Nucleic Acids Research, 2026, **54**, gkag634
https://doi.org/10.1093/nar/gkag634
Gene regulation, Chromatin and Epigenetics

Dynamics and master transcription factors dependence of intestinal super-enhancers during differentiation and oncogenesis

Wenxin Fang[1], Shuya Huang[1], Ruoxuan Xiao[1], Guangling Wei[1], Mingyang Zhang[1], Xia Qiu[2], Li Zou[1], Michael P. Verzi[2], Yan Wang[3,*], Lei Chen[1,4,*]

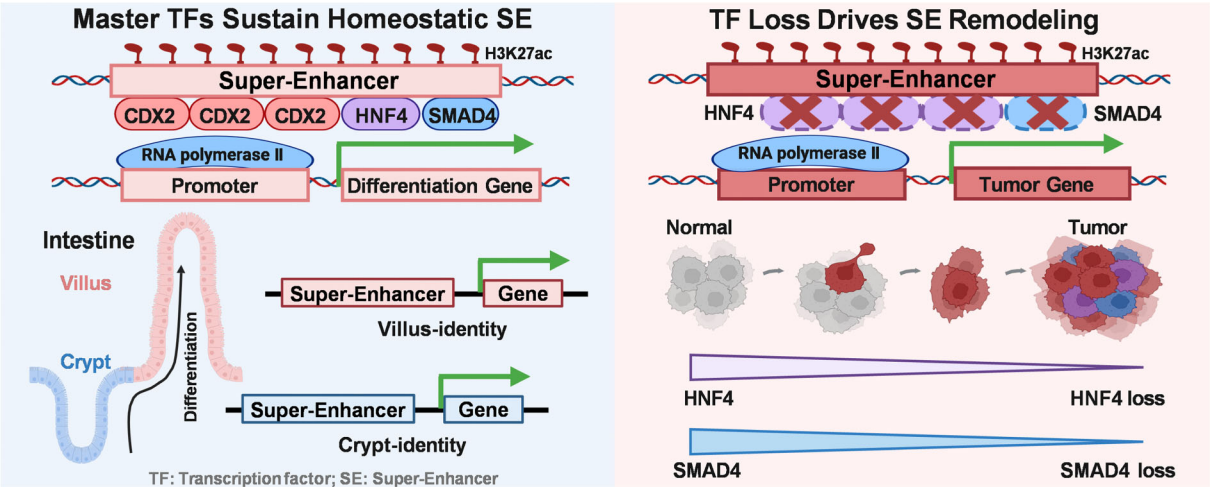
[1]Jiangsu Provincial Key Laboratory of Critical Care Medicine, Key Laboratory of Developmental Genes and Human Disease, School of Life Science and Technology, Southeast University, Nanjing 210031, China
[2]Department of Genetics, Rutgers University, Piscataway, NJ 08854, United States
[3]Center for Translational Medicine Research and Development, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China
[4]Institute of Microphysiological Systems, Southeast University, Nanjing 211189, China

*To whom correspondence should be addressed. Email: leichen@seu.edu.cn
Correspondence may also be addressed to Yan Wang. Email: yan.wang@siat.ac.cn

Abstract

The intestinal epithelium is a highly regenerative tissue organized along the crypt–villus axis, where spatially compartmentalized gene expression governs stem cell renewal, proliferation, and differentiation. Super-enhancers (SEs) are large clusters of regulatory elements densely bound by transcription factors (TFs) and cofactors that drive high-level expression of genes controlling cell identity and fate, yet their roles in intestinal epithelial identity and differentiation remain unclear. Here, we generate a spatiotemporal map of SEs in the small intestine, identifying compartment-specific SEs that define crypt and villus programs. Using mouse genetic models, we identify CDX2, HNF4, and SMAD4 as core TFs orchestrating SE-driven transcriptional networks essential for epithelial differentiation. CDX2 is required for SE integrity, and its loss causes widespread SE collapse and silencing of intestinal identity genes. We further demonstrate SE remodeling during colorectal cancer, in which HNF4 and SMAD4 function as SE-associated tumor suppressors that restrain oncogenic enhancer programs. Together, our findings establish SEs as central regulators of intestinal architecture, epithelial fidelity, and tumor progression.

Graphical abstract



Received: January 9, 2026. **Revised:** May 30, 2026. **Accepted:** June 4, 2026

Introduction

Gene expression in multicellular organisms is tightly controlled by intricate networks of transcriptional regulatory elements, ensuring that each cell maintains its identity and functions appropriately. Among these elements, super-enhancers (SEs) have emerged as central regulatory hubs that drive high-level transcription of key genes crucial for defining and maintaining cell identity [1]. SEs are extensive genomic regions densely occupied by transcription factors (TFs), cofactors, the mediator complex, RNA polymerase II (RNA Pol II), and active histone marks such as H3K27ac. Compared to typical enhancers (TEs), SEs exhibit superior transcriptional activation potency, orchestrating gene expression programs essential for cell fate specification and functional integrity [2, 3].

In mammals, SEs serve as core components of transcriptional regulatory circuits, where master TFs engage in autoregulatory and feedforward loops to coordinate developmental and homeostatic processes [4]. Dysregulation of SEs has been increasingly implicated in various pathophysiological processes, including cancer, immune dysfunction, and drug resistance [5]. In tumors, aberrant formation of oncogenic SEs—through chromosomal rearrangements, focal amplifications, or abnormal expression of TF—can drive constitutive expression of oncogenes, sustaining malignant phenotypes and disease progression [6]. These insights have led to growing interest in targeting SEs and their associated transcriptional machinery as a promising therapeutic avenue [7, 8].

Technological advancements such as chromatin immunoprecipitation sequencing (ChIP-seq) have enabled genome-wide profiling of enhancer and SE landscapes with high resolution. Complementary approaches—ranging from 3C-seq [9], 4C-seq [10, 11], Hi-C [12, 13], and Capture Hi-C [14] to chromatin accessibility assays like DNase-seq [15] and ATAC-seq [16]—have further enhanced our ability to study these regulatory elements in detail. Despite these advances, the regulatory landscape of SEs in the intestinal epithelium—a highly regenerative tissue organized along a crypt–villus axis—remains poorly defined. It remains unclear how SEs contribute to intestinal architecture, epithelial identity, and disease susceptibility.

In this study, we present the first comprehensive analysis of SEs in the small intestine. Using integrated multi-omics approaches [ChIP-seq, RNA-sequencing (RNA-seq)], we mapped the SE landscape and identified caudal type homeobox 2 (CDX2), hepatocyte nuclear factor 4 (HNF4), and SMAD family member 4 (SMAD4) as core TFs that modulate SE activity to maintain intestinal identity or function. We show that HNF4 loss activates latent oncogenic enhancers. The SEs emerging upon the loss of HNF4 or SMAD4 also shows stronger H3K27ac signal in late-stage colorectal tumors. These findings not only fill a critical knowledge gap in intestinal enhancer biology but also provide a molecular framework for understanding disorders of epithelial differentiation and developing SE-targeted therapies for intestinal cancers.

Materials and methods

Mice

To investigate gene function within the intestinal epithelium, conditional compound-mutant mice and relevant controls were generated by integrating the following alleles: the *Villin-Cre*^{ERT2} transgene [17] for tamoxifen-inducible, epithelium-specific recombination, *Cdx2*$^{f/f}$ (ref. [18]), *Hnf4α*$^{f/f}$ (ref. [19]), *Hnf4γ*$^{Crispr/Crispr}$ (ref. [20]), and *Smad4*$^{f/f}$ (ref. [21]). Experimental mice (aged 8–12 weeks) received intraperitoneal injections of tamoxifen (Sigma T5648) to induce recombination. For histologic analysis of wild-type (WT) and *Cdx2*^{KO}, mice received four consecutive daily injections of tamoxifen (50 mg kg^{-1}d^{-1}). Tissues were collected at day 8 after tamoxifen treatment. For RNA-seq analysis of WT and *Hnf4αγ*^{DKO} samples, epithelial cells of duodenum were collected after two or three consecutive days of tamoxifen treatment (50 mg kg^{-1}d^{-1}). ChIP-seq of H3K27ac (*Hnf4αγ*^{DKO} versus WT and *Smad4*^{KO} versus WT) assays utilized tissue harvested after three consecutive days of tamoxifen treatment (50 mg kg^{-1}d^{-1}) [20]. The *Apc*$^{Min/+}$ mice [22] (Jackson Laboratory, Bar Harbor, ME) were euthanized at disease onset for tumor and tissue collection. All mouse procedures were conducted in compliance with institutional guidelines and approved by the Animal Care Committee.

Isolation of intestinal epithelium, crypts, and villi

Fresh intestinal tissue was flushed with cold phosphate buffered saline (PBS), opened longitudinally, and divided into 1-cm segments. Fragments underwent sequential incubation in 3 mM ethylenediaminetetraacetic acid (EDTA)/PBS at 4°C with rotation: 5, 10, and 25–40 min, with EDTA refreshed between incubations. Epithelium was dissociated by vigorously shaken. The supernatant containing whole epithelium was collected or filtered through a 70-μm cell strainer: villi were retained on the strainer while crypts passed through. Cells were pelleted by centrifugation (170 g, 4°C, 3 min) and washed with cold PBS. The epithelial, crypt, or villus pellets were collected for RNA extraction or ChIP-seq as described in later sections.

ChIP-seq

One percent formaldehyde cross-linking (4°C, 10 min; RT, 35–50 min) was used for HNF4A/G ChIP, while cross-linking with 2 mM disuccinimidyl glutarate (DSG, 45 min, RT) followed by formaldehyde (20 min, RT) was employed for SMAD4 ChIP. Cells were lysed [1% sodium dodecyl sulfate (SDS), 10 mM EDTA, 50 mM Tris, and protease inhibitors], sonicated to 200–500 bp fragments, and incubated overnight at 4°C with antibody-coupled Dynabeads (HNF4A: Santa Cruz sc-6556 X; HNF4G: Santa Cruz sc-6558 X; SMAD4: Cell Signaling 46535) in binding buffer (final concentration of SDS: 0.125% for HNF4 ChIP-seq, 0.167% for SMAD4 ChIP-seq). After standard washes and reverse cross-linking, purified DNA was quantified with enrichment validation. ChIP-seq libraries were prepared from 5 ng DNA (Rubicon ThruPLEX Kit) and sequenced on Illumina platforms [20].

For micrococcal nuclease (MNase)-chromatin immunoprecipitation followed by sequencing (ChIP-seq), the villus or crypt cells were suspended in digestion buffer containing 50 mM Tris–HCl (pH 7.6), 1 mM $CaCl_2$, 0.2% Triton X-100, 5 mM sodium butyrate, and complete protease inhibitor cocktail. Chromatin was digested with 0.2 U micrococcal nuclease (MNase) for 8 min at 37 °C. Immunoprecipitation was performed using antibodies against H3K27ac (Abcam, ab4729) or H3K4me2 (Millipore, 07-030). Library preparation and high-throughput sequencing followed established protocols [18, 23].

