

# HMGA1 acts as an epigenetic gatekeeper of ASCL2 and Wnt signaling during colon tumorigenesis

Li Z. Luo,<sup>1</sup> Jung-Hyun Kim,<sup>1,2</sup> Iliana Herrera,<sup>1</sup> Shaoguang Wu,<sup>3</sup> Xinqun Wu,<sup>3</sup> Seong-Sik Park,<sup>2</sup> Juyoung Cho,<sup>2</sup> Leslie Cope,<sup>4</sup> Lingling Xian,<sup>1</sup> Bailey E. West,<sup>1,5</sup> Julian Calderon-Espinosa,<sup>1,6</sup> Joseph Kim,<sup>1</sup> Zanshé Thompson,<sup>1</sup> Isha Maloo,<sup>1,7</sup> Tatianna Larman,<sup>8</sup> Karen L. Reddy,<sup>9</sup> Ying Feng,<sup>10</sup> Eric R. Fearon,<sup>10</sup> Cynthia L. Sears,<sup>3,11,12</sup> and Linda Resar<sup>1,4,5,6,7,8,11</sup>

<sup>1</sup>Division of Hematology, Department of Medicine, the Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. <sup>2</sup>Research Institute, National Cancer Center, Goyang-si, Gyeonggi-do, Republic of Korea. <sup>3</sup>Division of Infectious Diseases, Department of Medicine, <sup>4</sup>Sidney Kimmel Comprehensive Cancer Center, Division of Biostatistics, <sup>5</sup>Pathobiology Graduate Program, Department of Pathology, and <sup>6</sup>Human Genetics Graduate Program, Department of Genetics and Molecular Medicine, the Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. <sup>7</sup>Biochemistry and Molecular Biology Program, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA. <sup>8</sup>Department of Pathology and <sup>9</sup>Department of Biological Chemistry, the Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. <sup>10</sup>Department of Oncology, University of Michigan, Ann Arbor, Michigan, USA. <sup>11</sup>Department of Oncology, Sidney Kimmel Comprehensive Cancer Center, and <sup>12</sup>Molecular Immunology, the Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

**Mutated tumor cells undergo changes in chromatin accessibility and gene expression, resulting in aberrant proliferation and differentiation, although how this occurs is unclear. HMGA1 chromatin regulators are abundant in stem cells and oncogenic in diverse tissues; however, their role in colon tumorigenesis is only beginning to emerge. Here, we uncover a previously unknown epigenetic program whereby HMGA1 amplifies Wnt signaling during colon tumorigenesis driven by inflammatory microbiota and/or *Adenomatous polyposis coli* (*Apc*) inactivation. Mechanistically, HMGA1 “opens” chromatin to upregulate the stem cell regulator, *Ascl2*, and downstream Wnt effectors, promoting stem and Paneth-like cell states while depleting differentiated enterocytes. Loss of just one *Hmga1* allele within colon epithelium restrains tumorigenesis and Wnt signaling driven by mutant *Apc* and inflammatory microbiota. However, HMGA1 deficiency has minimal effects in colon epithelium under homeostatic conditions. In human colon cancer cells, HMGA1 directly induces ASCL2 by recruiting activating histone marks. Silencing HMGA1 disrupts oncogenic properties, whereas reexpression of ASCL2 partially rescues these phenotypes. Further, HMGA1 and ASCL2 are coexpressed and upregulated in human colorectal cancer. Together, our results establish HMGA1 as an epigenetic gatekeeper of Wnt signals and cell state under conditions of APC inactivation, illuminating HMGA1 as a potential therapeutic target in colon cancer.**

## Introduction

Exquisite control of transcriptional networks that regulate plasticity and other stem cell properties allow for tissue specification during embryogenesis and tissue maintenance after birth (1–9). Often referred to as “stemness” networks, genes involved in self-renewal and plasticity are largely silenced in differentiated tissues, although a subset remain active in adult stem cells where they contribute to tissue regeneration during homeostatic conditions or following injury and other stressors (3–6). By contrast, neoplastic cells frequently corrupt these same transcriptional networks to foster aberrant growth and differentiation (1–7, 9, 10). Moreover, tumor progression is associated with increased expression of genes controlling

stem cell properties, which may foster the emergence of highly plastic tumor cells capable of metastatic progression, immune evasion, and therapy resistance (1–7, 9, 10). While mechanisms responsible for cell state during tissue regeneration and tumorigenesis remain incompletely understood, it is clear that chromatin reorganization and epigenetic alterations contribute to plasticity, self renewal, and other stem cell properties (4, 11).

As a highly regenerative tissue and frequent site for cancer, the colon epithelium provides a unique paradigm to study plasticity and cell state during tumorigenesis. Colon epithelium comprises an intricately organized hierarchy of epithelial cells maintained by proliferative stem cells that reside at the base of crypts (3, 4, 6–8). Moreover, it is among the most regenerative tissues of the body, renewing itself every 3–5 days to maintain nutrient absorption essential for life and provide a protective barrier from gut pathogens and toxins. Stem cells at the crypt base in colon and small intestinal epithelium are marked by the serpentine coreceptor for Wnt signals, leucine-rich, repeat-containing G-protein-coupled receptor 5 (LGR5) (7, 8). Recent studies in murine small intestine also identified distinct populations of upper crypt cells, marked by fibroblast growth factor binding protein 1 (FGFBP1) or LGR4, that regenerate all lineages, including LGR5+ cells, the latter of which requires the Wnt agonist, R-spondin (12–14). In both small intestine and

► **Related Commentary:** <https://doi.org/10.1172/JCI187442>

**Conflict of interest:** CS reports grant funding administered through her institution from Janssen and Bristol Myers Squibb. She reports unrelated royalties from Up to Date. LR reports investigator-initiated grant funding administered through her institution from PharmaEssentia for an unrelated project.

**Copyright:** © 2025, Luo et al. This is an open access article published under the terms of the Creative Commons Attribution 4.0 International License.

**Submitted:** July 11, 2024; **Accepted:** November 27, 2024; **Published:** February 3, 2025.

**Reference information:** J Clin Invest. 2025;135(3):e184442.

<https://doi.org/10.1172/JCI184442>.

colon, a Wnt gradient maintains LGR5+ cells by repressing differentiation at the base where Wnt levels are highest, while allowing cells to differentiate as they move up the crypt with decreasing Wnt. Thus, tightly regulated Wnt signaling is fundamental to epithelial regeneration in the gut.

Not surprisingly, mutations that activate Wnt signals are common in colon adenomas and adenocarcinomas (15–22). Inactivating mutations in the gene encoding the Adenomatous Polyposis Coli (APC) tumor suppressor protein, first described in the familial adenomatous polyposis (FAP) syndrome, are the most common genetic lesions found in colon adenomas and carcinomas (15–22). APC normally restrains Wnt function by maintaining  $\beta$ -catenin in the cytoplasm, thereby preventing  $\beta$ -catenin entry into the nucleus to activate Wnt target genes together with the TCF-4 transcription factor. Genomic studies established a model whereby colon carcinomas develop from polyps harboring APC mutations after the stepwise accumulation of mutations that inactivate additional tumor suppressor genes and/or activate protooncogenes (15–16). Neoplastic polyps are thought to arise in colon stem cells expressing *LGR5*, although targeting mutated LGR5+ cells in therapy has not been feasible (7, 8). Importantly, colon cancer is the third leading cause of cancer-related deaths in the US, and the incidence is rising globally, particularly in younger individuals (23–26). Thus, studies are warranted to decipher changes in cell state occurring during colon tumorigenesis to identify mechanisms that could be targeted to intercept the transition from mutant cells and localized tumors to advanced disease.

High Mobility Group A (HMGA1) proteins are architectural transcription factors that bind AT-rich sequences where they modulate chromatin structure and gene expression (4, 27–55). The *HMGA1* gene is highly expressed during embryogenesis (4, 9, 34, 37) and in adult stem cells (4, 29, 30, 47), but is silenced in most differentiated cells. *HMGA1* becomes reexpressed in aggressive cancer cells and high levels portend adverse clinical outcomes (28–38, 40–43, 56–64). In colon cancer, *HMGA1* is among the genes most highly overexpressed compared with nonmalignant colon epithelium (4, 30, 57, 62). While mechanisms upregulating *HMGA1* in cancer are incompletely understood, increasing evidence suggests that diverse oncogenic pathways, including growth factors (65, 66), mutations — such as mutant *Apc* (67, 68), *KRAS* (28, 36, 69), or mutant *JAK2* (29) — and oncogenic transcription factors, like cMYC (70) or cJUN (71, 72), converge on *HMGA1* to induce its expression in distinct settings. In transgenic mouse models, *Hmga1* overexpression leads to tumorigenesis (35, 41, 42, 58, 64). For example, transgenic mice overexpressing *Hmga1* in lymphoid cells develop clonal expansion with evolution to leukemia by upregulating transcriptional networks active in proliferating stem cells, poorly differentiated cancer cells, and inflammation (37, 38, 41, 58, 64). In experimental models of pancreatic cancer and myeloproliferative neoplasms, HMGA1 activates gene networks in a cell-intrinsic fashion to drive aberrant proliferation and differentiation, while inducing signals within the tumor microenvironment that promote fibrosis, culminating in tumor progression (28, 29). HMGA1 also upregulates genes involved in an epithelial-to-mesenchymal transition in colon cancer cell lines (35). In small intestinal stem cells, HMGA1 induces *Sox9* and Wnt signals from the stromal and epithelial niches to maintain the stem cell compartment and Paneth

cell niche during epithelial regeneration (4, 30). However, HMGA1 function in colon stem and progenitor cells during tumorigenesis was previously unknown.

Here, we uncover a previously unknown role for HMGA1 in modulating transcriptional networks to expand LGR5+ stem cells and Paneth-like cells during tumorigenesis driven by *Apc* deficiency. Strikingly, loss of just a single *Hmga1* allele disrupts tumorigenesis while prolonging survival in two different models of colon tumorigenesis with APC inactivation, including mice with biallelic deletion of colonic epithelial *Apc* (*CDX2P-CreER<sup>T2</sup>Apc<sup>fl/fl</sup>*) (73, 74) and mice harboring monoallelic mutant *Apc* (*Apc<sup>Min/-</sup>* or Min mice) colonized with the inflammatory human symbiote, enterotoxigenic *Bacteroides fragilis* (ETBF) (75–79). Single cell RNA sequencing (scRNA-seq) in *Apc*-deficient crypt epithelium reveals that HMGA1 maintains colon crypt cells in a stem and Paneth-like cell state while depleting differentiated enterocytes. Integration of transcriptomic analyses with assays of chromatin accessibility demonstrate that HMGA1 activates Wnt signals by “opening” chromatin at gene loci governing Wnt signaling, including the stem cell regulator, *Achaete-Scute Family BHLH Transcription Factor 2* (*Ascl2*), in addition to Wnt agonist receptors (*Lgr5*, *Lrp5*) and downstream effectors. We focus on the gene encoding ASCL2 as a master regulator of cell fate in the small intestine, although its role in the colon was previously unknown. In human colon cancer cells, HMGA1 directly induces *ASCL2* by recruiting activating histone marks. Further, silencing *HMGA1* disrupts oncogenic properties (proliferation and clonogenicity), while reexpression of *ASCL2* partially rescues oncogenic phenotypes in HMGA1-depleted human colon cancer cells. Most importantly, both *HMGA1* and *ASCL2* are coexpressed and upregulated in human colon cancer. Surprisingly, HMGA1 depletion has minimal effects on colon epithelial regeneration under homeostatic conditions. Our results establish HMGA1 as an epigenetic gatekeeper of *ASCL2* and Wnt signals in colon stem cells during tumorigenesis, but not steady state homeostasis, highlighting HMGA1 pathways as promising therapeutic targets for colon carcinogenesis.

## Results

*Loss of just a single Hmga1 allele is sufficient to decrease tumorigenesis and prolong survival in mice with colon tumors driven by biallelic Apc inactivation.* Because previous studies from our group and others showed that *HMGA1* is highly overexpressed in human colon cancer (30, 35, 62) and required for oncogenic properties in colon cancer cell lines (35), we sought to assess its role in colon tumorigenesis in vivo. Since *APC* is the most commonly mutated gene in human colon cancer (15), we examined *CDX2P-CreER<sup>T2</sup>Apc<sup>fl/fl</sup>* mice, an established model of colon tumorigenesis caused by inducible, biallelic loss of *Apc* within colon epithelium (73, 74). *CDX2P-CreER<sup>T2</sup>Apc<sup>fl/fl</sup>* mice were crossed to mice with global *Hmga1* deficiency (heterozygous or homozygous). Importantly, mice with heterozygous *Hmga1* have normal development and lifespans, whereas those with homozygous deficiency have partial embryonic lethality and develop premature aging phenotypes (kyphosis, bone loss, greying, and shortened lifespans) beginning after 10–12 months of age (29, 56). As expected, *CDX2P-CreER<sup>T2</sup>Apc<sup>fl/fl</sup>* mice with *Hmga1* heterozygous or homozygous deficiency have lower *Hmga1* gene expression and protein levels in colon epithelium (Supplemental Figure 1, A and B; supplemental material available

online with this article; <https://doi.org/10.1172/JCI184442DS1>). Following induction of Cre recombinase-mediated *Apc* deletion by tamoxifen (TAM), *CDX2P-CreER<sup>T2</sup>Apc<sup>fl/fl</sup>* mice with intact *Hmga1* alleles develop epithelial hyperplasia in the cecum, proximal, and midcolon regions (Figure 1A) associated with weight loss by 4 weeks (Figure 1B) and decreased survival (median survival 43 days after TAM;  $n = 16$ ) (Figure 1C). Strikingly, loss of just a single *Hmga1* allele in *CDX2P-CreER<sup>T2</sup>Apc<sup>fl/fl</sup>* mice mitigates weight loss while prolonging survival (median survival 61.5 days;  $P < 0.0001$ ,  $n = 12$ ; Figure 1, A–C). Survival is prolonged further (median survival 78 days;  $P < 0.0001$ ,  $n = 7$ ) (Figure 1C) in mice with *Hmga1* homozygous deficiency, indicating that *Hmga1* gene dosage modulates tumor progression in this model.

To determine more precisely how HMGA1 modulates tumorigenesis in *CDX2P-CreER<sup>T2</sup>Apc<sup>fl/fl</sup>* mice, we compared colon weights as a surrogate for tumor burden since extensive, contiguous tumors in the proximal colon precludes precise enumeration. Both the absolute and relative colon weight (% colon weight/body weight) increase in *CDX2P-CreER<sup>T2</sup>Apc<sup>fl/fl</sup>* mice with intact *Hmga1* compared with those with heterozygous or homozygous *Hmga1* deficiency (Figure 1D and Supplemental Figure 1C). At 21 days following *Apc* inactivation, colon epithelium becomes thickened and dysplastic with extensive adenomatous changes and increased crypt depth in *CDX2P-CreER<sup>T2</sup>Apc<sup>fl/fl</sup>* mice with intact *Hmga1* (Figure 1, A and E, and Figure 2, A and B). Intranuclear HMGA1 is prominent throughout the crypts up to the luminal enterocytes in *CDX2P-CreER<sup>T2</sup>Apc<sup>fl/fl</sup>* mice with intact *Hmga1* by IHC. By contrast, intranuclear HMGA1 is normally restricted to the crypt bases in WT mice lacking *Apc*-deficiency (Supplemental Figure 1D). Since APC restrains Wnt signals by sequestering  $\beta$ -catenin within inhibitory, cytoplasmic complexes, we compared  $\beta$ -catenin levels and localization in nuclei and cytoplasm in the *Apc*-deficient models. Following *Apc* inactivation (day 21), both nuclear and cytoplasmic  $\beta$ -catenin levels increase in colon epithelial cells of mice with intact HMGA1 compared with those with heterozygous or homozygous *Hmga1* deficiency, paralleling the distribution of intranuclear HMGA1 (Figure 2, A and B, and Supplemental Figure 1E). Cells staining positive for the proliferation marker Ki67 are also increased in *CDX2P-CreER<sup>T2</sup>Apc<sup>fl/fl</sup>* mice with intact HMGA1 compared with those with HMGA1 deficiency early in tumorigenesis, although they predominate at the crypt bases (Figure 2, A and B). In addition, HMGA1 increases in the adenomatous epithelium compared with nontumor, midcolon epithelium in mice with intact HMGA1 (Figure 3, A and B, and Supplemental Figure 1, A and B). Intriguingly, HMGA1 also increases in tumors from mice with *Apc* inactivation and *Hmga1* heterozygous deficiency compared with adjacent nontumor colon epithelium (Figure 3, A and B, and Supplemental Figure 1, A and B). These findings indicate that loss of just a single *Hmga1* allele in the setting to *Apc* inactivation decreases hyperproliferation,  $\beta$ -catenin levels, and tumorigenesis.

