



Zeb2 drives invasive and microbiota-dependent colon carcinoma

Karolina Slowicka^{1,2,3,21}, Ioanna Petta^{1,3,4,21}, Gillian Blancke^{1,3,4}, Esther Hoste^{1,2}, Emilie Dumas^{1,3,4}, Mozes Sze^{1,2}, Hanna Vikkula^{1,2}, Enrico Radaelli^{1,2,3,4}, Jody J. Haigh^{7,8}, Sven Jonckheere^{2,9}, Joachim Taminiau^{2,9}, Niels Vandamme^{1,2,9}, Andy Wullaert^{1,2,3,4}, Eugene Tulchinsky^{10,11}, David Nittner^{12,13}, Pieter Van Vlierberghe^{1,2,3,4}, Gert De Hertogh¹⁵, Pamela Baldin¹⁶, Emre Etlioglu¹³, Pratyaksha Wirapati¹⁷, Louis Boon¹⁸, Bart N. Lambrecht^{1,4}, Chris Callewaert¹⁹, Sabine Tejpar¹³, Steven Goossens^{2,9,14,20}, Geert Berx^{1,2,9,22} and Lars Vereecke^{1,3,4,22} and Geert van Loo^{1,2,3,9,22}

Colorectal cancer (CRC) is highly prevalent in Western society, and increasing evidence indicates strong contributions of environmental factors and the intestinal microbiota to CRC initiation, progression and even metastasis. We have identified a synergistic inflammatory tumor-promoting mechanism through which the resident intestinal microbiota boosts invasive CRC development in an epithelial-to-mesenchymal transition-prone tissue environment. Intestinal epithelial cell (IEC)-specific transgenic expression of the epithelial-to-mesenchymal transition regulator Zeb2 in mice (Zeb2^{IEC-Tg/+}) leads to increased intestinal permeability, myeloid cell-driven inflammation and spontaneous invasive CRC development. Zeb2^{IEC-Tg/+} mice develop a dysplastic colonic epithelium, which progresses to severely inflamed neoplastic lesions while the small intestinal epithelium remains normal. Zeb2^{IEC-Tg/+} mice are characterized by intestinal dysbiosis, and microbiota depletion with broad-spectrum antibiotics or germ-free rederivation completely prevents cancer development. Zeb2^{IEC-Tg/+} mice represent the first mouse model of spontaneous microbiota-dependent invasive CRC and will help us to better understand host-microbiome interactions driving CRC development in humans.

Over 1.2 million patients worldwide are diagnosed with colorectal cancer (CRC) each year, and approximately 50% will die from the disease, making CRC the third most common and fourth most deadly type of cancer^{1,2}. The majority of CRC cases are sporadic and develop slowly over several decades through a step-wise adenoma-to-carcinoma sequence, arising from adenomatous precursors that progressively accumulate genetic mutations and epigenetic alterations³. Hereditary CRC forms, such as familial adenomatous polyposis and hereditary non-polyposis colon cancer or Lynch syndrome, only account for about 5% of CRC cases. Environmental factors linked to a Western lifestyle contribute to the etiology of CRC, through mechanisms regulated by the intestinal microbiota⁴. The notion that CRC development is influenced by bacteria–host interactions is further supported by the identification of several CRC-promoting bacteria or ‘oncobacteria’, including *Fusobacterium nucleatum*^{5,6}. Chronic intestinal inflammation, fueled by the intestinal microbiota, involves both innate and adaptive immune responses with complex cytokine networks that orchestrate cancer development^{7,8}.

Epithelial-to-mesenchymal transition (EMT) is a cellular reprogramming process whereby epithelial cells lose their polarity and attachment to neighboring cells and acquire motile, migratory properties of mesenchymal cells. EMT regulates tissue morphogenesis during embryonic development, but also occurs in the adult organism in conditions of wound healing, fibrosis and cancer progression^{9,10}, where it correlates with acquisition of cancer stemness, suppression of cell death and resistance to therapy^{11–13}. EMT can be induced by a variety of growth factors as well as by components of the extracellular matrix and by hypoxia, eventually leading to the upregulation of one or several members of the SNAI, ZEB and TWIST families of EMT transcription factors¹⁴. ZEB2 is a member of the zinc finger E box-binding homeobox transcription factors, and was identified as a protein that interacts with SMAD proteins of the transforming growth factor- β (TGF- β) signaling pathway. High expression of ZEB2 is seen in many cancer cell lines and human cancer specimens, and correlates with increased cellular motility, invasion, cancer stem cell properties and tumor aggressiveness⁹.

¹VIB Center for Inflammation Research, Ghent, Belgium. ²Department of Biomedical Molecular Biology, Ghent University, Ghent, Belgium. ³Ghent Gut Inflammation Group, Ghent University, Ghent, Belgium. ⁴Department of Internal Medicine and Pediatrics, Ghent University, Ghent, Belgium.

⁵KU Leuven Center for Human Genetics, Leuven, Belgium. ⁶VIB Center for Brain and Disease Research, Leuven, Belgium. ⁷Research Institute in Oncology and Hematology, CancerCare Manitoba, Winnipeg, Manitoba, Canada. ⁸Department of Pharmacology and Therapeutics, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, Manitoba, Canada. ⁹Cancer Research Institute Ghent, Ghent, Belgium. ¹⁰University of Leicester, Leicester, UK.

¹¹Department of Biomedical Sciences, Nazarbayev University School of Medicine, Astana, Kazakhstan. ¹²Histopathology Expertise Center, VIB-KU Leuven Center for Cancer Biology, KU Leuven, Leuven, Belgium. ¹³Department of Biomolecular Medicine, Ghent University, Ghent, Belgium. ¹⁴Department of Pathology, University Hospital Leuven, KU Leuven, Leuven, Belgium. ¹⁵Department of Pathology, Cliniques Universitaires Saint-Luc, Université Catholique de Louvain, Brussels, Belgium. ¹⁶Department of Oncology, University Hospital Leuven, KU Leuven, Leuven, Belgium. ¹⁷Swiss Institute of Bioinformatics, Lausanne, Switzerland. ¹⁸Bioceros, Utrecht, the Netherlands. ¹⁹Center for Microbial Ecology and Technology, Ghent University, Ghent, Belgium.

²⁰Department of Diagnostic Sciences, Ghent University, Ghent, Belgium. ²¹These authors contributed equally: Karolina Slowicka, Ioanna Petta.

²²These authors jointly supervised this work: Geert Berx, Lars Vereecke, Geert van Loo. ✉e-mail: geert.berx@ugent.be; lars.vereecke@irc.vib-ugent.be; geert.vanloo@irc.vib-ugent.be

Importantly, ZEB2 expression in multiple human cancers is linked to metastasis and poor prognosis^{15–17}.

In steady state, ZEB2 is constitutively expressed in cells of mesenchymal origin, including immune cells and fibroblasts, but absent in epithelial cells. However, ectopic epithelial ZEB2 expression patterns are observed in CRC and are linked to poor oncologic outcome and distant recurrence¹⁸. How aberrant epithelial ZEB2 expression contributes to CRC malignancy remains unclear.

To study the role of intestinal epithelial Zeb2 *in vivo*, we generated intestinal epithelial cell (IEC)-specific Zeb2-expressing mice (Zeb2^{IEC-Tg/+} mice). Zeb2 expression in IECs induces intestinal barrier disruption and inflammation, leading to the development of invasive CRC in mice. However, microbiota depletion in Zeb2^{IEC-Tg/+} mice completely prevents CRC development, indicating that the intestinal microbiota is driving inflammation-associated neoplastic transformation of Zeb2-expressing IECs. Together, our study shows that aberrant epithelial Zeb2 expression is sufficient to induce invasive CRC development, which is driven by host–microbiota interactions.

Results

Mono-allelic expression of Zeb2 in the mouse intestinal epithelium induces invasive CRC. To investigate the effects of Zeb2 expression in the intestinal epithelium, we crossed conditional Rosa26-Zeb2 transgenic mice (Extended Data Fig. 1a,b)^{19,20} with a transgenic mouse line that expresses Cre under the control of the IEC-specific villin gene regulatory sequences (villin-Cre). The villin-Cre transgenic line targets all epithelial cell lineages of the distal small intestine, cecum and colon, and its expression leads to efficient Cre-mediated recombination in IECs starting before birth²¹. Upon Cre expression and excision of the floxed transcriptional stop cassette (Extended Data Fig. 1a) in IECs, a bicistronic transcript encoding Zeb2 and the enhanced green fluorescent protein reporter (eGFP) is expressed from the Rosa26 promoter. Heterozygous Zeb2 expression in IECs after villin-Cre-mediated excision was sufficient for a sixfold upregulation of Zeb2 and eGFP messenger RNA (mRNA) levels in the intestinal epithelium (Fig. 1a).

IEC-specific Zeb2-expressing mice (R26-Zeb2^{Tg/+}villin-Cre^{Tg/+}; referred to as Zeb2^{IEC-Tg/+}) are born according to Mendelian segregation and newborns are indistinguishable from their wild-type control littermates (R26-Zeb2^{Tg/+}villin-Cre^{+/+}). However, from approximately 18 weeks of age, Zeb2^{IEC-Tg/+} mice start losing body weight and start dying (Fig. 1b,c). The median lifespan of Zeb2^{IEC-Tg/+} animals is 25 weeks and none of these mice survives the age of 1 year (Fig. 1c). Already from a young age (5 weeks), Zeb2^{IEC-Tg/+} mice started suffering from diarrhea and started displaying blood in their feces—two parameters indicative of intestinal pathology (Fig. 1d)—and disease occurred independent of gender (Extended Data Fig. 1c,d). High-resolution endoscopy showed a severe and progressive carcinoma development in Zeb2^{IEC-Tg/+} mice, in contrast with wild-type littermate mice, which displayed a healthy mucosal surface and architecture (Fig. 1e,f). The colonic mucosal surface of Zeb2^{IEC-Tg/+} mice transformed from a visibly granular tissue with few bleeding spots in young animals into a highly nodular one in older mice, characterized by colon wall thickening and deposition of fibrin, gradually obstructing the intestinal lumen (Fig. 1e). Necropsy of Zeb2^{IEC-Tg/+} mice further revealed gross abnormalities of the colon, characterized by increased colon weight, severe dysplasia of the colonic epithelium and mucus accumulation in multiple tumor nodules, which developed mainly in the proximal colon, but never in the small intestine (Fig. 1g–j).

Intestinal epithelial Zeb2 expression drives severe invasive CRC in the absence of cooperating genetic defects. Histological analysis of colon sections from adult Zeb2^{IEC-Tg/+} mice showed gross changes in the epithelial tissue of these animals. Epithelial abnormalities started

appearing in young Zeb2^{IEC-Tg/+} mice at around 5 weeks of age (Fig. 2a). Older mice displayed a highly dysplastic colonic epithelium, which developed into an adenocarcinoma with formation of epithelial invasions and outgrowth of the colon muscle wall (Fig. 2b and Extended Data Fig. 2). More specifically, Zeb2^{IEC-Tg/+} mice showed a diffuse high-grade dysplasia of the mucosa associated with invasive mucinous adenocarcinoma in a background of atypical mucosal hyperplasia with multifocal intraepithelial neoplasia and intramucosal carcinoma (Fig. 2b and Extended Data Fig. 2). Using immunohistochemistry for eGFP—the reporter translated from the bicistronic Zeb2-IRES-eGFP transcript—we could demonstrate tissue-specific transgene expression throughout the colonic epithelial layer, including the neoplastic invasive lesions of Zeb2^{IEC-Tg/+} mice, while the colons of control animals were all negative for eGFP (Fig. 2c). Epithelial cell proliferation was determined by Ki67 immunohistochemistry and flow cytometry (Fig. 2c,d). Slow-cycling stem cells, which are only sporadically Ki67⁺, are positioned at the crypt bottom and give rise to active cycling (high Ki67⁺) transit-amplifying cells, which are located above the stem cell compartment in the colons of wild-type mice²². However, in Zeb2^{IEC-Tg/+} colons, proliferating (Ki67⁺) cells occupy the entire crypt length, including the crypt bottom, and also the intramucosal lesions contain mainly proliferating cells (Fig. 2c and Extended Data Fig. 3a,b). Activation of the transcription factor STAT3 has been demonstrated in many tumor types, including CRC²³. In agreement, Stat3 activation, shown by the presence of nuclear phospho-Stat3 immunostaining, could be demonstrated in Zeb2^{IEC-Tg/+} mice (Fig. 2c). Also, TGF- β signaling, linked with EMT induction, could be observed in the colon epithelium and tumor stroma of Zeb2^{IEC-Tg/+} mice, as demonstrated by phospho-Smad3 immunostaining (Fig. 2c). Increased expression of the stem cell surface marker CD44 could be observed in colon lysates and on tissue sections of Zeb2^{IEC-Tg/+} mice (Fig. 2e,f and Extended Data Fig. 3a). Interestingly, elevated expression of CD44 is a hallmark of highly tumorigenic CRC cells with metastatic properties^{24–26}. Surprisingly, no structural abnormalities nor pathology could be observed in the small intestines of Zeb2^{IEC-Tg/+} mice, although the Zeb2 transgene was clearly expressed, as demonstrated by positive eGFP staining on tissue sections and expression of Zeb2 and eGFP in tissue lysates of Zeb2^{IEC-Tg/+} mice (Fig. 2g and Extended Data Fig. 3c). In addition, no difference in cell proliferation (Fig. 2g,h and Extended Data Fig. 3d), nor differences in the expression of inflammatory mediators, CD44 or Paneth cell markers, could be observed in the small intestines between Zeb2^{IEC-Tg/+} and control animals (Extended Data Fig. 3e).

Next, we orthotopically transplanted single-cell suspensions isolated from Zeb2 transgenic tumor tissue into the cecum of immunocompromised NOD *scid* gamma (NSG) mice, to assess whether these cells can induce the formation of tumors. Twenty weeks after transplantation, small Zeb2 transgenic lesions could be identified in the ceca of recipient NSG mice, as identified by the expression of GFP, demonstrating the tumorigenic property of Zeb2 transgenic tissue (Extended Data Fig. 4a).

Given the lethal and invasive character of Zeb2-driven cancer in Zeb2^{IEC-Tg/+} mice, we next evaluated whether we could detect additional genomic aberrations in cancer tissue of Zeb2^{IEC-Tg/+} mice. Copy number analysis by shallow whole-genome sequencing revealed no insertions or deletions in CRC tissue from Zeb2^{IEC-Tg/+} mice compared with wild-type colon, or to control lung tissue from Zeb2^{IEC-Tg/+} mice (Extended Data Fig. 4b). In addition, a hot-spot mutation analysis was performed to sequence genes that are recurrently mutated in CRC patients (*APC*, *BRAF*, *CTNNB1*, *KRAS*, *NRAS*, *TP53*, *PIK3CA* and *PTEN*). However, no mutations were identified in all genes analyzed (data not shown).

Zeb2 transgenic expression in IECs induces intestinal barrier disruption, bacterial infiltration and colon inflammation during CRC development. The monolayered intestinal epithelium

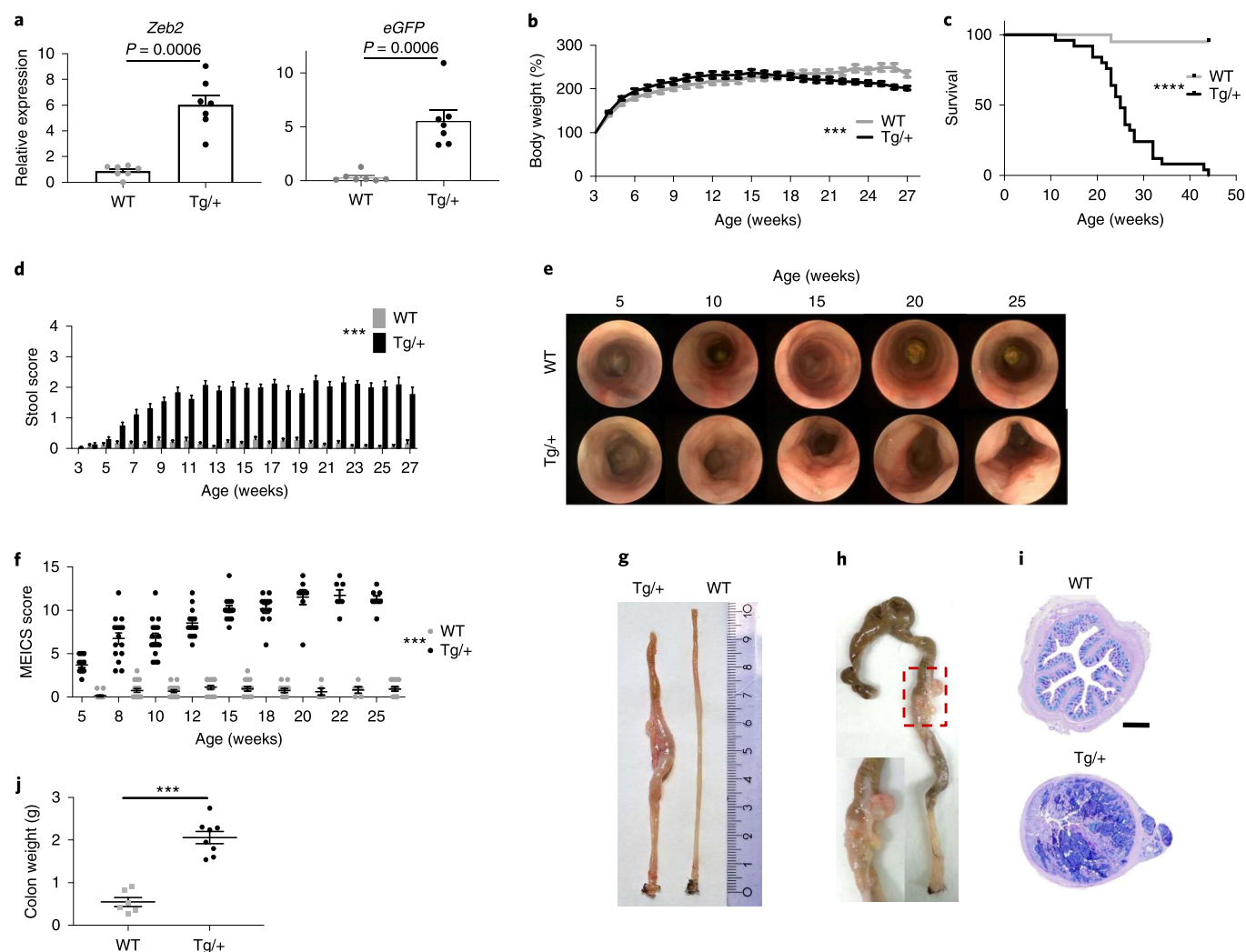


Fig. 1 | Mono-allelic expression of *Zeb2* in the mouse intestinal epithelium induces CRC. **a**, Quantitative PCR analysis for *Zeb2* and *eGFP* in colonic epithelial lysates from *Zeb2*^{IEC-Tg/+} (Tg/+) mice ($n=7$) and wild-type littermate controls (WT; $n=7$). **b**, Body weight of Tg/+ ($n=28$) and control mice ($n=27$) as a function of time. $***P<0.001$. **c**, Kaplan-Meier survival curve of Tg/+ ($n=25$) and control mice ($n=20$). $****P<0.0001$. **d**, Clinical stool scores of Tg/+ ($n=28$) and control mice ($n=27$) based on Hemocult assay. $***P<0.001$. **e**, Representative endoscopic pictures of 5-, 10-, 15-, 20- and 25-week-old Tg/+ (minimum $n=6$ mice per age) and control mice (minimum $n=5$ mice per age). **f**, Colon pathology detected by endoscopy and represented as MEICS scores. Each symbol represents one mouse. $***P<0.001$ (minimum $n=5$ mice per age for WT; minimum $n=6$ mice per age for Tg/+). **g**, Representative macroscopic images of the colon from a Tg/+ mouse and a control mouse (representative of $n=10$ mice of each genotype). **h**, Representative macroscopic image of a colon from a Tg/+ mouse (representative of $n=10$ Tg/+ mice). Inset: magnified image of the area shown by the red dashed box, depicting an invasive colon tumor. **i**, Representative Alcian Blue/PAS staining on colon sections of Tg/+ (top) and WT mice (bottom). Images are representative of $n=10$ mice of each genotype. Scale bar: 1,000 μm . **j**, Weights of Tg/+ and control colons. Each symbol represents one mouse. $***P=0.0007$. All data are presented as means $\pm$ s.e.m. Statistical significance was determined by two-tailed Mann-Whitney *U*-test (**a** and **j**), the residual maximum-likelihood method in Genstat (**b**, **d**, **f**) or log-rank Mantel-Cox test (**c**).

forms a tight physical barrier separating large amounts of commensal bacteria from underlying immune cells, ensuring intestinal immune homeostasis. Loss of barrier integrity allows translocation of microbes and their products across the epithelial lining, which activates underlying immune cells to mount local inflammatory responses. Structural intestinal barrier instability can lead to continuous exposure to bacterial products and chronic mucosal inflammation, which is now widely accepted to facilitate neoplastic transformation of the intestinal epithelium^{27–29}.