For MNase-ChIP in TF deficient models ($Hnf4\alpha\gamma^{DKO}$ and $Smad4^{KO}$) and their controls, isolated intestinal epithelial cells were suspended in digestion buffer (50 mM Tris–HCl, pH 7.6, 1 mM $CaCl_2$, 0.2% Triton X-100, 5 mM sodium butyrate, and protease inhibitors). Chromatin was fragmented using 0.67 U ml^{-1} MNase (Sigma, N3755) for 5 min at 37 °C. The reaction was quenched with 5 mM EDTA in 10 mM Tris, pH 7.6, and samples were dialyzed in RIPA chromatin buffer (10 mM Tris, pH 7.6, 1 mM EDTA, 0.1% SDS, 0.1% sodium deoxycholate, 1% Triton X-100). Overnight immunoprecipitation at 4°C employed an anti-H3K27ac antibody (Abcam, ab4729). Subsequent steps followed standard sonicated ChIP protocols, including adjustment to 0.13% SDS [20].

Histology and staining

Freshly harvested tissues, including small intestine, cerebellum, cortex, heart, kidney, liver, olfactory bulb, spleen, testis, and thymus, were fixed in 4% paraformaldehyde at 4°C overnight, washed with PBS, and dehydrated through ascending alcohols before paraffin embedding. The 5-μm sections were processed for immunohistochemistry using primary antibodies against: HNF4A (Abcam ab41898, 1:2000), HNF4G (Santa Cruz sc-6558 X, 1:2 000), EpCAM (ABclonal A25344, 1:1000), Cyclin D1 (CST 55506S, 1:600), Keratin 20 (CST 13063T, 1:2500), and SMAD4 (Santa Cruz sc-7966, 1:500). Detection was performed using species-appropriate biotinylated secondary antibodies with the Vectastain ABC HRP Kit (Vector Labs), followed by DAB development (ZSGB-BIO ZLI-9018) and hematoxylin counterstaining. Mounted slides were imaged on a Nikon Eclipse E800 microscope equipped with a Retiga 1300 CCD (charge-coupled device) camera (QImaging) and QCapture acquisition software [20].

RNA-seq data analysis

Epithelial cells were isolated from the mouse duodenum and processed for RNA extraction using TRIzol (Invitrogen) according to the manufacturer's instructions. Illumina's TruSeq RNA Library Prep kit v.2 was used for RNA-seq. All RNA-seq datasets were reprocessed using a standardized computational pipeline. Transcript-level quantification was performed with Kallisto [24] (v0.44.0), employing the pseudoalignment algorithm for efficient and accurate abundance estimation. Differential expression analysis was subsequently conducted using DESeq2 [25] (v1.36.0) within the R/Bioconductor environment. Transcript abundance was normalized using the Fragments Per Kilobase of transcript per Million mapped reads (FPKM) metric. Gene set enrichment analysis (GSEA) [26] was employed to examine the concordance patterns of SE-associated genes across the different sample groups, assessing whether these functionally related gene sets showed coordinated expression changes. Visualization and clustering of normalized expression levels for genes of interest were executed using Morpheus (https://software.broadinstitute.org/morpheus) and pheatmap (v1.0.13) (https://CRAN.R-project.org/package=pheatmap.). This included the generation of heatmaps and the application of clustering algorithms to identify coherent expression patterns within the dataset.

ChIP-seq data analysis

Raw sequencing reads (FASTQ format) underwent quality control using FastQC (v0.12.1). Reads were then aligned to mouse (mm9) reference genome using Bowtie2 [27] (v2.5.1) with default parameters, generating BAM format alignment files. BigWig coverage files were created from BAM files via deepTools [28] (v3.5.4) bamCoverage with RPKM normalization, read extension to fragment length, and exclusion of duplicate reads. Genomic interval operations (merge/intersect/subtract) were performed using BEDTools [29] (v2.31.1). Chromatin state dynamics were profiled with Haystack [30] (v0.5.5) using quantile-normalized BigWigs. Peak calling was executed with MACS2 [31] (v2.2.6) to generate BED-format peak files. Finally, normalized BigWig tracks were visualized in the Integrative Genomics Viewer [32] (IGV v2.17.1).

ChIP-seq peaks for (i) H3K27ac of the small intestine and nine other mouse tissues were called using MACS2 with a *P*-value cutoff of .05; (ii) H3K27ac and H3K4me2 of intestinal crypts and villi were called with a *q*-value cutoff of 0.05; (iii) CDX2 of intestine were called with a *P*-value cutoff of .01; (iv) H3K27ac and RNA Pol II in WT and $Cdx2^{KO}$ intestines were called with a *P*-value cutoff of .01; (v) HNF4 and SMAD4 of intestines were called with a *P*-value cutoff of .05; (vi) H3K27ac and RNA Pol II in WT and $Hnf4\alpha\gamma^{DKO}$ intestines were called with a *P*-value cutoff of .01; (vii) H3K27ac in WT and $Smad4^{KO}$ intestines were called with a *q*-value cutoff of 0.01; (viii) H3K27ac in Ctrl and azoxymethane (AOM)–dextran sulfate sodium (DSS) model intestines were called with a *P*-value cutoff of .01.

Identifying super-enhancers

SEs were identified from ChIP-seq data using the rank ordering of SEs (ROSE) [4] algorithm. The analysis was performed with a stitching distance of 12.5 kb (-s 12 500) and a transcription start site (TSS) exclusion zone size of 2.5 kb (-t 2500). Enhancer elements within 12.5 kb of each other were first stitched into contiguous domains. All enhancers were then rank-ordered based on H3K27ac signal intensity. The classification threshold between SEs and TEs was determined from a stitching plot, where the point of transition was defined as the position on the rank-ordered plot where the tangent slope equals 1. Regions above this threshold were classified as SEs, and those below as TEs. Downstream analyses were performed using DeepTools [28]. Principal component analysis (PCA) of identified SEs was conducted with the plotPCA function. Sample correlations were analyzed and visualized using plotCorrelation. Signal matrices were generated with computeMatrix, and aggregate signal profiles across SE regions were created with plotProfile.

scRNA-seq data analysis

Single-cell RNA sequencing (scRNA-seq) data from tumor and adjacent normal tissues of CRC patients were processed using Scanpy [33] (v1.10.2) in Python (v3.12.5). The processed single-cell data (.csv file from GSE200997) were loaded, and the expression levels of genes were examined within the epithelial cell population.

Statistical analysis

RNA-seq data are expressed as mean ± SEM; statistical significance was determined by DESeq2, with thresholds defined as $P < .05$ (*), $P < .01$ (**), and $P < .001$ (***). For comparative box plots visualizing log(FPKM + 1) values between two groups, statistical significance was assessed using the two-sided Mann–Whitney U-test. The Kolmogorov–Smirnov test was employed for GSEA. Bioinformatics statistical evaluations utilized embedded algorithms within respective analyti-

cal packages (GSEA, DAVID [34]). All analyses maintained a 95% confidence interval, with $P < .05$ considered statistically significant.

Results

Super-enhancers in intestinal tissue exhibit marked tissue specificity

To systematically decipher tissue-specific gene regulatory networks, we mapped the enhancer landscape across ten tissues—small intestine, cerebellum, cortex, heart, kidney, liver, olfactory bulb (OlfactBulb), spleen, testis, and thymus—by integrating H3K27ac ChIP-seq and RNA-seq data curated from published datasets [35]. SEs were identified using the standardized ROSE algorithm [4]. Briefly, H3K27ac peaks called by MACS2 and located within 12.5 kb of each other were stitched together to form enhancer clusters. SEs were defined as clusters above the inflection point (slope = 1) in the rank-ordered plot of H3K27ac signal density, distinguishing them from TEs (Fig. 1a) [36].

Cross-tissue comparison revealed that both SEs and TEs display strict tissue specificity. SEs derived from a given organ, small intestine, showed robust H3K27ac enrichment, with markedly diminished signals across the nine other tissues (Supplementary Fig. S1a and Supplementary Table S1). TEs exhibited a similar distribution but with much lower signals (Supplementary Fig. S1b). These patterns suggest that enhancers are intrinsically tissue-specific, driving distinct transcriptional programs. SEs exhibit stronger H3K27ac enrichment and more restricted tissue-specific activity than TEs, underscoring their pivotal role in defining tissue identity. PCA of the SE landscapes across nine tissues further illustrated this specificity. Neural tissues (cerebellum, cortex, and olfactory bulb) clustered tightly, indicating a conserved SE program. Meanwhile, intestinal SEs were highly distinct from these neuronal tissues (Fig. 1b and c), but it is not evident that there are substantial differences between intestine and several other tissues (e.g. kidney). This divergence suggests that intestinal epithelium is governed by a unique SE-driven regulatory architecture.

Gene Ontology analysis showed that small intestine-specific SE-associated genes were involved in key epithelial processes, such as cell migration and brush border (Fig. 1d). RNA-seq analysis confirmed that their expression was enriched in the intestine (Fig. 1e), aligning with their functions and supporting a tight link between SE activation, gene regulation, and tissue identity.

ChIP-seq profiling (Fig. 1f) revealed classic SE features, including broad, high-intensity H3K27ac peaks at the gene locus of epithelial cell adhesion molecule (EpCAM), an epithelial marker. Both transcript (Fig. 1g) and protein levels (Fig. 1h) of EpCAM were markedly higher in the intestine than in other tissues. HNF4A, a key TF critical for intestinal, liver, and kidney function [20, 37–39], also exhibited SE activity (Supplementary Fig. S1c) and showed elevated expression across all three tissues (Supplementary Fig. S1d and e).

Spatiotemporally restricted super-enhancers orchestrate axial patterning along the crypt-villus axis

To identify epigenetic drivers of intestinal epithelial differentiation, we mapped enhancer landscapes of crypts (stem/progenitor zone) and villi (terminally differentiated compartment) using H3K27ac and H3K4me2 ChIP-seq (Fig. 2a). ROSE analysis identified 445 crypt-specific and 895 villus-specific SEs (Fig. 2b and Supplementary Table S2), indicating a spatially partitioned regulatory program. The activity of SEs exhibits a highly localized pattern. Venn analysis revealed distinct SE sites: 122 crypt-specific, 572 villus-specific, and 323 shared sites (Fig. 2c). Crypt SEs showed significantly higher H3K27ac signals in crypts than villi (Fig. 2d), while villus SEs were enriched in villi (Fig. 2e). This bidirectional gradient highlights the capacity of SEs to mediate precise spatial regulation at the sub-tissue scale. RNA-seq confirmed functional compartmentalization—crypt SE genes were highly expressed in crypts, while villus SE genes were enriched in villi (Fig. 2f). Functional enrichment linked crypt SE genes to proliferation and cell cycle (Fig. 2g), while villus SE genes were associated with differentiation and metabolism (Fig. 2h).

