(Q3-Q1), and the upper whisker at the highest datum below Q3 + 1.5*(Q3-Q1), where Q1 and Q3 are the first and third quartiles. **(C)** Schematic of the AOM/DSS model in which mice are injected with AOM followed by 3 cycles of DSS-induced injury and recovery. *i.p.*, intraperitoneal injection. **(D)** Survival curves compared using Log-rank (Mantel-Cox) test. Data are pooled from 2 independent experiments with initial n = 45 WT, 38 *Mtg16*$^{T/T}$. **(E)** Weight loss compared using repeated measures 2-way ANOVA (initial n = 26 WT, 25 *Mtg16*$^{T/T}$), representative of 2 independent experiments. **(F-O)** Gross and histologic findings following sacrifice on Day 63, including **(F)** colon length, **(G)** representative images of dissected colon (red outlines indicate tumors), **(H)** gross tumor multiplicity (tumors/mouse visualized upon sacrifice), **(I)** histologic tumor multiplicity (tumors/mouse visualized in H&Es), and **(J)** tumor volume (average mm^3/mouse measured using calipers at sacrifice). **(K)** Individual tumors graphed by size and distance from anus normalized to colon length (distance from anus/colon length*100%). **(L)** Average tumor position (distance from anus) per mouse normalized to colon length. **(M)** Percent of tumors located in the middle and proximal colon (proximal two-thirds colon) by tumor size divided into quartiles. **(L)** Histologic injury scores (scoring described in Table S2). **(O)** Percent of tumors in each grade of dysplasia. **(F-O)** Data are pooled from 2 independent experiments with n = 35 WT, 26 *Mtg16*$^{T/T}$ remaining at sacrifice. *p < 0.05 by Mann-Whitney test **(D, F-J, L)**, least-squares regression **(I)**, or Chi-square test **(K, M)**.

References

1. Allaire JM, Crowley SM, Law HT, et al. The Intestinal Epithelium: Central Coordinator of Mucosal Immunity. Trends Immunol 2018;39:677–696.

2. Kiela PR, Ghishan FK. Physiology of Intestinal Absorption and Secretion. Best Pract Res Clin Gastroenterology 2016;30:145–159.

3. Beumer J, Clevers H. Cell fate specification and differentiation in the adult mammalian intestine. Nat Rev Mol Cell Bio 2021;22:39–53.

4. Beumer J, Puschhof J, Bauzá-Martinez J, et al. High-Resolution mRNA and Secretome Atlas of Human Enteroendocrine Cells. Cell 2020;181:1291-1306.e19.

5. Fothergill LJ, Callaghan B, Hunne B, et al. Costorage of Enteroendocrine Hormones Evaluated at the Cell and Subcellular Levels in Male Mice. Endocrinology 2017;158:2113–2123.

6. Fothergill LJ, Furness JB. Diversity of enteroendocrine cells investigated at cellular and subcellular levels: the need for a new classification scheme. Histochem Cell Biol 2018;150:693–702.

7. Beumer J, Gehart H, Clevers H. Enteroendocrine dynamics – New tools reveal hormonal plasticity in the gut. Endocr Rev 2020;41:bnaa018-.

8. Yan KS, Gevaert O, Zheng GXY, et al. Intestinal Enteroendocrine Lineage Cells Possess Homeostatic and Injury-Inducible Stem Cell Activity. Cell Stem Cell 2017;21:78-90.e6.

9. Parikh K, Antanaviciute A, Fawkner-Corbett D, et al. Colonic epithelial cell diversity in health and inflammatory bowel disease. Nature 2019;567:49–55.

10. Birchenough GMH, Nyström EEL, Johansson MEV, et al. A sentinel goblet cell guards the colonic crypt by triggering Nlrp6-dependent Muc2 secretion. Science 2016;352:1535–1542.

11. Nyström EEL, Martinez-Abad B, Arike L, et al. An intercrypt subpopulation of goblet cells is essential for colonic mucus barrier function. Science 2021;372:eabb1590.

12. McCauley HA, Guasch G. Three cheers for the goblet cell: maintaining homeostasis in mucosal epithelia. Trends Mol Med 2015;21:492–503.

13. Birchenough GMH, Johansson ME, Gustafsson JK, et al. New developments in goblet cell mucus secretion and function. Mucosal Immunol 2015;8:712–719.

14. Worthington JJ, Reimann F, Gribble FM. Enteroendocrine cells-sensory sentinels of the intestinal environment and orchestrators of mucosal immunity. Mucosal Immunol 2018;11:3–20.