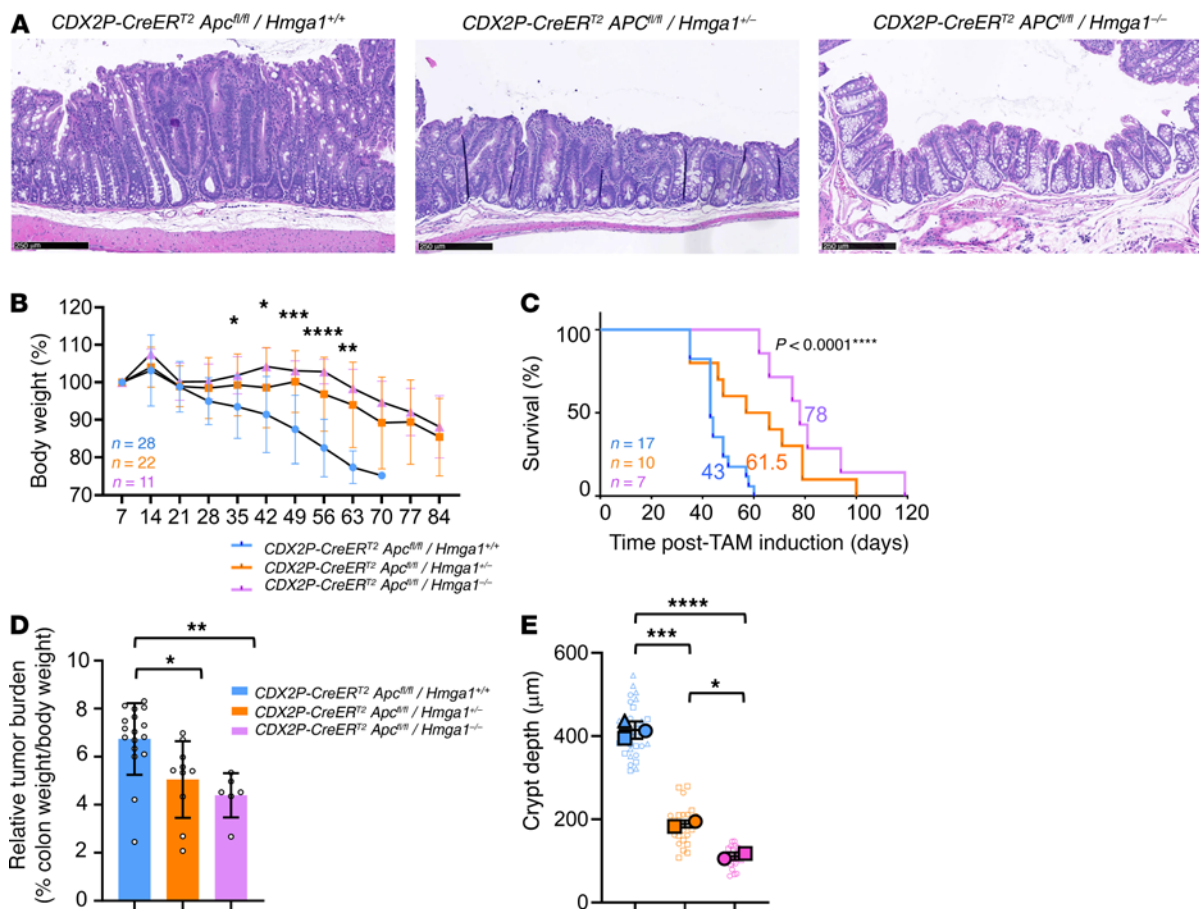
*HMGA1 hemizygous deficiency mitigates colon tumorigenesis induced by ETBF in Min mice.* Next, we investigated HMGA1 function in the multiple intestinal neoplasia (Min) model, which harbors a heterozygous *Apc* loss-of-function mutation (*Apc<sup>Min/+</sup>* or *Min<sup>+/-</sup>*) and develops distal colon tumors following inoculation with the human symbiotic bacterium ETBF. This model recapitulates salient features of human colon tumors with respect to the mutational status, location in the distal colon, and histopathology (75–79).

Furthermore, ETBF colonization is common in colon cancer (up to 90%) and epidemiologic studies suggest that it increases the risk of carcinogenesis (75–85). Following inoculation with ETBF at 5–6 weeks of age, Min mice with intact *Hmga1* exhibit poor weight gain and robust distal colon tumorigenesis by 11–12 weeks with a median survival of 17 weeks; by contrast, Min mice with global *Hmga1* hemizygosity gain more weight, develop fewer tumors, and exhibit prolonged survival (Figure 4, A–D). Histologic examination shows hyperproliferative colon epithelium and adenomatosis with increased crypt depth in the distal colons of Min mice with intact HMGA1 compared with those with HMGA1 deficiency (Figure 4, E and F). As expected, HMGA1 mRNA and protein levels by IHC are increased in mice with intact HMGA1 compared with *Hmga1* haploinsufficient Min mice (Figure 4, E and F, Figure 5, A–C, and Supplemental Figure 2A). Similarly, the proportion of cells staining positive for intranuclear  $\beta$ -catenin and cytoplasmic  $\beta$ -catenin are greater in colon epithelium and tumors of *Apc<sup>Min/+</sup>* mice with intact HMGA1, although Ki67 was unchanged in mice with intact or haploinsufficient HMGA1 (Figure 4, E and F, and Supplemental Figure 2B). Intriguingly, HMGA1 protein staining is similar in tumors from *Apc<sup>Min/+</sup>* mice with intact HMGA1 and HMGA1 haploinsufficiency, suggesting that mice with HMGA1 haploinsufficiency can upregulate the intact *Hmga1* allele to increase HMGA1 levels within their tumors (Figure 5, A–C). The small intestinal tumor burden is also greater in *Apc<sup>Min/+</sup>* mice with intact HMGA1 (Supplemental Figure 2C), indicating that HMGA1 contributes to tumorigenesis in both the colon and small intestine of Min mice.

*Loss of a single Hmga1 allele within the colon epithelium is sufficient to reduce colon tumorigenesis induced by ETBF in Min mice.* To determine whether HMGA1 deficiency within colon epithelium is sufficient to mitigate tumorigenesis in the Min-ETBF model, we generated Min mice with *Hmga1* deficiency (hetero- and homozygous genetic deletion) restricted to colon and small intestinal epithelium by crossing *Apc<sup>Min/+</sup>* mice with *Hmga1<sup>fl/fl</sup>* mice on a *Vil-cre* background. Notably, Min mice with tissue-specific HMGA1 deficiency developed fewer colon tumors and decreased crypt depth compared with Min mice with intact HMGA1 (Figure 6, A–E). Surprisingly, Min mice with *Hmga1* haploinsufficiency had a similar decrease in colon tumor number as Min mice with homozygous *Hmga1* loss, suggesting that a relatively modest decrease in HMGA1 within the epithelial compartment alone is sufficient to mitigate tumorigenesis. Small intestinal tumors also decrease modestly in this model, but only with homozygous loss of *Hmga1* (Supplemental Figure 3A). By contrast, tissue-specific biallelic loss of *Hmga1* in colon crypts from WT mice lacking *Apc* mutation show no significant changes in crypt depth, suggesting that HMGA1 deficiency under steady state, homeostatic conditions (no ETBF colonization) is not deleterious to colon epithelial regeneration (Supplemental Figure 3B). Together, these findings demonstrate that HMGA1 within the crypt epithelium drives tumorigenesis, and, moreover, tissue-specific, *Hmga1* haploinsufficiency is sufficient to impair colon tumor formation driven by mutant *Apc* and inflammatory ETBF, highlighting HMGA1 as a promising potential therapeutic target.

*HMGA1 expands colon stem cells and Paneth-like cells while depleting more differentiated enterocytes in Apc-deficient colon crypts.* To investigate molecular mechanisms underlying HMGA1 in *Apc*-deficient



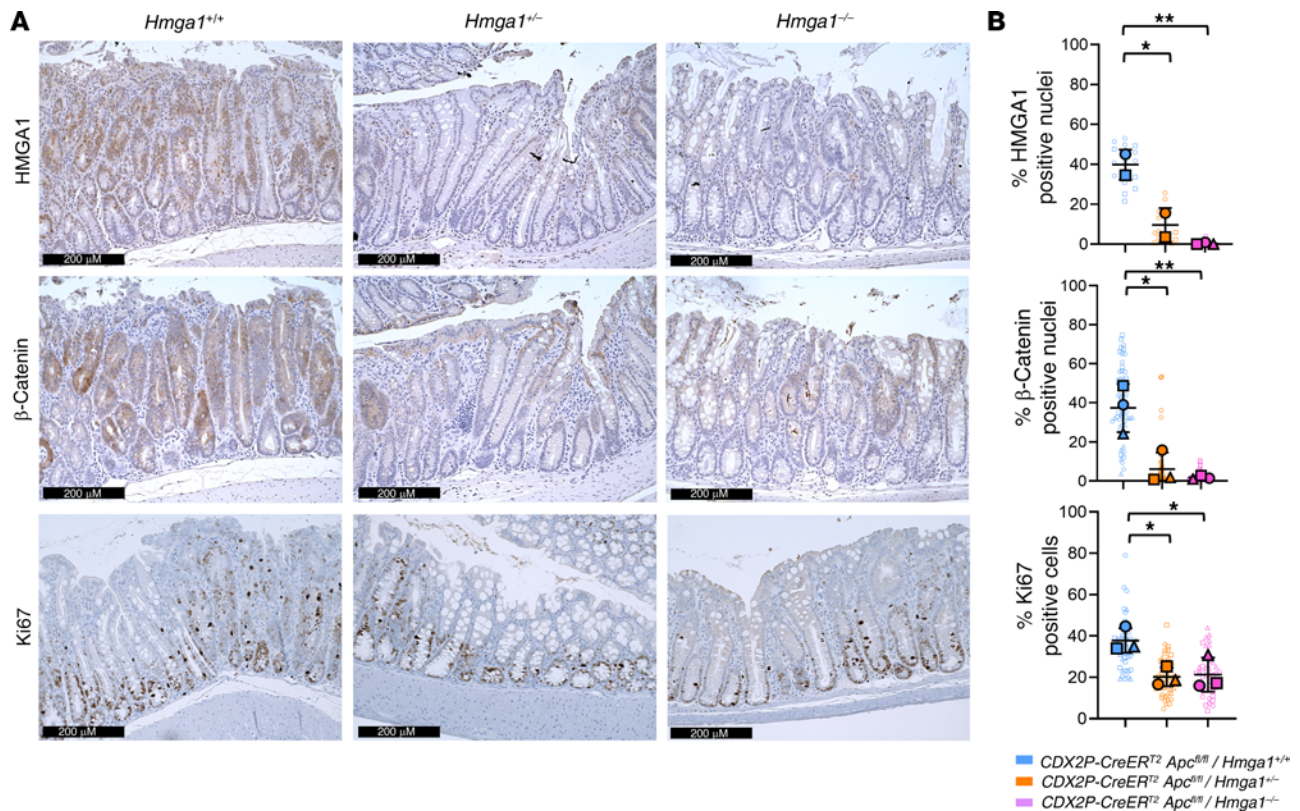


**Figure 1. Loss of a single *Hmga1* allele mitigates colon tumorigenesis and prolongs survival in *CDX2P-CreER<sup>T2</sup>/Apc<sup>fl/fl</sup>* mice. (A)** Representative images (H&E) of proximal colon in *CDX2P-CreER<sup>T2</sup>/Apc<sup>fl/fl</sup>* mice with *Hmga1* intact (*Hmga1<sup>+/+</sup>*, top), heterozygous deletion (*Hmga1<sup>+/-</sup>*, middle), or homozygous deletion (*Hmga1<sup>-/-</sup>*, bottom) at survival endpoint necropsy (top: day 35 after TAM; middle: day 57; bottom: day 81). Scale bars: 250  $\mu$ m. **(B)** Relative weight changes in *CDX2P-CreER<sup>T2</sup>/Apc<sup>fl/fl</sup>* models after TAM. (\* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ , 1-way ANOVA). **(C)** Kaplan-Meier plot showing survival in *CDX2P-CreER<sup>T2</sup>/Apc<sup>fl/fl</sup>* mice with *Hmga1<sup>+/+</sup>*, *Hmga1<sup>+/-</sup>*, or *Hmga1<sup>-/-</sup>*. (\*\*\*\* $P < 0.0001$ ; Mantel-Cox test). **(D)** Relative colon weight to body weight in *CDX2P-CreER<sup>T2</sup>/Apc<sup>fl/fl</sup>* mice with *Hmga1<sup>+/+</sup>*, *Hmga1<sup>+/-</sup>*, or *Hmga1<sup>-/-</sup>* (\* $P < 0.05$ ; *Hmga1<sup>+/+</sup>* versus *Hmga1<sup>-/-</sup>*, \*\* $P < 0.01$ , *Hmga1<sup>+/+</sup>* versus *Hmga1<sup>+/-</sup>*; Tukey's multiple comparison test following significance by 1-way ANOVA). **(E)** Proximal colon crypt depth in *CDX2P-CreER<sup>T2</sup>/Apc<sup>fl/fl</sup>* models (\* $P < 0.05$ , \*\*\*\* $P < 0.0001$ ; 1-way ANOVA with Tukey's multiple comparison test). Each shape (circle, square, triangle) corresponds to a different mouse ( $n = 2-3$ /genotype). The solid shapes show the mean from each mouse; the open, smaller shapes represent individual measurements/crypt (range = 9-13 crypts/mouse) at  $\times 20$  magnification.

colon crypts, we performed scRNA-seq in proximal colon crypt cells from *CDX2P-CreER<sup>T2</sup>/Apc<sup>fl/fl</sup>* mice with intact *Hmga1* compared with those with heterozygous or homozygous *Hmga1* deficiency. We examined transcriptomes at early stages in tumorigenesis (21 days following *Apc* inactivation via TAM) to identify mechanisms involved in tumor initiation. Single-cell transcriptomes depicted by uniform manifold approximation and projection (UMAP) reveal differences in overall distribution in cells with or without HMGA1 (Figure 7A). Unsupervised hierarchical clustering of transcripts (via Seurat) revealed 12 clusters (Figure 7B) from which cell identities were imputed using established markers (Supplemental Table 1). Of these clusters, five are comprised of epithelial crypt cells (denoted epithelial island) based on expression of the colon epithelial cell adhesion marker gene (*Epcam*), colon stem and progenitor cell genes (*Lgr5*, *Sox9*, *Ctnnb1*), and proximity by UMAP (Figure 7C). The remaining clusters are comprised of immune cells (Figure 7B). Within the immune cell islands, we identified  $Cd4^+$  and  $Cd8^+$  T cells with smaller populations of B cells, macrophages, and other

myeloid lineages (macrophage-like, mast cells, and neutrophils) (Supplemental Figure 4A).