Cyclin D1 (*Ccnd1*), a crypt-specific SE associated with proliferation, showed crypt-restricted H3K27ac enrichment (Fig. 2i), elevated RNA expression (Fig. 2j), and protein localization to the crypts (Fig. 2k). Conversely, keratin 20 (*Krt20*), a differentiation marker, displayed villus-specific SE activity (Fig. 2l), elevated transcript expression (Fig. 2m), and apical localization in terminally differentiated enterocytes (Fig. 2n). These findings suggest that spatially confined SEs act as epigenetic switches along the crypt–villus axis, coordinating proliferation and differentiation, and establishing regional identity through precise transcriptional regulation.

Building on the compartment-specific roles of H3K27ac-marked SEs (Fig. 2), we extended our profiling to H3K4me2, a mark of transcriptionally permissive chromatin, to assess broader sub-tissue epigenetic regulation in the intestinal epithelium. H3K4me2 ChIP-seq identified 1 122 crypt and 1 084 villus SEs (Supplementary Fig. S2a and b; Supplementary Table S2). H3K4me2-marked SEs exhibit spatial restriction—crypt SEs were enriched in crypts, and villus SEs were abundant in villi (Supplementary Fig. S2c). Functional enrichment analysis revealed that crypt SE-associated genes are linked to stem/progenitor programs (e.g. Notch signaling, chromatin remodeling), while villus SE-associated genes are enriched in differentiation pathways (e.g. absorption, metabolism) (Supplementary Fig. S2d and e). Importantly, SE-regulated genes showed higher overall expression than TE-associated genes in both crypts and villi (Supplementary Fig. S2f). Integration with RNA-seq further confirmed that SE-associated genes maintained consistent compartment-specific expression across histone marks (Supplementary Fig. S2g and h).

We utilized H3K4me2 mapping in conjunction with H3K27ac to enhance the confidence in SE identification. This dual approach provided a more stringent definition of active enhancer clusters and offered insight into their maturation state, as H3K4me2 marks primed chromatin landscapes that may persist beyond transient H3K27ac fluctuations. Most H3K27ac SE genes were co-marked by H3K4me2 within the same tissue (Supplementary Fig. S2i–n), indicating convergent epigenetic regulation. Dual-marked SE genes further supported compartment-specific regulation. *Myb* (crypt-specific, co-marked by H3K27ac and H3K4me2) exhibited SE activation and crypt-specific expression (Supplementary Fig. S2j and k). Conversely, solute carrier family 2 member 2 (*Slc2a2*, villus-specific, co-marked) displayed villus-restricted SE activity and expression (Supplementary Fig. S2m and n). While most SEs were co-marked by both histone modifications, we also found a subset of SEs marked exclusively by H3K4me2.

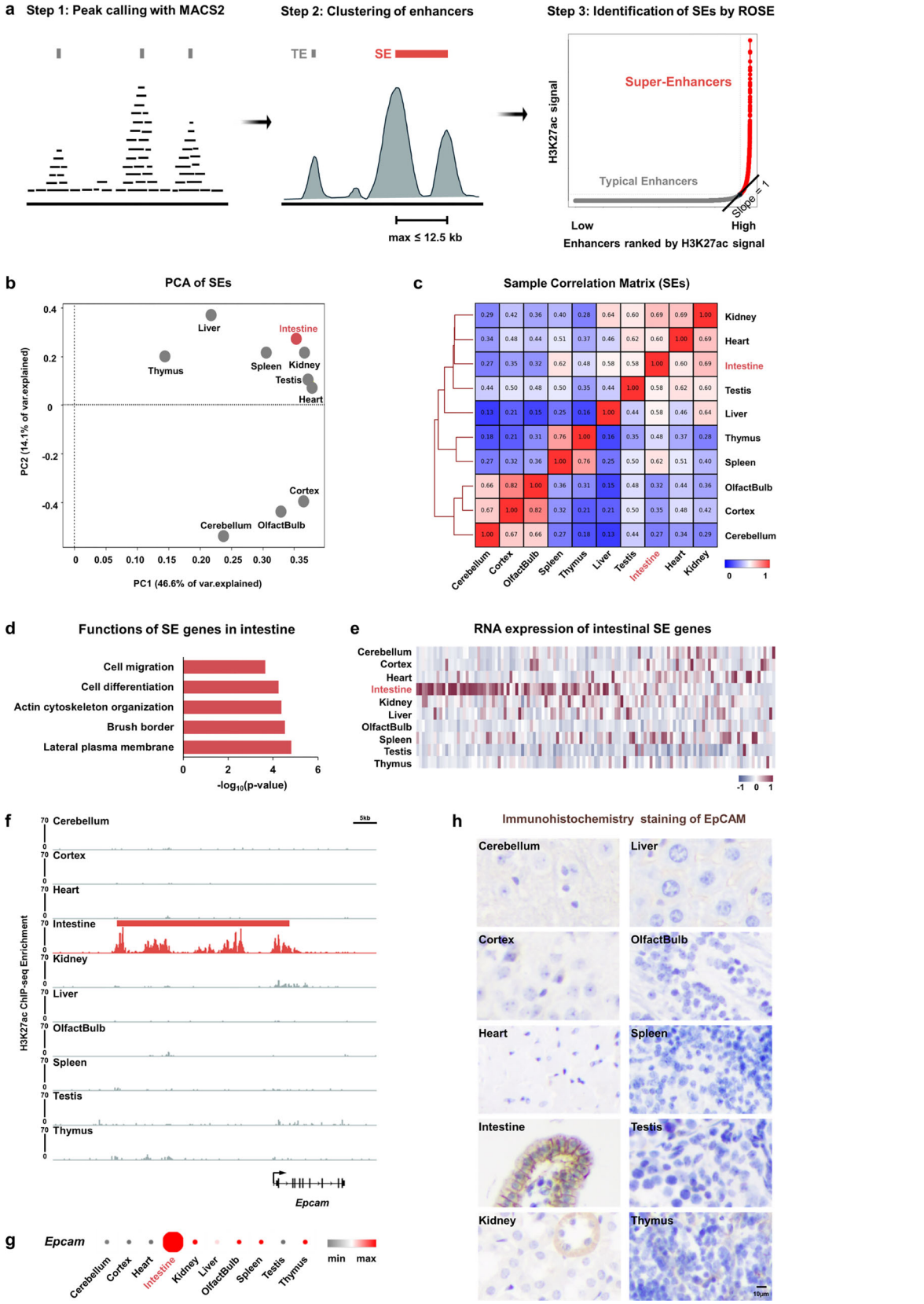


Figure 1. SE landscape defines tissue-specific epigenetic architecture of the intestine. (**a**) Workflow for SE identification using the ROSE algorithm. (**b**) PCA of SE landscapes across 10 tissues (H3K27ac ChIP-seq). (**c**) Hierarchical clustering and correlation matrix of SE profiles. (**d**) Functional annotation of small intestine-specific SE-associated genes by DAVID. (**e**) Heatmap of intestinal-specific SE-linked genes. (f–h) Representative examples of small intestinal SE-associated genes. (**f**) The epithelial marker *Epcam* exhibits SE activity in the small intestine, with IGV tracks showing the characteristic SE architecture of broad, clustered H3K27ac peaks. Consistent with SE activity, *Epcam* shows intestine-enriched expression, as evidenced by RNA-seq (**g**) and IHC staining ($n = 3$ biologically independent mice) (**h**). SE loci and SE-associated genes with intestinal functional annotations are provided in Supplementary Table S1. H3K27ac ChIP-seq (GSE49847): MACS2 $P < .05$.