15. Holst JJ. Enteroendocrine secretion of gut hormones in diabetes, obesity and after bariatric surgery. Curr Opin Pharmacol 2013;13:983–988.

16. Skibicka KP, Dickson SL. Enteroendocrine hormones—central effects on behavior. Curr Opin Pharmacol 2013;13:977–982.

17. Yamane S, Inagaki N. Control of intestinal stem cell fate: A novel approach to treating diabetes. J Diabetes Invest 2016;7:166–168.

18. Aliluev A, Tritschler S, Sterr M, et al. Diet-induced alteration of intestinal stem cell function underlies obesity and prediabetes in mice. Nat Metabolism 2021;3:1202–1216.

19. Yu Y, Yang W, Li Y, et al. Enteroendocrine Cells: Sensing Gut Microbiota and Regulating Inflammatory Bowel Diseases. Inflamm Bowel Dis 2019;26:11–20.

20. Kanke M, Kennedy MM, Connelly S, et al. Single-cell analysis of colonic epithelium reveals unexpected shifts in cellular composition and molecular phenotype in treatment-naïve adult Crohn's disease. Biorxiv 2021:2021.01.13.426602.

21. Murata K, Jadhav U, Madha S, et al. Ascl2-Dependent Cell Dedifferentiation Drives Regeneration of Ablated Intestinal Stem Cells. Cell Stem Cell 2020;26:377-390.e6.

22. Jadhav U, Saxena M, O'Neill NK, et al. Dynamic Reorganization of Chromatin Accessibility Signatures during Dedifferentiation of Secretory Precursors into Lgr5+ Intestinal Stem Cells. Cell Stem Cell 2017;21:65-77.e5.

23. Liu Y, Chen Y-G. Intestinal epithelial plasticity and regeneration via cell dedifferentiation. Cell Regen 2020;9:14.

24. Chen B, McKinley ET, Simmons AJ, et al. Human colorectal pre-cancer atlas identifies distinct molecular programs underlying two major subclasses of pre-malignant tumors. Biorxiv 2021:2021.01.11.426044.

25. Steinauer N, Guo C, Zhang J. Emerging Roles of MTG16 in Cell-Fate Control of Hematopoietic Stem Cells and Cancer. Stem Cells Int 2017;2017:1--12. Available at: https://www.hindawi.com/journals/sci/2017/6301385/.

26. Chyla BJ, Moreno-Miralles I, Steapleton MA, et al. Deletion of Mtg16, a Target of t(16;21), Alters Hematopoietic Progenitor Cell Proliferation and Lineage Allocation ▾ †. Mol Cell Biol 2008;28:6234–6247.

27. Cai Y, Xu Z, Xie J, et al. Eto2/MTG16 and MTGR1 are heteromeric corepressors of the TAL1/SCL transcription factor in murine erythroid progenitors. Biochem Bioph Res Co 2009;390:295--301. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19799863 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2774815 http://linkinghub.elsevier.com/retrieve/pii/S0006291X09019287.

28. Ghosh HS, Liu K, Hiebert S, et al. Eto2/MTG16 Regulates E-Protein Activity and Subset Specification in Dendritic Cell Development. Blood 2012;120:1229–1229.

29. Engel ME, Nguyen HN, Mariotti J, et al. Myeloid translocation gene 16 (MTG16) interacts with Notch transcription complex components to integrate Notch signaling in hematopoietic cell fate specification. Mol Cell Biol 2010;30:1852--63. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20123979 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2838074.

30. Gelmetti V, Zhang J, Fanelli M, et al. Aberrant recruitment of the nuclear receptor corepressor-histone deacetylase complex by the acute myeloid leukemia fusion partner ETO. Mol Cell Biol 1998;18:7185--91. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9819405 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC109300.

31. Wang J, Hoshino T, Redner RL, et al. ETO, fusion partner in t(8;21) acute myeloid leukemia, represses transcription by interaction with the human N-CoR/mSin3/HDAC1 complex. Proc National Acad Sci 1998;95:10860--10865. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.95.18.10860.

32. Moore AC, Amann JM, Williams CS, et al. Myeloid Translocation Gene Family Members Associate with T-Cell Factors (TCFs) and Influence TCF-Dependent Transcription. Mol Cell Biol 2007;28:977–987. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18039847 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2223385.