To analyze the stability of the intestinal barrier in the phenotype of *Zeb2*^{IEC-Tg/+} mice, we assessed barrier permeability by oral gavage of fluorescein isothiocyanate (FITC)-labeled dextran and measurement of fluorescence in serum 4 h after FITC-labeled dextran administration. Surprisingly, and already at young age,

FITC-labeled dextran translocation into the circulation was observed in *Zeb2*^{IEC-Tg/+} mice, but not in control littermate mice, demonstrating intestinal barrier disruption in *Zeb2*^{IEC-Tg/+} mice (Fig. 3a). Next, we performed bacterial 16S fluorescence in situ hybridization (FISH) on Carnoy-fixed colon samples to visualize the bacteria-mucus-epithelium interface. Intestinal bacteria are usually separated from the intestinal epithelium by an impermeable inner mucus layer³⁰. Colon tissues of young (7-week-old; precancerous stage) *Zeb2*^{IEC-Tg/+} mice display an inner mucus layer that is abnormally thin and discontinuous, compromising the spatial segregation of bacteria from the epithelial surface, and allowing bacterial mucosal association and infiltration (Fig. 3b).

Epithelial barriers depend on cell-cell interactions and the activity of intercellular junctions for permeability control. Indeed,

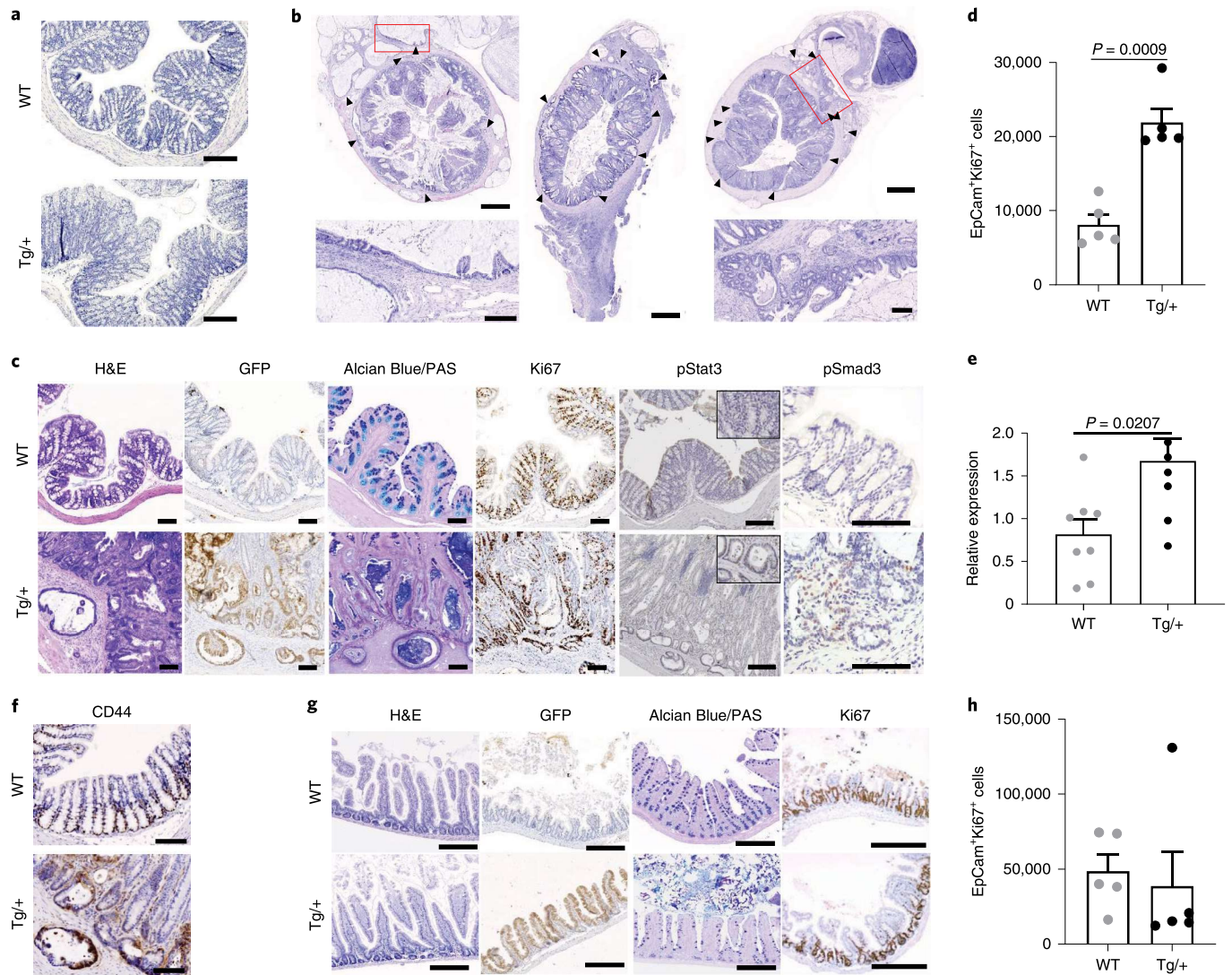


Fig. 2 | Zeb2^{IEC-Tg/+} mice develop severe pathology in the colon but not in the small intestine. **a**, Representative images of colon sections of 5-week-old Tg/+ and control mice, showing early stages of pathology (representative of $n=5$ mice of each genotype). Scale bars: 200 μ m. **b**, Top: hematoxylin and eosin-stained sections from the colons of 30- to 34-week-old Tg/+ mice ($n=7$), showing multicentric invasive adenocarcinoma in a background of atypical mucosal hyperplasia with multifocal intraepithelial neoplasia and intramucosal carcinoma (invasive fronts and transmural lesions are marked with an arrowhead). Bottom: magnified images of the areas shown by the red dashed boxes, showing transmural epithelial invasion. The intramucosal carcinoma is characterized by tumultuous proliferation of irregular glandular units and of the invading adenocarcinomatous component, with abundant mucus collection and transmural infiltration of the intestinal wall. Scale bars: 1,000 μ m (main) and 200 μ m (inset). **c**, Representative images of colon sections of Tg/+ ($n=5$) and control mice ($n=5$), showing the expression of eGFP, the presence of mucus (visualized by Alcian Blue/PAS staining), cell proliferation (Ki67), and Stat3 and Smad3 phosphorylation. Scale bars: 100 μ m. H&E, hematoxylin and eosin. **d**, Quantification by flow cytometry of the number of Ki67⁺EpCam⁺ epithelial cells in colon tissue lysates of Tg/+ ($n=5$) and control mice ($n=5$). $***P < 0.001$. **e**, Quantitative PCR analysis for CD44 in colonic epithelial lysates from Tg/+ ($n=8$) and control mice ($n=8$). **f**, Representative images of colon sections of Tg/+ ($n=5$) and control mice ($n=5$) stained for CD44. Scale bar: 100 μ m. **g**, Representative images of sections from the small intestines of Tg/+ ($n=5$) and control mice ($n=5$). Scale bars: 200 μ m. **h**, Quantification by flow cytometry of the numbers of Ki67⁺EpCam⁺ epithelial cells in tissue lysates from the small intestines of Tg/+ ($n=5$) and control mice ($n=5$). All data are presented as means $\pm$ s.e.m. Statistical significance was determined by a two-sided Student's t -test (d), or two-tailed Mann-Whitney U -test (e).

a downregulation of the tight junction protein zona occludens 1 (ZO-1) can be demonstrated in colon tissues of Zeb2^{IEC-Tg/+} mice. In control mice, ZO-1 anchored membrane proteins to the cell cytoskeleton and could be visualized by a tight honeycomb-like staining pattern (Fig. 3c). However, no such staining pattern could be observed in Zeb2-expressing tumor tissue, indicating loss of epithelial cellular junction stability (Fig. 3c). In contrast, and complementary with its function in EMT, Zeb2 expression in Zeb2^{IEC-Tg/+} mice induced the upregulation of the mesenchymal markers fibronectin and vimentin in colonic lysates (Fig. 3d). In agreement, inducible

expression of ZEB2 in the CRC cell line DLD-1 induced marked loss of cell adhesion molecules involved in tight and gap junctions (Extended Data Fig. 5).

Since no pathology could be observed in the small intestine of Zeb2^{IEC-Tg/+} mice (Fig. 2g and Extended Data Fig. 3), we also quantified expression of tight junction proteins and desmosomal proteins in the small intestine of Zeb2^{IEC-Tg/+} mice. While most of these proteins showed a trend towards reduced expression, Cldn-4 and Dsc-3 were substantially downregulated in the small intestine of Zeb2^{IEC-Tg/+} mice compared with control mice (Extended Data Fig. 6).

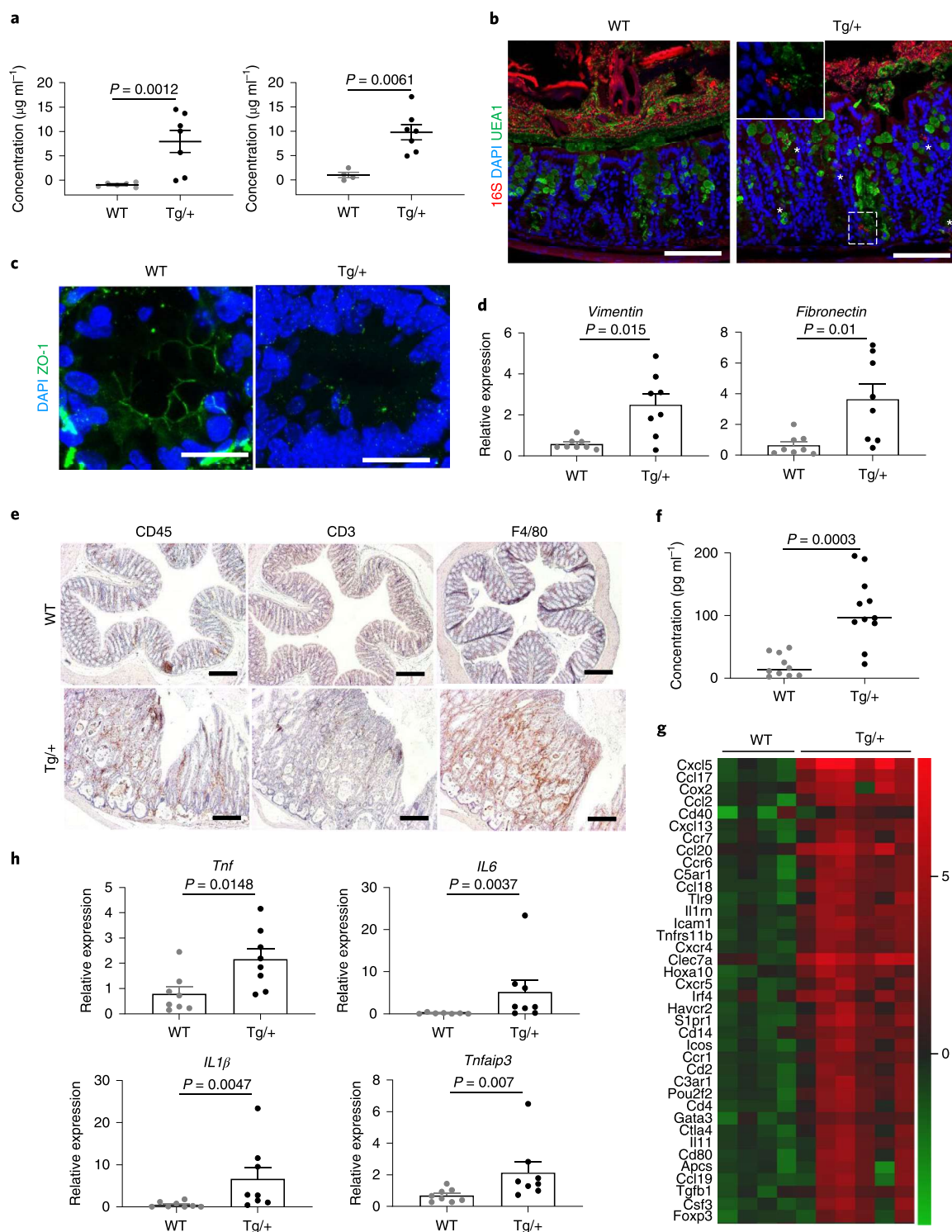


Fig. 3 | Zeb2 transgenic expression in IECs induces intestinal barrier disruption, bacterial infiltration and colon inflammation. **a**, Intestinal permeability assay using FITC-labeled dextran in 7-week-old (left) and 14-week-old (right) Tg/+ ($n=7$) and control mice ($n=6$ and 4 , respectively). Each symbol represents one mouse. Both datasets are representative of two independent experiments. **b**, Immunofluorescence staining for the mucus UEA1 lectin and 16S bacterial FISH on colon sections from 7-week-old Tg/+ and control mice (representative of $n=6$ mice of each genotype). Inset: magnified image of the area shown by a white dashed box, showing penetration of bacteria into the crypt region. Scale bars: $100\mu\text{m}$. **c**, ZO-1 immunostaining on colon sections from 25-week-old Tg/+ and control mice (representative of $n=4$ mice of each genotype). Scale bars: $20\mu\text{m}$. **d**, Quantitative PCR analysis for vimentin and fibronectin on IEC colon lysates from Tg/+ ($n=8$) and control mice ($n=8$). **e**, Immunostaining on colonic sections of Tg/+ and control mice for CD45⁺ leukocytes, CD3⁺ T cells and F4/80⁺ macrophages (representative of $n=5$ mice of each genotype). Scale bars: $100\mu\text{m}$. **f**, Serum levels of CXCL1 in Tg/+ ($n=11$) and control mice ($n=10$). **g**, Heatmap representing log₂-transformed quantitative PCR immune panel expression analysis on colon lysates from Tg/+ ($n=6$) and control mice ($n=4$). The log₂ (fold-change values) and the color scale are shown at the right of the heat map. **h**, Quantitative PCR analysis for inflammatory cytokines TNF, IL-1 β , IL-6 and A20 (*Tnfaip3*) measured in the colon lysates from Tg/+ ($n=8$) and control mice ($n=8$). All data are presented as means $\pm$ s.e.m. Statistical significance was determined by two-tailed Mann-Whitney *U*-test (**a**, **d**, **f** and **h**).

These data show that although barrier integrity is affected by Zeb2 expression, neither inflammation nor tumor pathology develops in the small intestine of Zeb2^{IEC-Tg/+} mice, probably due to the absence of sufficient numbers of bacteria in the small intestine.

Histological examination of colon sections revealed the presence of large numbers of immune cells infiltrating the colonic mucosa (Fig. 2a,b and Extended Data Fig. 2). To visualize the immune cell infiltrates in the colons of Zeb2^{IEC-Tg/+} mice, we performed a CD45 immunostaining and were able to confirm immune cell infiltrates in colonic sections of Zeb2^{IEC-Tg/+} mice (Fig. 3e). Furthermore, F4/80 staining revealed the presence of large numbers of macrophages in sections from the colons of Zeb2^{IEC-Tg/+} mice (Fig. 3e), but only small numbers of CD3⁺ T cells (Fig. 3e). Also, high levels of CXCL1 (also known as KC) could be detected in the sera of Zeb2^{IEC-Tg/+} mice (Fig. 3f). CXCL1 is mainly expressed by macrophages, neutrophils and epithelial cells and acts as a potent neutrophil chemoattractant³¹. Next, using an extensive quantitative reverse transcription PCR-based immune array for inflammatory mediators, we confirmed an inflammatory gene signature in Zeb2^{IEC-Tg/+} tumor tissue, demonstrating high expression of various immune and inflammatory genes in Zeb2^{IEC-Tg/+} colon tissue compared with control tissue (Fig. 3g). Quantitative PCR analysis also confirmed enhanced expression of the prototype inflammatory cytokines tumor necrosis factor (TNF), interleukin-1 β (IL-1 β) and IL-6, and of the nuclear factor- κ B (NF- κ B) response gene A20 (*Thfaip3*), in colonic lysates of Zeb2^{IEC-Tg/+} mice relative to control lysates (Fig. 3h).

Immune profiling of myeloid and T cell compartments in colon tissue using flow cytometry revealed strong accumulation of mainly myeloid cells, including neutrophils, monocytes, macrophages, dendritic cells and eosinophils in the colonic mucosa of Zeb2^{IEC-Tg/+} mice, as shown on t-distributed stochastic neighbor embedding (t-SNE) plots and quantification of specific myeloid subsets (Fig. 4a,b). Conversely, we found reduced levels of CD8⁺ T cells, reduced T_H1 (CD4⁺T-bet⁺) and increased regulatory T cell levels (CD4⁺FoxP3⁺) in colon lamina propria of Zeb2^{IEC-Tg/+} mice (Fig. 4a,b). CRCs characterized with high myeloid cell and low T cell infiltrates correlate with worse disease outcome and immunotherapy response^{32–34}. To evaluate the tumor-promoting role of myeloid cells in Zeb2^{IEC-Tg} mice, we treated mice with GR-1-depleting antibodies. Antibody-mediated depletion of GR-1⁺ myeloid cells induced significant reduction in lamina propria neutrophils (Fig. 4c), and induced significant protection from CRC development, as shown by endoscopic and histological analysis (Fig. 4d–f). These data indicate that the myeloid cell compartment significantly contributes to tumor development in Zeb2^{IEC-Tg/+} mice.

Together, our data show that Zeb2^{IEC-Tg/+} mice display a loss in intestinal barrier integrity, which allows penetration of microbial products into the underlying tissue. Loss of barrier stability and bacterial translocation in Zeb2^{IEC-Tg/+} mice induces strong local myeloid immune cell infiltration and inflammatory cytokine

production, which installs a state of chronic inflammation known to facilitate neoplastic transformation.

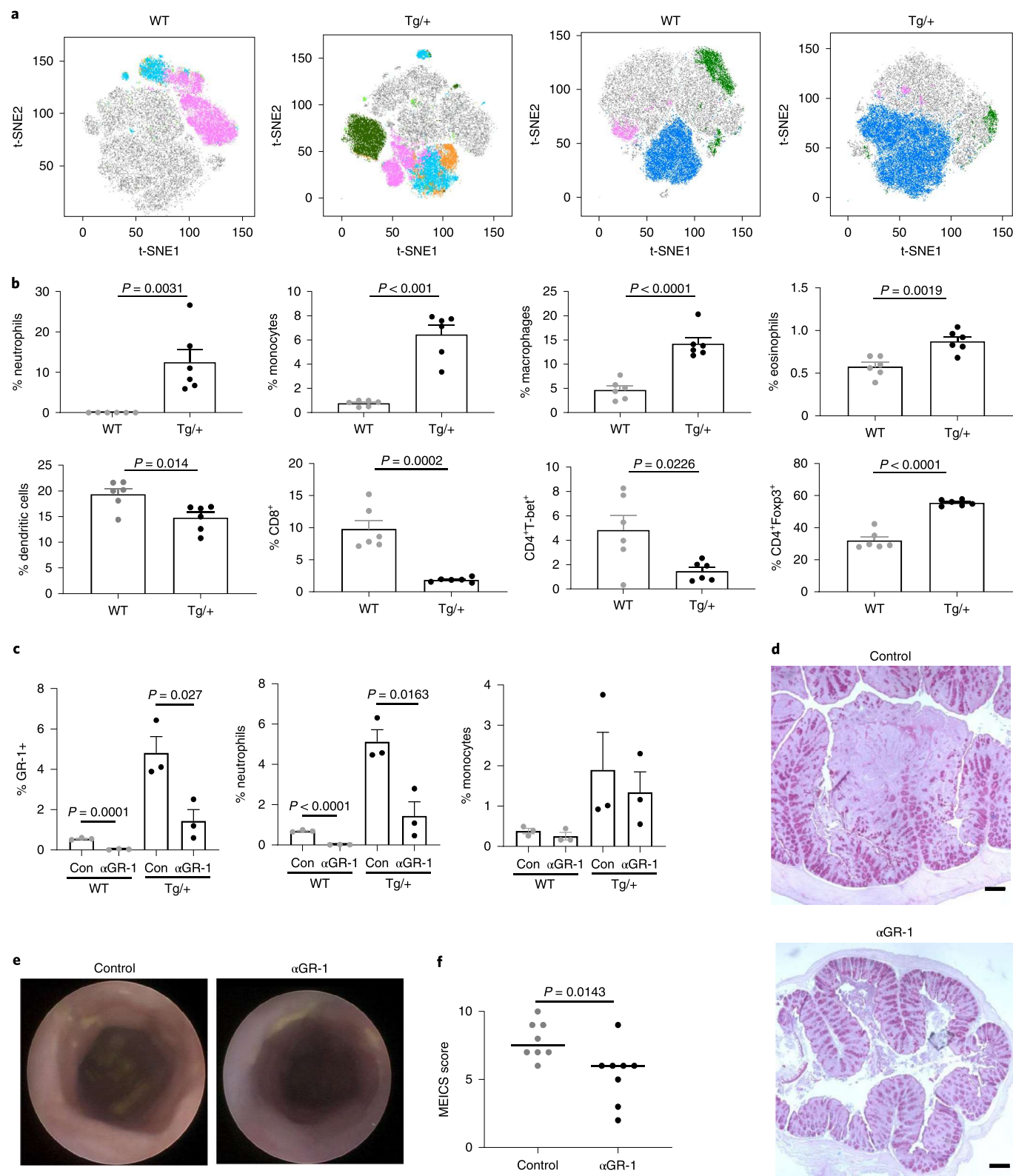
Zeb2 IEC transgenic mice are characterized by microbial dysbiosis. Loss of mucosal barrier integrity and microbial mucosal penetration in young Zeb2^{IEC-Tg/+} animals (Fig. 3) suggests that the microbiota play a role in the development of chronic mucosal inflammation and colon carcinoma. To characterize the microbiota composition in Zeb2^{IEC-Tg/+} mice versus wild-type littermate control mice, we performed 16S ribosomal RNA and ITS2 gene amplicon sequencing from stool samples of 25-week-old mice, to examine the bacterial and fungal community structure, respectively. By exclusively using heterozygous breeding couples and cohousing of the different genotypes, breeding and caging effects on the microbiota can be excluded. Regarding the bacterial diversity, Zeb2^{IEC-Tg/+} mice displayed increased α -diversity at the genus level compared with wild-type littermate controls (Fig. 5a). The β -diversity analysis showed two distinct communities between wild-type and Zeb2^{IEC-Tg/+} mice (Fig. 5b,c). These communities were also different from blank controls (data not shown), ruling out the possibility of any major contamination event during the generation of the data. To further explore key phylotypes that may contribute to the observed differences in microbial communities, linear discriminant analysis effect size (LEfSe) analysis was performed. The Zeb2^{IEC-Tg/+} fecal microbiota was characterized by a lower abundance of the S24-7 family and the *Lactobacillus* and *AF12* genera. Conversely an increase was observed in several taxa including: *RF32* and *Bacteroidales* orders; *Clostridiaceae*, *Desulfovibrionaceae*, *Enterobacteriaceae* and *Helicobacteraceae* families; *Bacteroides*, *Prevotella*, *Turicibacter*, *Parabacteroides*, *Enterococcus* and *Coprobacillus* genera; and *Bacteroides acidifaciens*, *Bacteroides distans* and *Akkermansia muciniphila* species (Fig. 5c,d and Extended Data Fig. 7). In colonic mucosal samples, the same dysbiotic trends were observed: increased bacterial diversity in Zeb2^{IEC-Tg/+} mice (Extended Data Fig. 8a) and differences in bacterial community structure (Extended Data Fig. 8b). We again identified a decreased abundance of the *Lactobacillus* genus and an increased abundance of the *Helicobacter* genus in the colonic mucosa of Zeb2^{IEC-Tg/+} mice (Extended Data Fig. 8c–e). In contrast with bacterial communities, we found no differences in fungal α - and β -diversity between genotypes in feces (Extended Data Fig. 8f,g). Taken together, these results showed that Zeb2^{IEC-Tg/+} mice at the age of established disease are characterized by marked bacterial dysbiosis.