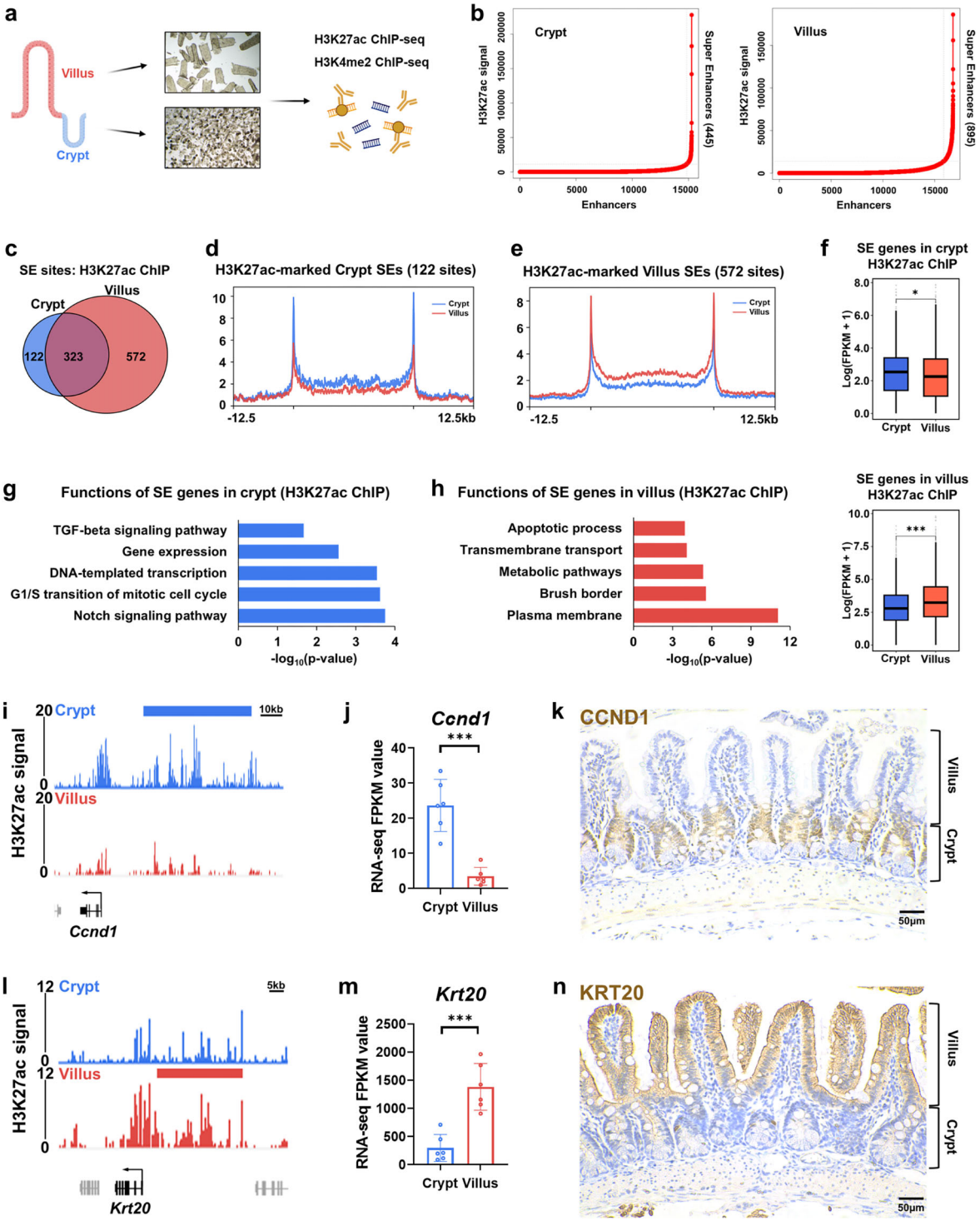


Figure 2. H3K27ac-marked compartment-specific SEs orchestrate crypt-villus axis functions. (**a**) Schematic of murine intestinal crypt-villus axis isolation. Crypts (stem/progenitor-enriched proliferative zone) and villi (terminally differentiated absorptive compartment) were separated for epigenomic profiling. (**b**) H3K27ac-marked SEs identified by the ROSE algorithm: 445 crypt-enriched versus 895 villus-enriched SEs. (**c**) Venn diagram of H3K27ac-marked SE sites of crypts and villi. (**d**) Crypt-specific SEs show significantly higher H3K27ac enrichment in crypts (blue) compared with villi (red). (**e**) Villus-specific SEs display preferential H3K27ac enrichment in villus (red) relative to crypt (blue), confirming compartment-specific activity. (**f**) RNA-seq analysis of SE-linked genes shows compartment-specific expression: crypt-associated SE genes (top) and villus-associated SE genes (bottom). The middle line represents the median (Mann–Whitney U test, $*P < .05$, $***P < .001$, $n = 6$ biological replicates per group). Functional enrichment of H3K27ac-marked crypt-specific (**g**) and villus-specific SE genes (**h**). Example of crypt-specific SE gene: (**i**) H3K27ac ChIP-seq profile showing crypt-specific SE enrichment at the *Ccnd1* locus; (**j**) RNA-seq confirming significantly higher *Ccnd1* expression in crypts than villi; (**k**) IHC demonstrating CCND1 protein localized to proliferative crypt cells ($n = 3$ biologically independent mice). Example of villus-specific SE gene: (**l**) H3K27ac ChIP-seq profile showing villus-specific SE enrichment at the *Krt20* locus; (**m**) RNA-seq confirming villus-restricted *Krt20* expression; (**n**) IHC showing KRT20 protein localized to villi ($n = 3$ biologically independent mice). All RNA-seq (crypt versus villus; GSE133949) data are presented as mean ± SEM ($n = 6$ biological replicates), DESeq2 $P < .001$ (***). H3K27ac ChIP-seq (crypt versus villus; GSE310534 and GSE148691): MACS2, $q < 0.05$. SE loci and SE-associated genes with functional annotations of crypts and villi are provided in Supplementary Table S2.

For example, colorectal cancer associated 2 (*Colca2*), a gene has been implicated in an increased risk of colorectal cancer (CRC) [40, 41], showed crypt-restricted H3K4me2 enrichment (Supplementary Fig. S2o and p). Conversely, carcinoembryonic antigen related cell adhesion molecule 20 (*Ceacam20*), a gene is associated with the progression of colitis [42], showed villus-restricted SE activity and expression (Supplementary Fig. S2q and r).

Together, these findings highlight SEs as key epigenetic regulators of intestinal compartmentalization, orchestrating crypt–villus patterning, stem cell maintenance, and differentiation. Several SE-associated genes are linked to colorectal tumorigenesis, underscoring SEs' dual role in tissue homeostasis and disease.

CDX2 coordinates super-enhancer architecture in the intestine

Building on the paradigm that SEs govern cell identity [4, 6], we hypothesized that intestinal master TFs hierarchically regulate epithelial differentiation via SE engagement. Integrative epigenomic profiling identified that CDX2 functions as the dominant SE regulator. In WT murine intestines, CDX2 ChIP-seq signals showed striking overlap with H3K27ac at SEs (Fig. 3a–c and Supplementary Table S3). ROSE identified 2497 H3K27ac-marked SEs and 1850 CDX2-bound SEs, with 1367 shared SE sites (Supplementary Fig. S3a, the left panel). Gene Ontology analysis showed strong functional concordance: 80% of top CDX2-SE and H3K27ac-SE genes converged on intestinal-specific pathways, including differentiation, metabolism, brush border assembly, and apical membrane organization (Fig. 3b and Supplementary Fig. S3a, the right panel). For example, SEs display nearly identical peak profiles of CDX2 ChIP-seq and H3K27ac ChIP-seq at gene loci of Claudin-7 (*Cldn7,* a core tight junction component) and *Epcam* (Fig. 3c).

CDX2 is highly expressed in the small intestine (Fig. 3d). To test whether CDX2 is required for SEs, we analyzed the intestinal epithelium-specific knockout mice, $Cdx2^{KO}$. H3K27ac ChIP-seq and ROSE analysis revealed a dramatic loss of SEs: from 2497 in WT to only 263 in $Cdx2^{KO}$—a ∼90% reduction (Fig. 3e and f, the left panel). In contrast, the numbers of TEs remained unchanged (WT: 25 993; KO: 25 854; Fig. 3f, the right panel). Loss of CDX2 resulted in the collapse of the majority of SEs (Fig. 3g and h; Supplementary Fig. S3b). Combined epigenomic and transcriptomic analysis confirmed that CDX2 loss leads to global SE depletion and transcriptional downregulation of SE-associated genes (Fig. 3i and Supplementary Fig. S3c). Notably, most downregulated genes were related to cell differentiation (Supplementary Fig. S3c). Consistently, villus-specific SEs exhibited reduced signals upon CDX2 loss (Supplementary Fig. S3d).

For example, in WT intestines, *Krt20*—a villus differentiation marker—exhibited classic SE features with broad H3K27ac enrichment, and CDX2 bound to this gene locus (Fig. 3j). The SE architecture at the *Krt20* locus was lost in *Cdx2* knockout ($Cdx2^{KO}$) epithelium, alongside reduced H3K27ac signals (Fig. 3j). This was accompanied by corresponding decreases in mRNA (Fig. 3k) and protein expression (Fig. 3l), supporting the model that CDX2 activates differentiation genes by establishing SEs. *Hnf4g* (another CDX2-bound gene), which encodes a nuclear receptor involved in epithelial differentiation, also showed reduced SE architecture and H3K27ac signals upon CDX2 loss (Fig. 3m) and showed downregulation of mRNA and protein levels (Fig. 3n and o).

RNA Pol II not only transcribes coding genes but also engages active enhancers, including SEs, to produce enhancer RNAs and support transcriptional activity [6, 43, 44]. To further study CDX2's regulatory role, we analyzed RNA Pol II ChIP-seq in WT and $Cdx2^{KO}$ intestines (Supplementary Fig. S3e–l). ROSE analysis identified 558 RNA Pol II-enriched SEs in WT, whereas only 38 were found in $Cdx2^{KO}$—a reduction of ∼93% (Supplementary Figs S3e and f, the left panel). In contrast, TE numbers remained stable (WT: 11 414; KO: 10 959; Supplementary Fig. S3f, the right panel). In $Cdx2^{KO}$, Pol II signal density was reduced at SEs (Supplementary Fig. S3g), and these SE-associated genes were also downregulated (Supplementary Fig. S3h). For example, CDX2 binds to dipeptidyl peptidase-4 (*Dpp4*) and cubilin (*Cubn*) in the intestine. SE loss and concomitant downregulation of mRNA expression were observed in the $Cdx2^{KO}$ (Supplementary Fig. S3i–l).

Altogether, these findings suggest that CDX2 is required for SE maintenance, driving high-level expression of epithelial identity genes and maintaining intestinal homeostasis.

HNF4 depletion reshapes the super-enhancer landscape and activates oncogenic programs

We further explored how other key intestinal TFs regulate the epithelium through SEs. HNF4 ChIP-seq and ROSE analysis identified 300 SEs anchored by the nuclear receptor HNF4 (HNF4α/γ) (Supplementary Fig. S4a and Supplementary Table S4). These HNF4-bound SEs were linked to both epithelial homeostasis (e.g. brush border, mineral absorption) and pathways in cancer (Fig. 4a).

Based on these findings and HNF4's known tumor-suppressive role [45–48], we proposed a model in which HNF4 occupies SE regions near oncogenes to maintain transcriptional silencing. To test this, we employed *Hnf4α/γ* double-knockout ($Hnf4\alpha\gamma^{DKO}$) mice. H3K27ac ChIP-seq revealed similar total SE numbers in WT and $Hnf4\alpha\gamma^{DKO}$ (1465 versus 1596; Fig. 4b and c; Supplementary Table S4), suggesting that HNF4 loss does not globally reduce SEs. Venn analysis revealed 615 SE sites unique to WT, 746 unique to $Hnf4\alpha\gamma^{DKO}$, and 850 shared between the two conditions (Fig. 4d). In $Hnf4\alpha\gamma^{DKO}$, H3K27ac signals were downregulated at the 615 SE sites unique to WT and upregulated at the 746 SE sites unique to $Hnf4\alpha\gamma^{DKO}$ (Fig. 4e). Genes associated with WT-specific SE genes were related to intestinal functions, while genes associated with $Hnf4\alpha\gamma^{DKO}$-specific genes were enriched in cancer pathways (Fig. 4f). Knockout of HNF4 led to downregulation of most HNF4-bound, H3K27ac-marked SE genes involved in intestinal differentiation (Supplementary Fig. S4b), suggesting that HNF4 also functions as a transcriptional activator in the context of SEs. For instance, the solute carrier family 19 member 1 (*Slc19a1*) locus lost its H3K27ac-marked SE upon HNF4 loss (Fig. 4g), resulting in transcript downregulation (Fig. 4h). In contrast, the locus encoding MDS1 and EVI1 complex locus (*Mecom*), an oncogenic TF implicated in colorectal cancer [49–52], acquired a SE (Fig. 4i) and was robustly upregulated in $Hnf4\alpha\gamma^{DKO}$ intestines (Fig. 4j). We found similar findings when we analyzed RNA Pol II ChIP-seq and identified SEs in WT and $Hnf4\alpha\gamma^{DKO}$ (Supplementary Fig. S4c–f). $Hnf4\alpha\gamma^{DKO}$-specific SEs were again enriched for oncogenic