33. Baulies A, Angelis N, Foglizzo V, et al. The Transcription co-Repressors MTG8 and MTG16 Regulate Exit of Intestinal Stem Cells From Their Niche and Differentiation into Enterocyte vs Secretory Lineages. Gastroenterology 2020;159:1328-1341.e3.

34. Williams CS, Bradley AM, Chaturvedi R, et al. MTG16 contributes to colonic epithelial integrity in experimental colitis. Gut 2013;62:1446--55. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22833394 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3663894.

35. McDonough EM, Barrett CW, Parang B, et al. MTG16 is a tumor suppressor in colitis-associated carcinoma. Jci Insight 2017;2. Available at: http://www.ncbi.nlm.nih.gov/pubmed/28814670 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5621889.

36. Short SP, Barrett CW, Stengel KR, et al. Kaiso is required for MTG16-dependent effects on colitis-associated carcinoma. Oncogene 2019;38:5091–5106.

37. Hunt A, Fischer M, Engel ME, et al. Mtg16/Eto2 contributes to murine T-cell development. Mol Cell Biol 2011;31:2544--51. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21536648 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3133377.

38. Acharya P, Sampathi S, Flaherty DK, et al. Mtg16-dependent repression of E protein activity is required for early lymphopoiesis. Biorxiv 2020:2020.07.24.220525.

39. Ghosh HS, Ceribelli M, Matos I, et al. ETO family protein Mtg16 regulates the balance of dendritic cell subsets by repressing Id2. J Exp Medicine 2014;211:1623--35. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24980046 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4113936.

40. Wang L-H, Baker NE. E Proteins and ID Proteins: Helix-Loop-Helix Partners in Development and Disease. Dev Cell 2015;35:269--280. Available at: https://linkinghub.elsevier.com/retrieve/pii/S153458071500684X.

41. Slattery C, Ryan MP, McMorrow T. E2A proteins: Regulators of cell phenotype in normal physiology and disease. Int J Biochem Cell Biology 2008;40:1431–1436.

42. Peng V, Georgescu C, Bakowska A, et al. E proteins orchestrate dynamic transcriptional cascades implicated in the suppression of the differentiation of group 2 innate lymphoid cells. J Biol Chem 2020;295:14866–14877.

43. Zhuang Y, Cheng P, Weintraub H. B-lymphocyte development is regulated by the combined dosage of three basic helix-loop-helix genes, E2A, E2-2, and HEB. Mol Cell Biol 1996;16:2898–2905.

44. Cisse B, Caton ML, Lehner M, et al. Transcription Factor E2-2 Is an Essential and Specific Regulator of Plasmacytoid Dendritic Cell Development. Cell 2008;135:37–48.

45. Beck K, Peak MM, Ota T, et al. Distinct roles for E12 and E47 in B cell specification and the sequential rearrangement of immunoglobulin light chain loci. J Exp Med 2009;206:2271–2284.

46. Bhattacharya A, Baker NE. A Network of Broadly Expressed HLH Genes Regulates Tissue-Specific Cell Fates. Cell 2011;147:881–892.

47. Dréau GL, Escalona R, Fueyo R, et al. E proteins sharpen neurogenesis by modulating proneural bHLH transcription factors' activity in an E-box-dependent manner. Elife 2018;7:e37267.

48. Bouderlique T, Peña-Pérez L, Kharazi S, et al. The Concerted Action of E2-2 and HEB Is Critical for Early Lymphoid Specification. Front Immunol 2019;10:455.

49. Rao C, Malaguti M, Mason JO, et al. The transcription factor E2A drives neural differentiation in pluripotent cells. Development 2020;147:dev.184093.

50. Rozenblatt-Rosen O, Regev A, Oberdoerffer P, et al. The Human Tumor Atlas Network: Charting Tumor Transitions across Space and Time at Single-Cell Resolution. Cell 2020;181:236–249.

51. Herring CA, Banerjee A, McKinley ET, et al. Unsupervised Trajectory Analysis of Single-Cell RNA-Seq and Imaging Data Reveals Alternative Tuft Cell Origins in the Gut. Cell Syst 2018;6:37-51.e9.

52. Liu Q, Herring CA, Sheng Q, et al. Quantitative assessment of cell population diversity in single-cell landscapes. Plos Biol 2018;16:e2006687.

53. Banerjee A, Herring CA, Chen B, et al. Succinate Produced by Intestinal Microbes Promotes Specification of Tuft Cells to Suppress Ileal Inflammation. Gastroenterology 2020;159:2101-2115.e5.

54. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 2002;30:207–210.

55. Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets—update. Nucleic Acids Res 2013;41:D991–D995.