The intestinal microbiota is required for CRC development in Zeb2 IEC transgenic mice. To determine the extent to which commensal intestinal microbes contribute to CRC development in Zeb2^{IEC-Tg/+} mice, we depleted the resident microbiota by treating these mice with a cocktail of broad-spectrum antibiotics (ampicillin, ciprofloxacin, metronidazole and vancomycin) from weaning age onwards. Interestingly, antibiotic treatment completely

Fig. 4 | Myeloid cell infiltrates contribute to CRC in Zeb2 IEC transgenic mice. **a,b**, t-SNE plots showing myeloid (blue: macrophages; yellow: monocytes; dark green: neutrophils; pink: dendritic cells; light green: eosinophils) (left) and T cell infiltration (blue: CD4⁺FOXP3⁺; pink: CD4⁺T-bet⁺; green: CD8⁺) (right) in 16-week-old Tg/+ ($n = 6$) versus wild-type mice ($n = 6$) (**a**), and corresponding quantifications of colonic lamina propria leukocyte subsets by flow cytometric analysis (**b**). Each symbol represents one mouse. **c**, Flow cytometry analysis confirming GR-1⁺ cell depletion in the colons of wild-type and Tg/+ mice after treatment with neutralizing anti-GR-1 (α -GR-1; $n = 3$ mice) versus isotype control antibodies (con; $n = 3$ mice). Lamina propria cells were stained with either anti-Ly6G and anti-Ly6C antibodies or with an anti-GR-1 antibody. The experiment was performed twice with similar results. Each symbol represents one mouse. **d**, Representative Alcian Blue/PAS-stained sections from the colons of 8-week-old Tg/+ mice treated with neutralizing anti-GR-1 or isotype control antibodies. Scale bars: 200 μ m. Images are representative of two independent experiments with similar results. **e**, Representative endoscopic images of 8-week-old Tg/+ mice treated with neutralizing anti-GR-1 or isotype control antibodies. Images are representative of two independent experiments with similar results. **f**, Colon pathology in 8-week-old Tg/+ mice treated with neutralizing anti-GR-1 ($n = 8$ mice) or isotype control antibodies ($n = 8$ mice), detected by endoscopy and represented as MEICS scores. Data from two independent experiments were pooled. Each symbol represents one mouse. All data are presented as means $\pm$ s.e.m. Statistical significance was determined by two-sided Student's *t*-test (**b**, **c** and **f**).

prevented CRC development and lethality in $Zeb2^{IEC-Tg/+}$ mice (Fig. 6a). No CRC pathology relating to loss of body weight (Fig. 6b), presence of fecal blood or diarrhea (Fig. 6c), changes in colon appearance and weight (Fig. 6d,e), expression of inflammatory markers (Fig. 6f), colon pathology based on colonoscopy (Fig. 6g), murine endoscopic index of colitis severity (MEICS) analysis (Fig. 6h) or

histology (Fig. 6i) could be detected in antibiotic-treated $Zeb2^{IEC-Tg/+}$ mice, in contrast with non-treated $Zeb2^{IEC-Tg/+}$ mice, which all developed CRC. Although antibiotic-treated $Zeb2^{IEC-Tg/+}$ mice are completely protected from the development of CRC, they still presented intestinal barrier leakage, comparable to the non-treated $Zeb2^{IEC-Tg/+}$ mice (Fig. 6j). This suggests that cell-intrinsic effects resulting from



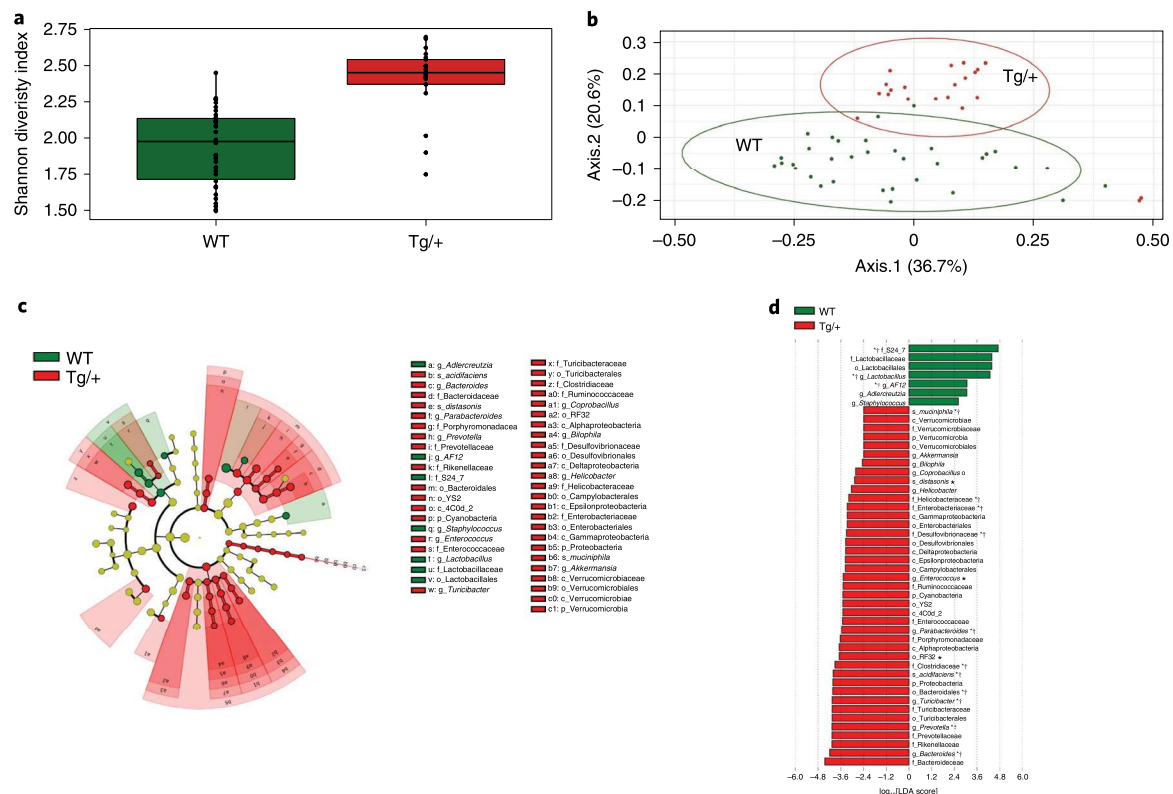


Fig. 5 | Microbiota dysbiosis in Zeb2 IEC transgenic mice. **a,b**, Altered bacterial diversity/community structure in fecal samples from 25-week-old wild-type (green; $n = 21$) and Tg/+ mice (red; $n = 34$). In **a**, α -diversity was measured using the Shannon diversity index at the genus level ($P = 3.5 \times 10^{-8}$; two-sided Wilcoxon test). Each box represents the third (upper) and first (lower) quartiles. Lines inside the boxes represent median values. In **b**, β -diversity is represented by a principal coordinates analysis of Bray–Curtis dissimilarity distances of bacterial composition. **c,d**, Differentially abundant taxa between wild-type and Tg/+ mice. **c** shows an LfSe cladogram of differentially abundant taxa in fecal samples from WT (green; $n = 21$) and Tg/+ mice (red; $n = 34$). **d** shows linear discriminant analysis (LDA) scores of differentially abundant taxa combined with multiple testing corrections (analysis associated to **c**): MaAsLin test ($^1P_{adj} < 0.05$) and t -test with false discovery rate correction ($^*P_{adj} < 0.05$). Letters before underscores show the taxonomic rank: c, class; f, family; g, genera; o, order; p, phylum; s, species.

transgenic Zeb2 expression directly impact on intestinal barrier stability, but are not sufficient to drive hyperproliferation or malignant transformation of the intestinal epithelium in the absence of the microbiota. In addition to treating mice with a four-component broad-spectrum antibiotic cocktail, we treated Zeb2^{IEC-Tg/+} mice with selective antibiotics to kill either Gram-positive (vancomycin) or Gram-negative (aztreonam) bacteria, from weaning age onwards. Interestingly, only vancomycin treatment provided protection against CRC development in Zeb2^{IEC-Tg/+} mice, suggesting that Gram-positive bacteria drive the pathology in Zeb2^{IEC-Tg/+} mice (Extended Data Fig. 9).

To rule out any side effect of antibiotics that may affect CRC development, Zeb2^{IEC-Tg/+} mice were rederived under germ-free conditions by embryo transfer in axenic recipient mice, and raised under germ-free conditions. Germ-free Zeb2^{IEC-Tg/+} mice were completely protected from CRC development and reached old age without exhibiting any intestinal abnormality (Fig. 7a–c and Extended Data Fig. 10). Also, germ-free Zeb2^{IEC-Tg/+} mice retained increased intestinal barrier permeability but did not display inflammatory gene expression in the colon (Fig. 7d,e). Collectively, these data show that Zeb2 expression in IECs causes barrier destabilization independent of the intestinal microbiota. However, the intestinal microbiota is required to cause colonic mucosal inflammation and subsequent CRC development in Zeb2^{IEC-Tg/+} mice.

In conclusion, we hypothesize that infiltration of commensal bacteria and their molecular products, caused by a more permeable epithelial barrier, is necessary to drive chronic inflammation

and IEC malignant transformation, initiating invasive colon cancer development in Zeb2^{IEC-Tg/+} mice.

Intestinal epithelial Zeb2 expression is associated with human CRC. The consensus molecular subtype (CMS) represents a robust classification system for CRC stratification, in which CMS4 is characterized by the enrichment of mesenchymal markers³⁵. We analyzed bulk gene expression profiles of tumor tissue from public online expression datasets comprising >3,500 patients with CRC^{36–40}. Unsupervised hierarchical clustering analysis showed clear segregation of the main CMS subtypes, including CMS4, which was characterized by high expression of EMT regulators including ZEB1, ZEB2, SNAIL1/SNAIL and SNAIL2/SLUG (Fig. 8a). Remarkably, several of the highly expressed immune genes in Zeb2^{IEC-Tg/+} mouse colon tissue were also prominently expressed in human CMS4 CRC (including *Cxcl5*, *Ptgs2/Cox2*, *Cxcl13*, *C5AR1*, *CCL2*, *CD40* and *CCR7*; Figs. 3g and 8a). ZEB2 expression, along with other mesenchymal markers in CRC tumors, is mainly attributed to infiltrating stromal and immune cells rather than to epithelial cells undergoing EMT. This observation is supported by transcriptome datasets on human colon cancer^{41–43}. However, although they are masked by strong mesenchymal cell infiltration, EMT mechanisms may still contribute to CRC in humans. We evaluated ZEB2 expression in mucosal biopsies of patients with CRC of different CMS classes, and could demonstrate that in the majority of all analyzed biopsies ZEB2 was expressed in the intestinal epithelium, particularly at invasive sites (Fig. 8b), in agreement with previous observations¹⁸.

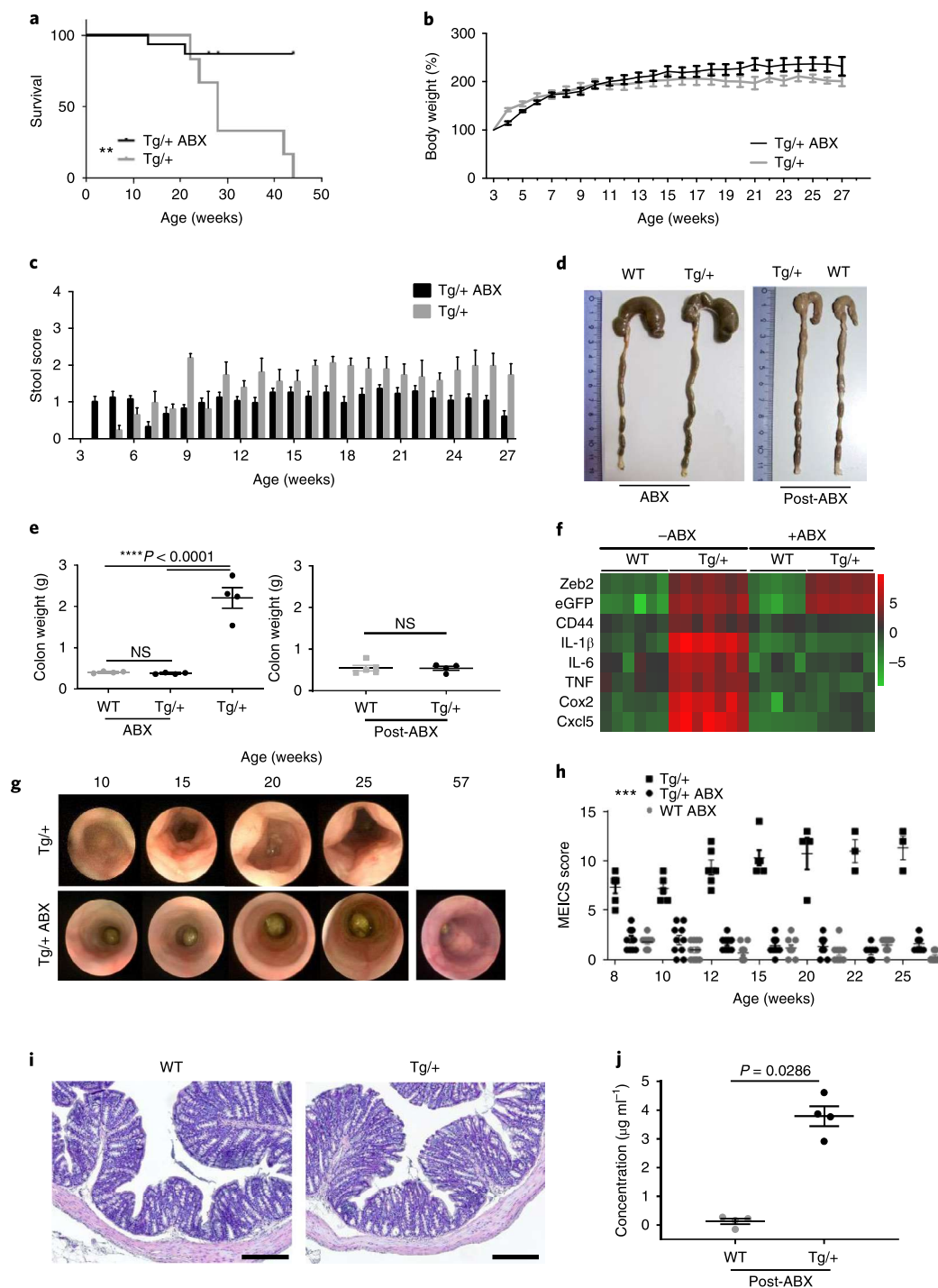


Fig. 6 | Microbiota dependency for CRC development in Zeb2 IEC transgenic mice. **a**, Kaplan-Meier survival curves of antibiotic-treated (ABX) Tg/+ mice ($n=10$) and untreated Tg/+ controls ($n=6$). $**P=0.0081$. **b**, Body weights of ABX Tg/+ mice ($n=10$) and untreated Tg/+ controls ($n=6$) as a function of time. **c**, Clinical stool scores of ABX Tg/+ mice ($n=10$) and untreated Tg/+ controls ($n=6$) based on Hemocult assay. **d**, Representative macroscopic images of the colons of ABX Tg/+ and wild-type control mice at the end of the treatment (left) and 30 weeks after cessation of antibiotics (right). Images are representative of a minimum of $n=4$ mice per genotype. **e**, Colon weights of ABX Tg/+, untreated Tg/+ and ABX wild-type littermates at the end of the treatment (left; $n=4$ mice per group) and 30 weeks after cessation of antibiotics (right; $n=5$ mice for wild type; $n=4$ mice for Tg/+). Each symbol represents one mouse. $****P<0.0001$. NS, not significant ($P=0.9882$). **f**, Heatmap representing $\log_2$ -transformed quantitative PCR expression analysis on colon lysates from non-treated (–ABX) versus antibiotic-treated (+ABX) Tg/+ and wild-type control mice. The $\log_2$ (fold-change values) and the color scale are shown at the right of the heatmap. **g**, Representative endoscopic pictures of 10-, 15-, 20-, 25- and 57-week-old ABX Tg/+ ($n=10$) and untreated Tg/+ control mice ($n=6$). **h**, Colon pathology detected by endoscopy and represented as MEICS scores. Each symbol represents one mouse ($n=10$ mice for wild type ABX and Tg/+ ABX; $n=6$ mice for Tg/+). $***P<0.001$. **i**, Representative hematoxylin and eosin-stained sections from the colons of ABX Tg/+ and untreated Tg/+ mice 30 weeks after cessation of the treatment (representative of $n=4$ mice of each genotype). Scale bars: 200 μm . **j**, Intestinal permeability assay using FITC-labeled dextran in ABX Tg/+ ($n=4$) and ABX wild-type mice ($n=4$) 20 weeks after cessation of the treatment. Each symbol represents one mouse. All data are presented as means $\pm$ s.e.m. Statistical significance was determined by log-rank Mantel-Cox test (**a**), one-way ANOVA with Tukey's correction (**e**), the residual maximum-likelihood method in Genstat (**h**) or two-tailed Mann-Whitney U -test (**j**).