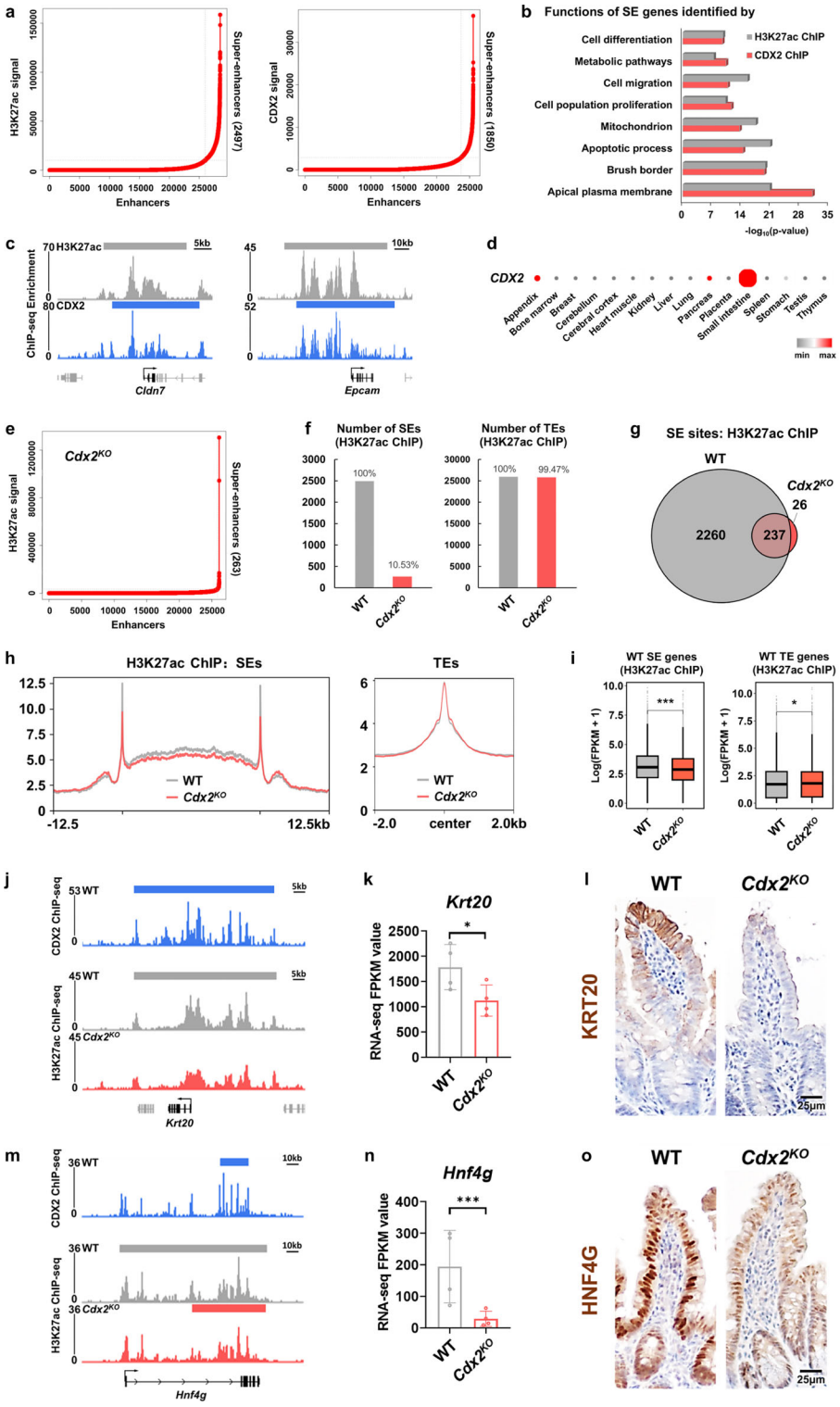


Figure 3. CDX2 governs the intestinal SE landscape to regulate epithelial transcription. (**a**) ROSE algorithm identifies 2497 H3K27ac-marked SEs (left) and 1850 CDX2-bound SEs (right) in the intestinal villus cells. (**b**) SE-associated genes identified by H3K27ac and CDX2 ChIP-seq show highly consistent functional enrichment through DAVID. (**c**) Representative examples of SEs co-occupied by H3K27ac and CDX2 ChIP-seq signals. (**d**) Tissue-specific expression analysis confirms *CDX2* enrichment in small intestine (The Human Protein Atlas). (**e**) H3K27ac-marked SEs in *Cdx2*^{KO}. (**f**) Number of H3K27ac-marked SEs (WT: 2 497; *Cdx2*^{KO}: 263) and TEs (WT: 25 993; *Cdx2*^{KO}: 25 854). (**g**) Venn diagram of H3K27ac-marked SE sites of WT and *Cdx2*^{KO}. (**h**) H3K27ac ChIP-seq shows selective attenuation at SE regions in *Cdx2*^{KO} (left), while TE regions remain unaffected (right). (**i**) Transcriptome analysis of SE- and TE-associated genes in WT and *Cdx2*^{KO}. The middle line represents the median (Mann–Whitney U test, $*P < .05$, $***P < .001$, $n = 4$ biological replicates per group). CDX2-dependent SE regulation: complete SE loss at the *Krt20* locus (**j**) with corresponding transcriptional (**k**) and protein (**l**) downregulation; partial retention of H3K27ac-marked SE at the *Hnf4g* locus (**m**) is insufficient to sustain messenger RNA (mRNA) (**n**) or protein (**o**) levels. CDX2 and H3K27ac ChIP-seq (*Cdx2*^{KO} versus *WT*; MACS2, $P < .01$; GSE98724): $n = 2$ biological replicates; RNA-seq (*Cdx2*^{KO} versus *WT*; GSE98724 and GSE115541): $n = 4$ biological replicates, DESeq2 $P < .05$ (*), $P < .001$ (***). IHC staining: $n = 2$ biologically independent mice. SE loci and SE-associated genes with functional annotations of WT and *Cdx2*^{KO} are provided in Supplementary Table S3.

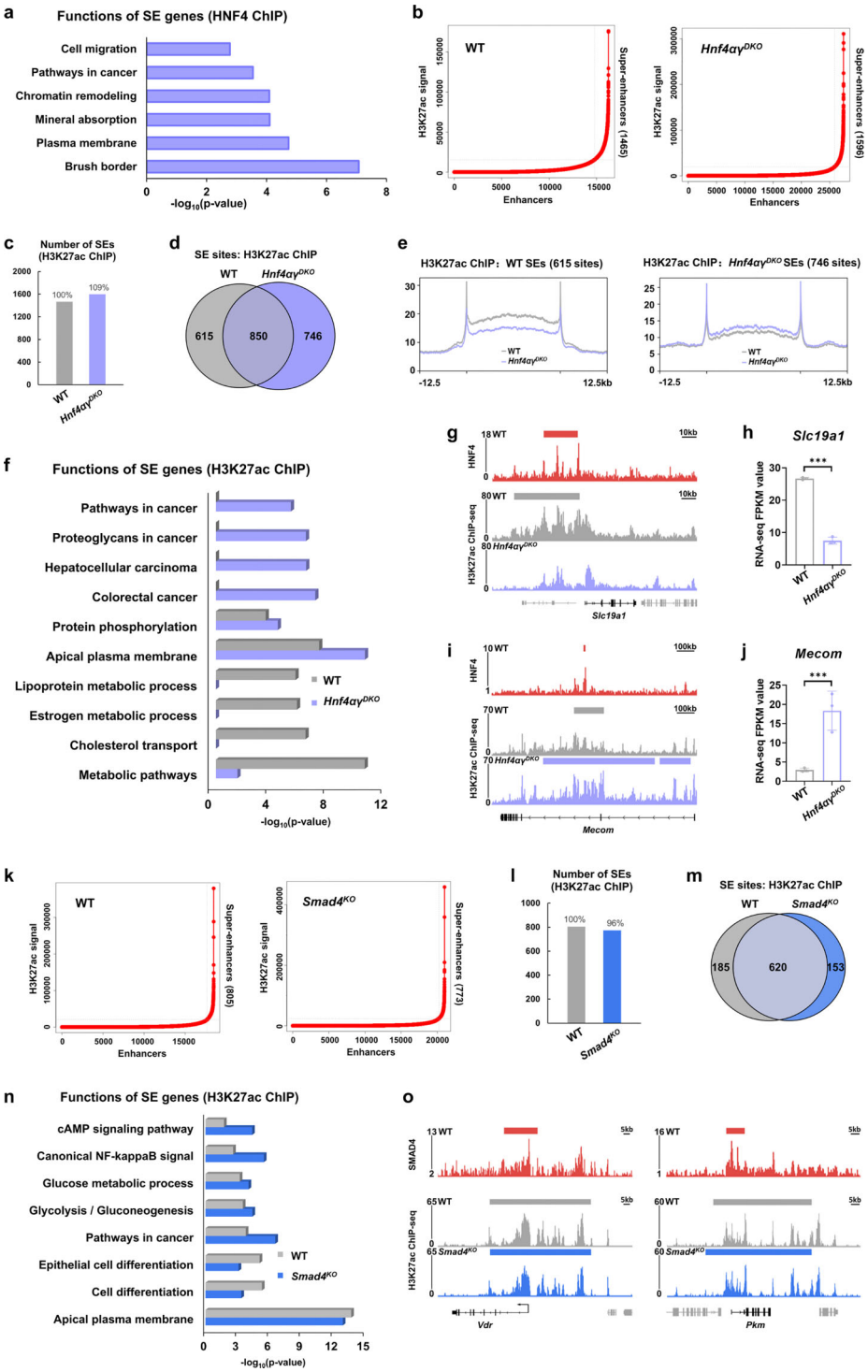


Figure 4. HNF4 loss, but not SMAD4 loss, remodels the SE landscape and activates oncogenic programs. (**a**) Functional enrichment of HNF4-bound SE genes in intestinal epithelial cells reveals dual roles in homeostasis and oncogenesis. (**b**) ROSE-defined H3K27ac SEs in WT and *Hnf4α*γ^{DKO} intestines. (**c**) Numbers of SE identified by H3K27ac ChIP-seq (WT: 1 465, *Hnf4α*γ^{DKO}: 1 596). (**d**) Venn diagram of H3K27ac-marked SE sites of WT and *Hnf4α*γ^{DKO}. (**e**) H3K27ac signals at SEs unique of intestinal epithelial cells in WT and *Hnf4α*γ^{DKO}. (**f**) Functional annotation of genotype-specific SE genes in WT and *Hnf4α*γ^{DKO}, showing enrichment of cancer-related pathways after HNF4 loss. (**g**) IGV tracks of *Slc19a1* locus showing H3K27ac-defined SE in WT. (**h**) Downregulation of *Slc19a1* in *Hnf4α*γ^{DKO}, DEseq2 $P < .001$ (***). (**i**) IGV shows enhanced SE signal at the *Mecom* locus in DKO mice. (**j**) *Mecom* was significantly upregulated in *Hnf4α*γ^{DKO}, DESeq2 $P < .001$ (***). (**k**) ROSE-defined H3K27ac SEs in WT and *Smad4*^{KO} intestines. (**l**) Numbers of H3K27ac-marked SEs in WT and *Smad4*^{KO} (WT: 805, *Smad4*^{KO}: 773). (**m**) Venn diagram of H3K27ac-marked SE sites of intestinal epithelial cells in WT and *Smad4*^{KO}. (**n**) Functional analysis of SE genes in WT and *Smad4*^{KO}. (**o**) Examples of conserved SE architecture and H3K27ac signal at the gene loci of intestinal epithelium in WT and *Smad4*^{KO}. HNF4 ChIP-seq: GSE310534, GSE34566, and GSE112946 (MACS2 $P < .05$); SMAD4 ChIP-seq (MACS2 $P < .05$; GSE310534 and GSE112946): $n = 2$ biological replicates; H3K27ac ChIP-seq (*Hnf4α*γ^{DKO} versus WT; MACS2, $P < .01$; GSE310534 and GSE112946): $n = 2$ biological replicates; H3K27ac ChIP-seq (*Smad4*^{KO} versus WT; MACS2, $q < 0.01$; GSE310534); RNA-seq (*Hnf4α*γ^{DKO} versus WT; GSE309833 and GSE112946): $n = 3$ biological replicates. Complete lists of SE loci and SE-associated genes with functional annotations are provided in Supplementary Table S4 (*Hnf4α*γ^{DKO} versus WT) and Supplementary Table S5 (*Smad4*^{KO} versus WT).

pathways (Supplementary Fig. S4f). These data suggest that HNF4 loss permits oncogene SE activation.