56. Wang F, Flanagan J, Su N, et al. Technical Advance RNAscope A Novel in Situ RNA Analysis Platform for Formalin-Fixed, Paraffin-Embedded Tissues. J Mol Diagnostics 2012;14:22--29. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3338343/pdf/main.pdf.

57. Whittem CG, Williams AD, Williams CS. Murine Colitis Modeling using Dextran Sulfate Sodium (DSS). J Vis Exp Jove 2010:1652.

58. Faustino-Rocha A, Oliveira PA, Pinho-Oliveira J, et al. Estimation of rat mammary tumor volume using caliper and ultrasonography measurements. Lab Animal 2013;42:217–224.

59. Boivin GP, Washington K, Yang K, et al. Pathology of mouse models of intestinal cancer: Consensus report and recommendations. Gastroenterology 2003;124:762–777.

60. Chen S, Zhou Y, Chen Y, et al. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 2018;34:i884–i890.

61. Patro R, Duggal G, Love MI, et al. Salmon: fast and bias-aware quantification of transcript expression using dual-phase inference. Nat Methods 2017;14:417–419.

62. Frankish A, Diekhans M, Ferreira A-M, et al. GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res 2018;47:gky955-.

63. Anon. RNA-seq workflow: gene-level exploratory analysis and differential expression. Available at: https://www.bioconductor.org/packages/devel/workflows/vignettes/rnaseqGene/inst/doc/rnaseqGene.html [Accessed February 5, 2021].

64. Love MI, Soneson C, Hickey PF, et al. Tximeta: Reference sequence checksums for provenance identification in RNA-seq. Plos Comput Biol 2020;16:e1007664.

65. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.

66. H P, M C, S F, et al. Bioconductor - AnnotationDbi. AnnotationDbi: Manipulation of SQLite-based annotations in Bioconductor R package version 1541 2021. Available at: https://bioconductor.org/packages/release/bioc/html/AnnotationDbi.

67. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. P Natl Acad Sci Usa 2005;102:15545–15550.

68. Liberzon A, Subramanian A, Pinchback R, et al. Molecular signatures database (MSigDB) 3.0. Bioinformatics 2011;27:1739–1740.

69. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 2013;14:178–192.

70. Hyams JS, Thomas SD, Gotman N, et al. Clinical and biological predictors of response to standardised paediatric colitis therapy (PROTECT): a multicentre inception cohort study. Lancet 2019;393:1708–1720.

71. Haberman Y, Karns R, Dexheimer PJ, et al. Ulcerative colitis mucosal transcriptomes reveal mitochondriopathy and personalized mechanisms underlying disease severity and treatment response. Nat Commun 2019;10:38.

72. Fenton CG, Taman H, Florholmen J, et al. Transcriptional Signatures That Define Ulcerative Colitis in Remission. Inflamm Bowel Dis 2020;27:94–105.

73. Carey R, Jurickova I, Ballard E, et al. Activation of an IL-6:STAT3-dependent transcriptome in pediatric-onset inflammatory bowel disease. Inflamm Bowel Dis 2008;14:446–457.

74. Chang K, Creighton CJ, Davis C, et al. The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet 2013;45:1113–1120.

75. Hoadley KA, Yau C, Hinoue T, et al. Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. Cell 2018;173:291-304.e6.

76. Rahman M, Jackson LK, Johnson WE, et al. Alternative preprocessing of RNA-Sequencing data in The Cancer Genome Atlas leads to improved analysis results. Bioinformatics 2015;31:3666–3672.

77. Rajamäki K, Taira A, Katainen R, et al. Genetic and epigenetic characteristics of inflammatory bowel disease associated colorectal cancer. Gastroenterology 2021.

78. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47–e47.

79. Goardon N, Lambert JA, Rodriguez P, et al. ETO2 coordinates cellular proliferation and differentiation during erythropoiesis. Embo J 2006;25:357–366.

80. Steinauer N, Guo C, Zhang J. The transcriptional corepressor CBFA2T3 inhibits all-trans-retinoic acid–induced myeloid gene expression and differentiation in acute myeloid leukemia EDITORS' PICK: CBFA2T3 inhibits ATRA-mediated myeloid maturation. J Biol Chem 2020;295:8887–8900.

81. Poindexter SV, Reddy VK, Mittal MK, et al. Transcriptional corepressor MTG16 regulates small intestinal crypt proliferation and crypt regeneration after radiation-induced injury. Am J Physiol-gastr L 2015;308:G562--71. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25573176 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4360050.