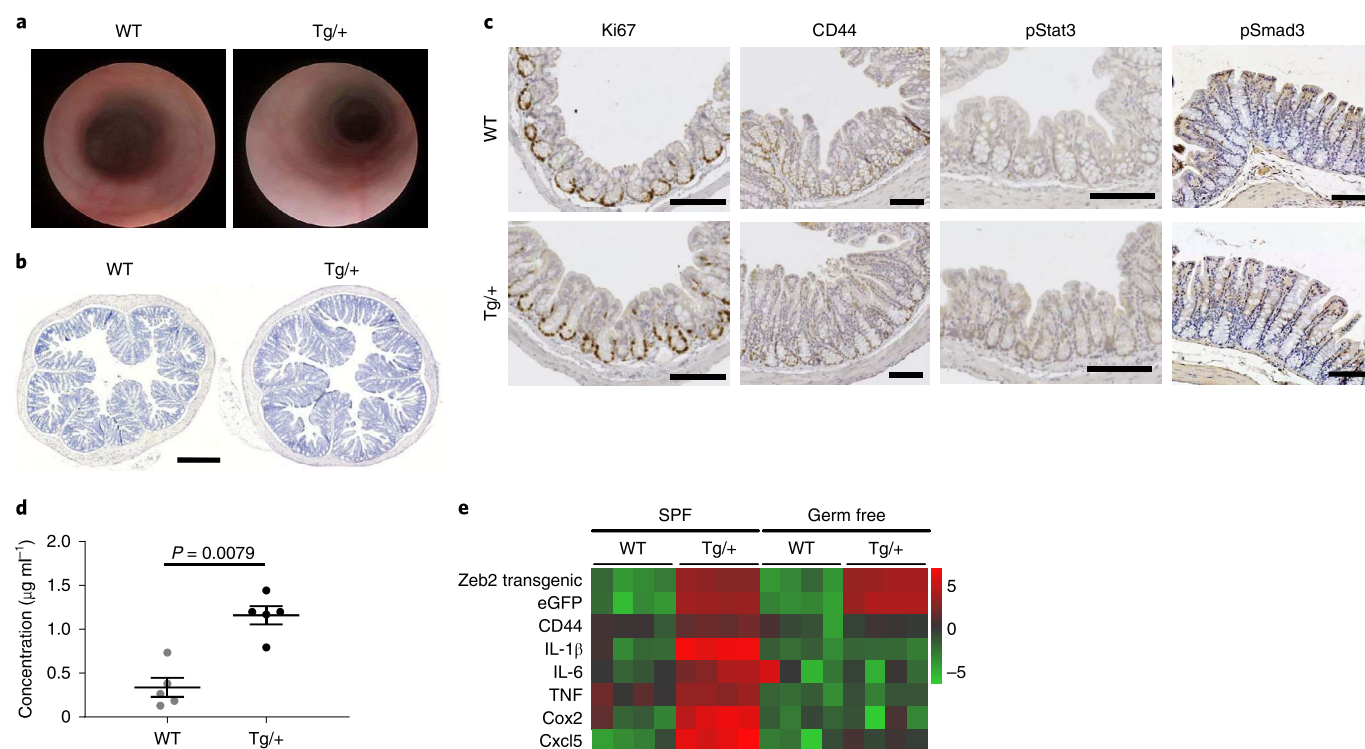


Fig. 7 | Axenic *Zeb2*^{IEC-Tg/+} mice are protected from CRC development. **a**, Representative endoscopic images of 55-week-old wild-type and Tg/+ mice. **b**, Representative hematoxylin and eosin-stained sections from the colons of 55-week-old germ-free wild-type and Tg/+ mice. Scale bar: 1,000 μm . **c**, Representative images of colon sections of 20-week-old germ-free wild-type and Tg/+ mice, showing normal cell proliferation (Ki67), CD44 expression, Stat3 phosphorylation and Smad3 phosphorylation (representative of $n = 4$ mice of each genotype). Scale bars: 100 μm . **d**, Intestinal permeability assay using FITC-labeled dextran in 12-week-old germ-free wild-type ($n = 5$) and Tg/+ mice ($n = 5$). Each symbol represents one mouse. Data are presented as means $\pm$ s.e.m. and were analyzed by two-tailed Mann–Whitney U -test. **e**, Heatmap representing $\log_2$ -transformed quantitative PCR expression analysis on colon lysates from 12-week-old specific pathogen-free (SPF) and germ-free WT and Tg/+ mice. The $\log_2$ (fold-change values) and the color scale are shown at the right of heat map.

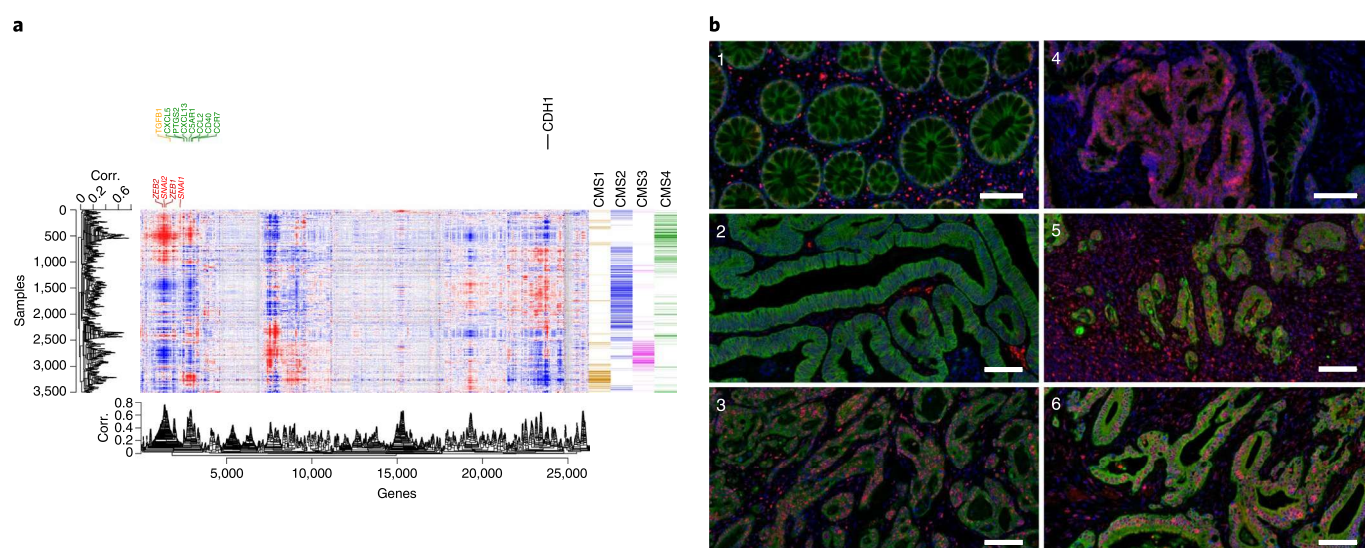


Fig. 8 | ZEB2 expression in human CRC. **a**, Heatmap depicting the results of hierarchical clustering of human CRC gene expression. Rows represent patients ($n = 3,511$); columns represent genes. Colors represent decreased (blue) or increased (red) normalized expression. CMS subtypes are denoted on the right-hand side of the heatmap as colored bars. The dendrogram scales correspond to average linkage of Pearson correlations as a similarity measure ('Corr.'). **b**, ZEB2 expression in human CRC biopsies. Images are representative of immunofluorescence staining for ZEB2 (red) and an epithelial pan-keratin marker (green) on colon sections from: (1) healthy colon with ZEB2 expression in lamina propria mesenchymal cells; (2) ZEB2-negative colonic CMS2 adenocarcinoma; (3) and (4) ZEB2-positive colonic adenocarcinoma (CMS subtype unknown); (5) ZEB2-positive colonic CMS4 adenocarcinoma; and (6) ZEB2-positive colonic CMS1 adenocarcinoma. Scale bars: 100 μm ($n = 4$ samples).

In contrast, ZEB2 expression was absent in the epithelium of healthy tissue sections, and restricted to mesenchymal cells in the lamina propria (Fig. 8b).

Together, we conclude that ZEB2 expression is prominent in the colon epithelium of human patients with CRC, and demonstrate profound similarities between human CRC and the Zeb2^{IEC-Tg/+} mouse model in terms of expression profiles, immune components and clinical CRC manifestation.

Discussion

CRC is a devastating and deadly disease, with increasing incidence in Western society, and which develops in response to genetic, immunological and environmental cues. High ZEB2 expression has been demonstrated in many CRC cell lines, and the expression of ZEB2 was shown to accelerate tumor growth and induce metastasis in xenotransplantation models, whereas ZEB2 silencing impairs tumor growth^{15,16}. Importantly, ZEB2 expression in human tumor colonic epithelium was shown by immunohistochemistry, and was apparent in over 40% of all CRC samples analyzed¹⁸. Epithelial ZEB2 expression was associated with poor oncologic outcome and distant recurrence, and was suggested as a novel biomarker for patient stratification. Using human CRC tissues from different CMS subtypes, here we confirmed prominent ZEB2 expression not only in the stromal tissue fraction, but also in the intestinal epithelium. Moreover, we now prove that intestinal epithelial Zeb2 expression in mice is sufficient to drive invasive CRC development, which we demonstrate to depend on microbiota–host interactions, since antibiotic treatment or germ-free rederivation completely protects Zeb2^{IEC-Tg/+} mice from CRC development. Hence, we propose a model in which Zeb2-mediated partial mesenchymal reprogramming of epithelial cells, in combination with host responses to the resident microflora upon barrier destabilization, synergistically drives the development of invasive and lethal CRC development in mice. In striking contrast with the colon cancer phenotype, no abnormalities were observed in the small intestines of Zeb2^{IEC-Tg/+} mice. This observation was in line with the requirement of intense epithelium–microbiota interactions for tumor development, since bacteria reach much higher densities in the colon compared with the small intestine. Finally, we also demonstrated that orthotopic transplantation of Zeb2 transgenic tumor tissue into immunocompromised mice could induce the formation of small Zeb2 transgenic lesions. However, these lesions did not evolve to the characteristic aggressive CRC that was observed in Zeb2^{IEC-Tg/+} mice, suggesting that the inflammatory context drastically amplifies the tumor-promoting potential of these Zeb2-expressing transplants, reinforcing a synergistic microbiota-dependent inflammation/EMT model of how Zeb2 promotes tumorigenesis.

CRC is typically considered to arise from adenomatous precursors that develop into carcinomas through progressive accumulation of gain-of-function mutations in oncogenes and loss-of-function mutations in tumor suppressor genes⁷. However, in our model, increased expression of Zeb2 in IECs was sufficient to induce the spontaneous development of CRC, in the absence of additional CRC-inducing mutations. Indeed, we could not identify mutations in oncogenes or tumor suppressor genes linked to CRC nor other genomic defects based on specific hot-spot mutation analysis and shallow whole-genome sequencing of Zeb2^{IEC-Tg/+} tumors. This observation is of high clinical relevance and proves that the classical progressive adenoma–carcinoma cascade can be bypassed by conditions that lead to aberrant epithelial Zeb2 expression. However, deep sequencing and analysis for epigenetic modifications should be performed to rule out the involvement of genetic and epigenetic alterations in the development of CRC in Zeb2^{IEC-Tg/+} mice.

Given the necessity of host–microbiota interactions for the development of spontaneous CRC in Zeb2 transgenic mice, this CRC mouse model represents a unique tool with which to study

interactions between the intestinal microbiota, host immunity and CRC development. Indeed, accumulating data show that the intestinal microbiome strongly influences colon cancer development in humans^{44,45}. A variety of gut microbial species, including *Fusobacterium* and *Bacteroides fragilis*, are known to trigger inflammatory pathways and oncogenic signaling, promoting CRC development and metastasis. Strikingly, recent data indicate that microbiota components can drive metastasis by traveling along tumor cells to distant sites⁴⁶. Moreover, the composition of the microbiome was recently shown to drastically influence the efficacy of anti-cancer therapy^{47–52}. Hence, it will be highly interesting to define distinct bacterial classes that regulate CRC development in Zeb2^{IEC-Tg/+} mice. The identification of these species and their metabolites will be crucial to better understand the mechanisms behind CRC and the possibilities to treat the disease.

Although many studies have shown that intestinal barrier disruption and chronic inflammation are associated with CRC, the Zeb2^{IEC-Tg} mouse model goes beyond that association. In the Zeb2^{IEC-Tg} model, an epithelial-restricted initiating event induces barrier defects, inflammation and a very fast, fully penetrant and aggressive CRC, which develops independent of exogenous carcinogens or additional CRC-associated genetic mutations. The classic azoxymethane/dextran sodium sulfate model of inflammation-associated CRC crucially relies on combining chronic dextran sodium sulfate-induced inflammation with the azoxymethane carcinogen, indicating that chronic inflammation alone is not sufficient to efficiently induce CRC development. Moreover, numerous mouse models have been described in which an epithelial-restricted defect leads to barrier destabilization and subsequent chronic intestinal inflammation, such as in IEC-specific NEMO-, FADD-, caspase-8-, RIPK1- or XBP1-deficient mice^{53–58}. While all of these models display barrier loss and intestinal inflammation associated with increased regenerative epithelial proliferation, none of these mouse models spontaneously develops intestinal tumors, demonstrating that barrier loss and chronic intestinal inflammation do not automatically induce carcinogenesis but require additional tumor-promoting factors. This study identifies epithelial Zeb2 expression as such a CRC-promoting factor. Although CRC in Zeb2^{IEC-Tg} mice does not involve genetic mutations in the typical human CRC-associated genes, we believe this model is nonetheless very relevant for human CRC, since epithelial Zeb2 expression could be confirmed in human CRC biopsies.

In conclusion, we have identified a novel CRC-promoting mechanism that depends on synergistic interactions between EMT-prone epithelial cells, the intestinal microbiota and the host immune system. Epithelial Zeb2 initiates loss of epithelial differentiation inducing barrier instability, bacterial infiltration and mucosal inflammation. In the absence of bacteria and subsequent inflammation, partial EMT is not sufficient to induce CRC development, but only upon synergistic interactions with microbial and myeloid immune components does aggressive invasive CRC develop. This Zeb2^{IEC-Tg/+} CRC mouse model represents a new and unique tool with which to study the mechanisms by which the intestinal microbiota contributes to the development of CRC, and will enable the validation of new microbiota-based therapeutic strategies for the treatment of invasive CRC.

Methods

Animal experiments. Rosa26-Zeb2 mice have been generated and described elsewhere¹⁹. To induce specific and mono-allelic overexpression of the Zeb2 transgene in the intestinal epithelium, Rosa26-Zeb2 mice were crossed with villin-Cre transgenic mice²¹. Axenic Rosa26-Zeb2 mice were generated by embryo transfer in axenic recipients at the germ-free mouse facility of the University of Bern and were successfully imported into the germ-free facility of Ghent University for further studies. Axenic mice were housed in positive-pressure flexible film isolators (North Kent Plastics). All experiments were performed on mice of C57BL/6 genetic background. Mice were housed in individually ventilated cages

at the VIB Center for Inflammation Research in a specific pathogen-free animal facility. All experiments on mice were performed according to institutional (Ethical Committee for Animal Experimentation at Ghent University's Faculty of Sciences and Faculty of Medicine and Health Sciences), national and European animal regulations.

Stool score. Occult or gross blood lost per rectum, and stool consistency were determined weekly. The presence of fecal blood was determined using Hemocult SENSa (Beckman Coulter) analysis. For bleeding, a score of 0 was assigned for no blood, a score of 2 was assigned for positive Hemocult and a score of 4 was assigned for gross bleeding. For stool consistency, a score of 0 was assigned for well-formed pellets, a score of 2 was assigned for pasty and semiformal stools and a score of 4 was assigned for liquid stools. These scores were added together and divided by two, resulting in a total stool score ranging from 0 (healthy) to 4 (maximum pathology). To eliminate any diagnostic bias, mice were scored blindly.

Endoscopic analysis. High-resolution mouse endoscopy and MEICS scoring was performed, as previously described³⁹, using a Coloview endoscopic system (Karl Storz). Mice were anesthetized with 2–2.5% isoflurane in oxygen during the procedure.

Anti-GR-1 antibody treatment. Zeb2^{IEC-Tg/+} mice were treated from weaning age onwards by intraperitoneal injection every 2 d with neutralizing anti-GR-1 antibodies (RB6.8C5 monoclonal antibody; 40 mg kg⁻¹ body weight), and clinical pathology was scored by MEICS analysis, as described above.

Depletion of commensal intestinal bacteria. To deplete commensal bacteria, mice were treated either with a cocktail containing 200 mg l⁻¹ ciprofloxacin (Sigma–Aldrich), 1 g l⁻¹ ampicillin (Sigma–Aldrich), 1 g l⁻¹ metronidazole (Sigma–Aldrich) and 500 mg l⁻¹ vancomycin (Labconsult) in drinking water, or with vancomycin (500 mg l⁻¹; Labconsult) alone, or with aztreonam (0.1 g l⁻¹; Cayman Chemical) alone. The presence of colonic microflora was regularly determined by culturing fecal pellets in both brain heart infusion (BD) and thioglycollate medium (Sigma–Aldrich). To restore commensal microbiota, dirty bedding from control wild-type (Rosa26-Zeb2^{IEC-Tg/+} villin-Cre^{+/+}) mice was transferred into the cages.

FITC–dextran intestinal permeability assay. Intestinal permeability was assessed by oral gavage of FITC-labeled dextran (Sigma–Aldrich). Mice were administered 600 mg FITC-labeled dextran per kg body weight in sterile phosphate-buffered saline (PBS) by oral gavage. After 4 h, blood was collected from the facial vein, and FITC-labeled dextran levels were measured in 100 µl serum by fluorometry (485 nm). Mice were anesthetized during blood collection. Serial dilutions of FITC-labeled dextran in PBS were used each time to generate a standard curve, and serum from PBS-gavaged mice was used as blanks.

Histology. Mouse organs were fixed in formalin (10%; neutral buffered; Sigma–Aldrich). Samples were dehydrated, embedded in paraffin, sectioned at 5 µm, stained with hematoxylin and eosin and examined by light microscopy. For combined Alcian Blue and periodic acid Schiff stainings, dewaxed sections were rehydrated and incubated in Alcian Blue for 20 min. Sections were subsequently washed with PBS before incubation in 1% periodic acid for 10 min followed by incubation in Schiff's reagent for 10 min. Sections were counterstained with Mayer's hematoxylin for 30 s, washed and dehydrated before mounting with Depex. For immunohistochemistry, sections were dewaxed and incubated in Dako antigen retrieval solution, boiled for 20 min in a Pick cell cooking unit and cooled down for 2.5 h. Endogenous peroxidase activity was blocked with 3% H₂O₂ for 30 min at room temperature. Slides were incubated in blocking buffer (fish skin gelatin, bovine serum albumin and 5% normal goat serum) for 1 h at room temperature. Sections with primary antibodies (rabbit anti-GFP (Cell Signaling; dilution 1/200), rabbit anti-Ki67 (Cell Signaling; 1/1,000), rabbit anti-Phospho-Smad3 phosphoS423-425 (Abcam; 1/500), anti-Phospho-Stat3 (Cell Signaling; 1/500), rat anti-CD45 (BD Pharmingen; 1/50), rabbit anti-CD3 (Abcam; 1/500) and rat anti-F4/80 (Serotec; 1/100)) were incubated overnight in blocking buffer at 4 °C. After excessive washing, slides were incubated with secondary antibody (biotinylated anti-primary antibody species), followed by avidin–biotin complexes, and peroxidase activity was detected with diaminobutyric acid substrate (Vector Laboratories). Slides were counterstained with Mayer's hematoxylin, dehydrated and mounted with Depex and imaged on an Olympus BX51 microscope or Zeiss Axio Scan Digital Slide Scanner (Carl Zeiss Microscopy). For cryosections, tissues were fixed in 1% paraformaldehyde and 4% sucrose overnight at 4 °C, followed by 20% sucrose until sedimented, then embedded in O.C.T., stored at –80 °C and sectioned at 8 µm. For immunofluorescence, slides were air-dried for 1 h at room temperature, rehydrated, incubated in blocking buffer and stained overnight with rabbit anti-ZO-1 antibody (Thermo Fisher Scientific; 1/100). After excessive washing, slides were incubated with goat anti-rabbit 488 DyLight (1/300), counterstained with DAPI, mounted and imaged on a Leica SP5 confocal microscope.

FISH. Stainings were performed as described previously⁴⁰. Briefly, colon tissue with fecal pellets was fixed in cold methanol–Carnoy fixative (60% methanol,

30% chloroform and 10% glacial acetic acid) overnight at 4 °C. The next day, tissues were transferred to pure methanol and processed using standard methods and embedded in paraffin. Paraffin sections of 5 µm were processed for mucin 2 staining with fluorescein-labeled UEA1 lectin (Vector Laboratories), and bacteria were detected with the universal bacterial FISH probe EUB338 (50-CGTGCCTC CCGTAGGAGT-30), with 5'-Cy3 labeling. Slides were counterstained with DAPI, mounted and imaged on a Leica SP5 confocal microscope.