Our previous work demonstrates a reinforcing HNF4–SMAD4 feed-forward module stabilizes enterocyte identity [20]. We then further mapped the landscape of SMAD4-bound SEs and found 353 SEs (Supplementary Fig. S4g and Supplementary Table S5). These SMAD4-bound SEs were linked to both homeostasis (e.g. cell differentiation) and stress responses (e.g. glucose starvation, TNF signaling; Supplementary Fig. S4h). However, SMAD4 knockout did not markedly affect SE number (Fig. 4k and l, WT: 805 versus $Smad4^{KO}$: 773), genomic distribution (Fig. 4m), the functional output of SE-associated genes (Fig. 4n and o), and H3K27ac signal intensity (Supplementary Fig. S4i), suggesting that SMAD4 exerts a weaker influence on SE regulation compared with CDX2 or HNF4. In summary, we propose that HNF4 and SMAD4 regulate intestinal SEs through distinct mechanisms: HNF4 maintains SE-mediated repression of oncogenes, and its loss promotes oncogenic SE reprogramming. By contrast, SMAD4 largely preserves global SE architecture, though its knockout is accompanied by a modest downregulation of cell differentiation and upregulation of cancer-associated pathways (Fig. 4n and o).

Dynamic super-enhancer reprogramming reveals oncogenic mechanisms in colorectal carcinogenesis

Building on the connection between cancer and TF (Fig. 4), we examined their regulatory functions using the AOM/DSS model. The AOM/DSS model is a chemically induced murine model of colitis-associated colorectal cancer [53–55]. In this dataset, a single intraperitoneal injection of AOM was followed by a single 7-day cycle of 2.5% DSS in drinking water to trigger tumor progression. H3K27ac ChIP-seq and RNA-seq profiling [56, 57] were then performed on the distal colon of control mice and experimental mice at 4, 7, and 10 weeks post–AOM–DSS induction (Fig. 5a). We found a dynamic remodeling of SEs: SE numbers dropped early (Control: 1025; 4-weeks: 723), recovered at 7-weeks (1023), and expanded by 10-weeks (1788) (Fig. 5b, Supplementary Fig. S5a, and Supplementary Table S6). PCA and correlation matrices of these SEs showed a clear difference during tumor progression (Fig. 5c and Supplementary Fig. S5b), reflecting progressive SE landscape reprogramming. To investigate the dynamic changes of SEs during cancer progression, we compared SE landscapes between control and samples collected at 4, 7, and 10 weeks in the AOM/DSS model. This analysis identified 1101 tumor-specific SEs that were present exclusively in the models but absent in the controls. Notably, the activity of these tumor-specific SEs progressively increased throughout tumor development, reaching its peak at the late stage (Fig. 5d).

Integrated analysis of these SE landscapes with RNA-seq data identified four distinct gene clusters with stage-specific functions during tumorigenesis (Fig. 5e and f). Cluster 1, specific to control colon tissues, was enriched in genes regulating epithelial structure, and these genes were suppressed in tumors. Cluster 2, predominant at the 4-week stage, was enriched in genes involved in apoptosis, chromatin remodeling, and cell migration—key processes initiating early transformation. Cluster 3, present across all stages of tumorigenesis, comprised genes associated with cancer, metabolism and mitochondrial function. Cluster 4 emerged at late stages, featuring genes related to cytoskeletal organization and cell adhesion. GSEA confirmed sustained activation of SE-associated genes across tumor stages (Supplementary Fig. S5c).

We next assessed how TF loss affects this trajectory. At SE loci unique to either $Hnf4\alpha\gamma^{DKO}$ or $Smad4^{KO}$ compared with their respective controls, the 10-week AOM–DSS model exhibited a marked increase in H3K27ac signal intensity relative to controls (Fig. 5g). Coincidentally, decreased protein levels of HNF4A and SMAD4 were observed in the tumors of $Apc^{Min/+}$ mice, a spontaneous colorectal cancer model (Fig. 5h), suggesting TF silencing during tumor progression. This observation was also recapitulated in human colorectal cancer, where scRNA-seq analysis showed downregulation of HNF4A and HNF4G in tumor epithelial cells compared to matched normal tissues (Fig. 5i). *Mecom*, the cancer-associated gene, exhibited enhanced SE signal upon HNF4 knockout (Fig. 4i). We also observed SE formation at the gene locus of *Mecom* in $Smad4^{KO}$ (Fig. 5j) and in late-stage (7 and 10 weeks) tumors (Fig. 5k), with transcriptional upregulation at these timepoints (Fig. 5l). Conversely, C-X-C chemokine receptor 4 (*Cxcr4*, promoting EMT via PI3K/AKT) [58–60] exhibited SE activation at an earlier stage (4 weeks) of AOM–DSS model and showed a corresponding transcriptional induction (Supplementary Fig. S5d and e), implicating *Cxcr4* in earlier stages of tumor initiation.

Altogether, this study delineates a biphasic model of SE-driven tumorigenesis in colorectal cancer. Early SEs, such as at the *Cxcr4* locus, may act as initiators of malignant transformation, while late-stage SEs, such as *Mecom*, may sustain tumor progression. HNF4 and SMAD4 function as critical transcriptional gatekeepers, constraining oncogenic transitions. These findings support strategies aimed at targeting early SE hubs to prevent tumor initiation, as well as restoring TF activity (homeostasis) or inhibiting late-stage SEs to suppress disease progression. Moreover, integrating dynamic SE profiling with TF loss signatures could facilitate precision staging and personalized treatment in colorectal cancer.

Discussion

This study systematically defines the central regulatory role of SEs in organizing intestinal tissue architecture, establishing a foundational framework for understanding epithelial patterning, identity, and disease. By mapping spatiotemporally restricted SE landscapes across the crypt-villus axis, we uncover precise regional enhancer activation that mirrors functional epithelial compartmentalization. Compartment-specific SEs govern gene expression programs essential for stemness/proliferation (crypts) and differentiation (villi), positioning SEs as key epigenetic hubs orchestrating intestinal identity and spatial organization.

Our findings underscore a hierarchical SE–TF framework in the intestinal epithelium. CDX2 acts as a master regulator, maintaining SE integrity and epithelial identity; its loss disrupts SE networks and impairs differentiation. HNF4 is a key intestinal TF that governs epithelial differentiation [20, 23], and its loss in gastrointestinal cancers has been associated with dedifferentiation and oncogenic reprogramming [45–48]. In contrast, SMAD4 serves as a central tumor suppressor within the TGF-β signaling pathway, where its frequent inactivation in colorectal cancers promotes invasion

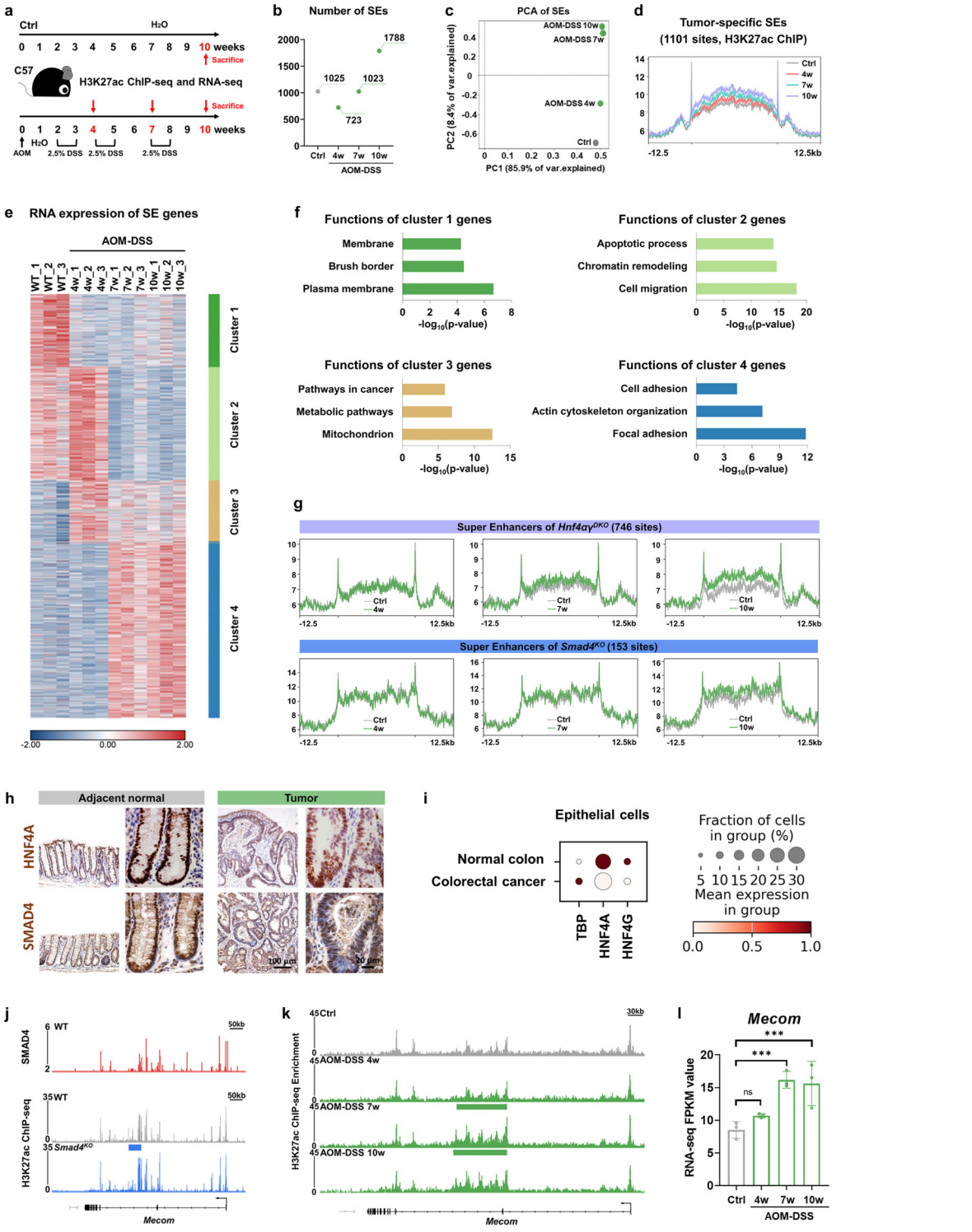


Figure 5. Dynamic remodeling of SE landscapes in colorectal tumorigenesis. (**a**) Schematic of experimental design using the distal colon fragments of a colorectal tumorigenesis model (Control, 4, 7, and 10 weeks of the AOM–DSS model). (**b**) Temporal dynamics of SE numbers: Control (1025), 4-week (723), 7-week (1023), 10-week (1788). (**c**) PCA of SE landscapes revealing stage-specific epigenetic clustering of colorectal tumorigenesis. (**d**) SitePro analysis reveals a time-dependent gain in signal intensity of H3K27ac from 4 to 10 weeks for 1101 tumor-specific SEs, which are absent in controls but emerge during cancer progression. (**e**) Heatmap of SE-associated gene expression across different time points of distal colon in the AOM–DSS model. (**f**) Functional annotations of identified SE-associated gene clusters. (**g**) H3K27ac signals of epithelial cells (AOM–DSS versus control) at unique SE regions of *Hnf4αγ*^{DKO} or *Smad4*^{KO}. (**h**) Reduced protein levels of HNF4A and SMAD4 in *Apc*$^{min/+}$ intestinal tumors, as evidenced by IHC ($n = 3$ biologically independent mice). (**i**) Dot plot of scRNA-seq (GSE200997) shows expression levels of HNF4 in epithelial cells from CRC patients. (**j**) Representative IGV tracks of SMAD4 ChIP-seq and H3K27ac ChIP-seq signals at the *Mecom* locus. (**k**) H3K27ac ChIP-seq profiles reveal SEs at the *Mecom* locus after 7 and 10 weeks of AOM–DSS treatment. (**l**) *Mecom* is upregulated during the progression of AOM–DSS tumors. All RNA-seq data are presented as mean ± SEM ($n = 3$ biological replicates), DEseq2 $P < .001$ (***), ns: not significant. H3K27ac ChIP-seq (MACS2 $P < .01$) and RNA-seq (AOM–DSS model, GSE178214): $n = 3$ biological replicates. All SE loci and SE-associated genes with functional annotations of the AOM–DSS model are provided in Supplementary Table S6.