To dissect HMGA1-dependent changes in the cell of origin for colon tumors, we focused on the epithelial island. With intact HMGA1, the  $Lgr5^+$  stem cell population, defined by high levels of *Lgr5*, *Msi1*, *Bmi1*, and other stem cell transcripts (Supplemental Table 1), comprise the majority of cells (40.1%) within *CDX2P-CreER<sup>T2</sup>/Apc<sup>fl/fl</sup>* colon crypts (Figure 7D). Transit amplifying (TA) cells are the next most abundant population, constituting 26% of epithelial crypt cells, whereas Paneth-like cells, based on Paneth cell markers (*Lyz1*, *Mmp7*, *Sox9*, *Retnlb*, *Chil3*, *Reg3g*, and *Deta*; Supplemental Table 1), comprise 20% of crypt cells. Intriguingly, while Paneth cells are not present in normal colon epithelium, Paneth cell "metaplasia" has been reported in proximal colon epithelium in adenomas, adenocarcinoma, and inflammatory bowel disease, and ectopic Paneth cells were observed in colon epithelium of *Apc*-deficient mouse models (73, 74, 85, 86). The terminally differentiated enterocyte (EC) and goblet cell



**Figure 2.** *Hmga1* haploinsufficiency decreases  $\beta$ -catenin and Ki67 in *CDX2P-CreER<sup>T2</sup>/Apc<sup>fl/fl</sup>* mice. (A) Representative IHC images of nuclear HMGA1 (top),  $\beta$ -catenin (middle), and Ki67 (bottom) in *CDX2P-CreER<sup>T2</sup>/Apc<sup>fl/fl</sup>* models at 3 weeks after TAM. Scale bar: 200  $\mu$ m. (B) Quantitative comparisons of IHC images (\* $P < 0.05$ , \*\* $P < 0.01$ ; Tukey's multiple comparison test following significance by 1-way ANOVA). Each shape (circle, square, triangle) corresponds to a different mouse [top bar graph ( $n = 2$ –3/genotype), middle bar graph ( $n = 3$ /genotype), bottom bar graph ( $n = 3$ /genotype)]. The solid shapes show the mean from each mouse; the open, smaller shapes represent individual values/field (range = 8–24 fields/mouse) at  $\times 20$  magnification.

clusters comprise the least frequent crypt cell types in this model (9.7% and 4.1%, respectively). Strikingly, HMGA1 deficiency decreases the proportion of stem and Paneth-like cell clusters by about 50% ( $P < 0.0001$ ) together with a concurrent expansion in the proportion of differentiated ECs (from 9.7% to 28.9%;  $P < 0.0001$ ), TA cells (from 26% to 34.9%;  $P < 0.0001$ ), and goblet cells (from 4.1% to 8%;  $P < 0.0001$ ) within the crypt epithelium. In *Apc*-deficient crypts with *Hmga1* heterozygous deficiency, the changes in most clusters are intermediate between crypts with intact or homozygous deficiency of *Hmga1* (Supplemental Figure 4, B and C). Together, these results indicate that intact HMGA1 is required to maintain the colon stem and Paneth-like cells in the setting of *Apc* deficiency while depleting more differentiated cells (ECs and goblet cells).

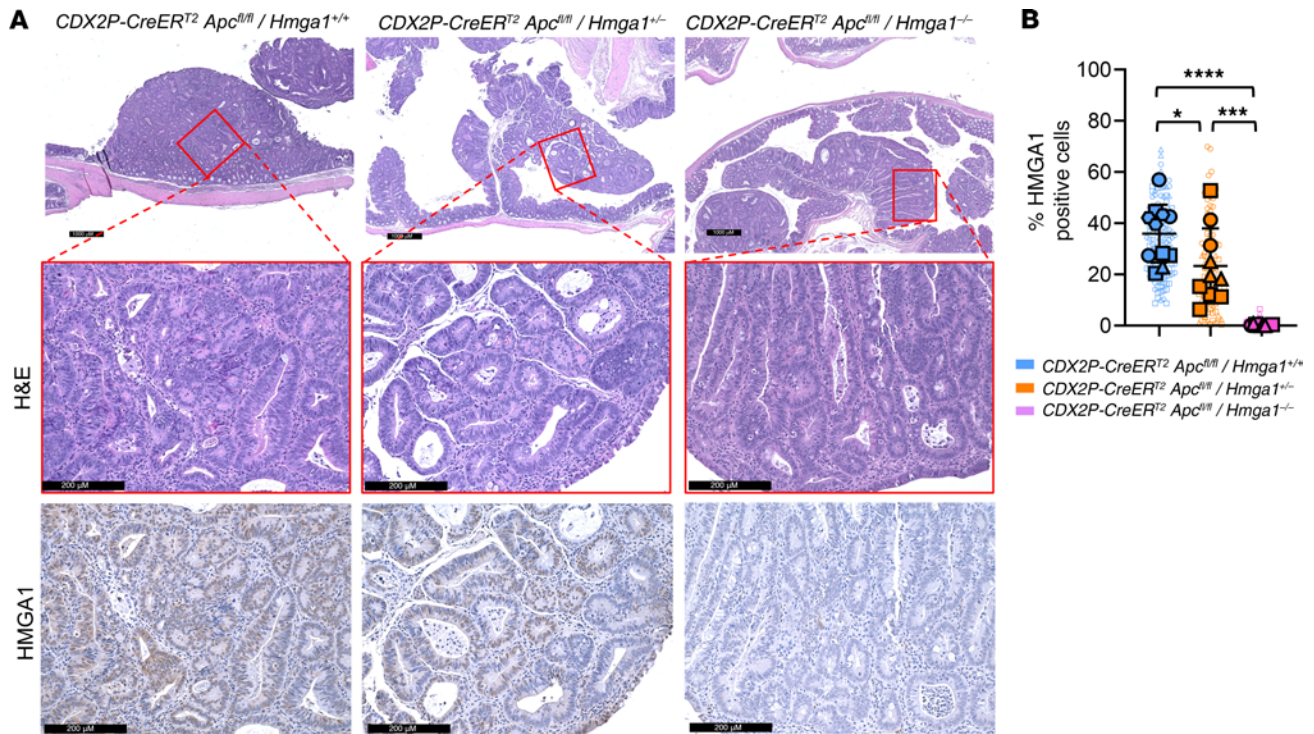
Within the immune cell islands, Cd4<sup>+</sup> and Cd8<sup>+</sup> T cells increase in frequency in the HMGA1-deficient crypt cells (Supplemental Figure 4), the latter of which could reflect an increase in tumor-infiltrating T lymphocytes. Both Cd4<sup>+</sup> and Cd8<sup>+</sup> T cells with *Hmga1* genetic deletion also exhibit a shift on UMAP, indicating that HMGA1 loss within these T cell populations alters their underlying transcriptomes (Supplemental Figure 4A).

*Trajectory and cell state analyses show that HMGA1 maintains an earlier cell state in Apc-deficient crypt cells.* To delineate HMGA1 function in differentiation dynamics in *Apc*-deficient colon crypt epithelium, we performed pseudotime trajectory analyses, assigning

time = 0 to the most dedifferentiated, stem cell cluster (via Seurat, Monocle 2). Intact HMGA1 results in a greater proportion of cells at earlier stages in development (time = 0; undifferentiated stage) whereas HMGA1 deficiency leads to more cells in later stages (time = 12; more differentiated stage) within the differentiation trajectory (Figure 8A). Next, we applied cell state analysis (Seurat, Monocle 2) as a more static assessment of differentiation status of each cell along the trajectory in *Apc*-deficient epithelial crypt cells (Figure 8B). Cell states (defined by the top 200 most differentially expressed genes within 5 groups with distinct transcriptomes) were assigned to individual cells along the trajectory. Similar to our cluster analysis, *Apc*-deficient crypt cells with intact *Hmga1* include a greater proportion of cells in an undifferentiated stem cell state (state 0) or Paneth-like state (state 1) compared to later, more differentiated cell states (states 3–4) (Figure 8B) or ECs (Figure 8C).

*Single cell transcriptomes suggest that HMGA1 accelerates proliferation by inducing gene networks involved in cell cycle progression.* Next, we inferred cell cycle status of each cell in the epithelial cluster from scRNA-seq (Seurat; standard settings). Transcriptomic changes suggest that intact HMGA1 in *Apc*-deficient crypt epithelial cells function by increasing proliferation, as evidenced by decreases in the proportion of cells in G0/G1 concurrent with increases in the proportion reaching G2/M; the proportion of S phase cells were similar in *Apc*-deficient crypt cells with or without HMGA1





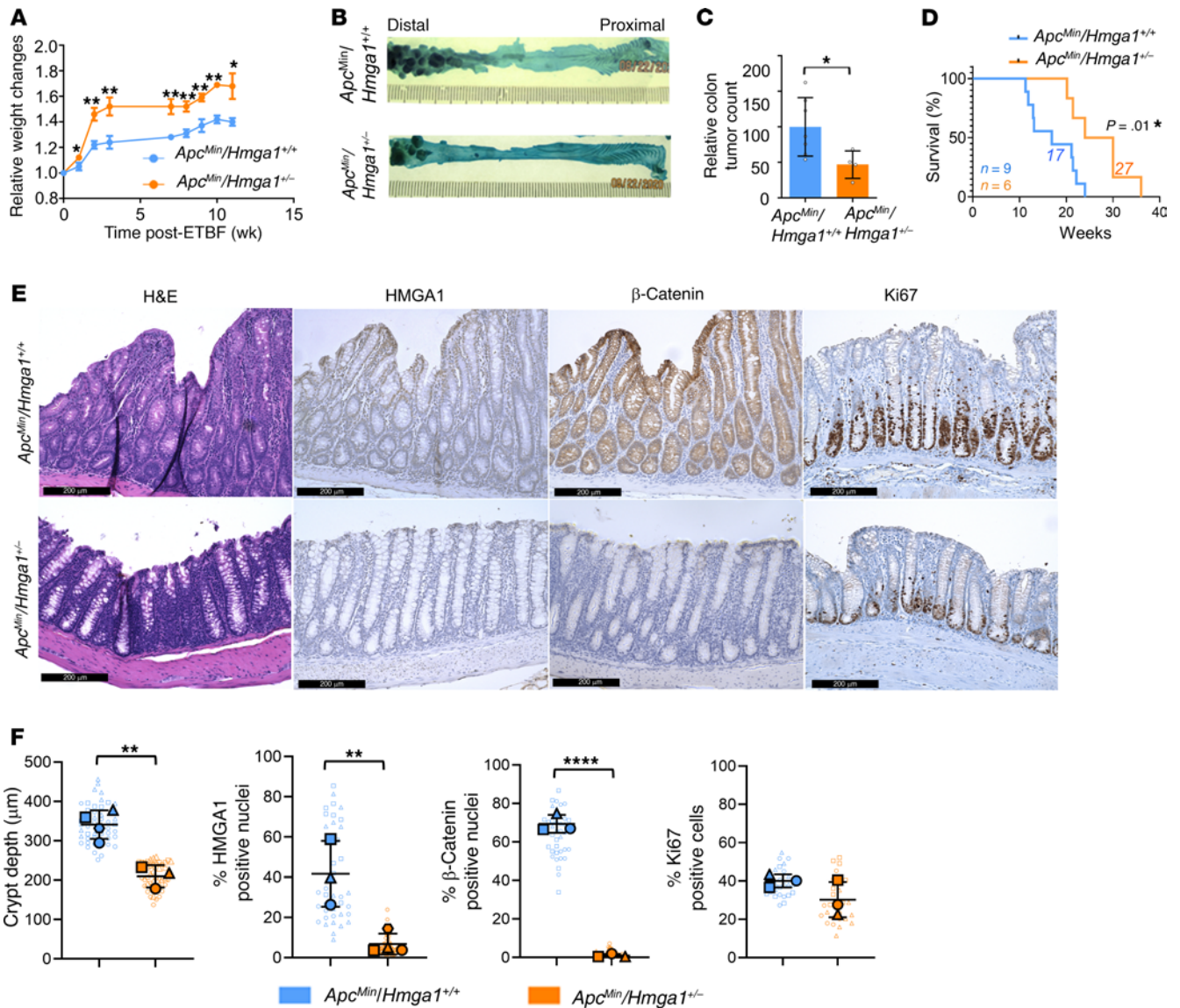
**Figure 3. *Hmga1* deficiency decreases colon tumorigenesis. (A)** Representative IHC images of nuclear HMGA1 (top),  $\beta$ -catenin (middle), and Ki67 (bottom) in *CDX2P-CreERT2/Apc<sup>fl/fl</sup>* models at 3 weeks after TAM. Scale bars: 1,000  $\mu$ m (top panel); 200  $\mu$ m (lower panel). **(B)** Quantitative comparisons of IHC images ( $*P < 0.05$ ,  $***P < 0.001$ ,  $****P < 0.0001$ ; Tukey's multiple comparison test following significance by 1-way ANOVA). Each shape (circle, square, triangle, hexagon) corresponds to a different mouse ( $n = 3\text{--}4/\text{genotype}$ ). The solid shapes show the mean values from each tumor; the open, smaller shapes represent individual values/field (range = 6–17 fields/tumor from 1–5 tumors/mouse) at  $\times 20$  magnification.

deficiency (Supplemental Figure 5). The proliferation marker gene encoding Ki67 is among the most upregulated genes of all G2/M genes. Collectively, our single-cell transcriptomes, together with increases in crypt depth, Ki67 protein staining, and tumorigenesis in mice with colon epithelial *Apc* inactivation and intact HMGA1 (Figure 1) are consistent with a model whereby HMGA1 increases proliferation restricted to cells at the earliest developmental stages, leading to expansion in stem and Paneth-like cells at the expense of more differentiated cells within the crypt epithelium.

*HMGA1* activates gene networks within crypt epithelial cells involved in *IFN* signaling, inflammation, DNA repair, proliferation, and *Wnt* signaling. To elucidate mechanisms underlying HMGA1 in *Apc*-deficient crypt cells, we performed gene set enrichment analysis (GSEA; MSigDB) with Hallmark and Curated gene sets (87, 88). GSEA with transcripts from all clusters (epithelial and immune cells) reveal that HMGA1 upregulates heterogeneous pathways, including those associated with metabolism (oxidative phosphorylation and glycolysis), proliferation (MYC Targets V1 and MYC Targets V2), and inflammation (IFN- $\alpha$  Response and IFN- $\gamma$  Response) (Supplemental Figure 6). HMGA1 also activates multiple WNT networks, including Wnt Pathway requiring MYC, Degradation of the  $\beta$ -catenin Destruction Complex, TCF Dependent Signaling in Response to WNT, and APC Targets (Table 1). By contrast, transcriptional networks repressed by HMGA1 include allograft rejection, IL2-STAT5 signaling, and mitotic spindle genes (Hallmark). Notably, repression in allograft rejection and IL2-STAT5 gene networks have been implicated in immune escape and decreases in cytotoxic

tumor infiltrating lymphocytes (Supplemental Figure 6) (89). The heterogeneity in these pathways is consistent with the diverse cell populations (epithelial and immune) within the crypts.