Quantitative real-time PCR. Total RNA was isolated using TRIzol reagent (Invitrogen) and an Aurum Total RNA Isolation Mini Kit (Bio-Rad), according to the manufacturer's instructions. The synthesis of complementary DNA (cDNA) was performed using the iScript Advanced cDNA synthesis kit (Bio-Rad) according to the manufacturer's instructions. cDNA was amplified for quantitative PCR in a total volume of 5 µl with a SensiFAST SYBR No-ROX Kit (Bioline) and specific primers on a LightCycler 480 (Roche). The reactions were performed in triplicates and the results were analyzed with qbase+ software (Biogazelle). As reference targets, *Cox4i1*, *Eef1a1* and *Matr3* were used (*Cox4i1* forward: AGAATGTTGGCTTCCAGAGC; *Cox4i1* reverse: TTCACAACACTCCCATGTGC; *Eef1a1* forward: TCGCCTTGGACGTTCTT TT; *Eef1a1* reverse: GTGGACTTGGCCGAATCTAC; *Matr3* forward: TGGACC AAGAGGAAATCTGG; *Matr3* reverse: TGAACAACCTCGGCTGGTTTC). The following mouse-specific primers were used: *Thnf* forward: ACCCTGGTATGAG CCAATATAC; *Thnf* reverse: ACACCCATCCCCTACAGAG; *IL1β* forward: CACCTCACAAAGCAGAGCACAAG; *IL1β* reverse: GCATTAGAAACAGTCCA GCCATAC; *IL6* forward: GAGGATACCACTCCCAACAGACC; *IL6* reverse: AAGTGCATCATCGTTGTTTCATACA; *vimentin* forward: TGGTTGACACC CACTCAAAA; *vimentin* reverse: GGTACCTGTCATGCTGAGAA; *fibronectin* forward: AGACTCGAGGCGGAAATTCC; *fibronectin* reverse: CCCTGCGACC CTCAGAAGT; *Zeb2* transgene forward: GATTACAAGGATGACGATGAC; *Zeb2* transgene reverse: CCTCTGGGATTGGCTTGT; *eGFP* forward: CGACAA CCCTACCTGAGCAG; *eGFP* reverse: CTTGTACAGCTCGTCCATGC; *A20/Thnf3* forward: AAACCAATGGTGATGGAAGACTG; *A20/Thnf3* reverse: GTTGTCCCATTCGTCATTCC; *CD44* forward: TCTGCCATCTAGCACTA AGAGC; *CD44* reverse: GTCTGGGTATTGAAAGGTGTAGC; *F11R* forward: ACCTTCTCATCCAGTGGCAT; *F11R* reverse: CAGAGGAGGGGACACTGATC; *Cldn4* Forward: ACCCACCACCTACCCTACTA; *Cldn4* reverse: TCCCCAG CCCTCCCCAAACCA; *Cldn6* forward: TCCGAATATCCCAACCAAGAT; *Cldn6* reverse: AGACAACCCCAAACTACTCC; *Cldn7* forward: AGGGTCTGC TCTGGTCTCT; *Cldn7* reverse: GTACGCAGCTTGTCTTCA; *Ocln* forward: CTGTGAAAACCCGAAGAAAGA; *Ocln* reverse: ATGTCCAGGCTCCCAAG ATAA; *Tjp1* forward: TGAATTGCAATCTCTGGTG; *Tjp1* reverse: TTTTCC TGAGCTGTCTCTCA; *Dsc1* forward: CCTGGTCTGTGTTATTTTGT; *Dsc1* reverse: CCAGATTCATTTGCCTACAG; *Dsc3* forward: TGGGTCTTCAG TTTTGTGTGG; *Dsc3* reverse: CGATGATCTTGTCTGCCTCTA; *Dsg2* forward: CCTTTCGGCATATTCTGCTCTT; *Dsg2* reverse: AGTCCAATGCATAGCCT GTCA; *Pkp1* forward: GTCAAACGGCAGAAAGTCC; and *Pkp1* reverse: TGGTGCCGTAGTTGTGTGAAT. In addition, the expression of 204 immune-related genes was analyzed by quantitative PCR using gene-specific primers. Expression values were measured in triplicates on a Roche LightCycler 480 and normalized to *Cox4i1*, *Matr3* and *Sdhα*. The results are expressed as the log₂[fold-change] expression compared with respective control genotypes.

Quantification of cytokines. Mouse blood was collected postmortem by cardiac puncture. Serum was isolated by 15 min cold centrifugation at maximum speed and stored at –80 °C. Mouse tissue was collected postmortem and snap-frozen. Tissue was lysed on a TissueLyser II (Qiagen) and proteins were isolated in NP-40 lysis buffer with protease inhibitors. Cytokines were quantified using the Bio-Plex Pro kit (Bio-Rad) on the Bio-Plex 200 system (Bio-Rad).

Flow cytometry analysis of colonic leukocytes. Lamina propria isolation from colonic samples of specific pathogen-free and germ-free Zeb2^{IEC-Tg} mice and their wild-type littermates was performed as described previously⁴¹. The isolated cells were used for extracellular and intracellular staining for representative markers of T and myeloid cells, in two separated panels (1 × 10⁶ cells per sample), then analyzed by flow cytometry using a five-laser – BD LSRFortessa. For the analysis of the T cell compartment, the cells were stained with 7-AAD (1/1,000) for live/dead separation and then extracellularly with the following antibodies: anti-CD3 APC (eBioscience; 17-0031-83; 1/100), anti-CD4 APC–Cy7 (BD; 552051; 1/200) and anti-CD8 V500 (BD; 560776; 1/100). After the extracellular staining, the cells were fixed/permeabilized using the Foxp3 Transcription Factor Staining Buffer Set (eBioscience; 00-5523-00). Finally, the cells were stained intracellularly with anti-Foxp3 Alexa Fluor 488 (eBioscience; 53-5773-82; 1/100), anti-T-bet PE–Cy7 (eBioscience; 25-5825-82; 1/100), anti-RORγt BV421 (BD; 562894; 1/100) and anti-GATA3 PE (eBioscience; 12-9966-42; 1/100). For the analysis of the myeloid compartment, the cells were stained with Fixable Viability Dye eFluor 506 (eBioscience; 65-0866-14; 1/300) for live/dead separation and then only extracellularly with the following antibodies: lineage (CD19 (eBioscience; 15-0193-82; 1/400), CD3 (eBioscience; 15-0031-82; 1/200) and NK1.1 (BioLegend; 108716; 1/200) PE–Cy5), anti-CD45 Alexa Fluor 700 (eBioscience; 56-0451-82; 1/800),

anti-Ly6G PercpCy5.5 (BD; 560602; 1/200), anti-Ly6C-APC (eBioscience; 17-5932-80; 1/200), or anti-GR-1 where mentioned (eBioscience; 45-5931-80; 1/200), anti-Siglec F BUV395 (BD; 740280; 1/200), anti-CD11b BV605 (BD; 563015; 1/600), anti-CD64 BV711 (BioLegend; 139311; 1/100), anti-F4/80 Biotin (eBioscience; 13-4801-82; 1/100), Streptavidin BV421 (BioLegend; 405226; 1/1000), anti-CD11c PE-eFluor 610 (eBioscience; 61-0114-82; 1/300) and anti-MHC Class II APC-eFluor 780 (eBioscience; 47-5321-80; 1/800). The flow cytometry data were analyzed using FlowJo software and a sequential gating strategy. The graphs showing the percentages of myeloid cells represent proportions of every parent population. In the analysis of T cells, CD8⁺ are presented as a percentage of the CD3⁺ parent population, and CD4⁺T-bet⁺ and CD4⁺Foxp3⁺ are presented as percentages of the CD4⁺ population. Regarding the t-SNE plots, 5,000 cells per sample were exported from six mice per genotype. The export was on the CD3⁺ lineage for the T cells and on the CD45⁺ lineage (excluding CD19⁺, Nk1.1⁺ and CD3⁺ fraction) for the myeloid cells. The samples were then concatenated and then analyzed with the t-SNE plugin of FlowJo (10.6.1) (iterations = 1,000; perplexity = 30), so the representative plots show a total of 30,000 cells. Following dimensional reduction, coordinates for each t-SNE dimension (i.e., t-SNE1 and t-SNE2) in the two-dimensional plots were determined and integrated as novel parameters.

Flow cytometry analysis for assessment of epithelial cell proliferation. Colonic and ileal epithelial cells were isolated according to the following protocol. The colon and ileum were flushed with PBS and cut into small pieces, followed by incubation in dissociation solution (DMEM supplemented with 10% fetal calf serum, 0.20% collagenase and 0.125 μ l ml⁻¹ dispase II) for 40 min, on a shaker at 37°C. To obtain single cells, the digested tissues were passed through a 100- μ m filter, followed by a wash with DMEM + 10% fetal calf serum. The pellet was resuspended in 30% Percoll and centrifuged for 5 min at 400g. Red blood cells were removed by incubation with ACK lysis buffer for 1 min at room temperature. Finally, cells were resuspended in PBS, and 1 $\times$ 10⁶ epithelial cells per sample were stained with the antibodies eFluor780 Live/Dead staining (eBioscience; 65-0865-14), EpCam PE-Cy7 (G8.8; eBioscience; 25-5791-80) and Ki67 eFluor660 (Sola15; eBioscience; 50-5698-82), and analyzed with a BD five-laser LSRFortessa flow cytometer.

Microbiome analysis. A total of 55 fecal samples were collected from 25-week-old wild-type and transgenic mice (34 and 21 mice, respectively). Mucosal colonic tissues were also collected from the same mice near the tumor area. Fresh feces and mucosal tissues were snap-frozen and immediately stored at -80°C until use. DNA was extracted using an RNeasy PowerMicrobiome kit (Qiagen), as instructed by the manufacturer, but with modifications, including the addition of 10 min at 90°C before vortexing and the exclusion of the DNase I step. No more than ten samples were prepared at once, and each set of extractions included a negative control (blank). To amplify the V4 region of the 16S ribosomal RNA gene, the 341F and 805R primers were used (CCTACGGGNGGCWGCAG⁶² and GACTACHVGGGTATCTAATCC⁶³, respectively). To amplify the *ITS2* gene (only in feces), the gITS7 and ITS4 primers were used (GTGAATCATCGARTCTTTG⁶⁴ and TCCTCCGCTTATTGATATGC, respectively). Amplification, indexing and Illumina MiSeq sequencing (2 $\times$ 250 base pairs (bp)) have been performed by Biofidal, according to the manufacturer's protocol. Overlapping paired-end reads were merged and processed with a LotuS pipeline⁶⁵, with clustering of sequences into operational taxonomic units using the Greengenes (gg_13.5) database for taxonomy assignment. Further analysis was performed using Phyloseq⁶⁶. Except for the α -diversity analysis, operational taxonomic units with an absolute abundance of <0.01% were removed from the analysis. Graphs were done with ggplot2. The LefSe algorithm (<http://huttenhower.sph.harvard.edu/galaxy>) was performed in combination with a multivariate analysis by linear models (MaAsLin) and a *t*-test with false discovery rate correction.

RNA sequencing analysis of the DLD-1 CRC cell line. DLD-1 cells were lentivirally transduced with a construct allowing doxycycline-inducible expression of the EMT transcription factor ZEB2. Cells were stimulated with doxycycline (1.5 μ g ml⁻¹) in RPMI growth medium supplemented with 10% fetal calf serum, sodium pyruvate and non-essential amino acids. Total RNA was extracted from DLD-1 cells using the NucleoSpin RNA Plus kit (Machery-Nagel) according to the manufacturer's instructions. The integrity of the samples was confirmed using an Agilent Bioanalyzer 2100. The mRNA library was prepared with the QuantSeq 3' mRNA FWD kit (Lexogen) followed by sequencing with the NextSeq 500 high-output kit (Illumina).

Orthotopic transplantation experiment. Single-cell solutions derived from tumor tissue of the Zeb2^{IEC-Tg/+} mice (10⁶ cells) were orthotopically inoculated in the cecum of NSG mice (The Jackson Laboratory), together with 2 $\times$ 10⁶ cells that were injected subcutaneously in the same mice. These injections were done in a mixture of 50%:50% Matrigel:serum-free medium. Mice were monitored on a weekly basis by measuring the body weight, and twice a week by palpation, for a subcutaneous tumor.

Shallow sequencing and hot-spot mutation screening. Tissues from 32-week-old R26-Zeb2^{Tg/+} (*n* = 4) and wild-type control littermates (*n* = 2) were snap-frozen

in liquid nitrogen and stored at -80°C until genomic DNA extraction. Genomic DNA was isolated using a QIAamp DNA Mini Kit (Qiagen) according to the manufacturer's instructions. Briefly, 25 mg tissue was homogenized in a TissueLyser II (Qiagen) for 2 min at 15 Hz, followed by digestion with Protease K at 56°C and 1,100 r.p.m. for 3 h. RNA was digested with RNase A (1 mg ml⁻¹; Invitrogen) and genomic DNA was purified on manufacturer's columns. The DNA concentration and quality were measured on a NanoDrop (Thermo Fisher Scientific) and stored at -20°C to avoid acidic hydrolysis. DNA from Zeb2 tumors was compared with the corresponding colon region of a healthy wild-type littermate, and internally with DNA isolated from the lungs. Shallow-depth whole-genome sequencing of tumor and control DNA was performed using an Illumina HiSeq 3000 instrument, starting from 200 ng input DNA. Tumor DNA was sheared to 200 bp with a Covaris Ultrasonicator. Afterwards, library preparation was performed with a NEXTFlex Rapid DNA Sequencing Kit (BioScientific), automated on a Hamilton STAR robot. Libraries were pooled to equimolar concentrations and samples were pooled per lane on the HiSeq 3000. Data analysis and visualization were performed with the online analysis suite Vivar⁶⁷. Each genome profile (line view and chromosome view) was manually checked for aberrations. In parallel, a hot-spot mutation analysis was performed. For this, the most frequently observed mutations in patients with CRC were identified using the Colorectal Cancer Atlas. Primers were designed spanning a region of approximately 1,000 bp around the mutation in the *Apc*, *Braf*, *Ctnnb1*, *Kras*, *Nras*, *p53*, *Pik3ca* and *Pten* genes. Genomic DNA was amplified with the Phusion High-Fidelity DNA Polymerase kit (Thermo Fisher Scientific) according to the manufacturer's instructions, using the gene-specific primers. After purification with the QIAquick PCR purification kit (Qiagen), the amplified region of interest was sequenced and analyzed for the presence of gene mutations. The following primers were used: *Apc* forward: GCTTCTGGCTTATCTTCA; *Apc* reverse: GCTGGCAGAAGTTTATACA; *Braf* forward: CCCAGCAGAGTTAAAGTGTA; *Braf* reverse: GCTTCTACCTTCTGAGTG; *Ctnnb1* forward: GCCTTTGTTC TCTTCCCT; *Ctnnb1* reverse: CGCGTCATCCTGATAGTTA; *Kras* forward: GGGTAGGTGTTGGGATAG; *Kras* reverse: GGACTTCTTTCTCTGTGG; *Nras* forward: GCATCATCTGCAAGGAA; *Nras* reverse: GCACGAACCTTGTT TACTG; *p53* forward: GCCGAACAGGTGGAATAT; *p53* reverse: GGAAGGAGAGAGCAAGAG; *Pik3ca* forward: GGGAGTCACTGTTCTACTT; *Pik3ca* reverse: GCTCACAACCTAAGCTGGA; *Pten* forward: GCAGTTCACAA AAGTCCCA; and *Pten* reverse: GCCAACAACTAAGGGTC.

Human CRC datasets and survival analysis. The human CRC expression profiling dataset is a multi-cohort collection similar to what was used by Guinney et al.³⁵, with additional datasets from the National Center for Biotechnology Information Gene Expression Omnibus (GSE41258, GSE24551 and GSE87211) and ArrayExpress (E-MTAB-863). Expression datasets were pooled after cohort-specific centering and batch correction using naive RUV2 (ref. ⁶⁸). Hierarchical clustering was performed using average linkage of Pearson's correlation. CMS labels were generated using the published CMS Classifier³⁵.

Human tissue samples. For immunohistochemistry, colonic biopsies were obtained from patients with CRC who had given written informed consent to participate in the Institutional Review Board-approved Biobank of University Hospitals Leuven, Belgium (S50887-ML4707).

Immunofluorescence with tyramide signal amplification. The following antibodies were used to detect the respective proteins: anti-Zeb2 (rabbit, 1:900)⁶⁹ and anti-panCK (mouse; 1:300; DAKO; M3515; clone AE1/AE3). The PerkinElmer Opal 4-Color Manual IHC Kit (PerkinElmer/Akoya; NEL810001KT) was used for the tyramide signal amplification according to the manufacturer's protocol. For introduction of the secondary HRP, the Envision+/HRP goat anti-Rabbit (Dako Envision+ Single Reagents; HRP; Rabbit; code K4003) was used for antibody raised in rabbit (Zeb2) and the OPAL Polymer HRP Ms+Rb (Akoya/Perkin Elmer; ARH1001EA) was used for the antibody raised in mouse (panCK). The various proteins were detected using the OPAL 570 (Zeb2) and OPAL 690 (panCK) reagents. Images were acquired on the Zeiss Axio Scan.Z1 using a $\times$ 20 objective and ZEN 2 software.

Statistics and reproducibility. All repeated independent experiments showed similar results. No statistical method was used to predetermine sample sizes, and no data were excluded from the analyses. Data were analyzed using GraphPad Prism 7 software. The results are expressed as means $\pm$ s.e.m. For survival data analysis, Kaplan-Meier estimates of the survivor function for both the control and transgenic mice were calculated using survival time data of individual mice. Differences in survival distribution between groups were tested using the log-rank Mantel-Cox test. Bio-plex, quantitative PCR and colon weight data were analyzed using unpaired, non-parametric Mann-Whitney *U*-tests. Body weight, stool score and MEICS data over time were analyzed as repeated measurements using the residual maximum likelihood as implemented in Genstat version 19. Briefly, a linear mixed model of the form response = μ + genotype + time + (genotype $\times$ time) + (subject $\times$ time) was fitted to the repeated measurements. The term subject $\times$ time represents the residual error term with dependent errors because the repeated measurements are taken

in the same individual, causing correlations among observations. Times of measurement were set as equally spaced, and the best model for correlation on repeated measurements (for example, uniform, autoregressive, antedependence or unstructured) was selected based on the likelihood ratio test statistic and the Akaike information criterion. The significance of genotype effects over time (that is, genotype $\times$ time) and changes in differences between genotype effects over time were assessed using an approximate *F*-test. For analysis of human colonic biopsies, analyses were carried out using R version 3.5.0 (R Development Core Team), and differential gene expression was assessed using the DESeq2 package.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Statistical source data for Figs. 1–7 and Extended Data Figs. 1, 3 and 5–10 are provided with the paper. RNA sequencing data on ZEB2-overexpressing DLD-1 cells have been deposited in the Gene Expression Omnibus under accession code [GSE148823](#). Sequencing data from the shallow whole-genome sequencing on Zeb2^{IEC-Tg/+} and control colon tumor tissue have been deposited in the Sequence Read Archive database under accession code [PRJNA626401](#). All other data supporting the findings of this study are available from the corresponding author on reasonable request.