and metastasis [61–63]. We demonstrate stage-specific SE reprogramming in colorectal cancer, particularly upon these TFs loss. Collectively, these findings highlight the critical role of SEs in coordinating epithelial identity, homeostasis, and malignancy.

While this study establishes a foundational framework for understanding SE regulation in intestinal identity and tumor progression, several key questions remain. Notably, the acquisition of cancer-associated SEs was observed only upon HNF4 loss. In contrast, CDX2 loss led to a sharp reduction in SE numbers, which may underlie its association with poor prognosis in colon cancer [64, 65]. SMAD4 loss had limited effect on SE maintenance but is also linked to worse clinical outcomes [66], suggesting that SMAD4 may influence tumor progression through mechanisms beyond SE remodeling. Future studies are needed to identify additional SE regulatory factors that may cooperate with or function independently of CDX2, HNF4, and SMAD4 in orchestrating intestinal homeostasis and oncogenesis. Moreover, the role of SEs in governing stem cell fate decisions, lineage plasticity, and epithelial regeneration after injury remains largely unexplored.

A critical next step will be to elucidate how SE landscapes respond to dynamic microenvironmental cues. Building on our findings that SE architecture is both spatially constrained and progressively deregulated upon TF loss, we propose that SEs exhibit context-dependent plasticity in response to niche-derived signals. Dissecting the mechanisms of microenvironment-responsive SE remodeling—particularly under conditions such as hypoxia, inflammation, and exposure to microbial metabolites—will be essential for understanding how extrinsic stimuli shape epigenomic identity and drive malignant transformation.

Acknowledgements

This research work supported by the Big Data Computing Center of Southeast University and the Center for Fundamental and Interdisciplinary Sciences of Southeast University.

Author contributions: W.-X.F. conducted experiments, data analysis, and manuscript writing. S.-Y.H., R.-X.X., G.-L.W., and M.-Y.Z. assisted with the sequencing data processing. X.Q. and L.Z. contributed to bench work. M.P.V. supported this study and revised the manuscript. Y.W. and L.C. conceived, designed, and supervised the study; contributed to bench work, manuscript writing, and revision.

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

The authors disclose no conflicts.

Funding

L.C. is supported by grants from the National Natural Science Foundation of China (32270830, U25A2016), the National Key R&D Program of China (2023YFA1801500), the Nature Science Foundation of Jiangsu Province (BK20230026), the Southeast University Interdisciplinary Research Program for Young Scholars (2024FGC1003), the Fundamental Research Funds for the Central Universities (2242025F10003), and the University's Start-up Research Fund (RF1028623015). M.P.V. is supported by grant R01DK126446 from the National Institutes of Health (NIH). Funding to pay the Open Access publication charges for this article was provided by the National Natural Science Foundation of China.

Data availability

The sequencing data we generated and used in this study have been deposited at GEO under the following accessions: GSE310534 and GSE148691 (H3K27ac and H3K4me2 ChIP-seq of small intestinal villi and crypts), GSE310534 and GSE112946 (HNF4 and SMAD4 ChIP-seq in WT; H3K27ac ChIP-seq in WT and $Hnf4\alpha\gamma^{DKO}$), GSE309833 and GSE112946 (RNA-seq in WT and $Hnf4\alpha\gamma^{DKO}$), and GSE310534 (H3K27ac ChIP-seq in WT and $Smad4^{KO}$). This paper also analyzes existing, publicly available data. These accession numbers for the datasets are including: GSE133949 (RNA-seq of small intestinal villi and crypts), GSE49847 [35] (H3K27ac ChIP-seq and RNA-seq data for 10 mouse organs, including the small intestine), GSE98724 [67] (CDX2 ChIP-seq in WT, H3K27ac and RNA Pol II ChIP-seq in WT and $Cdx2^{KO}$), GSE98724 [67] and GSE115541 [68, 69] (RNA-seq in WT and $Cdx2^{KO}$), GSE34566 [70, 71] (HNF4 ChIP-seq in WT), GSE244918 [72] (RNA Pol II ChIP-seq in WT and $Hnf4\alpha\gamma^{DKO}$), GSE178214 [56, 57] (H3K27ac ChIP-seq and RNA-seq in Ctrl and AOM–DSS-induced colorectal cancer models), and GSE200997 [73] (scRNA-seq on tumor and adjacent normal tissues of CRC patients).

References

1. Allen BL, Taatjes DJ. The Mediator complex: a central integrator of transcription. *Nat Rev Mol Cell Biol* 2015;16:155–66. https://doi.org/10.1038/nrm3951
2. Buecker C, Wysocka J. Enhancers as information integration hubs in development: lessons from genomics. *Trends Genet* 2012;28:276–84. https://doi.org/10.1016/j.tig.2012.02.008
3. Tang F, Yang Z, Tan Y *et al.* Super-enhancer function and its application in cancer targeted therapy. *NPJ Precis Oncol* 2020;4:2. https://doi.org/10.1038/s41698-020-0108-z
4. Whyte WA, Orlando DA, Hnisz D *et al.* Master transcription factors and Mediator establish super-enhancers at key cell identity genes. *Cell* 2013;153:307–19. https://doi.org/10.1016/j.cell.2013.03.035
5. Li XP, Qu J, Teng XQ *et al.* The emerging role of super-enhancers as therapeutic targets in the digestive system tumors. *Int J Biol Sci* 2023;19:1036–48. https://doi.org/10.7150/ijbs.78535
6. Hnisz D, Abraham BJ, Lee TI *et al.* Super-enhancers in the control of cell identity and disease. *Cell* 2013;155:934–47. https://doi.org/10.1016/j.cell.2013.09.053
7. Lovén J, Hoke HA, Lin CY *et al.* Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* 2013;153:320–34.
8. Liu Q, Guo L, Lou Z *et al.* Super-enhancers and novel therapeutic targets in colorectal cancer. *Cell Death Dis* 2022;13:228. https://doi.org/10.1038/s41419-022-04673-4
9. Jin F, Li Y, Dixon JR *et al.* A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* 2013;503:290–4. https://doi.org/10.1038/nature12644
10. Jiang Y-Y, Jiang Y, Li C-Q *et al.* TP63, SOX2, and KLF5 establish a core regulatory circuitry that controls epigenetic and transcription patterns in esophageal squamous cell carcinoma cell lines. *Gastroenterology* 2020;159:1311–27. https://doi.org/10.1053/j.gastro.2020.06.050