To focus our analysis on the tumor-initiating cells, we performed GSEA exclusively on transcripts from the crypt epithelial island. Further, this island is comprised of the majority of cells from the crypt isolates and cell numbers are sufficient for pathway analyses. In the remaining immune islands, cell numbers were insufficient for further GSEA. Strikingly, HMGA1 activates transcriptional networks involved in inflammation, including IFN- $\alpha$  and IFN- $\gamma$  response genes and proliferation (MYC targets V1) within the epithelial island (Figure 9A). DNA repair genes are also induced, which is a frequent transcriptional response when quiescent stem cells are triggered to cycle and proliferate (Figure 9A) (90–92). Among the IFN networks, multiple IFN-induced genes that mediate inflammatory signals are upregulated by HMGA1, including *IFN-induced transmembrane proteins 1, 2, 3* (*Ifitm 1–3*), *IFN stimulated gene 15* (*Isg15*), *Stat1*, *Stat2*, and cytokines (*Ccl5*, *Cxcl9/10*) (Figure 9B). Wnt pathway genes are also prominent among the networks activated by HMGA1 within the epithelial crypt cells (Figure 9A and Table 1). By contrast, HMGA1 represses gene networks involved in fatty acid metabolism and adipogenesis, metabolic pathways used extensively by differentiated ECs in intestinal epithelium (Figure 9A) (93). HMGA1 also represses genes controlling protein secretion (Figure 9A), an important cellular function of differentiated ECs, which secrete digestive enzymes (93). Together, our single-cell transcriptomes



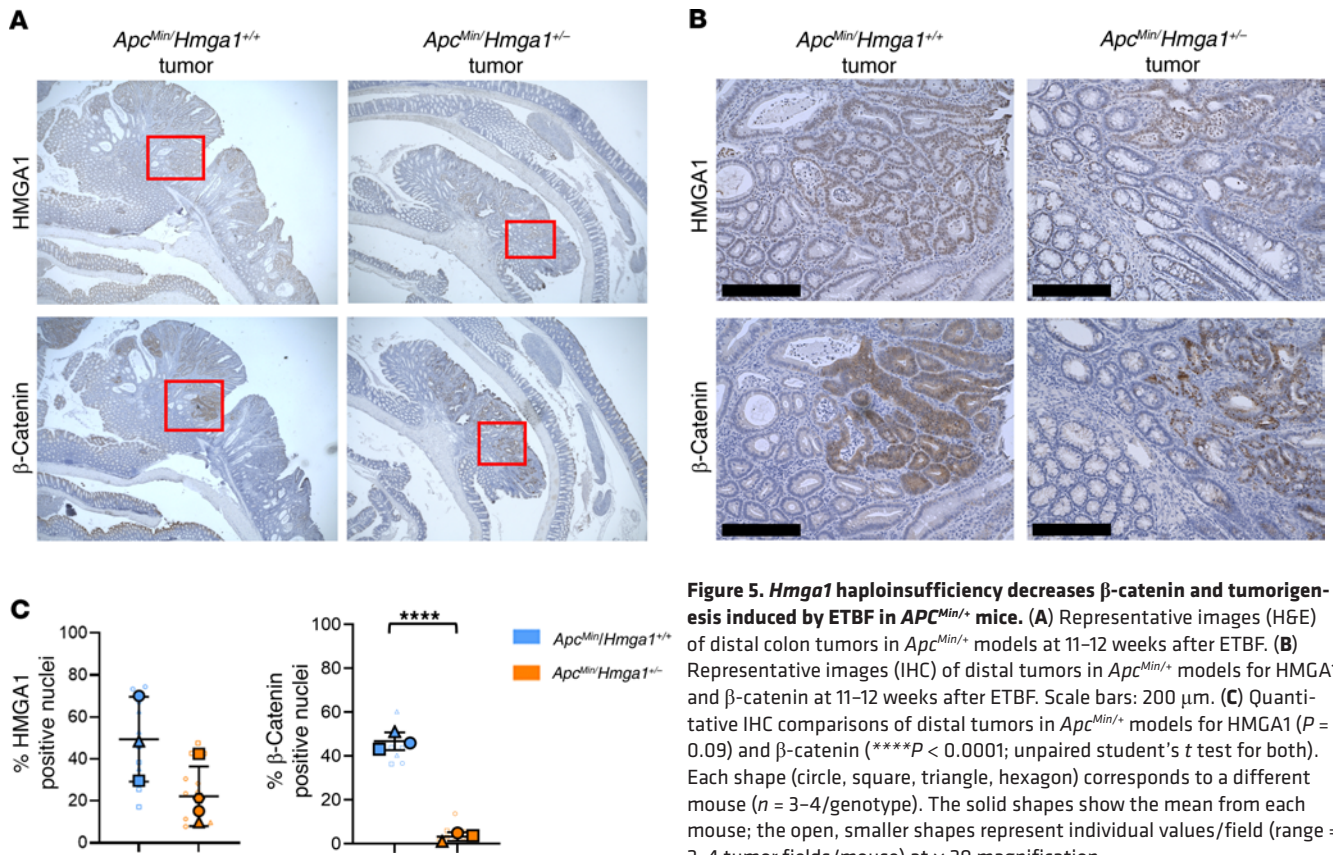
**Figure 4. *Hmga1* haploinsufficiency disrupts colon tumorigenesis induced by ETBF in *Apc<sup>Min</sup>/+* mice.** (A) Body weights at necropsy after ETBF in *Apc<sup>Min</sup>/+* mice with intact *Hmga1* or heterozygous *Hmga1* (\**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001; student's *t* test). (B) Representative images of methylene-blue stained colons to visualize tumors in *Apc<sup>Min</sup>/Hmga1<sup>+/-</sup>* mouse (top) compared with *Apc<sup>Min</sup>/Hmga1<sup>+/+</sup>* mouse (bottom) at 11-12 weeks after ETBF. (C) Normalized tumor numbers in *Apc<sup>Min</sup>/+* models (\**P* < 0.05; Mann-Whitney test). (D) Kaplan-Meier plot showing survival in in *Apc<sup>Min</sup>/+* mice with intact *Hmga1* or heterozygous (\**P* < 0.05; Mantel-Cox test). (E) Representative images (H&E left; IHC right; Scale bars: 200 μm) for HMGA1 (second column), β-catenin (third column), and Ki67 (right) in distal colon of *Apc<sup>Min</sup>/+* models at 11-12 weeks after ETBF. (F) Comparison of crypt depths (\*\**P* < 0.01) and IHC for nuclear HMGA1 (\*\**P* < 0.01), nuclear β-catenin (\*\*\*\**P* < 0.0001) and Ki-67 (*P* = 0.16, unpaired student's *t* test for each comparison) in *Apc<sup>Min</sup>/+* models. For crypt depth (left), each shape (circle, square, triangle, hexagon) corresponds to a different mouse (*n* = 3-4/genotype). The solid shapes show the mean from each mouse; the open, smaller shapes represent individual measurements/crypt (range = 9-16 crypts/mouse). For the IHC comparisons, each shape (circle, square, triangle, hexagon) corresponds to a different mouse (*n* = 3-4/genotype), the solid shapes show the mean value from each mouse; the open, smaller shapes represent individual values/field (range=9-19 fields/mouse) at x20 magnification.

and pathway analysis demonstrate that HMGA1 expands the stem and Paneth-like cells while driving transcriptional networks involved in proliferation, inflammation, Wnt signaling, and DNA repair. Conversely, HMGA1 restrains differentiation and represses metabolic gene networks active in differentiated ECs.

*HMGA1* enhances chromatin accessibility at gene loci involved in proliferation, DNA repair, inflammation, and Wnt signaling. Because HMGA1 is an architectural transcription factor that modulates chromatin structure, we performed assays to detect accessible chromatin mediated by HMGA1 via assays of transposase-accessible

chromatin sequencing (ATAC-seq) in *Apc*-deficient crypt cells with or without HMGA1 deficiency (94). Notably, overall chromatin accessibility is enhanced in *Apc*-deficient crypt cells with intact HMGA1 compared with crypts lacking HMGA1 (Figure 10, A-C). HMGA1 results in both more peaks and longer stretches of accessible chromatin overall (Figure 10, A-C). Focusing on promoter regions (up to -3 kb from the transcription start sites), we also identified more peaks and longer stretches of open chromatin within these regulatory regions with intact HMGA1 (Figure 10, B and C). Similar to our scRNA-seq results, gene networks associated





**Figure 5. *Hmga1* haploinsufficiency decreases  $\beta$ -catenin and tumorigenesis induced by ETBF in *APC<sup>Min/+</sup>* mice.** (A) Representative images (H&E) of distal colon tumors in *Apc<sup>Min/+</sup>* models at 11–12 weeks after ETBF. (B) Representative images (IHC) of distal tumors in *Apc<sup>Min/+</sup>* models for HMGA1 and  $\beta$ -catenin at 11–12 weeks after ETBF. Scale bars: 200  $\mu$ m. (C) Quantitative IHC comparisons of distal tumors in *Apc<sup>Min/+</sup>* models for HMGA1 ( $P = 0.09$ ) and  $\beta$ -catenin ( $****P < 0.0001$ ; unpaired student's  $t$  test for both). Each shape (circle, square, triangle, hexagon) corresponds to a different mouse ( $n = 3$ –4/genotype). The solid shapes show the mean from each mouse; the open, smaller shapes represent individual values/field (range = 2–4 tumor fields/mouse) at  $\times 20$  magnification.

with HMGA1-mediated accessible chromatin included pathways involved in proliferation (MYC Targets V1, E2F Targets, G2M Checkpoint genes) and inflammation (IFN- $\gamma$ , TNF- $\alpha$  signaling via NF- $\kappa$ B (Figure 10, D and E, and Supplemental Figure 7). Accessible chromatin is also enriched at Wnt signaling gene networks (Table 1). Intersecting pathways identified by both ATAC-seq and scRNA-seq (epithelial island) revealed that HMGA1 increased chromatin accessibility and expression of genes involved in proliferation (MYC Targets V1), DNA repair, inflammation (IFN- $\gamma$  response genes), and Wnt signaling (Figure 10E and Table 1).

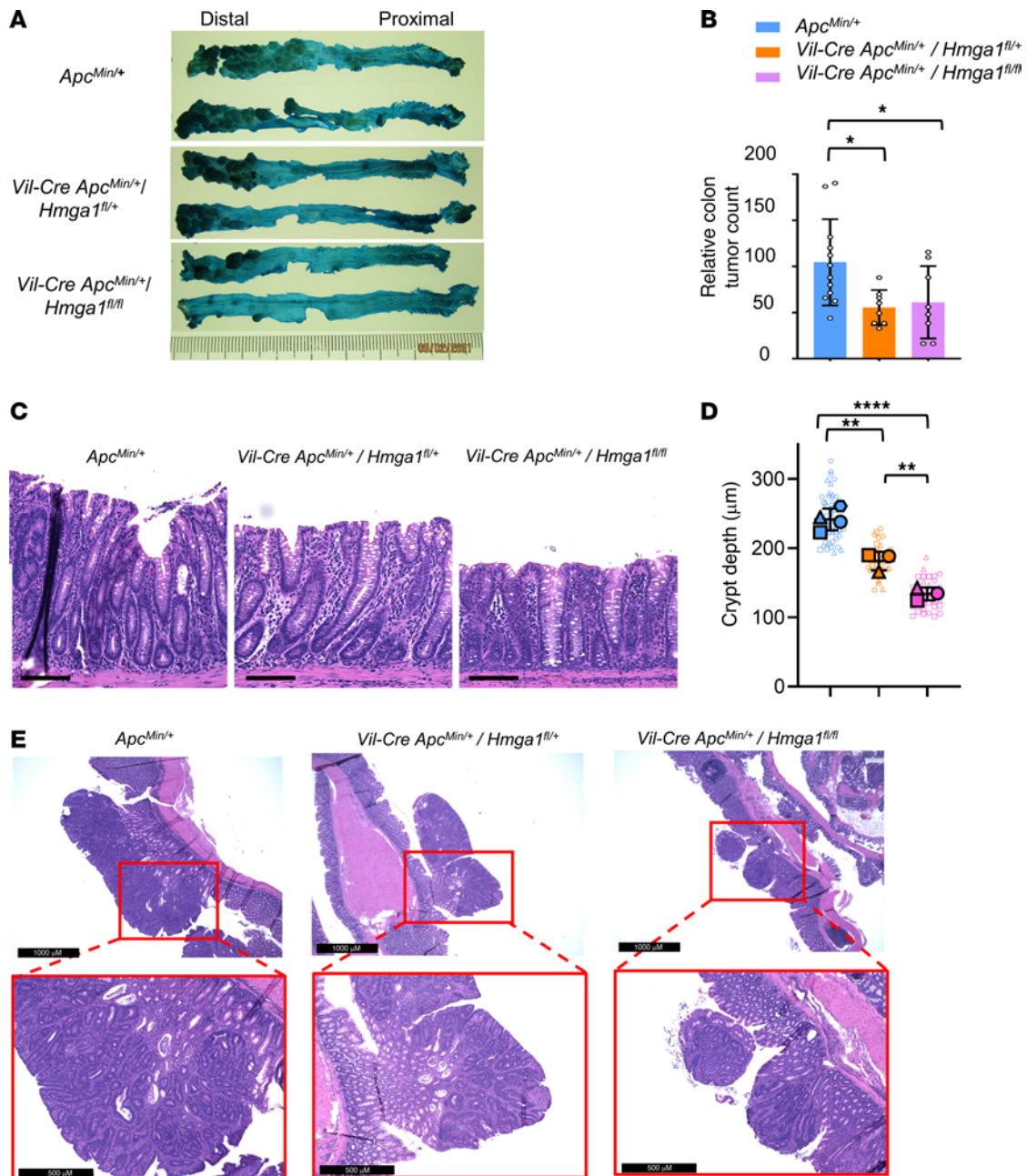
**HMGA1 amplifies expression of Wnt pathway genes in *Apc*-deficient colon crypts.** Given the fundamental role for Wnt signaling in colon tumorigenesis and HMGA1-dependent upregulation of Wnt genes and  $\beta$ -catenin levels in our tumor models, we further examined the relationship between HMGA1 and Wnt genes. We focused on canonical Wnt pathway genes, including Wnt effectors (*Cttnb1*, *Tcf4*, *Axin2*, *Cd44*, *Ets2*, *Ephb2*, *Ascl2*, *cMyc*, *Prom1*, *Sox9*) and Wnt receptors (*Lgr5*, *Lrp5*, *Lrp6*, *Fzd5*, *Fzd7*). Remarkably, all Wnt effector genes were upregulated at the level of single cells in the setting of intact HMGA1 and *Apc* deficiency (Figure 11A). Of the Wnt receptors, both *Lgr5* and *Lrp5* transcripts are upregulated in crypt cells with intact HMGA1. We also found significant positive correlations between *Hmga1* and multiple Wnt effectors (Figure 11B) with the strongest correlations ( $r > 0.68$ ;  $P < 0.05$ ) for *Ascl2*, *Axin2*, *Tcf4*, *Cttnb1*, *Ephb2*, and *cMyc*. To determine whether HMGA1 enhances chromatin accessibility at promoter regions for these genes, we examined our ATAC-seq results, which revealed increased chromatin accessibility at promoter regions for *Ascl2*,

*Tcf4*, *Prom1*, *Lgr5*, and *Lrp5* (Supplemental Figure 8). Intriguingly, HMGA1 is also associated with accessible chromatin at the *Hmga1* promoter (Supplemental Figure 8), suggesting that high levels of HMGA1 induce its own expression in *Apc*-deficient crypts by opening chromatin at its promoter region. Together, these results demonstrate that HMGA1 enhances chromatin accessibility to activate Wnt agonist receptor signaling and Wnt effector genes. In small intestinal epithelium, *ASCL2* activates Wnt genes (95–98), and upregulation in *Ascl2* could trigger a feed-forward loop whereby HMGA1 activates *Ascl2*, which, in turn, amplifies Wnt gene expression in colon epithelium with *Apc* inactivation.

**HMGA1 and ASCL2 are upregulated and coexpressed in human colorectal cancer.** To determine which HMGA1 pathways are relevant to human colon tumorigenesis, we queried the Cancer Genome Atlas (TCGA) for expression of HMGA1 and Wnt genes (Figure 11, C and D). As we previously reported, *HMGA1* and *SOX9* are upregulated in colon cancer compared with nonmalignant epithelium (4, 30, 62). Strikingly, most of the Wnt genes upregulated by HMGA1 in our murine model are also upregulated in human colon cancer, including the WNT effectors, *ASCL2*, *AXIN2*, *CTNNB1*, *MYC*, *EPHB2*, *CD44*, and *ETS2* and the WNT receptors, *LGR5*, *LRP5*, and *LRP6* (Figure 11, C and D). Further, both *ASCL2* and *cMYC* are upregulated and positively correlated with *HMGA1*, suggesting that HMGA1 may directly induce their expression in human colon tumorigenesis (Figure 11C).