82. Gunawardene AR, Corfe BM, Staton CA. Classification and functions of enteroendocrine cells of the lower gastrointestinal tract. Int J Exp Pathol 2011;92:219–231.

83. Akiyama S, Mochizuki W, Nibe Y, et al. CCN3 Expression Marks a Sulfomucin-nonproducing Unique Subset of Colonic Goblet Cells in Mice. Acta Histochem Cytoc 2017;50:17027.

84. Billing LJ, Larraufie P, Lewis J, et al. Single cell transcriptomic profiling of large intestinal enteroendocrine cells in mice – Identification of selective stimuli for insulin-like peptide-5 and glucagon-like peptide-1 co-expressing cells. Mol Metab 2019;29:158–169.

85. Schaum N, Karkanias J, Neff NF, et al. Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. Nature 2018;562:367–372.

86. Gehart H, Es JH van, Hamer K, et al. Identification of Enteroendocrine Regulators by Real-Time Single-Cell Differentiation Mapping. Cell 2019;176:1158-1173.e16.

87. Flier LG van der, Gijn ME van, Hatzis P, et al. Transcription Factor Achaete Scute-Like 2 Controls Intestinal Stem Cell Fate. Cell 2009;136:903–912.

88. Wice BM, Gordon JI. Forced Expression of Id-1 in the Adult Mouse Small Intestinal Epithelium Is Associated with Development of Adenomas*. J Biol Chem 1998;273:25310–25319.

89. O'Brien CA, Kreso A, Ryan P, et al. ID1 and ID3 Regulate the Self-Renewal Capacity of Human Colon Cancer-Initiating Cells through p21. Cancer Cell 2012;21:777–792.

90. Zhang N, Yantiss RK, Nam H, et al. ID1 Is a Functional Marker for Intestinal Stem and Progenitor Cells Required for Normal Response to Injury. Stem Cell Rep 2014;3:716–724.

91. Biyajima K, Kakizaki F, Shen X, et al. Id2 deletion attenuates Apc-deficient ileal tumor formation. Biol Open 2015;4:993–1001.

92. Zhang N, Subbaramaiah K, Yantiss RK, et al. Id1 Deficiency Protects against Tumor Formation in ApcMin/+ Mice but Not in a Mouse Model of Colitis-Associated Colon Cancer. Cancer Prev Res 2015;8:303–311.

93. Rockman SP, Currie SA, Ciavarella M, et al. Id2 Is a Target of the β-Catenin/T Cell Factor Pathway in Colon Carcinoma*. J Biol Chem 2001;276:45113–45119.

94. Khaloian S, Rath E, Hammoudi N, et al. Mitochondrial impairment drives intestinal stem cell transition into dysfunctional Paneth cells predicting Crohn's disease recurrence. Gut 2020;69:1939–1951.

95. Gersemann M, Becker S, Kübler I, et al. Differences in goblet cell differentiation between Crohn's disease and ulcerative colitis. Differentiation 2009;77:84–94.

96. Sünderhauf A, Hicken M, Schlichting H, et al. Loss of mucosal p32/gC1qR/HABP1 triggers energy deficiency and impairs goblet cell differentiation in ulcerative colitis. Cell Mol Gastroenterology Hepatology 2021.

97. Treveil A, Sudhakar P, Matthews ZJ, et al. Regulatory network analysis of Paneth cell and goblet cell enriched gut organoids using transcriptomics approaches. Mol Omics 2019;16:39–58.

98. Perse M, Cerar A. Dextran Sodium Sulphate Colitis Mouse Model: Traps and Tricks. J Biomed Biotechnol 2012;2012:1--13. Available at: http://www.hindawi.com/journals/bmri/2012/718617/.

99. Dieleman, Palmen, Akol, et al. Chronic experimental colitis induced by dextran sulphate sodium (DSS) is characterized by Th1 and Th2 cytokines. Clin Exp Immunol 1998;114:385–391.

100. Rajamäki K, Taira A, Katainen R, et al. Genetic and epigenetic characteristics of inflammatory bowel disease associated colorectal cancer. Gastroenterology 2021.

101. Andrews N, Serviss JT, Geyer N, et al. An unsupervised method for physical cell interaction profiling of complex tissues. Nat Methods 2021;18:912–920.

102. Schonhoff SE, Giel-Moloney M, Leiter AB. Neurogenin 3-expressing progenitor cells in the gastrointestinal tract differentiate into both endocrine and non-endocrine cell types. Dev Biol 2004;270:443–454.