11. Davies JOJ, Oudelaar AM, Higgs DR *et al.*. How best to identify chromosomal interactions: a comparison of approaches. *Nat Methods* 2017;14:125–34. https://doi.org/10.1038/nmeth.4146
12. Huang J, Li K, Cai W *et al.* Dissecting super-enhancer hierarchy based on chromatin interactions. *Nat Commun* 2018;9:943. https://doi.org/10.1038/s41467-018-03279-9
13. Kong S, Zhang Y. Deciphering Hi-C: from 3D genome to function. *Cell Biol Toxicol* 2019;35:15–32. https://doi.org/10.1007/s10565-018-09456-2
14. Mifsud B, Tavares-Cadete F, Young AN *et al.* Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nat Genet* 2015;47:598–606. https://doi.org/10.1038/ng.3286
15. Boyle AP, Davis S, Shulha HP *et al.* High-resolution mapping and characterization of open chromatin across the genome. *Cell* 2008;132:311–22. https://doi.org/10.1016/j.cell.2007.12.014
16. Buenrostro JD, Wu B, Chang HY *et al.*. ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr Protoc Mol Biol* 2015;109:21.29.1–9. https://doi.org/10.1002/0471142727.mb2129s109
17. el Marjou F, Janssen KP, Chang BH *et al.* Tissue-specific and inducible Cre-mediated recombination in the gut epithelium. *Genesis* 2004;39:186–93.
18. Verzi MP, Shin H, He HH *et al.* Differentiation-specific histone modifications reveal dynamic chromatin interactions and partners for the intestinal transcription factor CDX2. *Dev Cell* 2010;19:713–26. https://doi.org/10.1016/j.devcel.2010.10.006
19. Hayhurst GP, Lee YH, Lambert G *et al.* Hepatocyte nuclear factor 4alpha (nuclear receptor 2A1) is essential for maintenance of hepatic gene expression and lipid homeostasis. *Mol Cell Biol* 2001;21:1393–403. https://doi.org/10.1128/MCB.21.4.1393-1403.2001
20. Chen L, Toke NH, Luo S *et al.* A reinforcing HNF4–SMAD4 feed-forward module stabilizes enterocyte identity. *Nat Genet* 2019;51:777–85. https://doi.org/10.1038/s41588-019-0384-0
21. Yang X, Li C, Herrera PL *et al.*. Generation of Smad4/Dpc4 conditional knockout mice. *Genesis* 2002;32:80–1. https://doi.org/10.1002/gene.10029
22. Khor TO, Cheung WK, Prawan A *et al.* Chemoprevention of familial adenomatous polyposis in Apc(Min/+) mice by phenethyl isothiocyanate (PEITC). *Mol Carcinog* 2008;47:321–5. https://doi.org/10.1002/mc.20390
23. Chen L, Cao W, Aita R *et al.* Three-dimensional interactions between enhancers and promoters during intestinal differentiation depend upon HNF4. *Cell Rep* 2021;34:108679. https://doi.org/10.1016/j.celrep.2020.108679
24. Bray NL, Pimentel H, Melsted P *et al.*. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* 2016;34:525–7. https://doi.org/10.1038/nbt.3519
25. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550. https://doi.org/10.1186/s13059-014-0550-8
26. Subramanian A, Tamayo P, Mootha VK *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci* 2005;102:15545–50. https://doi.org/10.1073/pnas.0506580102
27. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;9:357–9. https://doi.org/10.1038/nmeth.1923
28. Ramírez F, Ryan DP, Grüning B *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* 2016;44:W160–5.
29. Quinlan AR. BEDTools: the Swiss-Army tool for genome feature analysis. *Curr Protoc Bioinformatics* 2014;47:11.12.1–34. https://doi.org/10.1002/0471250953.bi1112s47
30. Pinello L, Farouni R, Yuan G-C. Haystack: systematic analysis of the variation of epigenetic states and cell-type specific regulatory elements. *Bioinformatics* 2018;34:1930–3. https://doi.org/10.1093/bioinformatics/bty031
31. Zhang Y, Liu T, Meyer CA *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 2008;9:R137. https://doi.org/10.1186/gb-2008-9-9-r137
32. Robinson JT, Thorvaldsdóttir H, Winckler W *et al.* Integrative genomics viewer. *Nat Biotechnol* 2011;29:24–6. https://doi.org/10.1038/nbt.1754
33. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* 2018;19:15. https://doi.org/10.1186/s13059-017-1382-0
34. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009;4:44–57. https://doi.org/10.1038/nprot.2008.211
35. Yue F, Cheng Y, Breschi A *et al.* A comparative encyclopedia of DNA elements in the mouse genome. *Nature* 2014;515:355–64. https://doi.org/10.1038/nature13992
36. Pott S, Lieb JD. What are super-enhancers? *Nat Genet* 2015;47:8–12. https://doi.org/10.1038/ng.3167
37. Chen L, Luo S, Dupre A *et al.* The nuclear receptor HNF4 drives a brush border gene program conserved across murine intestine, kidney, and embryonic yolk sac. *Nat Commun* 2021;12:2886. https://doi.org/10.1038/s41467-021-22761-5
38. Chen L, Toke NH, Luo S *et al.* HNF4 factors control chromatin accessibility and are redundantly required for maturation of the fetal intestine. *Development* 2019;146:dev179432. https://doi.org/10.1242/dev.179432
39. Battle MA, Konopka G, Parviz F *et al.* Hepatocyte nuclear factor 4alpha orchestrates expression of cell adhesion proteins during the epithelial transformation of the developing liver. *Proc Natl Acad Sci USA* 2006;103:8419–24. https://doi.org/10.1073/pnas.0600246103
40. Peltekova VD, Lemire M, Qazi AM *et al.* Identification of genes expressed by immune cells of the colon that are regulated by colorectal cancer-associated variants. *Int J Cancer* 2014;134:2330–41. https://doi.org/10.1002/ijc.28557
41. Koshyk O, Dehner CA, van den Hout MFCM *et al.* EWSR1::POU2AF3(COLCA2) sarcoma: an aggressive, polyphenotypic sarcoma with a head and neck predilection. *Mod Pathol* 2023;36:100337. https://doi.org/10.1016/j.modpat.2023.100337
42. Kelleher M, Singh R, O'Driscoll CM *et al.*. Carcinoembryonic antigen (CEACAM) family members and inflammatory bowel disease. *Cytokine Growth Factor Rev* 2019;47:21–31. https://doi.org/10.1016/j.cytogfr.2019.05.008
43. Kim T-K, Hemberg M, Gray JM *et al.* Widespread transcription at neuronal activity-regulated enhancers. *Nature* 2010;465:182–7. https://doi.org/10.1038/nature09033
44. Lam MTY, Cho H, Lesch HP *et al.* Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. *Nature* 2013;498:511–5. https://doi.org/10.1038/nature12209
45. Radi SH, Vemuri K, Martinez-Lomeli J *et al.*. HNF4α isoforms: the fraternal twin master regulators of liver function. *Front Endocrinol (Lausanne)* 2023;14:1226173. https://doi.org/10.3389/fendo.2023.1226173
46. Bonzo JA, Ferry CH, Matsubara T *et al.* Suppression of hepatocyte proliferation by hepatocyte nuclear factor 4α in adult mice. *J Biol Chem* 2012;287:7345–56. https://doi.org/10.1074/jbc.M111.334599
47. Lazarevich NL, Shavochkina DA, Fleishman DI *et al.* Deregulation of hepatocyte nuclear factor 4 (HNF4)as a marker of epithelial tumors progression. *Exp Oncol* 2010;32:167–71.
48. Qu N, Luan T, Liu N *et al.* Hepatocyte nuclear factor 4 a (HNF4α): a perspective in cancer. *Biomed Pharmacother* 2023;169:115923. https://doi.org/10.1016/j.biopha.2023.115923
49. Liang B, Wang J. EVI1 in leukemia and solid tumors. *Cancers (Basel)* 2020;12:2667. https://doi.org/10.3390/cancers12092667
50. Birdwell C, Fiskus W, Kadia TM *et al.* EVI1 dysregulation: impact on biology and therapy of myeloid malignancies. *Blood Cancer J* 2021;11:64. https://doi.org/10.1038/s41408-021-00457-9

51. Deng X, Cao Y, Liu Y *et al.* Overexpression of Evi-1 oncoprotein represses TGF-β signaling in colorectal cancer. *Mol Carcinog* 2013;52:255–64. https://doi.org/10.1002/mc.21852
52. Chen Y, Ying Y, Wang M *et al.* A distal super-enhancer activates oncogenic ETS2 via recruiting MECOM in inflammatory bowel disease and colorectal cancer. *Cell Death Dis* 2023;14:8. https://doi.org/10.1038/s41419-022-05513-1
53. De Robertis M, Massi E, Poeta ML *et al.* The AOM/DSS murine model for the study of colon carcinogenesis: from pathways to diagnosis and therapy studies. *J Carcinog* 2011;10:9.
54. Parang B, Barrett CW, Williams CS. AOM/DSS model of colitis-associated cancer. *Methods Mol Biol* 2016;1422:297–307.
55. Dzhalilova D, Zolotova N, Fokichev N *et al.*. Murine models of colorectal cancer: the azoxymethane (AOM)/dextran sulfate sodium (DSS) model of colitis-associated cancer. *PeerJ* 2023;11:e16159. https://doi.org/10.7717/peerj.16159
56. Chen L, Luo Z, Zhao C *et al.* Dynamic chromatin states coupling with key transcription factors in colitis-associated colorectal cancer. *Adv Sci (Weinh)* 2022;9:e2200536. https://doi.org/10.1002/advs.202200536
57. Lin X, Chen JD, Wang CY *et al.* Cooperation of MLL1 and Jun in controlling H3K4me3 on enhancers in colorectal cancer. *Genome Biol* 2023;24. 268. https://doi.org/10.1186/s13059-023-03108-3
58. Ma K-T, Wu Y-J, Yang Y-X *et al.* A novel phthalein component ameliorates neuroinflammation and cognitive dysfunction by suppressing the CXCL12/CXCR4 axis in rats with vascular dementia. *J Ethnopharmacol* 2024;328:118117. https://doi.org/10.1016/j.jep.2024.118117
59. Sarallah R, Jahani S, Soltani Khaboushan A *et al.* The role of CXCL12/CXCR4/CXCR7 axis in cognitive impairment associated with neurodegenerative diseases. *Brain Behav Immun Health* 2025;43:100932. https://doi.org/10.1016/j.bbih.2024.100932
60. Lou H, Xia Y, Shao S *et al.* CXCR4/CXCL12 axis promotes lymphatic metastasis in tongue squamous cell carcinoma via PI3K/AKT signaling pathway. *J Transl Med* 2025;23:757. https://doi.org/10.1186/s12967-025-06707-9
61. Miyaki M, Iijima T, Konishi M *et al.* Higher frequency of Smad4 gene mutation in human colorectal cancer with distant metastasis. *Oncogene* 1999;18:3098–103. https://doi.org/10.1038/sj.onc.1202642
62. Massagué J, Blain SW, Lo RS. TGFβ signaling in growth control, cancer, and heritable disorders. *Cell* 2000;103:295–309. https://doi.org/10.1016/S0092-8674(00)00121-5
63. Papageorgis P, Cheng K, Ozturk S *et al.* Smad4 inactivation promotes malignancy and drug resistance of colon cancer. *Cancer Res* 2011;71:998–1008. https://doi.org/10.1158/0008-5472.CAN-09-3269
64. Bae JM, Lee TH, Cho NY *et al.* Loss of CDX2 expression is associated with poor prognosis in colorectal cancer patients. *World J Gastroenterol* 2015;21:1457–67. https://doi.org/10.3748/wjg.v21.i5.1457
65. Dalerba P, Sahoo D, Paik S *et al.* CDX2 as a prognostic biomarker in stage II and stage III colon cancer. *N Engl J Med* 2016;374:211–22. https://doi.org/10.1056/NEJMoa1506597
66. Wasserman I, Lee LH, Ogino S *et al.* SMAD4 loss in colorectal cancer patients correlates with recurrence, loss of immune infiltrate, and chemoresistance. *Clin Cancer Res* 2019;25:1948–56. https://doi.org/10.1158/1078-0432.CCR-18-1726
67. Saxena M, Roman AKS, O'Neill NK *et al.* Transcription factor-dependent 'anti-repressive' mammalian enhancers exclude H3K27me3 from extended genomic domains. *Genes Dev* 2017;31:2391–404. https://doi.org/10.1101/gad.308536.117
68. Banerjee KK, Saxena M, Kumar N *et al.* Enhancer, transcriptional, and cell fate plasticity precedes intestinal determination during endoderm development. *Genes Dev* 2018;32:1430–42. https://doi.org/10.1101/gad.318832.118
69. Kumar N, Tsai YH, Chen L *et al.* The lineage-specific transcription factor CDX2 navigates dynamic chromatin to control distinct stages of intestine development. *Development* 2019;146:dev172189. https://doi.org/10.1242/dev.172189
70. Verzi MP, Shin H, San Roman AK *et al.* Intestinal master transcription factor CDX2 controls chromatin access for partner transcription factor binding. *Mol Cell Biol* 2013;33:281–92. https://doi.org/10.1128/MCB.01185-12
71. San Roman AK, Aronson BE, Krasinski SD *et al.* Transcription factors GATA4 and HNF4A control distinct aspects of intestinal homeostasis in conjunction with transcription factor CDX2. *J Biol Chem* 2015;290:1850–60.
72. Vemuri K, Kumar S, Chen L *et al.*. Dynamic RNA polymerase II occupancy drives differentiation of the intestine under the direction of HNF4. *Cell Rep* 2024;43:114242. https://doi.org/10.1016/j.celrep.2024.114242
73. Khaliq AM, Erdogan C, Kurt Z *et al.* Refining colorectal cancer classification and clinical stratification through a single-cell atlas. *Genome Biol* 2022;23:113. https://doi.org/10.1186/s13059-022-02677-z

Received: January 9, 2026. **Revised:** May 30, 2026. **Accepted:** June 4, 2026