**HMGA1 upregulates ASCL2 and promotes oncogenic properties in human colon cancer cells.** The *ASCL2* transcription factor is critical to cell fate in the small intestine (95–98) although its

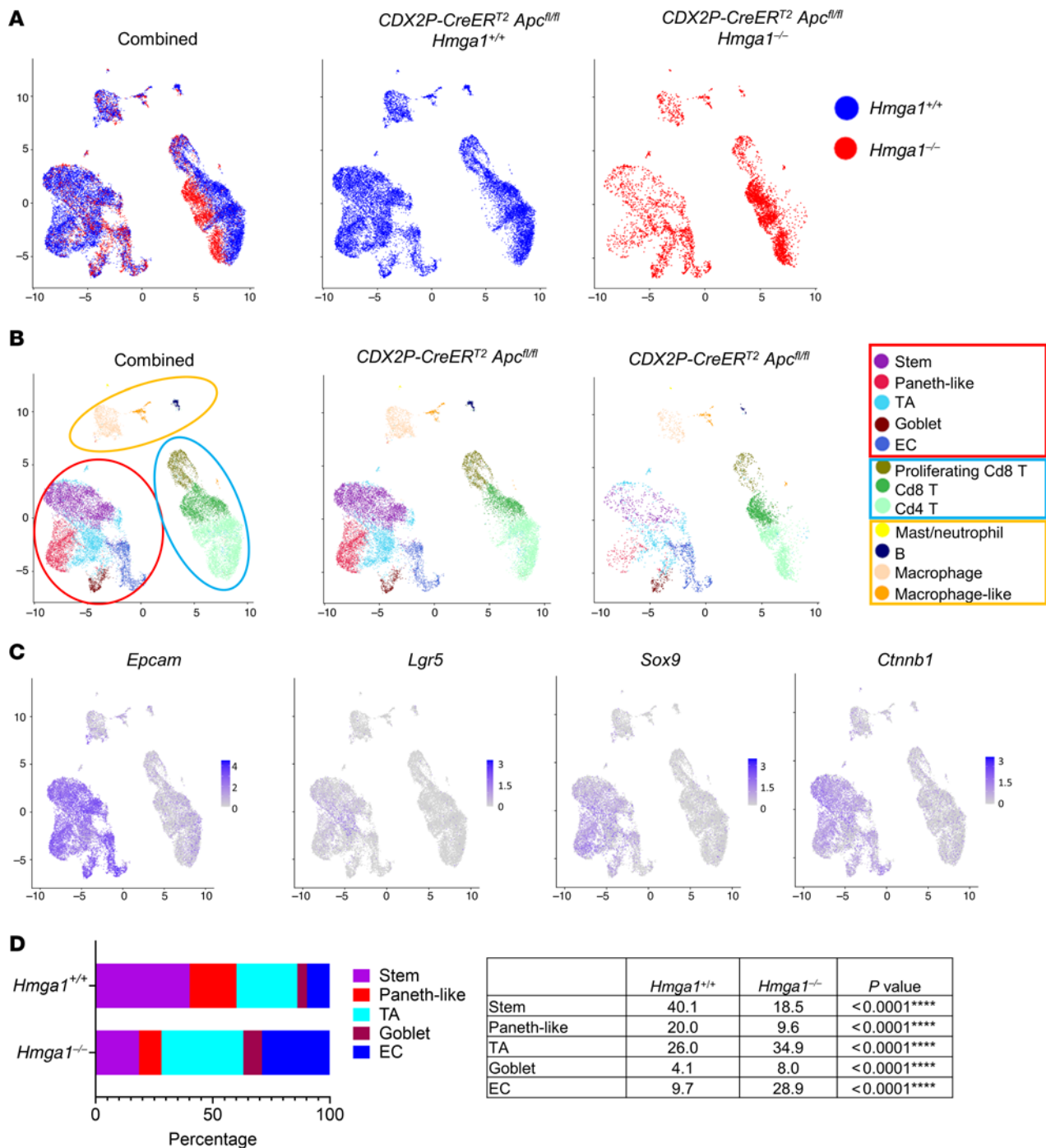




**Figure 6. Loss of *Hmga1* allele within colon epithelium decreases colon tumorigenesis induced by ETBF in *Apc<sup>Min/+</sup>* mice.** (A) Representative images of methylene-blue–stained colons of *Apc<sup>Min/+</sup>* mice (top) compared with *Apc<sup>Min</sup>* with tissue-specific heterozygous *Hmga1* deletion (middle) and tissue-specific homozygous *Hmga1* deletion (bottom) at 11–12 weeks after ETBF. (B) Relative tumor numbers (%) in *Apc<sup>Min</sup>* mice with intact *Hmga1*, tissue-specific heterozygous *Hmga1* deletion, or tissue-specific homozygous *Hmga1* deletion from 3 separate experiments; tumor numbers in control were assigned a value of 100 (\* $P < 0.05$ ; Mann-Whitney test for both comparisons). (C) Representative images (H&E) of distal colon of *Apc<sup>Min/+</sup>* with or without tissue-specific *Hmga1* deficiency models. Scale bars: 100  $\mu\text{m}$ . (D) Distal colon crypt depths in *Apc<sup>Min/+</sup>* mice with or without tissue-specific *Hmga1* deficiency. (\*\* $P < 0.01$ , \*\*\*\* $P < 0.0001$ ; Tukey's multiple comparisons test following significance by 1-way ANOVA). Each shape (circle, square, triangle, hexagon) corresponds to a different mouse ( $n = 3\text{--}4/\text{genotype}$ ). The solid shapes show the mean value from each mouse; the open, smaller shapes represent individual measurements/crypt (range = 9–19 crypts/mouse) at  $\times 20$  magnification. (E) Representative images (H&E) of distal colon tumors of *Apc<sup>Min/+</sup>* with or without tissue-specific *Hmga1* deficiency.

role in the colon has not been studied in detail. Because our results strongly link *HMGA1* to *ASCL2* in colon tumorigenesis in humans and mice, we tested whether *HMGA1* directly activates *ASCL2* expression in human colon cancer cells. Silencing *HMGA1* in 2 human colon cancer cell lines (SW620, SW480) by lentiviral-mediated delivery of short hairpin RNA (shRNA)

or CRISPR/Cas9 represses *ASCL2*, demonstrating that *ASCL2* expression depends on *HMGA1* (Figure 12A and Supplemental Figure 9A). Next, we tested whether *HMGA1* deficiency affects oncogenic properties in these cells. We previously reported that *HMGA1* knockdown decreases clonogenicity in SW480 cells using plasmid-mediated gene silencing (35). Here, we found that

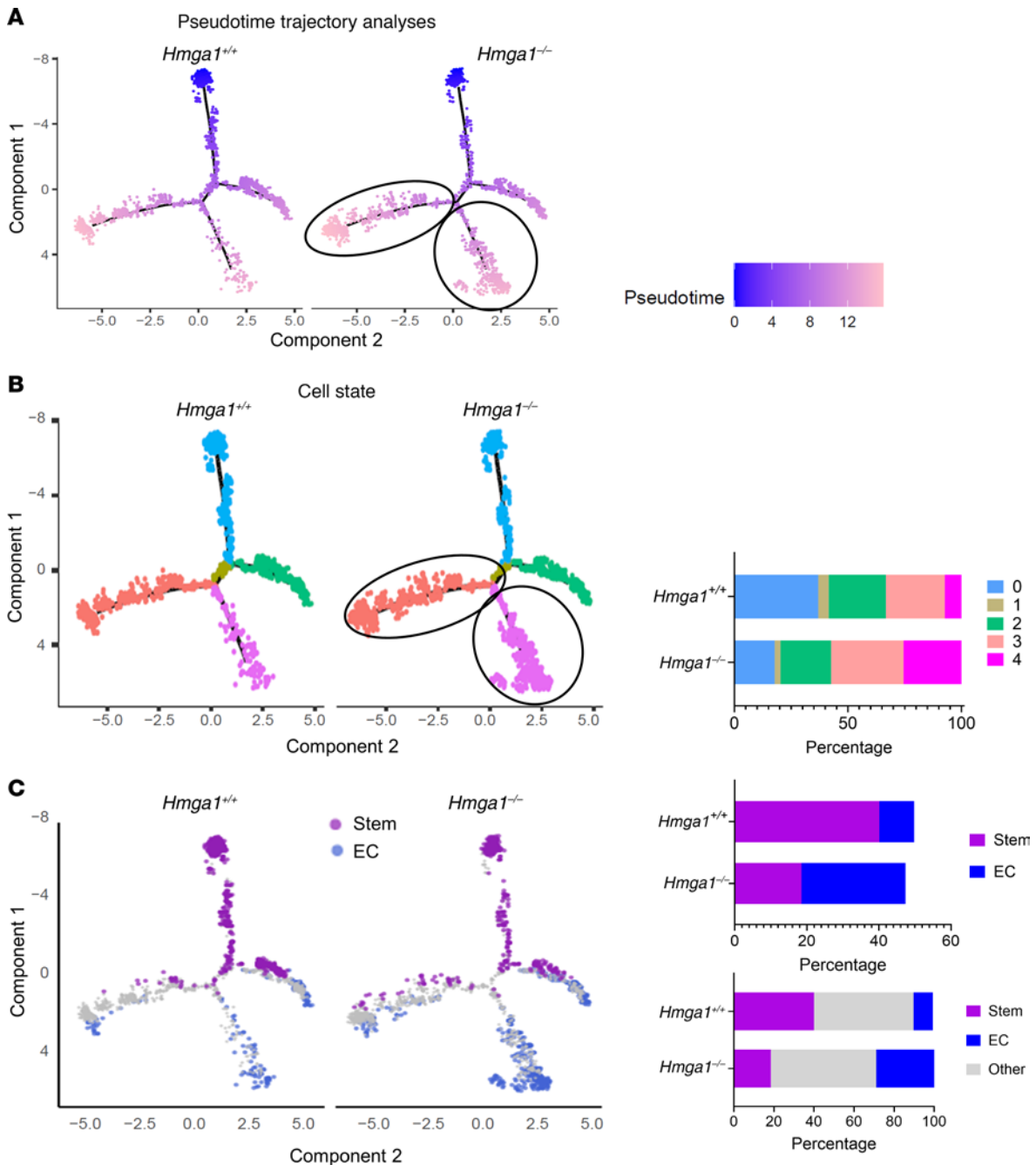


**Figure 7. HMGA1 expands colon stem cells and Paneth-like cells while depleting more differentiated cells in *Apc*-deficient colon crypts.** (A) UMAPs from scRNA-seq of crypt cells from *CDX2P-CreERT2 Apc<sup>fl/fl</sup>* mice with *Hmga1*<sup>+/+</sup> or *Hmga1*<sup>-/-</sup>; shown together (left) or separately to highlight differences (center and right). (B) UMAP from scRNA-seq by cluster. Three distinct islands capture epithelial cell types (red circle), T cells (blue circle), and other immune cells (yellow). Imputed cell identities are designated by separate colors. TA, transit amplifying cells; EC, enterocytes. (C) *Epcam*, *Lgr5*, and other Wnt genes (*Sox9*, *Ctnnb1*) are enriched in the epithelial island. Single cell transcripts from both genotypes are shown. (D) Relative proportion of cell types in crypt cells by genotype (bar graph, left; Table, right). (Association between cell and HMGA1 status was evaluated by  $\chi^2$  test for each cell type versus all others).

*HMGA1* silencing (via shRNA or CRISPR) disrupts proliferation and clonogenicity similarly in SW620 and SW480 cells (Figure 12, A and B, and Supplemental Figure 9, A and B), demonstrating that these in vitro, oncogenic phenotypes depend on high levels of HMGA1. To determine whether *ASCL2* restoration will rescue these phenotypes, we reexpressed *ASCL2* in both colon

cancer cell lines (SW620 and SW480) with *HMGA1* silencing (Supplemental Figure 9C). Restoration in *ASCL2* levels results in a partial rescue of proliferation and full rescue of clonogenicity in both cell lines, suggesting that *ASCL2* mediates some, but not all, effects of HMGA1 in these colon cancer cells (Supplemental Figure 9, C and D).

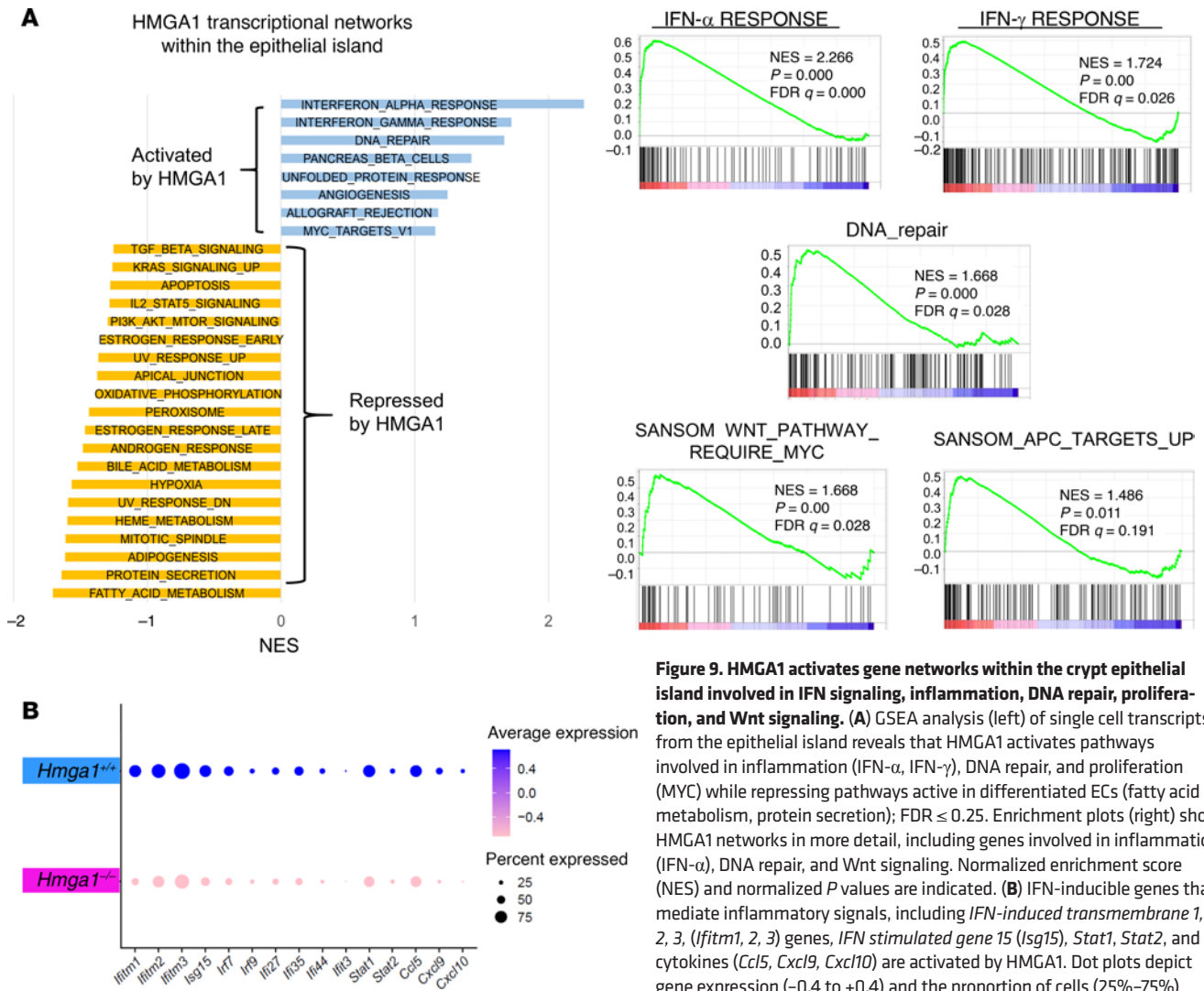




**Figure 8. *Hmga1* deficiency alters cell state, decreasing stem and Paneth-like cell populations while expanding more differentiated cell populations in *Apc* deficient crypt cells.** (A) Pseudotime trajectory analysis estimated from scRNA-seq of *CDX2P-CreER<sup>T2</sup> Apc<sup>fl/fl</sup>* crypt cells from the epithelial island with *Hmga1*<sup>+/+</sup> or *Hmga1*<sup>-/-</sup>. HMGA1 deficient cells are more prominent in later stages of pseudotime (indicated by black ovals) compared with time 0 cells. (B) Cell states defined by the top 200 most differentially expressed genes on the trajectories from pseudotime analysis were assigned 0–4 and indicated by color on a trajectory plot (left) or bar graph (right). Note the skewing to cell states 3 and 4 in HMGA1 deficient cells. (C) Stem cells and enterocytes (ECs) imputed from scRNA-seq are shown on the trajectories to highlight the major differences between *CDX2P-CreER<sup>T2</sup> Apc<sup>fl/fl</sup>* cells with intact HMGA1 or HMGA1 deficiency. HMGA1 deficient cells have increased ECs (blue) with decreased stem cells (violet). Bar graphs show relative cell frequencies (right); the top graphs show only stem and ECs, the bottom includes all cells with grey depicting cells that are not stem cells nor ECs.