Received: 4 March 2020; Accepted: 28 April 2020;

Published online: 15 June 2020

References

- Brenner, H., Kloor, M. & Pox, C. P. Colorectal cancer. *Lancet* **383**, 1490–1502 (2014).
- Bray, F. et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **68**, 394–424 (2018).
- Vogelstein, B. et al. Cancer genome landscapes. *Science* **339**, 1546–1558 (2013).
- O’Keefe, S. J. Diet, microorganisms and their metabolites, and colon cancer. *Nat. Rev. Gastroenterol. Hepatol.* **13**, 691–706 (2016).
- Bashir, A., Miskeen, A. Y., Hazari, Y. M., Asrafuzzaman, S. & Fazili, K. M. *Fusobacterium nucleatum*, inflammation, and immunity: the fire within human gut. *Tumour Biol.* **37**, 2805–2810 (2016).
- Brennan, C. A. & Garrett, W. S. *Fusobacterium nucleatum*—symbiont, opportunist and oncobacterium. *Nat. Rev. Microbiol.* **17**, 156–166 (2019).
- West, N. R., McCuaig, S., Franchini, F. & Powrie, F. Emerging cytokine networks in colorectal cancer. *Nat. Rev. Immunol.* **15**, 615–629 (2015).
- Jin, C. et al. Commensal microbiota promote lung cancer development via $\gamma\delta$ T cells. *Cell* **176**, 998–1013.e16 (2019).
- De Craene, B. & Berrx, G. Regulatory networks defining EMT during cancer initiation and progression. *Nat. Rev. Cancer* **13**, 97–110 (2013).
- Thiery, J. P., Acloque, H., Huang, R. Y. & Nieto, M. A. Epithelial–mesenchymal transitions in development and disease. *Cell* **139**, 871–890 (2009).
- Nieto, M. A., Huang, R. Y., Jackson, R. A. & Thiery, J. P. EMT: 2016. *Cell* **166**, 21–45 (2016).
- Stemmler, M. P., Eccles, R. L., Brabletz, S. & Brabletz, T. Non-redundant functions of EMT transcription factors. *Nat. Cell Biol.* **21**, 102–112 (2019).
- Goossens, S., Vandamme, N., Van Vlierberghe, P. & Berrx, G. EMT transcription factors in cancer development re-evaluated: beyond EMT and MET. *Biochim. Biophys. Acta Rev. Cancer* **1868**, 584–591 (2017).
- Skrypek, N., Goossens, S., De Smedt, E., Vandamme, N. & Berrx, G. Epithelial-to-mesenchymal transition: epigenetic reprogramming driving cellular plasticity. *Trends Genet.* **33**, 943–959 (2017).
- Kahlert, C. et al. Overexpression of ZEB2 at the invasion front of colorectal cancer is an independent prognostic marker and regulates tumor invasion in vitro. *Clin. Cancer Res.* **17**, 7654–7663 (2011).
- Li, M. Z. et al. ZEB2 promotes tumor metastasis and correlates with poor prognosis of human colorectal cancer. *Am. J. Transl. Res.* **9**, 2838–2851 (2017).
- Rokavec, M. et al. IL-6R/STAT3/miR-34a feedback loop promotes EMT-mediated colorectal cancer invasion and metastasis. *J. Clin. Invest.* **124**, 1853–1867 (2014).
- Sreekumar, R. et al. Assessment of nuclear ZEB2 as a biomarker for colorectal cancer outcome and TNM risk stratification. *JAMA Netw. Open* **1**, e183115 (2018).
- Tatari, M. N. et al. ZEB2-transgene expression in the epidermis compromises the integrity of the epidermal barrier through the repression of different tight junction proteins. *Cell. Mol. Life Sci.* **71**, 3599–3609 (2014).
- Goossens, S. et al. ZEB2 drives immature T-cell lymphoblastic leukaemia development via enhanced tumour-initiating potential and IL-7 receptor signalling. *Nat. Commun.* **6**, 5794 (2015).
- Madison, B. B. et al. *Cis* elements of the villin gene control expression in restricted domains of the vertical (crypt) and horizontal (duodenum, cecum) axes of the intestine. *J. Biol. Chem.* **277**, 33275–33283 (2002).
- Barker, N., van de Wetering, M. & Clevers, H. The intestinal stem cell. *Genes Dev.* **22**, 1856–1864 (2008).
- Bollrath, J. & Greten, F. R. IKK/NF- κ B and STAT3 pathways: central signalling hubs in inflammation-mediated tumour promotion and metastasis. *EMBO Rep.* **10**, 1314–1319 (2009).
- Zeilstra, J. et al. CD44 expression in intestinal epithelium and colorectal cancer is independent of p53 status. *PLoS ONE* **8**, e72849 (2013).
- Dalerba, P. et al. Phenotypic characterization of human colorectal cancer stem cells. *Proc. Natl Acad. Sci. USA* **104**, 10158–10163 (2007).
- Merlos-Suarez, A. et al. The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse. *Cell Stem Cell* **8**, 511–524 (2011).
- Clevers, H. At the crossroads of inflammation and cancer. *Cell* **118**, 671–674 (2004).
- Taniguchi, K. & Karin, M. NF- κ B, inflammation, immunity and cancer: coming of age. *Nat. Rev. Immunol.* **18**, 309–324 (2018).
- Asquith, M. & Powrie, F. An innately dangerous balancing act: intestinal homeostasis, inflammation, and colitis-associated cancer. *J. Exp. Med.* **207**, 1573–1577 (2010).
- Johansson, M. E. et al. The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proc. Natl Acad. Sci. USA* **105**, 15064–15069 (2008).
- De Filippo, K. et al. Mast cell and macrophage chemokines CXCL1/CXCL2 control the early stage of neutrophil recruitment during tissue inflammation. *Blood* **121**, 4930–4937 (2013).
- Germann, M. et al. Neutrophils suppress tumor-infiltrating T cells in colon cancer via matrix metalloproteinase-mediated activation of TGF β . *EMBO Mol. Med.* **12**, e10681 (2020).
- Jackstadt, R. et al. Epithelial NOTCH signaling rewires the tumor microenvironment of colorectal cancer to drive poor-prognosis subtypes and metastasis. *Cancer Cell* **36**, 319–336.e7 (2019).
- Grivennikov, S. I. et al. Adenoma-linked barrier defects and microbial products drive IL-23/IL-17-mediated tumour growth. *Nature* **491**, 254–258 (2012).
- Guinney, J. et al. The consensus molecular subtypes of colorectal cancer. *Nat. Med.* **21**, 1350–1356 (2015).
- Marisa, L. et al. Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Med.* **10**, e1001453 (2013).
- Laibe, S. et al. A seven-gene signature aggregates a subgroup of stage II colon cancers with stage III. *OMICS* **16**, 560–565 (2012).
- De Sousa, E. M. F. et al. Methylation of cancer-stem-cell-associated Wnt target genes predicts poor prognosis in colorectal cancer patients. *Cell Stem Cell* **9**, 476–485 (2011).
- Smith, J. J. et al. Experimentally derived metastasis gene expression profile predicts recurrence and death in patients with colon cancer. *Gastroenterology* **138**, 958–968 (2010).
- Budinska, E. et al. Gene expression patterns unveil a new level of molecular heterogeneity in colorectal cancer. *J. Pathol.* **231**, 63–76 (2013).
- Li, H. et al. Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nat. Genet.* **49**, 708–718 (2017).
- Calon, A. et al. Stromal gene expression defines poor-prognosis subtypes in colorectal cancer. *Nat. Genet.* **47**, 320–329 (2015).
- Isella, C. et al. Stromal contribution to the colorectal cancer transcriptome. *Nat. Genet.* **47**, 312–319 (2015).
- Irrazabal, T., Belcheva, A., Girardin, S. E., Martin, A. & Philpott, D. J. The multifaceted role of the intestinal microbiota in colon cancer. *Mol. Cell* **54**, 309–320 (2014).
- Nakatsu, G. et al. Gut mucosal microbiome across stages of colorectal carcinogenesis. *Nat. Commun.* **6**, 8727 (2015).
- Bullman, S. et al. Analysis of *Fusobacterium* persistence and antibiotic response in colorectal cancer. *Science* **358**, 1443–1448 (2017).
- Gopalakrishnan, V. et al. Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science* **359**, 97–103 (2018).
- Routy, B. et al. Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science* **359**, 91–97 (2018).
- Vetizou, M. et al. Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. *Science* **350**, 1079–1084 (2015).
- Viaud, S. et al. The intestinal microbiota modulates the anticancer immune effects of cyclophosphamide. *Science* **342**, 971–976 (2013).
- Matson, V. et al. The commensal microbiome is associated with anti-PD-1 efficacy in metastatic melanoma patients. *Science* **359**, 104–108 (2018).
- Iida, N. et al. Commensal bacteria control cancer response to therapy by modulating the tumor microenvironment. *Science* **342**, 967–970 (2013).
- Nenci, A. et al. Epithelial NEMO links innate immunity to chronic intestinal inflammation. *Nature* **446**, 557–561 (2007).
- Welz, P. S. et al. FADD prevents RIP3-mediated epithelial cell necrosis and chronic intestinal inflammation. *Nature* **477**, 330–334 (2011).

55. Takahashi, N. et al. RIPK1 ensures intestinal homeostasis by protecting the epithelium against apoptosis. *Nature* **513**, 95–99 (2014).
56. Kaser, A. et al. XBP1 links ER stress to intestinal inflammation and confers genetic risk for human inflammatory bowel disease. *Cell* **134**, 743–756 (2008).
57. Dannappel, M. et al. RIPK1 maintains epithelial homeostasis by inhibiting apoptosis and necroptosis. *Nature* **513**, 90–94 (2014).
58. Gunther, C. et al. Caspase-8 regulates TNF- α -induced epithelial necroptosis and terminal ileitis. *Nature* **477**, 335–339 (2011).
59. Becker, C. et al. In vivo imaging of colitis and colon cancer development in mice using high resolution chromoendoscopy. *Gut* **54**, 950–954 (2005).
60. Johansson, M. E. & Hansson, G. C. Preservation of mucus in histological sections, immunostaining of mucins in fixed tissue, and localization of bacteria with FISH. *Methods Mol. Biol.* **842**, 229–235 (2012).
61. Scott, C. L. et al. The transcription factor ZEB2 is required to maintain the tissue-specific identities of macrophages. *Immunity* **49**, 312–325.e5 (2018).
62. Muyzer, G., de Waal, E. C. & Uitterlinden, A. G. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl. Environ. Microbiol.* **59**, 695–700 (1993).
63. Herlemann, D. P. et al. Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J.* **5**, 1571–1579 (2011).
64. Ihrmark, K. et al. New primers to amplify the fungal ITS2 region—evaluation by 454-sequencing of artificial and natural communities. *FEMS Microbiol. Ecol.* **82**, 666–677 (2012).
65. Hildebrand, F., Tadeo, R., Voigt, A. Y., Bork, P. & Raes, J. LotuS: an efficient and user-friendly OTU processing pipeline. *Microbiome* **2**, 30 (2014).
66. McMurdie, P. J. & Holmes, S. PhyloSeq: a bioconductor package for handling and analysis of high-throughput phylogenetic sequence data. *Pac. Symp. Biocomput.* 235–246 (2012).
67. Sante, T. et al. ViVar: a comprehensive platform for the analysis and visualization of structural genomic variation. *PLoS ONE* **9**, e113800 (2014).
68. Gagnon-Bartsch, J. A. & Speed, T. P. Using control genes to correct for unwanted variation in microarray data. *Biostatistics* **13**, 539–552 (2012).
69. Sayan, A. E. et al. SIP1 protein protects cells from DNA damage-induced apoptosis and has independent prognostic value in bladder cancer. *Proc. Natl Acad. Sci. USA* **106**, 14884–14889 (2009).

Acknowledgements

We thank M. Gomez, S. Ganai, J. Kirundi, K. McCoy and A. MacPherson for generating axenic Zeb2^{IEC-Tg/+} mice at the University of Bern, Switzerland. We thank L. Bellen and K. Barbry for animal care and the VIB Flow Core for training, support and access to the instrument park. K.S. was a predoctoral fellow with the Institute for the Promotion of Innovation by Science and Technology (IWT) and was supported by a 'Kom op tegen Kanker' (Stand Up To Cancer) grant from the Flemish Cancer Society. Work in the G.v.L. laboratory is supported by research grants from the FWO, Strategic Basic Research (SBO) program, Geneeskundige Stichting Koningin Elisabeth (GSKE), CBC Banque Prize, Charcot Foundation, Stichting Tegen Kanker, Cancer Research Institute Ghent and Concerted Research Actions (GOA) of Ghent University. The laboratory of G.B. is supported by grants from the FWO, SBO, GOA, Vlaamse Liga Tegen Kanker and Stichting Tegen Kanker. The laboratory of L.V. is supported by grants from the FWO, SBO and Cancer Research Institute Ghent.

Author contributions

K.S., I.P., G. Blancke, E.H., E.D., M.S., H.V., S.J., J.T., N.V. and D.N. performed the experiments. K.S., I.P., E.H., E.D., E.R., J.J.H., A.W., D.N., P.V.V., G.D.H., P.B., E.E., P.W., B.N.L., C.C., S.T., S.G., G. Berx, L.V. and G.v.L. analyzed the data. E.T. and L.B. provided reagents. G. Berx, L.V. and G.v.L. provided ideas and coordinated the project. K.S., G. Berx, L.V. and G.v.L. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s43018-020-0070-2>.

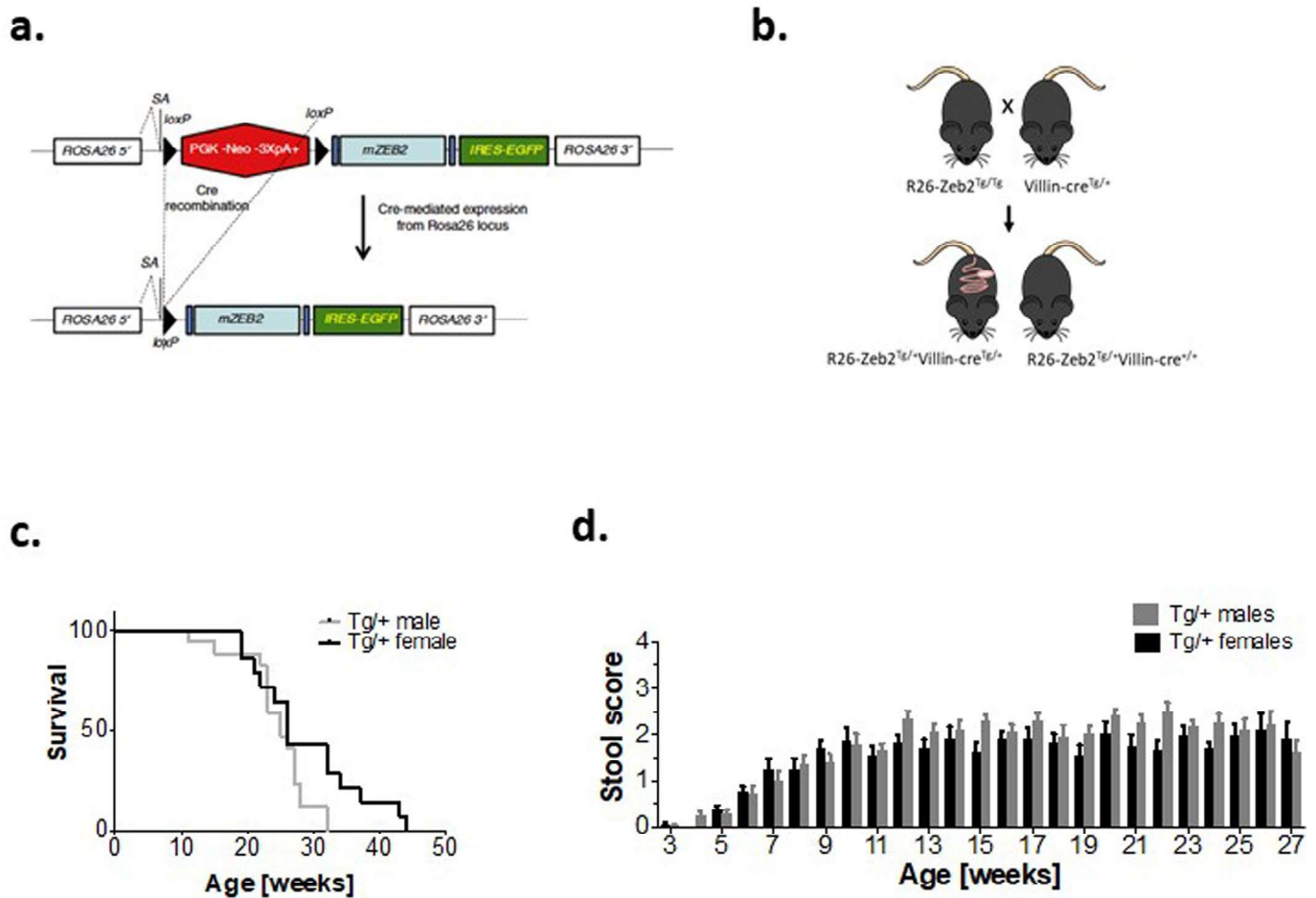
Supplementary information is available for this paper at <https://doi.org/10.1038/s43018-020-0070-2>.

Correspondence and requests for materials should be addressed to G.B., L.V. or G.v.L.

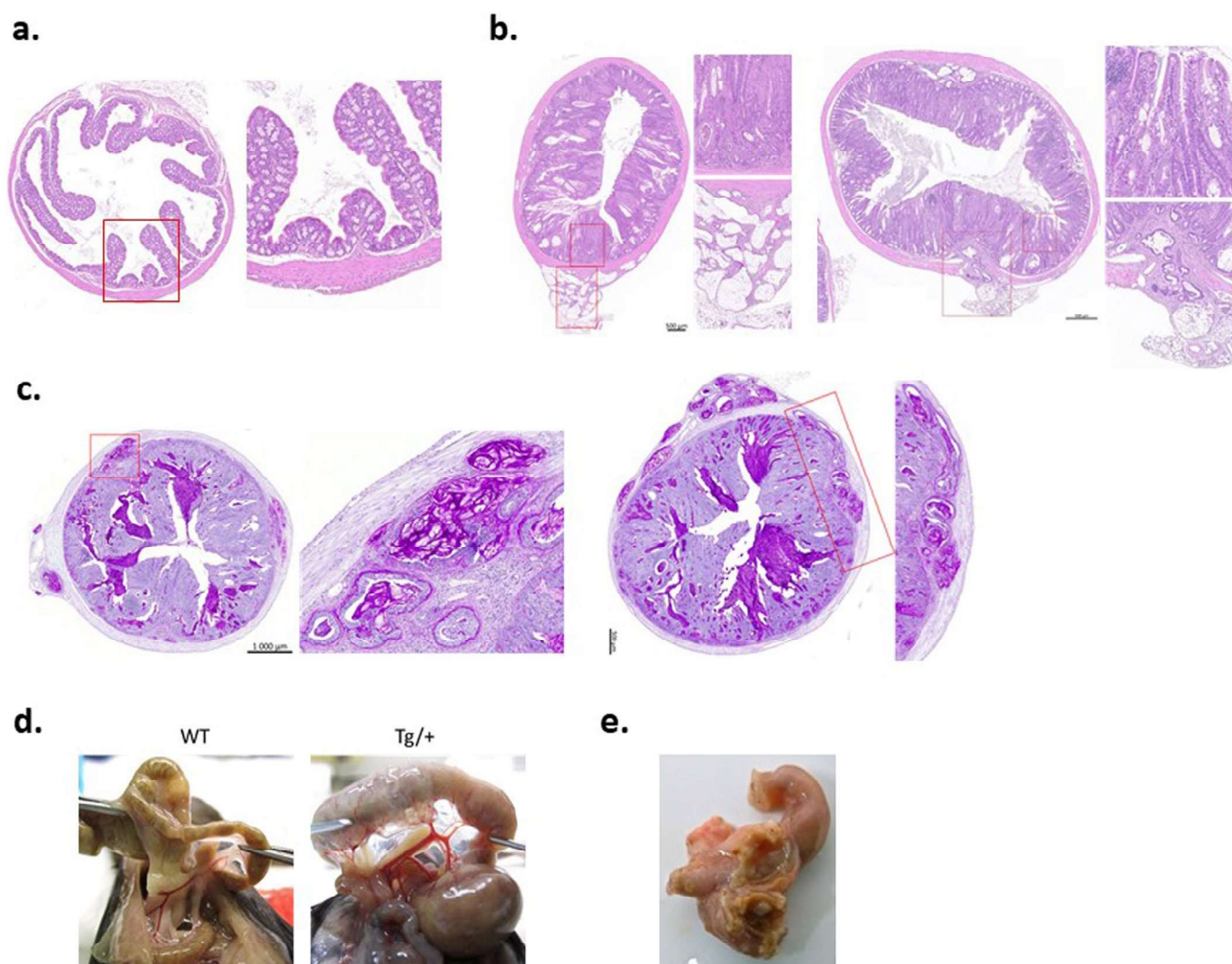
Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

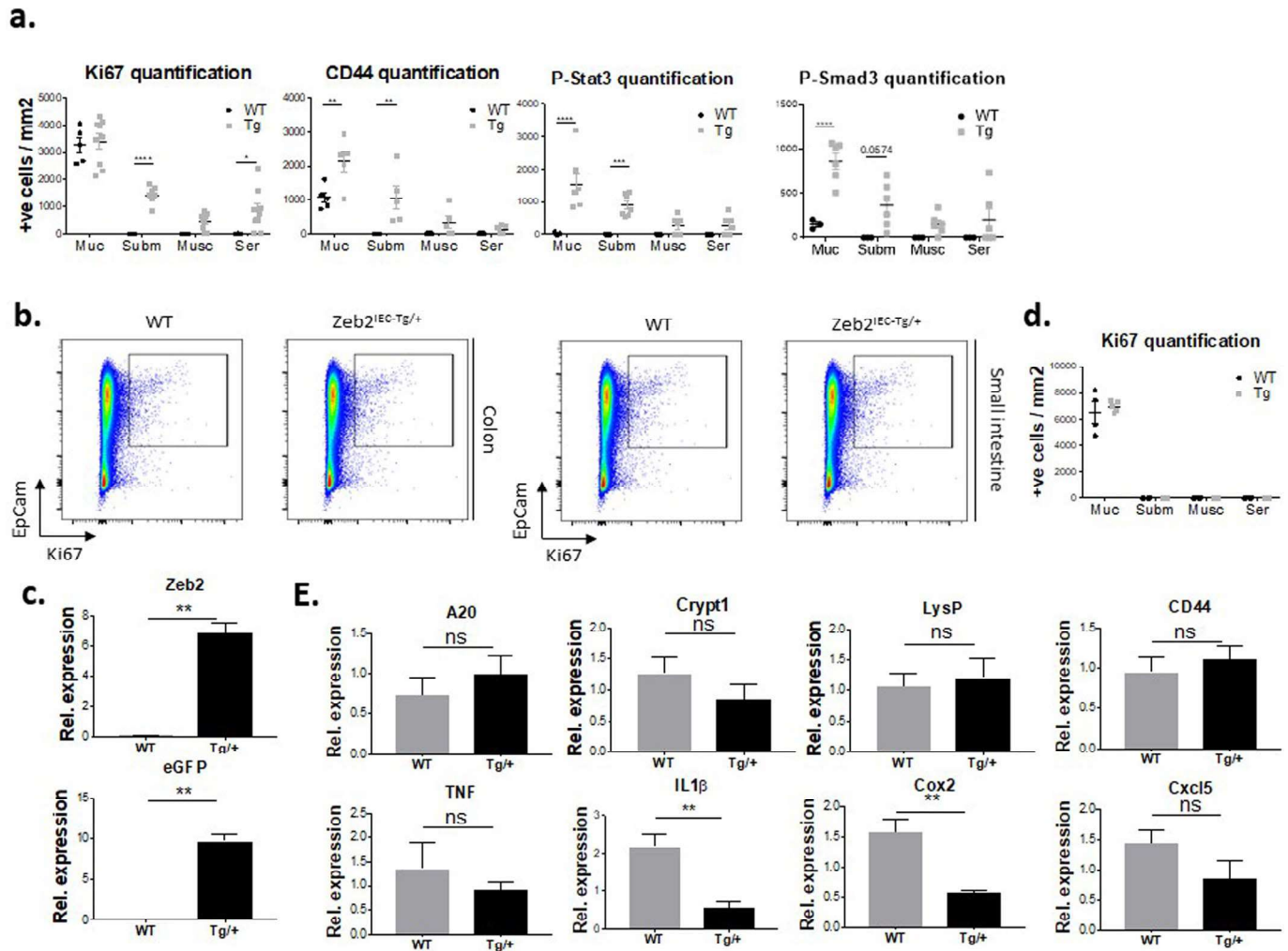
© The Author(s), under exclusive licence to Springer Nature America, Inc. 2020