103. Wang Y, Song W, Wang J, et al. Single-cell transcriptome analysis reveals differential nutrient absorption functions in human intestine. J Exp Med 2019;217:e20191130.

104. Jenny M, Uhl C, Roche C, et al. Neurogenin3 is differentially required for endocrine cell fate specification in the intestinal and gastric epithelium. Embo J 2002;21:6338–6347.

105. López-Díaz L, Jain RN, Keeley TM, et al. Intestinal Neurogenin 3 directs differentiation of a bipotential secretory progenitor to endocrine cell rather than goblet cell fate. Dev Biol 2007;309:298–305.

106. Li HJ, Ray SK, Kucukural A, et al. Reduced Neurog3 Gene Dosage Shifts Enteroendocrine Progenitor Towards Goblet Cell Lineage in the Mouse Intestine. Cell Mol Gastroenterology Hepatology 2020.

107. Li HJ, Ray SK, Singh NK, et al. Basic helix-loop-helix transcription factors and enteroendocrine cell differentiation. Diabetes Obes Metabolism 2011;13:5–12.

108. Hartenstein V, Takashima S, Hartenstein P, et al. bHLH proneural genes as cell fate determinants of entero-endocrine cells, an evolutionarily conserved lineage sharing a common root with sensory neurons. Dev Biol 2017;431:36–47.

109. Huang H-P, Liu M, El-Hodiri HM, et al. Regulation of the Pancreatic Islet-Specific GeneBETA2 (neuroD) by Neurogenin 3. Mol Cell Biol 2000;20:3292–3307.

110. Longo A, Guanga GP, Rose RB. Crystal Structure of E47−NeuroD1/Beta2 bHLH Domain−DNA Complex: Heterodimer Selectivity and DNA Recognition † , ‡. Biochemistry-us 2008;47:218–229.

111. Fujiwara T, O'Geen H, Keles S, et al. Discovering Hematopoietic Mechanisms through Genome-wide Analysis of GATA Factor Chromatin Occupancy. Mol Cell 2009;36:667–681.

112. Yang Y, Mlodzik M. Wnt-Frizzled/Planar Cell Polarity Signaling: Cellular Orientation by Facing the Wind (Wnt). Annu Rev Cell Dev Bi 2015;31:623–646.

113. Böttcher A, Büttner M, Tritschler S, et al. Non-canonical Wnt/PCP signalling regulates intestinal stem cell lineage priming towards enteroendocrine and Paneth cell fates. Nat Cell Biol 2021;23:23–31.

114. Rees WD, Tandun R, Yau E, et al. Regenerative Intestinal Stem Cells Induced by Acute and Chronic Injury: The Saving Grace of the Epithelium? Frontiers Cell Dev Biology 2020;8:583919.

115. Shivdasani RA, Clevers H, Sauvage FJ de. Tissue regeneration: Reserve or reverse? Science 2021;371:784–786.

116. Ayyaz A, Kumar S, Sangiorgi B, et al. Single-cell transcriptomes of the regenerating intestine reveal a revival stem cell. Nature 2019;569:121–125.

117. Beumer J, Clevers H. Regulation and plasticity of intestinal stem cells during homeostasis and regeneration. Development 2016;143:3639–3649.

118. Vega PN, Lau KS, Goldenring JR. Not All Insults Are Created Equal for Awakening Dormant Stem Cell Abilities. Cell Mol Gastroenterology Hepatology 2019;7:619–621.

119. Leedham SJ. Reserving the Right to Change the Intestinal Stem Cell Model. Cell Stem Cell 2020;26:301–302.

120. Yui S, Azzolin L, Maimets M, et al. YAP/TAZ-Dependent Reprogramming of Colonic Epithelium Links ECM Remodeling to Tissue Regeneration. Cell Stem Cell 2018;22:35-49.e7.

121. Qu M, Xiong L, Lyu Y, et al. Establishment of intestinal organoid cultures modeling injury-associated epithelial regeneration. Cell Res 2021:1–13.

122. Lee C-C, Chen W-S, Chen C-C, et al. TCF12 Protein Functions as Transcriptional Repressor of E-cadherin, and Its Overexpression Is Correlated with Metastasis of Colorectal Cancer*. J Biol Chem 2012;287:2798–2809.

123. Zhao H, Huang A, Li P, et al. E2A suppresses invasion and migration by targeting YAP in colorectal cancer cells. J Transl Med 2013;11:317.