*HMGA1* directly induces *ASCL2* by binding to its promoter and recruiting activating histone marks in human colon cancer cells. To ascertain whether HMGA1 binds directly to the *ASCL2* promoter to activate its expression, we used an in silico prediction algorithm (TRAP) (99), which identified 7 potential HMGA1 binding sites within the *ASCL2* promoter-enhancer region (labeled 1, 2, 3, 4, 5, 6, and 7; Supplemen-

tal Figure 10A). Because sites 6 and 7 are within 10 base pairs of each other, they could not be resolved by ChIP-PCR, and were therefore denoted region 6–7. Intriguingly, all of these sites are positioned near the homologous regions of HMGA1-dependent accessible chromatin in the mouse *Ascl2* promoter (Figure 12C). By chromatin immunoprecipitation-PCR (ChIP-PCR), we assessed HMGA1 chromatin



**Figure 9. HMG1 activates gene networks within the crypt epithelial island involved in IFN signaling, inflammation, DNA repair, proliferation, and Wnt signaling.** (A) GSEA analysis (left) of single cell transcripts from the epithelial island reveals that HMG1 activates pathways involved in inflammation (IFN- $\alpha$ , IFN- $\gamma$ ), DNA repair, and proliferation (MYC) while repressing pathways active in differentiated ECs (fatty acid metabolism, protein secretion); FDR  $\leq$  0.25. Enrichment plots (right) show HMG1 networks in more detail, including genes involved in inflammation (IFN- $\alpha$ ), DNA repair, and Wnt signaling. Normalized enrichment score (NES) and normalized P values are indicated. (B) IFN-inducible genes that mediate inflammatory signals, including *IFN-induced transmembrane 1, 2, 3*, (*Ifitm1, 2, 3*) genes, *IFN stimulated gene 15 (Isg15)*, *Stat1*, *Stat2*, and cytokines (*Ccl5*, *Cxcl9*, *Cxcl10*) are activated by HMG1. Dot plots depict gene expression (-0.4 to +0.4) and the proportion of cells (25%-75%) expressing each transcript within the epithelial island.

occupancy at these sites in SW620 cells, as these cells have higher levels of both *HMG1* and *ASCL2* compared with SW480 cells. There was robust HMG1 occupancy throughout the *ASCL2* promoter compared with the IgG antibody as a negative control (Supplemental Figure 10, A and B). Next, we compared HMG1 chromatin binding in SW620 cells with or without *HMG1* silencing, which showed enrichment for HMG1 occupancy at the same regions in control cells and depletion of HMG1 with *HMG1* silencing, validating the specificity of our HMG1 antibody (Figure 12D and Supplemental Figure 10C). Because HMG1 recruits active histone marks to upregulate developmental genes in other settings (28, 29), we tested whether HMG1 binding associates with activating histones, including histone H3 lysine 4 trimethylation (H3K4me3) and histone H3 lysine 27 acetylation (H3K27Ac), which mark active promoters and enhancers, respectively. These marks also associate with HMG1 chromatin binding in other tumor settings (28, 29). We found enrichment for both activating histone marks (H3K4me3, H3K27Ac) in the regions of HMG1 binding in SW620 and SW480 cells from public databases (GSE10692) (Figure 12, C and D). By ChIP-PCR, we found that both H3K4me3 and H3K27Ac bind to

the *ASCL2* promoter, with greatest enrichment near the HMG1 binding site number 1 (located near the transcription start site [TSS]) and these active marks are depleted with *HMG1* silencing (Figure 12D and Supplemental Figure 10C). By contrast, there was no change with *HMG1* silencing in the positive control, histone H3, which is a ubiquitous histone that does not modulate gene expression (Figure 12E and Supplemental Figure 10C). We also tested whether HMG1 depletion enables repressive histones to bind to the *ASCL2* promoter as a mechanism of downregulating *ASCL2* with *HMG1* silencing. Because the repressive histone 3 lysine 27 trimethyl (H3K27me3) was identified in a colon cancer cell line (HCT116) from a public database (GSE171817), we assessed its binding relative to that of HMG1. In control SW620 cells with high HMG1, there was minimal binding of the repressive mark, H3K27me3; however, HMG1 depletion results in modest, yet significant increases in the H3K27me3 repressive mark at the *ASCL2* promoter enhancer region (Figure 12E and Supplemental Figure 10D).

To determine if HMG1 activates the *ASCL2* promoter, we cloned the human *ASCL2* promoter sequence (-2.5 kb from the TSS) upstream of the luciferase reporter gene and transfected



**Table 1. Curated WNT pathways upregulated by HMGA1**

Pathway	scRNA-seq (all clusters)				scRNA-seq (epithelial clusters)				ATAC-seq			
	Size	NES	NOM P value	FDR q value	Size	NES	NOM P value	FDR q value	Size	NES	NOM P value	FDR q value
Sansom APC Targets Up	129	2.055	0.000	0.002	129	2.055	0.000	0.002	42	1.676	0.020	0.102
Sansom APC Targets Require MYC	220	1.792	0.000	0.043	220	1.463	0.000	0.200	49	1.751	0.000	0.089
Sansom WNT Pathway Require MYC	56	1.707	0.008	0.059	56	1.493	0.024	0.187	24	-0.663	0.836	1.000
Reactome Degradation of $\beta$ -catenin by the Destruction Complex	78	1.605	0.000	0.097	78	1.383	0.076	0.249	17	1.913	0.017	0.060
Reactome $\beta$ -catenin Independent WNT Signaling	110	1.421	0.030	0.187	110	1.219	0.097	0.358	22	1.326	0.109	0.265
Reactome TCF Dependent Signaling in Response to WNT	136	1.395	0.021	0.207	136	1.033	0.319	0.532	36	0.917	0.553	0.700
Sansom APC Targets	196	1.390	0.005	0.210	196	1.002	0.390	0.581	70	1.371	0.100	0.246

Size, gene number per network; NES, normalized enrichment score; NOM P-value, normalized P-value.

this construct into colon cancer cell lines (SW620, SW480) (Supplemental Figure 11A). In both cell lines with abundant levels of HMGA1, the *ASCL2* promoter is induced compared with control vector lacking the *ASCL2* promoter sequence (Supplemental Figure 11B). By contrast, *HMGA1* silencing decreases *ASCL2* promoter activity, consistent with HMGA1-dependent activation of the *ASCL2* promoter (Supplemental Figure 11B). Together, our results support a model whereby HMGA1, present in high levels, binds to the *ASCL2* promoter, enhances chromatin accessibility, and recruits activating histones to induce *ASCL2* and downstream Wnt genes, thereby driving tumorigenesis in the setting of *Apc* deficiency. Moreover, our findings further highlight HMGA1 as a promising potential therapeutic target, particularly since loss of HMGA1 in colon epithelium has only subtle effects on epithelial regeneration under homeostatic conditions.

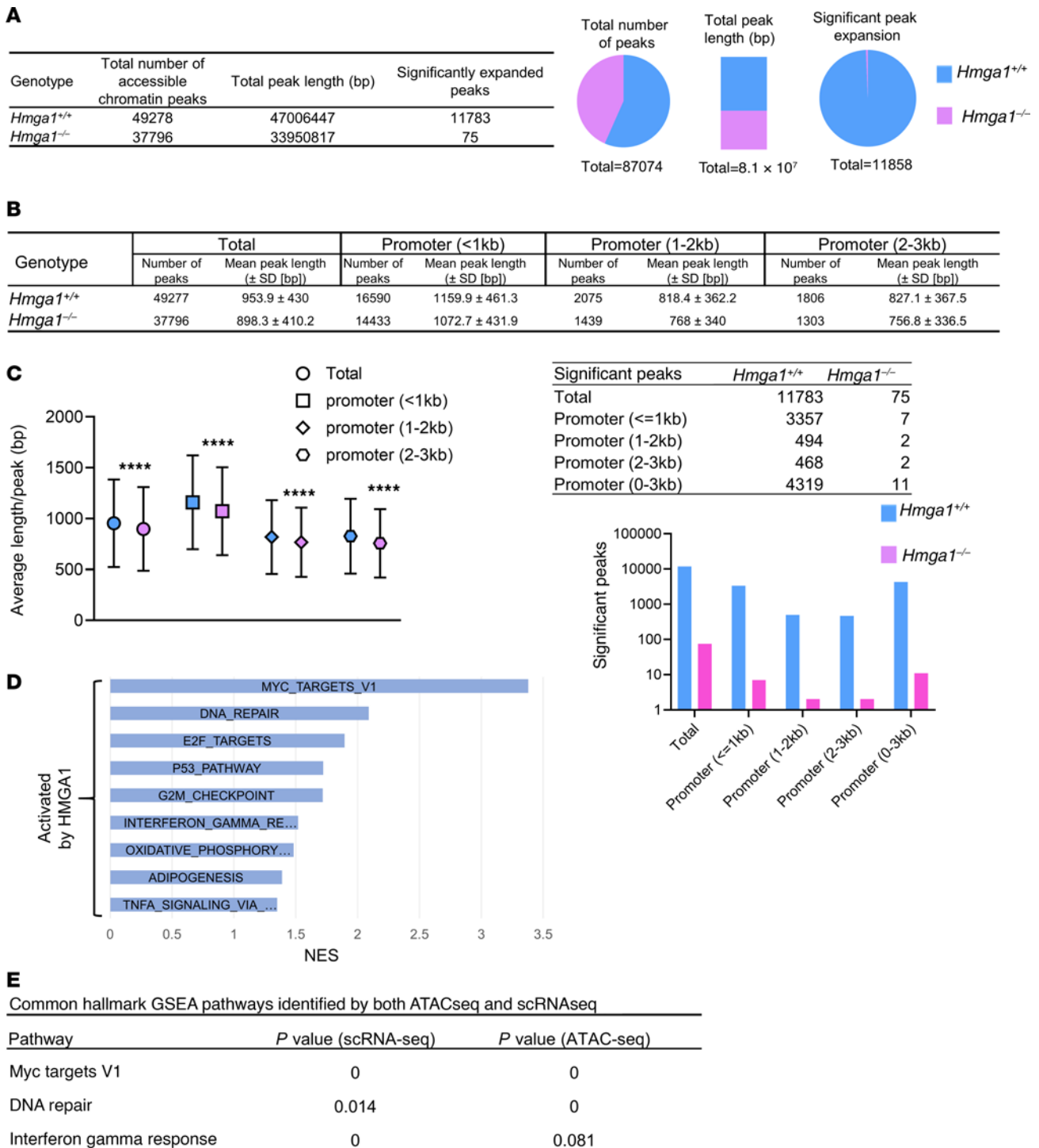
## Discussion

Changes in nuclear structure and function are required for normal development, tissue regeneration, and tumorigenesis (4, 11). While underlying mechanisms remain incompletely understood, chromatin state has emerged as a fundamental player required for diverse cell fate decisions in tumor biology. Embryonic stem cells and tissue-specific, adult stem cells have large nuclei harboring “open” accessible chromatin, which is thought to endow these cells with developmental potency or the capacity to differentiate into diverse progeny with distinct functions (4, 11). Similarly, nuclei in aggressive cancer cells are often enlarged and irregular (4, 11), whereas nuclear compaction accompanies differentiation in normal tissues (4, 11). Though somatic mutations accumulate in adult stem cells over time, particularly in highly proliferative tissues, such as the colon crypts, most mutated cells do not evolve into tumors. Thus, changes in chromatin structure and cell state provide a plausible requisite for tumor development. Indeed, pathologists distinguish cancer cells from nonmalignant cells primarily by alterations in nuclear architecture. These observations suggest that understanding mechanisms underlying chromatin structure and cell state during tumor evolution could reveal strategies to intercept the transition from early neoplasia to invasive cancers. Colon tumorigenesis offers a unique opportunity to study cell state, adult stem cells, and tumorigenesis given the hierarchical organization of stem and progenitors

within the crypts along with evidence implicating mutated LGR5+ colon stem cells as a tumor initiating cell (100). Moreover, the incidence of colon cancer is increasing globally, particularly in younger individuals, highlighting the significance of this work (24–26).

Here, we discover that HMGA1 acts as an epigenetic regulator that imposes a stem-like chromatin state within *Apc*-deficient crypt epithelial cells. HMGA1 enhances chromatin accessibility at key loci, leading to activation of gene networks involved in Wnt signaling, proliferation, and inflammation early in tumorigenesis. As an architectural transcription factor, HMGA1 binds to DNA and recruits histones and other chromatin complexes to modulate gene expression, rather than acting on its own. While HMGA1 drives clonal expansion, aberrant differentiation, and transformation in diverse settings, its function in colon tumorigenesis has not been studied in detail despite the fact that it is among the most overexpressed genes in colon cancer compared with nonmalignant epithelium (30, 62). We found that *Hmga1* haploinsufficiency dampens colon tumor development and prolongs survival in 2 models. Importantly, mice with *Hmga1* heterozygosity (and WT *Apc*) have normal development and lifespans (28, 29, 56). By contrast, *Apc* deletion together with intact HMGA1 results in increasing HMGA1 and nuclear  $\beta$ -catenin protein levels, not only at their normal location at the crypt base, but throughout the crypt extending toward the luminal epithelium. Intriguingly, HMGA1 protein levels increase within the colon tumors compared with nontumor colon epithelium, even in the setting of *Hmga1* haploinsufficiency. Precisely how this occurs will require further investigation, although these results underscore the importance of HMGA1 in tumorigenesis in the *CDX2P-CreER<sup>T2</sup>Apc<sup>fl/fl</sup>*. While prior studies show that *HMGA2* is overexpressed in colon cancer and *Hmga2* drives tumorigenesis in mouse models with *Let-7* deficiency (101), we focus on *HMGA1* since transcripts are approximately 100-fold higher than *HMGA2* in colon cancer datasets (TCGA) and in many other human tumors (28–30).