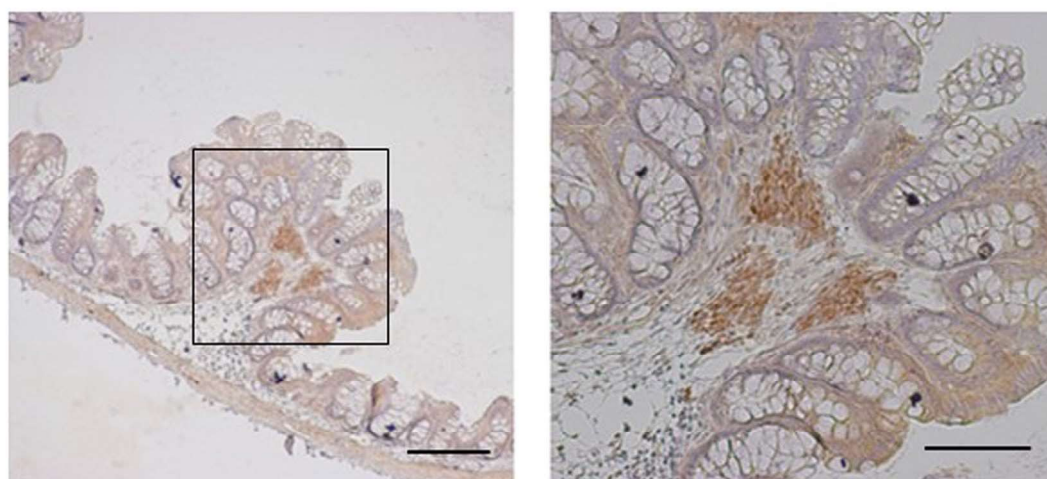
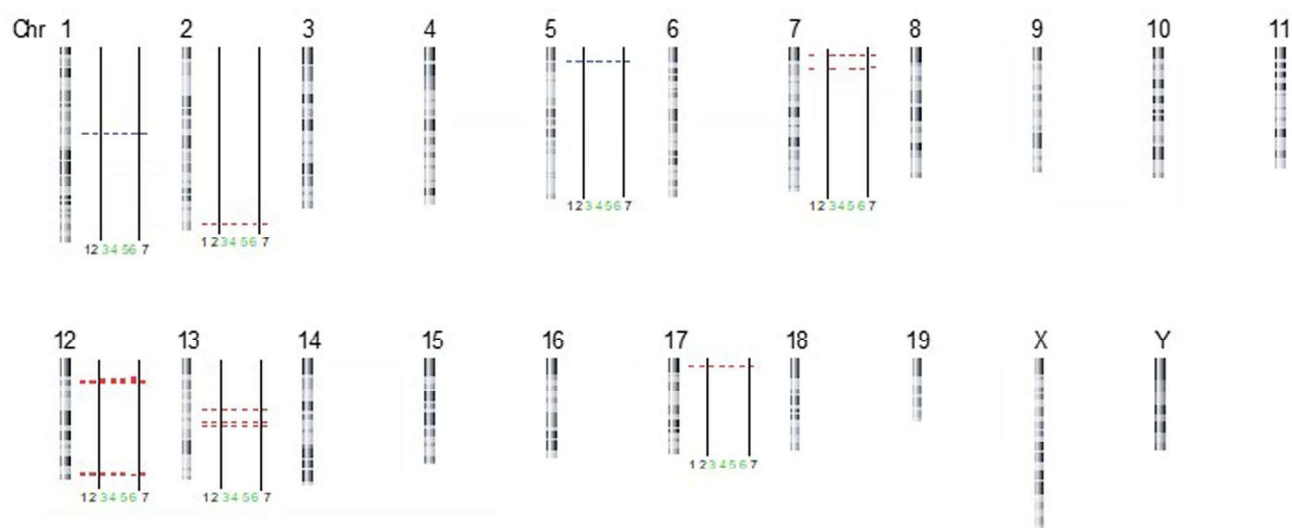
Extended Data Fig. 1 | Expression of Zeb2 in the mouse intestinal epithelium. **a**, Schematic representation of the conditional ROSA26-based Zeb2 overexpression mouse model. The mouse Zeb2 open reading frame was targeted to the ROSA26 locus preceded by a floxed (fl) transcriptional stop (PGK-Neo-3XpA) cassette and followed by IRES-EGFP reporter sequence. **b**, R26-Zeb2^{Tg/Tg} mice were crossed with a Cre transgenic line under the control of Villin promoter, generating R26-Zeb2^{Tg/+} villin-cre^{Tg/+} transgenic (Tg/+) mice and R26-Zeb2^{Tg/+} villin-cre^{+/+} (WT) controls. **c-d**, Kaplan-Meier survival curve (**c**) and clinical stool score (**d**) of male (n=12 mice) and female (n=10 mice) Zeb2^{EC-Tg/+} mice based on Hemocult assay. Data are presented as mean ± SEM.



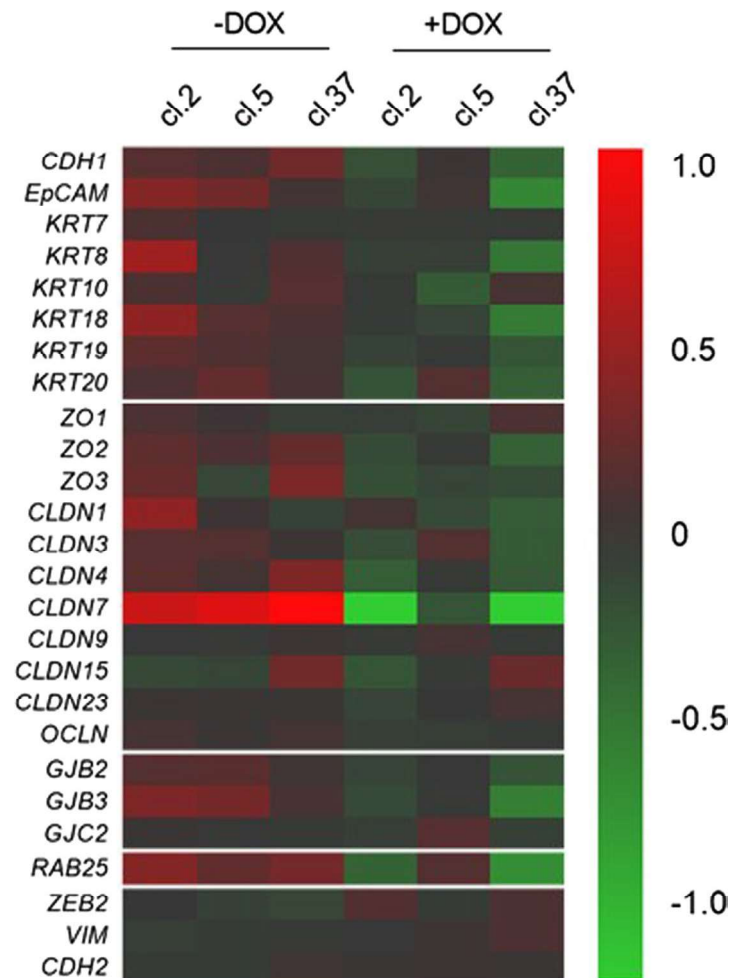
Extended Data Fig. 2 | Histopathology of $Zeb2^{IEC-Tg/+}$ mice. **a**, Representative H&E stained colon section from wild-type control mouse showing normal proximal colon (representative for $n=10$ mice). Left panel, magnification of boxed area. **b**, Representative H&E stained sections of $Zeb2^{IEC-Tg/+}$ mouse (representative for $n=7$ mice) demonstrating multicentric invasive adenocarcinoma in a background of atypical mucosal hyperplasia with multifocal intraepithelial neoplasia and intramucosal carcinoma. Also sparse inflammatory cell infiltrates, predominantly lymphocytes and plasma cells, can be observed in the lamina propria and submucosa. Left panel, higher magnification showing severe atypical hyperplasia/intramucosal carcinoma characterized by tumultuous proliferation of irregular glandular units (upper), and of the invading adenocarcinomatous component characterized by irregular glandular profiles with abundant mucus collection, transmural infiltration of the intestinal wall and spread along the mesenteric ligament (lower). **c**, PAS staining of section of distal colon and rectum of $Zeb2^{IEC-Tg/+}$ mouse (representative for $n=7$ mice). Abundant collection of PAS-positive mucus is evident in the irregular neoplastic glandular units transmurally invading the intestinal wall and infiltrating along the mesenteric ligament. **d**, Macroscopic view on proximal colon tissue dissected in 44-week old WT and $Zeb2^{IEC-Tg/+}$ mice. **e**, Macroscopic dissected proximal colon in 31-week old $Zeb2^{IEC-Tg/+}$ mouse displaying transmurular cancerous outgrowth.



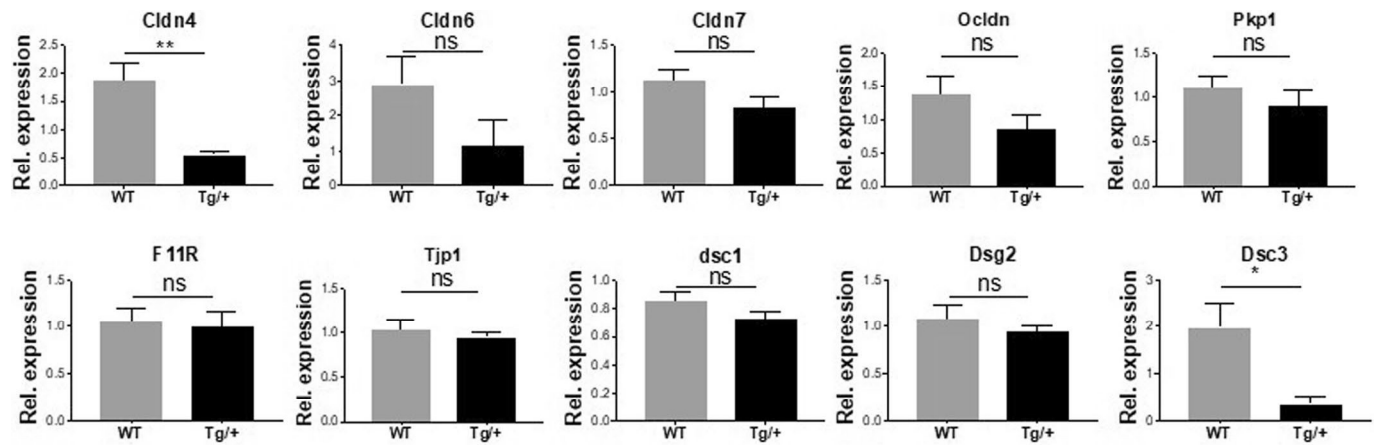
Extended Data Fig. 3 | Zeb2^{IEC-Tg/+} mice develop severe pathology in colon but not in small intestine. **a**, Quantification of Ki67-, CD44-, P-Stat3- and P-Smad3-positive cells in different regions of the colon (mucosa, submucosa, muscle or serosa) of 25-30-week old wild-type (WT) and Zeb2^{IEC-Tg/+} mice (Tg) (Ki67 Tg/+, n=9 mice and WT, n=6 mice; CD44 Tg/+, n=5 mice and WT, n=5 mice; P-Stat3 Tg/+, n=7 mice and WT, n=5 mice; P-Smad3 Tg/+, n=6 mice and WT, n=3 mice). For each tissue region and mouse, 3 separate areas were analyzed and averages per mouse per region were plotted in GraphPad Prism. Data are presented as mean $\pm$ SEM. ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. Two way ANOVA, corrected for multiple comparisons using Sidak test in GraphPad Prism. **b**, Representative FACS plots of Ki67+EPCam+ cells in colon (left panel) and small intestine (right panel) in 5 week old wild-type (WT) and Zeb2^{IEC-Tg/+} mice. The experiment was performed twice with similar results. **c**, Quantitative PCR analysis for Zeb2 and eGFP in lysates from the small intestine of Zeb2^{IEC-Tg/+} (Tg/+, n=6 mice) and wild-type littermate controls (WT, n=5 mice). Data are presented as mean $\pm$ SEM and were analyzed with two-tailed Mann Whitney test, p -value = 0.0043 for both. **d**, Quantification of Ki67-positive cells in different regions of the small intestine of 25-30-week old wild-type (WT, n=4 mice) and Zeb2^{IEC-Tg/+} mice (Tg, n=5 mice). For each tissue region and mouse, 3 separate areas were analyzed and averages per mouse per region were plotted in GraphPad Prism. Data are presented as mean $\pm$ SEM and were analyzed with two-way ANOVA with Sidak correction. **e**, Quantitative PCR analysis for inflammatory markers (TNF, IL-1 β , Cox2, Cxcl5 and A20), CD44, and Paneth cell markers (Crypt-1 and LysP) in lysates from the small intestine of Zeb2^{IEC-Tg/+} (Tg/+, n=6 mice) and wild-type littermate controls (WT, n=5 mice). Data are presented as mean $\pm$ SEM and were analyzed with two-tailed Mann Whitney test, p -value = 0.0087 for IL-1 β and 0.0043 for Cox2; ns, non-significant.

a.**b.**

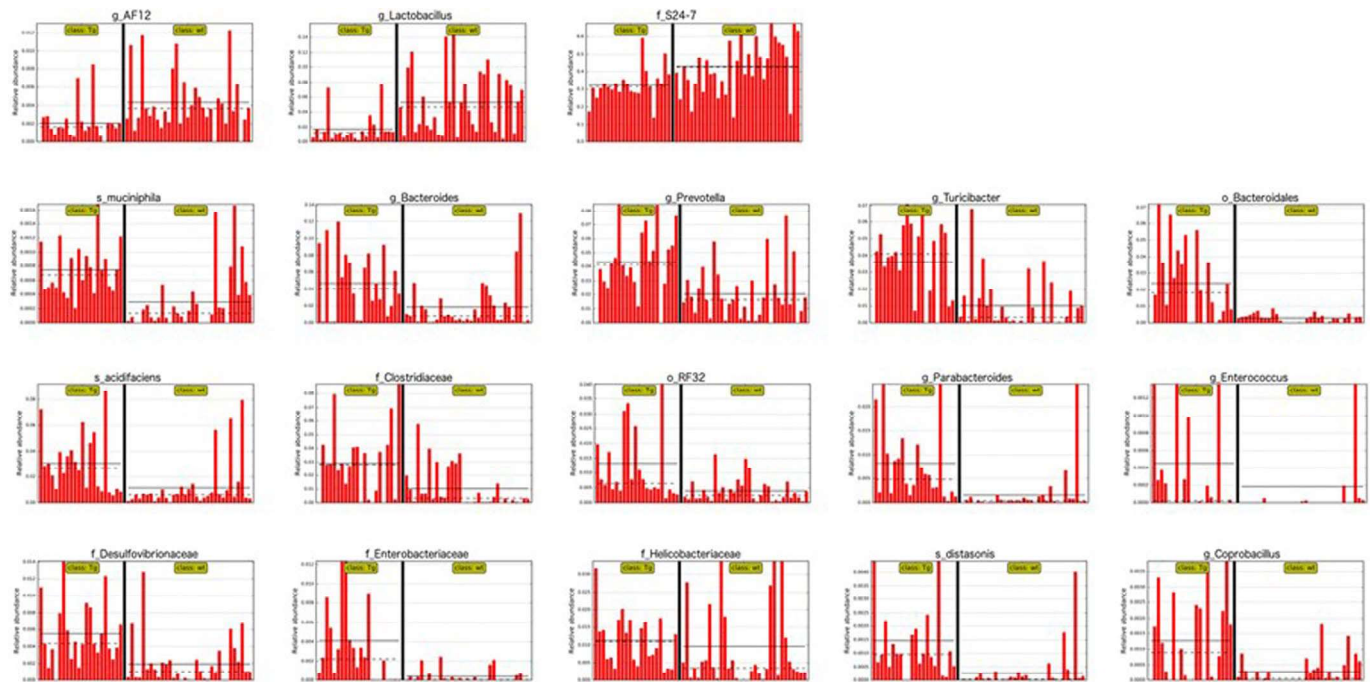
Extended Data Fig. 4 | Tumorigenic potential and copy number profiling of Zeb2 transgenic tissue. a, Orthotopic transplantation of intestinal tumor cells from Zeb2 transgenic mice into the cecum of NSG mice. Zeb2 transgenic cells could be identified by immunohistochemical staining for the GFP transgene. Scale bar: 200µm; boxed area Scale bar: 100µm. Data representative for one experiment on 5 NSG mice. **b,** Copy number analysis by shallow whole genome sequencing of Zeb2^{IEC-Tg/+} colon tumor tissue (n=4 mice, lanes 3-6), normal lung tissue (n=2 mice, lanes 1-2) and normal colon tissue from a wild-type littermate control mouse (lane 7). Genomic aberrations are positioned along the murine chromosomes and visualized in red for deletions and blue for amplifications.



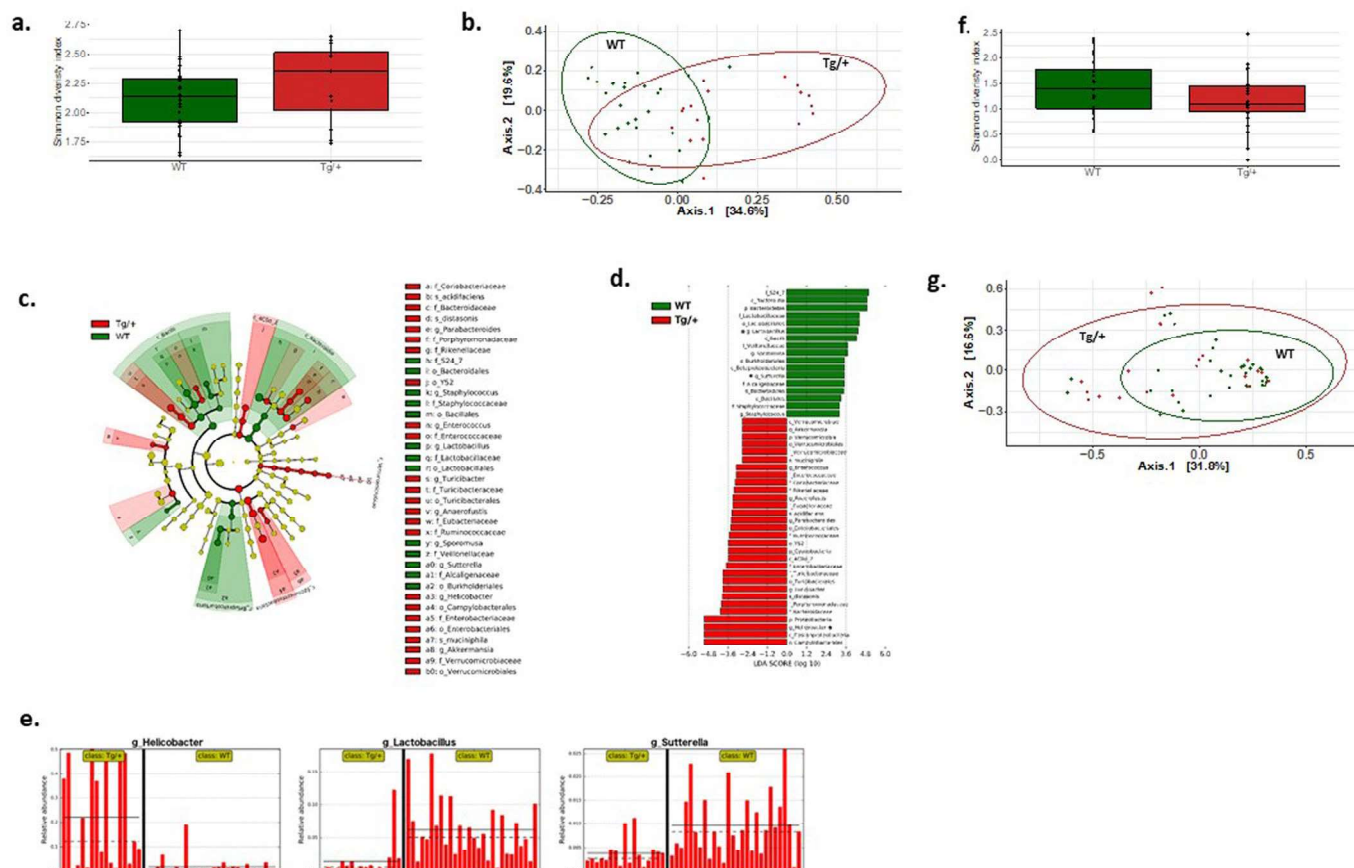
Extended Data Fig. 5 | Induction of ZEB2 expression leads to a loss of cell adhesion complexes such as tight and gap junctions. RNA-Seq analysis on Doxycycline (Dox)-inducible ZEB2-overexpressing DLD-1 colorectal cancer cell line clones. Data represents the mean centered rlog-transformed counts (n=3 biological repeats).