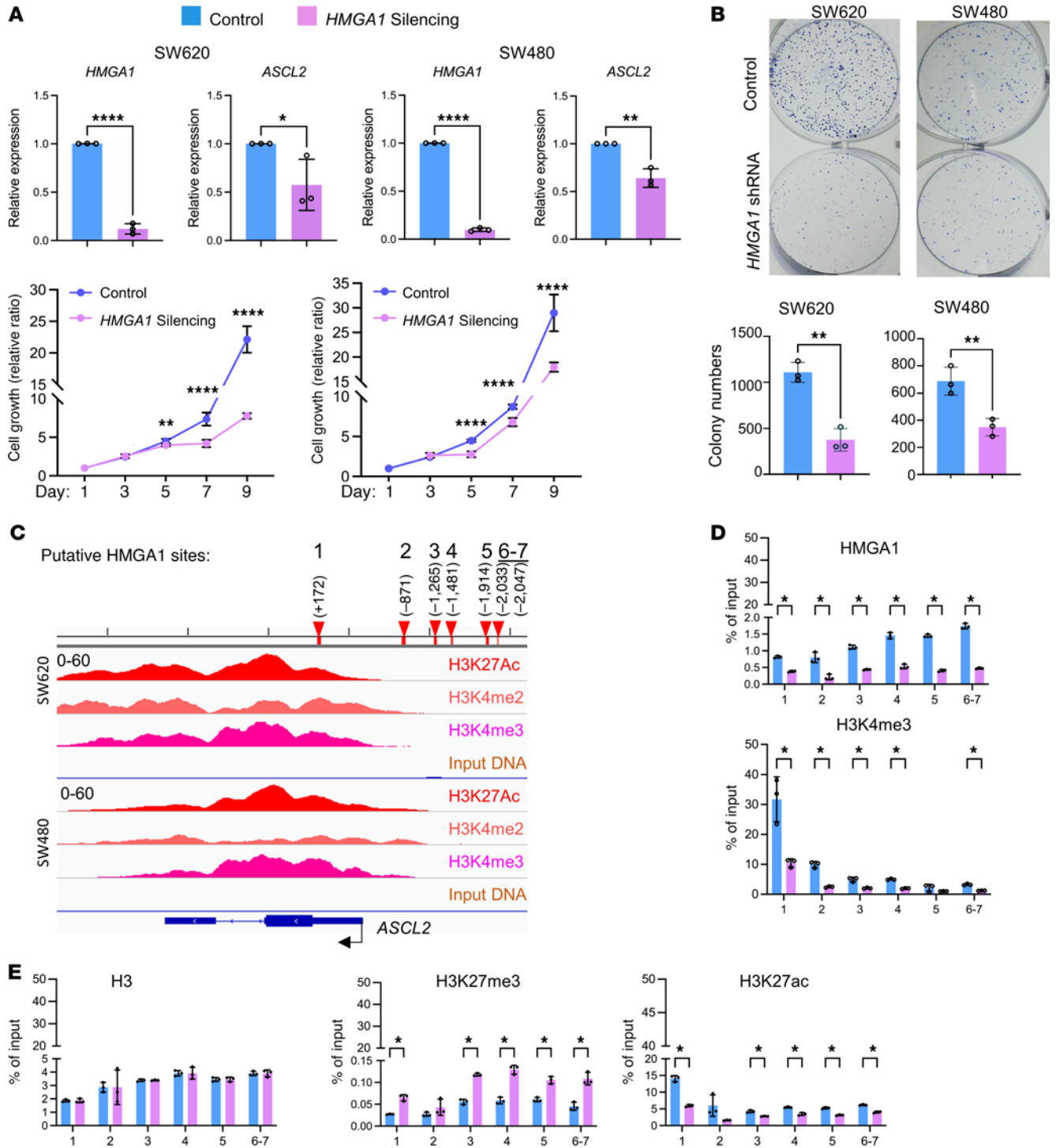
In the Min model following inoculation with ETBF, *Hmga1* haploinsufficiency globally or within the colon epithelium is sufficient to decrease tumorigenesis. Surprisingly, complete loss of *Hmga1* (homozygous deficiency) from colon epithelium decreases tumor incidence similar to that of haploinsufficiency. This was unexpected, since global deletion of *Hmga1* in the biallelic



**Figure 10. HMGA1 enhances chromatin accessibility at gene loci involved in proliferation, DNA repair, and inflammation.** (A) HMGA1 increases chromatin accessibility in crypt cell nuclei globally in *CDX2P-CreER<sup>2</sup> Apc<sup>fl/fl</sup>* mice. (B) HMGA1 enhances chromatin accessibility in promoter regions ranging from 0 to -3 kb upstream of the transcription start sites shown by average peak lengths. (\*\*\*\**P* < 0.0001; student's *t* test). (C) HMGA1 enhances chromatin accessibility in promoter regions ranging from 0 to -3 kb upstream of the transcription start sites shown by number of significantly expanded peaks. (*P* < 0.0001;  $\chi^2$ ). (D) HMGA1 enhances chromatin accessibility in gene sets involved in proliferation, inflammation, and metabolism. (E) GSEA pathways identified by intersecting ATAC-seq and scrNA-seq pathways with associated *P* values.







**Figure 12. *HMGA1* induces *ASCL2* by directly binding to the promoter and recruiting activating histone marks. (A)** Silencing *HMGA1* represses *ASCL2* and decreases proliferation in SW620 and SW480 cells. Control cells were transduced with empty lentiviral vector versus shRNA targeting *HMGA1*. (\**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001, \*\*\*\**P* < 0.0001; student's *t* test). **(B)** Silencing *HMGA1* decreases clonogenicity in SW620 and SW480 (\*\**P* < 0.01, \*\*\*\**P* < 0.0001; student's *t* test). **(C)** Predicted *HMGA1* binding sites 1–5 and region 6–7 in the *ASCL2* promoter region shown with activating histone marks from SW620 and SW480 (GSE106921). **(D)** ChIP assay results at sites 1–5 and region 6–7 in SW620 cells from one representative biological replicate for *HMGA1* and activating histone marks (H3K4me3, and H3K27ac). **(E)** ChIP assay results at sites 1–5 and region 6–7 in SW620 cells from one representative biological replicate for H3 and the repressive histone (H3K27me3). (\**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001, \*\*\*\**P* < 0.0001; student's *t* test following significance by ANOVA).

*Apc*-deficient model led to the greatest impact on tumor development and survival. However, in models of pancreatic tumorigenesis (28), we found a similar relationship whereby tissue-specific loss of just one *Hmga1* allele was sufficient to dampen tumorigenesis and prolong survival, akin to results with tissue-specific loss of both *Hmga1* alleles. Intriguingly, crypt depth is similar in the colon with intact *Hmga1* or complete loss of *Hmga1* in the WT *Apc* epithelial compartment. Based on our transcriptomic data showing that HMGA1 deficiency fosters differentiation of stem cells to enterocytes, we surmise that the crypt depth is maintained in the *Hmga1* deficient model lacking *Apc* mutation through skewing of more quiescent stem cells towards proliferating enterocytes. Unfortunately, there are no pharmacologic inhibitors to directly disrupt HMGA1 function in the clinics, although modulating HMGA1 function or levels by approximately 50% is likely to be a more feasible therapeutic goal than a more comprehensive disruption of its function.

To identify potential therapies to disrupt HMGA1 function, we focused on the epigenetic landscape downstream of HMGA1. HMGA1 enhances chromatin accessibility globally, in addition to “opening” regulatory regions of the genome important for activation of proliferation, inflammation, and Wnt signaling genes, including *Ascl2*. ASCL2 functions as a master regulator of stemness in small intestinal epithelium where it activates its own expression and that of downstream Wnt effector and receptor genes (95–97), and in esophageal cancer (98), although its role in colon epithelium had not been studied in detail. A recent study also links *ASCL2* expression to early-onset colorectal cancer in a Japanese cohort (102).

We also found that HMGA1 binds directly to the *ASCL2* promoter region and recruits activating histones (H3K4me3, H3K27Ac) to upregulate its expression. Indeed, *Hmga1* and *Ascl2* are the most tightly coregulated genes in murine crypt epithelium (Figure 7). *Hmga1* is also coregulated with other Wnt effectors (*Axin2*, *Tcf4*, *Ctmb1*, *cMyc*, *Sox9*) and Wnt agonist receptors (*Lgr5*, *Lrp6*, *Fzd7*, and *Ephb2*). In addition, HMGA1 enhances chromatin accessibility at promoter regions for *Tcf4*, *Prom1*, *Lgr5*, and *Lrp5*, suggesting that it may directly induce these Wnt genes or modify chromatin to facilitate their expression. In small intestinal stem cells and Caco2 cells (a human colon cancer cell line), HMGA1 directly induces *SOX9*, and upregulation of *SOX9* by HMGA1 could enhance expansion in Paneth-like cells, since Paneth cell differentiation in small intestine depends on *SOX9* (30). While classical Paneth cells are not present in the colon, Paneth cell metaplasia occurs in the proximal colon in *Apc* mice (73, 74) and in humans with inflammatory bowel disease (103), although their role in tumorigenesis is not yet clear. *HMGA1* also correlates with *MYC* in human colon cancer and prior work in other settings shows that *MYC* directly induces HMGA1 (70). HMGA1 also binds to the *MYC* promoter to induce its expression in embryonic stem cells (34). In colon crypt cells, HMGA1 enhances chromatin accessibility at the *Hmga1* promoter, suggesting that HMGA1 induces its own expression, a feature common to many stemness transcriptional regulators, such as *ASCL2* (95). Given the link between *ASCL2* and early onset colon cancer (102), further studies to explore *HMGA1* and *ASCL2* are warranted. Restoration of *ASCL2* only partially rescues proliferation in colon cancer cell lines with *HMGA1* silencing, indicating that HMGA1 regulates additional networks during colon tumorigenesis.

We also identified inflammatory and proliferative networks that are induced by HMGA1 and associated with HMGA1-dependent accessible chromatin. Both IFN- $\alpha$  and  $\gamma$  signaling networks are upregulated in crypt epithelial cells by HMGA1, leading to activation in IFN-stimulated genes and inflammatory networks, including *signal transduction and activator of transcription 1 and 2* (*Stat1/2*), and chemokine genes encoding C-X-X motif chemokine ligands 9 and 10 (*CXCL9/10*). Importantly, inflammatory cytokines and/or their receptors are often amenable to pharmacologic blockade. Our scRNA-seq results show that tumor-infiltrating T-lymphocytes increase in the setting of HMGA1 deficiency (Figure 7), suggesting that HMGA1 in colon tumor cells may foster an immunologically “cold” tumor microenvironment to facilitate tumor progression. Alternatively, the changes in T cell number could reflect altered transcriptomes from HMGA1 deficiency and associated changes in cell behavior, including proliferation and motility. Given the immune pathways identified from colon epithelial cells, studies focusing on HMGA1 inflammatory networks and immune escape are warranted and could identify new therapeutic strategies.

In summary, we discovered that HMGA1 acts as a molecular key that “opens” chromatin to activate transcriptional networks that maintain a stem and Paneth-like cell state early in colon tumorigenesis. Within crypt epithelial cells, HMGA1 enhances chromatin accessibility to activate the *Ascl2* master regulator gene, additional Wnt genes, and inflammatory networks in murine models with *Apc* inactivation. Further, in human colon cancer, *HMGA1* and *ASCL2* are coexpressed and upregulated along with downstream Wnt pathway genes. Together, our results establish HMGA1 as an epigenetic gatekeeper of ASCL2 and Wnt signals, inflammation, and a stem-like state in colon cells with *APC* inactivation, highlighting HMGA1 as a promising potential therapeutic target in colon cancer.

## Methods

*Sex as a biologic variable.* All studies were carried out on male and female mouse populations and similar findings were observed for both sexes.

Detailed methods, statistical analyses, and reagents are provided in the supplemental material section, including culture medium, primers, antibodies, and in silico approaches (Supplemental Table 5). Sequencing data were deposited into the Gene Expression Omnibus (GSE) with accession numbers GSE279070 (scRNA-seq) and GSE278871 (ATAC-seq).

*Animal models.* *CDX2P-CreER<sup>T2</sup>Apc<sup>fl/fl</sup>* (73, 74) or *Apc<sup>Min/+</sup>* (Min mice) (75–79) mice were previously described. The *CDX2P-CreER<sup>T2</sup>Apc<sup>fl/fl</sup>* were generated and provided in house at the University of Michigan (73, 74). The *Apc<sup>Min/+</sup>* (Min mice) were originally obtained from Bert Vogelstein at Johns Hopkins University who developed this model (104). Both were crossed to mice with global deficiency of one or both *Hmga1* alleles (all on C57Bl6 backgrounds) (28, 29). Tissue-specific *Hmga1*-deficient models were generated by crossing to mice with floxed *Hmga1* alleles. Additional details are provided in the supplement (Supplemental Data Set 1).

*Statistics.* To compare continuous variables across 2 groups, statistical significance was determined using a 2-tailed student's *t* test when normally distributed (ascertained by Ryan-Joyner and D'Agostino-Pearson tests). If not normal, the Mann-Whitney test was used. To compare more than 2 groups, we used a 1-way ANOVA with Dunnett's or Turkey's multiple comparisons (Prism 10, GraphPad Software) after which 2 groups were compared via 2-tailed student's *t* test if normally distributed or Mann-Whitney if not. For categorical data, association with condition



was evaluated by Fisher's exact test. We compared survival analyses under the assumption of Cox proportional hazards using the log-rank test.  $P < 0.05$  was considered significant. All code for the scRNA-seq analysis was performed using Seurat at the indicated resolutions; code will be made available from the corresponding author upon request.

**Study approvals.** All mouse studies were approved by the Johns Hopkins University Institutional Animal Care and Use Committee (IACUC).

**Data availability.** As above, metadata are available in the NCBI GEO database (access numbers: scRNA-seq: GSE279070 and ATAC-seq: GSE278871); the remaining data are provided in the Supporting Data Values file.

## Author contributions

LR and CLS conceptualized the project; LZL, IH, and BEW drafted parts of the manuscript, and LR wrote the final draft, which was reviewed by all authors prior to submission. LZL,

JHK, IH, SW, XW, SSP, JC, LC, LX, BEW, JCE, JK, ZT, IM, KLR, YF, ERF, CLS, and LR performed experiments and analyzed data. TL interpreted histology.

## Acknowledgments

This research was supported by the National Institutes of Health (R01 CA293602, R01 CA232741, R01 HL145780, R01 DK 102943, R01 HL143818), the Maryland Stem Cell Research Fund, and National Cancer Center grants (Korea; NCC-2311410 and 2310390).

Address correspondence to: Linda Resar, Johns Hopkins University SOM, 720 Rutland Avenue, Baltimore, Maryland 21205, USA. Phone: 410.614.0712; Email: lresar@jhmi.edu. Or to: Cynthia Sears, Johns Hopkins University SOM, 1550 Orleans Street, CRB2, Suite 1M.05, Baltimore, Maryland 21231, USA. Phone: 410.614.8378; Email: csears@jhmi.edu.

- Hanahan D. Hallmarks of cancer: new dimensions. *Cancer Discov.* 2022;12(1):31–46.
- Paul R, et al. Cell plasticity, senescence, and quiescence in cancer stem cells: biological and therapeutic implications. *Pharmacol Ther.* 2022;231:107985.
- Beumer J, Clevers H. Cell fate specification and differentiation in the adult mammalian intestine. *Nat Rev Mol Cell Biol.* 2021;22(1):39–53.
- Resar L, et al. Lessons from the crypt: HMGA1-amping up Wnt for stem cells and tumor progression. *Cancer Res.* 2018;78(8):1890–1897.
- Shibue T, Weinberg RA. EMT, CSCs, and drug resistance: the mechanistic link and clinical implications. *Nat Rev Clin Oncol.* 2017;14(10):611–629.
- Battle E, Clevers H. Cancer stem cells revisited. *Nat Med.* 2017;23(10):1124–1134.
- Shimokawa M, et al. Visualization and targeting of LGR5<sup>+</sup> human colon cancer stem cells. *Nature.* 2017;545(7653):187–192.
- Koo BK, Clevers H. Stem cells marked by the R-spondin receptor LGR5. *Gastroenterology.* 2014;147(2):289–302.
- Ben-Porath I, et al. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet.* 2008;40(5):499–507.
- Frau C, et al. Deciphering the role of intestinal crypt cell populations in resistance to chemotherapy. *Cancer Res.* 2021;81(10):2730–2744.
- Reddy KL, Feinberg AP. Higher order chromatin organization in cancer. *Semin Cancer Biol.* 2013;23(2):109–115.
- Malagola E, et al. Isthmus progenitor cells contribute to homeostatic cellular turnover and support regeneration following intestinal injury. *Cell.* 2024;187(12):3056–3071.
- Capdevila C, et al. Time-resolved fate mapping identifies the intestinal upper crypt zone as an origin of Lgr5<sup>+</sup> crypt base columnar cells. *Cell.* 2024;187(12):3039–3055.
- Li ML, Sumigray K. Redefining intestinal stemness: the emergence of a new ISC population. *Cell.* 2024;187(12):2900–2902.
- Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell.* 1990;61(5):759–767.
- Vogelstein B, et al. Genetic alterations during colorectal-tumor development. *N Engl J Med.* 1988;319(9):525–532.
- Olkinuora AP, et al. From APC to the genetics of hereditary and familial colon cancer syndromes. *Hum Mol Genet.* 2021;30(r2):R206–R224.
- Herrera L, et al. Gardner syndrome in a man with an interstitial deletion of 5q. *Am J Med Genet.* 1986;25(3):473–476.
- Leppert M, et al. The gene for familial polyposis coli maps to the long arm of chromosome 5. *Science.* 1987;238(4832):1411–1413.
- Bodmer WF, et al. Localization of the gene for familial adenomatous polyposis on chromosome 5. *Nature.* 1987;328(6131):614–616.
- Leppert M, et al. Genetic analysis of an inherited predisposition to colon cancer in a family with a variable number of adenomatous polyps. *N Engl J Med.* 1990;322(13):904–908.
- Kinzler KW, et al. Identification of FAP locus genes from chromosome 5q21. *Science.* 1991;253(5020):661–665.
- Queen J, et al. Understanding the mechanisms and translational implications of the microbiome for cancer therapy innovation. *Nat Cancer.* 2023;4(8):1083–1094.
- Stoffel EM, Murphy CC. Epidemiology and mechanisms of the increasing incidence of colon and rectal cancers in young adults. *Gastroenterology.* 2020;158(2):341–353.
- Keum N, Giovannucci E. Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. *Nat Rev Gastroenterol Hepatol.* 2019;16(12):713–732.
- Siegel RL, et al. Cancer statistics, 2022. *CA Cancer J Clin.* 2022;72(1):7–33.
- Reeves R, Beckerbauer L. HMGI/Y proteins: flexible regulators of transcription and chromatin structure. *Biochim Biophys Acta.* 2001;1519(1-2):13–29.
- Chia L, et al. HMGA1 induces FGF19 to drive pancreatic carcinogenesis and stroma formation. *J Clin Invest.* 2023;133(6):e151601.
- Li L, et al. HMGA1 chromatin regulators induce transcriptional networks involved in GATA2 and proliferation during MPN progression. *Blood.* 2022;139(18):2797–2815.
- Xian L, et al. HMGA1 amplifies Wnt signalling and expands the intestinal stem cell compartment and Paneth cell niche. *Nat Commun.* 2017;8:15008.
- Hillion J, et al. The high mobility group A1 (HMGA1) gene is highly overexpressed in human uterine serous carcinomas and carcinosarcomas and drives matrix metalloproteinase-2 (MMP-2) in a subset of tumors. *Gynecol Oncol.* 2016;141(3):580–587.
- Sumter TF, et al. The high mobility group A1 (HMGA1) transcriptome in cancer and development. *Curr Mol Med.* 2016;16(4):353–393.
- Shah SN, et al. HMGA1: a master regulator of tumor progression in triple-negative breast cancer cells. *PLoS One.* 2013;8(5):e63419.
- Shah SN, et al. HMGA1 reprograms somatic cells into pluripotent stem cells by inducing stem cell transcriptional networks. *PLoS One.* 2012;7(11):e48533.
- Belton A, et al. HMGA1 induces intestinal polyposis in transgenic mice and drives tumor progression and stem cell properties in colon cancer cells. *PLoS One.* 2012;7(1):e30034.
- Hillion J, et al. The HMGA1-COX-2 axis: a key molecular pathway and potential target in pancreatic adenocarcinoma. *Pancreatology.* 2012;12(4):372–379.
- Schuldenfrei A, et al. HMGA1 drives stem cell, inflammatory pathway, and cell cycle progression genes during lymphoid tumorigenesis. *BMC Genomics.* 2011;12:549.
- Resar LM. The high mobility group A1 gene: transforming inflammatory signals into cancer? *Cancer Res.* 2010;70(2):436–439.
- Baron RM, et al. Distamycin A inhibits HMGA1-binding to the P-selectin promoter and attenuates lung and liver inflammation during murine endotoxemia. *PLoS One.* 2010;5(5):e10656.
- Hillion J, et al. Upregulation of MMP-2 by HMGA1 promotes transformation in undifferentiated, large-cell lung cancer. *Mol Cancer Res.* 2009;7(11):1803–1812.
- Hillion J, et al. The high-mobility group A1a/signal transducer and activator of transcription-3 axis: an achilles heel for hematopoietic malignan-

- cies? *Cancer Res.* 2008;68(24):10121–10127.
42. Tesfaye A, et al. The high-mobility group A1 gene up-regulates cyclooxygenase 2 expression in uterine tumorigenesis. *Cancer Res.* 2007;67(9):3998–4004.
  43. Treff NR, et al. Human KIT ligand promoter is positively regulated by HMGA1 in breast and ovarian cancer cells. *Oncogene.* 2004;23(52):8557–8562.
  44. Attema JL, et al. The human IL-2 gene promoter can assemble a positioned nucleosome that becomes remodeled upon T cell activation. *J Immunol.* 2002;169(5):2466–2476.
  45. Massaad-Massade L, et al. HMGA1 enhances the transcriptional activity and binding of the estrogen receptor to its responsive element. *Biochemistry.* 2002;41(8):2760–2768.
  46. Kim J, et al. The non-histone chromosomal protein HMG-I(Y) contributes to repression of the immunoglobulin heavy chain germ-line epsilon RNA promoter. *Eur J Immunol.* 1995;25(3):798–808.
  47. Chou BK, et al. Efficient human iPS cell derivation by a non-integrating plasmid from blood cells with unique epigenetic and gene expression signatures. *Cell Res.* 2011;21(3):518–529.
  48. Thanos D, Maniatis T. The high mobility group protein HMG I(Y) is required for NF-kappa B-dependent virus induction of the human IFN-beta gene. *Cell.* 1992;71(5):777–789.
  49. Zhao K, et al. SAR-dependent mobilization of histone H1 by HMG-I/Y in vitro: HMG-I/Y is enriched in H1-depleted chromatin. *EMBO J.* 1993;12(8):3237–3247.
  50. Saitoh Y, Laemmli UK. Metaphase chromosome structure: bands arise from a differential folding path of the highly AT-rich scaffold. *Cell.* 1994;76(4):609–622.
  51. Geierstanger BH, et al. Short peptide fragments derived from HMG-I/Y proteins bind specifically to the minor groove of DNA. *Biochemistry.* 1994;33(17):5347–5355.
  52. Thanos D, Maniatis T. Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome. *Cell.* 1995;83(7):1091–1100.
  53. Falvo JV, et al. Reversal of intrinsic DNA bends in the IFN beta gene enhancer by transcription factors and the architectural protein HMG I(Y). *Cell.* 1995;83(7):1101–1111.
  54. Maher JF, Nathans D. Multivalent DNA-binding properties of the HMG-I proteins. *Proc Natl Acad Sci U S A.* 1996;93(13):6716–6720.
  55. Huth JR, et al. The solution structure of an HMG-I(Y)-DNA complex defines a new architectural minor groove binding motif. *Nat Struct Biol.* 1997;4(8):657–665.
  56. Gorbounov M, et al. High mobility group A1 (HMGA1) protein and gene expression correlate with ER-negativity and poor outcomes in breast cancer. *Breast Cancer Res Treat.* 2020;179(1):25–35.
  57. Williams MD, et al. HMGA1 drives metabolic reprogramming of intestinal epithelium during hyperproliferation, polyposis, and colorectal carcinogenesis. *J Proteome Res.* 2015;14(3):1420–1431.
  58. Di Cello F, et al. Inactivation of the Cdkn2a locus cooperates with HMGA1 to drive T-cell leukemogenesis. *Leuk Lymphoma.* 2013;54(8):1762–1768.
  59. Roy S, et al. HMGA1 overexpression correlates with relapse in childhood B-lineage acute lymphoblastic leukemia. *Leuk Lymphoma.* 2013;54(11):2565–2567.
  60. Nelson DM, et al. Flavopiridol induces BCL-2 expression and represses oncogenic transcription factors in leukemic blasts from adults with refractory acute myeloid leukemia. *Leuk Lymphoma.* 2011;52(10):1999–2006.
  61. Karp JE, et al. Phase I and pharmacokinetic study of bolus-infusion flavopiridol followed by cytosine arabinoside and mitoxantrone for acute leukemias. *Blood.* 2011;117(12):3302–3310.
  62. Grade M, et al. Gene expression profiling reveals a massive, aneuploidy-dependent transcriptional deregulation and distinct differences between lymph node-negative and lymph node-positive colon carcinomas. *Cancer Res.* 2007;67(1):41–56.
  63. Liao S-S, et al. HMGA1 is a determinant of cellular invasiveness and in vivo metastatic potential in pancreatic adenocarcinoma. *Cancer Res.* 2006;66(24):11613–11622.
  64. Xu Y, et al. The HMG-I oncogene causes highly penetrant, aggressive lymphoid malignancy in transgenic mice and is overexpressed in human leukemia. *Cancer Res.* 2004;64(10):3371–3375.
  65. Lanahan A, et al. Growth factor-induced delayed early response genes. *Mol Cell Biol.* 1992;12(9):3919–3929.
  66. Holth LT, et al. Effects of epidermal growth factor and estrogen on the regulation of the HMG-I/Y gene in human mammary epithelial cell lines. *DNA Cell Biol.* 1997;16(11):1299–1309.
  67. Bush BM, et al. The Wnt/ $\beta$ -catenin/T-cell factor 4 pathway up-regulates high-mobility group A1 expression in colon cancer. *Cell Biochem Funct.* 2013;31(3):228–236.
  68. Zeitels LR, et al. Tumor suppression by miR-26 overrides potential oncogenic activity in intestinal tumorigenesis. *Genes Dev.* 2014;28(23):2585–2590.
  69. Veite-Schmahl MJ, et al. HMGA1 expression levels are elevated in pancreatic intraepithelial neoplasia cells in the Ptf1a-Cre; LSL-KrasG12D transgenic mouse model of pancreatic cancer. *Br J Cancer.* 2017;117(5):639–647.
  70. Wood LJ, et al. HMG-I/Y, a new c-Myc target gene and potential oncogene. *Mol Cell Biol.* 2000;20(15):5490–5502.
  71. Dhar A, et al. Dominant-negative c-Jun (TAM67) target genes: HMGA1 is required for tumor promoter-induced transformation. *Oncogene.* 2004;23(25):4466–4476.
  72. Hommura F, et al. HMG-I/Y is a c-Jun/activator protein-1 target gene and is necessary for c-Jun-induced anchorage-independent growth in Rat1a cells. *Mol Cancer Res.* 2004;2(5):305–314.
  73. Feng Y, et al. Tissue-specific effects of reduced  $\beta$ -catenin expression on adenomatous polyposis coli mutation-instigated tumorigenesis in mouse colon and ovarian epithelium. *PLoS Genet.* 2015;11(11):e1005638.
  74. Feng Y, et al. Sox9 induction, ectopic Paneth cells, and mitotic spindle axis defects in mouse colon adenomatous epithelium arising from conditional biallelic Apc inactivation. *Am J Pathol.* 2013;183(2):493–503.
  75. Allen J, et al. Colon tumors in enterotoxigenic bacteroides fragilis (ETBF)-colonized mice do not display a unique mutational signature but instead possess host-dependent alterations in the APC gene. *Microbiol Spectr.* 2022;10(3):e0105522.
  76. Wu S, et al. Bacteroides fragilis enterotoxin cleaves the zonula adherens protein, E-cadherin. *Proc Natl Acad Sci U S A.* 1998;95(25):14979–14984.
  77. Wu S, et al. A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. *Nat Med.* 2009;15(9):1016–1022.
  78. Geis AL, et al. Regulatory T-cell response to enterotoxigenic bacteroides fragilis colonization triggers IL17-dependent colon carcinogenesis. *Cancer Discov.* 2015;5(10):1098–1109.
  79. Housseau F, et al. Redundant innate and adaptive sources of IL17 production drive colon tumorigenesis. *Cancer Res.* 2016;76(8):2115–2124.
  80. Dejea CM, et al. Microbiota organization is a distinct feature of proximal colorectal cancers. *Proc Natl Acad Sci U S A.* 2014;111(51):18321–18326.
  81. Boleij A, et al. The Bacteroides fragilis toxin gene is prevalent in the colon mucosa of colorectal cancer patients. *Clin Infect Dis.* 2015;60(2):208–215.
  82. Llosa NJ, et al. The vigorous immune microenvironment of microsatellite instable colon cancer is balanced by multiple counter-inhibitory checkpoints. *Cancer Discov.* 2015;5(1):43–51.
  83. Bullman S, et al. Analysis of Fusobacterium persistence and antibiotic response in colorectal cancer. *Science.* 2017;358(6369):1443–1448.
  84. Sears CL. Microbes and cancer: disease drivers, passengers, biomarkers, or therapeutics? *Cancer Metastasis Rev.* 2022;41(2):247–248.
  85. White MT, Sears CL. The microbial landscape of colorectal cancer. *Nat Rev Microbiol.* 2024;22(4):240–254.
  86. Vega PN, et al. Cancer-associated fibroblasts and squamous epithelial cells constitute a unique microenvironment in a mouse model of inflammation-induced colon cancer. *Front Oncol.* 2022;12:878920.
  87. Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005;102(43):15545–15550.
  88. Mootha V, et al. PGC-1 $\alpha$ -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet.* 2003;34(3):267–273.
  89. Gerace D, et al. Engineering human stem cell-derived islets to evade immune rejection and promote localized immune tolerance. *Cell Rep Med.* 2023;4(1):100879.
  90. Jackson BT, Finley LWS. Metabolic regulation of the hallmarks of stem cell biology. *Cell Stem Cell.* 2024;31(2):161–180.
  91. Chandel NS, et al. Metabolic regulation of stem cell function in tissue homeostasis and organismal ageing. *Nat Cell Biol.* 2016;18(8):823–832.
  92. Burgess RJ, et al. Metabolic regulation of stem cell function. *J Intern Med.* 2014;276(1):12–24.
  93. Ko CW, et al. Regulation of intestinal lipid metabolism: current concepts and relevance to disease. *Nat Rev Gastroenterol Hepatol.* 2020;17(3):169–183.
  94. Buenostro JD, et al. Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature.* 2015;523(7561):486–490.
  95. Murata K, et al. Ascl2-dependent cell dedifferentiation drives regeneration of ablated intestinal stem cells. *Cell Stem Cell.* 2020;26(3):377–390.
  96. Schuijers J, et al. Ascl2 acts as an R-spondin/Wnt-responsive switch to control stemness in intes-

- tinal crypts. *Cell Stem Cell*. 2015;16(2):158–170.
97. Van der Flier LG, et al. Transcription factor achaete scute-like 2 controls intestinal stem cell fate. *Cell*. 2009;136(5):903–912.
98. Hamilton M, et al. ASCL2 is a key regulator of the proliferation-differentiation equilibrium in the esophageal epithelium. *Biol Open*. 2024;13(1):bio059919.
99. Thomas-Chollier M, et al. Transcription factor binding predictions using TRAP for the analysis of ChIP-seq data and regulatory SNPs. *Nat Protoc*. 2011;6(12):1860–1869.
100. Ohta Y, et al. Cell-matrix interface regulates dormancy in human colon cancer stem cells. *Nature*. 2022;608(7924):784–794.
101. Madison BB, et al. Let-7 represses carcinogenesis and a stem cell phenotype in the intestine via regulation of Hmga2. *PLoS Genet*. 2015;11(8):e1005408.
102. Yokota K, et al. WnTRLINC1/ASCL2/c-Myc axis characteristics of colon cancer with differentiated histology at young onset and essential for cell viability. *Ann Surg Oncol*. 2019;26(13):4826–4834.
103. Tanaka M, et al. Spatial distribution and histogenesis of colorectal Paneth cell metaplasia in idiopathic inflammatory bowel disease. *J Gastroenterol Hepatol*. 2001;16(12):1353–1359.
104. Su K, et al. Multiple intestinal neoplasia caused by a mutation in the murine homolog of the APC gene. *Science*. 1992;256(5057):668–670.