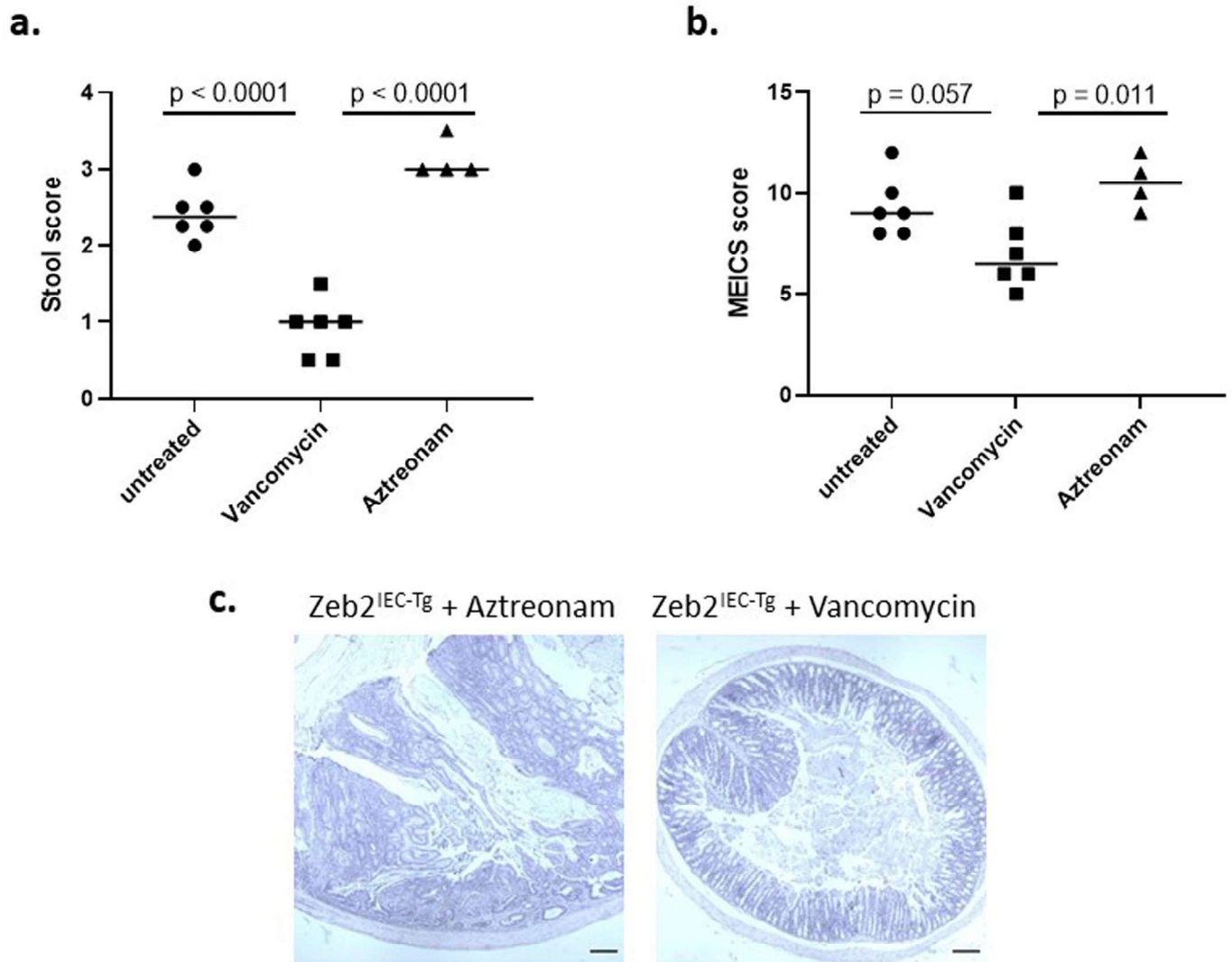
Extended Data Fig. 6 | Zeb2 expression and intestinal barrier integrity in the small intestine. Quantitative PCR analysis for expression of tight junction proteins (Claudin-4, -6 and -7, Occludin, Tjp-1, F11R/JAM-1) and desmosomes (Desmocollin-1 and -3, Desmoglein-2, Plakophilin-1) in lysates from the small intestine of *Zeb2*^{EC-Tg/+} (Tg/+, n=6 mice) and wild-type littermate controls (WT, n=5 mice). Data are presented as mean $\pm$ SEM and were analyzed with two-tailed Mann Whitney test. *, p-value = 0.0173; ** p-value = 0.0043; ns, non-significant.



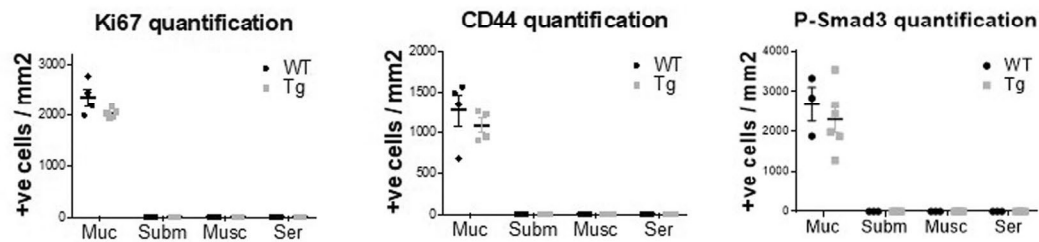
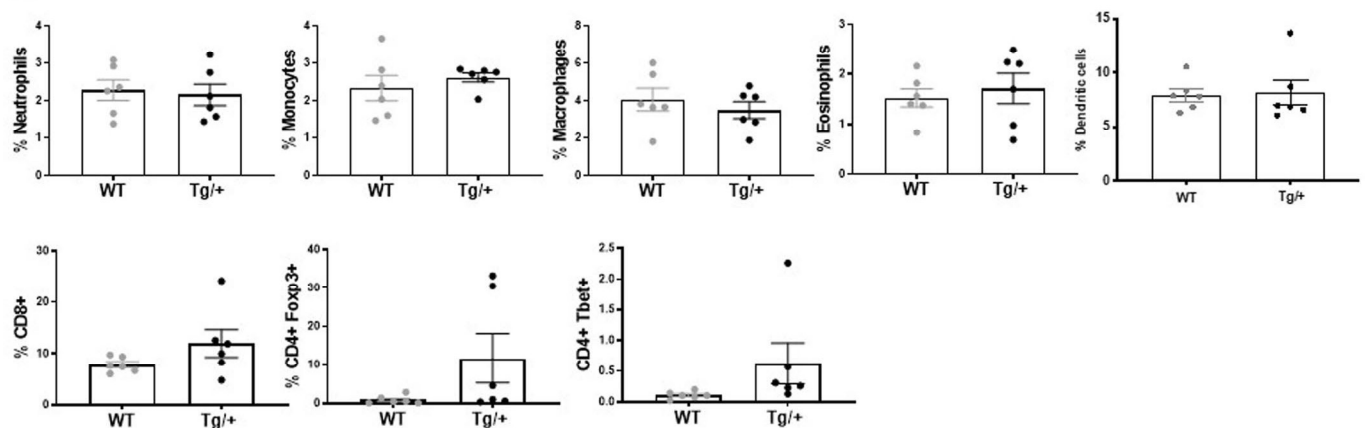
Extended Data Fig. 7 | Microbiota dysbiosis in *Zeb2*^{IEC-Tg/+} mice. Abundance of significantly different taxa between wild-type (WT, n=21 mice) and *Zeb2*^{IEC-Tg/+} (Tg/+, n=34 mice) fecal samples based on MaAslin and t-test FDR correction analysis. The horizontal straight line indicates the group means and the dotted line indicates the group medians.



Extended Data Fig. 8 | Microbiota dysbiosis in Zeb2^{IEC-Tg/+} mice. **a, b,** Altered bacterial diversity/community structure in mucosal samples from 25 week-old wild-type (WT, in green, n=28 mice) and Zeb2^{IEC-Tg/+} (Tg/+, in red, n=17 mice) mice. α -diversity measured by the Shannon Diversity Index at the genus level ($p=0.1439$, Welch two Sample t-test) (**a**). β -diversity represented by a PCoA of Bray-Curtis dissimilarity distances of bacterial composition (**b**). **c, d,** Differentially abundant taxa between wild-type (WT) and Zeb2^{IEC-Tg/+} (Tg/+). Linear Discriminant Analysis Effect Size (LefSe) cladogram of differentially abundant taxa in mucosal samples from WT (in green, n=21 mice) and Tg/+ (in red, n=34 mice) mice (**c**). Linear Discriminant Analysis (LDA) scores (**d**) of differentially abundant taxa combined with multiple testing corrections (analysis associated with Extended Data Fig. 8c): t-test with FDR correction (*: $p_{adj} < 0.05$) (no significant differences were identified based on MaAslin analysis). [phylum (p), class (c), order (o), family (f), genera (g), species (s)]. **e,** Abundance of significantly different taxa between wild-type (WT) and Zeb2^{IEC-Tg/+} (Tg/+) mucosal samples based on t-test FDR correction analysis. The horizontal straight line indicates the group means and the dotted line indicates the group medians. **f, g,** Mycobiota profiling in Zeb2^{IEC-Tg/+} mice. Fungal diversity/community (WT n=33 mice, in green, and Tg/+ n=22 mice, in red). α -diversity measured by the Shannon Diversity Index at the genus level ($p=0.0717$, Welch two Sample t-test) (**f**). β -diversity represented by a PCoA of Bray-Curtis dissimilarity distances of fungal composition (**g**). Each box represents the third (upper) and first (lower) quartiles. Lines inside the boxes represent the median.



Extended Data Fig. 9 | ABX treated Zeb2^{IEC-Tg/+} mice are protected from CRC development. **a**, Clinical stool score of untreated (n=6 mice), vancomycin-treated (n=6 mice) and aztreonam-treated (n=4 mice) Zeb2^{IEC-Tg/+} mice based on Hemocult assay. Data are presented as mean ± SEM. **b**, Colon pathology detected by endoscopy and represented as MEICS (murine endoscopic index of colitis severity) score. Each symbol represents one mouse. Data are presented as mean ± SEM. **c**, Colon histology of Aztreonam (left panel) and Vancomycin (right panel) treated Zeb2^{IEC-Tg/+} mice by H&E staining. Scale bar: 200µm. Representative images from one experiment with n=4 mice.

a.**b.**

Extended Data Fig. 10 | Axenic Zeb2^{IEC-Tg/+} mice are protected from CRC development. **a**, Quantification of Ki67-, CD44-, P-Stat3- and P-Smad3-positive cells in different regions of the colon (mucosa, submucosa, muscle or serosa) of 20 week-old germfree wild-type (WT) and Zeb2^{IEC-Tg/+} mice (Tg). For each tissue region and mouse, 3 separate areas were analyzed and averages per mouse per region were plotted in GraphPad Prism. Data are presented as mean $\pm$ SEM and were analyzed with two-way ANOVA with Sidak correction (Ki67 WT, n=4 mice and Tg/+, n=5 mice; CD44 WT, n=4 mice and Tg/+, n=4 mice; P-Smad3 WT, n=3 mice and Tg/+, n=6 mice). **b**, Flow cytometry analysis of colonic lamina propria immune cells in 12 week-old germfree WT (n=6 mice) and Zeb2^{IEC-Tg/+} (n=6 mice). Data are presented as mean $\pm$ SEM, two-sided student's t test. Each symbol represents one mouse.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- ☐ ☒ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- ☐ ☒ A description of all covariates tested
- ☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☒ ☐ For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☒ ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☒ ☐ Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

No software was used

Data analysis

Statistical analysis was performed with GraphPad Prism V8. qPCR data were analyzed with qBase+ software version 3.1 (Biogazelle). Flow cytometry data were analyzed using FlowJo software version 10.6.1.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The authors declare that the data supporting the findings of this study are available within the article and Supplementary Information Files or available from the authors upon reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Generally accepted sample sizes were used, with a significant difference between conditions, indicating that the sample size is sufficient.
Data exclusions	No data was excluded from the analysis
Replication	We analyzed a sufficient number of animals to ensure that the results obtained are consistent. All data in the manuscript are independently repeated at least once. Every figure states how many times an experiment has been repeated. Fully independent experiments have been performed, as defined by commonly accepted standards. All attempts at replication were successful.
Randomization	No specific method of randomization had been used to select animals. We compared groups of mice with different genotypes to assess the effect of specific genetic mutations in the phenotype. Group allocation was thus determined by the genotype of the mice.
Blinding	No blinding was done during group allocation as the group allocation was determined by the genotype of the mice. Histological scoring and counting of positive cells was performed blindly.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	rabbit anti-GFP (Cell Signaling, #2596, dilution 1/200), rabbit anti-Ki67 (Cell Signaling, #12202, dilution 1/1000), rabbit anti-Phospho-Smad3 phosphoS423-425 (1/500, Abcam, ab52903), anti-Phospho-Stat3 (Cell Signaling, #9145S, dilution 1/500), rat anti-CD45 (BD Pharmingen, 560501, 1/50), rabbit anti-CD3 (Abcam, ab5690, 1/500), rat anti-F4/80 (Serotec, MCA497G, 1/100), rabbit anti-ZO-1 antibody (1/100, ThermoFisher, 61-7300), neutralizing anti-GR1 antibodies (RB6.8C5 Mab, non commercial), anti-CD3 APC (eBioscience, 17-0031-83, 1/100), anti-CD4 APC-Cy7 (BD, 552051, 1/200), anti-CD8 V500 (BD, 560776, 1/100), anti-Foxp3 Alexa Fluor 488 (eBioscience, 53-5773-82, 1/100), anti-Tbet PE-Cy7 (eBioscience, 25-5825-82, 1/100), anti-RORgt BV421 (BD, 562894, 1/100) and anti-GATA3 PE (eBioscience, 12-9966-42, 1/100), Fixable Viability Dye eFluor™ 506 (eBioscience, 65-0866-14, 1/300), CD19 (eBioscience, 15-0193-82, 1/400), CD3 (eBioscience, 15-0031-82, 1/200), NK1.1 (Biolegend, 108716, 1/200) PE-Cy5], anti-CD45 Alexa Fluor 700 (eBioscience, 56-0451-82, 1/800), anti-Ly6G PercpCy5.5 (BD, 560602, 1/200), anti-Ly6C-APC (eBioscience, 17-5932-80, 1/200), [or anti-GR1 where mentioned (eBioscience, 45-5931-80, 1/200)], anti-Siglec F BUV395 (BD, 740280, 1/200), anti-CD11b BV605 (BD, 563015, 1/600), anti-CD64 BV711 (Biolegend, 139311, 1/100), anti-F4/80 Biotin (eBioscience, 13-4801-82, 1/100), Streptavidin BV421 (Biolegend, 405226, 1/1000), anti-CD11c PE-eFluor 610 (eBioscience, 61-114-82, 1/300), anti-MHC Class II APC-eFluor 780 (eBioscience, 47-5321-80, 1/800), EpCam PE-Cy7 (G8.8, eBioscience, 1/400, 25-5791-80), Ki67 eFluor660 (SolA15, eBioscience, 1/200, 50-5698-82), anti-Zeb2 (rabbit, 1/900, non-commercial) and anti-panCK (mouse, 1/300, DAKO, M3515, Clone AE1/AE3)
Validation	All antibodies used have been obtained from commercial vendors (with the exception of the neutralizing anti-GR1 antibody and the anti-ZEB2 antibody which has been validated for IHC on human samples). Validation information can be found on the respective Vendor websites.

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	DLD-1 cells were obtained from ATCC and lentivirally transduced with ZEB2cDNA using doxycyclin-inducible expression system
Authentication	DLD1 cell line from ATCC, not authenticated
Mycoplasma contamination	Cells were tested negative for mycoplasma.
Commonly misidentified lines (See ICLAC register)	DLD1 line not in ICLAC register

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	C57BL/6J Rosa26-Zeb2 transgenic mice (generated by the Geert Berx lab) were crossed with Villin-cre transgenic mice. Both male and female mice, between 1 and 55 weeks of age, were used in these studies. NSG mice were obtained from The Jackson Laboratory (NOD.cg-Prkdcscidll2rgtm1Wjl/SzJ).
Wild animals	The study did not involve wild animals
Field-collected samples	The study did not involve samples collected from the field.
Ethics oversight	All experiments on mice were conducted according to institutional, national and European animal regulations. Animal protocols were approved by the ethical committee of Ghent University Faculty of Sciences and Faculty of Medicine and Health Sciences (EC2017-042)

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	Colonic biopsies were obtained from CRC patients with following characteristics: patient 1 : f, age 82y, stage IV, moderately differentiated adenocarcinoma, BRAF_exon15_p.Val600Glu; TP53_exon7_p.Arg248Gln; TP53_exon5_p.Arg175His; PIK3CA_exon2_p.Arg108His; FBXW7_exon12_p.Gly687Glu; FBXW7_exon11_p.His580Asn, no treatment pre- or post-surgery patient 2 : f, age 86y, stage III, moderately differentiated adenocarcinoma, TP53_exon7_p.Cys242Trp; APC_exon15_p.Ser1356*; MET_exon2_p.Glu168Asp, no treatment pre- or post-surgery patient 3 : f, age 49y, stage IV, moderately differentiated adenocarcinoma, KRAS_exon2_p.Gly12Asp; TP53_exon5_p.Arg175His; APC_exon15_p.Gln1429*; FBXW7_exon9_p.Arg473Glufs*25, no treatment pre- or post-surgery patient 4 : 3, age 85y, stage IV, well-differentiated adenocarcinoma, KRAS_exon2_p.Gly12Val; APC_exon15_p.Gln406*, no treatment pre- or post-surgery
Recruitment	Colonic biopsies were obtained from CRC patients that had given written informed consent to participate.
Ethics oversight	Institutional Review Board approved Biobank of University Hospitals Leuven, Belgium (S50887-ML4707).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

- ☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☐ All plots are contour plots with outliers or pseudocolor plots.
- ☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Lamina propria isolation from colonic samples of SPF and GF Zeb2IEC-Tg and their wild type littermates was performed as described previously (Scott et al., Immunity, 2018). The isolated cells were used for extracellular and intracellular staining for representative markers of T and myeloid cells, in two separated panels (1x10 ⁶ cells/sample), and then analyzed by flow
--------------------	---

cytometry.
for assessment of epithelial cell proliferation: Colonic and ileal epithelial cells were isolated according to the following protocol: Colon and ileum were flushed with PBS and cut in small pieces, followed by incubation in dissociation solution (DMEM supplemented with 10 % FCS, 0.20 % collagenase and 0.125 µl/ml Dispase II) for 40 min, on a shaker at 37°C. To obtain single cells, the digested tissues were passed through a 100 µm filter, followed by a wash with DMEM +10 % FCS. The pellet was resuspended in 30 % percoll and centrifuged for 5 min at 400 x g. Red blood cells were removed by incubation with ACK lysis buffer for 1 min at room temperature.

Instrument amples were measured on a FACS Fortessa 5 laser or BD FACSymphony (BD Biosciences) and data were analyzed using FlowJo.

Software FlowJo was used to analyze the data.

Cell population abundance All shown cell populations were significantly abundant. All cell population were alive by using a live/dead marker. The abundance was calculated by adding counting beads to the sample just before measurement.

Gating strategy The flow cytometry data were analyzed using FlowJo software and a sequential gating strategy. The graphs demonstrating percentages of myeloid cells represent proportions of every parent population. In the analysis of T cells, CD8+ are presented as percentage of the CD3+ parent population and CD4+Tbet+ and CD4+Foxp3+ are presented as percentages of the CD4+ population. Regarding the tSNE plots, 5000 cells/sample were exported from 6 mice/genotype. The export was on the CD3+ for the T cells and on CD45+lineage (CD19, Nk1.1, CD3) for the myeloid cells. The samples were then concatenated and then analyzed with tSNE (iterations=1000, perplexity=30), so the representative plots show the a total of 30000 cells.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.